



HAL
open science

Quelles données environnementales pour territorialiser les limites planétaires? Un cas d'étude sur l'évaluation environnementale absolue de l'intégrité des cours d'eau sur un territoire

Quentin Dassibat

► To cite this version:

Quentin Dassibat. Quelles données environnementales pour territorialiser les limites planétaires? Un cas d'étude sur l'évaluation environnementale absolue de l'intégrité des cours d'eau sur un territoire. Sciences de l'environnement. École Nationale Supérieure des Mines de Saint-Étienne, 2024. Français. NNT : 2024EMSEM050 . tel-04958701

HAL Id: tel-04958701

<https://hal.science/tel-04958701v1>

Submitted on 20 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

THÈSE DE DOCTORAT de l'Institut Mines-Télécom Ecole Nationale Supérieure des Mines de Saint-Etienne

**Ecole Doctorale N°488
(Sciences, Ingénierie, Santé)**

Spécialité de doctorat :
Sciences et génie de l'environnement

Soutenue publiquement le 9 décembre 2024 par
Quentin Dassibat

Quelles données environnementales pour territorialiser les limites planétaires ?

Un cas d'étude sur l'évaluation environnementale absolue
de l'intégrité des cours d'eau sur un territoire

Devant un jury composé de :

AISSANI Lynda	Ingénieure de recherche, INRAE Rennes	Rapportrice
JUNQUA Guillaume	Maître-assistant, IMT Mines Alès	Rapporteur
GAUTIER Mathieu	Professeur, INSA Lyon	Examineur
LIPEME KOUYI Gislain	Professeur des universités, INSA Lyon	Président
PIATYSZEK Eric	Maître de conférences, Mines Saint-Etienne	Examineur
PIEGAY Hervé	Directeur de recherche, ENS de Lyon	Examineur
MARSAUCHE Maud	Chargée d'étude, Epures	Membre invitée
BARDET Fabrice	Directeur de recherche, ENTPE	Co-directeur de thèse
GONDRAN Natacha	Professeure, Mines Saint-Etienne	Co-directrice de thèse

Affidavit

Je soussigné, Quentin Dassibat, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Natacha Gondran et Fabrice Bardet, dans le respect des principes d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect de la charte nationale de déontologie des métiers de la recherche.

Ce travail n'a pas été précédemment soumis dans sa globalité en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Lyon, le 7 octobre 2024.

Quentin Dassibat

Ce travail de thèse est une œuvre de l'esprit, protégée par le droit d'auteur, tel que prévu aux articles L111-1 du CPI et suivants disposant que "L'auteur d'une œuvre de l'esprit jouit sur cette œuvre, du seul fait de sa création, d'un droit de propriété incorporelle exclusif et opposable à tous. [...]" Il est rappelé que par exception au droit d'auteur, la loi française autorise l'utilisation d'une œuvre divulguée, sans autorisation de son auteur, suivant les conditions définies dans l'article L122-5 du CPI disposant que "Lorsque l'œuvre a été divulguée, l'auteur ne peut interdire [...] la représentation ou la reproduction d'extraits d'œuvres, [...] sous réserve que soient indiqués clairement le nom de l'auteur et la source [...] les analyses et courtes citations justifiées par le caractère critique, polémique, pédagogique, scientifique ou d'information de l'œuvre à laquelle elles sont incorporées [...]"

Remerciements

"L'hypothèse retenue ici est que la quantification, entendue comme l'ensemble formé des conventions socialement admises et des opérations de mesure, crée une nouvelle façon de penser, de représenter, d'exprimer le monde et d'agir sur lui."

(DESROSIÈRES 2008).

Si cette hypothèse a pu également constituer le point de départ de mon travail de thèse, c'est qu'il me faut d'abord et conjointement remercier ma directrice et mon directeur de thèse, Natacha Gondran et Fabrice Bardet, pour leur ouverture à un dialogue entre leurs disciplines, entre sciences de l'environnement et sciences sociales. A cet égard, je remercie également feu l'École Urbaine de Lyon qui, au-delà d'être le financeur de la thèse, a constitué un espace propice à l'interdisciplinarité sur la place académique de Lyon-Saint-Étienne, comme manière de faire reconnaître une autre manière de connaître. L'UMR Environnement Ville Société a su prendre le relai de l'École Urbaine de Lyon, et je tiens à saluer la bienveillance des collègues rencontrés çà et là lors des séminaires et des "journées" de l'UMR. De même, je remercie l'ensemble des membres de mon jury pour leur ouverture disciplinaire et leur bienveillance à l'égard de ce travail exploratoire, certes, mais que la relecture particulièrement attentive d'Eric Piatyszek a rendu scientifiquement robuste. Je remercie les collègues de mes deux laboratoires d'accueil, RIVES à l'ENTPE et GEO à l'École des Mines de Saint-Étienne, qui ont su créer une ambiance de travail stimulante et rassurante à laquelle je me suis particulièrement accrochée dans les moments difficiles. Je remercie mes ami-e-s de longue date, celles et ceux qui le sont devenu-e-s, celles et ceux aussi que la divergence des lieux, des temporalités ou des trajectoires éloigne. Je remercie mes proches à qui j'ai trop peu dit, je crois, que cette aventure n'aurait pas abouti sans elles et eux.

Note aux lecteur-rice-s

Bien que nous reconnaissons son utilité et sa nécessité, l'utilisation d'une forme d'écriture inclusive ou non binaire n'a pas été retenue pour la rédaction de ce manuscrit. Dans un souci de clarté du propos, que la manipulation d'un certain nombre de concepts peut déjà rendre difficilement accessible, nous avons fait le choix conservateur du masculin pour désigner le neutre. Nous invitons donc les lecteur-rice-s à adopter, à la lecture de tout neutre masculin, la forme d'inclusivité ou de non-binarité qui leur plaît.

Table des matières

Affidavit	i
Remerciements	iii
Note aux lecteur-rice-s	v
Table des matières	x
Liste des figures	xiv
Liste des tableaux	xv
Liste des abréviations	xvii
Introduction générale	1
La Terre a-t-elle dépassé ses limites?	1
Institutionnaliser les limites planétaires : plusieurs dynamiques de portage	7
Hypothèses directrices du travail de thèse	15
Organisation du manuscrit	18
1 Le cadre des limites planétaires : du système Terre au territoire	21
Introduction	21
1.1 Le cadre global des limites planétaires	22
1.1.1 L’esprit du cadre : préserver un état environnemental de référence pour le système Terre	22
1.1.2 Les évolutions d’un cadre toujours plus robuste	28
1.1.3 Les limites planétaires : quels processus environnementaux, quels indicateurs, quel état?	32
1.2 Un nouveau cadre pour penser l’évaluation environnementale des organisations	40

1.2.1	Un renouvellement paradigmatique dans le champ de l'évaluation environnementale	40
1.2.2	Une introduction à l'évaluation absolue de la soutenabilité environnementale	42
1.2.3	Le principe d'assignation au cœur d'une évaluation AESA basée sur les limites planétaires	43
1.3	Utiliser les limites planétaires à l'échelle des organisations : les méthodes principales	46
1.3.1	Les applications à l'Analyse de Cycle de Vie dans le champ scientifique . .	46
1.3.2	Un nouveau défi pour l'intégration des enjeux environnementaux dans la prise de décision	50
1.4	Conclusion et introduction des hypothèses de recherche	56
1.4.1	Les modalités d'atterrissage du cadre des limites planétaires	56
1.4.2	Resserrer l'étude	60
1.4.3	Conduire une évaluation environnementale ou questionner l'évaluation environnementale ?	62
1.4.4	La donnée environnementale au cœur de la question de l'échelle d'étude .	64
1.4.5	Organisation du manuscrit	69
2	Compter l'eau : panorama des données environnementales sur l'eau dans les travaux sur les limites planétaires	71
	Introduction	71
2.1	L'eau dans le cadre des frontières du système Terre	72
2.1.1	Évolution de l'indicateur et des modèles au fil des mises à jour du cadre des limites planétaires	72
2.1.2	La frontière du système Terre relative à l'eau de surface	90
2.2	Les applications sous-globales de la limite relative à l'eau	102
2.2.1	Méthodologie pour la revue de la littérature	102
2.2.2	Les jeux de données représentatifs de la littérature : remonter la chaîne d'élaboration de la donnée	111
2.3	Conclusion et questions de recherche	129
2.3.1	Une échelle de la donnée ?	129
2.3.2	Vers un gradient d'approches <i>bottom-up</i>	132
3	Utiliser des données de mesure de débits dans une AESA : deux propositions méthodologiques	137
	Introduction	137

3.1	Le modèle hydrographique élémentaire	140
3.1.1	Préambule : notions de système d'information géographique	140
3.1.2	Intérêt du modèle hydrographique élémentaire	143
3.1.3	Méthode numérique pour la construction du MHE	144
3.2	L'évaluation environnementale de l'intégrité des cours d'eau par les données de mesure	153
3.2.1	Définir un indicateur de variabilité : d'une variabilité "de référence" à une variabilité "normale"	154
3.2.2	Le potentiel analytique de la variabilité statistique	158
3.3	L'assimilation des sorties d'un modèle hydrologique global à des données de mesure	169
3.3.1	Un modèle de rééchantillonnage de données hydrologiques spatialisées . . .	171
3.3.2	Formalisation mathématique du modèle	175
	Conclusion	191
4	Application numérique au territoire du SCoT Sud-Loire	195
	Introduction	195
4.1	Cas d'application : les données de mesure du réseau Vigicrues sur le territoire du SCoT du Sud-Loire	196
4.1.1	Identification des données disponibles	196
4.1.2	Les données hydrométriques du réseau Vigicrues	197
4.2	Validation du modèle hydrographique élémentaire	202
4.2.1	Remarque préalable : modélisation réaliste <i>versus</i> modélisation construc- tivistique	202
4.2.2	Procédure de validation	203
4.3	Évaluer l'intégrité des cours d'eau sur un territoire par les données de mesure : le logiciel HydroTrendsFrance	207
4.3.1	Description du logiciel HydroTrendsFrance : application numérique de la méthode d'évaluation environnementale	208
4.3.2	Résultats	213
4.4	Assimiler le modèle hydrologique global WBM aux données de débits Vigicrues : le logiciel relio	234
4.4.1	Description du logiciel relio : application numérique du modèle de ré- échantillonnage hydrologique	235
4.4.2	Mise en œuvre de l'algorithme invdist pour la comparaison à relio . . .	238
4.4.3	Résultats	240

Conclusion	255
Conclusion générale	263
Bibliographie	298
Annexes	299
1 Annexes au chapitre 3	299
1.1 Annexe à la section 4.2.2	299
1.2 Annexe à la section 4.2.3	303
2 Annexes au chapitre 4	306
2.1 Annexe à la section 4.2	306
2.2 Annexe à la section 5.1	311
Résumé	335
Abstract	337

Liste des figures

0.1	L'état de franchissement des différentes limites planétaires dans la dernière mise à jour de septembre 2024 (CAESAR et al. 2024)	2
0.2	Les trajectoires du système Terre en dehors de l'Holocène (STEFFEN et al. 2018)	4
0.3	La représentation de l'espace de fonctionnement sûr et juste par le <i>donut</i> (RAWORTH 2012)	11
0.4	Description géographique du SCoT Sud-Loire (SYNDICAT MIXTE DU SUD-LOIRE 2022)	14
0.5	Tableau bilan de l'évaluation environnementale du SCoT Sud-Loire conduite au regard du cadre des limites planétaires (EPURES 2021)	16
1.1	Les grandes tendances socioéconomiques depuis le début de l'ère industrielle (1750) au regard de l'évolution des variables de fonctionnement du système Terre. (Traduit de (STEFFEN et al. 2015a))	24
1.2	La représentation standard des limites planétaires en graphique radar issue de la version 2023 du cadre (Traduit de RICHARDSON et al. 2023).	26
1.3	Espace de fonctionnement sûr et juste issu de la version 2023 du cadre (Traduit de ROCKSTROM et al. 2023)	32
1.4	Principe de partage de l'espace de fonctionnement sûr proposé par Hjalsted et ses collègues (Traduit de HJALSTED et al. 2021).	44
1.5	Résultats de l'étude menée par le JRC européen sur l'impact de la consommation des Européens comparée à la part assignée des limites planétaires. (Traduit de (SALA et al. 2020)	49
1.6	La France face aux limites planétaires. Diagnostic réalisé par le Commissariat Général au Développement Durable (MTES 2023).	52
1.7	Le "donut" de Kate Raworth (RAWORTH 2017), qui s'appuie sur les limites planétaires pour définir son "plafond" environnemental.	54
1.8	Les différentes modalités d'atterrissage du cadre global vers un système sous-global	58

1.9	Le rôle de l'échelle de la donnée environnementale dans une approche <i>bottom-up</i> .	67
2.1	Représentation des flux constitutifs de l'empreinte eau bleue et de l'empreinte eau verte (ALDAYA et al. 2012)	75
2.2	Les 6 limites planétaires relatives à l'eau dans GLEESON et al. 2020b et leurs interactions avec le reste des limites planétaires	84
2.3	Construire une variable de contrôle pour l'eau dans le cadre des limites planétaires qui tiennent compte de l'hétérogénéité spatiale des phénomènes (GLEESON et al. 2020a)	88
2.4	Les quatre profils d'évolution d'une variable de contrôle qui définissent quatre types de limites planétaires (DEARING et al. 2014)	90
2.5	Réponse des macroinvertébrés, des espèces ripariennes et des poissons à l'altération quantitative des cours d'eau (POFF et al. 2010)	94
2.6	Schéma de l'approche par <i>percent-of-flow</i> (RICHTER et al. 2012)	95
2.7	Erreur d'estimation du ruissellement par le modèle WBM selon la surface du bassin versant modélisé (FEKETE et al. 2002)	102
2.8	Un gradient d'approche <i>bottom-up</i> par le croisement de l'échelle de l'étude et de l'échelle de la donnée	133
3.1	Schéma conceptuel de la méthodologie en 3 parties : méthode pour le MHE (1), méthode pour l'AESA (2), méthode pour le MRH (3)	139
3.2	Description d'une image raster	141
3.3	Description d'une couche vectorielle (ici, 4 types de géométrie sont représentés) .	142
3.4	Schéma séquentiel des traitements effectués par le MHE	147
3.5	Perturbation du cycle de l'eau en milieu urbain par les activités humaines (SØRUP et al. 2020)	158
3.6	Rééchantillonnage des données de modèle hydrologique global par le modèle MRH	173
3.7	Rééchantillonnage des données de mesure de débits par le modèle MRH	174
3.8	Validation croisée du modèle MRH par l'algorithme GDAL de distance inverse . .	175
3.9	Intervention des paramètres α_a , α_b et β_p dans le modèle MRH et lien avec le modèle hydrographique élémentaire (MHE)	177
3.10	Vue d'ensemble du modèle de rééchantillonnage hydrographique	180
3.11	Interpolation du débit par l'algorithme de convolution hydrographique linéaire . .	182
3.12	Représentation schématique de l'algorithme de dérasterisation et lien avec l'algorithme de convolution	185

4.1	Le parcours de production des données hydrométriques Vigicrues au sein du Système d'Information sur l'Eau (SIE)	199
4.2	Échelle limnimétrique du réseau Vigicrues installée sur l'Allier (Prades, Haute-Loire, 2024) (c) Gondran N.	199
4.3	Le bassin versant de la Loire à partir de la station Vigicrues du barrage de Villerest obtenu par le MHE (rose) et par le modèle WatershedWPS (bleu)	206
4.4	Schéma de fonctionnement du logiciel HydroTrendsFrance	209
4.5	Localisation des stations hydrométriques utilisées pour la chronique 2000-2020.	214
4.6	Valeurs mensuelles de l'indicateur d'anomalie de variabilité pour la station la plus à l'aval du territoire du Sud-Loire (station de Villerest, Loire, 2000-2020)	219
4.7	Fréquences mensuelles de dépassement des stations Vigicrues du Sud-Loire (moyenne annuelle 2000-2020, visualisation par les zones d'apport aux stations)	221
4.8	Comparaison des résultats entre l'indicateur de (ROCKSTROM et al. 2023) et l'indicateur d'anomalie de variabilité	224
4.9	Évolution tendancielle des débits mensuels (Sud-Loire, 2000-2020)	227
4.10	Évolution tendancielle des débits moyen mensuels (station de Villerest : 1921-2023; station de Bas-en-Basset : 1918-2023)	228
4.11	Évolution tendancielle des écarts-types moyens mensuels (station de Villerest : 1921-2023; station de Bas-en-Basset : 1918-2023)	229
4.12	Matrice de corrélation entre le coefficient de variation annuel moyen des débits et les éléments anthropiques du paysage (Sud-Loire, 2000-2020)	231
4.13	Isolement des variables corrélées à la variabilité des débits (Sud-Loire, 2000-2020)	232
4.14	Schéma des différents traitements effectués par le MRH et mis en œuvre dans relio	235
4.15	Séquences des traitements effectués pour le rééchantillonnage d'un raster WBM selon l'algorithme de distance inverse.	240
4.16	Performances de relio par rapport à des données de contrôle	243
4.17	Distribution du MBE de relio	244
4.18	Descripteurs statistiques des distributions du MBE pour les trois principaux bassins versants du territoire	245
4.19	Distribution du résidu additif de relio pour tous les mois et toutes les stations	246
4.20	Distribution du résidu multiplicatif de relio pour tous les mois et toutes les stations	247
4.21	Performances comparées de relio et invdist par rapport à des données de contrôle	248
4.22	Distribution du MBE de invdist	249
4.23	Distributions comparées du résidu additif de relio et invdist	250

4.24	Distributions comparées du résidu multiplicatif de relio et invdist	250
4.25	Visualisation du rééchantillonnage des données carroyées WBM selon notre méthode hydrographiquement contrainte (relio) et selon une méthode continue (invdist)	251
4.26	Comparaison des valeurs de débits WBM rééchantillonnées par rapport aux données de mesure Vigicrues, pour relio et invdist	252
4.27	Comparaison des valeurs de débits WBM rééchantillonnées par relio par rapport aux données de mesure Vigicrues	253
4.28	Distribution du MBE pour les valeurs de débits du WBM rééchantillonnées par relio	254
4.29	Distribution comparée du résidu additif pour les valeurs de débits du WBM rééchantillonnées par relio et invdist	255
4.30	Distribution comparée du résidu multiplicatif pour les valeurs de débits du WBM rééchantillonnées par relio et invdist	255
A.1	Représentation schématique du lien entre un modèle de ruissellement et un modèle d'écoulement	306
A.2	Anomalie de variabilité sur le territoire du SCoT du Sud-Loire	309
A.3	Anomalie de variabilité sur le territoire de la Métropole du Grand Lyon	310
A.4	Anomalie de variabilité sur le territoire de la communauté d'agglomération du Cotentin	311

Liste des tableaux

1.1	Évolutions du cadre des limites planétaires, de 2009 à 2023.	30
2.1	Tableau de synthèse des données utilisées dans le corpus de publications	107
2.2	Les différentes sources hiérarchiques de la donnée que nous utilisons pour qualifier l'une des deux dimensions de l' <i>échelle de la donnée</i>	131
3.1	Protocoles de traitement des différents types d'erreur	161
3.2	Comparaison des algorithmes de dérasterisation entre le modèle MRH et un mo- dèle SIG quelconque	187

Liste des abréviations

AESA	Absolute Environmental Sustainability Assessment
API	Application Programming Interface
BRGM	Bureau de recherches géologiques et minières
DOO	Document d'orientations et d'objectifs
DREAL	Direction régionale de l'environnement, de l'aménagement et du logement
EE-IO	Environmentally Extended Input/Output (table ou analyse)
EE-MRIO	Environmentally Extended Multiregionalized Input/Output (table ou analyse)
EFR	Environmental Flow Requirements
EPCI	Établissement public de coopération intercommunale
FAO	Food and Agriculture Organization
GES	Gaz à effet de serre
GHM	Global Hydrological Model
GIEC	Groupe d'experts intergouvernemental sur l'évolution du climat
GRDC	Global Runoff Data Center
IHE	Delft Institute for Water Education
ILCD	International Life Cycle Data system
JRC	Joint Research Center
LCI	Life Cycle Inventory
LCIA	Life Cycle Impact Assessment
LPJmL	Lund-Potsdam-Jena managed Land (modèle)
MAESA	Méthode pour l'AESA

MHE	Modèle hydrographique élémentaire
MIT	Massachusetts Institute of Technology
MNT	Modèle numérique de terrain
MRH	Modèle de rééchantillonnage hydrologique
MRIO	Multiregionalized Input/Output (table ou analyse)
MTES	Ministère de la transition écologique et solidaire
ODD	Objectifs de développement durable
OMD	Objectifs du millénaire pour le développement
ONU	Organisation des nations unies
OSM	Open Street Map
PAS	Projet d'aménagement stratégique
PhyC	Plateforme d'hydrométrie centrale
PIK	Potsdam Institute for Climate Impact Research
POF	Percent-of-Flow
PTGE	Projet de territoire pour la gestion de l'eau
relio	REsampling Large Inputs/Outputs (modèle)
SAGE	Schéma d'aménagement et de gestion de l'eau
SCoT	Schéma de cohérence territoriale
SDAGE	Schéma directeur d'aménagement et de gestion des eaux
SIE	Système d'information environnementale ou Système d'information sur l'eau (selon contexte)
SIG	Système d'information géographique
SOS	Safe operating space
SPC	Service de prévision des crues
SRC	Stockholm Resilience Centre
STN	Simulated Topological Network
UH	Unité d'hydrométrie
UNESCO	United Nations Educational, Scientific and Cultural Organization

WBM Water balance Model

Introduction générale

La Terre a-t-elle dépassé ses limites ?

En septembre 2009, paraît la première étude sur les *limites planétaires* qui fait état du dépassement de trois limites (ROCKSTRÖM et al. 2009a). Trois mécanismes de régulation de la Terre sont alors décrits comme en situation de risque élevé : le changement climatique, l'intégrité de la biosphère et la perturbation du cycle des nutriments. En janvier 2015, le dépassement d'une quatrième limite est annoncé, qui concerne le changement d'affectation des sols (STEFFEN et al. 2015b). En mai 2023, quatre limites sont toujours dépassées mais le cycle de l'eau serait davantage à risque que le climat (ROCKSTROM et al. 2023). En septembre 2023, les scientifiques alertent sur le dépassement de six limites : la perturbation du cycle de l'eau douce et le changement climatique sont désormais tous deux en risque élevé, ainsi que l'intégrité de la biosphère, la perturbation du cycle des nutriments, le changement d'affectation des sols et l'introduction de molécules chimiques nouvelles dans l'environnement (RICHARDSON et al. 2023). Un an plus tard, en septembre 2024, une septième limite est sur le point d'être dépassée : l'acidification des océans (voir les limites indiquées en gras sur la Figure 0.1) (CAESAR et al. 2024). Cet enchaînement resserré de publications qui font état d'un nombre toujours croissant de limites planétaires dépassées fait écho à ce que certains auteurs ont appelé la *Grande Accélération* (STEFFEN et al. 2015a). Ce terme désigne cette période de l'histoire qui commence aux environs des années 1850, très courte dans l'histoire de l'humanité, où une série d'indicateurs socio-économiques (richesse des nations, population mondiale, etc) et environnementaux (gaz à effet de serre, acidité des océans, etc) se sont mis à croître à une allure exponentielle, marquant une rupture sans précédent avec leur stabilité historique.

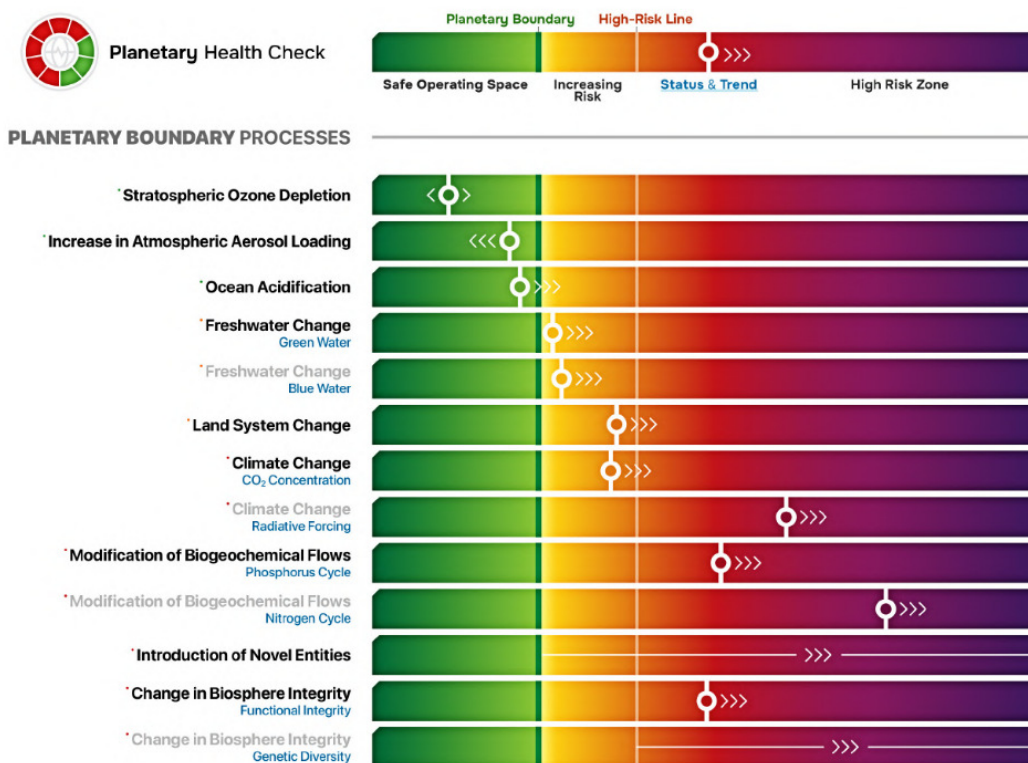


Figure 0.1 – L'état de franchissement des différentes limites planétaires dans la dernière mise à jour de septembre 2024 (CAESAR et al. 2024)

Pour répondre à la question que nous avons posée en introduction, il nous faut distinguer entre une réponse physique, à savoir ce que les limites planétaires disent de l'état physique des mécanismes de régulation de la Terre; et une réponse conceptuelle, à savoir ce qu'elles disent en tant que cadre conceptuel qui propose une certaine façon de quantifier notre rapport à l'environnement. Autrement dit, la hausse du nombre des limites planétaires dépassées peut être vue d'un côté comme une détérioration croissante des mécanismes environnementaux de régulation globale, et de l'autre comme l'effet d'une évolution de la façon de mettre la Terre en indicateurs.

Le dépassement des limites planétaires dans une perspective physique

Pour ce qui est de la réponse physique, les limites planétaires sont à envisager comme une synthèse des principaux mécanismes de régulation de la Terre. Plus précisément, elles décrivent les conditions qui permettent de maintenir l'état stable de ce qu'il convient d'appeler le *système Terre* (STEFFEN et al. 2020b). La Terre peut en effet être décrite comme un système dont on peut caractériser la stabilité ou l'instabilité, comme résultat de l'équilibre entre ses différents compartiments que sont l'atmosphère, la biosphère, l'hydrosphère, la lithosphère (LOVELOCK

1972). L'époque géologique de l'Holocène correspond précisément à une période de stabilité inédite du système Terre. Cette période, que les géologues font débuter avec la fin des dernières glaciations du Pléistocène il y a plus de 10 000 ans, se caractérise notamment par une stabilité exceptionnelle des concentrations de gaz à effet de serre dans l'atmosphère et, par conséquence, des températures de surface autour d'une valeur normale tempérée (STEFFEN et al. 2005). Cette stabilité constitue l'un des facteurs explicatifs de l'expansion des foyers de peuplement d'Homo sapiens sur la surface du globe au Néolithique, du développement concomitant de l'agriculture sédentaire et des formes d'organisation sociale associées (LUSSAULT et al. 2021). L'Holocène a ainsi été retenu comme époque géologique de référence par les concepteurs de la notion de limites planétaires en 2009 (ROCKSTRÖM et al. 2009a). Une limite planétaire quantifie donc l'écart entre l'état actuel d'une composante environnementale et son état stable sous l'Holocène. Les limites planétaires dans leur ensemble constituent alors un nombre réduit d'indicateurs qui décrivent l'éloignement actuel du système Terre par rapport à l'époque de référence de l'Holocène. Garder les mécanismes de régulation planétaire à leur état stable sous l'Holocène, c'est garantir ce que les publications nomment un *espace de fonctionnement sûr pour l'humanité*. Au-delà de cet état en revanche, c'est une trajectoire incertaine et instable qui se dessine (voir Figure 0.2), qui pourrait aboutir à ce que certains auteurs appellent la planète-serre (STEFFEN et al. 2018) ou que recouvre plus généralement le terme d'*Anthropocène*, qui caractérise cette période de l'histoire de la Terre où les activités humaines se sont hissées au rang de force géologique capable de modifier les caractéristiques de la période géologique actuelle (CRUTZEN 2002). En cela, c'est moins la Terre elle-même qui franchirait ses limites qu'une époque particulière de l'histoire géologique de la Terre. Le dépassement des limites planétaires représente ainsi une menace pour les conditions d'habitabilité de la Terre particulièrement propices au vivant depuis l'Holocène.

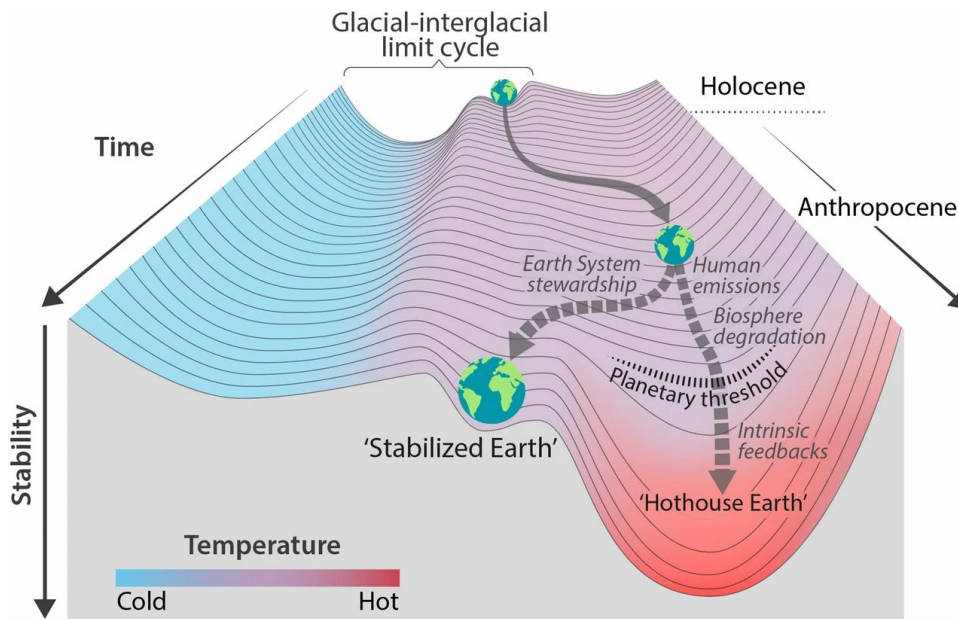


Figure 0.2 – Les trajectoires du système Terre en dehors de l’Holocène (STEFFEN et al. 2018)

Pour identifier, parmi tous les mécanismes de régulation du système Terre ceux qui sont déterminants dans le maintien de la stabilité de l’Holocène, les scientifiques se sont basés sur au moins deux critères : l’influence des mécanismes sur la résilience du système Terre (sa capacité à revenir à un état stable après une perturbation) et leur fonction de support au vivant. Le recours au concept de résilience fait d’ailleurs des limites planétaires une notion assez proche de celle des *points de bascule* (LENTON et al. 2008), qui sont pour leur part des mécanismes environnementaux davantage locaux, par rapport à l’échelle globale des mécanismes décrits par les limites planétaires, dont la déstabilisation peut entraîner un "effet domino" sur les mécanismes globaux et ainsi menacer la stabilité du système Terre lui-même (par exemple, la fonte du permafrost en Sibérie, le déboisement de la forêt amazonienne ou le régime des moussons en Afrique centrale ou en Asie du Sud-Est). Une dizaine de mécanismes environnementaux ont ainsi été identifiés comme déterminants pour la résilience et le vivant sur Terre. Ce nombre varie légèrement selon les publications que nous avons citées plus haut, mais conserve le socle stable des mécanismes suivants : le changement climatique, l’intégrité de la biosphère, le changement d’affectation des sols, la perturbation du cycle de l’eau douce, la perturbation du cycle des nutriments, l’introduction de molécules chimiques nouvelles dans l’environnement et le rejet d’aérosols dans l’atmosphère. Quant à la valeur stable que prennent ces mécanismes sous l’Holocène, les scientifiques s’appuient sur les travaux des communautés académiques auxquelles se rapporte chaque mécanisme (climatologie, chimie de l’atmosphère, océanologie, biologie, écologie, hydrologie, etc). La valeur stable sous l’Holocène, c’est-à-dire le seuil dont il ne faudrait pas s’éloigner, ou encore la *limite* à proprement parler, est ainsi

déterminée par consensus au sein de la communauté scientifique.

Le dépassement des limites planétaires dans une perspective conceptuelle

Il est tentant d'établir la parenté des limites planétaires avec la série des travaux qui, depuis plusieurs décennies déjà, ont sonné l'alarme sur l'insoutenabilité environnementale des activités humaines sur Terre. Les limites planétaires semblent en effet spontanément s'inscrire dans la lignée des rapports du GIEC sur le climat (quel est le niveau maximum d'augmentation des températures que l'on puisse supporter ?) (IPCC 2022), des travaux sur l'empreinte écologique (quel est le nombre de planètes Terre consommées chaque année pour le besoin des activités humaines ?) (GLOBAL FOOTPRINT NETWORK 2023) ou du plus ancien, et sans doute à l'origine de cette lignée, rapport du Club de Rome *Halte à la croissance* (MEADOWS et al. 1972). Si toutes ces démarches racontent la même histoire (la Terre ne peut pas soutenir indéfiniment les activités humaines dans la forme qu'elles ont progressivement et presque globalement prise depuis les révolutions industrielles), elles ne sont pas conceptuellement si proches lorsque l'on considère toutes ces démarches, précisément, comme des cadres conceptuels, c'est-à-dire comme des façons de penser, décrire et quantifier notre relation à l'environnement (DAHAN 2007).

Le rapport *Halte à la croissance* est construit sur la base des résultats obtenus par la 3ème génération du modèle *World*. Ce modèle est issu des travaux initiés dans les années 1950 par le MIT sur la dynamique des systèmes et à partir desquels a été forgé le concept de *modélisation globale* (FORRESTER 1971), devenu rapidement un champ scientifique à part entière. La démarche des modèles globaux consiste à décrire par des variables liées entre elles le fonctionnement d'un système, et à simuler informatiquement leur comportement (on parle de *rétroactions*) sur une période passée, présente ou future. C'est ce que fait le modèle *World3*, l'un des premiers modèles globaux, en simulant le lien entre la population mondiale, le niveau de richesse et la consommation des ressources naturelles. Les résultats de ce modèle sont à la base du rapport *Halte à la croissance*, qui envisage alors conceptuellement la relation à l'environnement par la question des ressources, ce qui, nous l'avons vu, n'est pas l'angle retenu par le cadre des limites planétaires.

Le cadre conceptuel de l'empreinte écologique, qui a été développé au début des années 1990, s'inscrit pour sa part dans une autre histoire, celle du champ scientifique de l'évaluation environnementale. L'empreinte écologique repose sur le concept de capacité de charge qui

est issu du champ de l'écologie et qui correspond à la population maximum qu'un milieu peut accueillir compte tenu des ressources disponibles. L'empreinte écologique n'est autre que l'inverse de cette capacité de charge : combien de ressources sont-elles nécessaires pour supporter une population humaine donnée et son mode de vie en termes d'utilisation des ressources et d'assimilation des déchets ? (REES 1992). Méthodologiquement, il ne s'agit plus ici d'effectuer des simulations informatiques mais une comptabilité environnementale, qui convertit les flux physiques de matières liés aux activités humaines en impacts environnementaux, et en l'occurrence en surfaces bioproductives nécessaires à ces activités (hectares globaux) (BORUCKE et al. 2013). Si l'empreinte écologique élargit la prise en compte des impacts anthropiques sur l'environnement au-delà des seules ressources disponibles (ou plutôt, elle convertit en "équivalent ressources" des pressions qui ne sont pas d'emblée considérées comme telles, comme les émissions de gaz à effet de serre), elle ne permet pas conceptuellement de penser les relations entre les différents mécanismes globaux de régulation planétaire parce qu'elle les synthétise dans un indicateur unique. Toutefois, un tel indicateur, en normalisant les résultats de l'évaluation environnementale sur une référence environnementale, ouvre la voie à ce qui deviendra plus tard l'évaluation environnementale dite *absolue* (BJØRN et al. 2013).

Si la filiation scientifique et conceptuelle de ces différents travaux est bien différente, une histoire commune se dessine toutefois dans le champ politique, par l'effort de portage politique et d'institutionnalisation dont ces travaux ont fait l'objet, en particulier au niveau des instances de gouvernance internationale.

Certains auteurs ont en effet proposé de relire cette histoire des cadres de gouvernance internationale de l'environnement par le prisme d'un effort de traduction dans le champ de l'action des travaux scientifiques sur l'environnement. Différentes lectures de cette traduction ont été ainsi proposées. Pour (ANDRÉ et al. 1999), il s'agit d'une traduction par le champ de l'*évaluation des impacts environnementaux*, un champ qui émerge dans les années 1970 aux États-Unis, au Canada et en France comme traduction normative et politique des méthodes d'évaluation environnementale développée à cette époque dans le champ scientifique, dont l'analyse de cycle de vie et, plus tard, l'empreinte écologique. D'autres auteurs ont adopté une lecture davantage conceptuelle et moins normative de l'origine de la gouvernance mondiale de l'environnement, en l'associant à la diffusion des méthodes de comptabilité économique et financière dans les sciences de l'environnement pour certains (RICHARD 2012), ou à une "mise en modèle" (ou scénarisation) de l'action publique pour d'autres (DAHAN 2007).

Initiée par la Conférence des Nations Unies de Stockholm en 1972, la gouvernance mondiale de l'environnement s'est rapidement traduite dans des cadres d'action publique, dont celui de *développement durable* proposé par la commissaire norvégienne Gro Harlem Brundtland en 1987 (WCED 1987) et qui s'est concrétisé dans des objectifs de politique publique par la Déclaration de Rio en 1992,. Le développement durable s'institutionnalisera encore davantage avec le cadre des *Objectifs du Millénaire pour le Développement* (OMD) qui est proposé en 2000 lors de la Déclaration du millénaire des Nations Unies et qui devient en 2015 le cadre de l'*Agenda 2030* où les huit OMD seront repris et étoffés en dix-sept *Objectifs de Développement Durable*.

L'adoption en 2015 des Accords de Paris par les parties de la Convention cadre des Nations Unies sur le climat est un exemple notoire de traduction directe du champ scientifique dans le champ de la gouvernance internationale. En effet, au-delà de constituer le prolongement du Protocole de Kyoto de 1997 qui arrivait à échéance en 2020, les Accords de Paris modifient en profondeur la façon même de concevoir l'impact des activités humaines sur le climat. Le Protocole de Kyoto engageait les États signataires à une réduction de leurs émissions de gaz à effet de serre par rapport à leurs émissions de 1990 (UNITED NATIONS 1997). Or choisir l'année 1990 comme référence est une décision arbitraire dont la pertinence environnementale est questionnable. Les travaux du GIEC ont en effet montré qu'un seuil de hausse des températures moyennes de surface gagnerait à être fixé à +2°C voire +1.5°C par rapport à l'ère pré-industrielle pour éviter d'enclencher des effets désastreux et incontrôlables sur l'ensemble du système Terre (IPCC 2018). Les Accords de Paris appellent ainsi à la responsabilité des États pour que les *contributions déterminées au niveau national* de réduction des émissions respectent, collectivement, le budget carbone global associé au seuil de hausse des températures recommandé par le GIEC (UNITED NATIONS 2015).

Institutionnaliser les limites planétaires : plusieurs dynamiques de portage

Dans cette histoire scientifique et institutionnelle, le cadre des limites planétaires cherche encore sa place, comme en témoignent ces propos de Johan Rockström, l'un des concepteurs du cadre, en préface de la dernière actualisation du cadre en septembre 2024 :

Notre but est d'opérer un changement radical dans la fréquence, la synthèse et la profondeur des recherches en cours sur la science des limites planétaires — en

passant d'une mise-à-jour [du cadre des limites planétaires] tous les 6 à 8 ans à des rapports annuels. Cela nous permettra d'intégrer les dernières données d'observation de la Terre et de produire des analyses actualisées. [...] À l'avenir, notre objectif est de traduire la science en action, allant de la gouvernance des limites planétaires à l'orientation des investissements dans la finance et les entreprises. Nous sommes en mission scientifique pour développer un tableau de bord des limites planétaires — une salle de contrôle pour la planète Terre.

Traduit de la préface de Johan Rockström à (CAESAR et al. 2024).

L'affiliation des limites planétaires à un unique champ scientifique n'est pas aisée, puisqu'elles relèvent à la fois des sciences du système Terre pour ce qui est du cadre global et des sciences de l'évaluation environnementale pour ce qui est de ses applications à des échelles locales : un champ nouveau se dessinerait-il à l'interstice de ces deux champs-ci, que Johan Rockström appelle la science des limites planétaires (*planetary boundaries sciences*) ? Le cadre des limites planétaires ne constitue en effet pas un modèle à part entière, notamment parce que les différentes variables qui servent à contrôler chaque limite ne sont pas liées entre elles par un modèle d'interaction global. Chaque limite ne correspond d'ailleurs pas nécessairement à un modèle, bien qu'au fil des révisions du cadre, on puisse identifier une tendance à la généralisation de l'utilisation des modèles pour la quantification des limites, ce qui n'était pas le cas dans la version originelle de (ROCKSTRÖM et al. 2009a). Apparaît ici tout l'enjeu de la méthode de quantification, puisque l'extrait mentionne également l'ambition "d'intégrer les dernières données d'observation de la Terre" : les limites planétaires peuvent à la fois être quantifiées par les résultats de modèles globaux, par la lecture des "données d'observation de la Terre" ou bien encore par une approche mixte dans laquelle les données d'observation sont intégrées aux modélisations.

Le cadre des limites planétaires n'a pas encore tout à fait sa place dans la gouvernance internationale de l'environnement, puisqu'il cherche encore les moyens de "se traduire en action". À cet égard, plusieurs dynamiques de portage du cadre des limites planétaires se sont dessinées depuis son portage initial en 2009 par le *Stockholm Resilience Center*, qui se présente comme un centre de recherche partenarial entre l'Université de Stockholm et le *Beijer Institute of Ecological Economics* rattaché à l'Académie suédoise royale des sciences. On peut distinguer entre un portage académique en dehors de la communauté des sciences du système Terre, un portage mixte entre des organismes académiques et des organismes non académiques, et un portage non académique.

Un portage académique des limites planétaires en dehors des sciences du système Terre

L'un des premiers portages du cadre des limites planétaires a été sa traduction dans le champ scientifique de l'évaluation environnementale, historiquement positionné sur l'étude de la soutenabilité des systèmes à une échelle locale. Le champ de l'évaluation environnementale y a en effet vu l'opportunité d'un renouvellement méthodologique, du moins, sinon paradigmatique, dans la possibilité d'une approche *absolue* généralisable à un panel large de catégories d'impacts (impacts sur le climat, l'eau, les sols, les nutriments, etc). L'idée de l'approche absolue est d'adosser toute évaluation des impacts environnementaux à une référence environnementale basée sur un état objectif et souhaitable du système Terre (BJØRN et al. 2013). L'*espace de fonctionnement sûr* constitue ainsi une référence absolue. Cette idée a rapidement débouché sur le développement d'un ensemble de méthodes de quantification pour parvenir à intégrer les seuils environnementaux du cadre des limites planétaires dans des indicateurs adaptés aux pratiques et méthodes habituelles du champ, en particulier l'analyse de cycle de vie et les analyses entrées-sorties environnementales (HEIJUNGS et al. 2014; BJØRN et al. 2015b; RYBERG et al. 2018; XUE et al. 2022).

La diversité de ces méthodes a pu être présentée comme autant de propositions méthodologiques intégrées au champ de l'*évaluation absolue de la soutenabilité environnementale* (AESA) (CHANDRAKUMAR et al. 2018; BJØRN et al. 2019). Il est d'autant plus pertinent de parler de champ que ces démarches d'intégration des limites planétaires dans l'évaluation environnementale ont d'emblée posé des questions d'ordre conceptuel, dont celle de l'échelle pertinente du cadre. Si l'*espace de fonctionnement sûr* se définit par rapport à l'état stable des mécanismes globaux de régulation du système Terre, on peut envisager que des mécanismes environnementaux locaux soient également à l'œuvre avec, chacun, leur seuil de stabilité. Les mécanismes globaux qui sont décrits par le cadre peuvent tout à fait être considérés comme d'abord locaux puis agrégés à l'échelle globale ou bien comme fondamentalement globaux? Par exemple, la ressource en eau globalement disponible pour les activités humaines peut être vue comme la somme des quantités localement disponibles une fois pris en compte les besoins écosystémiques de chaque bassin versant. Et inversement, les émissions de gaz à effet de serre peuvent être vues comme une pression qui ne fait sens qu'à l'échelle globale, c'est-à-dire une fois les gaz mélangés et dégradés dans l'atmosphère. Pour le champ de l'évaluation environnemen-

tale, cette question est celle de l'échelle à laquelle se définit la capacité de charge du système étudié. Aussi, très nombreux sont les auteurs à avoir réfléchi sur ce qu'ils appellent un *principe d'assignation*, qui permet de définir à l'échelle locale une capacité de charge initialement définie au niveau global (HJALSTED et al. 2021 ; RYBERG et al. 2020 ; HEIDE et al. 2023). Cette question de l'échelle de la capacité de charge va bien au-delà des questions méthodologiques sur les outils de quantification mobilisés, et a d'ailleurs questionné la communauté des sciences du système Terre, puisque dès la première actualisation du cadre des limites planétaires en 2015, les limites sont à la fois définies globalement et à une échelle locale jugée pertinente (typiquement le biome pour l'affectation des sols ou le bassin versant pour l'eau de surface) (STEFFEN et al. 2015b).

Outre ces considérations au sein du champ de l'évaluation environnementale, mentionnons l'intérêt porté au cadre des limites planétaires par le champ de l'économie hétérodoxe, en particulier la branche qui s'intéresse aux indicateurs alternatifs de richesse. L'équipe de recherche d'OXFAM Grande-Bretagne a en effet rapidement trouvé une traduction du cadre à des échelles locales dans ce qu'il convient désormais d'appeler l'*économie du donut* (RAWORTH 2017). Ce terme regroupe un ensemble de travaux visant à associer aux indicateurs environnementaux des limites planétaires, qui définissent l'*espace de fonctionnement sûr* pour l'humanité, des indicateurs socio-économiques (alternatifs à la seule richesse) qui définissent un *espace de fonctionnement juste*. L'articulation de l'espace juste et de l'espace sûr est schématiquement représentée par un *donut* (voir Figure 0.3), dans lequel l'anneau inférieur correspond aux indicateurs du *plancher social* (espace juste) et l'anneau supérieur à ceux du *plafond environnemental* (espace sûr) (RAWORTH 2012).

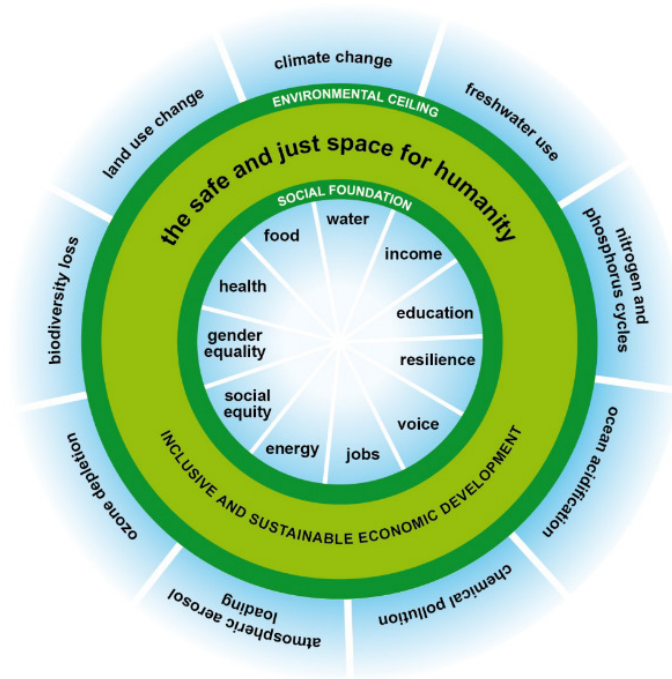


Figure 0.3 – La représentation de l’espace de fonctionnement sûr et juste par le *donut* (RAWORTH 2012)

Un portage mixte académique/non académique des limites planétaires

Une série d’organisations mixtes publiques-privées ont récemment vu le jour pour assurer le portage du cadre des limites planétaires en dehors de la sphère académique, à l’interface entre les mondes de la recherche, de l’entreprise et de l’action publique. On peut notamment citer la *Earth Commission*, qui se présente comme une coalition internationale de "scientifiques de la soutenabilité" et qui est l’organe académique du *Global Commons Alliance*, un consortium international qui mêle scientifiques, entreprises, organisations philanthropiques et représentants de la société civile. La *Earth Commission* est notamment à l’origine de l’une des deux révisions du cadre des limites planétaires en 2023 (ROCKSTROM et al. 2023). On peut également souligner la création récente d’un organisme qui porte le nom du champ scientifique que Johan Rockström appelle à créer, le *Planetary Boundaries Sciences*, qui est un organisme partenarial entre le centre de recherche allemand *Potsdam Institute for Climate Impact Research* (PIK), dont Johan Rockström est le directeur, et le consortium *Planetary Guardians* qui se présente comme un regroupement d’activistes environnementaux, de grands chefs d’entreprises et d’anciens chefs d’États. C’est dans ce contexte institutionnel qu’a été produite la dernière révision du cadre de septembre 2024 (CAESAR et al. 2024). Si ces organisations ne font pas encore partie de la gouvernance internationale de l’environnement, elles ambitionnent d’en orienter l’action et la direction.

Un portage non académique des limites planétaires

Enfin, le cadre des limites planétaires a fait l'objet de diverses traductions dans des contextes non académiques. En France, le cadre est à notre connaissance surtout porté par les acteurs de l'aménagement du territoire. La Fédération nationale des agences d'urbanisme y voit par exemple une façon de briser les silos de l'action publique en matière environnementale, où les enjeux environnementaux sont abordés indépendamment les uns des autres (FNAU 2024). Le commissariat général au développement durable a montré son intérêt pour la question dans une première tentative d'évaluation environnementale de la France par rapport aux limites planétaires en 2019 (MTES 2019), qui s'est affinée en 2023 et a fait l'objet d'une étude à part entière (MTES 2023). La ville de Grenoble, quant à elle, a trouvé dans la représentation schématique du donut un outil de participation citoyenne et de communication davantage engageant et stimulant pour la publication de son rapport annuel de développement durable (VILLE DE GRENOBLE 2022). Depuis 2021, le syndicat mixte en charge du pilotage du Schéma de cohérence territoriale (SCoT) du Sud-Loire, aux côtés d'Epures, l'agence d'urbanisme des territoires ligériens, s'est proposé comme territoire d'étude pour la mise en œuvre du cadre des limites planétaires dans les documents de planification. Cette expérimentation est conduite avec le laboratoire Environnement Ville Société de l'école des Mines de Saint-Etienne, de telle sorte que si le portage de l'initiative est bien assuré par des acteurs non académiques, son déploiement prend toutefois une forme partenariale académique-non académique.

Les limites planétaires sur le territoire du SCoT Sud-Loire

Un SCoT (Schéma de cohérence territoriale) est un type de territoire un peu particulier dans les différents échelons administratifs de la politique décentralisée de l'État en France. D'abord parce que les SCoT sont un échelon relativement récent que la loi Solidarité et renouvellement urbains (SRU) a créé en 2000. Ensuite parce qu'il s'agit d'un échelon administratif qui ne correspond à aucune base électorale directe : la gouvernance d'un SCoT n'est pas assurée par une représentation élue directement par les citoyens (comme les communes, les départements ou les régions), mais par un ou plusieurs établissements publics de coopération intercommunale (EPCI) selon l'envergure de son périmètre, généralement défini comme *bassin de vie* ou *bassin d'emploi*. Les EPCI en assurent le pilotage directement ou par l'intermédiaire d'un syndicat mixte dont la direction est nommée par un mode de suffrage représentatif des différents EPCI

constitutifs. Enfin, parce que le SCoT est essentiellement un "dragon de papier"¹, c'est-à-dire un ensemble de documents de planification qui s'imposent aux communes de son périmètre pour la définition de leur plan local d'urbanisme (PLU) ainsi qu'aux documents intercommunaux comme le plan local d'urbanisme intercommunal (PLUi), le plan local de l'habitat (PLH) ou le plan de déplacement urbain (PDU).

Le SCoT doit permettre d'articuler ces documents communaux avec les documents de planification supérieurs aux communes et intercommunalités, dont le schéma régional d'aménagement, de développement durable et d'égalité des territoires (SRADDET) qui est porté à l'échelle régionale, le schéma directeur d'aménagement et de gestion des eaux (SDAGE) qui est porté à l'échelle des grands bassins hydrographiques, et le schéma d'aménagement et de gestion des eaux (SAGE) qui est porté à l'échelle des sous-bassins hydrographiques. Il peut aussi porter d'autres initiatives, comme les plans climat air énergie territoire (PCAET) ou les plans de paysage.

L'ordonnance de modernisation des SCoT du 17 juin 2020 (loi ELAN) définit ainsi les SCoT par trois documents :

- un projet d'aménagement stratégique (PAS) qui rend compte de la vision prospective des élus du territoire sur un horizon de 20 ans ;
- un document d'orientation et d'objectifs (DOO) qui définit des objectifs localisés ou chiffrés sur les thématiques (1) du développement économique, agricole et commerciale, (2) du logement, des mobilités, des équipements et services, (3) des transitions écologique et énergétique, de la préservation des ressources naturelles ;
- des annexes, qui comprennent notamment un diagnostic territorial et un état de l'environnement.

Pour ce qui est du SCoT Sud-Loire (voir Figure 0.4), il regroupe quatre EPCI pour un total de 198 communes : Saint-Etienne Métropole (406 000 habitants), Loire-Forez Agglomération (113 000 habitants), la communauté de communes de Forez-est (65 000 habitants) et la communauté de communes des Monts du Pilat (15 000 habitants). Le SCoT Sud-Loire a vu le jour en 2013, où il ne comptait que 109 communes, et a initié une phase de révision en 2018, pour laquelle le comité syndical a souhaité s'inscrire dans le cadre de la loi

1. Expression que nous reprenons à un membre du syndicat mixte du SCoT Sud-Loire lors d'une réunion syndicale à laquelle nous étions conviés en tant qu'observateurs.

de modernisation de 2020. Les documents constitutifs du nouveau SCoT sont en préparation depuis 2021 : le PAS a été validé en 2022 et le DOO est en cours de construction depuis 2023.

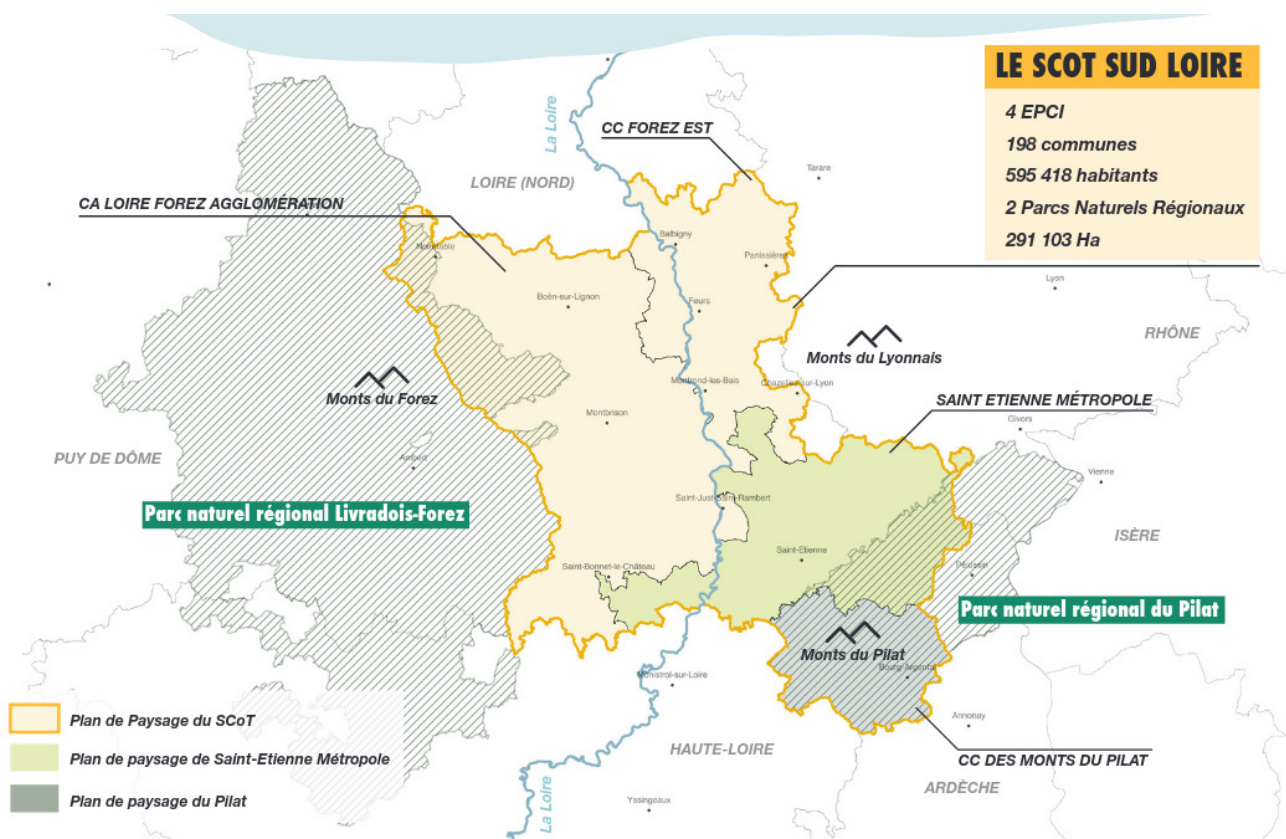


Figure 0.4 – Description géographique du SCoT Sud-Loire (SYNDICAT MIXTE DU SUD-LOIRE 2022)

C'est dans ce contexte de révision du SCoT que s'inscrit le présent travail de thèse de Doctorat, qui a initialement débuté en 2021 sous la forme d'un stage de fin d'étude financé par Epures. Si la question a demeuré inchangée (*le territoire du Sud-Loire dépasse-t-il ses limites?*), l'approche a évolué d'une perspective résolument opérationnelle (quantifier les pressions environnementales exercées par le territoire vis-à-vis des limites planétaires) vers une perspective davantage réflexive (que veut dire quantifier une limite planétaire à une échelle territoriale?). La première perspective a donné lieu à la publication (EPURES 2021), dont on peut également trouver un aperçu plus concis dans la partie dédiée aux applications territoriales du cadre des limites planétaires du rapport du Commissariat général au développement durable sur la contribution de la France aux limites planétaires (MTES 2023).

Hypothèses directrices du travail de thèse

Le projet de thèse a modifié la géométrie du portage du sujet, puisque la thèse a été financée par le programme Institut Convergences du Commissariat Général à l'Investissement (École Urbaine de Lyon), puis repris par l'Université de Lyon (UdL) suite à l'arrêt du financement². Cette modification du portage a transformé le contexte initial de *commande territoriale* vers un contexte que l'on pourrait qualifier d'*intervention territoriale*. En effet, les acteurs du SCoT Sud-Loire se sont portés volontaires pour "faire laboratoire" sans envisager une intégration directe des résultats de recherche dans la révision des documents du SCoT. Ce contexte d'intervention s'est traduit par l'invitation à de nombreuses réunions syndicales, comités techniques et comités de pilotage, auxquels nous avons assisté en tant qu'observateurs, passifs la plupart du temps, ou participants lorsqu'une invitation à présenter nos travaux nous était faite. Aussi le territoire sur lequel nous travaillons est-il davantage à envisager comme un cas d'étude en situation partenariale plutôt que comme commanditaire. Ce contexte s'est révélé particulièrement heuristique pour la formulation des hypothèses de recherche du travail de thèse.

Le travail initial réalisé sur le territoire (EPURES 2021) a notamment mis en avant la défaillance de la donnée environnementale pour la quantification des pressions territoriales au regard de la version (STEFFEN et al. 2015b) du cadre des limites planétaires. Si certaines limites planétaires telles que le changement climatique ou les émissions de particules fines sont spontanément quantifiables par les données du système d'information environnementale disponible sur le territoire (bases de données territoriale, départementale, régionale ou nationale), pour d'autres limites en revanche, telles que la perturbation des cycles de l'eau douce, de l'azote et du phosphore, les données disponibles au niveau territorial ne permettent pas une quantification directe des variables de contrôle correspondantes. En effet, en proposant des indicateurs alternatifs à ceux qui sont couramment en usage dans la politique environnementale (par exemple, s'intéresser au débit de réserve écologique ou aux quantités de phosphore non métabolisées par les cultures), le cadre des limites planétaires rend compte de ce que la production de la donnée environnementale est politiquement orientée, c'est-à-dire produite pour les besoins des indicateurs de gestion des milieux (DASSIBAT 2022). La publication initiale (EPURES 2021) s'était ainsi arrêtée à une qualification de la donnée selon une grille de fiabilité subjective : fiabilité bonne, moyenne ou faible (voir Figure 0.5). On peut dans ces conditions s'interroger sur la pertinence de construire une politique environnementale quantifiée sur la

2. Voir : <https://www.rue89lyon.fr/2022/04/27/pourquoi-nous-contestons-arret-brutal-ecole-urbaine-de>

base d'indicateurs dont la fiabilité n'est jugée que moyenne voire faible.









Limites planétaires		Echelle planétaire	Echelle Sud Loire	Fiabilité des données mobilisées
	Le changement climatique	Zone d'incertitude	Niveau dangereux	Bonne
	L'érosion de la biodiversité	Niveau dangereux	Zone d'incertitude	Moyenne
	La perturbation du cycle de l'eau douce	Espace de sécurité	Zone d'incertitude	Faible
	La perturbation du cycle du phosphore	Niveau dangereux	Non évalué	Faible
	La perturbation du cycle de l'azote	Niveau dangereux	Zone d'incertitude	Moyenne
	L'atteinte de la couche d'ozone	Espace de sécurité	Espace de sécurité	Bonne
	La diminution du couvert végétal	Zone d'incertitude	Non évalué	Moyenne
	L'augmentation des particules en suspension dans l'air	Non évalué	Espace de sécurité à niveau dangereux	Bonne

Figure 0.5 – Tableau bilan de l'évaluation environnementale du SCoT Sud-Loire conduite au regard du cadre des limites planétaires (EPURES 2021)

Si l'intérêt porté au cadre des limites planétaires par le syndicat mixte en charge du SCoT Sud-Loire dépasse la seule question de la bonne quantification des pressions environnementales, y voyant un intérêt réflexif sur les enjeux à prendre en compte dans les documents de planification et une stimulation nouvelle pour l'animation territoriale du SCoT (le chiffrage des objectifs n'est d'ailleurs pas le rôle premier du SCoT), cette question d'une quantification fiable demeure centrale dans une perspective scientifique d'évaluation environnementale, et en particulier d'évaluation environnementale absolue.

La question de la donnée environnementale est elle-même plus large que la question de la bonne quantification d'un indicateur. Les nombreux échanges que nous avons eus avec les acteurs du SCoT (en particulier les chargés de mission d'Epures et du bureau du syndicat mixte) ont révélé un fort besoin de connaissances environnementales du territoire pour stimuler et épauler l'animation territoriale auprès des élus et des acteurs du territoire (représentants de syndicats professionnels, représentants de services décentralisés de l'État, représentants de la société civile). Ce besoin de connaissances passe en partie par la disponibilité de la donnée environnementale, c'est-à-dire sa production, sa mise en forme et sa diffusion, avant même de songer à son élaboration dans des indicateurs de soutenabilité absolue. La ressource en

eau a souvent été mentionnée à cet égard. Décrite comme un enjeu majeur de préoccupation pour l'avenir, notamment dans la révision du PAS, l'eau est aussi un enjeu dont le SCoT doit juridiquement organiser l'animation sur son territoire, en s'assurant de la compatibilité entre les SDAGE et SAGE et les orientations du DOO. Or la donnée environnementale sur l'eau n'est pas produite par les animateurs du syndicat mixte du SCoT, mais d'abord par les services de l'Etat (notamment les Directions régionales de l'environnement, de l'aménagement et du logement) et, dans une moindre mesure, par les communes et EPCI au titre de l'exercice de leur compétence eau potable et eaux usées, ainsi que les comités de bassin au titre de la politique environnementale de l'eau qu'ils animent. Il y a ainsi une déconnexion pour le SCoT entre l'échelle où se produit la donnée et l'échelle où elle se mobilise.

Ce travail de thèse s'articule ainsi autour d'une hypothèse directrice sur le rôle de la donnée environnementale dans la territorialisation du cadre des limites planétaires. Dans la perspective territoriale où nous nous situons, le terme *territorialisation* renvoie au couple compétence/connaissance environnementale du territoire. En effet, la disponibilité de la donnée environnementale sur des enjeux portés par le cadre des limites planétaires et pour lesquels le territoire n'est pas compétent ou ne dispose pas des éléments de connaissance peut être vue comme une première modalité de territorialisation du cadre. Dans la perspective scientifique où nous nous situons, celle du champ de l'AESA, le terme *territorialisation* renvoie spontanément à des questions d'échelle géographique de la capacité de charge, comme nous l'évoquons plus haut, mais aussi à un choix méthodologique auquel toute évaluation environnementale (absolue ou non) est confrontée : celui du périmètre géographique. Parler d'*approche territoriale* en évaluation environnementale renvoie nécessairement à la distinction entre un périmètre géographique *orienté production* et un périmètre *orienté consommation* (LOISEAU et al. 2018). Un territoire est en effet un système qui génère des interventions anthropiques sur l'environnement soit au titre des biens et services produits à l'intérieur de ses frontières géographiques (le périmètre géographique est alors restreint à ces frontières), soit au titre de la consommation finale de ses agents économiques résidents, qui est en partie satisfaite par la production territoriale et en partie par des importations (le périmètre géographique s'étend alors aux différents lieux de production des biens consommés sur le territoire).

L'objet *territoire* permet ainsi de confronter le cadre des limites planétaires aux trois objets qui définissent une évaluation environnementale (SALA et al. 2015) : le contexte décisionnel (le niveau de compétence/connaissance du territoire en matière d'environnement), l'approche de

soutenabilité (la définition de la capacité de charge), et les choix méthodologiques (production ou consommation). Dans l'opération de mise en chiffres, ou de quantification, de ces trois objets intervient nécessairement la question de la disponibilité des données environnementales qui, pourtant, demeure sous-documentée dans la littérature disponible en AESA. Nous faisons ainsi l'hypothèse centrale qu'il ne se joue pas dans l'AESA d'un système territorial une unique question d'échelle géographique ni de périmètre géographique, mais aussi une question de source de la donnée environnementale.

Organisation du manuscrit

Ce travail de thèse entend ainsi proposer une réflexion sur le rôle de la donnée environnementale dans une approche territoriale des limites planétaires, en s'appuyant sur le cas d'étude de l'intégrité des cours d'eau sur le territoire du SCoT Sud-Loire. Le manuscrit s'organise pour cela en quatre chapitres.

Le chapitre 1 présente en détail le cadre des limites planétaires, à la fois dans ses développements à l'échelle globale et dans ses traductions aux échelles locales, en insistant sur les travaux portés par le champ de l'AESA. Ce chapitre se terminera par une formulation plus précise des hypothèses de recherche qui articulent la question centrale de cette thèse : comment faire coïncider la donnée environnementale avec l'échelle de l'évaluation AESA ?

Le chapitre 2 abordera cette question en examinant la littérature qui porte sur la limite planétaire relative à l'eau de surface. Nous chercherons à comprendre comment l'eau est quantifiée dans les travaux portant sur le cadre global et dans ceux qui portent sur des applications locales. Nous aboutirons alors à une grille analytique permettant de lier la question de *l'échelle spatiale de l'étude* à ce que nous appellerons *l'échelle de la donnée environnementale*. Une telle échelle de la donnée entend souligner l'importance de prendre en compte l'échelle spatiale à laquelle est produite la donnée, au-delà de la seule échelle spatiale du système étudié, ainsi que la qualité et la robustesse de la donnée, c'est-à-dire sa capacité à décrire le plus explicitement possible et avec le plus de fiabilité le phénomène environnemental quantifié. Nous proposerons alors de définir une approche territoriale des limites planétaires comme la situation où l'échelle de l'étude et l'échelle de la donnée environnementale coïncident.

Nous poserons alors deux questions de recherche :

1. Comment effectuer l'évaluation environnementale de l'intégrité des cours sur un territoire au regard de l'indicateur proposé par (ROCKSTROM et al. 2023) à partir de mesures de débits sur site systématisées ? Cette proposition étudie la faisabilité d'un indicateur AESA qui assure la cohérence méthodologique entre l'échelle de l'étude et l'échelle de la donnée à partir des données classées comme les plus spatialement explicites et robustes.
2. Comment coupler des données issues de mesures de débits sur site systématisées avec des données issues des sorties d'un modèle hydrologique global ? Autrement dit, comment coupler des données spatialement explicites avec des données non spatialement explicites mais spatialement différenciées ? Cette proposition étudie la faisabilité d'une assimilation entre sources de données différentes et étudie ainsi la possibilité de combiner des degrés différents d'approche *bottom-up*.

Les chapitres 3 et 4 constituent une proposition pour la mise en œuvre d'une approche d'AESA territoriale ainsi définie. Nous présenterons au chapitre trois deux méthodes de couplage entre les données environnementales sur l'eau et l'évaluation absolue de l'intégrité des cours d'eau. Une première méthode envisage de construire une méthode d'AESA par des données de mesures de débits sur site. Une seconde méthode propose un modèle d'assimilation des données de mesures de débits sur site à des données de modèle hydrologique global. Le chapitre quatre présentera les résultats obtenus selon ces deux méthodes.

Chapitre 1

Le cadre des limites planétaires : du système Terre au territoire

Introduction

Ce chapitre présente l'objet de recherche de ce travail de thèse, le cadre des limites planétaires, et son interaction avec l'échelle territoriale. Il reprend la structure et le contenu de l'article (DASSIBAT et al. 2024) qui a été écrit dans le cadre de ce travail et publié aux éditions *Techniques de l'Ingénieur*. Cet article a été rédigé dans un souci de vulgarisation et trouve ainsi naturellement sa place en introduction du présent manuscrit. Il nous permettra, dans une partie conclusive (4), de dégager les tensions propres à la déclinaison territoriale du cadre, et ainsi d'introduire les questions de recherche qui guideront les chapitres suivants.

La partie (1) de l'article commence par rappeler l'esprit du cadre des limites planétaires avant de passer en revue ses différentes versions, depuis sa publication originale en 2009 (ROCKSTRÖM et al. 2009a) jusqu'à ses deux dernières actualisations de 2023 (ROCKSTROM et al. 2023 ; RICHARDSON et al. 2023), en date au moment de la rédaction de l'article. À cette occasion, une description des mécanismes environnementaux associés à chacune des limites constitutives de la version (ROCKSTROM et al. 2023) du cadre sera donnée, permettant au lecteur de mieux appréhender le cadre dans sa globalité avant d'aborder la question de sa "descente d'échelle".

Les parties (2) et (3) de l'article abordent la déclinaison locale du cadre. Le terme *local* est à prendre ici dans un sens très large, désignant à la fois un territoire, une région, un

bassin versant ou plus généralement toute emprise géographique, mais aussi une organisation, une infrastructure, une filière technologique, un produit, un service, ou plus généralement toute fonction réalisée par un système. La partie (2) met ainsi l'accent sur la grande diversité de déclinaisons possibles au niveau *local*, ne réduisant pas le spectre aux seules déclinaisons *territoriales*. Précisément, cette diversité de déclinaisons est constitutive d'un champ de recherche à part entière que le cadre des limites planétaires a fait émerger depuis les années 2015 : l'évaluation absolue de la soutenabilité environnementale (AESA). Non seulement les applications sont-elles variées en termes de systèmes étudiés et d'échelles géographiques, mais également en termes de méthodologies d'évaluation environnementale, en particulier pour la construction des indicateurs basés sur les limites planétaires. C'est la raison pour laquelle la partie (2) entend moins constituer un guide pratique de la descente d'échelle que souligner l'émergence d'un champ de recherche nouveau pour l'évaluation environnementale, dont l'ambition d'adosser toute évaluation à une référence environnementale absolue est porteur d'un renouvellement paradigmatique fort au sein du champ.

La partie (3), par contraste avec le ton généraliste de la partie précédente, se propose d'évoquer des exemples de déclinaisons locales des limites planétaires, à la fois dans le cadre formel et normatif de l'AESA comme en dehors. Cette partie entend montrer que le cadre des limites planétaires n'est pas limité à l'évaluation environnementale, c'est-à-dire à la quantification des pressions anthropiques, mais peut aussi bien servir de référentiel, ou encore de cadre de pensée, dans une approche plus qualitative et, sans doute, plus facilement actionnable dans les organisations ou les politiques publiques.

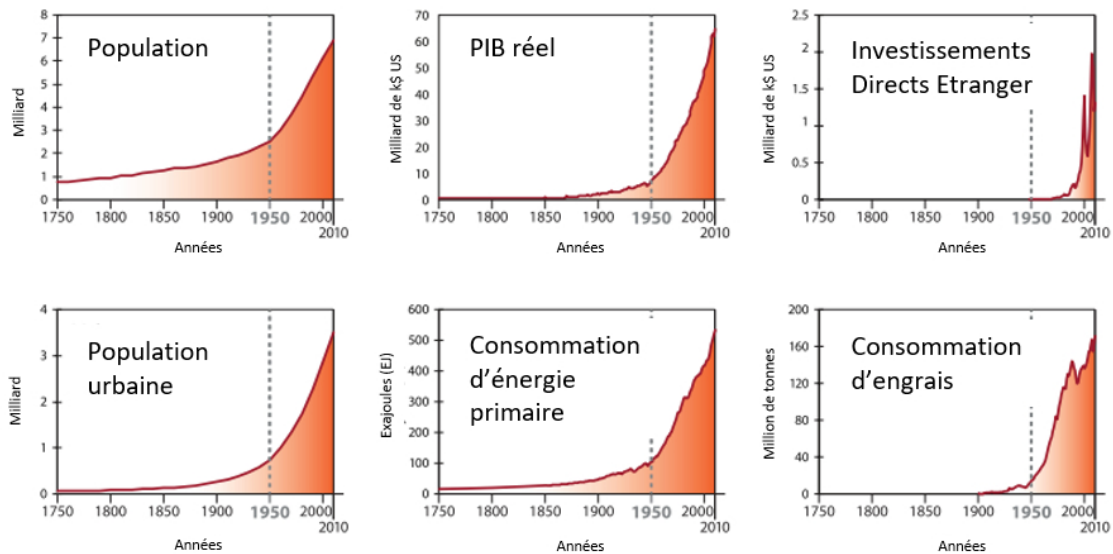
1.1 Le cadre global des limites planétaires

1.1.1 L'esprit du cadre : préserver un état environnemental de référence pour le système Terre

Le cadre des limites planétaires propose d'évaluer les impacts des activités humaines à l'échelle planétaire par rapport à une référence environnementale qui a été établie sur l'équilibre climatique de l'Holocène, ère climatique que l'Anthropocène tend à remplacer par l'influence des activités humaines sur les équilibres biogéophysiques. Les limites planétaires sont un cadre d'évaluation qui établit plusieurs seuils environnementaux à ne pas dépasser pour éviter de déstabiliser l'équilibre climatique actuel du système Terre. Ces limites sont mises à

mal par l'expansion de l'homme sur Terre depuis le début de l'Holocène, il y a un peu plus de 10 000 ans. Si certains travaux font remonter au Néolithique voire au Paléolithique les prémises d'une humanité qui deviendra force géologique (c'est ce que le terme d'Anthropocène recouvre (LUSSAULT et al. 2021)), la période d'après-guerre marque une rupture dans l'ampleur des impacts anthropiques sur l'environnement. Cette période est caractérisée de "Grande Accélération" par Will Steffen et ses collègues (STEFFEN et al. 2015a). Ce terme désigne l'évolution croissante des pressions anthropiques sur la période 1750-2000 (voir Figure 1.1). Cela va dans le sens de l'intuition formulée par le chimiste américain Paul Crutzen, au début des années 2000, pour qui l'humanité entraînerait la sortie progressive de la planète de son équilibre climatique de l'Holocène, la faisant entrer dans ce qu'il qualifie d'Anthropocène (CRUTZEN 2002; CRUTZEN et al. 2000).

Indicateurs socioéconomiques



Indicateurs du système Terre

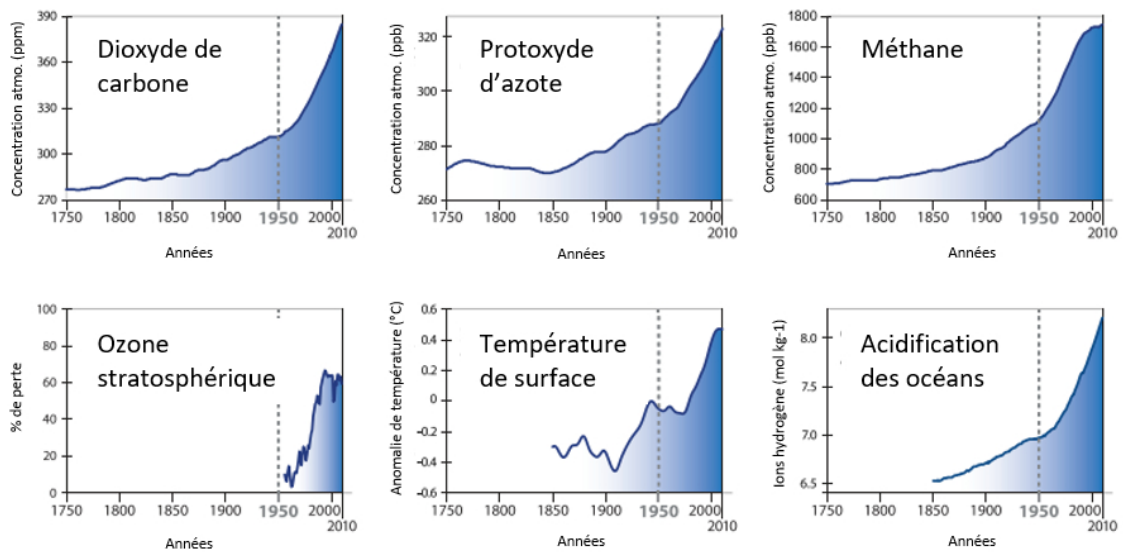


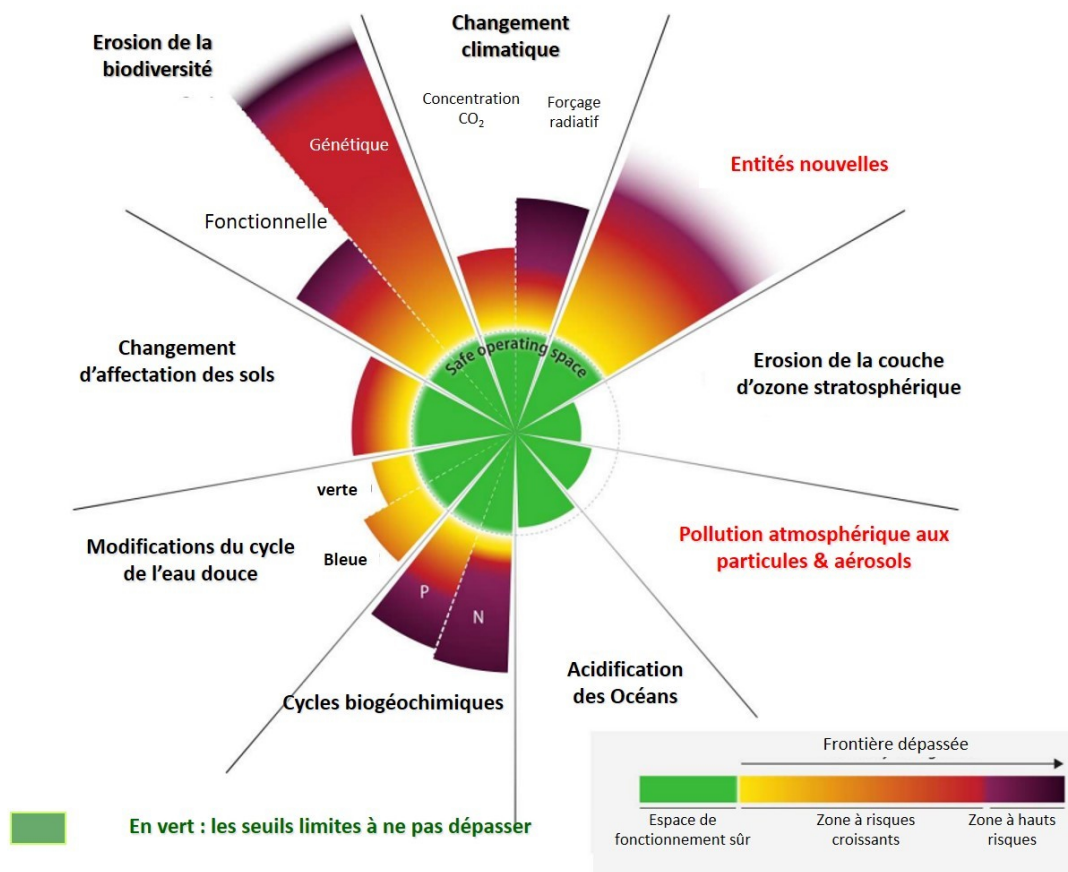
Figure 1.1 – Les grandes tendances socioéconomiques depuis le début de l'ère industrielle (1750) au regard de l'évolution des variables de fonctionnement du système Terre. (Traduit de (STEFFEN et al. 2015a))

L'Holocène s'est en effet révélé comme étant l'ère la plus propice à l'expansion de l'espèce humaine : "L'époque relativement stable de 11 700 ans qu'est l'Holocène est le seul état du système Terre dont nous sommes certains qu'il puisse supporter les sociétés humaines contemporaines" (STEFFEN et al. 2015b).

Les limites planétaires entendent ainsi déterminer des seuils environnementaux censés garantir la pérennité de cette ère climatique. Le respect des limites planétaires est présenté par les auteurs des articles fondateurs comme le plus grand défi auquel l'humanité n'ait jamais été confrontée, celui de maintenir ses conditions d'habitabilité.

Le cadre des limites planétaires apparaît également comme un cadre de référence qui précise ce que l'on entend communément par "préserver la nature" : il s'agit plus spécifiquement de maintenir les conditions d'habitabilité pour l'homme sur Terre (BJØRN et al. 2019). C'est dans cette logique que la notion d'*espace de vie préservé* ou *espace de fonctionnement sûr*¹ a été proposée (voir le corridor vert dans la Figure 1.2. Le respect des limites planétaires est en effet censé garantir à l'humanité d'évoluer dans un cadre préservé pour assurer la stabilité des équilibres biogéophysiques (ROCKSTRÖM et al. 2009a). On peut voir une représentation de l'espace de fonctionnement sûr en Figure 1.2.

1. traduit de l'anglais *Safe Operating Space*) par (BOUTAUD et al. 2020)



Source : d'après Richardson et al., Nature, septembre 2023

Figure 1.2 – La représentation standard des limites planétaires en graphique radar issue de la version 2023 du cadre (Traduit de RICHARDSON et al. 2023).

Le cadre des limites planétaires comportait neuf limites, dès 2009, afin de proposer un mode systémique d'appréhension des interactions entre les différentes composantes du système Terre et celles du système humain. Trois limites sont en effet jugées déterminantes pour caractériser l'état de santé du système climatique : le changement climatique, l'érosion de la couche d'ozone et l'acidification des océans. Les autres limites sont vues comme autant d'accélérateurs ou de ralentisseurs de la dégradation des trois limites principales : les limites ont des effets de rétroaction positive ou négative les unes sur les autres. En 2009, trois limites étaient déjà dépassées (ROCKSTRÖM et al. 2009a). Lors de sa mise à jour en 2015, cinq étaient dépassées (STEFFEN et al. 2015b). En 2022, des mises à jour individuelles pour l'eau verte (WANG-ERLANDSSON et al. 2022) et l'introduction d'entités nouvelles dans l'environnement (PERSSON et al. 2022) ont été proposées, concluant au dépassement de chacune de ces deux limites. En septembre 2023, la dernière actualisation du cadre (RICHARDSON et al. 2023) fait état de six limites dépassées : le changement climatique, l'intégrité de la biosphère, le changement d'utilisation des sols, les changements de cycle de l'eau douce, les flux biogéochimiques de l'azote et du phosphore et

l'introduction d'entités nouvelles dans l'environnement.

Alors qu'autant de limites ont été dépassées, comment se fait-il que nous soyons encore là ? Que veut dire dépasser une limite ? Deux axes de réponses seront présentés ci-dessous.

En termes de mécanisme environnemental, cela veut dire que la pression exercée par les activités humaines sur une composante environnementale dépasse la capacité de l'environnement à supporter cette pression sans se dégrader substantiellement. Cette capacité de charge de l'environnement vis-à-vis d'une composante donnée est estimée par une revue exhaustive de la littérature, puis simulée avec des modèles environnementaux éprouvés et des jeux de données actualisés. Ce travail permet d'estimer le niveau de dégradation à partir duquel des "points de bascule" se mettent en place, modifiant alors de façon irréversible l'équilibre existant. Une incertitude entoure le positionnement exact du point de bascule. Afin de répondre au principe de précaution, la "frontière planétaire" (traduction plus rigoureuse de l'anglais "*planetary boundary*") est ainsi positionnée au seuil bas de l'espace de vie préservé, c'est-à-dire au niveau de la valeur la plus faible de la zone d'incertitude dans laquelle se situe la limite.

En termes de conséquences environnementales, dépasser une limite signifie augmenter considérablement la variabilité du système Terre. Par exemple, pour le changement climatique, plus le forçage radiatif est fort, plus les effets sur les processus de régulation climatique sont importants et non linéaires (tant en fréquence qu'en intensité). Cela entraîne, par exemple, une augmentation de la fréquence et l'intensité des pluies décennales (qui ne se produisent en moyenne qu'une fois tous les 10 ans). C'est notamment à cause de cette variabilité non linéaire accrue que les États signataires des Accords de Paris (2015), se sont accordés sur l'objectif d'un forçage à $+2^{\circ}\text{C}$ maximum (seuil qui présente deux avantages : il s'agit d'un chiffre "rond" et qui nous laisse encore quelques décennies pour agir). Remarquons d'une part qu'à ce niveau, les pluies décennales sont tout de même 1,7 fois plus fréquentes et contiennent 14% d'humidité en plus qu'en l'absence de forçage (IPCC 2022) ; et d'autre part, que le cumul des engagements actuels des États signataires conduit pour l'instant à une augmentation de température moyenne mondiale de $+2,6^{\circ}\text{C}$ (ARMSTRONG MCKAY et al. 2022). La conséquence de cette variabilité est que certaines zones pourront devenir quasiment invivables pour l'homme mais pas la totalité de la planète. Franchir une limite ne signifie pas la disparition de l'espèce humaine mais l'apparition de territoires au mieux inconfortables, au pire invivables qui la force à se redistribuer géographiquement (GEMENNE 2021).

A cette variabilité accrue s'ajoute un autre type de conséquence, la transition du système vers un autre état (vraisemblablement moins stable) que celui de l'Holocène. C'est le concept des "points de bascule" (*tipping points*) : lorsqu'une variable environnementale s'éloigne trop d'une valeur qui garantissait l'état d'équilibre du système (les "frontières"), cet équilibre se déplace. On distingue plusieurs façons pour une variable de s'éloigner de sa valeur stable : les effets de seuil, les évolutions aléatoires, et l'incapacité à se régénérer au rythme où une ressource s'épuise. Cette notion de point de bascule est très documentée en écologie, à l'échelle locale. En 2008, Lenton et ses collègues proposent de l'appliquer à l'échelle planétaire pour le changement climatique, en lui ajoutant une dimension "politiquement pertinente" : si la réaction du système-Terre vient déplacer son état d'équilibre suite à la variation critique d'une de ses composantes, que cette réaction risque de se perpétuer pendant le millier d'années qui suit la perturbation et que cette perturbation (restreinte dans leur étude à une anomalie de températures moyennes de surface exprimée en écart par rapport à la normale préindustrielle) intervient dans le siècle actuel, alors la composante est considérée comme un point de bascule (LENTON et al. 2008). Steffen et ses collègues distinguent plusieurs types de points de bascule, influençant les trajectoires à venir du système Terre dans le contexte de l'Anthropocène (STEFFEN et al. 2018). Ils évoquent à ce titre la trajectoire "*Hothouse Earth*" (Terre-serre) si tous les points de bascule venaient à être franchis.

1.1.2 Les évolutions d'un cadre toujours plus robuste

Le cadre en tant que tel est à l'initiative d'une vingtaine de chercheurs issus de différentes sciences pour l'environnement, rassemblés autour du *Stockholm Resilience Center* (SRC). Le SRC est un centre de recherche conjoint sur le management et la gouvernance des socio-écosystèmes, co-porté par l'Université de Stockholm, le *Beijer Institute of Ecological Economics* et la *Royal Swedish Academy of Sciences*. En 2009, ce groupe de chercheurs, emmenés par Johan Rockström propose une première version du cadre des limites planétaires.

Le travail des experts rassemblés autour du cadre des limites planétaires est itératif : il procède par actualisations successives à mesure que les recherches progressent dans les différents champs disciplinaires embrassés par les limites planétaires. Par exemple, le travail médiatisé de Wang-Erlandsson en 2022 sur le franchissement d'une nouvelle limite sur l'"eau verte" (WANG-ERLANDSSON et al. 2022), a permis d'enrichir la limite sur le cycle de l'eau douce, aux côtés de la perturbation humaine sur les cycles de l'eau dite "bleu", qui représente les eaux

superficielles et souterraines. L'eau dite "verte" correspond à l'eau disponible pour les plantes, présente dans les précipitations, dans les végétaux eux-mêmes et dans les sols. Le "proxy" choisi pour représenter l'impact des activités humaines sur la modification de ces cycles consiste à quantifier l'écart entre l'humidité du sol actuelle et son niveau préindustriel, sur la base de modèles mondiaux. Notons que l'actualisation du cadre, en 2023, a fait le choix d'abandonner les approches basées sur les "consommations" d'eau au profit d'une estimation des "altérations" générées par les activités anthropiques. Ainsi, la variable de contrôle proposée par Rockström (ROCKSTROM et al. 2023) et Richardson (RICHARDSON et al. 2023), repose sur l'écart entre le débit actuel et le débit préindustriel, s'appuyant en cela sur les travaux des chercheurs Poff et Zimmerman (POFF et al. 2010).

Le tableau 1 donne à voir l'évolution du cadre des limites planétaires, de 2009 à 2023, et propose de mettre en relation les limites proposées avec leur variable de contrôle. Pour chaque limite, le tableau renseigne les variables de contrôle utilisées par les publications successives. Le code couleur des cellules est le suivant :

- en rouge, limite franchie ;
- en vert, limite non franchie ;
- en jaune, limite située dans la "zone d'incertitude" (STEFFEN et al. 2015b) ou "zone de risque accru" (RICHARDSON et al. 2023). Pour (ROCKSTRÖM et al. 2009a ; ROCKSTROM et al. 2023), cette terminologie n'est pas valable) ;
- en gris, limite non définie ou non quantifiée ;
- en blanc, limite non prise en compte.

Limite planétaire (*)	ROCKSTRÖM et al. 2009a	STEFFEN et al. 2015b	ROCKSTRÖM et al. 2023	RICHARDSON et al. 2023
Changement climatique	(a) concentration en CO ₂ atmosphérique [ppmv]	(a) concentration en CO ₂ atmosphérique [ppmv]	Probabilité de franchir les points de bascule pour trois seuils de hausse des températures moyennes de surface par rapport à l'ère préindustrielle (1850-1900) (4)	(a) concentration en CO ₂ atmosphérique [ppmv]
	(b) forçage radiatif [W.m-2]	(b) forçage radiatif [W.m-2]		(b) forçage radiatif [W.m-2]
Perte de biodiversité	Taux d'extinction [nombre d'espèces/Million d'espèces/an]	(a) diversité génétique :taux d'extinction [nombre d'espèces/Million d'espèces/an]	(a) surfaces des écosystèmes : taux de perte d'espaces naturels contribuant à la régulation du climat, du cycle de l'eau, du et à la biodiversité.	(a) diversité génétique :taux d'extinction [nombre d'espèces/Million d'espèces/an]
		(b) diversité fonctionnelle : Biodiversity Intactness Index (BII)	(b) intégrité fonctionnelle : taux de perte des zones naturelles contribuant aux personnes (" <i>Nature's Contribution to People</i> ") par km ² de surface.	(b) diversité fonctionnelle :énergie disponible pour les écosystèmes estimée en productivité primaire nette [rapport entre l'appropriation humaine de la productivité primaire nette et la productivité primaire nette totale]
Cycle du phosphore	Quantité de P ruisselée vers les océans [Mt.an-1]	(a) Global : Quantité deP ruisselée vers les océans [Tg.an-1]	(a) Sous-global : risque d'eutrophisation directe des eaux de surface, concentration de P[mg m-3]	(a) Global : Quantité de P ruisselée vers les océans [Tg.an-1]
		(b) Régional : Quantité de P apporté par les fertilisants aux sols érodables [Tg.an-1]	(b) Global : risque indirect d'eutrophisation, surplus agricole [Tg P an-1];minage [Tg P an-1]	(b) Régional : Quantité de P apporté par les fertilisants aux sols érodables [Tg.an-1]
Cycle de l'azote	Quantité de N ₂ atmosphérique fixée par l'homme [Mt.an-1]	Quantité de N fixée industriellement ou biologiquement par intention [Tg.an-1]	(a) risque d'eutrophisation des écosystèmes terrestres et eaux de surface, concentration de N [mg N l-1] [kg N ha-1] respectivement (4)	Quantité de N fixée industriellement ou biologiquement par intention [Tg.an-1]
			(b) global : risque indirect d'eutrophisation, surplus agricole [Tg N an-1]; intrants totaux [Tg N an-1]	
Épuisement de l'ozone stratosphérique	Concentration en ozone [unité Dobson]	Concentration en ozone [unité Dobson]	Non prise en compte	Concentration en ozone [unité Dobson]
Acidification des océans	Concentration globale moyenne des eaux de mer superficielles en aragonite [Ω]	Concentration globale moyenne des eaux de mer superficielles en aragonite [Ω]	Non prise en compte	Concentration globale moyenne des eaux de mer superficielles en aragonite [Ω]
Utilisation de l'eau douce	Consommation globale d'eau douce par l'homme [km3.an-1]	(a) Global : consommation totale d'eau bleue [km3.an-1]	(a) Eau superficielle : altération des débits moyens mensuels par rapport aux débits hors activités anthropiques (<i>pristine flow</i>) [%]	(a) Eau bleue :perturbation anthropique du débit de des cours d'eau au-delà deleur variabilité médiane en condition non-anthropisée sous l'Holocène [%]
		(b) Bassin : prélèvements d'eau bleue pour l'agriculture, l'industrie et les ménages par rapport à trois classes de débit mensuel moyen	(b) Eau souterraine : quantité annuelle de prélèvements par rapport à la quantité annuelle de recharge [%]	(b) Eau verte : écart de la variabilité de l'humidité du sol par rapport à la médiane de l'Holocène [%]
Changement d'utilisation des sols	Pourcentage des surfaces cultivées par rapport à la surface émergée totale	(a) Global : surface des couverts forestiers par rapport à leur surface originelle [%]	Intégrée à la limite sur la biosphère	(a) Global : surface des couverts forestiers par rapport à leur surface originelle [%]
		(b) Biome : surface de couverts forestiers par rapport à leur surface potentielle [%]		(b) Biome : surface de couverts forestiers par rapport à leur surface potentielle [%]
Émission d'aérosols atmosphériques	Concentration totale enparticules fines	(a) Global : Profondeur atmosphérique en particules fines (AOD)	(a) sous-global : risque de modification du régime des moussons, AOD (4)	(b) Régional : AOD annuel moyen
		(b) Régional : AOD annuel moyen	(b) global : écart annuel moyen d'AOD entre les deux hémisphères [%]	(a) Global : Écart annuel moyen d'AOD entre les deux hémisphères [%]
Pollution chimique	Prise en compte mais non définie	Prise en compte mais non définie (1)	Non prise en compte	Rythme d'introduction de nouvelles substances chimiques synthétiques par rapport au rythme auquel ces substances font l'objet d'un contrôle [%]

(*) Les formulations données dans cette colonne sont propres à cet article et s'appuient sur la traduction de Boutaud et Gondran (BOUTAUD et al. 2020).Elles sont reportées ici dans un souci de comparabilité entre les publications, mais peuvent différer de la formulation exacte adoptée par chacune.

(1) Précisée par (WANG-ERLANDSSON et al. 2022) : le rythme d'introduction de nouvelles substances chimiques synthétiques doit être inférieur au rythme auquel ces substances font l'objet d'un contrôle.

(2) La publication de 2023 ajoute au cadre essentiellement biogéophysique des limites planétaires une dimension sociale par la prise en compte de trois principes de justice (interspécifique, intergénérationnelle et intragénérationnelle) dans la fixation du seuil limite (GUPTA et al. 2023).Le seuil planétaire est fixé au niveau le plus exigeant entre la limite "sûre" (assurant la stabilité du système planétaire) et la limite "juste" (minimisant la population exposée à des dommages significatifs), avec la possibilité que les deux seuils soient alignés. Par souci d'homogénéité avec les autres publications, nous ne représentons ici que les frontières "sûres" et non les frontières "justes".

(3) Par défaut, les variables de contrôle proposées dans (ROCKSTRÖM et al. 2023) sont définies à l'échelle sous-globale et c'est l'agrégation des états environnementaux évalués localement qui donne la valeur de l'état environnemental global (sauf pour la limite changement climatique).

(4) Limite dont la *Just Boundary* est plus restrictive que la *Safe Boundary*. Par exemple, des dommages additionnels significatifs sur les populations se produisent à un niveau de réchauffement inférieur aux seuils de déclenchement des points de bascule.

(5) Cette publication consiste en une révision directe du cadre de 2015 en intégrant les mises à jour de (WANG-ERLANDSSON et al. 2022) et (PERSSON et al. 2022) pour, respectivement, la limite sur l'eau douce et les entités nouvelles.

(a), (b), (c) sont ajoutés lorsque qu'une, deux ou trois variables de contrôles sont proposées respectivement pour une même limite.

TABLE 1.1 – Évolutions du cadre des limites planétaires, de 2009 à 2023.

Notons que l'article (ROCKSTROM et al. 2023) de Rockström et ses collègues, co-écrit par des chercheurs rassemblés autour de la "Commission de la Terre", qui réunit chercheurs en sciences de l'environnement, mais aussi en sciences sociales, pousse la dimension systémique de l'analyse jusqu'à prendre en compte des critères de justice indépendants du fonctionnement propre du système Terre. Grâce à la participation parmi les auteurs de chercheurs en sciences sociales, dont J. Gupta, spécialiste des questions de gouvernance et développement inclusif, les limites sont alors fixées non seulement par rapport aux seuils environnementaux et aux points de bascule dont nous avons décrit le principe dans ce qui précède, mais aussi en vertu d'un principe qui n'a que peu d'importance pour le fonctionnement propre du système mais se révèle être d'une importance fondamentale pour les systèmes humains voire pour le vivant en général. Ce principe est celui de "l'absence de dommage significatif". Toute limite environnementale est ainsi évaluée par rapport aux risques pour la santé qu'elle fait courir à la fois aux générations présentes, aux générations futures ainsi qu'aux autres espèces qu'humaines (GUPTA et al. 2023). Si le seuil environnemental est associé à un risque de dommage significatif sur l'un de ces trois critères de justice, alors le seuil qui définira la limite "juste" est abaissé afin de ne pas exposer de populations à des dommages significatifs. Méthodologiquement, c'est donc bien à partir des mécanismes environnementaux que le cadre des limites planétaires établit ses limites, mais les seuils attribués à chacune sont déterminés selon le niveau le plus strict entre le niveau "sûr" et le niveau "juste". Ainsi, un "espace juste et sûr de fonctionnement" peut être défini (voir Figure 1.3). Une limite planétaire est désormais fixée au niveau le plus strict, pour un processus environnemental donné, entre la valeur qui garantit le fonctionnement "sûr" du système Terre et la valeur qui assure le respect de 3 principes de justice : intergénérationnelle, intragénérationnelle, interspécifique.

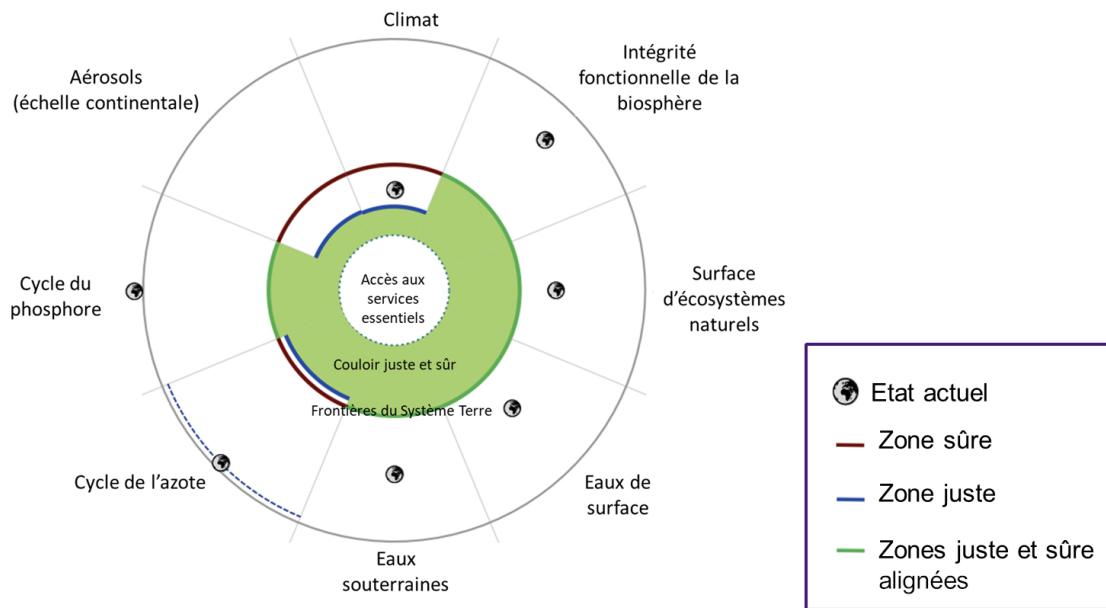


Figure 1.3 – Espace de fonctionnement sûr et juste issu de la version 2023 du cadre (Traduit de ROCKSTROM et al. 2023)

1.1.3 Les limites planétaires : quels processus environnementaux, quels indicateurs, quel état ?

Comme nous l'avons vu dans la partie précédente, le cadre a subi de nombreuses évolutions. Nous reprendrons dans cette partie les processus environnementaux identifiés dans la table 1, et nous en présenterons les variables de contrôle les plus actuelles qui leur sont associées. Nous allons présenter ci-dessous les "frontières du système Terre" telles que formalisées dans la publication de Rockström et ses collègues (ROCKSTROM et al. 2023) qui présente le double intérêt d'intégrer les enjeux de justice sociale et une volonté plus forte d'opérationnalisation du cadre et de suggérer des transformations nécessaires.

1.1.3.1 Les limites qui entraînent des risques de dépassement systémique

1.1.3.1.1 Le changement climatique

Le changement climatique désigne l'ensemble des modifications qu'entraînent les émissions de gaz à effet de serre et le changement d'utilisation des sols sur les processus de régulation du climat, notamment la température de surface (sol et troposphère), la circulation atmosphérique et le régime des précipitations. Plus précisément, les activités humaines ont un effet sur le bilan radiatif de la Terre, qui impacte les processus que nous venons de citer. Pour

une description plus approfondie des conséquences du réchauffement climatique sur les systèmes socio-écologiques, le lecteur peut se référer à Dumergues (DUMERGUES 2020).

La température moyenne de la surface de la Terre dépend de plusieurs paramètres qui sont pris en compte dans les modèles climatiques utilisés par le cadre des limites planétaires : la puissance du rayonnement solaire incident, l'absorption/la réflexion de ce rayonnement avant d'atteindre le sol, la concentration de l'atmosphère en GES.

La frontière "sûre", qui minimise les risques de franchissement de seuils de bascule a été estimée à 1,5°C d'augmentation de température moyenne mondiale (IPCC 2021). Différents points de bascule du climat sont associés à différents niveaux d'augmentation de la température moyenne mondiale. Cependant, une hausse de 1°C de la température moyenne mondiale expose des dizaines, voire des centaines, de millions de personnes à des événements météorologiques extrêmes ainsi qu'à des hausses de température si élevées que certains territoires deviennent inhabitables (IPCC 2022). Lorsque la hausse de température moyenne mondiale dépasse 1°C, des dizaines de millions de personnes risquent d'être exposées à divers phénomènes climatiques qui rendent impossible la vie sur certains territoires. Cela soulève d'évidentes préoccupations de justice intra et inter-générationnelle. Rockström et ses collègues préconisent donc une frontière "juste et sûre" correspondant à une augmentation de la température moyenne mondiale de 1°C... augmentation qui est inférieure à celle constatée en 2023 par les climatologues (1,2°C) (ROCKSTROM et al. 2023).

1.1.3.1.2 Les menaces sur l'intégrité de la biosphère

Une biosphère intègre remplit ses fonctions de régulation du système Terre grâce à la biodiversité qui la caractérise. Classiquement, cinq grands types de fonctions de régulation, ou "services rendus par la nature", sont considérés comme rendus par la biodiversité (MILLENNIUM ECOSYSTEM ASSESSMENT 2005) :

- Ressources : production de ressources (pêche, foresterie, pollinisation, fertilisation des sols, etc)
- Puits : décomposition des déchets, séquestration du CO₂ dans les sols et les océans, régulation des cycles biogéochimiques
- Site : fourniture d'espaces physiques pour la réalisation des activités humaines
- Support de la vie : préservation contre les risques de dégradation des conditions de vie

(régulation du cycle de l'eau par les zones humides pour limiter les inondations, fixation du sol par le couvert végétal pour limiter l'érosion et les glissements de terrain, etc)

- Paysage : qualités esthétiques d'un écosystème, dimension culturelle d'une espèce, par exemple

Au cours des dernières années, de nombreuses études ont alerté régulièrement sur les atteintes que portent les activités humaines à l'intégrité de la biosphère :

- Baisse de 60% de la population de vertébrés depuis les années 1970 dans le monde (WWF 2018)
- 6500 espèces menacées d'extinction, soit 27% des espèces étudiées, dont 40% des amphibiens, 33% des coraux, 31% des requins et des raies, 27% des crustacés, 25% des mammifères, 34% des conifères (IUCN 2020)
- Les services écosystémiques rendus par les pollinisateurs à l'agriculture en UE sont estimés à 3,7 milliards d'euros, 1 espèce pollinisatrice sur 3 (abeilles et papillons) est en déclin démographique et 1 espèce sur 10 en danger d'extinction en UE (VYSNA et al. 2021)
- Baisse de 76% de la biomasse d'insectes volants en Allemagne depuis 1990 (HALLMANN et al. 2017)
- Voire très localement, ces dégradations se constatent à des échelles réduites. Sur le bassin versant de la Loue en Franche-Comté par exemple, on observe une baisse de 50% de l'abondance aquatique en 10 ans et une perte des trois quarts de l'abondance entomique (DEGIORGI et al. 2020).

Deux variables ont été choisies par Rockström et ses collègues (ROCKSTROM et al. 2023) pour représenter l'intégrité de la biosphère : la surface d'écosystèmes naturels conservés majoritairement intacts et l'intégrité fonctionnelle de tous les écosystèmes, y compris urbains et agricoles.

Le maintien de larges zones d'écosystèmes naturels est nécessaire pour garantir les fonctions du système terrestre dont dépendent toutes les espèces vivantes, incluant les humains, aussi appelées les "contributions de la nature aux personnes", ainsi que la stabilité du Système Terre. En effet, ces écosystèmes permettent de réguler les stocks et cycles du carbone, de l'eau et des nutriments. Ils sont aussi garants du maintien de la diversité des espèces, qui assure une moindre vulnérabilité face aux changements écologiques à venir. Après avoir fait de nombreuses

modélisations et en observant la capacité de maintien de l'habitat pour la biodiversité et les différents cycles (carbone, eau et nutriments), Rockström et ses collègues proposent d'établir la frontière juste et sûre, pour la biosphère, à 50 à 60% de la surface terrestre couverte par des zones naturelles conservées en grande partie intactes. Aujourd'hui, seulement 45 à 50% de la surface terrestre de la planète est couverte par des zones naturelles intactes. Cette frontière est donc dépassée. Des mesures de restauration de l'habitat naturel permettraient de revenir à des valeurs de 50 à 60%. Notons également que cette frontière est définie pour les surfaces terrestres, mais que le maintien d'une même proportion d'océan est nécessaire pour maintenir la séquestration du carbone et minimiser l'extinction d'espèces marines supplémentaires.

L'intégrité fonctionnelle est définie comme "la capacité des écosystèmes urbains, agricoles ou autres écosystèmes modifiés par l'homme à fournir des fonctions écologiques et leurs contributions aux personnes" (ROCKSTROM et al. 2023). Cette variable doit être définie à l'échelle locale. Les auteurs proposent ainsi d'agréger les analyses obtenues sur des mailles de 1km*1km. Ils considèrent que chaque maille de 1km² doit comporter au moins 20 à 25% d'habitats semi-naturels diversifiés pour permettre aux diverses espèces indigènes d'assurer leurs multiples "contributions locales de la nature aux personnes" telles que la pollinisation, la lutte contre les ravageurs et les maladies, la régulation de la qualité de l'eau, la protection des sols, l'atténuation des risques naturels ou permettre des loisirs. La frontière planétaire serait donc que 100% des mailles de 1km² respectent cette surface minimale de 20 à 25% d'habitat semi-naturel en leur sein... or seulement un tiers (31 à 36%) de ces mailles remplissent cette condition en 2023... Encore plus inquiétant, certaines zones, particulièrement anthropisées, présentent des symptômes de pertes de résilience du fait de l'absence sur de grandes surfaces d'habitats semi-naturels. Ce constat plaide pour la mise en place de pratiques de régénération des habitats naturels pour restaurer les écosystèmes locaux et planétaires.

1.1.3.1.3 Érosion de la couche d'ozone et acidification des océans

Deux autres limites planétaires sont garantes de la stabilité du système Terre : l'érosion de la couche d'ozone et l'acidification des océans.

La première s'intéresse aux gaz de la famille des chlorofluorocarbures, qui se décomposent dans l'atmosphère sous l'effet du rayonnement solaire et empêchent la formation de l'ozone

dans la stratosphère. En conséquence, la stratosphère laisse davantage passer le rayonnement ultra-violet, néfaste pour la vie biotique terrestre. Grâce aux efforts de réglementation internationale depuis le Protocole de Montréal de 1987, l’ozone est en voie de reformation progressive, notamment au-dessus des pôles où il est attendu une reformation totale à horizon 2050 (CHIPPERFIELD et al. 2020).

L’acidification des océans, quant à elle, est liée à l’augmentation de la concentration de l’air en CO₂, qui diminue le pH des eaux océaniques (dissolution des molécules d’air en surface). La tendance est aujourd’hui à un déséquilibre du pH à la hausse, ce qui menace la capacité des organismes calcifiés à se maintenir et menace les chaînes trophiques océaniques qui en dépendent.

1.1.3.2 Les limites qui agissent par rétroaction sur les limites de dépassement systémique

1.1.3.2.1 La perturbation des cycles biogéochimiques du phosphore et de l’azote

Le phosphore participe à la formation des acides nucléiques et des phospholipides. Comme l’azote, il est un nutriment indispensable à la vie. Les phosphates du milieu intérieur ont un rôle physiologique, en particulier dans la régulation du pH (élimination rénale de phosphates acides).

Le phosphore se trouve exclusivement dans la lithosphère et la biosphère. Dans la lithosphère (abondance à 0,008% en masse dans la croûte terrestre), il est le résultat de l’enfouissement sans dégradation de la matière organique dans des roches calcaires. Les mouvements tectoniques ont pu faire remonter ces strates à la surface où leur lessivage entraîne la solubilisation d’une partie du phosphore contenu dans ces roches, qui est alors biodisponible pour les plantes.

L’azote est indispensable à la vie puisqu’il est impliqué dans la formation des acides aminés (protéines) et des acides nucléiques. Très abondant dans l’atmosphère sous la forme du diazote N₂ (78% du volume de l’air), il est toutefois difficilement assimilable.

Phosphates et nitrates sont épandus sur les sols sous formes d’engrais afin d’améliorer la productivité des sols. Cependant, ils sont peu retenus par le substrat et, sous l’effet de la

percolation de l'eau de pluie, sont emportés vers les nappes (risque de non-potabilité) ou, par ruissellement, vers les eaux superficielles. Cela crée alors un phénomène d'eutrophisation, qui crée une prolifération d'algues, ce qui réduit la quantité d'oxygène disponible pour les autres êtres vivants de l'écosystème.

Alors que les limites "planétaires" sur le phosphore et l'azote définies par les équipes de Steffen (2015) et Richardson (2023) correspondent au risque d'un évènement anoxique océanique (réduction généralisée du taux d'oxygène dissout dans les océans due à une surstimulation de la bioproduktivité primaire marine) (STEFFEN et al. 2015b ; RICHARDSON et al. 2023), les limites du "système Terre" de Rockström et ses collègues (ROCKSTROM et al. 2023) ne retiennent que le risque d'eutrophisation des hydrosystèmes terrestres.

Les limites fixées pour chacun de ces deux nutriments sont ainsi liées au cycle de l'eau. Cela génère d'importants dommages aux écosystèmes, pouvant produire l'effondrement des populations de poissons, libérer des composés toxiques par la prolifération d'algues (dans les eaux douces ou marines) ainsi qu'avoir des effets sur la santé humaine. En France, ce phénomène d'eutrophisation est bien connu tout le long de la Loire (de nombreux plans d'eau sont interdits à la baignade en été de ce fait), mais aussi sur les plages bretonnes avec le phénomène des algues vertes. La frontière planétaire est définie à travers une estimation du surplus de nutriments épandus sur les sols mais non absorbés par les cultures. Le surplus maximal qui serait acceptable est estimé à 61 millions de tonnes pour l'azote et de 4,5 à 9 millions de tonnes pour le phosphore. Ces deux frontières sont aujourd'hui largement dépassées. Des considérations de justice entre les pays s'ajoutent à cela. En effet, certains pays et communautés, à faibles revenus, n'ont pas accès aux engrais azotés et phosphorés, ce qui menace leur sécurité alimentaire. D'autre part, l'extraction minière de phosphate, dont les ressources sont limitées et localisées, expose les communautés voisines des exploitations, souvent pauvres et marginalisées, aux déchets miniers, aux terres détruites et à des violations des droits de l'homme.

1.1.3.2.2 La perturbation du cycle de l'eau douce

L'eau douce ne représente qu'une petite fraction de l'eau disponible sur Terre : elle est 40 fois moins abondante que l'eau salée. Et encore, seule une petite partie de cette eau douce est biodisponible pour les végétaux et accessible aux hommes et animaux. En effet, près

de 70% du volume d'eau douce est piégé dans les glaces et les neiges et près de 30% est stocké dans les aquifères. L'eau de surface, en lac ou cours d'eau, ne représente ainsi que 1% de l'eau douce disponible : 20% de cette eau est contenue dans le lac Baïkal (situé en Russie orientale) et 20% sont contenus dans les grands lacs d'Amérique du Nord. Si la Terre est une "planète bleue", l'eau réellement disponible pour les écosystèmes non-marins apparaît en fait marginale.

Alors que (STEFFEN et al. 2015b) raisonnaient en termes de stock d'eau douce disponible à l'échelle globale, (ROCKSTROM et al. 2023) proposent de réfléchir en termes de flux. La limite exprime un seuil critique de disponibilité de la ressource en eau pour le bon fonctionnement des écosystèmes à travers la définition d'un taux d'altération maximal du débit des cours d'eau.

La variable de contrôle sur les eaux de surface correspond alors à une altération maximale de 20% du débit moyen mensuel observé régionalement. Cette valeur de 20% correspond à un proxy empirique, basé sur une revue de la littérature, pour définir un seuil de bascule dans l'intégrité de la biodiversité aquatique. En 2020, 66% des régions ainsi évaluées se situaient en dessous de ce niveau d'altération. La variable de contrôle globale, qui agrège les différents débits mesurés, ne conduit alors pas un dépassement du stock d'eau douce maximum qu'il est possible de prélever. Les évaluations à l'échelle locale aboutissent quant à elles à un dépassement de la limite, car il est proposé que la frontière soit respectée si 100% des différentes régions connaissent une altération inférieure à 20%.

L'évaluation de l'altération du renouvellement des eaux souterraines est également proposée à travers la comparaison entre la quantité annuelle des prélèvements par rapport à la quantité annuelle rechargée. Les aquifères sont en effet indispensables aux hydrosystèmes dans la mesure où ils contribuent au soutien à l'étiage ainsi qu'au fonctionnement des zones humides et au maintien de la végétation terrestre. 47% des bassins versants dans le monde connaissent un rythme d'exploitation non soutenable de leurs aquifères et la limite globale est ainsi dépassée.

Une limite est également proposée pour l'eau verte, qui représente l'eau disponible pour les végétaux (contenue dans les précipitations, l'évaporation des plantes et l'humidité des sols). La variable de contrôle retenue pour l'eau verte consiste à quantifier l'écart moyen mensuel d'humidité racinaire du sol par rapport à sa valeur normale sous l'Holocène. Pour que la limite planétaire soit respectée, au moins 90% des zones ainsi étudiées ne doivent pas s'écarter de plus de 5% de la valeur normale de l'Holocène, valeur largement dépassée aujourd'hui selon les

modèles mobilisés.

1.1.3.2.3 L'accroissement de la charge atmosphérique en aérosols

Les aérosols comprennent une très grande variété de particules émises dans l'atmosphère par l'homme ou de façon naturelle. Leur point commun est d'interférer avec le climat mais de façon moins prévisible que les gaz à effet de serre, dont le principe de fonctionnement est bien connu (réémissions du rayonnement infrarouge terrestre).

L'effet des aérosols sur le bilan radiatif est complexe à mesurer dès lors que de nombreux effets s'opposent. Dans son dernier rapport publié en 2021 (IPCC 2021), le GIEC quantifie à l'échelle globale les effets nuancés des aérosols et aboutit à la conclusion qu'au bilan, sur la période 2010-2019, les aérosols ont eu un effet négatif sur le forçage radiatif par rapport à la période 1850-1900. Ainsi, sur un forçage total entre les deux périodes de $+1^{\circ}\text{C}$, le dioxyde de soufre a contribué à hauteur de $-0,5^{\circ}\text{C}$, les NOx à hauteur de $-0,2^{\circ}\text{C}$, les suies à hauteur de $+0,1^{\circ}\text{C}$ et les COV à hauteur de $+0,25^{\circ}\text{C}$.

Rockström et ses collègues (ROCKSTROM et al. 2023) insistent sur un autre effet environnemental lié aux aérosols : une différence croissante de pollution aux aérosols entre les hémisphères Nord et Sud (les émissions de particules tendant à se réduire dans l'hémisphère Nord), peut impacter les conditions hydrologiques régionales, et modifier des régimes tels que celui de la mousson. (ROCKSTROM et al. 2023) et (RICHARDSON et al. 2023) proposent donc d'utiliser comme frontière la différence de profondeur optique des aérosols (AOD) entre les deux hémisphères. Ces émissions d'aérosols peuvent être d'origine naturelle, anthropique (particules) ou être dues à l'injection volontaire d'aérosols stratosphériques, dans une perspective de géo-ingénierie solaire. Pour l'instant, cette frontière "sûre" est respectée, mais une frontière "juste" plus stricte est proposée pour prendre en compte les importants dommages causés à la santé humaine liés à l'exposition aux aérosols, tels que les particules. Certaines populations sont aujourd'hui exposées à des concentrations de particules fines (dont le diamètre est inférieur à $2,5\ \mu\text{m}$) supérieures à la frontière "juste" estimée à $15\ \mu\text{g}/\text{m}^3$ d'air.

1.1.3.2.4 L'introduction d'entités nouvelles dans l'environnement

L'introduction d'entités nouvelles était identifiée comme une limite planétaire, mais aucune variable de contrôle n'était quantifiée en 2015 dans l'article de (STEFFEN et al. 2015b). Cette limite représente la préoccupation quant au devenir des molécules (chimiques, radioactives ou biologiques, par exemple) lorsqu'elles sont introduites par les activités humaines dans les écosystèmes. Les processus de dégradation et de régulation de ces "entités" sont très complexes et multiples, impliquant différentes molécules ayant des effets variés, difficiles à modéliser à l'échelle planétaire par une seule "variable de contrôle". La non-quantification de ce thème, en 2015, présentait également l'avantage d'introduire, sous la forme d'un point d'interrogation, de nouvelles thématiques environnementales qui pourraient devenir de nouvelles candidates au statut de "limite planétaire", si l'évolution des connaissances scientifiques et de leur niveau de dégradation mettent en évidence de nouvelles sources de préoccupations impactant le "système Terre" et pouvant remettre en cause son équilibre.

Fin 2021, un article de L. Persson et ses collègues (PERSSON et al. 2022) a mis en évidence l'augmentation fulgurante de la production, des rejets et impacts des molécules chimiques et plastiques de synthèse. Par exemple, la production et l'utilisation mondiale de plastiques, qui augmente régulièrement depuis les années 1950, a augmenté de 79% de 2000 à 2015. Les auteurs en concluaient que la frontière planétaire pour l'introduction d'entités nouvelles était déjà dépassée dans la mesure où la production et les rejets annuels de molécules chimiques augmente à un rythme qui dépasse la capacité mondiale de leur évaluation et leur suivi.

1.2 Un nouveau cadre pour penser l'évaluation environnementale des organisations

1.2.1 Un renouvellement paradigmatique dans le champ de l'évaluation environnementale

Le cadre des limites planétaires n'a pas été pensé pour être appliqué localement, mais pour promouvoir une vision systémique du "système Terre" (ROCKSTRÖM et al. 2009a). Cependant, la communauté scientifique travaillant sur l'évaluation des pressions sur l'environnement, et en particulier sur l'analyse de cycle de vie, y a vu l'occasion d'appuyer ses résultats d'analyse sur une grille basée sur un état objectif du système Terre. Précisément, le cadre des limites

planétaires a permis au champ de l'évaluation environnementale de développer un nouveau concept, celui d'évaluation absolue de la soutenabilité environnementale (AESAs). Ce concept a été proposé en 2015 par Anders Bjørn suite à la publication d'un article dans lequel il proposait d'utiliser le cadre des limites planétaires comme facteur de normalisation des résultats d'une analyse du cycle de vie (ACV) (BJØRN et al. 2015a; BJØRN et al. 2015b). Pour reprendre sa définition :

"Les capacités de charge sont dérivées des seuils inhérents à la réponse de l'environnement, par exemple, aux concentrations croissantes de polluants ou à l'utilisation des ressources. [...] En utilisant la capacité de charge comme référence, les indicateurs de l'ACV peuvent devenir des indicateurs de soutenabilité environnementale absolue. Ces indicateurs sont absolus car les capacités de charge sont indépendantes du système produit évalué." Traduit de (BJØRN et al. 2015a)

L'idée initiale des méthodes d'évaluation absolue de la soutenabilité environnementale est de questionner les approches d'*éco-efficience* : réduire la quantité d'impacts environnementaux par unité ou par euro de bien produit ne suffit pas pour atteindre un état environnemental satisfaisant (BJØRN et al. 2013).

Passer de l'échelle planétaire à cette échelle locale, implique de définir une "capacité de charge", qui représenterait le "reste à émettre" (ou à polluer) qui peut être assigné à un individu, un pays ou une entreprise. Ce "reste à émettre" correspond à la quantité maximale d'émissions pouvant être générées par les activités humaines pour respecter les différentes limites planétaires. Cette capacité de charge peut être considérée au même titre que l'accès à des ressources naturelles limitées pour lesquelles les différents acteurs se font concurrence (HAUSCHILD et al. 2020). Cependant, si la définition des limites, à l'échelle planétaire, fait l'objet d'un consensus croissant, les différents principes et méthodes pour répartir ce "reste à émettre" entre les différents "acteurs sous-globaux" font l'objet de partis pris éthiques incontournables.

Le terme de *soutenabilité absolue* se trouve dans la littérature anglophone, mais nous préférons utiliser l'expression d'*évaluation absolue de la soutenabilité*, afin de rendre compte de la visée méthodologique et normative de l'évaluation.

Nous présenterons dans les parties suivantes ce que cette modalité "absolue" d'évaluation environnementale implique en termes méthodologiques (2.2) pour nous focaliser ensuite sur la question du "principe d'assignation" qui est au cœur de cette méthodologie (2.3).

1.2.2 Une introduction à l'évaluation absolue de la soutenabilité environnementale

Le cadre de l'évaluation absolue de la soutenabilité environnementale vise à répondre à la question suivante "la pression environnementale de cette activité est-elle suffisamment faible pour pouvoir la considérer comme soutenable sur le plan environnemental, et sinon, de combien la pression devrait-elle être inférieure?" (BJØRN et al. 2019). Il faut donc d'abord définir une "capacité de charge acceptable", ou "espace environnemental" qui permettrait de préserver les processus de régulation biogéophysiques du système Terre. L'enjeu est de réussir à maintenir les activités humaines dans cet espace environnemental.

Pour chaque processus de régulation du système Terre, le cadre des limites planétaires permet de définir cette "capacité de charge". D'autres "objectifs de soutenabilité environnementale" que les limites planétaires pourraient être retenus (BJØRN et al. 2019), du moment qu'ils décrivent des limites biophysiques basées sur la connaissance scientifique (HAUSCHILD et al. 2020).

La démarche globale pour l'évaluation environnementale absolue d'un système peut se décomposer en quatre étapes que nous détaillons ci-après (BJØRN et al. 2019) :

1. définir le système étudié, ses fonctions, son périmètre ;
2. quantifier les pressions environnementales exercées par ce système dans les mêmes métriques que la méthode d'AESA retenue ;
3. comparer les pressions du système à la capacité de charge allouée au système ;
4. communiquer et présenter les résultats.

Pour résumer, l'utilisation des limites planétaires pour mener une évaluation environnementale consiste à considérer chacune des limites comme un espace environnemental dans lequel les activités humaines peuvent évoluer et dont elles ne doivent pas sortir. Tout l'enjeu est de savoir quelle part de cet espace chacune des activités humaines actuelles et futures peuvent occuper. Cet enjeu renvoie à des considérations éthiques puisqu'il s'agit de privilégier un principe de

justice distributive sur un autre (RYBERG et al. 2020).

1.2.3 Le principe d'assignation au cœur d'une évaluation AESA basée sur les limites planétaires

Le principe d'assignation permet d'attribuer au périmètre de l'activité retenue pour l'évaluation une fraction de l'espace de fonctionnement sûr. La façon d'attribuer cet espace commun et planétaire à des systèmes sous-globaux est déterminante dans une évaluation AESA parce qu'elle fixe le niveau de pression que l'activité ne doit pas dépasser afin d'être considérée comme absolument soutenable. Le partage, entre acteurs et activités, de cet espace environnemental global pose un problème de deux types : quels principes de justice sous-tendent le choix de la méthode de désagrégation retenue ? quelle clé de désagrégation choisir pour représenter le périmètre particulier d'une activité au sein de l'espace environnemental ?

Six principes de partage peuvent être définis : l'égalité par personne (le droit à émettre d'un pays est proportionnel à sa population), les droits acquis ("*grandfathering*" : le droit à émettre d'un pays est proportionnel à ses émissions passées), le droit au développement (le principe d'égalité est différencié en faveur des économies "pauvres"), le principe pollueur payeur (la contribution du pays à l'effort d'atténuation est proportionnelle à sa responsabilité aux pressions), la capacité de payer (la contribution du pays à l'effort d'atténuation est proportionnelle à sa capacité à payer ou agir contre les pressions), et le volontarisme (chaque pays détermine sa contribution à l'atténuation selon ses capacités) (HÄYHÄ et al. 2016).

Le partage de l'espace de fonctionnement sûr se fait généralement en deux temps (HJALSTED et al. 2021) :

1. une assignation de l'espace planétaire à l'échelle individuelle, basée sur un principe d'assignation qui permet de répartir une ressource donnée entre des individus ;
2. une mise à l'échelle de cette répartition individuelle vers la part pouvant être attribuée à l'activité évaluée. La Figure 1.4 propose une vision schématique de cette démarche.

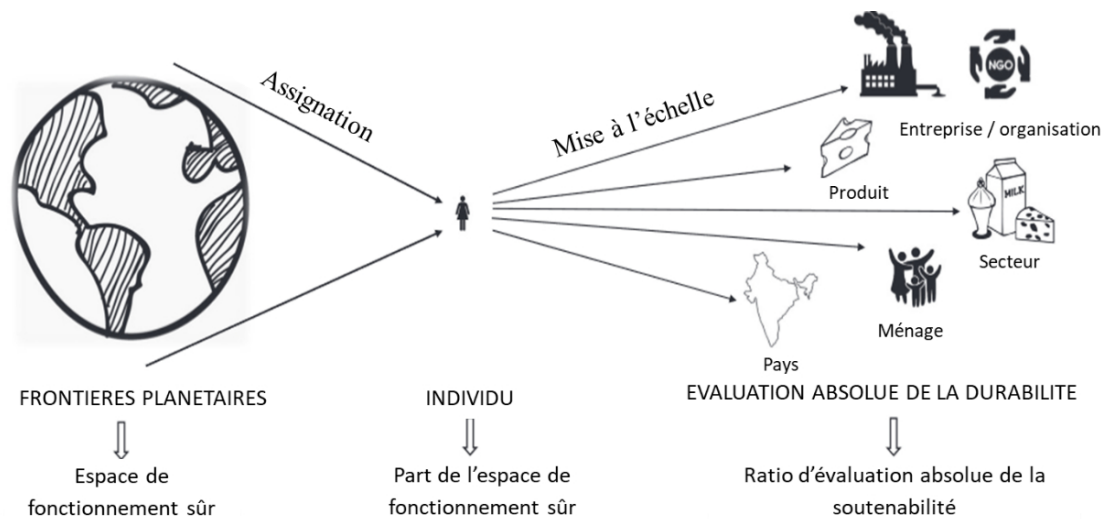


Figure 1.4 – Principe de partage de l’espace de fonctionnement sûr proposé par Hjalsted et ses collègues (Traduit de HJALSTED et al. 2021).

1.2.3.1 La mise à l’échelle individuelle

Le choix du postulat individuel se positionne dans la lignée de la définition du développement durable, dans le rapport Brundtland de 1987, qui se concentre sur la question de la satisfaction des besoins humains et constitue une vision anthropocentrée de la soutenabilité, centrée sur l’individu (HJALSTED et al. 2021). Tout système ou toute activité ne bénéficie d’une fraction de l’espace de fonctionnement sûr que dans la mesure où il contribue à satisfaire les besoins d’un, ou plusieurs, individus. Cette hypothèse anthropocentrée, si elle permet de définir un premier principe évident d’allocation – le ratio par tête (ou *per capita*) – pose la question de savoir comment caractériser la contribution d’une activité donnée à la satisfaction des besoins humains.

Plusieurs principes de modèles de distribution de l’espace de fonctionnement sûr peuvent également être identifiés (RYBERG et al. 2020) :

- utilitarisme : la somme des bien-être doit être maximisée.
- égalitariste : tout agent doit bénéficier de la même quantité de toute chose
- prioritarisme : la valeur morale d’une quantité attribuée est d’autant plus grande que la situation de l’agent à qui elle est attribuée est dégradée
- suffisantisme : tout agent doit être suffisamment doté pour pouvoir répondre à ses besoins essentiels
- droits acquis : un principe de justice selon lequel la distribution d’une quantité à venir entre agents est proportionnelle à la répartition actuelle déjà observée entre ces agents

(statu quo de la répartition réelle)

Trois de ces principes sont le plus souvent retenus et traduits en principe d'assignation. L'égalitarisme donne lieu au principe d'assignation égale par habitant. Le principe de "droits acquis" est souvent employé pour estimer le poids spécifique d'un secteur d'activité donné sur l'ensemble de la capacité de charge par personne, car il est celui pour lesquels des données sont le plus facilement accessibles. Mais ce choix reflète l'absence de critères moraux : une entreprise d'extraction du charbon recevrait par exemple davantage de droits à émettre du CO₂ qu'une entreprise de production d'énergie éolienne du fait que cette dernière est entrée sur le marché en émettant d'emblée moins de CO₂ (RYBERG et al. 2020). Enfin, le principe de prioritarisme est traduit en un principe de capacité à payer, où les pays les moins riches doivent pouvoir bénéficier d'une part plus importante de l'espace environnemental que les pays plus riches.

1.2.3.2 La mise à l'échelle du système étudié

La méthode de "mise à l'échelle" vise à répartir la capacité de charge par personne selon les différentes activités. Dix différents principes de partage entre les différentes activités peuvent être observés : répartition égale par tête, capacité à réduire ses impacts, dette historique, droits acquis, surface au sol, valeur ajoutée économique, coût d'efficacité, dépenses de consommation finale, contenu calorifique, extrant physique de production (BJØRN et al. 2020a).

Cette diversité des principes montre qu'il n'y a pas encore de stabilité dans les méthodologies proposées, ouvrant ainsi le champ à l'originalité et la créativité aux développeurs de méthodes. Pour l'instant, aucune méthode n'est recommandée, si ce n'est la nécessité de mener une étude de sensibilité afin de confronter les conclusions aux différents partis pris méthodologiques possibles. Les différents choix méthodologiques cachent ainsi des partis pris éthiques forts, qui ne doivent pas rester le seul choix d'experts en évaluation environnementale.

1.3 Utiliser les limites planétaires à l'échelle des organisations : les méthodes principales

1.3.1 Les applications à l'Analyse de Cycle de Vie dans le champ scientifique

Trois types d'étude visant à effectuer des évaluations absolues de la soutenabilité environnementale peuvent être observées (BJØRN et al. 2020a) :

- Les études basées sur les analyses d'impact du cycle de vie ("AESA-ACV") qui reprennent les catégories d'impacts référencés par l'ILCD (International Life Cycle Data), c'est-à-dire les indicateurs considérés comme des références par la communauté internationale européenne. Des valeurs de référence, définies à partir des capacités de charge estimées à partir des limites planétaires permettent de normaliser les différents résultats d'ACV (voir (BJØRN et al. 2015b) et (SALA et al. 2020) par exemple).
- Les études qui s'appuient à la fois sur les catégories d'impacts spécifiques à l'ILCD et sur celles du cadre des limites planétaires "*LCIA- and PB-based*" : trois catégories d'impacts de l'ILCD peuvent en effet s'exprimer avec des modèles permettant de passer des métriques portant sur les émissions aux variables de contrôle sélectionnées par les limites planétaires : le changement climatique, l'érosion de la couche d'ozone, l'utilisation de l'eau douce.
- Les études basées sur les variables de contrôle du cadre des limites planétaires "*PB-based only*" dont les auteurs ont développé de nouveaux jeux de facteurs de caractérisation exprimés dans la même unité que les limites planétaires (ppm CO₂) par exemple, plutôt que des kg CO₂eq (voir (RYBERG et al. 2018) par exemple).

Dans les parties suivantes, nous mettrons l'accent sur deux méthodologies utilisées en ACV pour l'application des limites planétaires à des systèmes locaux (organisation, produit, ou activité par exemple) :

- La méthode *LCIA-based AESA* : proposée par Anders Bjørn (BJØRN et al. 2015b), elle adapte le cadre des limites planétaires aux indicateurs de l'ACV ;
- La méthode *PB-based LCA* : proposée par Morten Ryberg (RYBERG et al. 2018), elle est basée sur l'application directe des limites planétaires, en modifiant les facteurs de caractérisation de l'ACV.

Nous poserons ensuite la question de la possibilité pour ces méthodes d'intégrer la prise de décision.

1.3.1.1 La méthode *LCIA-based* : une ACV normalisée par la capacité de charge

Cette section présente la méthode d'AESA basée sur l'ACV et les limites planétaires, initialement proposée par A. Bjørn pour l'évaluation d'un système de type produit ou secteur (BJØRN et al. 2015b).

L'idée est d'intégrer les limites planétaires à l'analyse d'impact ACV au moment de la normalisation des résultats de LCIA. Les catégories d'impacts étant inchangées par rapport aux catégories recommandées par le *ILCD handbook*, les métriques peuvent être différentes des métriques utilisées par les variables de contrôles du cadre des limites planétaires (STEFFEN et al. 2015b).

Le cadre des limites planétaires est ainsi utilisé pour définir des valeurs de référence permettant de normaliser les résultats d'ACV des différentes catégories d'impacts. Pour toute catégorie d'indicateurs orientés impacts (dit "*midpoint*"), une référence de normalisation lui est appliquée. Celle-ci se calcule comme la capacité de charge, estimée à partir des limites planétaires, ramenée au prorata de la population comprise dans le périmètre de l'étude ACV. Il s'agit donc d'une évaluation AESA basée sur un principe d'allocation égalitariste (partage égal par tête), pour reprendre les catégories d'analyse élaborées dans la partie précédente. La capacité de charge, exprimée par année, est en fait la valeur annualisée d'un seuil planétaire. Ceci ne vaut pas pour toutes les catégories d'impact utilisées dans cette méthode, car les métriques de l'ILCD et des limites planétaires ne peuvent pas toujours être mises en relation. Par exemple, l'eau douce dans le cadre des limites planétaires est regardée du côté quantitatif, c'est-à-dire l'altération apportée par l'homme au débit des cours d'eau du fait de ses prélèvements. L'aspect qualité de l'eau est pris en compte mais au travers des limites sur les cycles des nutriments Azote et Phosphore. L'ILCD de son côté propose de prendre en compte quatre indicateurs qui recourent ceux des limites planétaires mais sans s'y résoudre : l'eutrophisation marine, terrestre et de l'eau douce, ainsi que l'utilisation de l'eau. Bjørn et Hauschild (BJØRN et al. 2015b) se sont ainsi inspirés de la "pensée limites planétaires" mais n'en ont pas adopté systématiquement les métriques. Ils se sont appuyés sur des travaux distincts pour chacun d'eux qui quantifient une limite globale adaptée à leur métrique.

L'usage qui est fait de la notion de "capacité de charge" en AESA est quelque peu différent des définitions initiales des écologues mais s'inspire de la définition utilisée par l'empreinte écologique. L'objectif est d'évaluer les résultats pour un système donné, non plus par rapport à un autre état ou scénario de ce système (par exemple l'impact comparé de deux chaînes différentes de production de tomates en conserve), ni par rapport à un système proche (par exemple l'impact comparé de cette même chaîne de production entre deux pays différents), mais bien dans l'absolu ("per se") (HJALSTED et al. 2021). La capacité de charge, utilisée comme référence de normalisation, est alors définie comme "l'intervention soutenable maximum qu'un système naturel peut supporter sans que se produisent des changements négatifs dans sa structure ou son fonctionnement tels qu'ils soient irréversibles ou difficilement réversibles" (BJØRN et al. 2015b).

Cette méthode a par exemple été reprise par une équipe de chercheurs du Joint Research Center, laboratoire de recherche de l'Union européenne, pour l'appliquer au cas de la consommation Européenne (SALA et al. 2020). Ce projet vise à estimer les impacts générés par les consommations des ménages du territoire étudié (l'Union européenne à 28 Etats-membres) à partir des catégories d'impacts d'ACV normalisées par des valeurs de références basées sur les limites planétaires. Sur la Figure 1.5 sont ainsi représentés : en bleu, la part de l'impact "importée" (approche empreinte) et en gris, l'impact total (domestique + importé). Quand une catégorie d'impact dépasse la valeur 1, c'est que la pression environnementale dépasse la capacité de charge assignée. C'est par exemple le cas du changement climatique (CC).

Le principe général d'une évaluation basée sur les consommations (de type "empreinte") est d'additionner la quantité de biens et services produits sur le territoire avec la quantité importée d'autres territoires, et d'y retrancher les exportations vers d'autres territoires (BJØRN et al. 2015b).

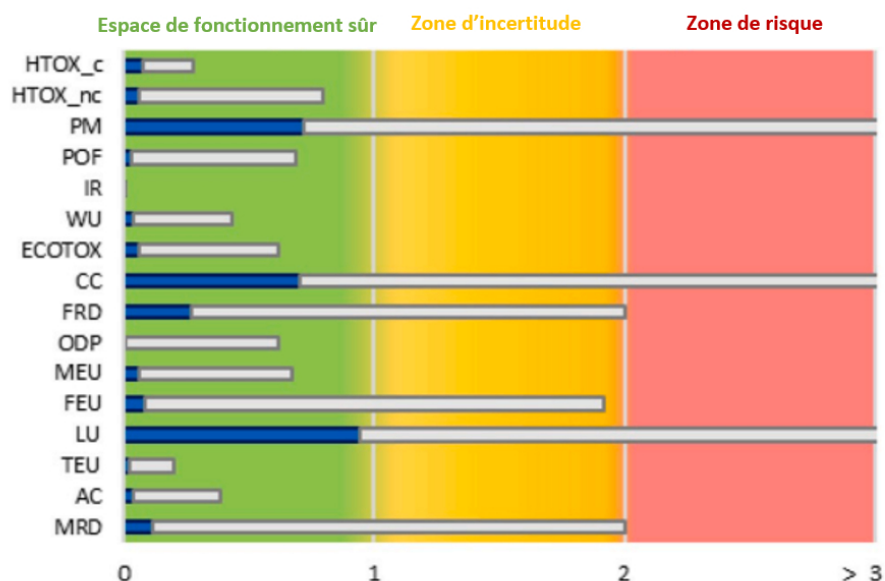


Figure 1.5 – Résultats de l'étude menée par le JRC européen sur l'impact de la consommation des Européens comparée à la part assignée des limites planétaires. (Traduit de (SALA et al. 2020))

1.3.1.2 La méthode *PB-based* : des indicateurs d'ACV convertis en indicateurs de limites planétaires

Selon Morten Ryberg (RYBERG et al. 2018), concepteur de cette méthode, mobiliser des indicateurs prédéfinis par les méthodologies communément reconnues en ACV est adapté pour le consommateur-citoyen qui souhaiterait s'informer sur les impacts de son mode de vie et le faire évoluer vers davantage de soutenabilité, mais se révèle mal approprié pour évaluer la façon dont une activité de production exerce une pression sur les limites planétaires. Ryberg propose donc d'intégrer les variables de contrôle des limites planétaires à l'évaluation des impacts de l'analyse de cycle de vie, en proposant de nouveaux facteurs de caractérisation représentant des catégories d'impacts.

Les nouveaux facteurs de caractérisation ainsi définis, visent à représenter la variation moyenne de la distance, pour chaque catégorie d'impact, entre l'état environnemental actuel et l'état de référence pour la variation d'une unité de flux élémentaire.

Alors que la méthode de Bjørn et Hauschild (BJØRN et al. 2015b) part des indicateurs classiques de l'ILCD pour leur appliquer un facteur de normalisation qui repose sur une "pensée limites planétaires" (i.e. si une métrique du cadre des limites planétaires correspond à une métrique d'un indicateur ILCD, elle est directement utilisée ; sinon, un seuil planétaire ou un

proxy qui s'exprime dans la bonne métrique est cherché dans la littérature sans que celui-ci appartienne *stricto sensu* au cadre des limites planétaires), Ryberg souhaite en quelque sorte "forcer" les indicateurs d'ACV à s'exprimer dans la métrique des limites planétaires en proposant une nouvelle méthode complète de caractérisation des impacts.

1.3.2 Un nouveau défi pour l'intégration des enjeux environnementaux dans la prise de décision

En octobre 2023, plus de 6000 entreprises avaient rejoint, de par le Monde, la démarche nommée "science-based target initiative", qui leur impose de se fixer des objectifs de réduction de leurs impacts environnementaux, fondés sur des bases scientifiques représentant les contraintes biogéophysiques. Pour l'instant, cette initiative concerne principalement la réduction des émissions de gaz à effet de serre en vue de respecter l'objectif de limiter à 1,5°C la hausse de la température moyenne mondiale. De tels engagements, ambitieux, peuvent être vus comme une source d'innovation. En effet, se fixer des objectifs de réduction permettant de respecter les seuils de l'espace de fonctionnement sécurisé impose des réductions beaucoup plus fortes que des objectifs de réduction basés sur des innovations incrémentales. Pour atteindre ces objectifs, les entreprises doivent mettre en place des innovations de rupture et réinterroger non seulement la façon dont elles produisent, mais également les objectifs de leur production, et en particulier en quoi cette production permet de répondre à des besoins (essentiels ou non ?) de leurs clients (RODRIGUES et al. 2021). Cela peut en théorie amener les acteurs à reconsidérer leurs activités.

1.3.2.1 Les leviers de soutenabilité dans les organisations que les limites planétaires viennent activer

Plusieurs leviers peuvent être actionnés au sein des organisations au regard du cadre des limites planétaires, qu'il s'agisse d'entreprises ou de collectivités territoriales. Nous en donnons ici quelques-uns :

- La réduction des émissions de carbone : Les entreprises peuvent travailler à réduire leurs émissions de gaz à effet de serre, en adoptant des pratiques plus durables et en investissant dans des technologies plus propres.
- La gestion durable de l'eau : Les entreprises peuvent mettre en œuvre des pratiques de gestion de l'eau responsables, réduire leur consommation d'eau et investir dans des technologies de traitement des eaux usées pour minimiser leur impact sur le cycle de l'eau douce.

- La conservation de la biodiversité : Les entreprises peuvent adopter des pratiques qui minimisent leur impact sur la biodiversité, comme la préservation des habitats naturels, la réduction de l'utilisation de produits chimiques nocifs et le soutien à des initiatives de conservation.
- La gestion des déchets : Les entreprises peuvent travailler à réduire la production de déchets, à recycler davantage et à adopter des pratiques de gestion des déchets qui minimisent leur impact sur les limites planétaires.
- Les approvisionnements responsables : Les entreprises peuvent examiner et améliorer leurs chaînes d'approvisionnement pour garantir qu'elles ne contribuent pas à des changements d'utilisation des terres non durables, à la déforestation ou à d'autres activités nuisibles pour la planète.
- L'innovation durable : Les entreprises peuvent investir dans la recherche et le développement de produits et de technologies qui respectent les limites planétaires, favorisant ainsi l'innovation durable.
- La communication transparente : La divulgation transparente des impacts environnementaux des entreprises à travers des rapports d'évaluation peut aider à sensibiliser les parties prenantes et à encourager à la responsabilité.

Pour l'heure, le cadre des limites planétaires est davantage envisagé en entreprise comme un "horizon désirable", que comme une méthode normalisée d'évaluation. C'est ainsi que des grosses multinationales, pour en citer quelques-unes à titre d'exemple, intègrent des enjeux environnementaux proches de ceux des limites planétaires, sans pour autant mettre en œuvre une méthodologie de soutenabilité "absolue" à proprement parler.

Par ailleurs, en 2023, le Ministère de la transition écologique en France a publié un rapport d'évaluation faisant état de l'environnement au regard des limites planétaires (MTES 2023), dont nous sommes co-auteurs (GONDRAN et al. 2023). L'approche ne vise pas à produire des indicateurs de suivi ou de monitoring de la politique environnementale nationale mais à produire un diagnostic à titre indicatif et incitatif (voir Figure 1.6). Cette utilisation du cadre comme outil de diagnostic fait suite au travail pionnier en France mené par l'agence d'urbanisme de la région stéphanoise, Epures, d'adaptation des limites planétaires au contexte territorial du Sud-Loire (EPURES 2021).

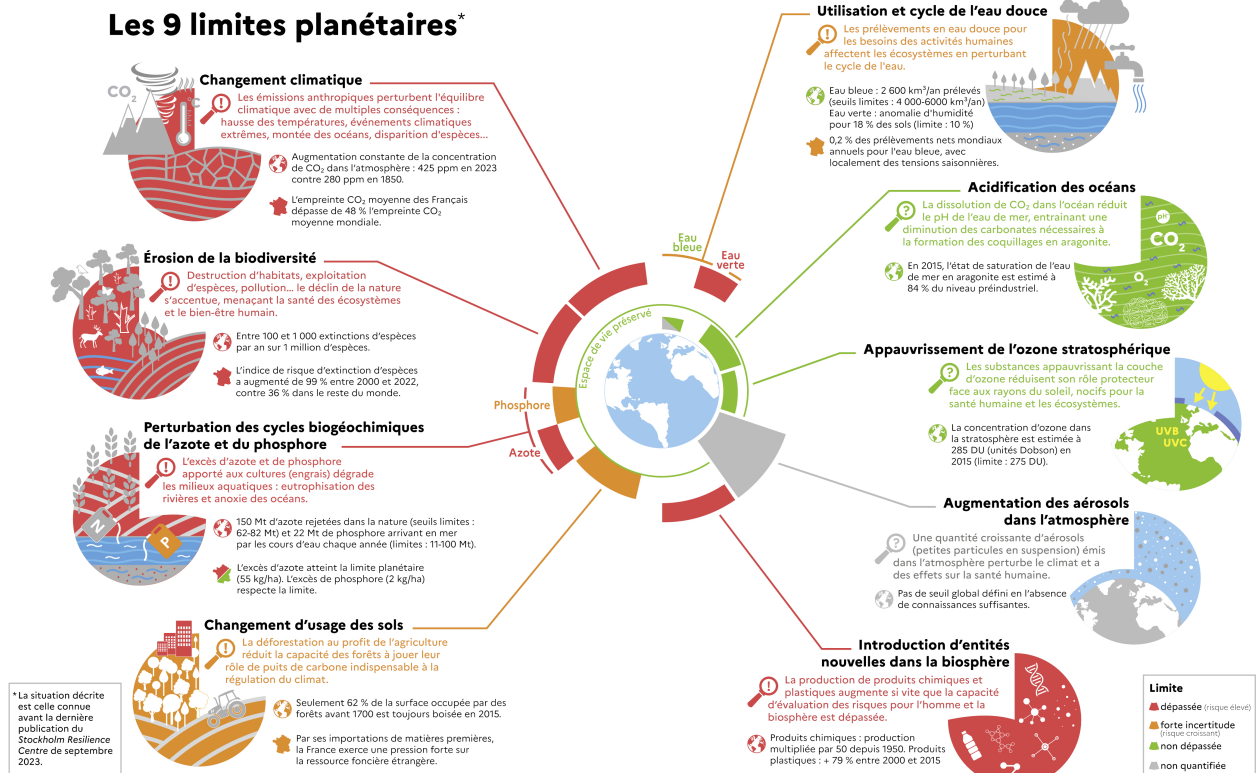


Figure 1.6 – La France face aux limites planétaires. Diagnostic réalisé par le Commissariat Général au Développement Durable (MTES 2023).

1.3.2.2 Prendre en compte les effets rebonds par une approche systémique

En 1971, P. Ehrlich et J. Holdren avaient proposé l'équation $I = PAT$ (EHRlich et al. 1971). L'impact environnemental total de l'humanité (ou d'un territoire donné) est ainsi représenté comme le produit de trois facteurs : la taille de la population (P), le niveau de consommation (A pour "Affluence", la composante matérielle du mode de vie) et le facteur technologique (T); c'est-à-dire l'intensité d'impact par valeur ou fonctionnalité créée. Les facteurs A et T sont souvent considérés comme indépendants, mais cette hypothèse s'est révélée fautive (HAUSCHILD et al. 2020). En effet, depuis les années 70, la majorité des politiques publiques environnementales portent sur le facteur technologique. Ces avancées technologiques ont pu conduire à des diminutions de consommations en énergie ou en matières par service rendu, mais elles ont aussi souvent entraîné une croissance dans l'utilisation du service en question, en réduisant le coût unitaire du service... suscitant alors une augmentation globale de la consommation. C'est ce que l'on nomme "l'effet rebond" (HAUSCHILD et al. 2020). Une évaluation environnementale classique, exprimée par fonction rendue (km parcouru, par exemple) ne permet pas d'anticiper cet effet rebond. Il apparaît donc indispensable de combiner l'amélioration de l'éco-efficacité à la définition de limites globales à notre consommation afin de

limiter l'effet rebond. En proposant un cadre d'analyse plus global, l'évaluation absolue de la soutenabilité pourrait permettre d'éviter, ou a minima, de mettre en évidence, les effets rebonds potentiels (HAUSCHILD et al. 2020).

Sans une évaluation absolue de la soutenabilité, les améliorations proposées par les entreprises sont majoritairement technologiques et incrémentales. Elles ne permettent pas d'atteindre les réductions d'impacts nécessaires pour rester dans l'espace de fonctionnement sécurisé (HAUSCHILD et al. 2020). Ce constat ne remet pas en cause la nécessité de développer des technologies plus éco-efficientes, mais appelle à intégrer des questionnements liés à la justice sociale et à l'éthique au sein même de l'évaluation environnementale. Cela semble nécessaire pour stimuler les changements de paradigmes, tant économiques que sociaux, bien plus radicaux qui sont nécessaires pour limiter l'effet rebond.

1.3.2.3 Le "donut" : une approche plus facilement actionnable par les organisations

Le "donut" économique de Kate Raworth est un concept introduit par l'économiste britannique Kate Raworth dans son livre *Doughnut Economics : Seven Ways to Think Like a 21st-Century Economist* (Économie du Donut : Sept manières de penser comme un économiste du 21e siècle) (RAWORTH 2017). Publié en 2017, ce modèle vise à redéfinir les objectifs de l'économie en s'appuyant sur des limites planétaires.

Le modèle du donut présente deux limites fondamentales qui définissent les domaines dans lesquels l'humanité doit évoluer pour assurer une prospérité durable. Ces deux limites sont représentées graphiquement sous la forme d'un donut avec deux anneaux concentriques (voir Figure 1.7) :

- L'anneau intérieur (le plancher social) : Il représente les besoins fondamentaux de l'humanité, tels que l'eau potable, l'alimentation adéquate, l'éducation, la santé, le logement et l'énergie. L'objectif est que personne ne devrait tomber en dessous de ces seuils. Ces objectifs sont à rapprocher des très connus Objectifs de Développement Durable de l'ONU (NATIONS 2015), pour ceux du moins du volet "social". Ce plancher est à mettre plus généralement en rapport avec l'approche de la pauvreté par la privation : les individus, les familles et les groupes de la population peuvent être considérés comme étant en situation de pauvreté lorsqu'ils manquent des ressources nécessaires pour obtenir les types d'alimentation, participer aux activités, et avoir les conditions de vie et les commodités

qui sont habituelles, ou du moins largement encouragées ou approuvées, dans les sociétés auxquelles ils appartiennent (TOWNSEND 1987).

- L'anneau extérieur (le plafond environnemental) : Il représente les limites environnementales au-delà desquelles l'activité humaine ne doit pas aller pour éviter de compromettre la stabilité de la planète. Ces limites comprennent le changement climatique, la perte de biodiversité, l'utilisation de l'eau douce, l'acidification des océans, etc.

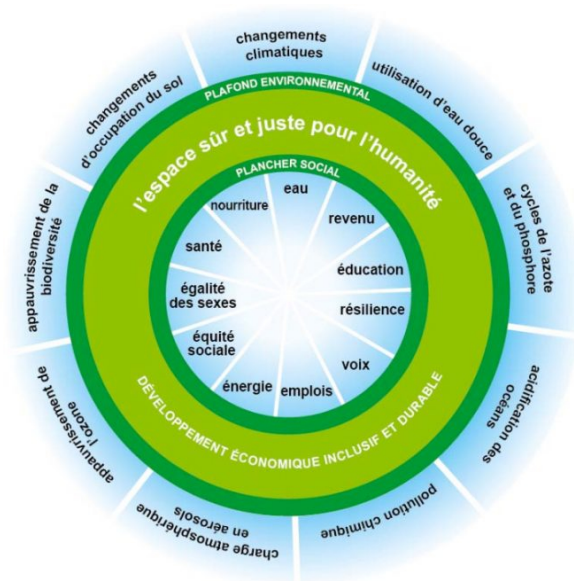


Figure 1.7 – Le "donut" de Kate Raworth (RAWORTH 2017), qui s'appuie sur les limites planétaires pour définir son "plafond" environnemental.

L'objectif économique idéal, selon le modèle du donut, est de maintenir l'activité économique de la société à l'intérieur du plafond environnemental tout en garantissant que personne ne tombe en dessous du plancher social. Cela représente un équilibre entre la prospérité humaine et la durabilité environnementale.

Le modèle du donut est devenu un outil visuel puissant pour repenser la croissance économique traditionnelle et promouvoir un modèle économique plus durable et éthique. Il a été adopté par certains gouvernements, organisations et communautés pour orienter leurs politiques et leurs actions vers un avenir plus durable.

En termes d'application du "donut", Amsterdam est devenue en 2015 l'une des premières villes à adopter le modèle du "donut" économique de Kate Raworth dans sa planification urbaine. La municipalité d'Amsterdam a intégré ce concept dans sa stratégie de développement

durable, en cherchant à créer une ville prospère tout en respectant les limites planétaires. La ville a notamment proposé un Plan pour une Économie Circulaire en élaborant une série de mesures ambitieuses pour devenir entièrement circulaire d'ici à 2050. Ce plan s'inscrit dans la lignée des principes de la "*Donut Economics*" en visant à réduire les déchets, à maximiser la réutilisation des ressources, favoriser les modes de transport durables, investir dans des bâtiments écoénergétiques, promouvoir l'agriculture urbaine, adopter des pratiques de gestion des déchets plus durables, ou encore favoriser les partenariats public-privé qui visent à développer des solutions innovantes et s'inscrivent dans des pratiques commerciales responsables. Plus opérationnellement, Amsterdam utilise le modèle du "donut" comme guide pour équilibrer les aspects sociaux et environnementaux de sa planification urbaine. L'objectif est de garantir une vie décente à tous les citoyens tout en restant dans les limites environnementales de la planète. Amsterdam a également mis l'accent sur l'engagement citoyen en encourageant la participation des résidents dans le processus de planification urbaine. Le donut a ainsi été l'occasion de réfléchir à de nouveaux dispositifs de participation citoyenne, avec cette question "quelle est votre vision du plancher social et du plafond environnemental", s'arrogant par la même occasion de la dimension purement scientifique du cadre des limites planétaires. Cela inclut des initiatives visant à recueillir les commentaires des citoyens et à intégrer leurs besoins dans la planification de la ville (RAWORTH et al. 2020).

Cette utilisation du donut est à rapprocher, en France, de l'initiative portée par la ville de Grenoble, qui a entrepris de réaliser son "portrait donut" (GRENOBLE 2023). L'ambition n'était pas de respecter à la lettre la définition des variables de contrôles donnée par le cadre des limites planétaires, mais d'envisager un diagnostic environnemental et social qui permette de positionner la ville par rapport à une référence, en l'occurrence la moyenne nationale. L'enjeu est d'en faire un document partageable propre à susciter le débat et la discussion, par les différents services de la ville, les institutions chargées des politiques urbaines, et les citoyens.

Des travaux récents en France comme ceux du chercheur Nicolas Desmoitier ont cherché à opérationnaliser davantage le "donut" dans un contexte de politique publique en proposant une batterie d'indicateurs actionnables et capables de rendre compte de la dimension "juste" de l'espace de sécurité sûr et juste (DESMOITIER et al. 2023). Les chercheurs identifient ainsi 15 dimensions constitutives du socle social d'un espace juste : l'alimentation, l'eau, accès aux soins, au logement, à l'énergie, à la sécurité et à la justice, au travail, le niveau de revenu, d'éducation, l'égalité de genre, le niveau de bien-être, la santé, l'équité sociale, les droits démocratiques.

1.4 Conclusion et introduction des hypothèses de recherche

Si l'article dont nous venons de reproduire le contenu dans les parties précédentes adopte un ton très généraliste, il a du moins le mérite de souligner les différentes perspectives depuis lesquelles l'objet "limites planétaires" peut être apprécié. Le cadre des limites planétaires peut en effet être considéré à l'aune des sciences du système Terre (partie 1 de l'article), du champ de l'évaluation environnementale (partie 2 et 3) et de ce que l'on peut appeler "le champ de l'action", c'est-à-dire les utilisations (déjà) opérationnelles du cadre (partie 3). Nous allons à présent développer ces différentes perspectives, que nous appellerons désormais des *modalités d'atterrissage*, afin de positionner notre travail de thèse dans la diversité de ces modalités d'atterrissage possibles (4.1). Nous présenterons ensuite la façon dont nous avons défini un cas d'étude, qui restreint l'étude à une unique limite et une unique version du cadre (4.2 et 4.3). À l'aune de ce positionnement et de ce cas d'étude, nous avancerons nos hypothèses directrices de recherche (4.4). Nous présenterons finalement l'organisation générale du manuscrit (4.5).

1.4.1 Les modalités d'atterrissage du cadre des limites planétaires

Nous proposons d'abord de requalifier le terme d'*échelle locale* (ou application/déclinaison *locale*) que nous avons utilisé jusqu'à présent sans trop de rigueur. Nous justifierons alors le choix du terme "système sous-global" et nous verrons quel cas particulier constitue le "territoire" dans le spectre large des systèmes sous-globaux. Nous proposerons ensuite une synthèse des différentes modalités d'atterrissage possibles pour les limites planétaires entre l'échelle globale et un système sous-global.

1.4.1.1 De l'échelle locale au système sous-global

Plutôt que de parler d'*échelle locale*, nous parlerons de système sous-global, en empruntant l'expression à (LI et al. 2020), qui nous paraît moins ambiguë car moins proche de la notion de "territoire" dans un contexte francophone et non réductible à une seule question d'échelle géographique. En effet, les parties précédentes l'ont mis en avant, les applications du cadre ne

portent pas exclusivement sur des approches géographiques (comment descendre d'échelle?) mais aussi sur des approches fonctionnelles (comment représenter une activité, un secteur, dans un indicateur basé les limites planétaires?). Aussi, nous proposons de définir un système sous-global selon deux dimensions : d'une part l'emprise géographique du système étudié, et d'autre part la fonction opérée par ce système (voir Figure 1.8). Cette montée en généralité par la notion de *système sous-global* nous permettra ensuite de mieux cerner ce que recouvre l'objet *territoire* qui est au coeur de ce travail de thèse.

L'emprise géographique est simplement définie comme le continuum allant d'un simple *site* (qui se résout à un point dans l'espace, comme une infrastructure) jusqu'à un ensemble de *zones* qui peuvent ne pas être contiguës (qui sont définies par un ensemble de surfaces, par exemple les frontières nationales des États-membres de l'Union Européenne). Évidemment, tout système dispose d'une marge de manoeuvre sur cette échelle : si le système étudié est un produit, on peut le positionner sur un unique site ou sur plusieurs sites selon que l'on considère une ou plusieurs étapes de sa chaîne de valeur. L'enjeu relève surtout de la distinction entre un *site* (ou ensemble de sites) et une *zone* (ou ensemble de zones), qui repose respectivement sur la nature ponctuelle ou diffuse des pressions environnementales étudiées (c'est-à-dire sur la possibilité d'attribuer les pressions à une source déterminée ou non) (BJØRN et al. 2019). Cette proposition ne vaut pas généralité, mais nous semble permettre de positionner l'échelle d'un système avec le moins d'ambiguïté possible.

Pour ce qui est de la deuxième dimension, la fonction opérée par le système, la distinction que nous proposons est également très générale, entre système unifonctionnel (qu'on pourrait assimiler à une unique unité fonctionnelle dans le langage de l'ACV) et système multifonctionnel, sachant qu'un système multifonctionnel peut toujours être vu comme un ensemble de systèmes unifonctionnels (RODRIGUES et al. 2021).

Dans notre spectre des systèmes sous-globaux, un territoire est alors un système multifonctionnel spatialement défini par une zone ou un ensemble de zones.

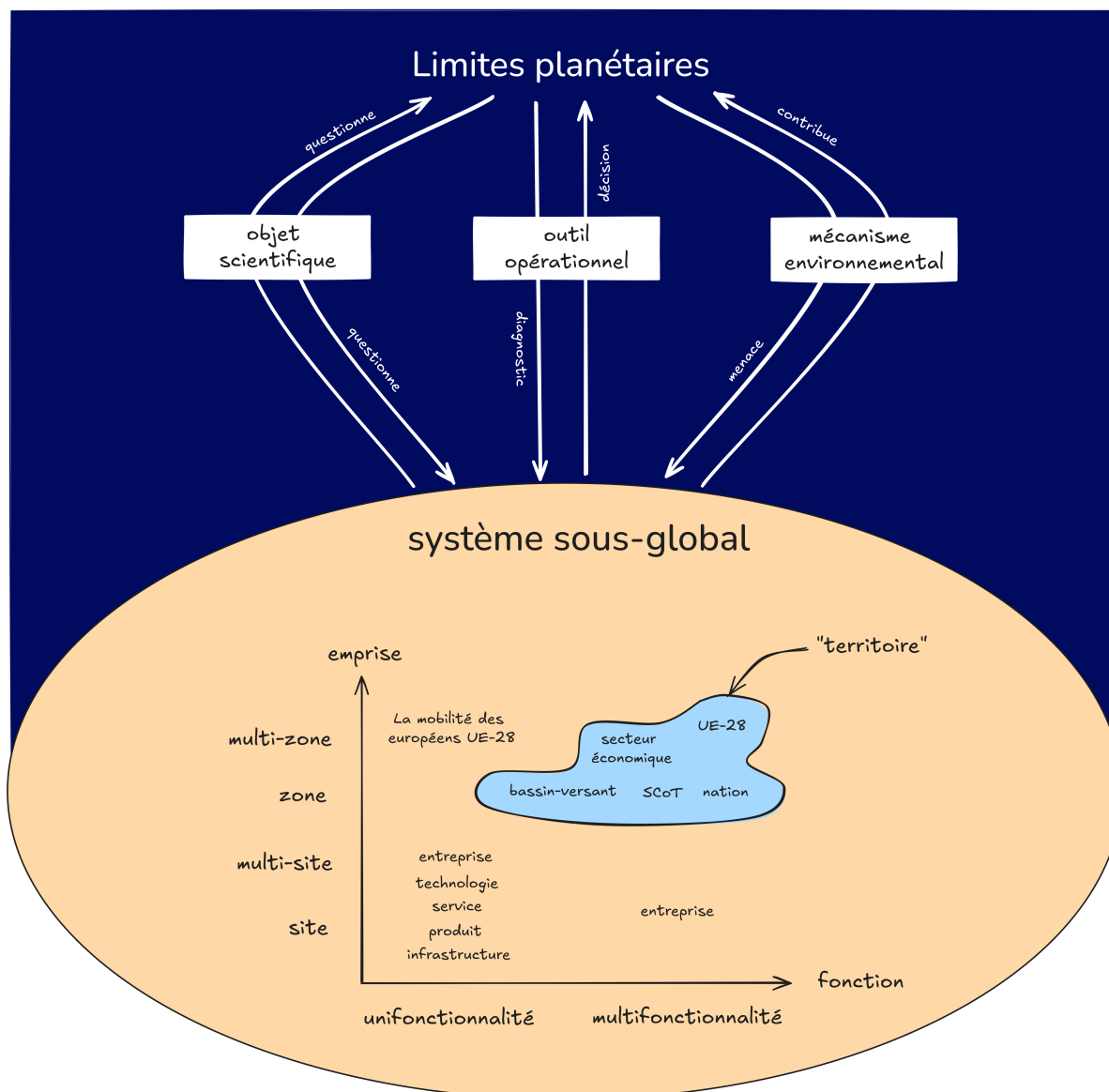


Figure 1.8 – Les différentes modalités d’atterrissage du cadre global vers un système sous-global

1.4.1.2 Les différentes modalités d’atterrissage du cadre des limites planétaires

Le type de système sous-global étudié n’a pas d’importance, à notre sens, sur les différentes modalités d’atterrissage du cadre global des limites planétaires vers ce système. Nous préférons d’ailleurs parler d’*atterrissage* plutôt que d’*application* car, précisément, toute utilisation du cadre au regard d’un système sous-global ne constitue pas nécessairement une application, ou déclinaison, du cadre à cette échelle. Nous proposons de réduire au nombre de trois ces différentes modalités d’atterrissage : le cadre global comme (1) "mécanisme environnemental", comme (2) "outil opérationnel" et comme (3) "objet scientifique" (voir Figure 1.8).

Dans la modalité (1), le cadre global des limites planétaires est envisagé comme un cadre de dynamique des systèmes dans lequel on viendrait modéliser la chaîne d'impact allant du système sous-global vers le système Terre (contribution du système sous-global à l'atteinte d'une limite) ou l'inverse (menace de l'instabilité du système Terre sur le système sous-global).

La modalité (2) englobe à proprement parler les applications du cadre, c'est-à-dire la projection des indicateurs du cadre global vers des indicateurs adaptés au système sous-global. Dans cette modalité, nous distinguons au moins trois variantes, qui sont issues de nos réflexions et discussions avec différents acteurs du "champ de l'action" pendant ce travail de thèse² :

- les limites planétaires comme "canevas" : le cadre des limites planétaires peut être envisagé dans une approche qualitative, comme outil de dialogue, de concertation, de sensibilisation ou encore d'éducation environnementale auprès des parties prenantes du système étudié. C'est typiquement dans cette catégorie que l'on peut positionner le travail de Raworth sur la construction du *donut* (RAWORTH 2017).
- les limites planétaires comme "observatoire environnemental" : le cadre est ici envisagé de façon quantitative dans un but d'évaluation environnemental. Il s'agit de développer des indicateurs fidèles et cohérents à ceux du cadre global mais adaptés au (ou représentatifs du) système étudié. L'application du cadre sert ici de diagnostic environnemental. C'est typiquement la démarche plébiscitée dans le champ de l'AESA, et qu'on retrouve également dans le champ des "*donut Economics*" (FANNING et al. 2022)³. Notre travail de thèse s'inscrit également dans cette démarche, qui se justifie par le contexte d'intervention territoriale pour le syndicat mixte en charge du pilotage du SCoT Sud-Loire.
- les limites planétaires comme "planification" : cette application du cadre peut être vue comme la suite logique de l'application précédente, où les indicateurs développés à partir du cadre pour le système étudié servent non seulement de diagnostic mais aussi d'outil de suivi dans une politique environnementale.

Enfin, la modalité (3) d'atterrissage, comme objet scientifique, n'a pour sa part pas été explicitement évoquée dans l'article reproduit plus haut. C'est qu'il s'agit d'une modalité transversale aux deux précédentes, et d'une modalité réflexive qui vient questionner à la fois le

2. Voir notamment la présentation donnée le 18/01/2024 pour le COMETE, la communauté écologie et territoires rattachée au Ministère de la transition écologique : <https://territoires-en-transition.ecologie.gouv.fr/#page.type.events.id.65a7e077a6f9334ec0473ab1>

3. Voir le projet des "tableaux de bord" nationaux porté par l'Université de Leeds : <https://goodlife.leeds.ac.uk/national-trends/country-trends/#GBR>

cadre des limites planétaires (les enjeux environnementaux qu'il prend en compte reflètent-ils les enjeux expérimentés localement ? peut-on se satisfaire d'un cadre qui ne prend en compte que les processus de régulation du système Terre et pas les ressources disponibles ?) et le système étudié (n'y a-t-il pas une taille ou échelle critique de système pour une application du cadre ? vaut-il mieux quantifier la capacité de charge associée au système étudié ou bien la contribution du système à la capacité de charge globale ?). C'est la tension entre modalité (2) et (3) que nous nous proposons d'étudier.

1.4.2 Resserrer l'étude

Si la thèse porte sur la territorialisation des limites planétaires, il nous faut resserrer davantage son objet d'étude pour aboutir non seulement à une hypothèse directrice de recherche, mais aussi se doter d'un cas d'étude à partir duquel mettre à l'épreuve cette hypothèse. Se posent d'emblée deux questions : à partir de quelle version du cadre des limites planétaires et à partir de quelle limite construire un cas d'étude ?

On pourrait penser *a priori* que le choix d'une limite et d'une version du cadre implique un positionnement disciplinaire spécifique (climatologie, agronomie, hydrologie, physico-chimie des milieux, biologie, etc). En réalité, un tel choix n'est pas à envisager sous un angle disciplinaire mais méthodologique, comme choix d'une méthode d'évaluation environnementale à l'intérieur d'un seul et même champ disciplinaire, celui de l'évaluation absolue de la soutenabilité environnementale (AESAs). C'est du moins dans ce sens que pousse, notamment, le chercheur Anders Bjorn, à qui l'on doit l'une des premières méthodes de couplage de l'analyse de cycle de vie et des limites planétaires (BJØRN et al. 2015b) et premier instigateur du concept de "soutenabilité absolue" (BJØRN et al. 2013), dans une réflexion plus conceptuelle sur ce qu'est l'évaluation absolue de la soutenabilité environnementale, et qui s'émancipe du cadre des limites planétaires à proprement parler (BJØRN et al. 2019). Anders Bjorn et ses collègues proposent ainsi en 2019 un cadre méthodologique à portée normative dont l'ambition est de structurer en champ scientifique l'évaluation absolue de la soutenabilité. En tant qu'indicateur, l'évaluation dite "absolue" est antérieure à ces travaux et, à notre connaissance, indépendante des travaux d'Anders Bjorn et ses collègues de la *Denmark Technical University* (DTU) puisque ce caractère absolu d'un indicateur de soutenabilité nous semble remonter à la publication de (MOLDAN et al. 2012) dans leur laboratoire de la *Charles University Environment Center* en République Tchèque.

[...] indicators should take into account absolute levels of at least the environmental variable (pressures). In other words, we need some targets to show policy makers if there is a threshold for reduction or a ceiling that allows a further increase in pressures. Unless a particular indicator is related to a threshold (i.e. a limit) of sustainable use, one cannot make the judgment that it shows sufficient relative or absolute decoupling to support the conclusion that more sustainable patterns of consumption and production have been achieved. (MOLDAN et al. 2012)

Si une définition de l'indicateur absolu est posée (il s'agit d'un indicateur basé sur une cible, un seuil, une limite environnementale) ainsi qu'une ambition d'aide à la décision, il n'est pas présenté comme une démarche scientifique et n'a pas l'ambition de constituer un champ à part entière de l'évaluation environnementale. C'est bien là l'intérêt de la publication d'Anders Bjorn et ses collègues (BJØRN et al. 2019) qui présentent une méthode (un cadre) générique pour l'évaluation absolue, et contribuent par là même à la constituer en champ scientifique.

La méthode propose de distinguer entre plusieurs étapes, ou composantes, d'une telle évaluation : d'une part, les choix méthodologiques propres à toute évaluation environnementale (objectifs de l'étude, périmètre, outil de quantification, collecte des données, etc) et, d'autre part et nouvellement, le choix d'une norme absolue, parmi une liste suggérée par les auteurs (*Context-based sustainability, Ecological footprint, Human appropriation of net primary production, planetary boundaries, science-based targets, water footprint*). À cet égard, le cadre des limites planétaires et ses différentes versions, constituent des candidats parmi d'autres pour le choix d'une norme absolue. Le choix d'une limite à l'intérieur de l'une des versions du cadre n'est alors plus à envisager sous l'angle disciplinaire que nous mentionnions plus haut, mais comme choix méthodologique pour l'évaluation environnementale : tout dépend de ce que l'on cherche à évaluer, et non pas du champ de recherche auquel on se rattache, puisque, précisément, le champ est celui de l'AESA.

Ainsi présenté, le choix d'une version du cadre des limites planétaires est avant tout un choix méthodologique ou normatif intégré dans une démarche d'AESA. Pour ce travail de thèse, nous faisons le choix d'étudier le cadre des frontières du système Terre (ROCKSTROM et al. 2023), qui se justifie d'une part par une considération temporelle : il s'agissait pendant le temps de la réalisation de cette thèse (2021-2024) de la version la plus actualisée des limites planétaires

(bien que lui ait ensuite succédé l'actualisation proposée par (RICHARDSON et al. 2023)); et d'autre part par la visée explicite de transformation de l'action environnementale visée par la publication de (ROCKSTROM et al. 2023) en définissant des variables de contrôle d'abord locales.

1.4.3 Conduire une évaluation environnementale ou questionner l'évaluation environnementale ?

Définir un cas d'étude ne s'arrête pas au choix d'une version du cadre et d'une limite, que nous n'avons d'ailleurs pas encore précisée. Puisque nous nous plaçons dans un contexte disciplinaire d'AESA, le choix de ce que l'on se propose d'évaluer revêt deux dimensions : une dimension pratique, où l'évaluation porte directement sur un système dont on cherche à qualifier ou quantifier la soutenabilité (par exemple l'intégrité des cours d'eau sur un territoire donné); et une dimension épistémologique, où l'évaluation d'un système permet d'évaluer la méthode d'évaluation elle-même, c'est-à-dire d'interroger le champ de l'AESA sur ses choix méthodologiques (pourquoi retenir tel seuil ? pourquoi utiliser telle méthode de quantification ? etc). Se joue ici une tension entre la modalité d'atterrissage (2- les limites planétaires comme outil opérationnel), et en particulier comme "observatoire environnemental" voire comme "outil de planification", et la modalité (3- réflexivité sur les objets manipulés).

Au regard de la littérature existante sur le cadre des limites planétaires, on peut isoler un certain nombre de points d'attention retenus par les chercheurs du champ quant aux questions qu'ils adressent à leur champ, soit sur un mode implicite (par exemple, dans les revues de la littérature, par les catégories utilisées pour classer les publications) soit explicitement par des propositions de réflexion ou des cadres analytiques questionnant les pratiques du champ. Plus précisément, on peut distinguer entre plusieurs types de publications :

1. les révisions du cadre, c'est-à-dire les publications qui proposent de revoir la formulation d'une ou plusieurs limites et qui aboutissent effectivement à de nouvelles versions du cadre des limites planétaires ;
2. les méta-analyses, qui sont les publications qui proposent la mise en perspective d'une ou plusieurs limites au regard de réflexions personnelles ou d'autres cadres d'analyse ;
3. les revues de la littérature, c'est-à-dire les publications qui établissent l'état de l'art des publications relatives à une ou plusieurs limites ;
4. les applications, qui constituent des évaluations environnementales à part entière dans

lesquelles un ou plusieurs systèmes se voient évaluer. Dans cette dernière catégorie, on peut encore distinguer :

- (a) les publications qui portent sur l'évaluation environnementale d'une unique limite (donc d'un unique compartiment environnemental) ;
- (b) les publications qui procèdent à une évaluation plus générale du cadre, dans un souci de systématiser l'évaluation sur le plus de limites possible.

Les publications du type (3- revues de la littérature) proposent des catégories pour le classement des publications produites dans le champ, qui nous semblent constituer un point de départ heuristique pour identifier les questionnements épistémologiques que le champ se pose à lui-même.

Mo Li et ses collègues par exemple (LI et al. 2021), dans une revue générique portant sur les applications "à différentes échelles" (*across scales*) du cadre des limites planétaires, distinguent : le processus planétaire sur le(s)quel(s) porte l'étude, l'échelle (e.g. "*global*", "*regional*", "*biome*", "*basin*", etc), la variable de contrôle utilisée, la valeur seuil de la limite, la valeur actuelle de la limite, le cadre de référence, mais aussi le champ de recherche auquel se rattache implicitement l'étude.

Dans une revue centrée sur la limite eau, Jonas Bunsen et ses collègues (BUNSEN et al. 2021) proposent les catégories de classement suivantes : la variable de contrôle utilisée, la formule de calcul de l'espace de fonctionnement sûr, la valeur de l'espace de fonctionnement sûr (ou valeur seuil de la limite), la valeur actuelle de la limite, la résolution spatiotemporelle) et la famille méthodologique de l'étude menée (révision, commentaire, analyse de cycle de vie, analyse entrées-sorties, méthodologie spécifique dite "*stand-alone application*").

On peut également citer deux revues de la littérature bien connues dans le champ, celle d'Anders Bjorn (BJØRN et al. 2020a) et celle de Morten Ryberg (RYBERG et al. 2020), qui sont des revues spécifiques au champ de l'AESA. Bjorn et ses collègues proposent ainsi les catégories suivantes : principe de calcul pour l'estimation de la capacité de charge (publication utilisée pour la référence environnementale, cible spécifique), composante de l'évaluation spatialement différenciée, principe d'assignation, cas d'étude. Ryberg et ses collègues pour leur part mettent en avant les catégories suivantes : le processus planétaire sur le(s)quel(s) porte l'étude, l'échelle du cas d'étude (e.g. pays, industrie, secteur, etc), le principe d'allocation.

De ces revues de la littérature, nous identifions certaines catégories communément utilisées : le cas d'étude, l'échelle de l'étude, la version du cadre des limites planétaires qui sert de méthode d'évaluation absolue, la famille méthodologique de l'étude menée, l'indicateur utilisé pour la variable de contrôle, l'indicateur utilisé pour la limite.

La question de l'échelle de l'étude, ou plus précisément des imbrications entre échelles, nous paraît primordiale pour ce travail de thèse dont l'objet tient à l'application *territoriale* du cadre des limites planétaires.

1.4.4 La donnée environnementale au cœur de la question de l'échelle d'étude

Le contexte d'intervention territoriale de cette thèse nous amène à être davantage sensible à la tension généralement relevée dans la littérature entre l'échelle globale et l'échelle locale : l'évaluation environnementale du territoire vise-t-elle à quantifier la contribution du territoire au dépassement des limites globales ? Ou bien vise-t-elle à quantifier les pressions exercées par le territoire sur les limites de son environnement local ?

Pour le moment, nous avons seulement relevé l'ambiguïté que revêt le terme d'*échelle locale* et lui avons préféré celui de *système sous-global*, qui ajoute à la seule question de l'échelle celle de la fonctionnalité et qui désambiguïse le raccourci trop simple entre le *local* et le *territoire*. Cependant, nous n'avons pas encore évoqué la question de la confrontation entre échelles, à savoir l'échelle globale et l'échelle du système sous-global. Est-ce que la limite associée à un système sous-global doit être considérée comme une fraction de la limite globale (par exemple une part de la consommation soutenable globale d'eau douce), ou est-ce que l'on peut associer au système sous-global une capacité de charge explicitement sous-globale (la disponibilité de la ressource en eau à l'échelle même du système sous-global) ?

Cette tension est généralement conceptualisée à travers la distinction entre approches *bottom-up* et *top-down* (RYBERG et al. 2020 ; LI et al. 2020 ; VEA et al. 2020). Certains auteurs n'ont toutefois pas recours à ces catégories conceptuelles, par exemple (DEARING et al. 2014) qui évoquent l'approche *bottom-up*, par le concept de *regional safe operating space*, ou (ZIPPER et al. 2020) qui parlent de *local safe operating space*. Ce que recouvrent généralement les

catégories *bottom-up* et *top-down* est bien résumé par (BAI et al. 2024).

The 2009 publication of planetary boundaries framework stimulated efforts to allocate shares of a global ‘safe operating space’ to actors operating at different scales. This process is what we term cross-scale translation, and it can follow top-down, bottom-up or combined approaches. Top-down allocation, also called the ‘fair shares’ approach, involves apportioning global budgets to national, sub-national, sectoral and individual business scales. [...] There are also efforts to adapt the planetary boundary frameworks to the national or regional scale using a bottom-up approach. A bottom-up approach involves determination of a sub-global or local safe operating space and is typically based on nationally or locally relevant variables and boundary values such as determining basin-scale environmental water-flow requirements. [...] Other studies have aggregated environmental pressures at a lower scale (industries) to produce environmental pressures at a higher scale (sub-national or national). [...] The combined approach utilizes both top-down and bottom-up approaches and suggests ways to reconcile differences in the values obtained. (BAI et al. 2024)

Pour contribuer à cette discussion, il nous semble nécessaire d’amener une distinction peu souvent faite explicitement, à savoir de distinguer chacune des deux approches *bottom-up* et *top-down* selon que l’on parle de l’indicateur de pression environnementale ou de l’indicateur de cible environnementale. En effet, un indicateur de type "limite planétaire" est un indicateur de type *distance-to-target* (HEIJUNGS et al. 2014) c’est-à-dire un indicateur composé d’une part d’un indicateur qui renseigne sur la pression environnementale exercée par un système et d’autre part d’un indicateur qui renseigne sur la cible environnementale à ne pas dépasser. Chacun de ces deux indicateurs peut être quantifié indépendamment (c’est en fait presque tout le temps le cas) et donc chacun par une méthode *bottom-up* ou *top-down*. C’est le cas des méthodes d’ACV qui peuvent être spatialement explicites pour la quantification de la pression (et donc de type *bottom-up*) mais qui utilisent des méthodes d’assignation du budget global pour définir les cibles à l’échelle du système (donc approche *top-down*), par exemple dans (SALA et al. 2020). (BAI et al. 2024) résolvent toutefois cette difficulté en évoquant la possibilité d’une "approche combinée" (*combined approach*).

Nous suggérons qu’une approche vraiment *bottom-up* (c’est-à-dire vraiment territorialisée) tienne à ce que la quantification de la pression et la quantification de la cible s’appuient sur

une donnée spatialement *explicite* et pas seulement spatialement *différenciée*. Par spatialement explicite, nous entendons le fait que la donnée reflète un état environnemental local réel (explicite) ; et par spatialement différenciée, une donnée qui estime l'état environnemental local sans pour autant prétendre statuer son état réel. Une donnée explicite est donc nécessairement différenciée, mais pas l'inverse. Autrement dit, nous faisons l'hypothèse que le critère discriminant pour distinguer entre approches *bottom-up* et *top-down* nous semble moins relever de l'échelle de l'étude que de ce que nous appelons l'*échelle des données environnementales*.

Sans égard à l'échelle de la donnée, une approche pourtant qualifiée de *bottom-up* peut échouer à qualifier l'état environnemental réel du système étudié, si la donnée n'est pas spatialement explicite (voir Figure 1.9). Par exemple, l'indicateur sur la frontière relative à l'eau de surface dans (ROCKSTROM et al. 2023) est un indicateur global par calcul de la fréquence des dépassements dits locaux. On rangerait alors classiquement cette méthode dans l'approche *bottom-up* (car les dépassements sont d'abord locaux) sauf que le modèle utilisé pour l'évaluation environnementale de l'unité élémentaire d'analyse (le pixel du maillage, ou niveau "local") n'a pas été calibré sur des données réelles produites sur chacune de ces unités élémentaires, et ne produit pas non plus un résultat validé à cette échelle (c'est-à-dire comparable à des données d'observation moyennant une erreur jugée acceptable) mais à l'échelle des grands bassins versants (FEKETE et al. 2002). L'approche est donc ici spatialement différenciée par modélisation mais n'est pas explicite pour autant, et donc ne constitue pas à notre sens une approche authentiquement *bottom-up*. Ceci va à l'encontre de l'exemple proposé par (BAI et al. 2024) dans l'extrait que nous avons reproduit ci-dessus : quantifier les "*basin-scale environmental water-flow requirements*" n'implique pas nécessairement une approche *bottom-up* si les données ne sont pas explicites pour l'échelle de l'indicateur proposé, en l'occurrence les bassins versants. Par exemple, dans de nombreux cas, les données disponibles sont explicites à l'échelle nationale, et les auteurs souhaitent différencier spatialement ces données à des échelles inférieures. La différenciation est alors souvent effectuée par interpolation de densité sur des grandeurs jugées corrélées (HUANG et al. 2018 ; HOEKSTRA et al. 2010a ; FLÖRKE et al. 2013). Une telle approche produit alors des données spatialement différenciées mais qui ne sont plus explicites pour la nouvelle échelle, et cette approche n'est pas authentiquement *bottom-up* car on ne peut pas conclure réellement sur le dépassement de capacités de charge locales.

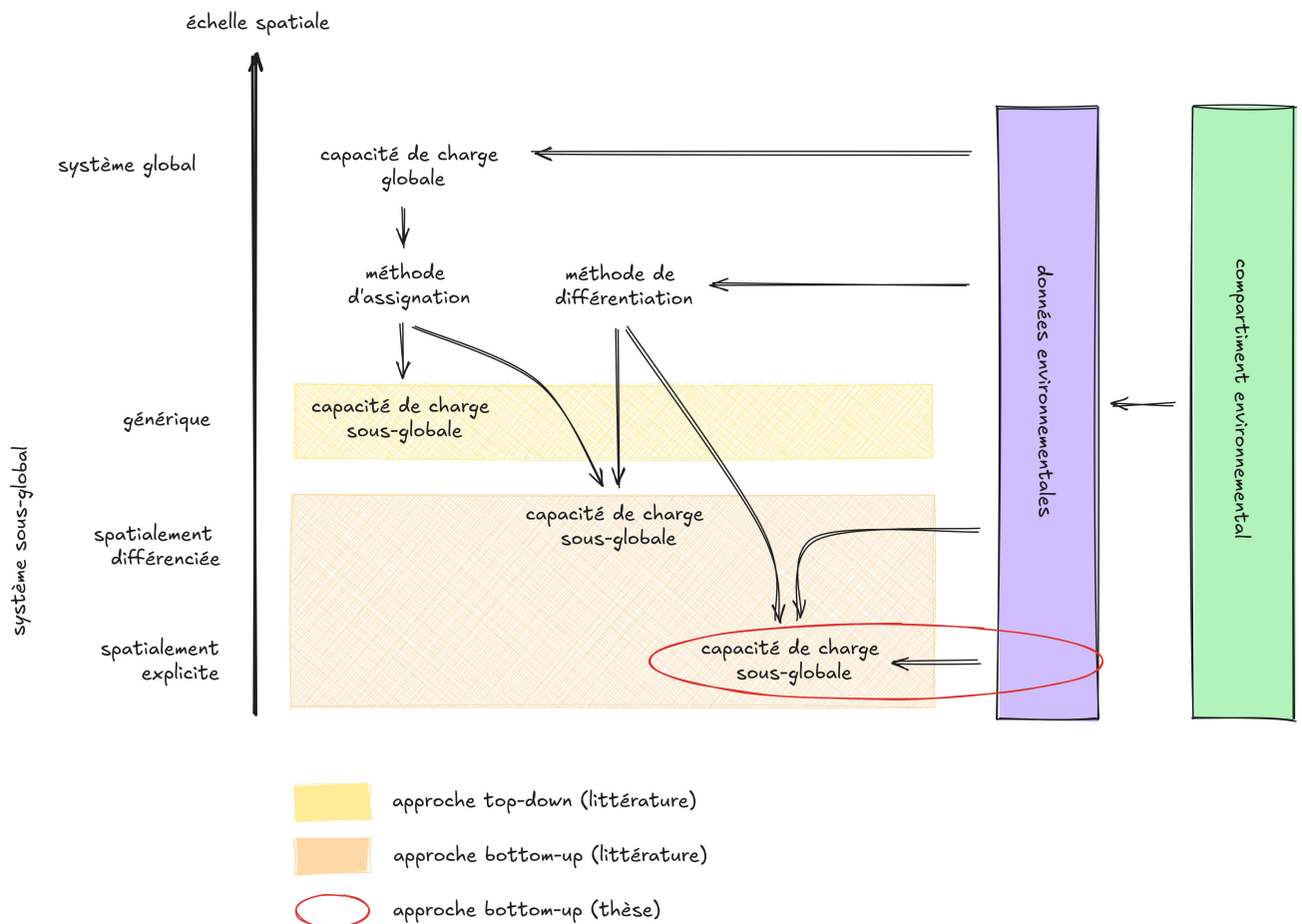


Figure 1.9 – Le rôle de l'échelle de la donnée environnementale dans une approche *bottom-up*

La question que nous adressons donc au champ de l'AESA porte sur la nécessité de prendre en compte l'échelle des données environnementales en plus de l'échelle géographique dans une démarche d'évaluation environnementale territoriale. Comment alors distinguer ce qui relève de l'échelle de l'étude et ce qui relève de l'échelle des données? En particulier, comment analyser les jeux d'échelle de la donnée où une donnée devient tour-à-tour spatialement explicite à une échelle et seulement spatialement différenciée à une autre échelle? Qu'en est-il des opérations d'agrégation et de désagrégation des données qui sont déjà passées par une opération de différenciation, par exemple lors de la mise à l'échelle des résultats au système étudié?

Notons que dans ce parti pris d'une approche "*bottom-up* basée sur des données spatialement explicites" se joue également la nécessité de *contextualiser* l'évaluation environnementale du territoire, où très souvent le territoire est réduit à une simple zone d'étude dans les publications (simple mention du nom et de la position géographique).

Par exemple, si le système étudié est la gestion des terres agricoles dans un Parc Naturel

National (PNN), la simple mention du nom du PNN ne suffit pas à contextualiser politiquement ou administrativement le territoire. Une façon de contextualiser, en termes d'action publique, le système étudié est par exemple la suivante :

[...] the Vanoise National Park promotes the use of organic farming in its territory and purchasing policy (for food, textiles, etc.). Therefore, it wanted to highlight the benefits of organic farming against intensive agriculture. However, organic farming generates a lower yield per unit of land and thus requires larger areas than intensively cultivated land to produce a similar quantity of products, so it has a larger crop and grazing EF [Ecological Footprint] [...] However, considering the negative impacts of intensive agriculture on soils (erosion, depletion of soil nutrients, etc.), it was not acceptable for a decision-maker wishing to promote organic farming to introduce into its follow-up tools, crop yields that would favor, on the cropland footprint side, intensive farming rather than organic. Therefore, the same crop yields were used regardless of the origin of the products. (GONDRAN 2012)

Dans cet exemple, où l'empreinte écologique (l'une de première méthode de ce qui ne s'appelait pas encore l'évaluation absolue de la soutenabilité) du système agricole du PNN de la Vanoise est évaluée, on peut noter l'infléchissement de la méthodologie d'évaluation au regard de la politique environnementale portée par le PNN, puisque le paramètre de productivité agricole à l'hectare est ajusté pour ne pas favoriser l'agriculture intensive au détriment de l'agriculture biologique : le but de l'évaluation étant de doter le territoire d'un outil de pilotage de sa politique environnementale, l'évaluation ne pouvait pas aller à l'encontre de la vision politique portée.

Précisément, un lien peut être fait entre la prise en compte du contexte territorial (ou périmètre d'action publique) et la caractérisation de la donnée que nous évoquons plus haut. Ce lien réside dans le système d'information environnementale (SIE) disponible sur le territoire, qui est l'émanation d'un certain mode d'action publique en matière environnementale, historiquement constitué. Le SIE peut être vu comme un intermédiaire entre une politique environnementale portée par des acteurs politiques et un territoire administrativement défini (mais qui peut être doté d'une cohérence écologique, par exemple en correspondant à la délimitation d'un bassin versant).

Aussi, la limite planétaire relative à l'eau de surface nous a semblé constituer un cas d'étude idéal pour répondre à l'interrogation que nous soulevons, qui s'intéresse à l'utilisation des données spatialement explicites pour produire une évaluation *bottom-up* du territoire, c'est-à-dire réellement territorialisée. En effet, parmi l'ensemble des limites, la limite relative à l'eau de surface décrit par nature un processus sous-global (écoulements de surface sur les bassins versants) et est sûrement celle qui trouve le plus d'écho territorial eu égard à la disponibilité d'un système d'information sur l'eau déjà mobilisé dans l'exercice des compétences eau par les communes du territoire et dans la politique de l'eau des Comités de bassins, Agences de l'Eau et services de l'État.

1.4.5 Organisation du manuscrit

Les chapitres suivants proposent ainsi une exploration du rôle des données environnementales dans une approche territoriale (c'est-à-dire *bottom-up*, dans le sens où nous l'avons définie) du cadre des limites planétaires.

Le chapitre 2 constitue une revue de la littérature sur notre cas d'étude, la frontière du système Terre relative à l'eau de surface, à la fois à l'échelle du cadre global et dans les travaux d'AESA dont le système sous-global correspond au type territorial (dans le sens où nous l'avons défini plus haut). Ces travaux, globaux et sous-globaux, seront présentés et confrontés par le prisme des données environnementales qu'ils mobilisent. Nous élaborerons alors une typologie des données environnementales en usage dans le champ, et soulignerons l'importance des données de type "observation" dans une approche *bottom-up*. Cette typologie nous permettra de préciser ce que recouvre le bloc "données environnementales" dans la Figure 1.9. Nous formulerons alors deux questions de recherche quant au rôle que peuvent jouer les données d'observation dans une approche territoriale d'AESA :

1. d'une part, les données d'observation pour l'évaluation environnementale. Schématiquement (voir Figure 1.9), cette proposition consiste à étudier précisément la flèche allant des données environnementales vers la définition d'une capacité de charge à une échelle spatialement explicite (cercle rouge) ;
2. d'autre part, les données d'observation pour la validation des données issues d'un modèle hydrologique global. Schématiquement (voir Figure 1.9), cette proposition se situe à l'intérieur même du bloc "données environnementales", par une flèche qui n'est pas représentée mais qui relierait deux types de données toutes deux définies à une échelle

spatialement explicite : les données d'observation et les données de modèle de simulation numérique.

Le chapitre 3 détaillera les propositions méthodologiques pour chacune des questions de recherche. Le chapitre 4 présentera finalement les résultats obtenus par l'application numérique de ces deux propositions à notre cas d'étude. Il s'agira de présenter la mise en œuvre numérique des traitements proposés au chapitre 3 à travers :

- le développement du logiciel **HydroTrendsFrance** pour la question (1) ;
- le développement de **relio**, une méthode numérique de *data assimilation* de données hydrologiques spatialisées pour la question (2)

Chapitre 2

Compter l'eau : panorama des données environnementales sur l'eau dans les travaux sur les limites planétaires

Introduction

Ce chapitre vise à préciser ce que nous avons qualifié de *données environnementales* dans nos hypothèses de recherche, qui porte sur le rôle des données environnementales dans la construction méthodologique d'une approche AESA *bottom-up*, faisant l'hypothèse que la seule échelle "sous-globale" ne suffit pas à définir une capacité de charge explicite pour le système étudié.

Nous proposons dans ce chapitre d'identifier les données environnementales sur l'eau mobilisées par les publications qui ont traité de la frontière du système Terre sur l'eau de surface, à la fois à l'échelle du cadre global et dans les travaux d'AESA dont le système sous-global correspond au type *territorial*. Ces travaux, globaux et sous-globaux, seront présentés et confrontés par le prisme des données environnementales qu'ils mobilisent. Nous élaborerons alors une typologie des données environnementales en usage dans le champ, et soulignerons l'importance des données de type "observation" dans une approche *bottom-up*.

Si notre cas d'étude porte sur le cadre des "frontières du système Terre" (*Earth system boundaries*), et en particulier sur la frontière relative à l'eau de surface (*surface water boundary*), ce chapitre fera également appel à la littérature portant sur le cadre des "limites planétaires" (*pla-*

netary boundaries). Nous replacerons en effet le cadre des frontières du système Terre, du moins la limite relative à l'eau dans la publication de (ROCKSTROM et al. 2023), dans sa trajectoire de parenté avec les différentes formulations de la limite relative à l'eau dans les différentes versions du cadre. Si l'esprit de cette limite dans (ROCKSTROM et al. 2023) est différente de celui des autres versions du cadre, cela signifie non pas qu'elle ait été pensée à part mais bien plutôt qu'elle en est l'émanation, comme résultat d'un processus de discussions et d'améliorations scientifiques.

Aussi le chapitre commencera-t-il par rappeler la trajectoire de la limite planétaire, puis frontière du système Terre, relative à l'eau (1.1). Une présentation détaillée de la limite sur l'eau de surface dans la version (ROCKSTROM et al. 2023) du cadre sera menée (1.2), qui sera l'occasion de présenter en détail notre cas d'étude ainsi que les données mobilisées pour le calcul de l'indicateur associé, et notamment le modèle hydrologique global (*Water Balance Model*) qui est utilisé. Sera ensuite présentée la littérature ayant entrepris l'AESA sur l'eau d'un système sous-global de type "territoire", avec une attention particulière portée au type de données mobilisées (2). Cette partie commencera par présenter la méthodologie suivie pour la constitution du corpus de publications et des jeux de données (2.1). Suivront ensuite une présentation détaillée des jeux des données représentatifs au sein du corpus (2.2) puis une réflexion sur le rôle de ce que nous appellerons l'*échelle de la donnée* dans une approche AESA *bottom-up*, au-delà du rôle de l'échelle de l'étude (3).

2.1 L'eau dans le cadre des frontières du système Terre

Dans un premier temps, nous nous intéresserons aux publications que nous avons, au chapitre précédent, catégorisées dans les types (I - révisions du cadre) et (II - méta-analyses du cadre). Ce sont en effet ces publications qui contribuent principalement au renouvellement dynamique du cadre des limites planétaires. Dans un second temps, nous présenterons en détail la limite eau dans le cadre de (ROCKSTROM et al. 2023).

2.1.1 Évolution de l'indicateur et des modèles au fil des mises à jour du cadre des limites planétaires

Depuis la première publication du cadre des limites planétaires en 2009, la variable de contrôle sur l'eau a connu des évolutions notoires, dans sa formulation, sa valeur, comme dans son périmètre. Ces évolutions peuvent être considérées comme internes à la communauté scientifique des

limites planétaires (une communauté *de fait*, qui rapproche, autour d'un même objet d'étude commun, des communautés épistémiques distinctes). Néanmoins, la façon dont est épistémologiquement et méthodologiquement abordée la question de l'eau dans les différentes versions du cadre des limites planétaires, et dans ses différentes applications sous-globales, relève d'une histoire plus longue et plus large que le cadre des limites planétaires. Nous proposons ici d'effectuer une brève histoire d'un cadre conceptuel qui émerge dans les années 1990, "la couleur de l'eau", en tant que matrice conceptuelle de l'indicateur relatif à l'eau de surface dans le cadre des limites planétaires, puis de raconter les évolutions de cet indicateur au fil des révisions du cadre.

2.1.1.1 La "grande" histoire de l'eau : un cadre conceptuel préalable au cadre des limites planétaires

Dans la version initiale du cadre des limites planétaires (ROCKSTRÖM et al. 2009b), la question de l'eau, ou plus précisément de l'*utilisation de l'eau douce (freshwater use)*, est abordée sous l'angle des interactions entre le cycle de l'eau, la régulation climatique et les services écosystémiques. Le lien entre ces trois dimensions repose sur le bilan de l'*eau verte* entre demande végétative du substrat terrestre et demande évaporative de l'atmosphère : un bilan déséquilibré en faveur de l'eau métabolisée par les végétaux (notamment pour l'agriculture) entraîne une instabilité du système climatique, notamment sur les régimes de pluie (ROCKSTRÖM et al. 2009b). Si l'accent est mis sur l'eau verte, l'indicateur finalement retenu porte sur l'eau bleue, formulé comme quantité maximale extractible du ruissellement global. Il nous semble que derrière cette position conceptuelle et cette proposition d'indicateur se jouent deux histoires (plus modestement, trajectoires) scientifiques de l'eau. L'une, plutôt d'ordre conceptuel, envisage de distinguer entre plusieurs "couleurs" de l'eau (en fait, entre plusieurs compartiments du cycle de l'eau). L'autre, davantage quantitative, entend développer des modèles computationnels visant à quantifier les stocks et les flux d'eau. Bien sûr, ces histoires et ces communautés de recherche sont liées et perméables, mais les distinguer permet, nous semble-t-il, de mieux comprendre pourquoi un tel indicateur est proposé en 2009.

2.1.1.1.1 La couleur de l'eau

En 1997, Johan Rockström soutient sa thèse de Doctorat à la faculté des sciences de l'université de Stockholm, au département d'écologie des systèmes. Sa thèse porte sur la compréhension des mécanismes agrohydrologiques à l'œuvre dans le système agraire subsaharien. L'accent est précisément mis sur les interactions entre pratiques agricoles, précipitations

et évapotranspiration (ROCKSTRÖM 1997). Dans les années qui suivent, il signe ou cosigne un certain nombre d'articles qui reprennent à son travail de thèse l'étude de l'évaporation comme paramètre décisif non plus du seul système agricole subsaharien, mais plus généralement de la production alimentaire, de la sécurité alimentaire, et des services écosystémiques à l'échelle planétaire. Parmi ces cosignataires, soulignons la présence de Malin Falkenmark. Falkenmark était une chercheuse suédoise, professeure d'hydrologie appliquée, et chercheuse émérite dans le laboratoire de Johan Rockström¹. Comme le rappelle Rockström lui-même dans un article publié en 2024 suite au décès de Falkenmark, elle est à l'origine de la "coloration" du cycle de l'eau au début des années 1990 (ROCKSTRÖM et al. 2024). En 1995, elle formalise ainsi les concepts d'eau verte et d'eau bleue dans un rapport pour la FAO intitulé "*Land-Water linkages : a synopsis*" (FALKENMARK 1995). L'idée essentielle est de souligner l'importance des précipitations incorporées dans les cultures et dans les écosystèmes naturels et anthropisés, pour décentrer la focale du simple débit des cours d'eau, qui était alors considéré comme la seule source d'eau directement interceptée. La fraction verte de l'eau est ainsi constituée de deux flux : l'évapotranspiration directe (fraction précipitée ou ruisselée de l'eau qui est évapotranspirée par les plantes) et l'évapotranspiration indirecte (fraction de l'eau qui est détournée par l'irrigation puis évapotranspirée par les plantes cultivées). Typiquement et à l'échelle de la balance hydrologique planétaire, si le cumul annuel des précipitations est de 100, alors l'eau bleue (ruisselée directement ou interceptée par la zone saturée) vaut 35 et l'eau verte (évaporée ou transpirée) vaut 65 (FALKENMARK et al. 2004).

Au-delà du cercle proche à Johan Rockström, la question de la couleur de l'eau a largement passionné les chercheurs du champ plus large des "*hydro-politics*" dans les années 1990, regroupant à la fois des géographes, des économistes et des politistes (ALLAN et al. 1993). Dans ce foisonnement de travaux, on peut trouver le concept d'*eau virtuelle*, qui se définit initialement comme la quantité d'eau nécessaire à la production d'un bien agricole (ALLAN 1998), mais qui peut être élargi à la production des biens manufacturiers (HOEKSTRA et al. 2002). La perspective initiale de ce concept n'était d'ailleurs pas tellement celle de l'évaluation environnementale ni de l'écologie des milieux, mais plutôt de l'économie politique, ou encore de la géopolitique, à savoir l'optimisation des échanges commerciaux internationaux pour limiter les déficits en eau des pays à faible niveau de ressources en eau. En augmentant l'importation des biens intensifs en eau et en réduisant l'exportation (donc la production nationale) de ces mêmes biens, c'est en effet le "déficit commercial en eau" qui est réduit

1. https://en.wikipedia.org/wiki/Malin_Falkenmark

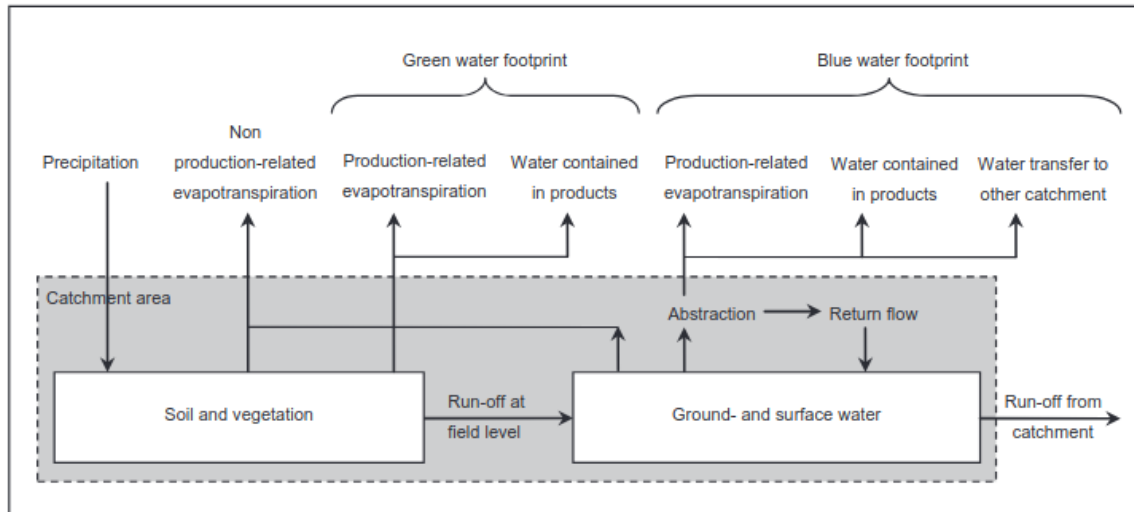


Figure 2.1 – Représentation des flux constitutifs de l’empreinte eau bleue et de l’empreinte eau verte (ALDAYA et al. 2012)

(ALLAN et al. 1993). Le concept a ensuite été repris au début des années 2000 par les chercheurs Hoekstra et Hung qui se sont proposés de procéder à une quantification systématique des flux d’eau virtuelle embarqués dans les biens agricoles (HOEKSTRA et al. 2002). Ce travail les a conduits à proposer le concept d’*empreinte eau* (voir Figure 2.1), qu’ils ont quantifié là aussi de façon systématique pour les différents pays du monde, reprenant ainsi l’esprit de l’*empreinte écologique*, un indicateur proposé au début des années 1990 (REES 1992). L’empreinte eau se définit par rapport à l’eau virtuelle : il s’agit, pour tout pays, de la somme des consommations d’eau réalisées sur le territoire national et des importations d’eau virtuelle (nettes d’exportations). Les consommations d’eau se définissent à leur tour comme les prélèvements nets de rejets (i.e. l’eau qui est effectivement incorporée dans un bien et qui n’est pas restituée). Ces prélèvements sont de type *eau bleue* pour désigner les eaux de surface et souterraines, et de type *eau verte* pour désigner les précipitations interceptées pour les besoins d’une production. À partir des années 2010, toujours sous l’impulsion du chercheur Hoekstra, le concept d’*eau grise* est ajouté à la liste des couleurs de l’eau qui sont comptabilisés dans l’indicateur d’empreinte eau (HOEKSTRA et al. 2011b). Il s’agit de la quantité d’eau nécessaire à la dilution des pollutions contenues dans les eaux usées (industrielles, agricoles et urbaines). C’est ce travail d’ampleur conduit dans les années 2000 sur la quantification des empreintes eau, à la fois à l’échelle des biens et des nations, qui a véritablement ouvert la voie aux nombreuses applications futures dans le champ de l’évaluation environnementale, nous le verrons par la suite.

2.1.1.1.2 La modélisation du cycle de l'eau

Au-delà de ces couleurs de l'eau qui ont débouché sur un champ de recherche à part entière, Rockström met en avant une tout autre histoire en soulignant l'influence des travaux pionniers dans le champ de la modélisation hydrologique globale sur les recherches entreprises par Falkenmark :

"[her] research was heavily inspired by her fellow global climatologists and hydrologists like Mark Isaakovich L'vovich, Michail Budyko, Peter Gleick, Charles Vörösmarty, and Sandra Postel. In particular, L'vovich's World Resources and their Future strongly influenced Falkenmark's thinking. Here, for the first time, freshwater resources were mapped across the entire world, presenting hydroclimatic differences between different regions, in particular in terms of the partitioning between stable runoff and vapour fluxes of evapotranspiration" (ROCKSTRÖM et al. 2024)

C'est qu'il y a en effet à cette époque un ensemble de chercheurs qui s'emploient à développer des modèles hydrologiques globaux. Sans en faire ici toute l'histoire, le nom de Vörösmarty qui est cité ici retient notre attention puisqu'il est à l'origine du *Water Balance Model* qui servira à l'actualisation par Johan Rockström du cadre des limites planétaires en 2023. Ces recherches s'inscrivent dans le mouvement plus global des sciences de la systémique (ou dynamique des systèmes). Dans leur article sur les origines des sciences du système terre (STEFFEN et al. 2020a), Will Steffen et Johan Rockström, avec d'autres collègues, décrivent en effet la tendance de fond à partir des années 1950 et jusqu'à aujourd'hui, qui a contribué à cesser d'étudier la Terre par une approche de géographie de terrain (*field geography*) au profit des outils de mesure globaux (*global sensing and monitoring*) et des modèles numériques. Les auteurs soulignent à ce titre les travaux marquants du consortium *International Geophysical Year*, constitué de 67 pays, et qui a donné naissance à la glaciologie, la météorologie et l'océanographie actuelles, ainsi que les publications, plus connues du grand public de James Lovelock ou Dennis et Melinda Meadows, à qui l'on doit respectivement la fameuse "Hypothèse Gaïa" (LOVELOCK et al. 1974) et le célèbre rapport *The Limits to Growth* (MEADOWS et al. 1972).

2.1.1.2 La petite histoire de l'eau dans les limites planétaires

2.1.1.2.1 La consommation globale d'eau douce comme variable de contrôle dans (ROCKSTRÖM et al. 2009a)

En 2009, l'évaluation quantitative menée par Rockström et ses collègues sur le stock d'eau de surface appropriable par les humains (en $km.yr^{-1}$) ne fait pas l'objet d'une étude spécifique mais s'appuie sur les résultats d'une littérature déjà existante. C'est là sans doute ce qui distingue le plus cette première version du cadre des limites planétaires des versions suivantes. Les deux publications centrales sur lesquelles repose le montant global annuel d'eau prélevable pour les activités humaines sont (DE FRAITURE et al. 2001) et (VOROSMARTY et al. 2000).

De Fraiture et ses collègues utilisent pour leur estimation du ruissellement annuel global le modèle PODIUM. Sur la base des données de ruissellement collectées par la FAO (qui deviendra plus tard la base AQUASTAT, nous y reviendrons plus loin) pour l'année centrale 1995, les auteurs utilisent un facteur d'utilisation potentielle pour déterminer la part du ruissellement accessible (qui est déterminé, pays par pays selon un proxy de variabilité saisonnière interannuelle des précipitations et une capacité de stockage en réservoirs). Ainsi, si le ruissellement annuel moyen est estimé à $42500 km^3$, le facteur d'utilisation potentielle n'étant en moyenne que de 36%, l'eau bleue consommable pour les activités humaines n'est estimée qu'à $15000 km^3$. Les auteurs s'appuient sur un seuil de rareté de l'eau (*water scarcity*) à 60%, c'est-à-dire lorsque les écosystèmes ne disposent plus que d'environ un tiers de l'eau disponible, s'opposant en cela à la vision orientée-humain des indices de rareté existant (par exemple le *water scarcity index* de Falkenmark selon lequel la situation est critique en-dessous de $1 km^3$ par tête et par an). Dans ces conditions, le volume annuel consommable par les activités humaines est estimé à $6000 km^3$.

Vörösmarty et ses collègues arrivent à des résultats similaires avec leur modèle WBM (*Water Balance Model*), ce qui s'explique par le fait que le WBM est calibré sur les stations de mesure hydrométrique du GRDC (Global Runoff Data Center) qui sont opérées nationalement, et donc constituent en partie les données d'entrée à partir desquelles le reporting à la FAO est effectué et sur lequel est basé le modèle PODIUM. La différence principale entre le WBM et PODIUM tient à ce que le WBM simule dynamiquement le ruissellement selon un modèle d'interaction sol-climat influencé par les réservoirs et l'occupation du sol à une résolution de 30 arcmin², alors que PODIUM part d'une valeur fixe de ruissellement (donnée collectée par la

2. Nous aurons l'occasion de présenter le modèle WBM plus en détails dans la section 1.2.2.2 de ce

FAO) qui est appropriée selon différents scénarios économiques et démographiques à une échelle nationale. À partir du WBM, le volume annuel consommable par les activités humaines est estimé à 4000 km³. Les valeurs obtenues par ces deux modèles définissent la zone d'incertitude correspondant à la limite sur l'eau de surface dans la première version du cadre.

Cette proposition initiale pour la formulation de la limite relative à l'eau a suscité des discussions et des interrogations parmi les scientifiques de la communauté. Ces discussions ont porté sur au moins trois points : le seuil global est-il correctement calculé ? Est-ce bien pertinent de définir un seuil au niveau global quand on sait que l'eau est surtout une problématique au niveau local ? Est-ce que l'indicateur ne devrait pas aussi porter sur d'autres compartiments du cycle de l'eau que le ruissellement de surface (en particulier l'eau verte et l'eau de souterraine) ?

Concernant la valeur du seuil global déterminée en 2009 (ROCKSTRÖM et al. 2009a), correspondant au volume maximum d'eau douce prélevable par les activités humaines, si l'on aurait pu s'attendre à une critique des modèles qui ont permis de l'établir, dans leurs hypothèses, leurs paramètres, leur validation, la littérature que nous avons consultée n'a pas tellement porté sa critique sur ce point. Si l'on prend par exemple (BOGARDI et al. 2013), dont les deux co-auteurs sont Fekete et Vörösmarty, tous deux concepteurs du WBM, leur point de vue ne porte pas explicitement sur une éventuelle amélioration ou utilisation ajustée du WBM, qui pourtant n'a pas été pensé spécifiquement pour la quantification d'une telle limite puisqu'il préexiste à l'apparition du cadre conceptuel des limites planétaires. Dans leur article, les auteurs font état d'un jeu d'équilibre, pour la détermination du seuil planétaire, entre plusieurs variables qui dépassent la seule recherche de la stabilité climatique. Certes, un tel seuil sur la disponibilité de l'eau de surface limite les effets dramatiques sur les services écosystémiques et le dérèglement du système des précipitations, mais son niveau devrait tenir compte des besoins des communautés locales en termes de ressources. Une question qui se pose en effet pour la limite relative à l'eau est de savoir si le respect de la limite au sens planétaire (c'est-à-dire dans un but de préservation de l'équilibre climatique de l'Holocène) est compatible avec la possibilité d'assurer les besoins élémentaires de tous, de même d'ailleurs que pour les deux limites portant sur le cycle de nutriments Azote et Phosphore. Cette question posée par (BOGARDI et al. 2013) n'a, à notre connaissance, pas trouvé de réponse avant la publication de (GERTEN et al. 2020), qui constatent qu'il est possible de nourrir une humanité de 10,2 milliards d'individus (leur question dépassait *stricto sensu* celle de

chapitre.

l'accès à l'eau, tout en réduisant l'eau à son seul but alimentaire) et de respecter quatre des limites planétaires dites "terrestres" (eau, sol, biosphère, azote). La question spécifique concernant la disponibilité de l'eau pour la diversité des besoins humains, et pas seulement l'alimentation, n'a été à notre connaissance que récemment traitée par (STEWART-KOSTER et al. 2024), dont Johan Rockström est coauteur. Cette question est donc plutôt restée en suspens entre les différentes actualisations du cadre, et avec elle d'ailleurs celle de la valeur du seuil global.

2.1.1.2.2 Un seuil global et des seuils locaux dans l'actualisation de (STEFFEN et al. 2015b)

La question qui a le plus suscité de discussions suite à la première proposition d'une variable de contrôle sur l'eau en 2009 est celle de l'hétérogénéité intrinsèque de la limite eau, qui admet une infinité de valeurs possibles selon l'échelle d'analyse (les hydrosystèmes locaux, de même que la stabilité hydroclimatique globale, ont chacun leur propre seuil de rupture) et selon la temporalité (ou plutôt saisonnalité, selon la cyclicité du régime hydrologique annuel). Pour ne citer qu'une publication dans laquelle cette question est abordée, parce qu'elle sera à l'origine de la nouvelle formulation de la limite relative à l'eau dans la version 2015 du cadre des limites planétaires (STEFFEN et al. 2015b), nous pouvons évoquer le travail de (PASTOR et al. 2014). Ce travail est préfiguré dans (GERTEN et al. 2013), mais dans une version plus succincte et qui légitime ce travail comme candidat pour la révision de l'indicateur relatif à l'eau dans le cadre des limites planétaires (cette dernière publication étant en effet cosignée par Johan Rockström).

Dans (PASTOR et al. 2014), la proposition est d'intégrer un cadre conceptuel déjà existant par ailleurs, appelé *environmental flow requirements* (EFR), et assez différent du cadre conceptuel des "couleurs de l'eau" par sa filiation, puisqu'il s'inscrit dans le champ de la gestion des milieux aquatiques, et non pas dans celui d'une comptabilité environnementale internationale et uniformisée. Les EFR sont des niveaux d'écoulement minimum dans les cours d'eau qui permettent le bon fonctionnement des écosystèmes ripariens. Les auteurs comparent dans leur article cinq indicateurs EFR (ou méthodes) "génériques" avec des valeurs réelles d'EFR calculées dans 11 contextes locaux différents (cas issus de la littérature). Parmi les indicateurs EFR testés, l'un est proposé directement par les auteurs, la méthode dite du *variable monthly flow*, qui consiste à définir *relativement* un régime de basses eaux, eaux intermédiaires

et hautes eaux. *Relativement*, c'est-à-dire en classant tout mois de l'année dans l'un de ces régimes en comparant l'écart relatif entre le débit moyen mensuel et le débit moyen annuel. Après ce classement qui tient compte de la saisonnalité, un certain pourcentage de "débit réservé" pour les écosystèmes est attribué à chaque classe de régime. Cet indicateur montre une bonne corrélation linéaire avec les valeurs réelles d'EFR calculées dans 11 contextes locaux. Les auteurs ont souhaité quantifier l'écart moyen entre ces différents indicateurs, au-delà de leur niveau de performance par rapport à des cas réels. Ils ont pour cela simulé des écoulements dits naturels et des écoulements dits actuels avec le modèle LPJmL à une résolution de 0.5 min, puis ils ont appliqué chacun des 5 indicateurs EFR sur chaque carreau. La comparaison des valeurs des débits "réservés" à l'embouchure des 14 plus gros fleuves est censée donner un ordre de grandeur de l'incertitude qu'il peut y avoir quant au choix de l'indicateur EFR.

Remarquons que la comparaison des débits réservés aux débits naturels simulés (c'est-à-dire sans interventions anthropiques) est effectué à titre de comparaison, ou informatif. Dans la version révisée de (STEFFEN et al. 2015b), il n'est pas indiqué clairement quel type de débits, naturel ou actuel, est simulé. L'enjeu n'est pas là, il s'agit d'abord d'asseoir la définition d'une limite relative à l'eau au niveau local. La question de la référence environnementale, et donc du débit "naturel", sera davantage mise en avant dans les révisions suivantes (ROCKSTROM et al. 2023; RICHARDSON et al. 2023). Toutefois, le *Supplementary Materials* de (STEFFEN et al. 2015b) indique que le calcul du dépassement de la limite à l'échelle locale³, par la méthode du *variable monthly flow*, a pris en compte distinctement ces deux types de débits : le débit naturel (appelé "*pristine*" dans le SM, c'est-à-dire que la simulation de l'écoulement ne tient pas compte du stockage en réservoirs, de l'irrigation ni de l'occupation anthropique du sol) a été utilisé pour déterminer la valeur des EFR sur chaque entité hydrographique étudiée (correspondant aux bassins versants du modèle *STN-30p drainage network*), tandis que le débit actuel a été simulé pour calculer la métrique du débit moyen mensuel. Pour les simulations, le modèle LPJmL a été utilisé avec un jeu de données climatiques (températures, précipitations) forcées (car pris sur la période 1981-2000) et une résolution de 0.5°. Le modèle LPJmL que les auteurs utilisent ne prenant pas en compte les prélèvements des secteurs industriels et domestiques, ces prélèvements sont ajoutés à chaque carreau de 0.5° selon les données de (FLÖRKE et al. 2013), que nous aurons l'occasion de présenter plus loin. Il s'agit donc bien d'une approche *bottom-up*

3. Notons que c'est là aussi une différence qui influencera les versions ultérieures du cadre : la limite relative à l'eau fait l'objet d'une étude spécifique et propre à la révision du cadre, là où dans la version initiale de (ROCKSTRÖM et al. 2009a), la limite avait été fixée par revue de la littérature, à l'instar des travaux du GIEC sur le climat.

au sens de l'échelle élémentaire de l'étude (carreau de résolution 0.5°) mais qu'en est-il au niveau des données utilisées? Sont-elles spatialement explicites et localement validées à cette échelle élémentaire d'étude?

Le modèle LPJmL utilisé par (STEFFEN et al. 2015b) correspond à sa version 3.2 proposée par (ROST et al. 2008). Il s'agit d'un modèle de végétation globale dynamique initialement développé pour simuler le cycle global du carbone terrestre et la réponse des modèles de carbone et de végétation au changement climatique. Le modèle permet ainsi de simuler différents scénarios d'anthropisation des surfaces cultivables, entre une fraction végétative dite naturelle (modélisée par 9 essences) et une fraction végétative dite anthropisée (modélisée par 12 essences). La demande en eau liées aux cultures est également simulée selon différents scénarios de répartition entre irrigation (*irrigated croplands*) et pluies naturelles interceptées (*rainfed croplands*). Les auteurs, pour la version 3.2 du modèle, ajoutent une distinction synthétique entre la somme des flux simulés correspondant à l'eau verte et ceux qui correspondent à l'eau bleue. Si les auteurs n'effectuent pas d'analyse d'incertitudes, ils effectuent toutefois une analyse de sensibilité sur 4 jeux de données météorologiques différents. La validation en tant que telle du modèle est supposée déjà effectuée par les précédentes publications liées aux versions antérieures du modèle. En particulier, le comportement simulé de chaque essence végétale a été validé à l'échelle micro (i.e. un individu) et à l'échelle macro (i.e. toute une population de l'essence), et ceci par comparaison entre les rendements simulés et les rendements observés sur quelques cas d'étude, de même que pour le cycle du carbone associé.

La validation du comportement hydrologique du modèle pour l'eau verte, c'est-à-dire la rétention d'eau par la végétation du sol, a été effectuée indépendamment du comportement hydrographique en termes de ruissellement et d'écoulements (eau bleue). Pour ce qui est de la validation du ruissellement donc, elle est effectuée dans (GORDON et al. 2004). Le but des auteurs est de comparer la performance en termes de simulation de ruissellement entre 5 modèles de simulation écologique/végétatif (dont le LPJmL, dans sa version plus ancienne où il s'appelait seulement LPJ). Cette comparaison est menée par rapport à des données d'observation issues du *Hydro Climatic Data Network* et opérées par le *US Geological Survey*. Les stations de mesure retenues pour l'évaluation correspondent aux critères suivants : au moins 50 ans de débits enregistrés, au moins 5200 km^2 de surface de bassin versant associé, jusqu'à 10% de la surface du bassin versant considérée comme non contributive. Treize bassins versants situés en Amérique du Nord sont ainsi retenus pour l'évaluation, allant de $5\,944$ à $28\,228 \text{ km}^2$.

Les débits journaliers des stations sont mensualisés puis convertis en unité de ruissellement en les rapportant à la surface du bassin versant. Pour la superposition spatiale entre l'image raster des données d'observation et l'image raster des données simulées, les auteurs ont choisi de pondérer chaque cellule du raster simulé par le pourcentage de son emprise qui intersecte le pixel en superposition dans le raster observé. Autrement dit, la correction de l'écart de projection est effectué par un proxy de surface, partant du principe que le ruissellement est proportionnel à la surface. Le LPJ présente ainsi une erreur absolue par rapport aux données observées (*mean absolute error*) au mieux de 1.24 et au pire à 37.04 mm, ce qui représente une erreur d'environ 10% ramenée au ruissellement annuel moyen observé⁴.

Une étude récente de (LIU et al. 2021) généralise à la fois la publication de (PASTOR et al. 2014) sur la comparaison des différentes méthodes d'EFR et celle de (GORDON et al. 2004) sur la comparaison de différents modèles hydrologiques globaux (GHM) en ajoutant une comparaison entre trois jeux de données météorologiques pour le forçage des GHM. Pour chaque carreau de raster ($0.5^\circ \times 0.5^\circ$), les auteurs évaluent les 18 résultats de ruissellement possibles (6 GHM et 3 jeux de données météorologiques). La médiane de ces résultats constitue alors le ruissellement naturel de référence (les simulations sont effectuées en régime non anthropisé). Les auteurs calculent ensuite le ruissellement disponible pour les activités humaines sur chaque carreau en retranchant le ruissellement écologique de réserve selon chacune des 7 méthodes EFR retenues⁵. Le ruissellement disponible est enfin comparé aux prélèvements effectués sur le carreau correspondant pour toutes les activités humaines, obtenus à partir des données de (HUANG et al. 2018) que nous décrivons plus tard.

Ainsi, dans leur actualisation de 2015, Steffen et ses collègues (STEFFEN et al. 2015b) proposent d'établir des variables de contrôle à deux niveaux :

- au niveau global, en fixant un seuil maximal de consommation totale d'eau bleue [$km.yr^{-1}$]
- au niveau de chaque grand bassin versant, en fixant des seuils de réserve écologique pour les prélèvements d'eau bleue pour l'agriculture, l'industrie et les ménages par rapport à trois classes de débit mensuel moyen (basses eaux, régime normal, hautes eaux).

4. Les valeurs brutes d'erreurs sont données en millimètres d'eau ruisselée. Celle-ci s'obtient en convertissant le débit, en m^3 , observé à une station ou modélisé, par la surface du bassin versant associé, en m^2 .

5. the flow duration curve (FDC) method (Q90 and Q50), Tennant method (1976), Tessmann method (1980), variable monthly flow (VMF) method (Pastor *et al* 2014), Q90/Q50 method (Pastor *et al* 2014), and Smakhtin method (2004)

2.1.1.2.3 Intégrer d'autres compartiments du cycle de l'eau que le ruissellement de surface

La question des compartiments du cycle de l'eau est en réalité déjà évoquée dans (ROCKSTRÖM et al. 2009a) mais la définition de l'indicateur (la consommation globale d'eau douce par l'homme) ne porte que sur une comparaison entre les consommations annuelles globales d'eau bleue et le ruissellement annuel total accessible (compartiment ruisselé du cycle de l'eau). Le compartiment eau verte (incorporé dans la végétation et évapotranspiré) est indirectement présent dans la version (STEFFEN et al. 2015b) que nous venons de détailler. En effet, si l'indicateur porte explicitement sur l'eau bleue, par une comparaison carreau-par-carreau entre le débit de réserve écologique en régime non anthropisé et le débit total en régime anthropisé, le modèle sous-jacent pour la simulation des débits modélise le bilan hydrologique pour les fractions verte et bleue de l'eau.

La définition explicite d'un indicateur portant sur l'eau verte intervient avec la publication de (GLEESON et al. 2020b) qui, d'ailleurs, élargissent le spectre des indicateurs planétaires sur l'eau à l'ensemble des compartiments du cycle de l'eau (voir Figure 2.2). Tim Gleeson et ses collègues proposent en effet d'associer à chacun des cinq compartiments du cycle de l'eau qu'ils distinguent une limite propre et une échelle d'évaluation adaptée :

- l'eau de surface : par la variable du ruissellement, elle soutient le fonctionnement des écosystèmes ripariens, en particulier les conditions d'habitat. La variable de contrôle pourrait être formulée comme le pourcentage de la superficie d'un bassin versant, ou le pourcentage de son linéaire hydrographique qui respecte le débit écologique de réserve. Cette proposition est très proche de l'indicateur retenu par (STEFFEN et al. 2015b).
- l'eau souterraine : par la variable du débit de base (ou de soutien à l'étiage), elle assure également le fonctionnement des écosystèmes ripariens mais aussi, pour les aquifères côtiers, la régulation du niveau des océans. L'échelle d'évaluation recommandée est celle des aquifères locaux. Pour souligner la fonction de soutien à l'étiage, la variable de contrôle proposée se base sur le débit d'étiage des cours d'eau du bassin versant associé à l'aquifère, qui ne doit pas être inférieur à la baisse anticipée à cause du forçage climatique.
- la glace : par la variable de la couche de glace globale, elle régule le niveau des océans. L'échelle d'évaluation est ici globale et passe par la mesure de l'élévation du niveau des mers. La variable de contrôle pourrait ainsi consister à calculer l'épaisseur de la couche de glace totale et à en déterminer le rythme de fonte tel qu'il ne soit pas plus rapide que les

effets anticipés du forçage climatique.

- la vapeur : par la variable l'évapotranspiration et du forçage radiatif (la vapeur d'eau étant un gaz à effet de serre), elle contribue respectivement au maintien des propriétés du sol (notamment humidité) et à la régulation du climat. L'échelle d'évaluation est là aussi globale et la variable de contrôle peut s'exprimer en pourcentage des terres émergées qui enregistrent un changement d'évapotranspiration plus rapide que le changement anticipé à cause du forçage climatique.
- l'humidité du sol : par la variable de la séquestration du carbone, elle contribue également à la régulation du climat. L'échelle d'évaluation préconisée est le biome (ou du moins une étendue hydroclimatiquement cohérente) et la variable de contrôle peut être définie comme le pourcentage des terres émergées qui enregistrent un changement dans leur régime de précipitation plus rapide que le changement anticipé à cause du forçage climatique.

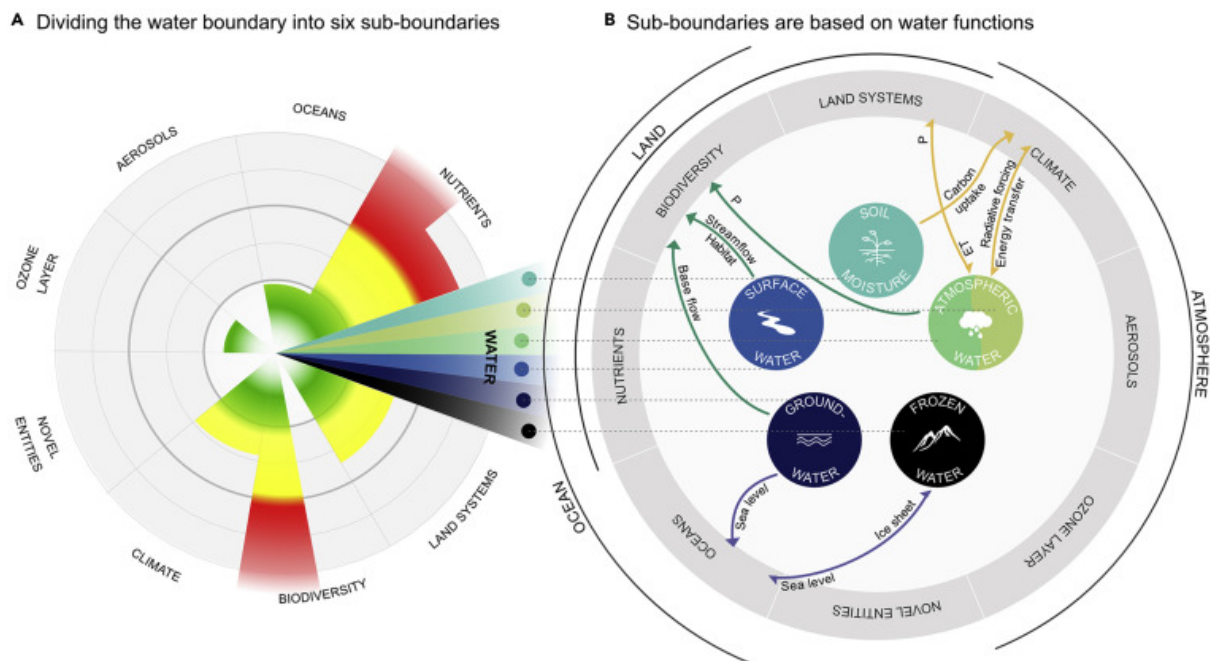


Figure 2.2 – Les 6 limites planétaires relatives à l'eau dans GLEESON et al. 2020b et leurs interactions avec le reste des limites planétaires

Il est beaucoup question dans la définition des variables de contrôle proposées par les auteurs de fixer le seuil par rapport à un changement anticipé causé par le forçage radiatif. En fait, leurs variables de contrôle ne partent pas du principe qu'il existe un état de référence environnementale, mais plutôt que nous nous éloignons de plus en plus de cet état (eût-il jamais existé) et qu'il vaut mieux baser l'évaluation sur un scénario futur souhaitable, c'est-à-dire sur un certain niveau de forçage radiatif acceptable/supportable et donc ne pas dégrader davantage

les compartiments du cycle de l'eau que le rythme fixé par le scénario de forçage retenu.

[...] we propose setting interim planetary sub-boundaries based on global normative standards for carbon and existing global data. Interim planetary boundaries for water could be set by quantifying the change in proposed control variables for each water component under the Representative Concentration Pathways (RCP) with related emissions and land-use scenarios consistent with the United Nations Framework Convention on Climate Change Paris Agreement.(GLEESON et al. 2020b)

Cette proposition alternative pour la fixation de la limite, par rapport à un scénario anthropisé plutôt qu'à un état de référence pré-Anthropocène, trouvera un certain écho dans les actualisations suivantes du cadre, sans pourtant apparaître très clairement. Ainsi dans (ROCKSTROM et al. 2023), la variabilité maximum autorisée des débits correspondra à 20% d'un débit de référence, calculé non pas sur un état d'écoulement des cours d'eau pré-Anthropocène mais en situation dite "*pristine*", c'est-à-dire une situation où le forçage radiatif est présent mais les sources de pressions directes sur les cours d'eau sont exclues (prélèvements pour l'irrigation et barrages). La question du rythme acceptable de changement (c'est-à-dire un changement qui ne s'opère pas plus vite que le scénario de référence) sera quant à elle reprise dans l'actualisation de (RICHARDSON et al. 2023). Toutefois, et tout récemment, cette idée de baser les conditions de référence sur un climat déjà forcé a été abandonnée (PORKKA et al. 2024). En effet, cette dernière actualisation en date de la limite planétaire relative à l'eau bleue et l'eau verte, d'ailleurs cosignée par un bon nombre des chercheurs dont nous venons de présenter les travaux (Tom Gleeson, Dieter Gerten, Martina Flörke et Johan Rockström), entend comparer à l'échelle locale (sur tout carreau d'un maillage de résolution 0.5°) l'écart entre la fréquence des anomalies de débit et d'humidité du sol en régime non-perturbé (période 1661-1860) et cette même fréquence en régime perturbé (période 1861-2005). Les résultats montrent une hausse tendancielle de la fréquence d'anomalie des débits et d'humidité du sol, qui excède depuis la fin du XXème siècle les plages de variabilité les plus extrêmes enregistrées en régime non-perturbé. Ce que les auteurs de cette publication entendent par "régime non-perturbé" correspond à un scénario sans forçage radiatif et avec une occupation humaine de la planète figée en 1860 :

" [...] while the pre-industrial simulations can be considered to describe an artificially stable state, they adequately correspond to long-term pre-industrial conditions

before major anthropogenic modifications affecting the distribution of freshwater."

(PORKKA et al. 2024)

Cette dernière publication confirme par ailleurs la tendance à l'utilisation de modèles hydrologiques globaux dans la quantification de la limite planétaire relative à l'eau de surface. L'approche retenue dans (PORKKA et al. 2024) est celle des ensembles de modèles (*model ensembles*) qui consiste à donner en entrée les mêmes jeux de données à des GHM différents, et à prendre le résultat moyen (ou médian) entre les sorties des différents modèles. Ici, les données d'entrée sont issues du projet ISIMIP 2b, qui se destine à regrouper, dans une nomenclature et une méthodologie cohérente, les travaux de la communauté scientifique sur la simulation des différents compartiments environnementaux dans un climat forcé à +1.5°C (FRIELER et al. 2017).

Ce que les versions du cadre postérieures à 2015 ont toutes repris à la version de (STEFFEN et al. 2015b), c'est la distinction entre les échelles spatiales et temporelles d'une part, et d'autre part la distinction entre compartiments du cycle de l'eau. Sans retenir les cinq compartiments pour autant, l'eau souterraine, l'eau de surface bleue et l'eau de surface verte sont distinguées dans (ROCKSTROM et al. 2023) et (RICHARDSON et al. 2023). Pour ce qui est de l'eau verte, mentionnons le travail intermédiaire de (WANG-ERLANDSSON et al. 2022) dans lequel les auteurs ont procédé à une évaluation planétaire et spatialement explicite des flux d'eau verte, contribuant ainsi à en déterminer un seuil de référence environnementale. Cette approche est conservée dans la dernière actualisation de (PORKKA et al. 2024).

Sur la question d'une approche alternative pour la fixation de la limite relative à l'eau, les auteurs de (GLEESON et al. 2020b) avaient insisté, la même année, en publiant un second article dans lequel ils formulaient de nouvelles propositions (GLEESON et al. 2020a). Si ce second article affirmait davantage l'importance du cycle de l'eau dans le fonctionnement du système Terre (fonctions de régulation hydroclimatique, de régulation hydroécologique, de stockage et de transport), et insistait sur les points de bascule déjà franchis (passant en revue les cas d'étude dans la littérature pour chacun des compartiments du cycle de l'eau), il suggère surtout une méthodologie nouvelle qui fait le pari de l'hétérogénéité spatiale plutôt que de l'agrégation spatiale (voir Figure 2.3). C'est là une perspective de soutenabilité forte : si l'indicateur vient, par construction, à compenser les dépassements de certains contextes locaux par le respect des limites dans certains autres contextes, alors on considère que le capital naturel est spatialement substituable.

To provide more methodology to water cycle modifications and water in the planetary boundary framework, we propose four new methods of spatial analysis that are based on the spatial distribution of impacts rather than summing of water fluxes. The current methodology for the freshwater use planetary boundary (or any other methodology that involves summing water fluxes) masks both the spatial and temporal heterogeneity of the water cycle and implies that resilience loss caused by water impacts in one place can be remediated by water abundance in another place. [...]

1. *The “unweighted approach” calculates the percentage of global land that has crossed a certain threshold;*
2. *The “weighted approach” calculates the percentage of the global land area that has crossed a certain threshold weighted by the importance of that hydrologic unit to the Earth system function;*
3. *The “keystone region approach” identifies regions where certain water stores are disproportionately important to specific Earth system components;*
4. *“Tipping elements” [...] is the corresponding critical point—in forcing and a feature of the system—at which the future state of the system is qualitatively altered. (GLEESON et al. 2020a)*

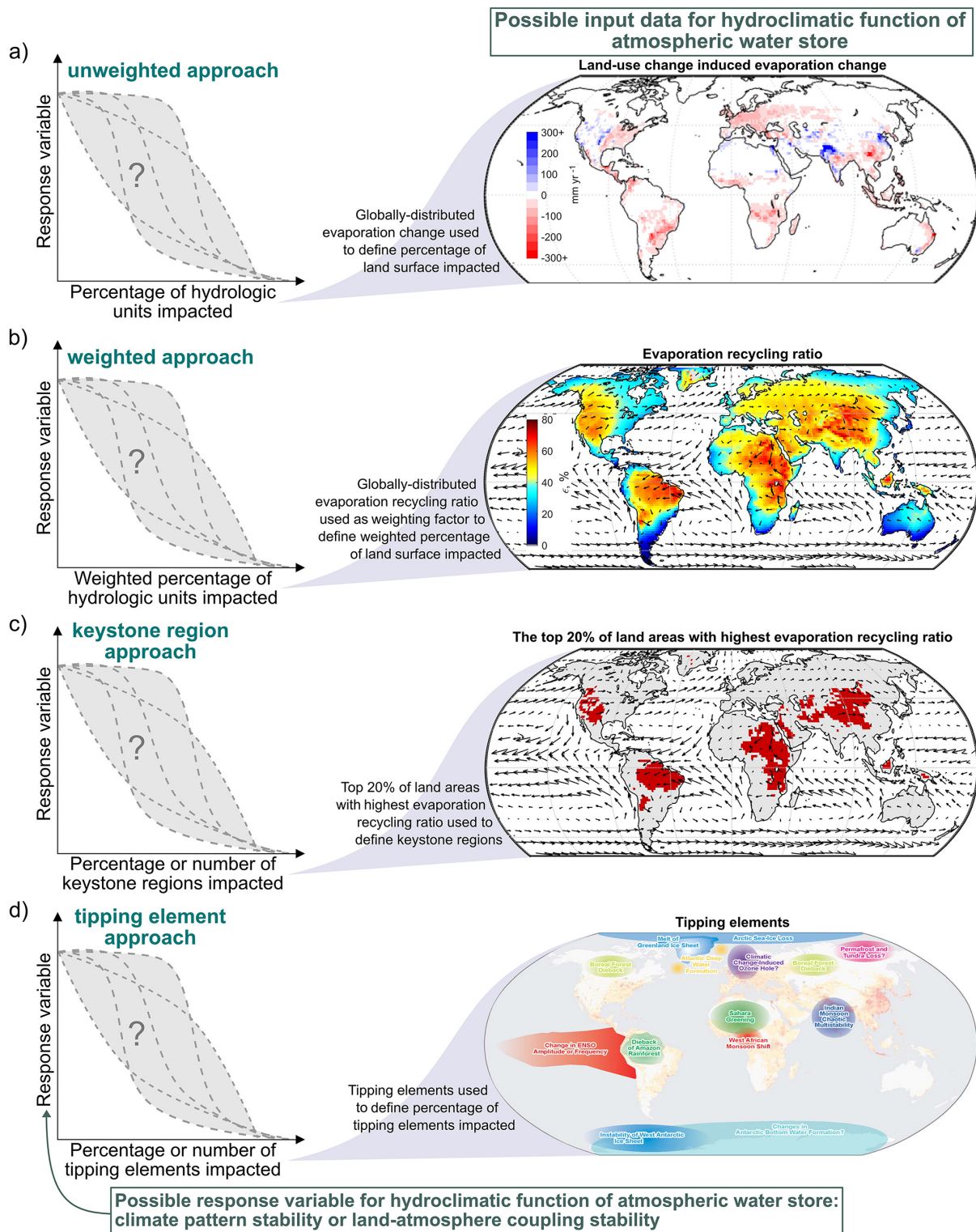


Figure 2.3 – Construire une variable de contrôle pour l'eau dans le cadre des limites planétaires qui tienne compte de l'hétérogénéité spatiale des phénomènes (GLEESON et al. 2020a)

Cette réflexion sur ce qui fait, au fond, la nature d'une limite, entre un effet localisé ou agrégé, un effet discret dans le temps, saisonnier ou une tendance de long terme, se retrouve quelques années plus tôt dans un article de (DEARING et al. 2014) qui, s'il n'a pas

particulièrement influencé la limite relative à l'eau dans les évolutions successives du cadre, a influencé un bon nombre de chercheurs qui ont tenté d'appliquer cette limite à des échelles locales, en particulier dans des contextes d'écologie aquatique (SZABÓ et al. 2020 ; BAI et al. 2020 ; RAMÍREZ et al. 2021 ; GREEN et al. 2017).

John Dearing et ses collègues ont ainsi proposé de distinguer entre quatre types de limites planétaires selon l'allure que suit sur le temps long la variable de contrôle (voir Figure 2.4) :

1. Type 1 - Tendance linéaire : l'évolution mesurée ou observée d'un phénomène suit un profil linéaire dans le temps long. Une limite peut être définie soit par rapport à un seuil de rupture connu dans cette tendance, soit lorsque l'écart au niveau d'arrière-plan naturel (*background level*) devient trop important ;
2. Type 2 - Tendance non linéaire : l'évolution du phénomène suit un profil très variable (par exemple de nombreuses variations saisonnières sur la chronique) et le profil tendanciel (courbe de lissage de tendance) finit par devenir non linéaire. Cela peut s'évaluer soit en calculant de façon glissante le coefficient de variation, qui tend à s'éloigner d'une valeur stable ; soit en déterminant une plage ou enveloppe de variabilité en régime normal qui, si elle devient dépassée, signifie qu'il y a dépassement de la limite. Remarquons que ce concept d'*enveloppe de variabilité de long terme* est toutefois à la base de la formulation des limites relatives à l'eau dans les versions (ROCKSTROM et al. 2023) et (RICHARDSON et al. 2023) du cadre des limites planétaires, sous l'appellation de *safe corridor* dans (ROCKSTROM et al. 2023) par exemple ;
3. Type 3 - Effets de seuils (*thresholds*) : le phénomène est plutôt stable sur le temps long sans variabilité saisonnière et enregistre soudainement une rupture dans la chronique ;
4. Type 4 - Signaux faibles (*early warning signals*) : il s'agit d'un mélange entre les types 2 et 3 : le phénomène est naturellement variable sur le temps long mais contenu dans une enveloppe jusqu'à un point de rupture brutal dans la chronique où le signal envoyé par le phénomène devient particulièrement instable et monotone en tendance.

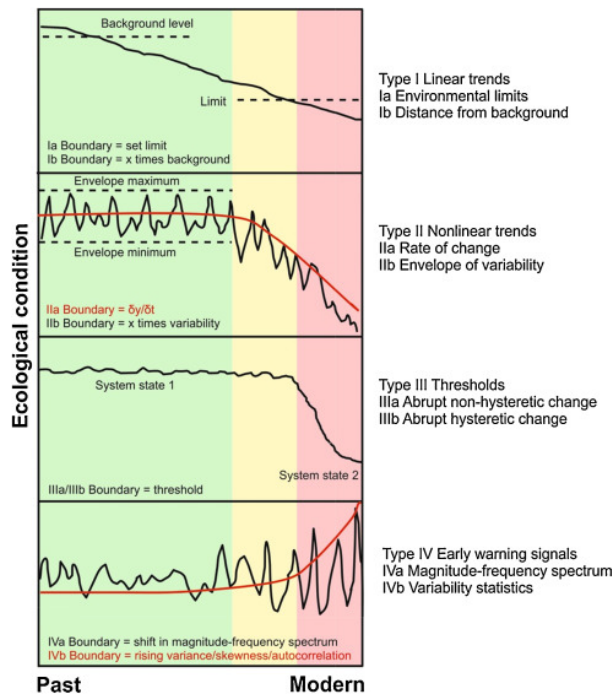


Figure 2.4 – Les quatre profils d'évolution d'une variable de contrôle qui définissent quatre types de limites planétaires (DEARING et al. 2014)

Nous avons retracé brièvement l'histoire, petite et grande, de la limite relative à l'eau dans le cadre des limites planétaires et avons essayé d'identifier les facteurs ayant contribué à sa formulation finale (pour l'heure), à la fois d'un point de vue conceptuel et quantitatif. Nous allons désormais présenter dans le détail la version (ROCKSTROM et al. 2023) de la limite relative à l'eau, qui nous servira de cas d'étude dans la suite des chapitres.

2.1.2 La frontière du système Terre relative à l'eau de surface

Comme nous l'avons présenté au chapitre 1, le cadre des frontières du système Terre proposé par (ROCKSTROM et al. 2023) a la particularité par rapport au reste des versions du cadre de comporter, pour chaque variable environnementale, un niveau *sûr* en termes de résilience écologique et un niveau *juste* en termes de préjudice sur le vivant. Notre étude portant sur l'évaluation environnementale d'un territoire, nous laisserons la frontière *juste* de côté. À son tour, la frontière *sûre* se décompose en deux frontières. D'une part une frontière portant sur les eaux de surface, d'autre part une frontière portant sur les eaux souterraines, dans la continuité des évolutions conceptuelles du cadre que nous évoquions avant. La frontière sur les eaux souterraines reconnaît le service rendu par les aquifères pour le soutien à l'étiage (cf. notre discussion sur (GLEESON et al. 2020b)) mais reconnaît également son rôle dans la préservation des zones humides et des habitats associés. Un niveau global et régional est défini : au niveau

global, le volume de rechargement annuel moyen de tous les aquifères doit être supérieur au volume annuel moyen de prélèvements; et au niveau régional, la métrique est la même mais désagrégée au niveau de chaque aquifère. Remarquons que l'étude de (WANG-ERLANDSSON et al. 2022) est envisagée pour définir une limite sur les eaux vertes, mais (ROCKSTROM et al. 2023) ne conduisent pas d'étude en propre sur ce compartiment du cycle de l'eau. Nous allons présenter plus longuement la frontière sur l'eau de surface, qui nous intéresse comme cas d'étude pour l'évaluation de l'intégrité des cours d'eau.

L'indicateur retenu par (ROCKSTROM et al. 2023) pour la frontière relative à l'eau de surface consiste à évaluer si le débit actuel d'un cours d'eau est altéré de plus de 20% à la hausse ou à la baisse par rapport à sa valeur en régime pré-industriel. La valeur de 20%, qui est le seuil (ou la frontière) pour l'écart relatif entre le débit perturbé et préservé, est obtenue par la méthode des *environmental flow requirements*. Pour comprendre comment cet indicateur a été construit, nous allons distinguer, d'une part, les bases conceptuelles qui ont présidé à sa construction, et notamment le choix de cette valeur *a priori* arbitraire de 20%, puis nous verrons ensuite comment sont calculés et à quoi correspondent ces "débits perturbés" et "débits pré-industriels".

2.1.2.1 Les *environmental flow requirements* comme cadre conceptuel

Johan Rockström et ses collègues reprennent le cadre conceptuel des *environmental flow requirements* (EFR) issu des améliorations précédentes apportées à la formulation de la limite eau (PASTOR et al. 2014). Ce cadre n'apparaît pas avec celui des limites planétaires, il est issu de plusieurs décennies de recherche, surtout portée par le champ de l'écologie et par les gestionnaires des milieux aquatiques, sur les interactions entre caractéristiques hydrologiques et écosystèmes. Ce cadre a également bénéficié d'un portage institutionnel important puisque, dès 2007, la nécessité d'utiliser les EFR dans la gestion des cours d'eau figure dans la Déclaration de Brisbane (SYMPOSIUM 2007). Cette déclaration a un statut un peu particulier parce qu'elle émane d'une organisation qui ne fait pas partie des instances officielles supraétatiques de gouvernance internationale, comme les Nations Unies, la Banque Mondiale ou le Fonds Monétaire International, mais d'une fondation australienne nommée *International River Foundation* qui voit le jour en 2003 à l'initiative d'un partenariat public-privé, entre de la ville de Brisbane et la région du Queensland d'une part, et les compagnies australiennes d'extraction minière Thiess et BMA (*BHP Mitsubishi Alliance*)⁶ d'autre part. Si le périmètre de la fondation, et donc de sa déclaration, se veut international, la représentativité des pays

6. <https://riverfoundation.org.au/our-story/>

(par leurs délégations de scientifiques, ONG, représentants des États) n'est pas aussi exhaustive que dans les instances supranationales de gouvernance. Par exemple, parmi les signataires de la Déclaration de 2007, la France ne figure qu'à travers son représentant à l'UNESCO. Cette fondation est l'émanation d'une première série de conférences internationales sur la gestion des rivières (*International River Symposium*) tenues à Brisbane à partir de l'année 1998⁷. Ces conférences sont désormais opérées par la *International River Foundation*. À l'issue de la conférence de 2017, la Déclaration de Brisbane de 2007 est mise à jour et propose à cette occasion une redéfinition des EFR (ARTHINGTON et al. 2018). Angela Arthington et ses collègues retiennent ainsi la définition suivante en justifiant le changement apporté à la définition de 2007 :

Environmental flows describe the quantity, timing, and quality of freshwater flows and levels necessary to sustain aquatic ecosystems which, in turn, support human cultures, economies, sustainable livelihoods, and well-being.” In this definition, “Aquatic ecosystems include rivers, streams, springs, riparian, floodplain and other wetlands, lakes, coastal waterbodies, including lagoons and estuaries, and groundwater-dependent ecosystems”. By altering the original wording from “quantity, timing, and quality of water flows required to sustain freshwater and estuarine ecosystems.” to “quantity, timing, and quality of freshwater flows and levels necessary to sustain aquatic ecosystems”, the revised definition meets the call to embrace flowing (lotic), standing (lentic) and GDEs [groundwater-dependant ecosystems], as well as aquatic ecosystems that may alternate between these states (e.g., ephemeral streams and intermittent rivers). (ARTHINGTON et al. 2018)

Si jusqu'à présent les métriques que nous avons évoquées pour l'évaluation des EFR portaient sur les débits ou les volumes d'eau ruisselés (PASTOR et al. 2014; LIU et al. 2021), rien n'empêche dans cette définition de retenir des métriques davantage qualitatives, comme la température de l'eau, son niveau de pollution organique ou chimique. Toutefois, étant donné l'ambition du cadre des limites planétaires de porter sur un périmètre précisément planétaire, les modèles disponibles à cette échelle se sont développés historiquement sur des métriques quantitatives, comme nous avons pu l'évoquer plus haut.

Johan Rockström et ses collègues s'appuient sur deux références principales pour justifier la métrique EFR qu'ils retiennent : les publications de (POFF et al. 2010) et (RICHTER et al. 2012).

7. <https://riversymposium.com/about/>

N. LeRoy Poff et ses collègues ont constitué une revue de la littérature sur 165 études ayant documenté la réponse de la biodiversité à une modification (ou altération) quantitative du régime hydrologique dans un contexte local. Cette relation flux-écologie (*flow-ecology relationships*) leur permet de légitimer le choix d'une métrique quantitative portant sur l'altération (ou la modification) des débits.

Poff et ses collègues identifient dans leur sélection d'articles plusieurs types de "variables de flux altéré" (*altered flow variables*) pour la quantification de l'altération des régimes hydrologiques : les variables d'ampleur (*magnitude*), de fréquence, de durée, de cadence (*timing*) et de vitesse de changement (*rate of change*). Étant donné la surreprésentation des variables d'ampleur dans leur sélection d'articles, les auteurs ne poursuivent leur étude que sur ce type de variable. Plusieurs grandeurs hydrologiques sont alors identifiées : le débit de pointe (*peak flow*), le débit moyen (*average discharge*), le débit d'étiage (*base flow*), la variabilité de court-terme (*short-term variation*). Afin de pouvoir comparer ces grandeurs différentes et définir des seuils critiques d'altération du régime hydrologique sur la biodiversité, les auteurs synthétisent ces grandeurs dans une métrique d'altération relative, définie, pour toute grandeur, comme l'écart observé de cette grandeur par rapport à son niveau avant impact, ou niveau de référence (*pre-impact conditions*). Les écarts relatifs supérieurs à 100% ou inférieurs à -100% sont ramenés respectivement à 100% et -100%.

Concernant la réponse de la biodiversité à cette altération relative, les auteurs décident de subsumer les différentes espèces étudiées dans les publications par groupe taxonomique : les poissons, les espèces ripariennes, les macroinvertébrés. Là aussi, étant donné la diversité des grandeurs écologiques pour mesurer la réponse de la biodiversité à l'intérieur de ces groupes, les auteurs regroupent d'abord les grandeurs en trois groupes de métriques (l'abondance spécifique, la richesse spécifique et la démographie spécifique), puis ils retiennent l'indicateur synthétique du pourcentage de changement (*percent-change*) par rapport aux conditions de référence (ou, à défaut, aux conditions à l'amont, i.e. en tête de bassin).

Leur étude montre qu'il n'y a pas de relation claire, c'est-à-dire pas de seuil de rupture évident, dans la réponse des macroinvertébrés ni des espèces ripariennes à l'altération quantitative du milieu (voir Figure 2.5). Pour le groupe des poissons en revanche, ils déterminent une relation négative entre l'altération quantitative du milieu et la réponse de la biodiversité,

mais qui s'explique vraisemblablement davantage par le fait que les études sélectionnées ont principalement documenté des altérations quantitatives du milieu à la baisse (e.g. baisse du débit). Le seuil auquel se produit une rupture dans la réponse de la biodiversité des poissons se situe au-delà de -50% à la baisse ou +75% à la hausse du débit.

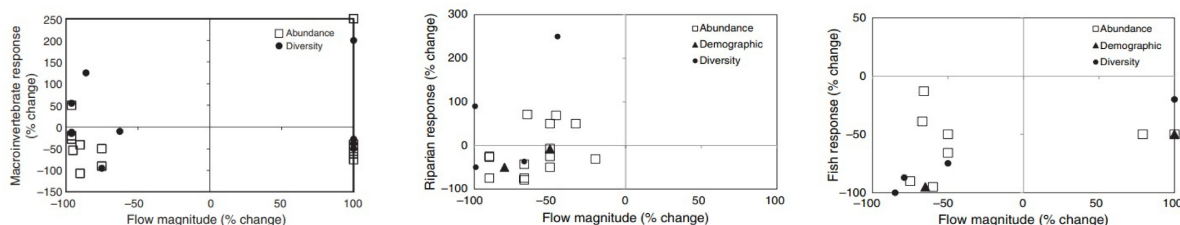


Figure 2.5 – Réponse des macroinvertébrés, des espèces ripariennes et des poissons à l'altération quantitative des cours d'eau (POFF et al. 2010)

C'est véritablement avec l'article de (RICHTER et al. 2012) que Johan Rockström et ses collègues vont arrêter une valeur d'altération quantitative du milieu. Si la revue de littérature effectuée par (POFF et al. 2010) ne permet pas de conclure définitivement sur une valeur de rupture écosystémique, elle permet à Johan Rockström et ses collègues de montrer que le choix d'une valeur unique, pour tous les groupes taxonomiques et tous les contextes locaux, est nécessairement arbitraire et provisoire. C'est bien là le postulat de Richter et son équipe, qui entendent proposer une "valeur par défaut" (*presumptive standard*) d'altération quantitative. L'approche retenue par Richter et ses collègues est de commencer par distinguer trois approches pour la détermination de l'EFR : l'approche par "débit-seuil minimum" (*minimum flow threshold*), l'approche par "standards statistiques" (*statistically-based standards*) et l'approche par "POF" (*percent-of-flow*). Richter et ses collègues recommandent pour leur part de préférer l'approche POF (voir Figure 2.6). L'approche par débit-seuil minimum présente en effet le risque que les gestionnaires d'ouvrage ne tiennent pas compte de la variabilité naturelle du cours d'eau en maintenant le débit à un niveau constant pour le besoin de leurs opérations. L'approche statistique présente le désavantage d'être souvent trop technique pour être implémentée et constituer un outil de suivi et de pilotage efficace. L'approche POF est censée répondre à ces deux critiques.

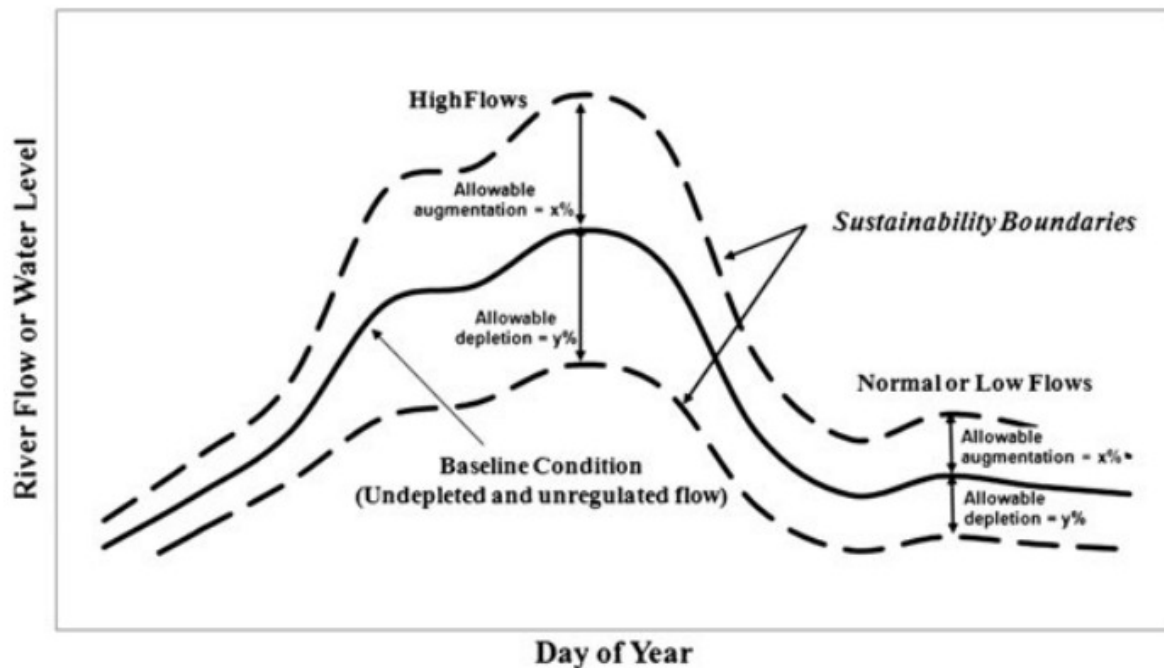


Figure 2.6 – Schéma de l'approche par *percent-of-flow* (RICHTER et al. 2012)

Afin de déterminer le pourcentage acceptable de variabilité du débit (POF), Richter et ses collègues proposent de passer en revue les valeurs en usage dans différents contextes de gouvernance des milieux. Ils se réfèrent alors davantage à une littérature grise, issues des normes et des documentations techniques produites par les gestionnaires, plutôt qu'une littérature proprement scientifique. Leurs territoires d'étude sont les suivants : *Southwest Florida Water Management District*, *Michigan's Water Withdrawal Assessment Tool*, *UK Application of the European Union Water Framework Directive*, *Maine Sustainable Water Use Rule*. Ce sont donc exclusivement des cas occidentaux et anglophones, pris aux États-Unis et au Royaume-Uni. Quoiqu'il en soit, ils identifient une plage de prélèvements autorisés allant de 8 à 35% du débit journalier moyen. Par ailleurs, ils recommandent de fixer un seuil supérieur à 20% pour une protection forte du milieu, et un seuil compris entre 11 et 20% pour une protection plus lâche. C'est cette valeur de 20% qui est retenue par (ROCKSTROM et al. 2023). Cette valeur retenue n'a donc pas véritablement de fondement écologique ou biologique. Il s'agit d'une norme consensuelle en usage parmi les gestionnaires sélectionnés dans leur étude.

2.1.2.2 Le *Water Balance Model* pour l'évaluation quantitative de l'indicateur

Nous en arrivons à la présentation de l'indicateur sur l'eau de surface proposé par (ROCKSTROM et al. 2023), que nous allons formaliser et dont nous allons décrire les données qui

ont permis de le calculer, lesquelles sont issues du modèle hydrologique global *Water Balance Model* (WBM).

2.1.2.2.1 Formalisation de l'indicateur

Formellement, l'indicateur sur les eaux bleues de surface proposé par (ROCKSTROM et al. 2023) peut s'écrire ainsi :

$$\text{Au niveau local}^8, \quad \theta_{i,t} = \begin{cases} 1 & \text{si } |MMF_{i,t,d} - MMF_{i,t,p}| \geq 1.2MMF_{i,t,p} \\ 0 & \text{sinon} \end{cases}$$

Où i désigne un pixel d'image raster (en l'occurrence, de résolution $6' \times 6'$), t désigne un mois de l'année ($1 \leq t \leq 12$) et $MMF_{t,i}$ le débit moyen mensuel (*mean monthly flow*) sur le pixel i et pour le mois t sur la période d'étude (en l'occurrence, 2000-2020), c'est-à-dire la moyenne du débit mensuel de tous les mêmes mois sur la période. La simulation effectuée à partir du WBM se fait soit en condition perturbée, i.e. *disturbed* ($MMF_{i,t,d}$) soit en condition pré-industrielles, i.e. *pristine* ($MMF_{i,t,p}$). Ainsi $\theta_{i,t}$ prend la valeur 1 si l'EFR n'est pas respecté (le débit est altéré au-delà des $\pm 20\%$) sur le carreau i pour le mois t , et 0 s'il est respecté (non-dépassement de la limite).

Une formulation alternative de l'indicateur au niveau local est proposée dans (GREEN 2023), qui est la publication technique associée à (ROCKSTROM et al. 2023) :

$$f_i(\theta_{i,t}) = \sum_{1 \leq t \leq 12} \theta_{i,t}$$

Où $f_i(\theta_{i,t})$ est la fréquence annuelle de non-respect de l'EFR. Autrement dit $0 \leq f_i(\theta_{i,t}) \leq 12$.

Enfin :

$$\text{Au niveau global,} \quad \Gamma = \begin{cases} 1 & \text{si } \sum_i f_i(\theta_{i,t}) \neq 0 \\ 0 & \text{sinon} \end{cases}$$

Où Γ vaut 0 si et seulement si l'ensemble des pixels i ont une fréquence annuelle de dépassement qui est nulle. Autrement dit, la limite planétaire est respectée au niveau global

8. Le "niveau local" correspond à un point de l'espace i défini par ses coordonnées $\langle x, y \rangle$ et appartenant à un réseau hydrographique. Ce réseau étant représenté par l'image raster issue du modèle STN-30p, le point i est en fait un pixel de l'image. Autrement dit, le "niveau local" peut correspondre à n'importe quel point d'un cours d'eau, ou encore n'importe quel pixel de l'image raster d'un réseau hydrographique.

à condition que l'ensemble des surfaces ne dépasse pas leur propre limite sur tous les mois de l'année (en moyenne sur la période d'étude).

2.1.2.2.2 Description du *Water Balance Model*

Pour déterminer les valeurs de débits en régime perturbé (MMF_d) et en régime pré-industriel (MMF_p), (ROCKSTROM et al. 2023) ont recours à une version enrichie du *Water Balance Model*, le WBM_{plus} (WISSER et al. 2010). Le WBM a été développé initialement à la fin des années 1990 par Charles Vörösmarty et ses collègues comme un modèle hydrologique global dédié à la simulation des écoulements de surface (*Water Transport Model* à l'origine, couplé dans les années 2000 à un modèle de bilan hydrologique, d'où le *Water Balance Model*). Le WBM_{plus} est enrichi d'un modèle de pressions anthropiques, qui simule les consommations liées à l'irrigation et les prélèvements liés aux barrages et réservoirs. Le modèle fonctionne pixel par pixel. Sur chaque pixel, l'espace est partitionné en une fraction irriguée (*irrigated*) et une fraction non irriguée (*rainfed*). Le modèle effectue un bilan hydrologique vertical en considérant les flux échangés entre la surface et l'atmosphère. Ce bilan est effectué pour différents compartiments.

(1) Les surfaces enneigées. Pour chaque pixel, il y a enneigement des précipitations si celles-ci adviennent lorsque la température de surface sur le pixel est inférieure à -1°C . La neige fond dès que la température devient supérieure à 1°C . Est ainsi récupéré en sortie un débit journalier issu de la fonte des neiges qui est fonction de la température et des précipitations sur le pixel.

(2) L'humidité du sol. Pour chaque pixel, l'humidité du sol [mm] est calculée à partir de l'évapotranspiration potentielle, du débit disponible pour le sol (somme des précipitations et du débit issu de la fonte des neiges), du déficit d'humidité du sol (qui va de pair avec l'évapotranspiration potentielle, comme facteurs mutuellement limitants), et un paramètre de forme selon le type de sol et de végétation en présence sur le pixel.

(3) La demande évaporative liée aux cultures irriguées. Pour chaque pixel, l'évapotranspiration spécifique liée aux cultures irriguées du pixel est calculée, en différenciant les types de cultures. Pour chaque culture, un coefficient spécifique pondère une évapotranspiration générique. L'irrigation est simulée dès que les précipitations ne suffisent plus à assurer une humidité

du sol critique compte tenu de l'évapotranspiration spécifique et d'un coefficient spécifique de stress physiologique (il y a irrigation dès que "la plante a soif"). La modélisation tient aussi compte d'une saisonnalité générique pour chaque type de cultures (seulement pendant une partie de l'année qui correspond à une certaine plage du couple température-précipitations). Cette demande irrigative ainsi calculée est dite "irrigation nette". Pour calculer la demande en eau totale liée à l'irrigation, il faut calculer l'irrigation brute, qui est égale à l'irrigation nette multipliée un coefficient de perte en ligne entre la source d'extraction et le pixel irrigué (en pratique, une telle distance n'est pas modélisée et des coefficients génériques d'efficacité par pays serviront de données d'entrée : voir plus bas).

(4) Les transferts souterrains. Ces transferts ne sont pas directement modélisés mais appréhendés comme une rétention (ou un retardement) de l'écoulement. Pour chaque pixel, lorsque l'humidité est à son équilibre, tout apport supplémentaire de précipitation est considéré comme un excès de débit qui contribue au ruissellement. Ce débit en excès est réparti entre une fraction libre (qui rejoindra l'exutoire) et une fraction infiltrée qui est représentée par un vase de rétention, lequel libère la fraction en excès retenue dans le vase à une vitesse donnée (un coefficient dont l'unité est l'inverse du temps).

(5) Le stockage dans les réservoirs (lacs de retenue). Pour les gros réservoirs, il s'agit de calculer le débit régulé en sortie d'ouvrage comme une fonction du débit entrant, en distinguant le débit entrant de long terme et de court-terme. Si le débit entrant de court-terme est plus grand que le débit de long-terme, alors le débit sortant est égal au débit entrant de court-terme *modulo* un coefficient de gestion. À l'inverse, le débit sortant est pondéré par l'écart absolu entre les débits entrants de long et court-terme. La quantité d'eau stockée dans le réservoir est calculée à partir du débit sortant et du débit entrant de court-terme. L'évaporation à la surface du réservoir n'est pas prise en compte. Pour les petits barrages, leur fonction est réduite à un stockage des précipitations en saison humide pour les besoins de l'irrigation en saison sèche. Pour chaque pixel, ce type de réservoir est modélisé. La quantité d'eau stockée est la plus petite quantité entre la demande brute en irrigation (calculée en 3) et la fraction libre du débit ruisselé (calculée en 4) pondérée par un facteur d'interception. En effet, tout le débit ruisselé sur le pixel ne peut pas être intercepté (ou capturé) par l'ouvrage. Ce facteur d'interception est égal à la surface telle que la somme du ruissellement libre sur cette surface soit égal à une unité d'irrigation. La quantité d'eau stockée est ensuite diminuée selon un coefficient d'évaporation.

Une fois ces différents compartiments du cycle hydrologique modélisés, ils sont mis en réseau par un modèle de transport horizontal (*routing*). Le modèle de *routing* utilisé correspond au cadre conceptuel FrAMES (*Framework for Aquatic Modeling in the Earth System*), pour lequel plusieurs méthodes sont possibles. Ici, les auteurs choisissent la méthode Muskingum-Cunge basée une solution numérique des équations de Saint-Venant pour la convection et les équations de Manning et Chezy pour les contraintes morphologiques. Sans trop détailler, le transport consiste à calculer la valeur d'un pixel d'une unité vers l'aval (pixel contigu aval) et à l'instant suivant (jour suivant) comme combinaison linéaire de l'état du pixel contigu amont au jour actuel et du pixel aval au jour d'avant. La connaissance de la direction d'écoulement (i.e. le réseau hydrographique) est issue du maillage hydrographique *Simulated Topological Network 30 arcmin* (STN-30p, VOROSMARTY et al. 2000). La quantité transférée de l'amont vers l'aval est paramétrisée selon des variables morphologiques/topographiques, comme la profondeur du lit de la rivière, sa largeur, sa pente. Avec ce modèle de transport de l'amont vers l'aval, et les modèles précédemment décrits étant exécutés à chaque pas de temps, on obtient le débit moyen journalier sur chaque pixel du maillage. Connaissant cette donnée, à chaque pas de temps, la demande irrigative est assurée, par ordre de priorité, par les petits réservoirs (5), par les rétentions (4) ou par l'abstraction d'une fraction du débit transporté. Cette demande irrigative est ensuite retirée à l'écoulement lors du transport vers l'aval.

Concernant les données utilisées par (WISSER et al. 2010), les données climatiques (températures de surface et précipitations) sont simulées à partir du jeu de données CRU TS 2.1 dont la résolution spatiale est de 30 arcmin, la résolution temporelle est mensuelle et la couverture temporelle va de 1901 à 2002 (l'ambition des auteurs, comme le titre de leur article l'indique, est de reconstituer les écoulements journaliers sur tout le XXème siècle). Les auteurs soulignent que les résultats d'un modèle hydrologique global varient considérablement selon le jeu de données climatiques fourni en entrée, car leur représentativité spatiale, les algorithmes de correction et d'interpolation, varient considérablement entre jeux de données. Les auteurs ont ainsi comparé la pluviométrie annuelle moyenne sur toute la surface terrestre obtenue avec 4 jeux de données de précipitation différents : 732 mm pour le GPCPmon, 797 mm pour le GPCP1dd, 778 mm pour le VASclimO et 785 mm pour le CRU TS 2.1. Un second enjeu lié aux données climatiques est le compromis entre résolution temporelle et couverture temporelle. Si le CRU TS 2.1 couvre bien la période qui intéresse les auteurs, la résolution temporelle mensuelle ne convient pas directement pour l'estimation des écoulements journaliers. Ils utilisent pour cela une méthode d'interpolation basée sur le jeu de données de résolution

journalière GPCP1dd et en récupèrent, pour chaque pixel du maillage, la distribution journalière des précipitations mensuelle, qui est ensuite appliquée au cumul mensuel du CRU TS 2.1.

Pour ce qui est des données sur la distribution spatiale des terres irriguées, les auteurs utilisent le *FAO Global Map of Irrigated Areas*, qui est statique sur l'année 2000. Pour évaluer dynamiquement l'évolution des terres irriguées au cours du XXème siècle, les auteurs utilisent la série temporelle de (FREYDANK et al. 2008), qui reconstitue l'évolution de la part des terres irriguées au cours du XXème siècle pays par pays. Les auteurs appliquent alors uniformément aux pixels du *FAO Global Map of Irrigated Areas* l'évolution des surfaces irriguées pour tous les pixels au sein d'un même pays. Le *FAO Global Map of Irrigated Areas* ne renseigne que sur les surfaces "équipées pour l'irrigation", ce qui ne signifie pas nécessairement qu'elles soient irriguées en permanence. C'est pourtant l'hypothèse que fait le modèle d'irrigation (3). La répartition spatiale des types de cultures est issue de la littérature (MONFREDA et al. 2008) et est supposée temporellement constante. L'efficacité des systèmes d'irrigation est issue de la base AQUASTAT (que nous aurons l'occasion de présenter en détails) pour l'année 2008 et est supposée constante sur la période. Quant aux propriétés pédologiques du sol pour le modèle d'humidité (2), elles sont issues de la *FAO Soil Map of the World* pour l'année 2003 et sont supposées constantes également sur la période. Quant aux données sur les gros réservoirs, elles sont issues du *Global Repository of Registered Impoundment* qui compile des données elles-mêmes issues du registre International Commission on Large Dam. Cette base regroupe 668 barrages à travers le monde dont la capacité du réservoir est supérieure à 0.5 km³ et dont la hauteur de chute est supérieure à 15m. L'attribution des barrages aux pixels du réseau STN-30p semble être déjà effectuée dans le *Global Repository of Registered Impoundment*.

2.1.2.2.3 Un modèle validé pour les grands bassins versants

Ce qui nous intéresse sûrement davantage réside dans la validation du modèle et, plus précisément, la plage de validité du modèle : sur quels bassins versants, définis par leur surface, le modèle est-il performant ? La performance est-elle évaluée sur chaque pixel ou seulement sur les pixels exutoires (en sortie de bassin versant) ? Connaître cette plage de validité nous permet de connaître les possibilités d'utilisation du modèle au niveau local. Remarquons que nous avons employé le terme "validation" du modèle, qu'il est important de distinguer du terme "calibration" du modèle. Le WBM_{plus} n'est pas un modèle calibré.

"It is important to note that in keeping with earlier approaches the model parameters in the WBM_{plus} model were assigned a priori and not further calibrated to match observed discharge." (WISSER et al. 2010).

Les auteurs du WBM_{plus} proposent de valider les sorties de leur modèle sur des données d'observations mesurées par des stations de mesure. Ces données sont regroupées dans le *Global Runoff Data Centre* (GRDC) et ont déjà fait l'objet d'un couplage (*coregistration*) avec le réseau hydrographique STN-30p dans (FEKETE et al. 2002). Les auteurs récupèrent ainsi les données dites *Composite Runoff Fields* issues de ce travail⁹, qui comprend 660 stations de mesure. Naturellement, la validation du modèle ne peut se faire qu'à partir d'une simulation d'écoulement perturbé, puisque les stations de mesure enregistrent un débit nécessairement perturbé. L'étude du biais moyen (moyenne des écarts entre données observées et données modélisées) montre qu'il y a une tendance générale à la sous-estimation des débits par le modèle, et que le modèle performe d'autant plus que la taille du bassin versant augmente.

"Averaged for all stations, the average MBE of -1.2 mm/month indicates that the model on average underestimates discharge and this negative bias can partly be explained with biases arising from errors in the precipitation input fields due to gauge undercatch. As the bias arising from uncertainties in the input data partly cancels out over large domains, the model performance generally increases with basin size." (WISSER et al. 2010).

Il y aurait donc une taille critique de bassin versant en dessous de laquelle le modèle performerait trop mal pour être utilisé. C'est davantage l'étude de (FEKETE et al. 2002) qui renseigne sur ce point. Dans leur étude, ils ont cherché à coupler les sorties du modèle WBM (dans sa version antérieure à WISSER et al. 2010) avec les données de mesure du réseau GDRC. La taille des bassins versants associée aux stations retenues du GDRC varie de 7000 km^2 à plus de 5 millions de km^2 (Gange, Mississippi, Amazone...). En considérant la distribution du biais moyen pour tous les bassins versants en fonction de leur superficie, les auteurs concluent que "*a 30min network cannot simulate the contributing area well on basins smaller than on the order of 10^4 km^2* ". On constate en effet sur la Figure 2.7 que plus la taille du bassin rétrécit, plus l'erreur

9. Nous présenterons plus dans le chapitre la base de données du GRDC ainsi que ce travail de production des *Composite Runoff Fields*

moyenne augmente, et ceci plutôt symétriquement.

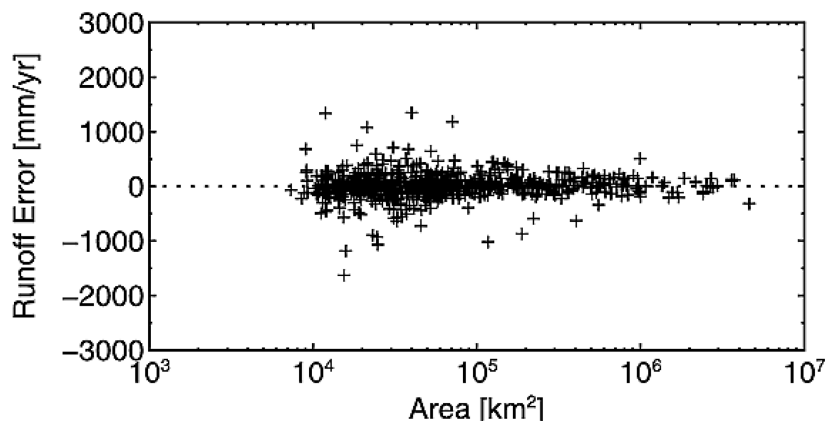


Figure 2.7 – Erreur d'estimation du ruissellement par le modèle WBM selon la surface du bassin versant modélisé (FEKETE et al. 2002)

2.2 Les applications sous-globales de la limite relative à l'eau

Bien que la partie précédente ait spécifiquement porté sur la version (ROCKSTROM et al. 2023) de la limite relative à l'eau de surface, en retraçant son histoire et en décrivant sa construction, nous n'allons pas réduire notre revue de la littérature sur les applications sous-globales à cette version car elle est trop récente pour constituer un corpus suffisamment fourni.

Dans cette partie, nous passerons en revue les publications du type (III - revues de la littérature) et (IV - applications sous-globales) afin d'identifier les types de données utilisées au sein du champ de l'AESA pour conduire une évaluation de la limite relative à l'eau de surface.

2.2.1 Méthodologie pour la revue de la littérature

2.2.1.1 Une échelle sous-globale, locale ou territoriale ?

Nous avons déjà évoqué dans le chapitre 1 notre préférence pour le qualificatif *sous-global*, plutôt que de *local* ou *territorial*. Cette préférence prend tout son sens dans le contexte d'une revue de la littérature où la recherche de publications par le mot clé "territorial" n'est guère

pertinente. Nous avons suggéré au chapitre 1 qu'un territoire pouvait être entendu comme un système multifonctionnel spatialement défini par une zone ou un ensemble de zones.

Le *territoire* est en effet un objet sémantiquement et culturellement marqué dans le contexte de la politique administrative française qui trouve mal sa correspondance dans la langue anglaise académique, où les termes de "*local scale*", "*local context*", "*region*", "*county*" et, plus rarement, "*territory*" peuvent s'y substituer sans le recouvrir. Nous avons donc retenu les publications qui mentionnent explicitement une zone (et non pas un site) comme cas d'étude, par exemple la consommation finale de l'Union européenne (SALA et al. 2020). Pour les publications moins explicites, nous avons choisi une définition "en négatif" du territoire : l'objet étudié n'est pas territorial s'il s'agit d'un système unifonctionnel, par exemple un produit ("les batteries lithium-ion des véhicules automobiles à motorisation hybride" (LAVISSE et al. 2024)) ou un secteur non explicitement rattaché à une zone géographique. Par exemple, sera retenue dans notre corpus une étude qui porte sur les entreprises de la chimie dans l'Union européenne (BARNOSELL et al. 2024) mais pas une étude qui porte sur le portefeuille de produits d'une entreprise en particulier (VARGAS-GONZALEZ et al. 2019).

Si cette restriction de ce qui fait un territoire reste assez arbitraire et sujette à discussion, elle nous permet néanmoins de dresser un portrait relativement exhaustif et représentatif de ce qui est publié dans le champ de l'AESA, car elle demeure finalement assez peu restrictive. Remarquons que les méthodes basées sur l'ACV (champ dans lequel se situent, la plupart du temps, les études d'AESA) sont principalement orientées sur des produits, des filières industrielles et technologiques ou des secteurs d'activité. Cependant, un nouveau champ émerge actuellement au sein de la communauté ACV, celui des *ACV territoriales*. Ces dernières peuvent être distinguées en deux catégories : d'une part, les analyses qui se concentrent sur un produit, une filière ou un secteur donnés, mais qui sont ancrées dans un territoire spécifique et, d'autre part, celles qui essaient d'estimer les impacts de l'ensemble des activités de production et consommation d'un territoire donné (LOISEAU et al. 2018).

2.2.1.2 Constitution du corpus de publications

En gardant ces remarques préliminaires en tête, qui permettront de discriminer parmi les publications trouvées celles qui seront retenues pour le corpus, nous restreignons la recherche des publications aux années de la période 2009-2024.

Les publications (IVa - applications sous-globales, eau uniquement) ont été identifiées par une requête sur le moteur de recherche ScienceDirect avec les mots clés suivants dans le champ "title" : "water" & "planetary boundaries" ($N = 11$). Quant aux publications (IVb - application sous-globale, toutes limites), elles sont identifiées par deux requêtes distinctes dans le champ "title" : la première avec les mots-clés "planetary boundaries" NOT "planetary boundary layer" ($N = 129$), et la seconde avec les mots-clés "absolute sustainability" OR "absolute environmental sustainability" OR "safe operating space" ($N = 44$). Les publications (III - revue de la littérature) sont identifiées par l'une et l'autre de ces deux requêtes. À noter que la recherche des publications (IVb) a permis d'enrichir la base des publications (IVa), et que les références incluses dans les publications (III), (IVa) et (IVb) mais non référencées dans ScienceDirect ont également permis d'augmenter le nombre des publications traitées. Cette recherche a été actualisée pour la dernière fois en juillet 2024.

En restreignant par la dimension territoriale, nous retenons 71 publications. Toutes ces études ne portent pas nécessairement sur des cas d'applications quantifiés, et donc toutes ne mobilisent pas nécessairement des données quantitatives, qui sont les données qui nous intéressent pour notre revue. Finalement, 39 publications constituent le corpus restreint qui nous intéresse parce que la limite relative à l'eau est appliquée à un système que l'on peut rapprocher de la notion de "territoire" et parce qu'une évaluation AESA est conduite sur ce système à partir de données quantitatives. Dans ce corpus, les publications du type (IVa - applications sous-globales, eau uniquement) sont au nombre de 20 et les publications du type (IVb - applications sous-globales, toutes limites) sont au nombre de 19.

2.2.1.3 Une typologie provisoire des données

Nous n'allons pas dans la suite présenter chacune de ces 31 publications, mais plutôt les présenter par type de données utilisées. À ce stade, nous ne proposons pas de grille d'analyse du corpus pour typologiser les données rencontrées. Une telle grille sera mobilisée dans la dernière partie de ce chapitre, qui proposera un retour analytique sur les types de données rencontrées et viendra discuter la question que nous formulions en introduction de chapitre, sur la place des données d'observation et des approches spatialement explicites. Les sections qui suivent sont construites "par dépliage". Nous avons pour cela d'abord isolé, pour chacune des 31 publications, les jeux de données utilisés (souvent plusieurs pour une même publication), en distinguant les

données pour le calcul de l'indicateur de pression et le calcul pour les indicateurs de seuil. Nous avons alors cherché les redondances et les cas particuliers dans les jeux de données mobilisés. L'identification des redondances a donné lieu à une première typologie induite par le corpus lui-même :

- (1) les bases de données d'inventaire (e.g. Ecoinvent, EnerFuture, etc) et les tables entrées-sorties environnementales (e.g. Exiobase, Eora, etc). Bien que différentes dans leur structure et leur contenu, nous rapprochons ces deux jeux de données car ils ont en commun d'être souvent mobilisés par les auteurs qui ont recours à l'ACV comme méthode d'évaluation environnementale. D'autres méthodes d'évaluations peuvent toutefois avoir recours à ces jeux de données.
- (2) les bases de données statistiques des organismes étatiques et supraétatiques de comptabilité nationale ou autres organismes d'études et de *reporting* statistique, que l'on peut distinguer selon leur périmètre géographique (régional, national, global), leur périmètre sectoriel (d'un unique secteur d'activité à l'ensemble des secteurs économiques) et leur unité de compte (unité monétaire ou unité matérielle comme la biomasse agricole, les émissions de gaz à effet de serre, les débits des cours d'eau, etc)
- (3) les données issues de publications scientifiques ou de documentations techniques spécifiques. C'est notamment une pratique répandue parmi les auteurs qui utilisent l'ACV comme méthode d'évaluation environnementale, où des données sectorielles spécifiques issues de la littérature grise ou scientifique viennent compléter les données d'inventaire, et où les facteurs de caractérisation sont généralement issus de la littérature¹⁰.
- (4) les données de terrain, issues d'enquêtes spécifiques, de campagnes de mesures réalisées *ad-hoc* pour l'étude ou d'une instrumentation de type *remote sensing*.
- (5) les données issues de modèles de simulation numérique. Nous précisons bien ici "simulation numérique" pour distinguer ce type de modèles des modèles statistiques, économiques, algébriques qui peuvent être utilisés pour produire les données mentionnées dans les catégories précédentes. Ce type de données n'a été identifié que pour les publications du groupe (IVa - applications sous-globales, eau uniquement), du fait qu'ils ont permis d'obtenir des débits d'eau.

Le tableau suivant synthétise, pour chaque catégorie de la typologie provisoire proposée

10. Nous parlons bien ici de l'ACV spécifiquement dans un contexte d'AESA, où la modification des facteurs de caractérisation comme méthode d'intégration des limites planétaires constitue l'une des voies possible pour le couplage entre l'ACV et les limites planétaires.

ci-dessus, le nombre de publications du corpus concernées (une même publication peut avoir recours à plusieurs types de données) ainsi que le nom des jeux de données associés à chacune ¹¹.

11. La liste des noms des jeux de données associés à chaque catégorie est rendue ici de façon exhaustive, sauf pour deux cas. (1) Lorsque les auteurs ne précisent pas clairement le jeu de données utilisé mais simplement sa famille. Par exemple, (TAN et al. 2022) mentionnent "*multi-regional input-output tables for 42 sectors in 31 provinces in China, 2015*" sans préciser le produit EE-MRIO utilisé, et dans ce cas nous ne reportons rien dans la colonne "nom des jeux de données". De même par exemple pour (M. ALGUNAIBET et al. 2019) qui mentionnent simplement "*other specific technically sound sources*" que nous classons dans le type de données 3 sans rien reporter dans la colonne "nom des jeux de données". (2) Lorsqu'il y a une trop forte redondance entre des jeux de données quasiment similaires. Par exemple, nous indiquons simplement *China and Provincial Statistical Yearbook* pour désigner en même temps toutes les déclinaisons de ce jeu de données dans les différentes provinces de Chine.

Type de données	Nombre de publications			Nom des jeux de données
	<i>total</i>	<i>dont groupe</i>	<i>dont groupe</i>	
		<i>IVa (N=20)</i>	<i>IVb (N=19)</i>	
(1) données d'inventaire et tables EE-IO	15	4	11	Ecoinvent :7, EnerFuture :1, Eora :2, Global MRIO Lab :1, Exiobase :1, ILCD :1
(2) données statistiques	28	15	11	ISIC :1, IEA's NZE scenarios :1, World Bank database :1, EU-28 database :3, China and Provincial Statistical Yearbook :6, China and Provincial Water Resources Bulletin :5, China Energy Statistical Yearbook :1, AHCES :1, RESDC :1, FAOSTAT :3, WRI :1, BTS :1, NASS :1, ETA-Lancet :1, NBS :1, AQUASTAT :3, WARMS database :1, Pacific Institute's WW database :1, GDRC :1
(3) publications scientifiques et techniques	17	7	11	(HOEKSTRA et al. 2011b; HOEKSTRA et al. 2011a; HOEKSTRA et al. 2010a) :4, (LI et al. 2020) :1, (RYBERG et al. 2018) :4,(YANG et al. 2017; YANG et al. 2018) :2, Miscellaneous technical reports from industries or administrations :7
(4) données de terrain	5	2	3	Inventory from farms :1, China Meteorological Data Network :1, Remote Sensing Data for the EU Copernicus Network :1, Field measurements on lakes :1, Qualitative Field Survey :1
(5) données des modèles de simulation	6	6	0	AquaCrop :1, WBM :1, CLM :1, LPJmL :1, WaterGAP3 :1

" : " indique le nombre de publications concernées (une publication peut utiliser plusieurs jeux de données)

ILCD : International Life Cycle Data System

ISIC : International Standard Industrial Classification (World)

IEA's NZE scenarios : International Energy Agency Net Zero Emission Scenarios (World)

AHCES : Annual Households Expenditures Survey (Algeria)

RESDC : Resources and Environment Sciences and Data Center (China)

FAOSTAT : Food and Agriculture Organization's Statistical database (World)

WRI : World Resources Institute (World)

BTS : Bureau of Transportation Statistics (US)

NBS : National Bureau of Statistics (US)

NASS : National Agricultural Statistics Service (US)

AQUASTAT : Food and Agriculture Organization's Water database (World)

WARMS = Water Authorization and Registration System (South-Africa)

WW : World's Water database (World)

GDRC : Global Runoff Data Center (World)

TABLE 2.1 – Tableau de synthèse des données utilisées dans le corpus de publications

Un tableau exhaustif¹², précise pour chaque publication du corpus le type de données utilisées selon cette typologie provisoire, et détaille d'autres dimensions (le cas d'étude et son

12. Accessible dans les formats (.pdf et .ods) à l'adresse de ce dépôt : https://drive.google.com/drive/folders/1R5BDtgR0FG119sb5yCh5E6yR6_MjaBxK?usp=sharing

échelle, la famille méthodologique d'AESA, le traitement des incertitudes, etc).

2.2.1.4 La typologie hiérarchique des données du JRC européen

En 2019 paraît dans la revue *The International Journal of Life Cycle Assessment* un article d'Eleonora Crenna et ses collègues de l'Ispra portant sur les sources de données et les choix méthodologiques lors d'une normalisation de résultats d'ACV (CRENNA et al. 2019). Si cet article ne porte pas au premier abord sur l'évaluation absolue, il est en fait très en lien. En effet, la normalisation de résultats d'ACV est l'angle par lequel l'une des premières méthodes de couplage entre les limites planétaires et l'ACV a été opéré, dans la publication de (BJØRN et al. 2015b). Une méthode alternative avait d'ailleurs été proposée un an plus tôt par les danois Reinout Heijungs, Arjan de Koning et Jeroen Guinée qui consistait à effectuer un tel couplage lors de la phase d'analyse des résultats, en proposant une méthode d'optimisation linéaire du résultat d'ACV sous contrainte de limites planétaires (HEIJUNGS et al. 2014). Jeroen Guinée est précisément le rédacteur en chef de la revue *The International Journal of Life Cycle Assessment*. Quant aux co-auteurs d'Eleonora Crenna, on retrouve Serenella Sala, avec qui elles publieront un an plus tard en 2020 une étude remarquable sur l'impact environnemental de la consommation des ménages dans l'Union européenne au regard des limites planétaires (SALA et al. 2020). Si donc l'article de (CRENNA et al. 2019) ne porte pas explicitement sur l'évaluation absolue, il a indéniablement une portée normative pour les praticiens de l'évaluation absolue.

Dans leur article, Eleonora Crenna et ses collègues proposent un protocole hiérarchique pour la sélection des données d'inventaire, en considérant que l'un des principaux freins à la robustesse des résultats d'ACV tient à l'incertitude inhérente à la construction des données ainsi qu'à la grande variété des données mobilisées dans un même inventaire selon les différents flux en présence. Leur étude consiste à ainsi à identifier les jeux de données les plus fiables pour couvrir les catégories d'impact ILCD. Les auteurs présentent ainsi leur protocole hiérarchique de sélection des données d'inventaire :

Considering the broad variety of available statistical and scientific sources, the detailed hierarchical procedure [...] was used to guide the selection of inventory data, especially when alternative options for the same inventory flows were available. Generally, official statistics based on measured values or officially accepted models were preferred since they ensure higher robustness of the inventories. This was

followed by (i) activity-based estimations, (ii) statistical proxies, and (iii) speculative assumptions based on reasonable correlation and/ or cause-effect models. (CRENNA et al. 2019)

On en dérive ainsi le classement suivant entre les différents types de données :

1. données issues des statistiques officielles basées sur des mesures
2. données issues de modèles officiellement acceptés
3. données issues d'estimations basées sur des données d'activité
4. données issues de proxies statistiques. Un proxy statistique étant défini comme une variable mesurable qui remplace la variable idéale mais non mesurable (STATISTICS 2008).
5. données issues d'hypothèses basées sur une corrélation vraisemblable et/ou sur un modèle de cause-à-effet

Pour assurer davantage la robustesse des résultats d'ACV, et en particulier leur normalisation, à ce protocole de sélection des données s'ajoute une procédure d'évaluation de la robustesse en tant que telle, qui consiste en trois points :

- complétude de la couverture de l'inventaire (*inventory coverage completeness*) : capacité de l'inventaire des données à couvrir la liste des flux élémentaires de l'ILCD pour chaque catégorie d'impact. Trois scores sont proposés selon la couverture de l'inventaire en % : haut, moyen et faible.
- robustesse de l'inventaire (*inventory robustness*) : qualité des données utilisées dans l'inventaire, à savoir que moins il y a de sources de données différentes mobilisées et moins il est fait usage de stratégies d'interpolation pour en augmenter la résolution spatio-temporelle ou en combler les trous, plus la robustesse est grande. Là aussi, trois scores sont proposés selon la part des données d'inventaire ayant subi une interpolation : haut, moyen et faible.
- robustesse de la méthode d'évaluation des impacts : cette robustesse dépend du classement des méthodes proposées par l'ILCD.

La question de la robustesse recoupe évidemment celle de l'évaluation des incertitudes, mais l'aborde de façon moins stricte. La robustesse ainsi évaluée par des scores qualitatifs est davantage une contrainte de transparence à l'égard des praticiens de l'ACV (et plus spécifiquement de l'ACV en contexte d'évaluation absolue) qui sont appelés à documenter leur pratique.

Pour ce qui est de notre cas d'étude sur l'eau et sur un territoire, cette typologie conforte l'hypothèse que nous formulons en introduction de ce chapitre quant à la place centrale qui devrait être accordée aux données d'observation spatialement explicites. Si la typologie hiérarchique du JRC est proposée dans un but généraliste (le type de système évalué ou les catégories d'impacts à évaluer ne sont pas spécifiés), l'évaluation quantitative de l'intégrité des cours d'eau sur un territoire implique d'évaluer un phénomène environnemental tangible à l'échelle du territoire, contrairement à des émissions de gaz, à effet de serre ou destructeurs de la couche d'ozone stratosphérique par exemple. Par *tangible*, il faut entendre *observable* et *mesurable* par une instrumentation dédiée. S'il est techniquement faisable de mesurer les émissions de gaz dans l'atmosphère ou la stratosphère à partir de sources identifiées, un tel dispositif n'est pas déployé systématiquement sur toutes les sources d'émission du territoire pour des raisons évidentes d'infrastructure et de coût. Les cours d'eau en revanche font l'objet d'une instrumentation systématisée pour leur surveillance, par exemple à travers les stations de mesure du réseau Vigicrues en France (dont nous reparlerons au chapitre 4).

Évidemment, les jeux de données mobilisés dans notre corpus de publications ne se présentent pas spontanément sous la forme de la typologie du JRC, qui est une typologie idéale. Il y a en effet dans la réalité une accumulation de la donnée, c'est-à-dire une transformation, depuis une donnée de type (1) qui serait proposée par un organisme officiel spécialisé dans la production d'une unique donnée (par exemple des débits des cours d'eau) jusqu'à sa mise à disposition dans une base de données composite (d'autres données territoriales sont disponibles que les seuls débits) et multi-régionale (d'autres territoires sont disponibles que celui de l'organisme déposant).

Pour les jeux de données qui ne s'inscrivent pas spontanément dans le type (1) de la typologie du JRC, nous allons donc proposer de remonter la chaîne d'élaboration de ces jeux de données. Nous n'allons pas effectuer ce travail pour l'intégralité des jeux de données du corpus, mais pour les plus représentatifs et/ou utilisés d'entre eux, à savoir :

- pour les données d'inventaire et tables EE-IO : les jeux de données Ecoinvent, Eora et Exiobase
- pour les données statistiques : certains jeux de données peuvent directement s'inscrire dans la classe (1) de la typologie du JRC, comme le *Chines Water Resources Bulletin*, et d'autres méritent une attention plus poussée, comme la base AQUASTAT que nous étudierons

- pour les données issues des publications scientifiques et techniques : nous étudierons principalement le travail des chercheurs Hoekstra et Mekonnen (d'autres publications pourront être évoquées indirectement)
- pour les données de terrain : nous considérons que ces données s'inscrivent déjà dans la classe (1) de la typologie du JRC, bien qu'à des degrés de robustesse différents selon le degré de validation scientifique dont a fait l'objet la méthodologie de collecte des données. Nous n'étudierons pas ces jeux de données.
- pour les données des modèles de simulation : nous étudierons le cas du WaterGAP 3. En effet, nous avons déjà décrit les modèles WBM et LPJmL dans la partie (2) de ce chapitre.

Les parties suivantes s'organisent ainsi selon les jeux de données mentionnés ci-dessus. L'enjeu est ici de les présenter avec suffisamment de précision pour pouvoir catégoriser les publications de notre corpus selon la typologie hiérarchique des données.

2.2.2 Les jeux de données représentatifs de la littérature : remonter la chaîne d'élaboration de la donnée

2.2.2.1 Les travaux de l'ONU sur l'empreinte eau

Si le titre du paragraphe mentionne l'ONU, c'est que les chercheurs Mekonnen et Hoekstra étaient (sic¹³) rattachés à l'IHE Delft, qui est le centre de recherche et d'enseignement de l'UNESCO sur les technologies, infrastructures et la gestion de l'eau. Créé en 1957 à Delft aux Pays-Bas par l'UNESCO sous le nom de International Institute for Hydraulic and Environmental Engineering (IHE), il est désormais sous tutelle de l'UNESCO mais n'est plus directement géré par lui.

À ces deux chercheurs, on doit les principaux travaux sur le concept d'*empreinte eau*, qui se sont constitués en jeux de données à part entière. Le premier millésime de ces données d'empreinte eau est publié en 2004 (CHAPAGAIN et al. 2004), avec une résolution par pays et en distinguant les secteurs agricole (culture et élevage) et industriel. La méthodologie pour le calcul des empreintes sectorielles s'est ensuite affiné et à donner lieu à quatre publications distinctes, qui sont celles que mobilisent aujourd'hui la majorité des articles qui effectuent une application

13. Arjen Hoekstra est décédé subitement le 18 novembre 2019 alors qu'il était sur son vélo de retour du travail, voir (VANHAM et al. 2020)

sous-globale de la limite planétaire relative à l'eau. Ces quatre articles sont donnés dans l'ordre de leur parution (et l'ordre a son importance car les travaux sont cumulatifs : pour calculer une empreinte d'un système il faut d'abord calculer l'empreinte liée à ses sous-systèmes). En 2010 paraissent ainsi un rapport sur l'empreinte eau bleue, verte et grise des cultures et produits dérivés (HOEKSTRA et al. 2010a), et un rapport sur l'empreinte eau bleue, verte et grise de l'élevage et produits dérivés (HOEKSTRA et al. 2010b).

Ces deux premiers rapports sont suivis par deux études qui élargissent le périmètre, l'un à l'empreinte eau des nations (HOEKSTRA et al. 2011b) et l'autre à l'empreinte eau des 405 plus grands bassins versants du monde (HOEKSTRA et al. 2011a). La démarche de construction des données d'empreinte eau pour l'élevage et pour les cultures est assez différente, à la fois dans les données sous-jacentes mobilisées et dans la résolution obtenue. Pour l'élevage, Hoekstra et Mekonnen utilisent un ensemble de données très hétérogènes principalement issues de la littérature pour tout ce qui concerne les caractéristiques des systèmes d'élevage, les apports nutritifs des animaux, la composition de la nutrition animale, etc et des données issues des organismes statistiques internationaux (notamment FAO) pour le nombre de têtes de cheptels et les quantités de viandes commercialisées internationalement. La résolution de leur étude sur l'élevage est ainsi nationale. Pour ce qui est de leur étude sur les cultures, les auteurs proposent un modèle spatialement explicite et de résolution bien plus fine que l'échelon national, par un carroyage de 5 arcmin. Le modèle spatial repose principalement sur un calculateur d'évapotranspiration journalière en fonction du type de culture positionnée sur chaque carreau. Il s'agit d'un modèle que les auteurs ont développé spécifiquement pour leur étude, bien qu'ils utilisent également le modèle CROPWAT développé par ailleurs pour certaines cultures spécifiques. Des données d'observation climatiques spatialisées (pluviométrie et température) sont récupérées au pas mensuel et rééchantillonnées au pas journalier grâce au modèle CRU-dGen (daily weather generator model). Les sorties de modèle en termes d'eau bleu, verte et grise sont ensuite sommées aux échelles nationales.

S'appuyant sur leurs résultats d'empreinte eau pour l'élevage et les cultures, Hoekstra et Mekonnen proposent dans leurs études suivantes d'élargir le calcul aux secteurs industriels et domestiques, ce qui permettra, en ajoutant l'empreinte du secteur agricole (distingué en culture et élevage donc) de calculer les empreintes eau totales liées à la production et à la consommation des biens et services économiques. Ce calcul est agrégé aux échelles des nations dans (HOEKSTRA et al. 2011b) et aux échelles des principaux bassins versants du globe dans (HOEKSTRA et al. 2011a). La démarche respectivement utilisée pour l'agrégation à ces deux échelles distinctes

méritent qu'on la décrive car elle fait intervenir des jeux de données supplémentaires, d'une part parce que le calcul est élargi aux secteurs domestique et industriel, d'autre part parce que l'échelle cible (nation ou bassin versant) demande l'utilisation de jeux de données spécifiques à la calibration spatiale. Pour ce qui est du calcul de l'empreinte eau des principaux bassins versants du globe, il faut d'abord souligner que le rapport dans lequel est documenté ce calcul ne se contente pas de calculer l'empreinte eau mais entend calculer une "rareté de l'eau" (*water scarcity*). Cet indicateur hybride celui de l'empreinte eau bleue (qui est utilisé comme indicateur de prélèvement ou de pression sur la ressource) et celui de la disponibilité de l'eau bleue.

Nous allons le voir, cette étude réalisée en 2011 annonce vraisemblablement les formulations de la limite eau à l'échelle régionale dans STEFFEN et al. 2015b et ROCKSTROM et al. 2023. Il s'agit en effet de faire le rapport, pour tout bassin versant, entre l'empreinte eau bleue calculée sur ce bassin et la disponibilité de l'eau bleue sur ce même bassin. La disponibilité de l'eau bleue se définit comme le ruissellement naturel/préserver (*natural/undepleted runoff*) moins un pourcentage de débit écologique préservé (*environmental flow requirement*). On est là très proche de la définition de la limite planétaire sur l'eau qui sera retenue dans STEFFEN et al. 2015b. C'est sans doute ce qui explique l'usage prépondérant de cette base de données parmi les auteurs qui appliquent la limite eau à des systèmes sous-globaux. Mais c'est ici la question des données employées pour le passage à l'échelle des bassins versants et l'élargissement au reste des secteurs qui nous intéresse.

L'indicateur d'empreinte eau bleue sur un bassin versant se décompose en trois empreintes intermédiaires : l'empreinte liée au secteur agricole, au secteur industriel et au secteur domestique. Pour le secteur agricole, les auteurs utilisent exclusivement les résultats de leur étude sur l'empreinte eau des cultures (HOEKSTRA et al. 2010a) : il s'agit d'un résultat spatialement explicite sur chacun des carreaux de l'étude car la simulation du modèle d'évapotranspiration sous-jacent est rejouée sur chacun. Pour l'industrie et le secteur domestique, les prélèvements des deux secteurs sont récupérés auprès des organismes internationaux (FAO) puis spatialement distribués selon la densité de population observée sur chaque carreau de 5 arcmin. Les prélèvements sont également distribués temporellement, en allouant uniformément sur les douze mois de l'année les valeurs annuelles de prélèvement. Cette distribution spatiale et temporelle en fait un modèle spatialement explicite mais par estimation. Enfin, les prélèvements des deux secteurs sont pondérés par un facteur générique de consommation, i.e. la fraction des prélèvements qui n'est pas restituée au milieu, respectivement 5 et 10% pour l'industrie et le secteur domestique.

Quant à l'indicateur de disponibilité de l'eau bleue, il s'obtient en calculant d'abord le ruissellement actuel/réel puis en ajoutant la somme des prélèvements spatialement superposés. Le ruissellement actuel est obtenu à partir du jeu de données *Composite Runoff V1.0 Database* (FEKETE et al. 2002, nous y reviendrons), qui fournit une valeur de débit moyenne mensuel pour tout carreau de 30 arcmin (données spatialement explicites). Ces carreaux ne sont pas rééchantillonnés vers le carroyage cible de résolution 5 arcmin (même si les auteurs ne le précisent pas, nous verrons au chapitre 5 que le rééchantillonnage de données spatialisées qui représentent des valeurs de débits n'est en effet pas évident). À ce stade, l'échelle d'analyse quitte donc les carreaux et passe aux bassins versants : les valeurs de débits actuels à 30 arcmin et les valeurs de prélèvement à 5 arcmin sont sommées à l'échelle (plus grande) des bassins versants. Le débit dit préservé est alors obtenu en sommant ces deux valeurs. Finalement, l'indice de "rareté de l'eau" est calculé comme le rapport entre les prélèvements sur un bassin versant et ce débit dit préservé.

Concernant le calcul de l'empreinte eau aux échelles nationales, effectué dans (HOEKSTRA et al. 2011b), sont distinguées l'approche production (eau contenue dans les productions nationales et exportées) et l'approche consommation (eau contenue dans les productions nationales et importées). Pour l'empreinte eau de la production, la démarche est identique à celle que nous venons de décrire pour les bassins versants. Les données à la résolution de 5 arcmin sont récupérées dans (HOEKSTRA et al. 2010a; HOEKSTRA et al. 2010b) pour le secteur agricole, et de la FAO pour les prélèvements industriels et domestiques nets de rejets, spatialement et temporellement projetés sur la grille de résolution 5 arcmin. À partir des données d'empreinte eau ainsi récupérées pour les différents secteurs industriels de la FAO, un coefficient d'empreinte eau par unité de valeur ajoutée pour chaque secteur est calculée, qui vient ensuite multiplier les données d'import-export pour chaque secteur et chaque pays.

Ce sont ces valeurs d'empreinte eau liées à la consommation des nations que reprennent par exemple Serenalla Sala et le JRC européen dans leur étude sur l'impact de la consommation des ménages européens par rapport aux limites planétaires (SALA et al. 2020), bien qu'indirectement à travers une base EE-MRIO, mais qui s'appuie sur les résultats de (HOEKSTRA et al. 2011b).

2.2.2.2 La ressource en eau dans les bases de données d'inventaire et les tables EE-IO

Les tables EE-IO (*environmentally extended input output*) sont des matrices de données croisant secteurs d'activités et impacts environnementaux. Elles permettent d'estimer les liens entre chaque secteur économique et ses impacts sur l'environnement : consommations de ressources et émissions. Les analyses de cycle de vie proposant une évaluation absolue de la soutenabilité peuvent faire appel à ces bases de données à plusieurs étapes : soit au niveau de l'inventaire du cycle de vie pour estimer des impacts d'arrière-plan, soit au niveau de la phase d'assignation de la capacité de charge, pour estimer la part que représente un secteur d'activité donné dans l'ensemble des émissions (voir chapitre 1).

Nous allons dans cette section étudier la façon dont est produite la donnée relative à la ressource (quantitative) en eau dans les tables entrées-sorties environnementales Eora et Exiobase et dans la base de données d'inventaire Ecoinvent. Bien que différentes dans leur structure et dans leur utilisation, nous considérons sur le même plan ces deux types de données (données d'inventaire et données des tables EE-IO) du fait de leur mode de construction : il s'agit de partir de données d'activités (à l'échelle du procédé pour les bases d'inventaire, et à l'échelle du secteur économique pour les tables EE-IO) et de les coupler avec données sur l'impact environnemental des interventions anthropiques (facteur de caractérisation pour les bases d'inventaire et coefficients environnementaux pour les tables EE-IO). Ces données sont très souvent mobilisées par les auteurs qui mettent en œuvre une méthodologie ACV pour l'évaluation AESA.

Ces deux données existent dans des versions régionalisées. Pour les bases d'inventaire, la régionalisation correspond à deux démarches bien distinctes : l'ACV régionalisée et l'ACV spatialement explicite. Pour les tables EE-IO (*environmentally-extended input-output*), la régionalisation des impacts anthropiques s'opère à partir des tables multi-régionalisées MR-IO et non plus directement à partir des tables nationales "isolées" IO. La multi-régionalisation ne signifie pas d'abord que la résolution des tables est augmentée à une échelle infranationale (régionale) mais que les différentes tables nationales sont liées entre elles et produisent des résultats agrégés cohérents : le terme *multi* est en cela sûrement plus proche de la démarche que le terme *régionalisé*.

Le produit Global MRIO Lab, qui apparaît dans notre corpus de publications, est en fait une plateforme en ligne qui met à disposition les trois principales bases EE-MRIO que sont

Wiod, Exiobase et Eora, mais il ne s'agit pas à proprement d'une méthode de régionalisation des tables EE-IO (LENZEN et al. 2017). Les tables EE-MRIO considérées comme les plus abouties et, dans notre sélection de publications, les plus utilisées, sont les tables Eora (LENZEN et al. 2012) et Exiobase 3 (STADLER et al. 2018). Nous n'allons pas ici effectuer une présentation des performances comparées de ces deux tables, en termes de couverture et granularité temporelle et sectorielle. C'est d'ailleurs dans cet esprit que les concepteurs de la table Wiod, moins populaire que les deux autres dans notre corpus, présentent leurs travaux :

[...] one database should not be seen as 'better' than another database. It may be better (or more appropriate) for answering some questions but not for other questions. For example, it must be a deliberate choice to include as many countries as possible or to strive for a high level of reliability by limiting the number of countries (knowing that for certain countries the quality of the data is poor or that data are not official). Another example is the choice of how to deal with discrepancies that exist between the export and import values recorded in the National Accounts Statistics (NAS) and in the International Trade Statistics (ITS) [...] One option is to take the absolute values from NAS and to assign these to countries-of-origin and countries-of-destination, using shares obtained from ITS. Other options are to take the values from ITS and adapt the product-level exports and imports from NAS, or to develop a construction algorithm that allows for specific choices at product or industry level (DIETZENBACHER et al. 2013)

Partant, la différence principale entre Exiobase et Eora tient sans doute dans l'ambition de chacune des deux tables : Exiobase est un projet financé par l'Union européenne (de même que la table Wiod par ailleurs) dans un but premier d'harmonisation de sa politique environnementale communautaire : une attention particulière est donc portée à l'harmonisation des nomenclatures entre les États membres de l'UE, tandis qu'Eora vise une granularité sectorielle accrue mais au possible détriment d'une comparaison directe entre nations.

Le principe de fonctionnement d'Eora et d'Exiobase est assez similaire. Si le principe de compatibilité macroéconomique de Leontieff est repris aux tables IO, pour assurer la cohérence entre elles de plusieurs tables régionalisées et construire les tables MR-IO, Eora et Exiobase fonctionnent avec des solveurs sous contraintes pour déterminer les meilleurs coefficients techniques (ou autres paramètres intermédiaires) lorsqu'il y a incohérence entre

les données brutes macroéconomiques (matrices de prix, taxes, imports-exports) agrégées et régionales. Sont ensuite ajoutées des *tables satellites* pour la jointure des tables MR-IO avec les tables environnementales, ce qui permet d'obtenir l'équivalent environnemental des coefficients techniques, à savoir les multiplicateurs environnementaux, et ainsi produire les tables EE-MRIO. Ce qui nous intéresse ici, c'est moins le fonctionnement algébrique des tables EE-MRIO que le type de données utilisées pour obtenir les multiplicateurs environnementaux, et en particulier le multiplicateur lié à la ressource en eau.

2.2.2.2.1 Eora

Le *supplementary materials* de (LENZEN et al. 2012) indique que la table satellite liée à l'eau est construite à partir de la base de données *WaterStat*. Cette base de données a depuis changé de nom et correspond à l'actuelle *WaterPub*, qui est la base de référence pour tous les travaux et données sur l'empreinte eau. *WaterPub* est entretenue par le *Water Footprint Network* qui est une fondation non gouvernementale de plus de 200 membres du secteur de la recherche, du public et du privé, fondée en 2008 à l'initiative de Arjen Hoekstra, suite à l'intérêt qu'ont suscité ses travaux fondateurs du concept d'empreinte eau et d'eau virtuelle au début des années 2000 (HOEKSTRA et al. 2002).

Les données d'empreinte eau mises à disposition par les chercheurs Arjen Hoekstra et Mesfin Mekonnen, que nous avons présentées dans la section précédente, sont ainsi les données sous-jacentes à celles qu'Eora met à disposition dans ses multiplicateurs environnementaux.

Remarquons que la table WIOD fait également appel aux données d'empreinte eau des cultures de (HOEKSTRA et al. 2010a). Dans la documentation technique associée à la version 2012 de table WIOD (DIETZENBACHER et al. 2013; GENTY et al. 2012) des comptes satellites dédiés à l'utilisation des ressources ont été ajoutés à WIOD, contrairement à la version 2000 (CORSATEA et al. 2000) où le couplage n'était effectué qu'avec des tables satellites d'émissions de polluants atmosphériques (surtout gaz à effet de serre). À partir du travail de Hoekstra et Mekonnen, les auteurs ont calculé, pour chaque pays et chaque type de culture, les empreintes eau bleue, verte et grise moyennes. Ces intensités eau spécifiques et nationales sont ensuite multipliées par la production pour chaque nation de chaque type de culture (données d'activité FAO, qui figurent également dans la ligne correspondante de la table WIOD). La valeur

d’empreinte eau moyenne sur la période et sur les types de culture est ensuite corrigée pour chaque pays en proportion de l’écart qu’il y a avec cette même valeur nationale moyenne dans (HOEKSTRA et al. 2010a) : il s’agit d’un coefficient de correction (*scaling factor*). La procédure est identique pour l’empreinte eau de l’élevage en suivant les données de (HOEKSTRA et al. 2010b), et très similaire pour les consommations industrielles et domestiques en suivant les données de (HOEKSTRA et al. 2011b).

2.2.2.2 Exiobase

La publication de (STADLER et al. 2018) qui documente la base de données Exiobase 3 précise que les comptes satellites relatifs à l’eau sont construits à partir de trois jeux de données différents : (HOEKSTRA et al. 2011b) et (PFISTER et al. 2014) pour les consommations d’eau bleue de l’agriculture (la résolution retenue est la nation, bien que (PFISTER et al. 2014) proposent une résolution plus importante, pour les 10000+ grands bassins versants du monde), et (FLÖRKE et al. 2013) pour les consommations d’eau bleue de l’industrie.

Les travaux de Stephan Pfister et ses collègues sur l’empreinte eau sont particulièrement utilisés en ACV, car l’ambition des auteurs, avant même leur publication de 2014 reprise par les concepteurs d’Exiobase 3, était de fournir aux praticiens de l’ACV des facteurs de caractérisation spatialement explicites sur les consommations d’eau dans le monde¹⁴. En 2009, ils publient ainsi leur première publication "aboutie" sur de tels facteurs de caractérisation spatialement explicites et applicables à la catégorie d’impact *midpoint* "water deprivation" (PFISTER et al. 2009). La démarche consiste à comparer, pour tout bassin versant élémentaire délimité dans le modèle WaterGAP (+10000 bassins), les quantités d’eau utilisées (*water use*) avec les quantités d’eau disponibles (*water available*). La publication est ambiguë sur la source d’où sont récupérées les valeurs de *water use*. Il est à la fois mentionné le jeu de données de (CHAPAGAIN et al. 2004), mais dont la résolution n’est que nationale. Aucune indication sur un éventuel rééchantillonnage spatial (pour arriver à l’échelle des bassins versants) n’étant donnée, il nous paraît plus vraisemblable que les valeurs de *water use* aient été récupérées de la seconde source mentionnée, le modèle WaterGAP, à une résolution annuelle. Cela dit, les données de (CHAPAGAIN et al. 2004) sont peut-être mobilisées pour passer des prélèvements en eau renseignés par le WaterGAP à la consommation d’eau (prélèvements nets des rejets).

14. Une démarche qui fait écho au groupe de travail international WULCA sur les facteurs de caractérisation spécifiques à l’eau en ACV. Voir notamment (BOULAY et al. 2018).

Après une prise en compte de la variabilité saisonnière du ratio prélèvements-sur-disponibilité (*withdrawals-to-availability*) par un proxy de précipitations, et après arrangement des valeurs sur une échelle logistique (valeurs de 0.01 à 1), les auteurs aboutissent à un "indice de stress eau" (*water stress index*) qui sert de facteur de caractérisation. Les facteurs de caractérisation sont également exprimés aux échelles nationales en sommant les facteurs des bassins versants à l'intérieur des emprises nationales (moyennant l'utilisation d'une pondération calculée comme la consommation d'eau bleue totale de l'emprise) et, pour les bassins versants transnationaux, une allocation à la surface relative des emprises nationales est effectuée (ce qui est très discuté car la superficie d'un bassin versant n'est pas nécessairement corrélé à son ruissellement).

S'appuyant sur leur étude de 2009, Pfister et ses collègues améliorent en 2014 la résolution temporelle de leurs facteurs de caractérisation en reprenant leur étude avec un pas mensuel (PFISTER et al. 2014). La mensualisation est effectuée à l'aide de deux facteurs correctifs qui viennent modifier la valeur du ratio *withdrawals-to-availability*. D'une part, en s'appuyant sur les données mensuelles et annuelles de *withdrawals-to-availability* récupérées dans un autre modèle (Water Balance Model, FEKETE et al. 2002), un facteur mensuel-sur-annuel est calculé qui vient corriger la valeur initiale de *withdrawals-to-availability* annuelle. Un second facteur intervient pour corriger les disparités climatiques spatiales. Ce facteur est calculé comme l'écart-type des précipitations moyennes mensuelles sur la période 1961-1990 : il s'agit d'approximer les disparités climatiques par la variabilité interannuelle des précipitations moyennes mensuelles (et donc par proxy du ruissellement). Il est également mentionné que l'empreinte eau bleue des cultures n'est plus, dans cette mise à jour de 2014, récupérée de (HOEKSTRA et al. 2011b) mais recalculé par la moyenne géométrique de 4 simulations différentes effectuées, pour tout bassin versant, avec le modèle CROPWAT.

2.2.2.2.3 Ecoinvent

Ecoinvent, dans sa version 3, est la base de données environnementales la plus mobilisée dans notre corpus parmi les publications adoptant une méthode d'évaluation par ACV.

La documentation source d'Ecoinvent 3 (WERNET et al. 2016) indique que la question de l'eau est traitée de la façon suivante. La publication de PFISTER et al. 2016 entre davantage

dans le détail des données sous-jacentes relatives à l'eau dans Ecoinvent 3. La base Ecoinvent, depuis sa première version, est dédiée à l'inventaire de cycle de vie (LCI), en particulier parce que les données d'arrière-plan (*background*) représentent généralement 99% de flux élémentaires dans une ACV. Ecoinvent n'est pas donc pas une base d'analyse d'impact (LCIA). Cette base a été développée dans un contexte partenarial de recherche public-privé en Suisse dans les années 1990 et les évolutions successives de la base se sont employées à régionaliser toujours davantage les données à d'autres territoires, comme le rappelle l'introduction à la présentation d'Ecoinvent 3 :

The ecoinvent LCI database was first published in 2003 with its version 1. It represented the results of many person-years of work and built on top of several existing LCI data collection efforts in Switzerland of the 1990s [...]. With the release of version 2 in 2007, the data basis was significantly expanded and enriched, adding several new economic sectors such as biofuels, consumer electronics, and others to the database. In addition, the geographical coverage increased, with more data covering areas outside of Europe. Sectors such as electronics were covered with data of production in Asia. [...] Due to this increasing international and global use, it was decided to further develop the database into the direction of a globally valid database instead of simply maintaining its status as a Swiss database. The increasingly international use of the database conflicted with some decidedly Swiss supply chains in the database. Users were found to commonly apply Swiss or European datasets to other regions due to lack of geographically appropriate data. This led to inconsistencies on several levels : first of all, conditions and supply chains for important factors, such as electricity mixes, vary with geography for many activities and, secondly, with the increasing evidence of the importance of regionalized impact assessment. (WERNET et al. 2016)

Concernant les données relatives à l'eau dans Ecoinvent 3, il s'agit d'un effort parallèle au développement propre d'Ecoinvent qui s'est structuré dans un projet dédié depuis 2010 :

The update of ecoinvent regarding the water inventory is based on the Quantis Water Database (WDB) project that was launched in collaboration with ecoinvent at the end of 2010 together with eight private companies (Veolia Environnement, Natura, Steelcase, Kraft, Danone, Molson Coors, Unilever, L'Oréal) and ended in early

2012. To quantify water inputs and outputs, the WDB was built on the ecoinvent 2.2 water data and typical water consumption rates for different industries based on literature data. Additionally, primary data was collected from literature for several hundred processes. The details are described in Vionnet et al. (2012). The resulting inventory database was used to update inventories that existed in ecoinvent v2.2 for ecoinvent v.3.0. Ecoinvent v3.1 updated all processes/activities of version 3.0 to the same standard based on the procedure of the WDB. (PFISTER et al. 2016)

La construction de la base WDB est documentée dans le rapport technique produit par le groupe de consultants et praticiens de l'ACV, Quantis International (VIONNET et al. 2012). Contrairement aux autres jeux de données que nous avons pu documenter plus haut, la WDB ne répond pas à un unique objectif de quantification (e.g. eau consommée par les cultures, eau consommée par l'élevage, eau prélevée par l'industrie, etc) mais vise à quantifier l'intégralité des flux élémentaires contenus dans Ecoinvent 2 (puis Ecoinvent 3). *In fine*, les flux élémentaires relatifs à l'eau sont distingués en 12 flux élémentaires d'intrants (e.g. `water`, `cooling`, `unspecified natural origin` ou `transformation, from inland water body, unspecified`) et 10 flux d'extrants (PFISTER et al. 2016).

Ces flux élémentaires sont dérivés du protocole suivant, dont nous reprenons la nomenclature :

- paramètre A : quelle est la quantité d'eau de surface prélevée par usage et par pays? Les données utilisées proviennent d'AQUASTAT 2011 et (HOEKSTRA et al. 2011b). Sont ensuite distingués les prélèvements entre les eaux de surface et les eaux souterraines, par pays et par usage si les données sont disponibles dans les deux bases sources mentionnées, soit en appliquant un ratio générique calculé comme la moyenne mondiale des prélèvements en eau de surface sur les prélèvements totaux à partir d'AQUASTAT. Pour les eaux de surfaces, les prélèvements en lac et rivière ne sont pas distingués. Pour l'irrigation agricole spécifiquement, eaux superficielles et souterraines sont distinguées d'après l'étude de (SIEBERT et al. 2010) où le modèle Global Crop Water Use est développé et utilisé et calibré sur les données AQUASTAT.
- paramètre B : quelle est la quantité d'eaux souterraines prélevées à un rythme non soutenable? Pour la fraction prélevée dans les eaux souterraines, ce paramètre évalue la proportion qui est prélevée à un rythme supérieur à la recharge des aquifères. Les ratios de renouvellement soutenables des eaux souterraines sont issues du (GLEICK et al. 2006).
- paramètre C : quelle est la quantité d'eaux souterraines fossiles (non rechargeables) ? Les

données sont issues des deux aquifères documentés : le *Nubian Sandstone Aquifer System* et le *Saudi Arabia Aquifer System*. Tout pays est supposé prélevé dans ces aquifères en proportion de son ratio prélèvement non renouvelable d'eaux souterraines sur prélèvement total d'eau souterraine.

- paramètre D : quelle est la quantité d'eaux souterraines prélevée dans les zones saturées (*shallow aquifers*) ? Cette donnée est récupérée de (SIEBERT et al. 2010) en comparant la surface des zones saturées à la surface totale des aquifères de tout type par pays.

Mis bout-à-bout, ces paramètres permettent d'estimer, pays par pays, les prélèvements en eau. Ils sont considérés comme des intrants (*inputs*) dans la nomenclature Ecoinvent. La WDB entend également renseigner sur les quantités d'eau consommées (prélèvements nets de restitution au milieu), qui sont considérés dans Ecoinvent comme des extrants (*outputs*). En particulier, la WDB rend compte de la fraction consommée des prélèvements par trois mécanismes : évaporation, rejet, incorporation. Pour ce faire, secteur par secteurs, les consommations sont modélisées :

- pour les procédés de l'industrie : l'incorporation est considérée nulle et l'évaporation est distinguée en évaporation pour refroidissement et autres sources d'évaporation. Les données ne sont pas régionalisées mais sectorisées par technologies et les taux d'évaporation sont issus de la documentation technique pour chaque procédé impliqué.
- pour l'agriculture : l'évapotranspiration des pluies naturelles et de l'irrigation est prise en compte, les rejets (trop d'eau) ne sont pris en compte que pour l'irrigation, l'incorporation par les plantes est prise en compte pour les pluies naturelles et l'irrigation, à partir de publications de références dans la littérature dont (PFISTER et al. 2009).
- pour l'hydroélectricité : l'évaporation est prise en compte pour les procédés de génération électrique à pompe et en lac de retenue, mais pas pour les procédés en eau courante. Les données d'évaporation sont issues de (PFISTER et al. 2009).

2.2.2.3 Un modèle de sciences économiques appliqué à l'environnement : le WaterGAP

Abordons à présent la description du modèle WaterGAP, souvent mentionné jusqu'ici, pour lequel on peut identifier au moins deux publications de références : (ALCAMO et al. 2003 ; FLÖRKE et al. 2013). Le WaterGAP est un modèle hydrologique global qui quantifie l'utilisation humaine des eaux souterraines et de surface ainsi que les flux et le stockage de l'eau sur toutes

les surfaces émergées de la Terre.

Florke et ses collègues proposent d'améliorer le modèle WaterGAP2 dans leur étude (FLÖRKE et al. 2013), en particulier en lui ajoutant un module pour l'évaluation des consommations d'eau domestiques et industrielles sur toutes les années de la période 1950-2010, alors que la version WaterGAP2 ne prenait en compte de façon dynamique que les consommations liées à l'agriculture, les consommations domestiques seulement pour quelques régions et les consommations industrielles seulement pour l'année 1995 (ALCAMO et al. 2003).

Pour la consommation d'eau domestique (prélèvements nets de rejets), les auteurs utilisent un proxy de richesse économique, en faisant l'hypothèse que plus le niveau de vie augmente (mesuré par le produit intérieur brut par tête), plus la consommation d'eau par tête augmente, fortement d'abord puis de moins en moins à mesure que les consciences écologiques se forment et que le progrès technique rend les technologies moins intensives en eau (typiquement, les auteurs considèrent un coefficient annuel de baisses des consommations de -2%/an pour les pays dits développés de l'ONU). Ces courbes sont obtenues par collecte directe de données statistiques sur les consommations d'eau auprès des organismes nationaux compétents pour 172 pays, avec une bonne couverture pour la période 1990-2010. Le niveau de richesse est récupéré de la même façon pour la période 1960-2010 puis à partir de la base de données *History Database of the Global Environment* (HYDE version 3.1) pour les chroniques antérieures. Cette même base de données fournit des cartes de densité de population à une résolution de 5 arcmin qui sont utilisées pour désagréger les prélèvements et niveaux de richesses du national vers un carroyage de 5 arcmin. L'intensité des consommations domestiques en $m^3/hab/an$ pour toute année de la chronique 1950-2010 est estimée pour les années manquantes en faisant l'hypothèse d'une relation sigmoïdale typique à partir des points connus. Ces intensités sont désagrégées de l'échelle nationale vers le carroyage de résolution 5 arcmin selon la procédure mentionnée.

Les auteurs modélisent également les prélèvements pour le refroidissement des centrales thermoélectriques, en distinguant 14 types de centrales et en récupérant une intensité de prélèvement en eau par MWh d'énergie générée pour chaque type (données issues de la littérature). L'emplacement et la puissance des centrales est récupérée de la base de données *World Electric Power Plants Data Set of the Utility Data Institute*. Pour ce qui est des prélèvements d'eau liés à l'industrie manufacturière, la démarche est sensiblement la même que pour la reconstruction des prélèvements domestiques : les prélèvements sont reconstitués à partir d'un proxy d'activité

réputé positivement corrélé, en l'occurrence la valeur ajoutée. Ces prélèvements sont également projetés de l'échelle nationale vers le carroyage de résolution 5 arcmin selon la densité de population. Finalement, à partir des prélèvements estimés de ces trois secteurs, domestique, thermoélectrique et manufacturier, et des données récoltées sur le volume total d'eau traité par les stations, les auteurs calculent les prélèvements nets (les consommations, qui ne sont pas retournées au milieu). La validation du modèle est effectuée par comparaison avec une étude similaire mais plus ancienne et méthodologiquement différente sur la consommation mondiale d'eau entre 1950 et 1995 (SHIKLOMANOV 2000), ainsi qu'avec les statistiques nationales agrégées pour la période 1970-2010.

Le modèle WaterGAP est en fait la combinaison d'un modèle d'utilisation de l'eau et d'un modèle de disponibilité de l'eau. L'amélioration apportée par Klorke et ses collègues dans la version WaterGAP3 concerne donc seulement le modèle d'utilisation de l'eau, qui a été spatialement explicité et temporellement étendu pour les consommations domestiques et industrielles. Quant au modèle de consommation de l'eau par l'agriculture et au modèle de disponibilité de l'eau, ils sont documentés dans la publication associée au WaterGAP2 de (ALCAMO et al. 2003). Pour les consommations d'eau liées à l'agriculture, sont distingués les consommations liées aux cultures et à l'élevage. Pour l'élevage, l'hypothèse faite est que les prélèvements d'eau bleue (le terme n'est pas utilisé car le concept n'a pas encore pénétré le champ au moment de la rédaction de l'article) sont intégralement consommées (métabolisées ou transpirées) par les animaux. Des coefficients de consommation d'eau par type d'animaux sont récupérés dans la littérature et multipliés par le nombre d'animaux d'élevage de chaque type (données GlobalARC 1996). Pour les consommations liées aux cultures, un modèle d'irrigation nette est construit, similaire à CROPWAT, qui simule les prélèvements en eau pour l'irrigation des cultures, selon le type de cultivar, et en retranche l'eau évaporée ou fuitée par le système d'irrigation lui-même (qui s'améliore au fil du temps ou de la richesse du pays selon un facteur de progrès technique). Le modèle a été calibré sur un jeu de données réelles d'irrigation disponible pour les États-Unis et l'année 1995, et les auteurs reconnaissent que "*despite the good agreement with these data sets, it is not expected that the model performs as well everywhere*".

Concernant le modèle de disponibilité de l'eau (plus simplement modèle hydrologique), il se compose de deux modèles intermédiaires : un modèle d'interception (la fraction précipitée qui est directement évaporée par la canopée) et un modèle de ruissellement (la fraction précipitée qui atteint le sol et qui n'est pas évapotranspirée en surface, donc ruisselée). Le modèle prend

donc n compte l'effet de l'occupation du sol par les paramètres d'interception et de ruissellement, ainsi que les paramètres climatiques (précipitations, températures). À ces paramètres physiques s'ajoute les résultats du modèle de consommation d'eau : l'eau ruisselée est réduite du montant de l'eau consommée pour tout carreau. Le WaterGAP2 a été calibré sur les données d'observation (débits annuels moyens) de 724 stations de mesure du *Global Runoff Data Center* (sur une période de 30 ans, plus ou moins selon la disponibilité temporelle des données aux stations). La calibration consiste à ajuster la valeur d'un coefficient exponentiel ajouté dans l'expression du débit ruisselé. Pour 385 stations (et donc les 385 bassins versants modélisés correspondants), une valeur raisonnable pour le coefficient de calibration a pu être trouvée. Pour le reste des bassins versants, le coefficient de calibration est estimé à partir d'une régression linéaire multiple sur les bassins calibrés entre leur valeur calibrée de ruissellement et leurs caractéristiques. Une comparaison intéressante est d'ailleurs faite par les auteurs du WaterGAP2 entre leur modèle hydrologique et le modèle de Vörösmarty et ses collègues, qui deviendra le *Water Balance Model* :

One difference between the approaches of Vörösmarty et al. (1998) and WaterGAP 2 is that they do not explicitly take into account interception. The macroscale model of Vörösmarty et al. (1998) was used by Fekete et al. (1999) to compute long-term average runoff at the global scale. They adjusted the modelled runoff by introducing a correction factor that minimizes the difference between modelled and measured discharge, rather than by calibrating model parameters. Thus, for basins with discharge measurements, the hydrological model is used for the spatial interpolation of runoff. However, the computed long-term average runoff distribution is not consistent as the time period of the climate input mostly does not coincide with the diverse time periods of available discharge measurements and because the reduction of river discharge by human water consumption was not taken into account.
(ALCAMO et al. 2003)

2.2.2.4 La collecte des données sur l'eau par la FAO

Enfin, il nous faut mentionner la très utilisée base AQUASTAT des Nations Unies, qui est opérée par la FAO (*Food and Agriculture Organization*).

AQUASTAT est le successeur du programme GWIS (*Global Water Information System*) de la FAO, né en 1993 du besoin de disposer sur l'eau d'un système d'information géographique

standardisé et mondial, dans la veine de ce qui existait déjà depuis les années 1980 pour l'occupation des sols et les ressources de la terre. Le programme GWIS a évolué l'année suivante, en 1994, en deux programmes distincts, un programme d'information géographique centré sur l'échelon national (AQUASTAT) et un programme centré sur les bassins versants et les études hydrologiques physiques¹⁵.

AQUASTAT est bâti sur une collecte de données auprès des organismes récoltants dans chaque pays où l'ONU opère. Depuis 2018, la fonction de collecte s'est incarné dans des correspondants directs de la FAO au sein des institutions nationales (*National Correspondents*), qui sont chargés annuellement du programme de collecte, de validation et de remontée des données au niveau supranational. En 2019, 107 pays parmi les 187 pays du périmètre des Nations Unies avaient nommé un tel correspondant. La remontée des statistiques s'effectue sur une base annuelle par l'intermédiaire du *Water and Agriculture Questionnaire*, qui comporte 35 variables hydrologiques à renseigner et doit être remonté à la FAO le premier semestre de chaque année¹⁶. Un questionnaire plus complet est remonté tous les cinq ans. La FAO effectue derrière un contrôle de qualité, standardisation et validation des données. Avant 2018 et les correspondants nationaux, la FAO se chargeait directement de compiler les données nationales à partir des sources disponibles publiquement et de la littérature scientifique.

Les données récoltées annuellement sont de cinq types :

- les ressources en eau : il s'agit de toutes les ressources disponibles dans chaque pays, en surface comme en souterrain, accessible comme non accessible, d'origine nationale ou extérieure (i.e. héritée de l'amont du territoire). Les questionnaires sont accompagnés d'une méthodologie standardisée pour le calcul des ressources en eau, mais différenciée selon les cas particuliers (notamment en situation de bassin versant transfrontalier). La métrique principale qui est calculée pour l'évaluation de la ressource est la disponibilité d'eau renouvelable totale (TRWR, *total renewable water resource*), qui se divise en une fraction interne au territoire (IRWR) et une fraction externe héritée de l'amont (ERWR). L'IRWR est le débit annuel moyen (sur une tendance longue) des eaux de surfaces et souterraines enregistrées nationalement, moins le débit de décharge des nappes d'accompagnement (typiquement débit d'étiage), plus le débit de recharge des rivières vers leurs aquifères. Pour calculer ces paramètres, la FAO n'est pas contre l'utilisation de modèles localement adaptés qui convertissent des données de pluie en données de débits, ou l'utilisation de

15. <https://www.fao.org/aquastat/en/overview/history>

16. <https://www.fao.org/aquastat/en/overview/methodology/>

références issues de la littérature. Si les données sont issues de mesure de terrain, l'IRWR doit être augmenté des consommations d'eau enregistrées au même endroit pour rendre compte de la ressource totale disponible. Pour ce qui est de l'ERWR, il doit être mesuré ou estimé à la frontière du territoire par où l'eau arrive sur le territoire. À cet endroit, le débit annuel moyen de long-terme est considéré comme une ressource extérieure (MARGAT et al. 2005).

- l'utilisation de l'eau : sont distingués les prélèvements industriels, agricoles, municipaux/domestiques et l'évaporation des lacs artificiels et réservoirs. Les données récoltées pour ces prélèvements ne sont pas issues du questionnaire Water And Agriculture mais récupérées auprès des organismes statistiques internationaux (Eurostat, OECD, etc) et complétées en cas de données manquantes par des régressions et des modèles de bilan hydrologique¹⁷.
- l'irrigation et le drainage : volumes annuels prélevés pour les champs irrigués de façon permanente et pour les champs irrigués de façon discontinue. Le principe de collecte de ces données n'est pas évident¹⁸.
- les eaux usées : volumes annuels rejetés par les stations de traitements des eaux usées, granularité nationale. AQUASTAT ne s'intéresse pas à la charge de polluants contenus dans les eaux usées. Les données proviennent des documentations techniques et rapports des organismes nationaux et de la littérature scientifique¹⁹.

Remarquons qu'à cette échelle de portage institutionnelle (ONU), un autre produit régulièrement mobilisé dans la littérature (plutôt dans la littérature de première-main, en particulier dans les modèles hydrologiques globaux) est la base de données du GRDC (*Global Runoff Data Center*). Cette base est portée par le *World Meteorological Organization* depuis 1988 et mandaté par l'ONU pour soutenir les programmes de recherches des Nations Unies en lien avec le changement climatique. Les données de débits des cours sont en effet identifiées dans le *United Nations Framework Convention for Climatic Change* (UNFCCC) comme l'une des variables climatiques essentielles (EVC, *Essential Climate Variable*)²⁰. La base comprend aujourd'hui plus de 10000 stations de mesure réparties dans 159 pays et les chroniques les plus anciennes remontent à 200 ans. La base de données du GRDC comprend le produit brut des chroniques de débits (débits mensuels moyens) et des produits élaborés : *Annual Characteristics*

17. <https://www.fao.org/aquastat/en/overview/methodology/water-use/>

18. <https://www.fao.org/aquastat/en/overview/methodology/irrig-drainage/>

19. <https://www.fao.org/aquastat/en/overview/methodology/wastewater/>

20. https://grdc.bafg.de/GRDC/EN/01_GRDC/grdc_node.html

and Long-Term Statistics qui élabore des métriques statistiques plus élaborées; *Freshwater fluxes into the World's Oceans* qui restreint la base aux stations des estuaires et embouchures; *Global Composite Runoff Fields* qui implémente un modèle hydrologique global (le *Water Balance Model*) calibré sur les données du GRDC pour accroître la résolution spatiale des données (FEKETE et al. 2002). La différence principale à cette échelle institutionnelle entre les données AQUASTAT et les données du GRDC tient donc aux métriques reportées par les pays : le GRDC compile les données brutes à l'échelle des stations de mesures et élaborent des métriques hydrologiques; AQUASTAT compile pour sa part des données agrégées à l'échelle nationale liées à la disponibilité de la ressource et aux prélèvements, et élabore des indicateurs de gestion.

Une étude de (HUANG et al. 2018) a proposé de rééchantillonner spatialement et temporellement les données nationales de prélèvements et de disponibilité de la ressource contenues dans AQUASTAT vers une grille de résolution spatiale 0.5° et temporelle le mois, pour la période 1971-2010. Pour la désagrégation spatiale, plusieurs techniques ont été utilisées. Concernant les prélèvements domestiques, la génération d'électricité, l'industrie minière et le secteur manufacturier, les données AQUASTAT sont spatialisées par des cartes densité de population (obtenues du produit *History Database of the Global Environment* (HYDE) pour la période 1970–1980 et du produit *Gridded Population of the World* (GPW) pour la période 1990–2010). De même, les prélèvements pour l'élevage sont spatialisés à partir des cartes de densité d'élevage (issues de la *FAO's Animal Production and Health Division*). Le rééchantillonnage temporel quant à lui (annuel vers mensuel) est effectué, pour les prélèvements domestiques et ceux liés à la génération d'électricité, à partir des températures mensuelles spatialisées (*WATCH Forcing Data Methodology*). Quant aux prélèvements pour l'irrigation, leur rééchantillonnage spatial et temporel est réalisé à partir des cartes d'irrigation issues de simulations sur quatre modèles hydrologiques globaux (dont le LPJmL et le WaterGAP, que nous avons déjà eu l'occasion de présenter).

il est intéressant de constater l'hybridation des types de données qui en résulte, où les données de modélisation et les données statistiques servent à rééchantillonner d'autres données statistiques qui elles-mêmes sont à la fois des données d'observation collectées par les différentes agences nationales et des données estimées/modélisées par la FAO.

2.3 Conclusion et questions de recherche

Le corpus de publications que nous avons présenté dans les parties (1) et (2) fait apparaître deux tensions, ou dimensions, pour la caractérisation d'une approche *bottom-up*. D'un côté, la classique *échelle de l'étude*, et de l'autre, l'*échelle de la donnée*, sur laquelle nous avons souhaité mettre l'accent dans ce travail. Au regard du corpus, comment peut-on qualifier ce que nous appelons une *échelle de la donnée* et, partant, définir un cadre analytique pour qualifier une approche AESA de *bottom-up*? Nous allons pour cela dans une première section discuter à nouveaux frais la typologie hiérarchique des données proposée par le JRC européen (CRENNA et al. 2019) au regard de notre corpus, afin d'apporter des éléments pour définir l'*échelle de la donnée* (3.1). Nous proposerons ensuite de croiser cette dimension avec la classique *échelle de l'étude*, et ceci afin d'aboutir à ce que nous appelons un *gradient d'approche bottom-up* (3.2).

2.3.1 Une échelle de la donnée ?

S'il va de soi que la donnée est toujours produite à une certaine échelle spatiale (et temporelle également, mais cette dimension nous intéresse moins ici), il n'est pas évident que cette échelle de production de la donnée coïncide avec l'échelle du système étudié. Nous proposons de définir l'échelle de la donnée selon une double dimension (voir Figure 2.8) :

- d'un côté l'échelle au sens spatial : nous parlerons de *granularité spatiale de la donnée*, c'est-à-dire l'échelle géographique à laquelle la donnée est produite qui définit l'ensemble des différenciations spatiales possibles ;
- de l'autre, l'échelle au sens hiérarchique : nous parlerons de *source hiérarchique de la donnée*, en reprenant le terme à la typologie hiérarchique des données du JRC européen, qui consiste à classer les données selon qu'elles sont plus ou moins spatialement explicites et selon leur niveau de robustesse.

2.3.1.1 Une source hiérarchique de la donnée

À partir de la typologie hiérarchique des données proposée par le JRC européen, nous avons suggéré que la donnée la plus spatialement explicite correspondait à leur catégorie (1) : "données issues des statistiques officielles basées sur des mesures". Rappelons à toutes fins utiles l'ensemble de leurs catégories, hiérarchiquement classées par ordre de robustesse :

1. données issues des statistiques officielles basées sur des mesures
2. données issues de modèles officiellement acceptés
3. données issues d'estimations basées sur des données d'activité
4. données issues de proxies statistiques. Un proxy statistique étant défini comme une variable mesurable qui remplace la variable idéale mais non mesurable (STATISTICS 2008).
5. données issues d'hypothèses basées sur une corrélation vraisemblable et/ou sur un modèle de cause-à-effet

Au regard du corpus que nous avons étudié, en particulier lorsque l'on s'intéresse aux liens sous-jacents entre les jeux de données mobilisés dans les publications du corpus, nous proposons de scinder la catégorie (1) "données issues des statistiques officielles basées sur des mesures" en plusieurs sous-catégories. En effet, le cas d'étude sur l'eau que nous avons abordé vient questionner le qualificatif "basées sur des *mesures*". Ce qui est appelé *mesure* a pu recouvrir dans notre corpus plusieurs réalités, ou plusieurs pratiques de la donnée. Nous avons pu voir dans la partie précédente qu'une valeur de débit est le résultat d'une grande variété méthodologique d'approches : si le *Global Runoff Data Center* compile les valeurs de débits observés par les stations de mesure sur les principaux fleuves du monde, la base AQUASTAT en fait de même mais avec des valeurs de débits agrégés par nation et en autorisant l'utilisation de modèles pour palier localement l'indisponibilité des données. Quant aux valeurs de débits issues de modèles, tels le WBM ou le WaterGAP, celles-ci peuvent être à la fois validées ou calibrées par rapport à des données de mesure sur site, donnant lieu à certains produits mixtes modélisation-observation comme les *Composite Runoff Fields* du GDRC. Pour l'obtention des données nationales sur l'empreinte eau, à savoir les quantités disponibles et consommées par nation, les chercheurs Mekonnen et Hoekstra ont eu abondamment recours à des conversions de données économiques sur les secteurs agricole, industriel et domestique. Ceci nous amène à proposer les catégories suivantes pour raffiner la catégorie (1) du JRC sur les données dites "mesurées" :

- données issues de campagnes de mesure sur site systématisées
- données issues de campagnes de mesure sur site isolées
- données issues de mesures distantes
- données à caractère environnemental issues de reportings
- données à caractère économique issues de reportings. Notons que cette catégorie est différente de la catégorie "données issues d'estimations basées sur des données d'activité" car, pour cette dernière, il s'agit des données non-reportées directement (typiquement,

lorsqu'elles ne sont pas à l'échelle souhaitée pour le système étudié) mais bien *estimées* à partir d'autres données d'activités. C'est par exemple le cas de l'étude reconnue de (LI et al. 2020) dans laquelle les auteurs ont produit un indicateur de *water exceedence* pour les 10 mégapoles chinoises à partir des données d'activités issues des tables MRIO.

La table 2.2 résume les différentes sources hiérarchiques de la donnée que nous proposons, que l'on reporte également sur la Figure 2.8.

Catégorie	Sous-catégorie	Qualité
estimations	proxies	non-explicites
	données d'activité	non-explicites
	modèles acceptés	semi-explicites
reportings	caractère économique	semi-explicites
	caractère environnemental	semi-explicites
mesures	remote sensing	explicites
	sur site isolées	explicites
	sur site systématisées	explicites

TABLE 2.2 – Les différentes sources hiérarchiques de la donnée que nous utilisons pour qualifier l'une des deux dimensions de l'échelle de la donnée

Nous suggérons aussi des critères de robustesse pour ces nouvelles catégories, qui précisent en réalité ce que recouvrait le terme "officiel" dans l'appellation originale de la catégorie que nous avons scindée. On peut pour cela se figurer un plan à deux axes avec en abscisses le degré de systématisation de la donnée et en ordonnées le degré de documentation de la donnée. Le plus faible degré de la systématisation est la production isolée (unique plage temporelle ou date, ou unique localisation géographique), tandis que le degré le plus fort est la standardisation (temporelle et spatiale, qui permet d'obtenir de nombreux répliqués dans des contextes spatiotemporels variés). Pour ce qui est de la documentation, le degré le plus faible est l'absence de documentation spécifique et le degré le plus fort une reconnaissance par les pairs dans la communauté scientifique concernée, par exemple lorsque la documentation méthodologique prend la forme d'un article scientifique publié.

2.3.1.2 Une granularité spatiale de la donnée

Parler d'échelle de l'étude, c'est désigner au moins deux dimensions : l'échelle du système étudié (son emprise géographique : un pays, une région, un SCoT, une entreprise, une infrastructure, etc) et la granularité du système étudié (c'est-à-dire la résolution à laquelle le système est considéré : un territoire peut être vu comme un ensemble de carreaux sur un maillage cartographique, une entreprise comme un ensemble de sites). Nous proposons de mettre la granularité du système étudié au regard de la granularité spatiale de la donnée. Cette dernière peut être vue comme la résolution spatiale des données à partir desquelles le système est quantifié, et qui ne correspond pas nécessairement à la granularité du système étudié. Par exemple, lorsque les données sont définies sur carroyage et que le système étudié est un territoire, si la granularité souhaitée pour le système est territoriale, une opération d'agrégation des carreaux sera nécessaire : la granularité de la donnée ne coïncide alors pas avec celle du système. Si la granularité souhaitée pour le territoire est un ensemble de points dans le réseau hydrographique du territoire, alors il est possible qu'un même carreau élémentaire recouvre plusieurs de ces points, et dans ce cas une opération de désagrégation est nécessaire.

Afin de qualifier, dans le cas d'une évaluation environnementale qui porte sur l'eau de surface, cette granularité spatiale (de la donnée et du système étudié), nous faisons référence à la typologie des échelles d'intervention pour la gestion de l'eau proposée par (DEREPASKO et al. 2021), que nous modifions pour mieux correspondre aux spécificités des échelles rencontrées dans notre corpus (voir Figure 2.8).

Dans une même évaluation, la granularité spatiale du système étudié est généralement unique (qui dépend de l'objectif de l'étude) mais la granularité spatiale de la donnée est généralement multiple car dépendante des jeux de données utilisés.

2.3.2 Vers un gradient d'approches *bottom-up*

Dans une AESA relative à la limite planétaire sur l'eau, nous suggérons de qualifier le degré *bottom-up* d'une approche en croisant deux dimensions : la classique "échelle de l'étude" et ce que nous avons décrit dans la section précédente comme une "échelle de la donnée". Nous proposons d'analyser le caractère *bottom-up* d'une évaluation AESA en deux étapes :

1. la granularité spatiale du système étudié correspond-elle à la granularité spatiale des jeux

de données mobilisés ?

2. à quelle source hiérarchique des données les jeux de données mobilisés correspondent-ils ?

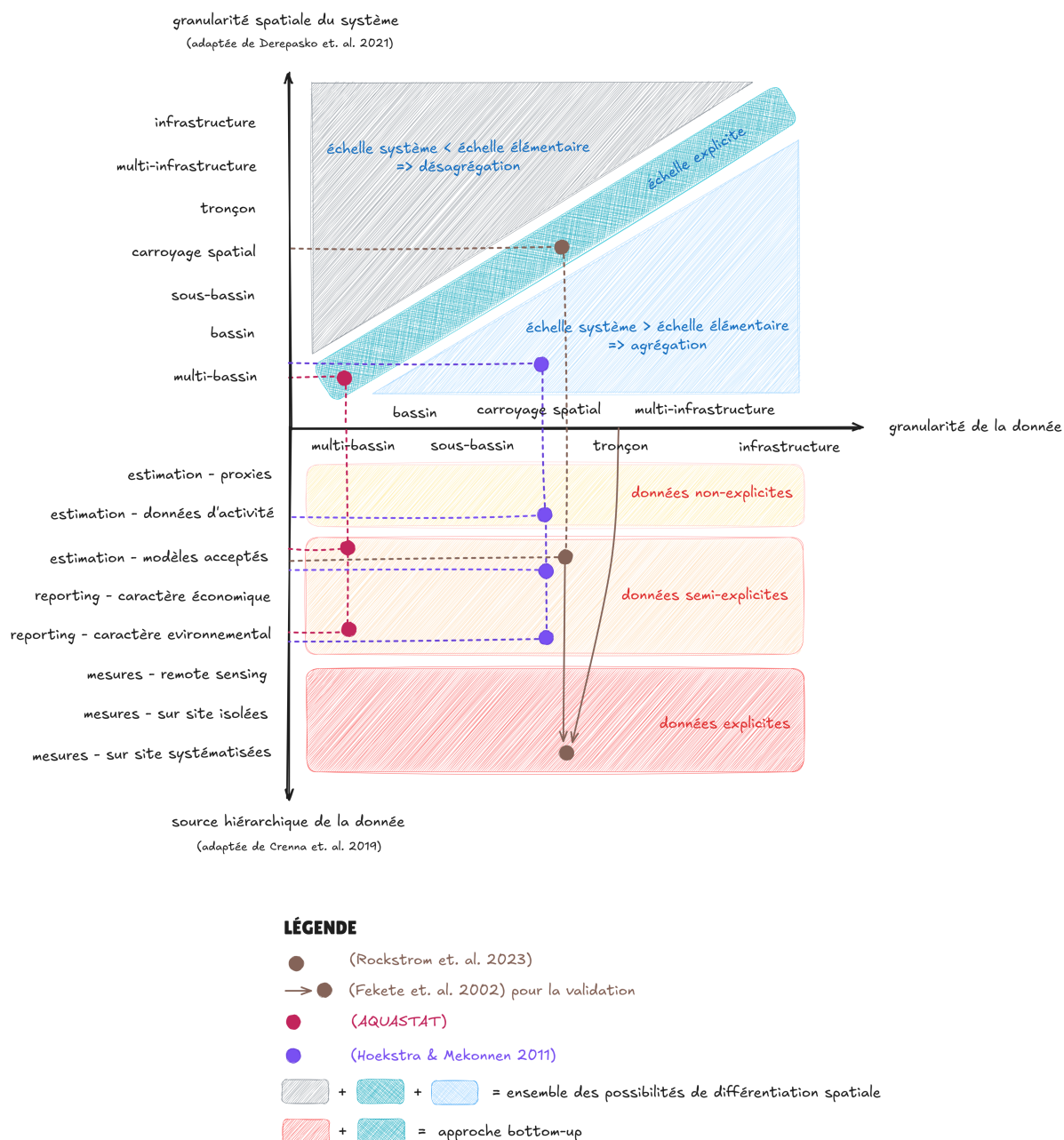


Figure 2.8 – Un gradient d’approche *bottom-up* par le croisement de l’échelle de l’étude et de l’échelle de la donnée

(1) La granularité spatiale du système étudié correspond-elle à la granularité spatiale des jeux de données mobilisés ?

Nous proposons de définir une relation projective entre la granularité spatiale du système étudié et celle de la donnée : lorsque l’on s’intéresse aux jeux de données mobilisés dans une

publication, on peut définir une projection de la granularité spatiale du système étudié vers la granularité spatiale de la donnée où chaque jeu de données est défini. Trois cas de figure peuvent se présenter (pour un jeu de données considéré) :

- la granularité spatiale du jeu de données correspond à celle du système étudié : nous disons alors que l'évaluation est spatialement explicite ;
- la granularité spatiale du jeu de données est plus petite que celle du système étudié : l'évaluation aura nécessairement recours à une stratégie d'agrégation pour remonter à la granularité spatiale du système et, de fait, l'évaluation n'est pas explicite puisque les données ne sont pas explicitement (ou nativement) définie à la granularité spatiale du système étudié ;
- enfin, la granularité spatiale du jeu de données est plus grande que celle du système : l'évaluation aura nécessairement recours à une stratégie de désagrégation pour descendre à la granularité du système étudié. Nous disons d'une telle évaluation qu'elle n'est pas explicite.

Ces trois cas de figure dessinent l'ensemble des combinaisons possibles de ce qui est généralement appelé la *différentiation spatiale* en évaluation environnementale.

(2) À quelle source hiérarchique de la donnée les jeux de données mobilisés correspondent-ils ?

La granularité spatiale des jeux de données mobilisés peut à son tour être projetée sur l'axe des sources hiérarchiques de la donnée. Les données que nous qualifions d'*explicitites* sont celles qui ont été produites par mesure, pour lesquelles nous distinguons trois sous-catégories selon le degré de robustesse de la méthode de mesure. Pour effectuer cette projection, il s'agit de se demander, pour chaque jeu de données, à quelle source hiérarchique de données ils correspondent. Par exemple, des données carroyées (granularité spatiale) peuvent être le résultat d'un calcul effectué à partir de données dont la source est un *reporting* environnemental, un modèle de simulation numérique ou une estimation basée sur des données d'activité (voir l'exemple de (HOEKSTRA et al. 2011b) dans la Figure 2.8).

Finalement, une fois ces deux projections effectuées (de la granularité du système vers la granularité de la donnée, puis de la granularité de la donnée vers la source hiérarchique de la donnée), nous proposons de qualifier une évaluation de *bottom-up* lorsque celle-ci est spatialement explicite (la granularité du système coïncide avec la granularité des données pour tous les jeux de données employés) et lorsqu'elle la source hiérarchique de la donnée est de type explicite pour tous les jeux de données employés. Cette approche simplifie nécessairement la réalité, où

un jeu de donnée peut lui-même être le résultat d'une stratégie d'agrégation/désagrégation d'échelles, ou bien peut n'être que partiellement basé sur des données de mesure (voir l'exemple de (ROCKSTROM et al. 2023) et (FEKETE et al. 2002) dans la Figure 2.8), mais elle souligne du moins la nécessité, pour qualifier une approche de *bottom-up* et donc de *territoriale*, de faire le lien entre l'échelle du système et l'*échelle de la donnée*, qui est à la fois spatiale et qualitative (hiérarchique).

Dans la suite du manuscrit, nous nous sommes proposés d'étudier deux manières de mobiliser des données environnementales explicites de classe 1 ("mesures - sur site systématisées") dans une approche AESA de la frontière du système Terre sur l'eau de surface. En partant de notre cas d'étude sur le territoire du SCoT Sud-Loire, nous avons souhaité mettre à l'épreuve deux questions de recherche :

- comment effectuer l'évaluation environnementale de l'intégrité des cours sur un territoire au regard de l'indicateur proposé par (ROCKSTROM et al. 2023) à partir de mesures de débits sur site systématisées ? Cette proposition étudie la faisabilité d'un indicateur AESA basé sur des mesures de débits sur site systématisées.
- comment coupler des données issues de mesures de débits sur site systématisées avec des données issues des sorties d'un modèle hydrologique global ? Autrement dit, comment coupler des données spatialement explicites avec des données non spatialement explicites mais spatialement différenciées ? Cette proposition étudie la faisabilité d'une assimilation entre types de données différents et étudie ainsi la possibilité de combiner des degrés différents d'approche *bottom-up*.

Le chapitre 3 présentera la méthodologie associée à ces deux questions de recherche, tandis que le chapitre 4 présentera les résultats obtenus.

Chapitre 3

Utiliser des données de mesure de débits dans une AESA : deux propositions méthodologiques

Introduction

Le chapitre 2 a mis en avant le rôle méthodologique joué par les données environnementales dans le but de conduire une AESA dans une approche *bottom-up*. Une échelle des données a ainsi été proposée au regard de la classique échelle de l'étude. Cette échelle des données permet de classer hiérarchiquement les jeux de données environnementales selon leur niveau d'explicitation spatiale. Dans cette échelle des données, nous avons identifié les données de mesure comme les plus spatialement explicites et donc les plus à même de permettre de quantifier une capacité de charge selon une approche *bottom-up*.

Face à l'absence de publications dans notre corpus qui mobilisent des données de mesure pour conduire une évaluation basée sur le cadre de (ROCKSTRÖM et al. 2009a), il a paru judicieux de documenter cette démarche. Nous avons souhaité rendre compte de deux utilisations possibles pour les données de mesure dans une AESA :

- comment effectuer l'évaluation environnementale de l'intégrité des cours d'eau sur un territoire au regard de l'indicateur proposé par (ROCKSTROM et al. 2023) à partir de mesures de débits sur site systématisées ? Ce type de données n'étant pas mobilisé par (ROCKSTROM et al. 2023), nous proposerons le développement d'une méthode d'AESA (MAESA) qui tienne compte à la fois de la spécificité des données de mesure et de la

spécificité de l'indicateur proposé dans cette version du cadre. Nous évaluerons alors la pertinence de la méthode AESA proposée.

- comment coupler des données ponctuelles issues de mesures de débits sur site systématisées avec des données carroyées issues des sorties d'un modèle hydrologique global? Autrement dit, comment coupler des données spatialement explicites avec des données non spatialement explicites mais spatialement différenciées? Il s'agit d'étudier la faisabilité d'une *assimilation* entre types de données différents (différents selon notre échelle des données et différents comme objets géomatiques). Nous présenterons un modèle de rééchantillonnage de données hydrologiques (MRH) que nous avons développé dans ce but, afin de permettre une comparaison directe entre des données explicites (mesures) et des données différenciées (sorties de modèle).

La première question, qui envisage la possibilité d'une utilisation directe dans une AESA des données de mesure de débits, c'est-à-dire non couplée avec d'autres types de données telles que nous les avons définies dans le chapitre précédent (données estimées, données d'activité, etc), amène à réfléchir sur ce que désigne la "variabilité excessive" du débit, qui est le concept central dans la formulation de l'indicateur dans (ROCKSTROM et al. 2023). En particulier, alors que la variabilité des débits est déterministe lorsque les données d'un modèle hydrologique sont utilisées (les pressions anthropiques, qui sont à l'origine de cette variabilité, sont données d'avance comme paramètres d'entrée du modèle), la variabilité observée par des données de mesure est le résultat de pressions multiples qui ne sont pas toutes ni identifiables ni quantifiables. Sans une spatialisation des données, l'AESA se réduit à une évaluation temporelle (voir Figure 3.1). Par ailleurs, des données de mesure de débits sont des objets spatialement ponctuels, qui sont définis dans l'espace par les coordonnées de la station de mesure ou de l'instrumentation hydrométrique. Rappelons que nous avons défini le système *territoire* sur lequel nous travaillons comme un système sous-global multi-fonctionnel et *dont l'emprise est une zone*, c'est-à-dire que les pressions anthropiques y sont diffuses sur toute la surface de l'emprise. Un premier enjeu émerge donc dans l'utilisation de données de mesure ponctuelles : comment spatialiser les pressions mesurées explicitement mais de façon ponctuelle? Pour identifier et quantifier ces sources de pressions diffuses, il est nécessaire de spatialiser l'étude de la variabilité. C'est pour cette raison qu'un modèle hydrographique élémentaire (MHE) est nécessaire (voir Figure 3.1), afin d'isoler les sources de pression spatialement diffuses pour reproduire l'effet explicatif d'un modèle, et afin de raisonner en termes de *zone*, ce qui est le propre du système territoire que nous étudions. Sans capacité explicative, la variabilité n'apporte que peu d'informations sur l'intégrité des cours

d'eau : où s'arrête la variabilité "naturelle" ou erratique d'un cours d'eau mesurée en un point et où commence la variabilité induite par les activités humaines sur l'ensemble du bassin versant ?

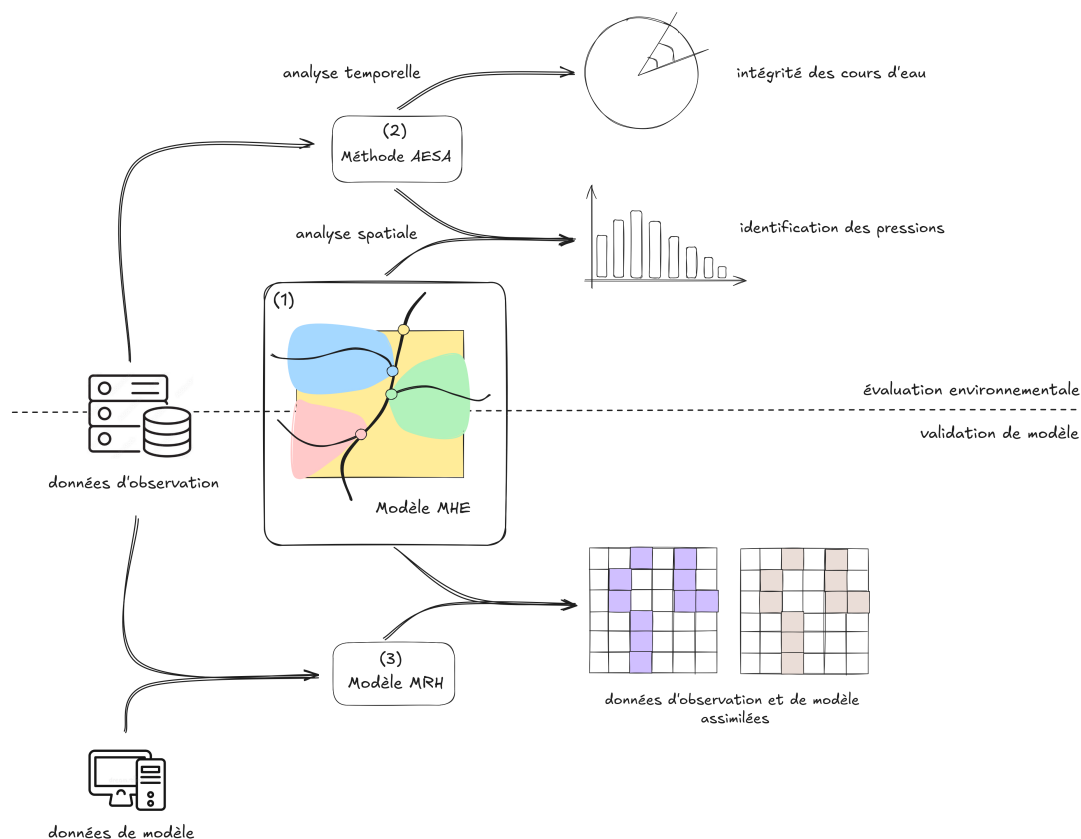


Figure 3.1 – Schéma conceptuel de la méthodologie en 3 parties : méthode pour le MHE (1), méthode pour l'AESA (2), méthode pour le MRH (3)

La seconde question de recherche envisage quant à elle une utilisation indirecte des données de mesure. Il ne s'agit plus de procéder à une AESA à partir des données de mesure, mais de voir comment les données simulées par un modèle hydrologique global (e.g. le *Water Balance Model*) peuvent être couplées, ou assimilées, aux données de mesure. Une telle assimilation, et ceci vaut d'ailleurs pour toute démarche de *data assimilation*, envisage au mieux de corriger, au moins de quantifier, les incertitudes liées à la fois aux données modélisées (biais, erreur paramétrique, erreur structurelle, etc) et aux données de mesure (non-représentativité/homogénéité, non-fiabilité, etc) (DELMOTTE et al. 2011). Une source d'incertitude importante dans une évaluation basée exclusivement sur des données de mesure relève de ce que l'*état de référence* environnemental n'est pas absolu, nous le verrons, mais dépendant de la série temporelle utilisée. Les données de modèles permettent en revanche de disposer d'un tel état de référence, au prix toutefois des incertitudes inhérentes à la modélisation elle-même. Tout l'enjeu est alors de quantifier l'écart entre les sorties du modèle et les données de mesure qui leur correspondent,

dans le but, d'abord, de valider les sorties du modèle, voire d'optimiser le paramétrage du modèle en entrée. Nous détaillerons donc le modèle d'assimilation que nous avons développé, que nous appelons modèle de rééchantillonnage hydrographique (MRH) (voir Figure 3.1).

Ce chapitre commence par présenter le MHE qui a été développé pour ce travail de thèse (1). Ce modèle permet à la fois de soutenir une analyse spatiale de la variabilité mesurée ponctuellement, mais aussi de servir de support à l'assimilation entre les données de mesure et les données de modèle hydrologique global (voir Figure 3.1). Nous détaillerons ensuite la méthode d'AESA que nous proposons pour évaluer l'intégrité des cours d'eau sur un territoire à partir de données de mesure (2). Nous présenterons dans un dernier temps le modèle MRH qui a été développé pour permettre d'assimiler des données de mesure à des sorties de modèle hydrologique global (3).

3.1 Le modèle hydrographique élémentaire

Le modèle hydrographique élémentaire (MHE) que nous avons développé permet de disposer des données hydrographiques spatiales qui serviront de support aux traitements effectués dans les parties suivantes. Par "modèle hydrographique", nous entendons tout ce qui relève du tracé topographique des cours d'eau, à savoir en particulier le sens d'écoulement des cours d'eau et la délimitation des bassins versants. Il ne s'agit pas à ce stade d'un modèle hydrologique, dans lequel la composante hydraulique est ajoutée à la composante topographique. Le modèle hydrographique que nous mettons en œuvre n'est autre qu'une description spatiale du réseau hydrographique.

3.1.1 Préambule : notions de système d'information géographique

La section mobilisera un vocabulaire propre au langage des SIG, raison pour laquelle il nous semble judicieux de commencer par une définition des principaux objets manipulés : les images rasters et les couches vectorielles.

Une image raster (voir Figure 3.2) n'est ni plus ni moins qu'une image numérique (comme pourrait l'être une photographie affichée sur un écran numérique) mais avec des propriétés

différentes :

- propriétés géospatiales : comme toute image numérique, les pixels ont une résolution et le nombre des pixels multiplié par la résolution de chaque pixel donne la dimension de l'image en hauteur et largeur. La particularité d'une image raster est de comporter des métadonnées (qui ne sont pas affichées mais lues par la machine à l'ouverture au travers d'un encodage spécifique et normalisé par le consortium OSGEO) qui convertissent la résolution des pixels et la dimension de l'image dans un système géoréférencé. Une métadonnée essentielle est ainsi le système de référence spatiale (SRS)
- propriétés des valeurs de pixels : contrairement aux images de type "photographie" dont les pixels portent une information de type colorimétrique (typiquement un encodage RGB), les pixels des images raster peuvent porter plusieurs types de valeurs possibles (dont l'encodage est lui aussi défini par les standards OSGEO). Par exemple, la librairie GDAL, qui s'appuie sur le standard OSGEO reconnaît trois formats de données possibles pour les pixels d'une image raster : le bit, le nombre entier et le nombre réel (flottant). Une propriété importante des images raster est qu'elles peuvent avoir plusieurs "bandes", c'est-à-dire différents jeux de données, à raison d'un par bande. Typiquement, si l'on souhaite ranger la moyenne mensuelle des débits sur une chronique, on pourra créer un raster à 12 bandes où chaque bande conserve les valeurs de débit d'un unique mois.

Ainsi lorsque nous parlons de "pixel" dans une image raster, nous faisons indistinctement référence à ces deux propriétés, position spatiale et valeur, sauf mention spécifique.

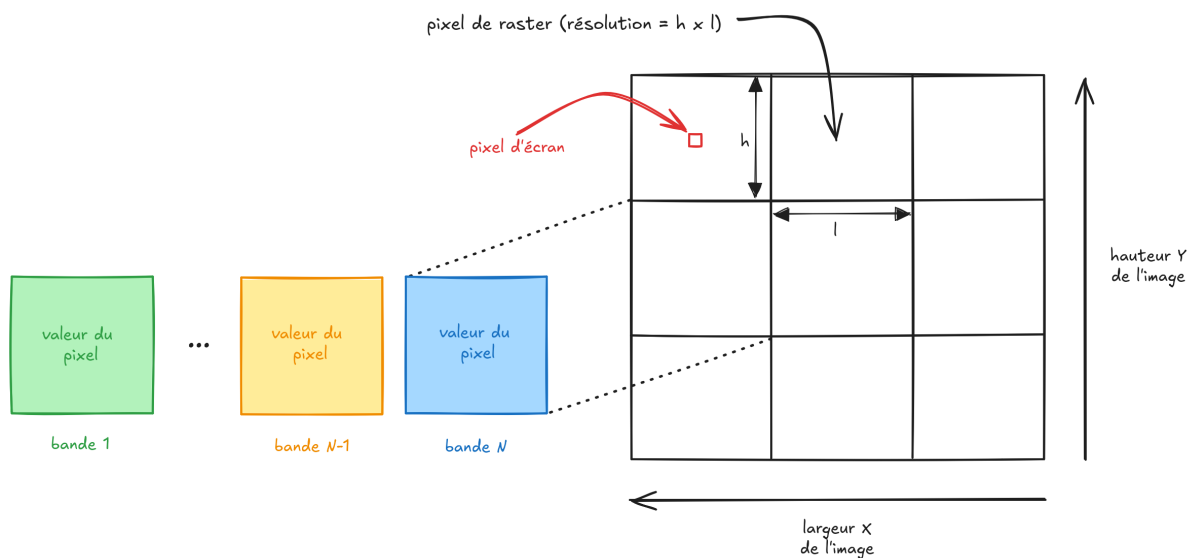


Figure 3.2 – Description d'une image raster

Quant à une couche vectorielle (voir Figure 3.3), la notion de résolution disparaît car les objets vectoriels sont des objets continus, contrairement aux pixels qui sont des objets discrets. Par exemple, il est impossible de représenter dans un espace géoréférencé une ligne avec une image raster : même en disposant de façon contiguë des pixels en ligne, il s'agira toujours d'une surface car les pixels sont toujours définis par une résolution, qui ne peut pas être nulle comme dans le cas d'une ligne au sens mathématique du terme. Un objet géomatique vectoriel est typiquement généré à partir des coordonnées de ses sommets, qui sont liés entre eux pour l'affichage graphique à travers un langage informatique approprié. Un encodage courant pour la géométrie des objets vectoriels est le Well-Known Text (WKT). Dans cet encodage, un point est défini par deux coordonnées $\langle x,y \rangle$, on note typiquement `Point(x,y)`. Un polygone est défini par les coordonnées de ses sommets, par exemple pour un triangle, on note `Polygon((x1,y1),(x2,y2),(x3,y3),(x1,y1))`. L'intérêt des objets vectoriels, au-delà de leur caractère spatialement continu, est qu'ils peuvent être associés à des attributs. Une couche vectorielle est à se figurer sous la forme d'un tableau de données où chaque ligne correspond à une entité spatiale et chaque colonne à un attribut, dont l'un de ces attributs est la géométrie de l'entité. Cette colonne "géométrie" est interprétée par un lecteur approprié à l'ouverture du fichier au regard des métadonnées qui comportent notamment un SRS. Le type de données possibles pour les attributs d'une couche vectorielle est bien plus important que pour les pixels de raster. Les attributs peuvent prendre n'importe quel format d'objet Python par exemple, et notamment la chaîne de caractères.

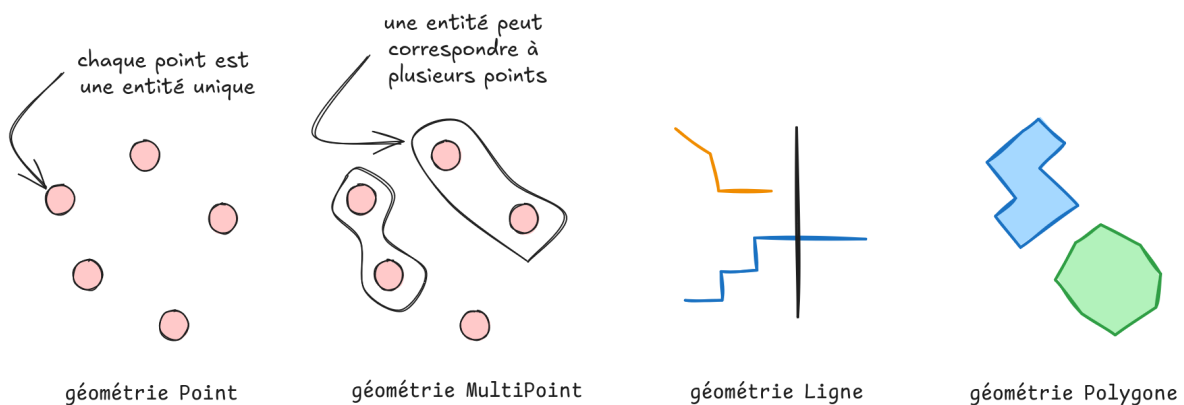


Figure 3.3 – Description d'une couche vectorielle (ici, 4 types de géométrie sont représentés)

3.1.2 Intérêt du modèle hydrographique élémentaire

L'intérêt de définir spatialement le réseau hydrographique réside dans la possibilité de générer le bassin versant associé à n'importe quel point positionné arbitrairement sur le réseau, qui pourront par exemple être l'emplacement d'un ensemble de stations hydrométriques en cours d'eau. Pourquoi alors ne pas utiliser les jeux de données déjà produits sur le territoire français, comme la BD TOPO ou la BD Carthage qui, toutes deux, proposent une représentation du réseau hydrographique au format vectoriel sur le territoire national? C'est précisément qu'il n'est pas possible à partir de ces deux produits de reconstituer le bassin versant à partir d'un point quelconque du réseau, et ceci pour au moins une raison principale : l'information sur le sens d'écoulement est sinon absente (BD TOPO) du moins très difficilement exploitable (BD Carthage)¹.

Une autre raison, qui relève davantage de la mise en œuvre numérique des traitements, est que le format vectoriel dans lequel ces produits sont disponibles rend leur traitement nettement plus compliqué que s'ils étaient disponibles au format raster, et la rasterisation de ces produits ne se fait pas sans bruit numérique (perte d'information ou ajout d'information par le traitement lui-même, ce qui pose ultérieurement des problèmes de continuité hydraulique).

De même, pourquoi ne pas utiliser les produits plus élaborés, comme les aires hydrographiques de la BD Carthage qui propose une délimitation des bassins versants à différents niveaux de confluence? La raison en est que ce produit propose une délimitation figée des bassins versants, qui ne correspond pas au besoin que nous aurons dans les parties 3 et 4 de délimiter dynamiquement des bassins versants à partir de points arbitraires. Les bassins versants de la BD Carthage sont en fait les bassins versants "naturels", du plus haut niveau (exutoire dans la mer ou l'océan) jusqu'à des niveaux très fins (premiers niveaux de confluence en tête de bassin). Les bassins versants qui nous intéressent ne sont pas "naturels" dans le sens où les exutoires ne correspondent pas à des nœuds hydrographiques "naturels" (confluence, affluence, delta) mais à des points arbitrairement placés. Ils ne sont pas non plus naturels dans le sens où ils sont "tronqués", et ceci de deux façons : d'une part, pour les bassins versants en bordure de zone d'étude, ceux-ci sont tronqués par la limite propre de l'étendue de la zone d'étude ; d'autre part, lorsqu'un bassin versant se superpose à un deuxième (par exemple dans

1. Parmi les attributs associés à la couche des tronçons, figure l'ordre amont/aval des tronçons. Les algorithmes des principaux fournisseurs de traitements SIG qui effectuent la délimitation des bassins versants à partir du sens d'écoulement ne fonctionnent pas avec un sens d'écoulement défini par un ordre amont/aval mais par une "direction de flux", que nous aurons l'occasion de présenter plus loin

le cas d'une station amont et d'une station aval), le bassin versant à l'aval est tronqué là où commence le bassin versant amont, bien qu'hydrographiquement le second fasse partie du premier.

Pour toutes ces raisons, il nous a ainsi été nécessaire de reconstituer à la main un réseau hydrographique qui permette une telle délimitation flexible des bassins versants.

La partie 2 utilisera le MHE pour ses sorties (les bassins versants délimités dynamiquement à partir de la position des stations de mesure en cours d'eau). La définition des bassins versants associés aux stations de mesure du territoire répond à deux objectifs au moins : un objectif de visualisation des données d'une part (en particulier les données élaborées issues de l'analyse des tendances de long terme) qui ne sont sinon représentées que par des points ; et d'autre part, un objectif d'analyse spatiale en permettant d'identifier les entités spatialisées influençant l'écoulement à l'exutoire (barrages, occupation des sols, stations d'épuration, etc).

La partie 3 montrera que l'utilité du MHE réside non seulement dans la sortie finale du modèle mais aussi dans les sorties intermédiaires (notamment la charge accumulée en tout point du réseau) qui portent une information hydraulique indispensable pour le développement du modèle MRH.

3.1.3 Méthode numérique pour la construction du MHE

La procédure numérique employée dans ce travail pour la reconstitution du réseau hydrographique n'est pas propre à ce travail de thèse mais s'appuie sur des développements algorithmiques déjà anciens, mais toujours en usage dans les logiciels SIG, remontant aux années 1980 aux États-Unis, époque et lieu où se structurent d'ailleurs les sciences de l'information géographique (cf. les travaux fondateurs de (TOMLIN et al. 1979) d'où émaneront les premiers logiciels SIG pour le traitement des images de raster). Avant les années 1980, les premiers travaux d'hydrologie numérique s'intéressent notamment à des procédures systématisées d'ordonnement des tronçons hydrographiques. C'est notamment le cas de Strahler dans les années 1950 qui développe une procédure de classement des tronçons par ordre de connectivité (STRAHLER 1957). Strahler, entre autres hydrologues, introduit également les éléments de vocabulaire spécifiques à l'hydrologie numérique, en s'appuyant sur le vocabulaire mathématique des graphes. Un réseau hydrographique numérique est défini par trois éléments constitutifs :

- un exutoire (*outlet*) : le point le plus à l'aval du réseau, ou racine du réseau
- un ensemble de sources : ensemble des points les plus à l'amont du réseau
- un ensemble de nœuds ou jonctions : ensemble des points définis par l'intersection entre deux segments amont qui, après ce point, forme un unique segment aval

La procédure décrite par Strahler ouvre la voie à des implémentations logicielles ultérieures mais n'est pas d'emblée destinée au traitement d'un jeu de données spécifique. Elle ne permet pas en elle-même d'obtenir un tracé mais de discriminer dans un tracé les tronçons selon leur importance. Au tournant des années 1970-1980, des méthodes spécifiques au traitement des modèles numériques de terrain (MNT), qui sont des images raster, se développent autour des concepts de *direction de flux* et d'*accumulation de flux* et permettent d'obtenir les premiers tracés hydrographiques (TARBOTON et al. 1991). La méthode qui fait date est celle décrite par (O'CALLAGHAN et al. 1984), s'articulant en trois étapes de traitement successives à partir d'un MNT :

- élimination des dépressions (i.e. trous topographiques où le ruissellement est piégé à moins que la hauteur d'eau ruisselée atteignent la hauteur de la dépression) et calcul de la matrice de direction de flux
- calcul de la matrice d'accumulation à partir d'une matrice de charge et de la matrice de direction de flux
- définition du réseau hydrographique à partir d'un seuil minimum d'accumulation

Nous allons revenir sur chacune de ces étapes en détail par la suite. Remarquons pour le moment que la procédure initiale prévoit en dernière étape l'utilisation d'un seuil minimum d'accumulation. Il s'agit en fait d'un seuil à partir duquel on peut considérer que la charge accumulée par un point du réseau (i.e. cellule de la matrice ou pixel de l'image raster) est suffisante pour correspondre dans la réalité à un écoulement permanent. Ce n'est en effet pas parce que l'eau ruisselle dans un sens donné sur une surface que de l'eau ruisselle en permanence sur cette surface. Le nombre de Strahler a pu également être implémenté en substitution des étapes 2 et 3 : il s'agit alors de déterminer le nombre de branches minimum dans le réseau pour lequel on considère qu'il y a écoulement permanent, autrement dit jusqu'à quel niveau on peut remonter l'arbre du réseau pour ne garder que les branches en charge permanente. Quelle qu'en soit la procédure de détermination numérique, ce seuil peut être physiquement décrit :

[...] *this break point [seuil minimum d'accumulation] represents a physical transition from channel erosion and transport mechanisms at large scale to hillslope erosion and transport mechanisms at small scale. The hillslope transport processes are diffusive in nature whereas the channel mechanisms rely on the power of concentrated flow in channels.* (TARBOTON et al. 1991)

Partant, (TARBOTON et al. 1991) proposent une procédure numérique pour la détermination de ce seuil physique ainsi décrit. Cette procédure est numériquement non-triviale et consiste à déterminer la charge accumulée minimum à partir de laquelle les propriétés hydrologiques du pixel associé ne respectent plus les lois de l'écoulement (définies par des travaux empiriques sur les relations entre caractéristiques morphologiques de l'écoulement et importance du tronçon dans le réseau, e.g. nombre de Strahler).

La procédure de conversion d'un MNT en un réseau hydrographique décrite par (O'CALLAGHAN et al. 1984) est notamment implémentée dans le logiciel sans interface graphique PCRaster développé depuis les années 2010 par l'Université d'Utrecht et l'Institut pour l'Environnement et la Soutenabilité du JRC Européen (KARSENBERG et al. 2010). Nous avons retenu ce logiciel en raison de son extension en Python (PCRaster Python) et de sa facilité d'utilisation en ligne de commande (PCRcalc). La génération du réseau hydrographique par accumulation de flux consiste à effectuer la chaîne de traitements suivante (voir Figure 3.4) :

- conversion du MNT en objet PCRaster (appelé *map*, qu'on traduira par *carte*)
- génération de la carte de direction de flux (opération qui intègre des paramètres pour la gestion des dépressions)
- génération de la carte de charge initiale
- génération de la carte d'accumulation de flux
- élimination des pixels du réseau dont l'accumulation est inférieure à un certain seuil

PCRaster embarque également des fonctions pour la détermination du nombre de Strahler, par laquelle l'opération de filtrage des pixels aurait également pu être conduite. Si cela n'aurait rien changé pour les sorties de notre modèle hydrographique (i.e. la délimitation des bassins versants), nous aurons besoin d'information sur la charge hydraulique du réseau dans la section

3 et, des deux méthodes, seule l'accumulation de flux permet de disposer d'une telle information.

Nous allons à présent détailler les procédures numériques de traitements.

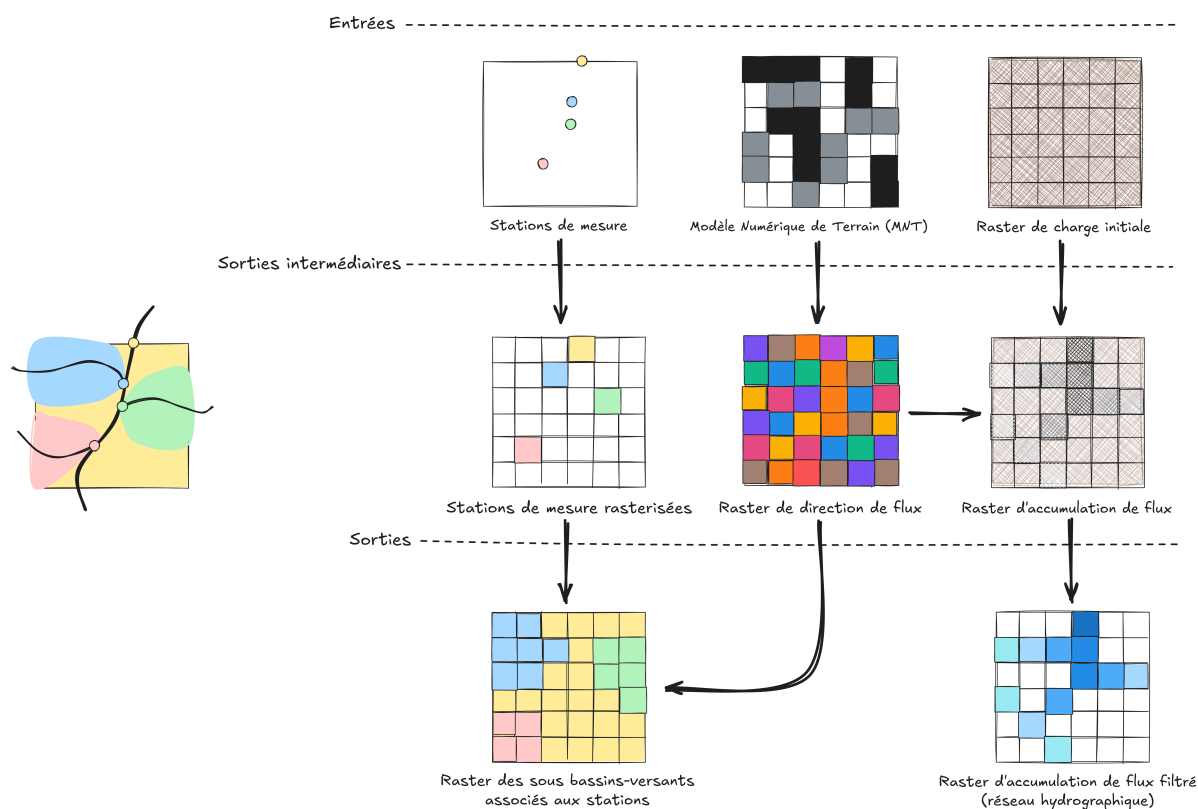


Figure 3.4 – Schéma séquentiel des traitements effectués par le MHE

3.1.3.1 Génération du réseau hydrographique par accumulation de flux

Nous décrivons ici l'implémentation numérique dans le langage de programme Python du MHE. Cette implémentation s'effectue en quatre étapes agencées dans un script dédié que reprendront les applications numériques de la MAESA et du MRH (voir Figure 3.1).

Étape 1 : Conversion du modèle numérique de terrain en objet PCRaster

On crée la fonction `convert_to_pcraster` qui met en œuvre la commande `gdalwarp` (émulée par la librairie `os` de Python) en forçant l'utilisation des propriétés `geoTransform` du MNT source pour la conversion au format PCRaster `.map` (sans quoi se posent des problèmes ultérieurs de consistance dans la projection spatiale et de définition des métadonnées des rasters intermédiaires générés). Cette étape étant purement numérique, nous ne l'avons pas représentée dans la Figure 3.4 mais elle se positionne entre le MNT et le raster de direction de flux.

Étape 2 : Génération du raster de direction de flux

Cette opération consiste à transformer chaque pixel du MNT en direction d'écoulement. Nous utilisons pour cela la fonction PCRaster `lddcreate`, en y ajoutant l'utilisation d'une couche de clone (le MNT) pour la définition propre des propriétés `geoTransform` du raster (notamment la résolution, la dimension et la projection). Le résultat obtenu est un raster de direction de flux (voir Figure 3.4).

La fonction `pcraster.lddcreate` est une fonction de la librairie PCRaster qui permet de créer un raster de direction de flux à partir d'un modèle numérique de terrain préalablement converti dans le format `.map` de PCRaster. Cette fonction appartient à la classe des opérations "fenêtre" (*window-operation*), c'est-à-dire que le pixel de la couche en sortie se voit attribuer une valeur qui est fonction d'un ensemble de pixels adjacents (ou "fenêtre de pixels") dans la couche en entrée. Plus particulièrement, `lddcreate` utilise une fenêtre de 9 pixels (le pixel central et ses 8 pixels adjacents) pour déterminer la valeur du pixel central en sortie. La valeur en sortie est encodée par un nombre entier allant de 1 à 9, encodage connu sous le nom d'*algorithme D8*, où chaque entier correspond à une direction d'écoulement. Il y a en effet 8 directions possibles et 1 position stationnaire pour le sens d'écoulement d'un pixel. Le sens d'écoulement est donné par la position du pixel adjacent par rapport au pixel principal avec qui il partage la pente la plus forte. Si toutes les pentes sont positives autour du pixel principal, le pixel principal est une dépression topographique. Si un ou plusieurs pixels autour du pixel principal ont une pente nulle, il s'agit d'une dépression répartie sur plusieurs pixels. Quatre paramètres de `lddcreate` permettent d'ajuster ce qui doit être considéré comme une dépression effective (en particulier pour identifier les dépressions formant des lacs ou des retenues, auquel cas il s'agit d'une région de pixels contigus de même altitude). Dans notre cas, nous ne souhaitons pas que les lacs et retenues soient confondus avec des exutoires topographiques (points les plus bas du réseau). En fixant une valeur très élevée pour ces 4 paramètres, typiquement 10^{31} (valeur qui, pour simplifier, correspond à la quantité de matière accumulée de l'amont et stockée dans un pixel), il devient très peu probable qu'un pixel du MNT satisfasse ces paramètres et qu'il soit considéré comme une dépression effective.

Étape 3 : Initialisation du transport de matière à l'origine

Un raster de charge initiale unitaire (voir Figure 3.4) est créé par la fonction `create_material` qui fonctionne en convertissant un vecteur `numpy` uniforme de valeur 1 en une couche raster dont les propriétés `geoTransform` sont empruntées à la couche de clone (le MNT). Ce raster est une couche de transport de matière à l'origine, qui peut être vu comme la traduction numérique d'une pluie unitaire sur l'ensemble de la zone étudiée : il s'agira ensuite de transporter cette pluie sur l'ensemble du bassin versant selon la direction de flux déterminée précédemment, puis de fixer un seuil d'accumulation à partir duquel une quantité suffisante de pluie transportée (i.e. de ruissellement) peut être considérée comme formant un cours d'eau actif. Notons que la valeur de transport à l'origine est déterminante pour certaines applications hydrologiques : si par exemple il s'agissait de déterminer la contribution relative d'un point du réseau à un autre, les valeurs de transports aux bornes (héritées d'un transport amont extérieur à l'emprise de l'étude) devraient être prises en compte. Ici, il s'agit simplement de déterminer quels pixels du territoire font partie d'un cours d'eau. L'approche est ici davantage booléenne que numérique : on ne cherche pas à savoir quelle est la valeur exacte de charge d'un pixel, mais à discriminer les pixels contributifs à l'écoulement permanent de ceux qui ne le sont pas. La partie 3 de ce chapitre reviendra sur cette question des conditions aux bornes car la valeur de charge des pixels dans le réseau sera au cœur du modèle proposé.

Étape 4 : Création de l'accumulation de flux

La fonction `create_accuflux` repose sur la fonction PCRaster `accuflux` qui prend en entrée le raster de direction de flux et le raster de matériel à l'origine. À partir de ces deux couches, les pixels de la couche en sortie prennent la valeur accumulée de leur matériel à l'origine (qui vaut 1) ajoutée à la valeur accumulée dont ils héritent à l'amont. La fonction `create_accuflux` force PCRaster à utiliser une couche de clone pour l'écriture des propriétés de `geoTransform` et spécifie le répertoire de destination pour la couche créée. À l'issue de cette première fonction, tous les pixels du territoire sont encore considérés comme contributifs à l'écoulement permanent de surface. L'opération suivante vient filtrer les pixels en ne conservant que ceux dont la valeur accumulée est supérieure à un certain seuil, nommé `accThreshold`. Au préalable, la couche d'accumulation de flux est reformatée depuis son format PCRaster `.map` vers le format universel `geoTiff .tif` afin d'assurer la robustesse de l'opération de filtrage. Cette opération est assurée par la commande `gdal_calc` et émulée par la librairie `os` de Python. Il s'agit simplement de ne retenir dans le raster en sortie que les pixels dont la valeur est supérieure à `accThreshold` et de remplacer les autres par une valeur nulle (`numpy.nan` plus précisément). À l'issue de cette

étape, le réseau hydrographique est modélisé.

3.1.3.2 Délimitation des sous bassins versants

La délimitation des bassins versants s'effectue en deux étapes, dont nous décrivons ici les implémentations dans le langage Python. Remarquons que nous définissons un bassin versant comme une zone d'apport à un point bas. En toute rigueur, un bassin versant correspond à une zone d'apport hydrographiquement close. Nous le verrons, une zone d'apport peut, dans le MHE, être non close si deux pixels d'un même réseau hydrographique se voit attribuer le rôle d'exutoire (par exemple, deux stations de mesure sur un même cours d'eau, l'une à l'amont et l'autre à l'aval).

Étape 1 : Rasterisation des stations de mesure

La fonction `join_points_to_pixels` réalise la première opération qui consiste à repositionner les coordonnées spatiales de chaque station de telle sorte à ce qu'elles se trouvent superposées à un pixel non nul du réseau. Cette opération est indispensable pour la suivante, qui ne fonctionne que si un pixel du réseau a été identifié comme exutoire (l'exutoire ne peut pas se trouver en dehors du réseau hydrographique!). Or la position spatiale des stations de mesure est déterminée de façon indépendante de la modélisation du réseau. Nous adoptons une jointure par plus proche voisin, c'est-à-dire que toute station se verra attribuer comme nouvelle coordonnée celle du pixel du réseau le plus proche de sa coordonnée actuelle. Si des fonctions existent déjà dans la librairie `geopandas` par exemple pour effectuer une jointure par plus proche voisin entre deux couches de points, nous n'avons pas eu connaissance de fonctions permettant d'effectuer une telle jointure entre une couche de points d'un côté et une couche de pixels de l'autre. La fonction `join_points_to_pixels` commence ainsi par convertir la couche de pixels en points, et ceci en itérant sur l'ensemble des pixels du raster fourni en entrée et, pour chacun, elle récupère le barycentre. La fonction `sjoin_nearest` de `geopandas` pourrait alors être appliquée pour attribuer aux coordonnées des stations la coordonnée du barycentre le plus proche dans l'ensemble des barycentres ainsi déterminés. Cependant, `sjoin_nearest` ne dispose pas d'un paramètre permettant de spécifier à l'algorithme d'attribuer à tout barycentre une unique station. Or l'attribution doit être bijective pour que le nombre d'exutoires corresponde au nombre de stations, sans quoi plusieurs stations peuvent se voir attribuer un même exutoire, faussant alors l'analyse spatiale. Cette non-bijectivité apparaît lorsque la résolution du raster

d'accumulation de flux est trop grande. Par exemple, si le MNT d'origine (dont découlent les propriétés `geoTransform` et donc la résolution du réseau hydrographique) est de résolution 1m, il est peu probable qu'un tel pixel se voit attribuer deux stations équidistantes. En revanche, pour une résolution de 90m, ce biais devient très probable. Notons que le recours à un MNT de résolution 90m peut se justifier par la réduction des temps de calcul, qui augmentent exponentiellement lorsque la résolution diminue.

On implémente donc dans la fonction `join_points_to_pixels` un algorithme "glouton" de jointure spatiale. Le caractère glouton de l'algorithme peut s'interpréter comme une logique du "premier arrivé, premier servi". On calcule pour cela une matrice des distances entre toute station et tout barycentre. Les stations sont ensuite triées dans un ordre aléatoire, puis la première de la liste se voit attribuer le barycentre le plus proche d'elle dans la matrice des distances. Ce barycentre est alors retiré de la liste des barycentres disponibles pour les autres stations. Et ainsi de suite pour le reste des stations.

La construction "gloutonne" de cet algorithme permet d'éviter les multiples associations entre un même pixel et plusieurs stations. En revanche, elle ne garantit pas qu'une station soit bien associée au pixel du réseau qui lui correspond dans la réalité. Par exemple, dans le cas où une station située sur un cours d'eau important se trouve plus proche du pixel d'un cours d'eau secondaire, elle sera associée à ce second cours d'eau. Une méthode de vérification aurait pu être ajoutée à la fonction, en partant du principe de corrélation entre rangs (typiquement ce que font les tests de Spearman) : si une station enregistre le $p^{\text{ième}}$ plus gros débit annuel moyen, par exemple, alors il est fortement probable que le pixel qui doit lui être associé enregistre la $p^{\text{ième}}$ plus grosse valeur de charge accumulée. Néanmoins, une telle méthode n'a pas été implémentée. La vérification peut alors s'effectuer manuellement dans un logiciel de SIG à interface graphique (e.g. QGIS), en superposant le raster d'accumulation de flux et la couche des stations projetées dans le réseau hydrographiques. Si une station enregistre un débit qui ne correspond vraisemblablement pas à la valeur d'accumulation à laquelle elle a été associée, alors la position du point est décalée en superposition d'un pixel proche et vraisemblable dans le raster d'accumulation de flux. Notons que cette correction manuelle est seulement menée pour les traitements relatifs à la section 3, où les valeurs de débits sont prises arbitrairement pour le mois de janvier. Pour les traitements de la section 2, une telle correction n'a pas été entreprise. Nous reviendrons donc plus en détail sur l'erreur associée à une mauvaise projection des stations dans la section 3.

Cette démarche de jointure en deux temps, automatisée puis manuelle, entre la position des stations et le réseau hydrographique modélisé est similaire à celle de (LEHNER 2012) où l'enjeu est de joindre les stations du GRDC (*Global Runoff Data Center*) avec le réseau hydrographique HydroSHEDS de résolution 15 arcsec. Les deux temps de la procédure sont :

1. Une jointure automatisée. Un rayon de recherche de 5 km autour de chaque station détermine les pixels à partir desquels sont délimités les bassins versants potentiels de chaque station. Leur surface est calculée et comparée à la valeur d'emprise du bassin amont de la station dans les données GDRC. Les pixels dont le bassin versant présente un écart de plus de 50% avec la valeur connue ne sont pas retenus. La distance euclidienne entre la station et les pixels dans le rayon de recherche sont également calculées. Le pixel retenu est celui dont la surface est la plus proche de la surface connue *et* qui est le moins distant de la station.
2. Une jointure manuelle : pour les stations n'ayant aucun pixel dont la surface du bassin versant est moins de 50% différente de la surface réelle, ou les stations qui n'ont pas de données de surface de bassin amont dans la base GDRC, ces stations sont attribuées manuellement à un pixel.

Cette démarche que le GDRC a implémentée est notamment reprise dans (BARBAROSSA et al. 2018) mais à partir du réseau hydrographique HydroSHEDS de résolution 30 arcmin.

Étape 2 : Délimitation des sous bassins versants

La fonction `create_subcatchments` permet de délimiter les sous bassins versants. Elle repose sur la fonction `subcatchments` de la librairie `PCRaster` qui prend en entrée une liste de coordonnées correspondant aux exutoires et une couche de direction de flux. La fonction `create_subcatchments` permet d'automatiser la conversion au format `PCRaster` de la couche vectorielle des stations projetées dans le réseau hydrographique modélisé (cf. opération précédente). Notons que la délimitation des bassins versants n'a pas besoin du raster d'accumulation de flux. Ce raster est en revanche indispensable pour l'opération précédente, car il permet de s'assurer que la position des stations est hydrographiquement cohérente. `PCRaster` dispose de deux fonctions de délimitation des bassins versants : une fonction `catchment` et une fonction `subcatchment` que nous avons préféré utiliser. La différence entre les deux tient au traitement des bassins versants superposés. La fonction `catchment` reconstitue en effet l'ensemble du bassin

versant associé à un exutoire, quand bien même celui-ci se superposerait à celui d'un autre exutoire à l'amont. La fonction `subcatchment` arrête la délimitation dès qu'il y a superposition : cette option nous a paru plus judicieuse dans un but d'analyse spatiale si l'on souhaite associer au comportement hydrologique d'une station (d'un exutoire) des éléments anthropiques présents sur sa surface contributive. La couche raster au format `.map` de PCRaster ainsi créée est ensuite converti en format vectoriel de géométrie `Polygon` grâce à la fonction `raster_to_polygons` qui fait un appel à la fonction `gdal_vectorize` et élimine les géométries inconsistantes générées par PCRaster et GDAL.

3.2 L'évaluation environnementale de l'intégrité des cours d'eau par les données de mesure

La formulation de l'indicateur sur l'eau de surface proposé par (ROCKSTROM et al. 2023) met l'accent sur un excès de variabilité du débit : au-delà d'une variation du débit mensuel moyen supérieure à 20% du débit mensuel moyen de référence (dit "non perturbé" par les activités humaines), la variation est considérée comme excessive. L'indicateur est fondamentalement un indicateur de variabilité, définie comme écart à une référence. L'utilisation de données de mesure pour le calcul d'un tel indicateur, c'est-à-dire des séries temporelles portant sur des chroniques de débits, entraîne nécessairement une confrontation entre la variabilité ainsi définie dans le cadre des frontières du système Terre, et la variabilité traditionnellement définie dans le champ de la statistique descriptive, que l'on pourrait qualifier de variabilité par rapport à une norme statistique (par exemple la moyenne, nous y reviendrons). Nous allons donc introduire cette section en présentant les définitions et formulations usuelles de la variabilité au sens statistique, c'est-à-dire les indicateurs spontanément utilisables pour quantifier la variabilité d'une série temporelle. Nous proposerons ensuite une reformulation de l'indicateur utilisé dans (ROCKSTROM et al. 2023), qui tire parti des définitions de la variabilité au sens statistique et reproduit "l'esprit" du cadre des frontières du système Terre en introduisant dans l'indicateur la référence à une norme. L'utilisation de données de mesure conduit ainsi à passer d'un indicateur basé sur une référence environnementale à un indicateur basé sur une norme statistique. Et c'est au fond la question de l'état de référence, au cœur de l'esprit du cadre des limites planétaires, qui est posée à nouveaux frais grâce à l'utilisation de données de mesure. Cette partie vise donc à proposer l'indicateur d'une MAESA basée sur des données de mesure.

3.2.1 Définir un indicateur de variabilité : d'une variabilité "de référence" à une variabilité "normale"

3.2.1.1 La variabilité statistique d'une série temporelle comme mesure de l'erreur

Le débit moyen d'un cours d'eau, lorsqu'il est obtenu à partir d'une série spatio-temporelle, est une donnée statistiquement agrégée, à la fois temporellement et spatialement, à partir d'observations réparties dans le temps et positionnées dans l'espace en différents points du cours d'eau. Ces observations constituent un échantillon du phénomène réel d'écoulement. Ceci induit des biais de construction statistique : c'est toute la question de l'écart entre la valeur *réelle* du débit moyen (moyen sur une fenêtre temporelle et moyen sur un tronçon) et sa valeur *estimée* par l'agrégat utilisé. Cet écart est largement documenté en statistique et quantifié par le calcul de l'*erreur standard*. Cette métrique renseigne sur l'effet lié à la taille de l'échantillon. La variabilité au sein d'un échantillon est mesurée par l'écart-type, tandis que la variabilité d'une même métrique entre échantillons est mesurée par l'erreur standard (ALTMAN et al. 2005). L'erreur standard vient donc quantifier l'erreur d'échantillonnage. En notant *sterr* l'erreur standard, on pose :

$$sterr = \frac{\sigma_X}{\sqrt{n}}$$

Où X désigne l'échantillon dans lequel sont rangées les n observations du phénomène. σ_X désigne alors l'écart-type empirique. Cette formule fait apparaître une relation quadratique inversement proportionnelle entre l'erreur standard et la taille de l'échantillon : par exemple, une augmentation de la taille de l'échantillon par 10 ne réduira l'erreur standard que par 3 environ (si l'écart-type est inchangé).

La variabilité à l'intérieur de X est quant à elle mesurée par l'écart-type empirique sur l'échantillon (i.e. normalisé par les $n - 1$ premières observations). La formule est la suivante :

$$\sigma_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Une métrique alternative à l'écart-type, souvent utilisée dans la littérature, est le coefficient de variation, qui est une valeur normalisée de l'écart-type par rapport à la moyenne sur le même échantillon, soit, pour cv qui désigne le coefficient de variation et \bar{X} la moyenne empirique :

$$cv = \frac{\sigma_X}{\bar{X}}$$

L'erreur standard se décline en différentes formulations selon le paramètre statistique à estimer (AHN et al. 2003). Pour ce qui est de la moyenne, l'erreur standard mesure la distance entre la moyenne théorique sur la population (inconnue) et la moyenne empirique sur un échantillon (connue). L'erreur standard est donc une mesure du biais au sens statistique, c'est-à-dire, si μ la moyenne théorique :

$$sterr = \mu - \bar{X}$$

En plus du biais, l'erreur standard permet de produire une autre mesure d'erreur, à savoir l'incertitude liée à la taille de l'échantillon, c'est-à-dire la probabilité que la moyenne théorique et la moyenne empirique se situe à une distance déterminée (*sterr*) l'une de l'autre. Pour déterminer cet intervalle de confiance, on fait l'hypothèse d'une distribution normale de *sterr* : si l'on comparait un grand nombre d'échantillons entre eux, la distribution des erreurs standard suivrait une gaussienne. On peut alors associer à un quantile de la distribution normale la valeur réelle associée. Typiquement, au niveau de confiance de 95% et avec U_X qui désigne l'incertitude de la moyenne pour l'échantillon X :

$$U_X = \frac{1.96 \cdot sterr}{\bar{X}}$$

3.2.1.2 Proposition d'un indicateur AESA basé sur des données de mesure : l'anomalie de variabilité

En s'appuyant sur les métriques pour la quantification la variabilité propres au champ de la statistique, et en reprenant l'esprit de l'indicateur défini par (ROCKSTROM et al. 2023), nous proposons de formuler un indicateur de variabilité maximum du débit par rapport à une norme, c'est-à-dire un indicateur d'anomalie de variabilité.

Le principe de l'indicateur défini par (ROCKSTROM et al. 2023) est de vérifier que le débit d'un cours d'eau reste dans une plage de variabilité raisonnable : le débit mensuel moyen ne doit pas être inférieur à 80% ni supérieur à 120% de sa référence. La variabilité jugée raisonnable est ainsi de +/-20%. Le débit de référence est calculé à partir d'un modèle hydrologique global (le WBM) qui simule les écoulements de surface en l'absence de certaines pressions anthropiques (pas de barrages ni d'irrigation agricole). Avec des données de mesure sur site, il n'est pas possible de calculer un tel débit "non perturbé" pour établir le débit de référence. En revanche, pour une période de plusieurs années, on peut regarder les variations interannuelles du débit

pour un même mois et les comparer à la variabilité du débit sur le mois. La variabilité du débit d'un même mois sur plusieurs années, qui est une norme statistique, est alors utilisée comme valeur de référence. L'indicateur que nous proposons mesure ainsi l'écart entre la variabilité enregistrée au cours d'un mois donné et la variabilité normale (ou moyenne) de ce même mois sur le temps long, par exemple 2000-2020, pour reprendre la période utilisée dans (ROCKSTROM et al. 2023) : il s'agit d'une anomalie de variabilité, et l'anomalie maximum autorisée est elle aussi fixée au seuil de +/-20%.

Pour formaliser mathématiquement cet indicateur d'anomalie, on adopte les notations suivantes :

- i désigne une station de mesure instrumentée sur un cours d'eau
- t désigne le mois d'une année donnée (e.g. le mois de janvier 2000)
- k désigne une journée donnée à l'intérieur du mois t
- Γ (gamma) désigne un mois générique sans année spécifique (e.g. un mois de janvier quelconque)
- K_t désigne le nombre de jours k dans le mois t : $K_t = \text{card}(k \in t)$
- K_Γ désigne le nombre de jours k de tous les mois t correspondant au mois générique Γ :

$$K_\Gamma = \sum_{t \in \Gamma} \text{card}(k \in t)$$

Distinguons d'abord deux métriques : le débit mensuel moyen et le débit mensuel moyen interannuel.

Le débit mensuel moyen (par rapport à t) :

$$\mu_{i,t} = \frac{\sum_{k \in t} \mu_{i,k}}{K_t}$$

Où $\mu_{i,k}$ est le débit moyen journalier du jour k à la station i .

Et le débit mensuel moyen interannuel (par rapport à Γ) :

$$\mu_{i,\Gamma} = \frac{\sum_{t \in \Gamma} \sum_{k \in t} \mu_{i,k}}{K_\Gamma}$$

De même, on définit les écarts-types associés à chaque débit :

$$\sigma_{i,t} = \sqrt{\frac{1}{K_t - 1} \sum_{k \in K_t} (\mu_{i,k} - \overline{\mu_{i,k}})^2}$$

Et :

$$\sigma_{i,\Gamma} = \sqrt{\frac{1}{K_\Gamma - 1} \sum_{k \in K_\Gamma} (\mu_{i,k} - \overline{\mu_{i,k}})^2}$$

À partir de ces deux métriques de débit (mensuel et mensuel-interannuel), nous allons définir trois métriques pour la variabilité du débit, toutes trois exprimée comme un coefficient de variation cv , c'est-à-dire le rapport entre l'écart-type et la moyenne :

- $cv_{i,t}$: la variabilité mensuelle du mois t , qui désigne la variabilité qu'il y a à l'intérieur d'un même mois (mois t), c'est-à-dire entre les débits moyens journaliers
- $\overline{cv_{i,t \in \Gamma}}$: la variabilité mensuelle moyenne d'un même mois, qui désigne la moyenne des variabilités mensuelles d'un même mois (mois Γ)
- $cv_{i,\Gamma}$: la variabilité interannuelle d'un même mois (mois Γ), qui désigne la variabilité qu'il y a sur plusieurs années entre les débits mensuels moyens d'un même mois.

Plus précisément, les trois métriques sont calculées de la façon suivante :

$$cv_{i,t} = \frac{\sigma_{i,t}}{\mu_{i,t}}$$

$$\overline{cv_{i,t \in \Gamma}} = \frac{\sum_{t \in \Gamma} cv_{i,t}}{\text{card}(t \in \Gamma)}$$

$$cv_{i,\Gamma} = \frac{\sigma_{i,\Gamma}}{\mu_{i,\Gamma}}$$

On peut alors définir l'indicateur d'anomalie de variabilité. Il s'agit d'un taux de variation entre le coefficient de variation du mois t et un coefficient de variation de référence, calculé comme le coefficient de variation moyen interannuel du mois t .

$$\Delta cv_{i,t \in \Gamma} = \frac{cv_{i,t} - \overline{cv_{i,t \in \Gamma}}}{\overline{cv_{i,t \in \Gamma}}}$$

Si $\overline{cv_{i,t \in \Gamma}}$ est supérieur à 0.2 ou inférieur à -0.2, on peut considérer que la station i enregistre une anomalie de variabilité excessive et qu'il y a un risque pour l'intégrité du cours d'eau correspondant. Cet indicateur permet d'alerter sur les trop fortes variabilités interannuelles, qui sont susceptibles de porter préjudice à la biodiversité aquatique.

3.2.2 Le potentiel analytique de la variabilité statistique

Au-delà de l'indicateur que nous proposons, la variabilité apporte en soi des éléments d'analyse et de compréhension du phénomène hydrologique étudié. En effet, si nous avons vu dans la section 3.1.1 que la variabilité est influencée par la construction de l'agrégat lui-même (la taille de l'échantillon), elle est également influencée par des facteurs "réels" que sont les caractéristiques physiques des bassins versants. Par exemple, les caractéristiques topographiques, géologiques et pédologiques influencent le ruissellement, de même que les caractéristiques anthropiques (les activités humaines étant responsables par exemple de l'altération/artificialisation des paysages et des prélèvements en eau, voir Figure 3.5), et de même bien sûr que les caractéristiques climatiques (intempéries, aléas climatiques). Il est donc envisageable de considérer la variabilité du débit comme un indicateur d'altération du débit, et donc d'altération de l'intégrité des cours d'eau.

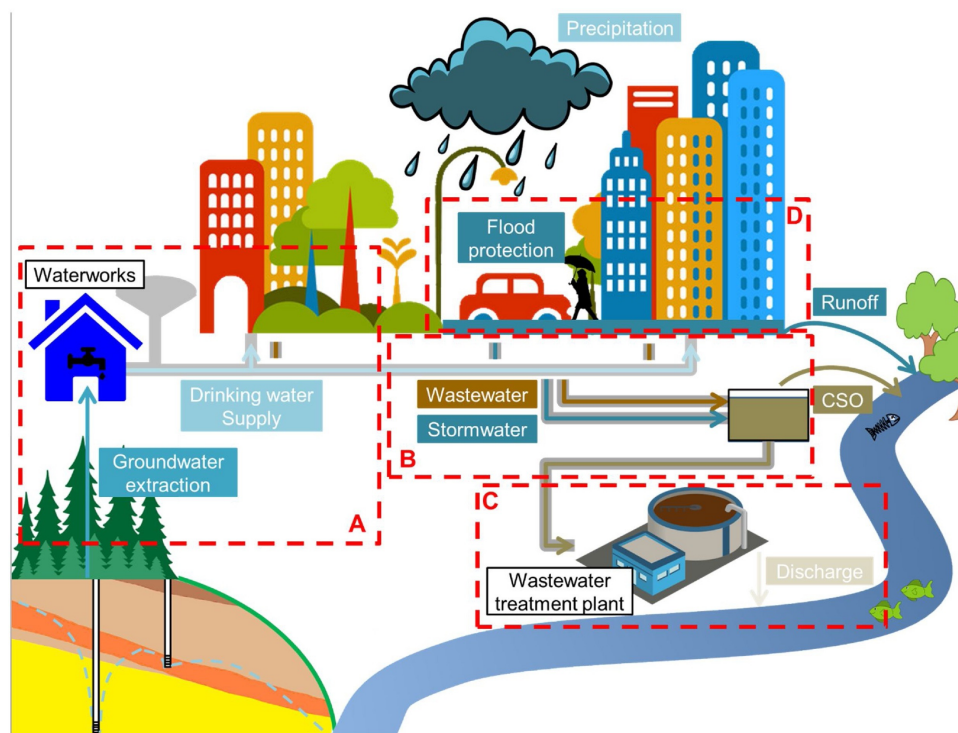


Figure 3.5 – Perturbation du cycle de l'eau en milieu urbain par les activités humaines (SØRUP et al. 2020)

3.2.2.1 Cadre analytique

La typologie des incertitudes mixte qualitative-quantitative proposée par (PASTOR et al. 2020) nous semble constituer un cadre analytique pertinent pour décomposer et analyser la variabilité du module empirique. Dans cette typologie, Amandine Pastor et ses collègues combinent une entrée qualitative en qualifiant les sources d'incertitudes associées à chaque type d'incertitude, et une entrée quantitative en présentant les méthodes de calcul généralement utilisées pour la quantification de chaque type d'incertitude. Cette typologie est retenue pour au moins trois raisons :

- Les auteurs ont construit leur typologie par une revue de la littérature sur les modèles intégrés d'évaluation des impacts environnementaux (*integrated impact assessment models*). Ils ont ainsi examiné la récurrence des outils quantitatifs et des cas d'application pour le traitement des incertitudes dans un champ scientifique proche de celui de l'évaluation absolue de la soutenabilité environnementale.
- Leur typologie mixte quantitative-qualitative correspond bien à notre double approche de la variabilité : la variabilité comme mesure statistique (quantitatif) et la variabilité comme outil d'analyse hydrologique (qualitatif). Cette approche complète des incertitudes, par la définition et par les outils de calcul, est particulièrement intéressante parce que rarement présente dans le champ de l'évaluation absolue de la soutenabilité environnementale. À cet égard, Samuel Zipper, un auteur majeur du champ ayant notamment contribué à la limite planétaire sur l'eau, propose dans (ZIPPER et al. 2020) une typologie (du moins un passage en revue) des incertitudes que l'on peut rencontrer lors d'un travail de descente d'échelle de la limite eau : les incertitudes qui relèvent des rétroactions complexes et non-linéaires du système Terre ; celles qui ont trait aux interactions entre variables de contrôle et variables de réponses ; celles qui sont inhérentes aux données de mesure et aux modèles utilisés pour quantifier les variables de contrôle ; celles enfin qui se rapportent aux approches retenues pour agréger et désagréger les variables de contrôle aux différentes échelles. Cette typologie, bien qu'elle identifie différentes sources d'incertitudes, n'offre pas de cadre d'analyse de ce qui fait précisément l'incertitude (ou la variabilité).
- Enfin, l'autrice principale de l'article, Amandine Pastor, était l'une des contributrices principales à la version 2015 de l'indicateur de la limite planétaire sur l'eau (STEFFEN et al. 2015b), légitimant encore davantage l'utilisation de ses travaux dans le champ de l'évaluation absolue.

Leur typologie répartit les incertitudes en trois familles : les incertitudes techniques, méthodologiques et épistémiques. Plus précisément :

- les incertitudes techniques relèvent de la qualité de la donnée d'entrée et sont souvent exprimées comme une valeur associée à une plage d'erreur
- les incertitudes méthodologiques relèvent de l'insuffisance des données d'entrée et sont exprimées par un intervalle de confiance
- les incertitudes épistémiques se décomposent en trois groupes :
 - les incertitudes structurelles, qui relèvent de la modélisation (équations, processus) et sont traitées par des tests de comportements sur différents jeux de fonctions
 - les incertitudes paramétriques, qui relèvent des données d'entrées et des paramètres de modèles, et qui sont traitées par des tests de sensibilité sur les variables d'entrées
 - les incertitudes numériques (*computational*) qui relèvent des capacités numériques et informatiques

Pour chaque type d'incertitude, la table 3.1 donne la traduction que nous en faisons pour le problème qui nous occupe ici : l'analyse de la variabilité des débits (qui est bien une mesure de l'erreur, ou incertitude).

Type d'incertitude	Traduction au cas d'étude	Procédure de traitement
Technique	Erreur de jaugeage des stations de mesure (conversion hauteurs d'eau en débits surtout en périodes de crue)	(non traité) Avis d'experts
Méthodologique	Le débit mensuel moyen agrégé spatialement et temporellement est une estimation du module réel sur la population	(non traité) Évaluation de l'erreur standard à 3 niveaux d'agrégation différents
	Choix arbitraire de la période 2000-2020 (ROCKSTROM et al. 2023)	(non traité) Analyse de sensibilité sur le coefficient de variation annuel moyen selon la longueur de la chronique en entrée
Epistémique		
Structurelle	Variabilité inter-annuelle stochastique du climat	(non traité) Matrice de corrélation pour chaque bassin versant entre le débit mensuel moyen et les précipitations mensuelles moyennes
	Variabilité causée par l'anthropisation	(traité) Matrice de corrélation pour chaque bassin versant entre le coefficient de variation annuel moyen et certains éléments d'anthropisation
	Tendance climatique de long terme due au forçage radiatif	(traité) Régression de Mann-Kendall sur l'évolution du débit moyen à long terme et test d'hypothèse sur une baisse tendancielle
Paramétrique	Ce type d'incertitude tient aux paramètres d'entrée d'un modèle : ici la donnée que nous produisons (débit mensuel) n'est pas obtenue de façon paramétrique (moyenne empirique et écart-type empirique sont non-paramétriques)	Non concerné
Numérique	Ce type d'incertitude tient à l'utilisation de modèles de simulation numérique, qui ne sont pas utilisés ici	Non concerné

TABLE 3.1 – Protocoles de traitement des différents types d'erreur

Le traitement des incertitudes structurelles liées à l'anthropisation et au forçage climatique nous semble d'un intérêt analytique supérieur aux autres types d'incertitudes, qui relèvent davantage de biais liés à la construction de l'agrégat statistique lui-même que de pressions environnementales exercées sur les cours d'eau.

3.2.2.2 Déceler l'impact de l'anthropisation du paysage

Nous avons identifié l'impact de l'anthropisation des paysages sur la variabilité du débit comme une erreur de type structurelle, autrement dit comme une composante structurante de la variabilité qui est une mesure de l'erreur. Nous allons ici décrire la procédure de quantification de cet impact, par matrice de corrélation entre la variabilité du débit sur un bassin versant et les éléments d'anthropisation présents sur ce bassin versant.

La matrice de corrélation est une forme standardisée de la matrice de covariance. Toute entrée de la matrice est constituée du coefficient de corrélation de Pearson entre plusieurs variables aléatoires deux à deux. L'intérêt de la standardisation (normalisation par rapport au produit des écart-types des deux variables aléatoires) est de ramener les coefficients dans un intervalle allant de -1 à 1 .

Dans notre cas, nous souhaitons identifier parmi certains éléments anthropiques du paysage ceux qui sont les plus corrélés à la variabilité des débits. Cela permet moins de quantifier la contribution de chaque source de variabilité à la variabilité totale que de déterminer les contributeurs les plus significatifs à la variabilité.

Nous avons pour cela identifié un ensemble d'éléments constitutifs de l'aménagement du territoire qui nous ont semblé être susceptibles d'avoir un impact significatif sur la variabilité. Nous donnons également le nom des variables associés à chaque élément pour la formalisation ultérieure :

- les éléments identifiés dans (ROCKSTROM et al. 2023) : les barrages (*dam*) et les lacs de retenue associés (*res*), ainsi que les terres agricoles irriguées (la classe correspondant dans les différents modes d'occupation du sol occ_i)
- des éléments supplémentaires : les stations d'épuration (*wwp*, dont le débit en sortie d'ouvrage peut représenter une part significative du débit pour des cours d'eau à faible débit, notamment en période d'étiage), les divers obstacles à l'écoulement (les seuils et les écluses, respectivement *weir* et *lock*), l'ensemble des modes d'occupation du sol (occ_i , autre que l'agriculture et notamment les sols urbanisés, dont nous faisons l'hypothèse qu'en modifiant les propriétés hydrauliques des bassins versants, notamment le temps de concentration, ils sont susceptibles d'avoir un impact significatif sur la variabilité).

Les bassins versants délimités par le MHE à partir des stations de mesure peuvent avoir des superficies et des débits qui varient grandement de l'un à l'autre. Si nous cherchions la corrélation entre l'écart-type du débit et l'occurrence des éléments anthropiques, nous biaiserions la corrélation en faveur des bassins versants enregistrant les plus gros débits et donc les plus gros écarts-types. Le coefficient de variation apparaît ainsi plus approprié puisqu'il est une mesure de l'écart-type normalisé par le débit moyen : les petits bassins versants sont ainsi mis à égalité avec les gros bassins versants. On reprendra ainsi les 12 coefficients de variation $cv_{i,t}$ calculés mensuellement pour chaque station, puis on effectuera la moyenne pour obtenir un coefficient de variation annuel moyen pour chaque station, que l'on peut noter \overline{cv}_i et qui se calcule comme :

$$\overline{cv}_i = \frac{\sum_{t=1}^{12} cv_{i,t}}{12}$$

On définit ainsi, en vue de la matrice corrélation, l'échantillon $(\overline{cv}_i)_{1 \leq i \leq I}$ des coefficients de variation annuels moyens de toutes les stations de la zone d'étude.

De même, l'occurrence des éléments anthropiques doit être rapportée à la surface des bassins versants pour assurer la cohérence avec ce que nous mentionnions avant. L'occurrence surfacique de ces éléments consiste simplement à les dénombrer sur chaque bassin versant et à rapporter cette valeur à la surface du bassin versant exprimée en mètres carrés. Ainsi par exemple, wwp_i est le nombre de stations d'épuration du bassin versant de la station i rapporté à la surface du bassin versant i , et ainsi on définit l'échantillon $(wwp_i)_{1 \leq i \leq I}$ et on utilise la même notation pour le reste des éléments anthropiques.

Enfin, pour définir l'échantillon associé à la variable aléatoire de l'occupation du sol, on commence par identifier toutes les classes d'occupation du sol en présence sur le territoire parmi les 44 qui constituent la nomenclature Corine Land Cover (CLC). Pour chaque bassin versant, on compte le nombre de pixels de chaque classe d'occupation du sol ainsi identifiée, qu'on rapporte au nombre total de pixels qui sont superposés au bassin versant. On note $occ_{i,c}$ la part de la classe d'occupation du sol c , dans le bassin versant i . Et on note p_j la classe d'occupation du sol du pixel j . On a alors :

$$occ_{i,c} = \frac{card(j \in i | p_j = c)}{card(j \in i)}$$

Où la fonction $card$ compte le nombre d'éléments qui correspondent à l'argument. En notant Γ le nombre de classes d'occupation du sol répertoriées sur le territoire, on définit les $I \times \Gamma$ échantillons $(occ_{i,c})$.

On a ainsi l'ensemble d'échantillons suivants à tester, classés en colonne par variable aléatoire :

$$\begin{pmatrix} \overline{cv}_1 & wwp_1 & res_1 & dam_1 & weir_1 & lock_1 & occ_{1,1} & \dots & occ_{1,c} & \dots & occ_{1,\Gamma} \\ \vdots & & & & & & & & & & \\ \overline{cv}_i & wwp_i & res_i & dam_i & weir_i & lock_i & occ_{i,1} & \dots & occ_{i,c} & \dots & occ_{i,\Gamma} \\ \vdots & & & & & & & & & & \\ \overline{cv}_I & wwp_I & res_I & dam_I & weir_I & lock_I & occ_{I,1} & \dots & occ_{I,c} & \dots & occ_{I,\Gamma} \end{pmatrix}$$

On peut alors formaliser la matrice de corrélation qui en sera dérivée. Pour simplifier l'écriture, X_j désigne une variable aléatoire parmi n (une colonne de la matrice ci-dessus parmi les n colonnes). La matrice de corrélation C s'écrit :

$$C := (c_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n} = \frac{cov(X_i, X_j)}{\sigma_{X_i} \sigma_{X_j}}$$

Cette matrice nous donne le coefficient de corrélation standardisé de toutes les paires de variables aléatoires. Ce qui nous intéresse en réalité, c'est seulement la corrélation entre la variabilité des débits (i.e. la variable aléatoire $(\overline{cv}_i)_{1 \leq i \leq I}$) et les éléments anthropiques du paysage (le reste des variables aléatoires). Si les (\overline{cv}_i) correspondent à la variable aléatoire X_1 , alors on ne s'intéressera qu'à la première ligne de C .

On effectue alors un test de p-value à deux queues qui permet de tester l'hypothèse nulle "le coefficient n'est pas significativement différent de zéro" (i.e. "il n'y a pas de corrélation entre les deux variables").

Remarquons que la délimitation des bassins versants issue du MHE peut entraîner un biais dans l'analyse. Le choix qui a été fait dans le MHE est en effet de délimiter les sous bassins versants plutôt que les bassins versants. Ce choix a son avantage et son inconvénient : s'il permet de différencier l'emprise sur laquelle sont observées certaines caractéristiques hydrologiques (débit moyen, variabilité moyenne), ces caractéristiques ne sont pas indépendantes des caractéristiques des potentiels bassins versants amont superposés. Par exemple, pour un cas à deux stations appartenant au même bassin versant global, si un barrage est positionné sur le bassin versant amont, il impacte nécessairement les caractéristiques du bassin versant amont et du bassin versant aval, mais plus la source de pression est lointaine, plus son impact est diffus. Cette diffusion de l'impact n'est pas prise en compte dans notre analyse de la corrélation.

3.2.2.3 Déceler l'impact de la tendance climatique de long terme

Une cause de la variabilité du module mensuel moyen est l'effet du forçage climatique à long terme, qui a pour conséquence sur le cycle de l'eau une augmentation de la quantité d'eau dans l'atmosphère (et donc une diminution proportionnelle de la quantité disponible dans le stock d'eau de surface) et, d'autre part, une augmentation des extrêmes climatiques (IPCC 2023). En analysant une série chronologique de longue durée, on peut identifier plusieurs composantes : une composante saisonnière ou cyclique, une composante tendance et une composante aléatoire (GILBERT 1987). La statistique de Mann-Kendall (KENDALL 1948) est largement utilisée pour l'évaluation de la composante de long terme en hydrologie, du fait qu'elle est particulièrement adaptée aux séries temporelles non linéaires (cyclicité des débits sur l'année, événements saisonniers) ainsi qu'aux situations de non-indépendance des observations dans les séries (le test de Mann-Kendall étant non paramétrique, il ne repose pas sur une hypothèse d'indépendance et de distribution identique des variables aléatoires) (SILVA et al. 2024).

Il y a deux hypothèses que nous voulons tester :

- y a-t-il une diminution significative à long terme du débit mensuel moyen de chaque station $\mu_{i,t}$?
- y a-t-il une augmentation significative à long terme de la variabilité du débit mensuel moyen de chaque station $\sigma_{i,t}$?

Le caractère *significatif* doit être compris au sens statistique, dans le sens où les hypothèses ci-dessus sont acceptées en comparaison avec l'hypothèse nulle – "pas de tendance monotone à la baisse (respectivement à la hausse)" et ce en raison d'un certain niveau de confiance jugé acceptable (typiquement supérieur à 95%). Au-delà de la tendance, on cherchera également à quantifier l'ampleur de cette tendance, en proposant pour cela de calculer un taux d'évolution du débit en se basant sur la droite de régression obtenue à l'aide de la pente de Sen (SEN 1968), qui est le coefficient directeur généralement calculé dans la littérature à l'issue d'une régression de Mann-Kendall.

La régression de Mann-Kendall vise à calculer une métrique généralement notée Z_{MK} , sur laquelle un test de p-value est effectuée. La régression consiste à identifier le signe (positif ou négatif) de la différence entre toutes les paires de valeur dans la série, puis à sommer ces

signes pour obtenir une valeur unique du signe global de la série. On effectue ensuite une correction paramétrique du signe qui tient compte de l'effet des valeurs dupliquées, c'est-à-dire qui apparaissent plusieurs fois dans la série (les répétitions saisonnières). On obtient ainsi une tendance globale corrigée des effets saisonniers qu'il ne reste plus qu'à normaliser (plus rigoureusement, à standardiser) pour obtenir la métrique Z_{MK} .

Formellement, on note x_j l'observation disponible au temps j et x_k toute observation antérieure au temps j . Dans notre cas, on remplacera x par μ ou σ et on procédera station par station (l'indice i de $\mu_{i,t}$ ou $\sigma_{i,t}$ disparaît et le temps j correspond par exemple à notre mois t). On a ainsi les ensembles $(x_j)_{1 \leq j \leq n}$ et $(x_k)_{1 \leq k \leq j}$. La valeur du signe des différences est alors calculée de la façon suivante :

$$\text{sign}(x_j - x_k) = \begin{cases} 1 & \text{si } x_j - x_k > 0, \\ 0 & \text{si } x_j - x_k = 0, \\ -1 & \text{si } x_j - x_k < 0. \end{cases}$$

La contrainte $k \leq j$ implique que la dimension de la matrice $\text{sign}(x_j - x_k)_{1 \leq k \leq n, 1 \leq j \leq n}$ n'est pas $n \times n$ mais seulement la matrice triangulaire supérieure (si les j sont rangés en colonne et les k en ligne) moins la diagonale. D'où une dimension $(n-1) \times n \times \frac{1}{2}$.

On peut maintenant calculer la somme S des différences :

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sign}(x_j - x_k)$$

On effectue ensuite la correction paramétrique de S pour les valeurs dupliquées dans la série. Cette correction permet d'obtenir ce qui est considérée comme la variance de la série, et qu'on note $\text{var}(S)$. Pour calculer $\text{var}(S)$ il faut d'abord identifier toutes les valeurs dupliquées puis compter, pour chaque valeur, le nombre d'occurrences de cette valeur dans la série. On définit ainsi des groupes de valeurs dupliquées. On note g le nombre total de ces groupes et t_p le nombre d'occurrences dans le groupe p .

$$\text{var}(S) = \frac{1}{18} \left(n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t_p+5) \right)$$

On peut enfin calculer la métrique Z_{MK} :

$$Z_{MK} = \begin{cases} \frac{S-1}{\sqrt{\text{var}(S)}} & \text{si } S > 0, \\ 0 & \text{si } S = 0, \\ \frac{S+1}{\sqrt{\text{var}(S)}} & \text{si } S < 0. \end{cases}$$

Cette métrique renseigne sur la tendance de la série : si Z_{MK} est positive, la tendance est à la hausse ; si elle est négative, la tendance est à la baisse et si elle nulle aucune tendance n'existe. Il faut cependant encore conduire un test d'hypothèse pour évaluer si la tendance ainsi calculée est statistiquement significative. On effectue pour cela un test de p-value.

Le test de la p-value consiste à vérifier que la probabilité associée à la métrique Z_{MK} dans la loi normale centrée réduite est supérieure à une probabilité donnée, appelée niveau de confiance. Ce test repose donc sur l'hypothèse qu'un échantillon de tirages de Z_{MK} suive une loi normale centrée réduite. Formellement, pour $X \hookrightarrow \mathcal{N}(0, 1)$, on détermine la probabilité $p = \mathcal{P}(X \leq Z_{MK})$ dans la fonction de distribution de X . On fixe ensuite un niveau de significativité α (typiquement $\alpha \geq 0.05$) et on procède au test d'hypothèse en tant que tel :

- Hypothèse de tendance à la baisse
 - H_0 : pas tendance monotone ; H_1 : tendance monotone décroissante
 - Si $p \leq \alpha$ alors on rejette H_0 , sinon on ne peut pas rejeter H_0
 - On détermine alors le niveau de confiance $c = 1 - \alpha$
- Hypothèse de tendance à la hausse
 - H_0 : pas tendance monotone ; H_1 : tendance monotone croissante
 - On définit $p' = 1 - p = \mathcal{P}(X > Z_{MK})$
 - Si $p' \leq \alpha$ alors on rejette H_0 , sinon on ne peut pas rejeter H_0
 - On détermine alors le niveau de confiance $c = 1 - \alpha$

Numériquement, on n'imposera pas de niveau de significativité α prédéfini mais on itérera sur $0.01 \leq \alpha \leq 1$ avec un pas de 0.01 tant que $p > \alpha$ (respectivement $p' > \alpha$). Pour l'analyse, on ne retiendra que les tendances dont le niveau de confiance est supérieur à 95%.

Comment à présent estimer l'ampleur de la tendance ? C'est-à-dire quel est le taux de variation du module mensuel moyen ou de sa variabilité sur la chronique étudiée ?

Pour répondre à cette question, on estime le taux de variation associée à la droite affine de régression entre le début et la fin de la chronique. La droite affine de régression s'obtient en connaissant la pente a et l'ordonnée à l'origine b .

On détermine a grâce à la procédure de Sen, qui consiste à prendre la valeur médiane des différences entre paires de valeurs pondérées par leur distance. On calcule donc à nouveau les $(n-1)n\frac{1}{2}$ paires $(x_j - x_k)_{1 \leq k \leq n, 1 \leq j \leq n}$ que l'on pondère cette fois-ci à la distance (temporelle) entre j et k . On définit ainsi l'ensemble des $(Q_i)_{1 \leq i \leq n}$ tel que :

$$Q_i = \frac{x_j - x_k}{j - k}$$

On trie les (Q_i) par ordre croissant et on récupère la médiane $med(Q_i)$ selon la procédure suivante :

$$med(Q_i) = \begin{cases} Q_{\frac{1}{2}(n+1)} & \text{si } Q \text{ est impair,} \\ \frac{1}{2} \left(Q_{\frac{n}{2}} + Q_{\frac{n+2}{2}} \right) & \text{si } Q \text{ est pair.} \end{cases}$$

On a ainsi $a = med(Q_i)$.

Pour ce qui est de l'estimation de b , on utilise la méthode médiane (YOUNG 2018). Il s'agit de calculer l'ensemble des ordonnées à l'origine possibles à partir de toutes les observations $(x_j)_{1 \leq j \leq n}$ de la série. La forme affine générique est $y = ax + b$. Ici, on remplace les coordonnées (x, y) par (j, x_j) . On définit ainsi l'ensemble des $(b_j)_{1 \leq j \leq n}$ tels que :

$$b_j = x_j - a \times j$$

On récupère ensuite la médiane $med(b_j)$ des (b_j) et on a $b = med(b_j)$.

Finalement, avec la droite de régression on calcule le taux de variation δ aux deux extrémités de la droite, c'est-à-dire la différence entre la valeur d'arrivée et la valeur de départ, rapportée à la valeur de départ.

$$\delta = \frac{(a \times n + b) - b}{b}$$

3.3 L'assimilation des sorties d'un modèle hydrologique global à des données de mesure

Cette section aborde la deuxième question de recherche que nous adressons aux données de mesure. La première section a envisagé d'incorporer les données de mesure dans un indicateur AESA dans le but de réaliser l'évaluation environnementale de l'intégrité des cours d'eau au regard du cadre des frontières du système Terre, et ceci en modifiant l'indicateur original de telle sorte à le mettre en compatibilité avec les méthodes d'analyse de la variabilité dans le champ de la statistique. La présente section vise à s'interroger sur la possibilité d'intégrer de telles données directement dans la pratique de la modélisation, c'est-à-dire en les incorporant au modèle hydrologique global utilisé dans (ROCKSTROM et al. 2023). Une telle incorporation, ou intégration, ou encore *assimilation* (ce dernier terme se révélant le plus approprié, nous y reviendrons) des données de mesures aux données de modélisation poursuit plusieurs buts :

- l'utilisation exclusive des données de mesure pour la construction d'un indicateur de type "limite planétaire" demande d'infléchir le caractère absolu de la référence environnementale utilisée. Comme nous l'avons vu en effet, une telle référence devient en réalité une norme statistique dont le défaut principal est d'être autoréférencée dans la chronique temporelle étudiée. Il ne s'agit plus alors d'un état de référence environnemental absolu mais d'une norme. Il y a donc une réflexion à mener sur la possibilité de se doter d'une méthode pour comparer point-par-point des données de mesure qui renseignent sur l'état actuel (anthropisé) du débit, avec des données de modélisation qui renseignent sur son état préservé.
- Le modèle hydrologique global utilisé dans (ROCKSTROM et al. 2023) a fait l'objet d'une validation sur un nombre relativement réduit de stations de mesure, correspondant aux plus grands bassins versants du monde (FEKETE et al. 2002). La question de la validité du modèle sur le reste des pixels de son domaine mérite donc d'être posée et, à cet égard, il apparaît nécessaire de se doter d'une méthode de comparaison entre des données de mesure définies comme des points dans l'espace sans correspondance directe avec les pixels du maillage du modèle.

La question se réduit alors à la possibilité de développer une méthode de couplage entre les sorties d'un modèle carroyé et des données ponctuelles d'observation. D'un point de vue géométrique, il s'agit de comparer des données définies sur des entités surfaciques (pixels à

4 paires de coordonnées) et réparties de façon continue sur l'ensemble du domaine (image raster en sortie de modèle), avec des données définies sur des entités ponctuelles (points à 2 coordonnées) et réparties de façon discrète sur l'ensemble du domaine (données de mesure des stations de mesure).

Dans le champ des SIG, des méthodes existent pour convertir des couches de données `Point` en des images raster. Ces méthodes sont dites de rééchantillonnage (*resampling*) et d'interpolation. Il s'agit de combiner ces deux opérations, rééchantillonnage et interpolation, pour projeter depuis un domaine d'origine les pixels (ou points) vers un domaine de destination et estimer la valeur des données manquantes dans le domaine de destination, respectivement ².

L'opération de rééchantillonnage consiste à distribuer un échantillon de valeurs définies dans un domaine spatial d'origine vers le domaine spatial de destination, sans changer la taille de l'échantillon. Plusieurs méthodes existent pour ce faire, dont les plus communes sont la méthode "par plus proche voisin" et la méthode "moyenne" ³. La méthode par plus proche voisin convient bien pour projeter des données ponctuelles (géométrie `Point`) vers une grille de raster, en attribuant à tout pixel du raster la valeur du point le plus proche. La méthode par moyenne convient mieux pour la projection de pixels de raster vers un nouveau raster de résolution différente, en procédant par opération focale : si plusieurs pixels d'origine sont contenus un pixel de destination, le pixel de destination reçoit la valeur moyenne des pixels d'origine qu'il contient ; inversement, si plusieurs pixels de destination sont contenus dans un pixel d'origine, chaque pixel de destination se voit attribuer la valeur du pixel d'origine, de telle sorte que la moyenne des pixels de destination donne bien la valeur du pixel d'origine.

L'opération d'interpolation vient modifier la taille de l'échantillon en estimant les valeurs

2. Remarquons qu'il n'existe pas de définition formelle de ces deux termes, et que les définitions que nous proposons ci-après reprennent celles qu'en donnent les principaux fournisseurs d'algorithmes SIG que nous utilisons. Par exemple, la documentation de `GDAL` (voir : https://gdal.org/en/latest/tutorials/gdal_grid_tut.html) n'évoque pas explicitement le terme de rééchantillonnage mais lui préfère celui de "carroyage" (*gridding*) : "Gridding is a process of creating a regular grid (or call it a raster image) from the scattered data. Typically you have a set of arbitrary data scattered over the region of survey measurements and you would like to convert them into the regular grid for further processing and combining with other grids". La documentation de `rasterio` quant à elle (voir : <https://rasterio.readthedocs.io/en/stable/topics/resampling.html>) ne fait pas explicitement la différence entre ces deux opérations, tout en les mentionnant chacune : "Resampling refers to changing the cell values due to changes in the raster cell grid. [...] When you change the raster cell grid, you must recalculate the pixel values. There is no "correct" way to do this as all methods involve some interpolation".

3. Voir par exemple les différents algorithmes de rééchantillonnage embarqués dans les fonctions `gdalwarp` (<https://gdal.org/en/latest/programs/gdalwarp.html>) ou `gdal_translate` (https://gdal.org/en/latest/programs/gdal_translate.html) par le paramètre `-r`

manquantes sur l'ensemble du domaine de destination à partir des valeurs connues qui viennent d'être rééchantillonnées (c'est-à-dire projetées dans l'espace de destination). Le principe est d'estimer les valeurs prises par les pixels situés entre des pixels dont la valeur est connue, suivant une relation dite de "pondération inverse à la distance". Cette technique, développée à la fin des années 1960 (SHEPARD 1968) et raffinée dans de nombreuses variantes par les fournisseurs de logiciels SIG⁴, fait l'hypothèse, souvent vérifiée en analyse spatiale, que plus un point est éloigné d'un second, moins la probabilité qu'il partage la même valeur (pour un phénomène donné) est forte ; et inversement plus les points sont proches. Cette hypothèse se comprend aisément dans le cas d'une pluie : plus on s'éloigne de l'endroit où il pleut, moins il est probable qu'il pleuve. Ce principe, ici présenté de façon très simpliste, peut prendre des formes très sophistiquées, notamment avec les méthodes de krigeage (voir (CHILÈS et al. 2018) pour un aperçu des méthodes). Notons également que l'opération d'interpolation est le pendant spatial des opérations de réanalyse (*reanalysis*) qui consistent à reconstituer des données manquantes dans des séries temporelles ; une technique particulièrement utilisée pour l'étude des phénomènes météorologiques sur le temps long (UPPALA et al. 2005).

Ces deux opérations ne prennent toutefois pas en compte la particularité physique d'un réseau hydrographique, qui est la continuité par morceau : l'ensemble du domaine n'est pas valable car les seuls pixels autorisés à être définis sont ceux qui appartiennent effectivement au réseau hydrographique. De même, les valeurs prises par les pixels du maillage de destination ne sont pas calculées sur une base hydrauliquement représentative, mais suivant le principe de distance inverse qui n'est pas valable dans le cas où les données appartiennent à un réseau hydrographique, car la distribution spatiale des valeurs suit alors le tracé hydrographique et non plus le principe (continu dans l'espace) de distance inverse. Partant de cette double difficulté rencontrée par les techniques de rééchantillonnage actuelles dans le champ des SIG, portant à la fois sur le rééchantillonnage des positions et des valeurs, ainsi que sur leur interpolation, cette partie propose un modèle de rééchantillonnage physiquement contraint par le réseau hydrographique.

3.3.1 Un modèle de rééchantillonnage de données hydrologiques spatialisées

Le modèle de rééchantillonnage hydrographique (MRH) que nous avons développé dans le cadre de ce travail de thèse consiste à projeter sur un maillage à forte résolution les données de

4. Voir par exemple les algorithmes d'interpolations proposés par le fournisseur GDAL dans sa fonction `gdal_grid` sous le paramètre `-a` : https://gdal.org/en/latest/programs/gdal_grid.html

sortie d'un modèle hydrologique global et des données ponctuelles d'observation. Ce maillage (image raster) doit être de forte résolution pour éviter la perte de précision liée à la projection des données ponctuelles d'observation. Il s'agit alors d'augmenter la résolution des images de sortie de MHG jusqu'à la nouvelle résolution souhaitée. L'augmentation de la résolution passe ici par une technique d'estimation statistique, et non pas par le développement d'un modèle hydrologique local qui rejouerait les simulations sur un maillage de résolution accrue. La technique d'estimation que nous proposons consiste, de même que les méthodes classiques d'interpolation en géostatistique, à utiliser un noyau d'interpolation basée sur une relation inverse à la distance entre les pixels inconnus (à forte résolution) et les pixels connus (à faible résolution), mais en ajoutant une contrainte de *routing* hydrographique. Cette technique répond au besoin d'augmenter la résolution de l'image raster d'un facteur très important (typiquement par un facteur 100 si la résolution d'origine est de 7000m et la résolution cible de 30m).

Le modèle MRH peut être décrit en trois algorithmes principaux :

- un algorithme d'interpolation qui fonctionne avec ce que nous appelons un noyau de convolution hydrographique linéaire
- un algorithme de dérasterisation des pixels (qui projette les pixels de l'image raster issue du MHG vers une image raster de destination)
- un algorithme de rasterisation de données ponctuelles (qui projette les entités de géométrie `Point` vers une image raster de destination)

Ces trois algorithmes sont mutuellement indépendants et peuvent être envisagés comme des méthodes à part entière pour les besoins spécifiques auxquels elles répondent. Pour ce qui nous intéresse en revanche, à savoir la possibilité de comparer des données ponctuelles de mesure avec des données carroyées de modélisation, nous agençons ces trois algorithmes dans deux séquences de traitements différentes : le rééchantillonnage des données de modèle hydrologique global d'une part, le rééchantillonnage des données de mesure de débits d'autre part.

Pour le rééchantillonnage des données de MHG, la séquence des traitements est la suivante (voir traitements (1) et (2) dans la Figure 3.6) :

- l'image raster des débits modélisés à faible résolution est projetée dans une image raster de haute résolution par l'algorithme de rasterisation
- l'image raster ainsi générée est interpolée selon une contrainte de continuité hydrographique (*routing*) au moyen de l'algorithme de convolution hydrographique linéaire

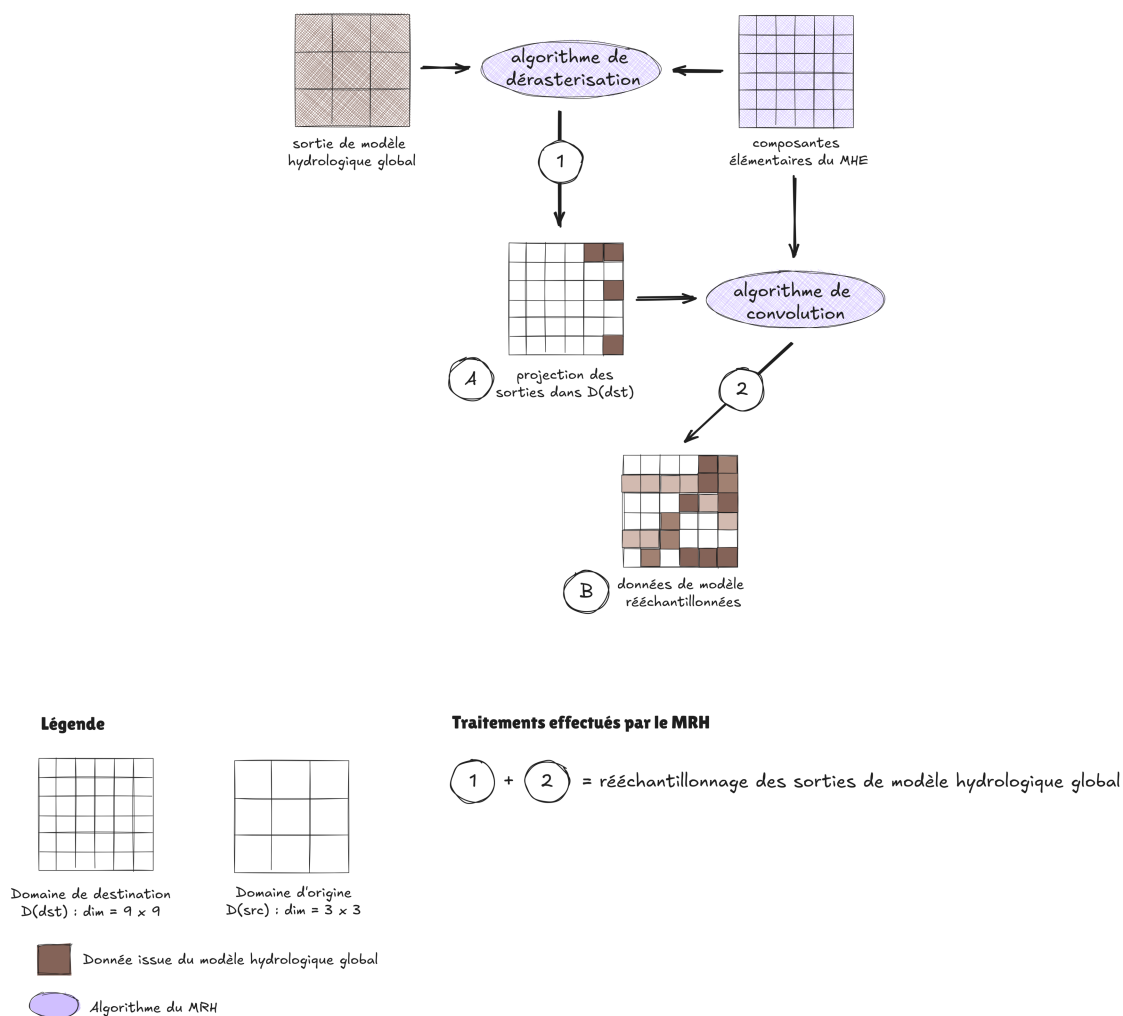


Figure 3.6 – Rééchantillonnage des données de modèle hydrologique global par le modèle MRH

Pour le rééchantillonnage des données de mesure, la séquence est la suivante (voir traitements (3) et (4) dans la Figure 3.7) :

- la couche vectorielle des données ponctuelles d’observation est projetée dans une image raster de haute résolution
- encore une fois, l’image raster ainsi générée est interpolée selon une contrainte de continuité hydrographique (*routing*) au moyen de l’algorithme de convolution hydrographique linéaire

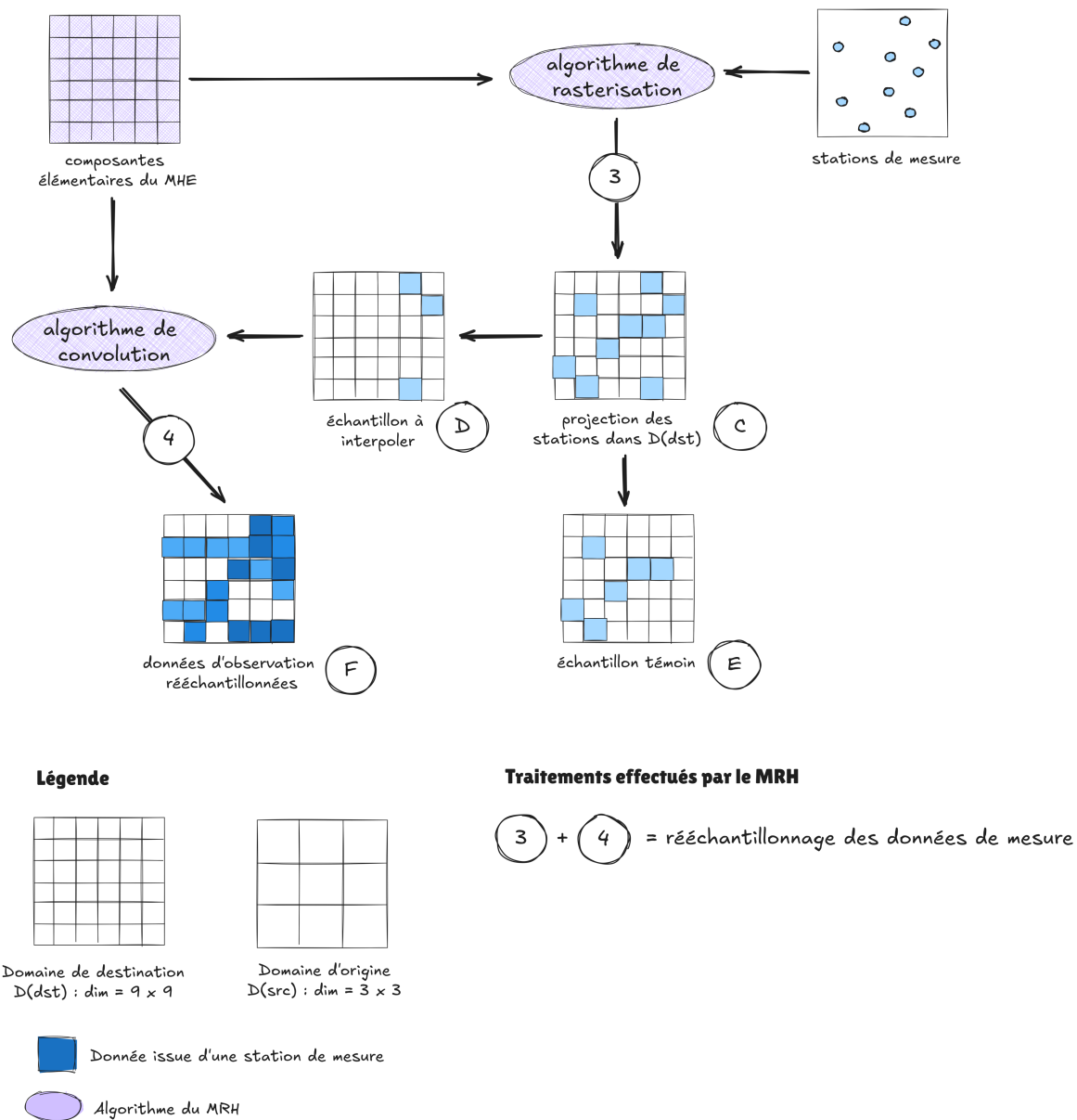


Figure 3.7 – Rééchantillonnage des données de mesure de débits par le modèle MRH

Ces deux séquences de traitements aboutissent au rééchantillonnage des données de modèle et des données de mesure dans une même image raster de haute résolution (typiquement $30\text{m} \times 30\text{m}$). À partir de là, plusieurs analyses sont envisageables :

- validation des données de modèle hydrologique global : les données de modèles rééchantillonnées peuvent être comparées aux données de mesures reprojctées, pour tout pixel en superposition dans chacune des images raster (voir comparaison de l'image (B) avec l'image (C) dans la Figure 3.8).
- validation du modèle MRH : un échantillon de données de mesure peut être interpolé selon le modèle MRH (image raster (D) interpolée en l'image (F) dans la Figure 3.8) et comparé à un deuxième échantillon témoin (image raster (E) dans Figure 3.8).

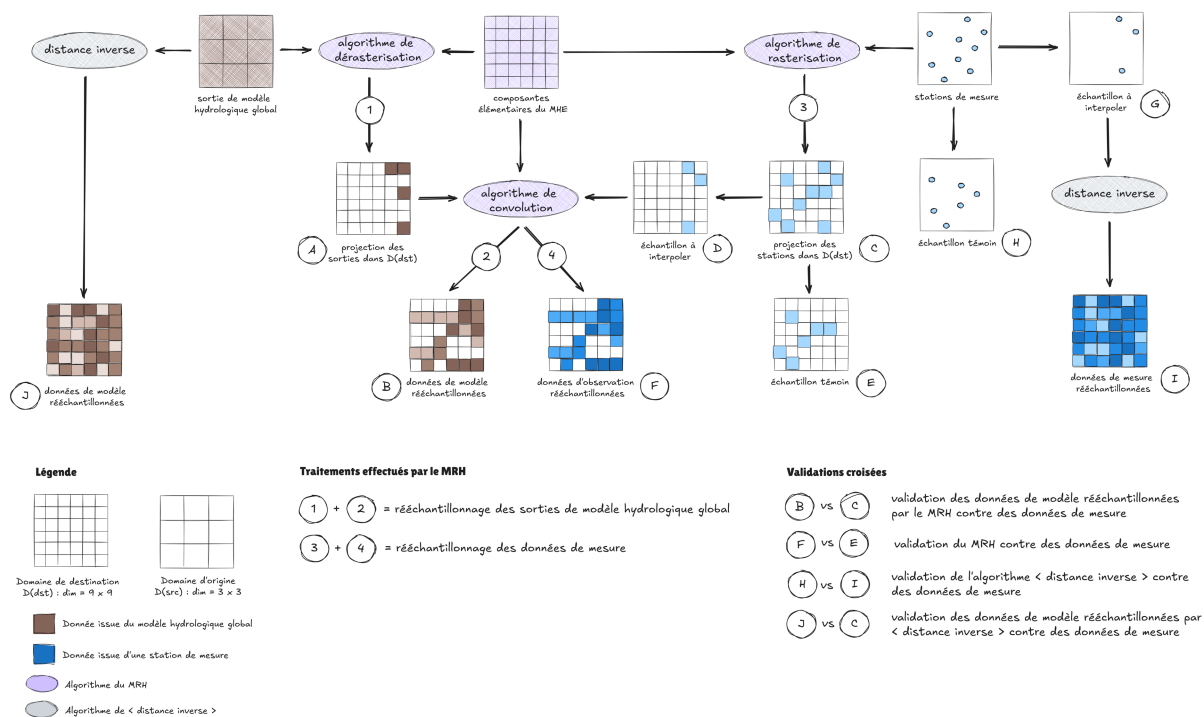


Figure 3.8 – Validation croisée du modèle MRH par l’algorithme GDAL de distance inverse

La validation du modèle MRH ne pouvant être conduite que sur son algorithme de convolution, une validation croisée avec un algorithme standard d’interpolation en SIG est menée (algorithme de distance inverse du fournisseur GDAL). Il s’agit alors d’effectuer les deux validations mentionnées ci-dessus mais à partir de l’algorithme de distance inverse plutôt du modèle MRH (voir comparaison entre les images (H) et (I) dans la Figure 3.8). De même, une comparaison entre les données de modèle rééchantillonnées par l’algorithme de distance inverse et les données de mesure reprojctées est menée (voir comparaison entre les images (J) et (C) dans la Figure 3.8).

3.3.2 Formalisation mathématique du modèle

Cette section présente la formalisation mathématique des différents algorithmes qui constituent le modèle MRH.

3.3.2.1 Notations générales

Le modèle MRH proposée peut se formaliser comme un modèle : une fonction vient convertir un échantillon passé en entrée au moyen d'un noyau d'interpolation. Commençons par une formalisation élémentaire du modèle :

- j désigne un pixel de l'image raster source, appartenant au domaine source \mathcal{D}_{src} (typiquement le raster issu du modèle hydrologique global)
- i désigne un pixel de l'image raster de destination, appartenant au domaine de destination \mathcal{D}_{dst}
- t désigne une période hydrologique élémentaire (e.g. le mois, l'année)
- $y_{i,t}$ désigne le débit estimé (ou rééchantillonné) du pixel i pour la période t (unité typique du pixel : km^3yr^{-1}).
- $x_{j,t}$ désigne le débit connu du pixel j pour la période t . En réalité, $x_{j,t}$ est connu lorsque que $j \in \mathcal{D}_{src}$ c'est-à-dire lorsque $x_{j,t}$ est spatialement référencé dans son domaine source \mathcal{D}_{src} mais devient estimé lorsqu'il est projeté dans un domaine de destination \mathcal{D}_{dst} , d'où la notation générique : $x_{j,t} \in \mathcal{D}_{src} \rightarrow \widehat{x}_{i,t} \in \mathcal{D}_{dst}$. Deux situations sont possibles pour le passage $\mathcal{D}_{src} \rightarrow \mathcal{D}_{dst}$:
 - la dérasterisation des $x_{j,t}$ lorsque \mathcal{D}_{src} est un domaine raster de géométrie élémentaire le pixel : cf. section sur l'algorithme de dérasterisation.
 - la rastérisation des $x_{j,t}$ lorsque \mathcal{D}_{src} est un domaine vectoriel de géométrie élémentaire **Point** : cf. section sur l'algorithme de rastérisation.
- χ_i désigne le noyau d'interpolation dans le domaine \mathcal{D}_{dst} : cf. section sur l'algorithme de convolution hydrographique linéaire.
- q est une fonction qui lie les variables $x_{j,t}$ et χ_i pour obtenir le débit rééchantillonné $y_{i,t}$.

Ainsi, en première formalisation, le modèle s'écrit :

$$y_{i,t} := q(\chi_i, x_{j,t})$$

Il s'agit d'un modèle paramétrique car la fonction q a recours à plusieurs jeux de paramètres :

- q transforme la variable $x_{j,t}$ en $\widehat{x}_{i,t}$ à l'aide des paramètres α_a , α_b et δ , dont nous allons ici donner une brève description, afin de les rappeler et les décrire plus en détail dans les sections suivantes.
- q invoque la fonction (ou noyau) χ_i qui est elle-même une fonction paramétrique dépendant des paramètres α_a , α_b et β_p (voir Figure 3.9).

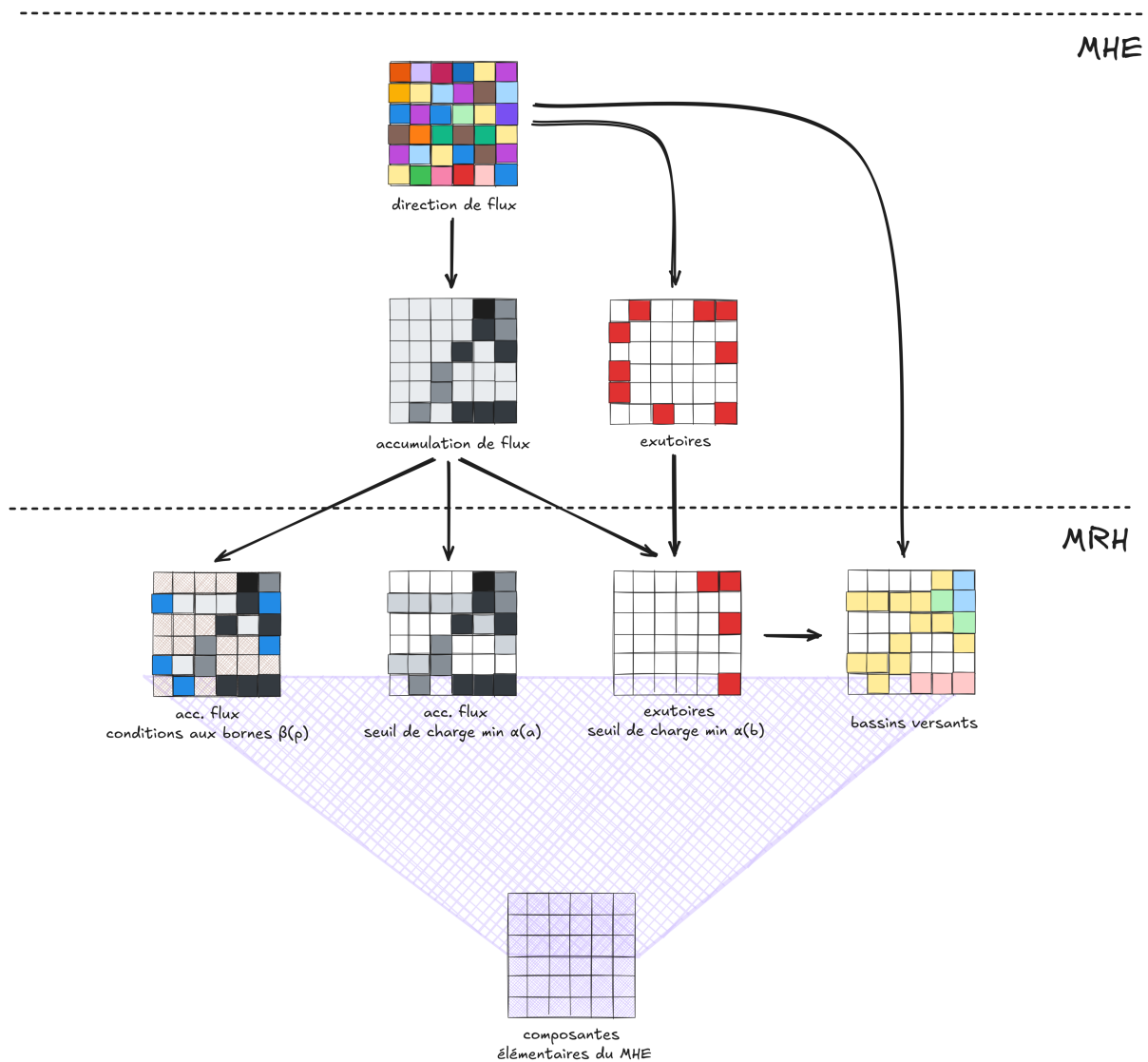


Figure 3.9 – Intervention des paramètres α_a , α_b et β_p dans le modèle MRH et lien avec le modèle hydrographique élémentaire (MHE)

Le noyau d'interpolation χ_i est mis en œuvre dans l'algorithme de convolution hydrographique linéaire. Cet algorithme fonctionne exclusivement dans le domaine \mathcal{D}_{dst} , et en particulier dans les pixels de \mathcal{D}_{dst} qui ont été déterminés comme contributifs à l'écoulement. Pour déterminer les pixels contributifs de \mathcal{D}_{dst} plusieurs traitements physiques sont nécessaires, lesquels requièrent une paramétrisation spécifique. Un paramètre physique important, dénoté α_a , est le seuil de mise en charge d'un pixel pour qu'il soit considéré en état d'écoulement permanent. Un autre paramètre déterminant consiste à déterminer les conditions initiales aux bornes de la zone d'étude, c'est-à-dire la charge hydraulique dont le réseau hérite par l'amont, paramètre que nous notons β_p , où p désigne un seuil de coupure en pourcentage de la charge hydraulique héritée de l'amont effectivement capturée par β (i.e. p est un quantile de la fonction

de distribution de la charge hydraulique amont). Le principe de fonctionnement du noyau est d'allouer une valeur de débit connue à l'exutoire le long du réseau, vers l'amont, en proportion de la charge relative de chaque pixel dans la charge totale du réseau. Ce fonctionnement entraîne l'utilisation d'un paramètre supplémentaire, noté α_b , qui correspond au seuil de mise en charge des pixels en bordure de la zone d'étude de telle sorte à les caractériser ou non comme exutoires. Ces pixels exutoires sont ceux qui initialisent la charge connue de débit à l'exutoire.

L'algorithme de dérasterisation nécessite lui aussi l'utilisation du paramètre α_b . Ce paramètre influence en effet le nombre de bassins versants qui seront délimités. Or un bassin versant se délimite à partir de deux données : un exutoire et une couche de directions de flux. Le principe de l'algorithme de dérasterisation est d'associer à chaque pixel de l'image raster dans \mathcal{D}_{src} un ensemble de points (en fait des pixels de petite taille) dans \mathcal{D}_{dst} qui correspondent aux exutoires sous-jacents. L'algorithme de dérasterisation implique donc également l'utilisation du paramètre α_a car il s'appuie sur la modélisation du réseau hydrographique sous-jacent. Ainsi, si $x_{j,t}$ est une valeur de débit connu dans \mathcal{D}_{src} (image raster) alors sa valeur estimée $\widehat{x}_{i,t}$ dans \mathcal{D}_{dst} (couche vectorielle de géométrie **Point**) dépend de α_a et α_b .

L'algorithme de rasterisation convoque pour sa part une fonction de jointure spatiale gloutonne par plus proche voisin qui permet de projeter une couche vectorielle de géométrie **Point** vers le domaine raster \mathcal{D}_{src} . Cette fonction crée un décalage de position spatiale entre la coordonnée $\langle \mathbf{x}, \mathbf{y} \rangle$ initiale et la coordonnée du pixel attribué. L'ensemble des écarts de position est dénoté par le vecteur δ . Ainsi, si $x_{j,t}$ est une valeur de débit connu dans \mathcal{D}_{src} (couche vectorielle de géométrie **Point**) alors sa valeur estimée $\widehat{x}_{i,t}$ dans \mathcal{D}_{dst} (image raster) dépend de δ .

Finalement, de façon générique et pour grouper les paramètres impliqués dans les deux algorithmes de rasterisation et de dérasterisation, on peut noter :

$$\widehat{x}_{i,t}(\alpha_a, \alpha_b, \delta)$$

De même, pour résumer les paramètres inhérents à l'algorithme de convolution, on peut noter :

$$\chi_i(\alpha_a, \alpha_b, \beta_p)$$

Le modèle MRH se formalise alors ainsi :

$$y_{i,t} := q(\chi_i, x_{j,t} : \alpha_a, \alpha_b, \beta_p, \delta)$$

La démarche complète entreprise par le modèle MRH est schématiquement reproduite dans la Figure 3.10.

Avant de présenter le fonctionnement de chacun des algorithmes, nous introduisons quelques notations génériques supplémentaires.

L'une de ces notations supplémentaires permet de désigner les points de l'espace dans le domaine \mathcal{D}_{dst} qui constituent effectivement le réseau hydrographique (c'est-à-dire les pixels de \mathcal{D}_{dst} qui peuvent être considérés comme partie d'un cours d'eau actif). Ainsi on note, $\forall i \in \mathcal{D}_{dst}, i \in \mathcal{G}_v$ si i est un pixel du réseau hydrographique \mathcal{G}_v où $v \in \llbracket 1, w \rrbracket$ et \mathcal{G}_v est un sous-ensemble de \mathcal{D}_{dst} . Autrement dit \mathcal{G}_v est l'ensemble des pixels i de \mathcal{D}_{dst} considérés comme formant partie d'un cours d'eau actif et v désigne tout à la fois un exutoire, un bassin versant et le réseau hydrographique de ce bassin versant dans \mathcal{D}_{src} .

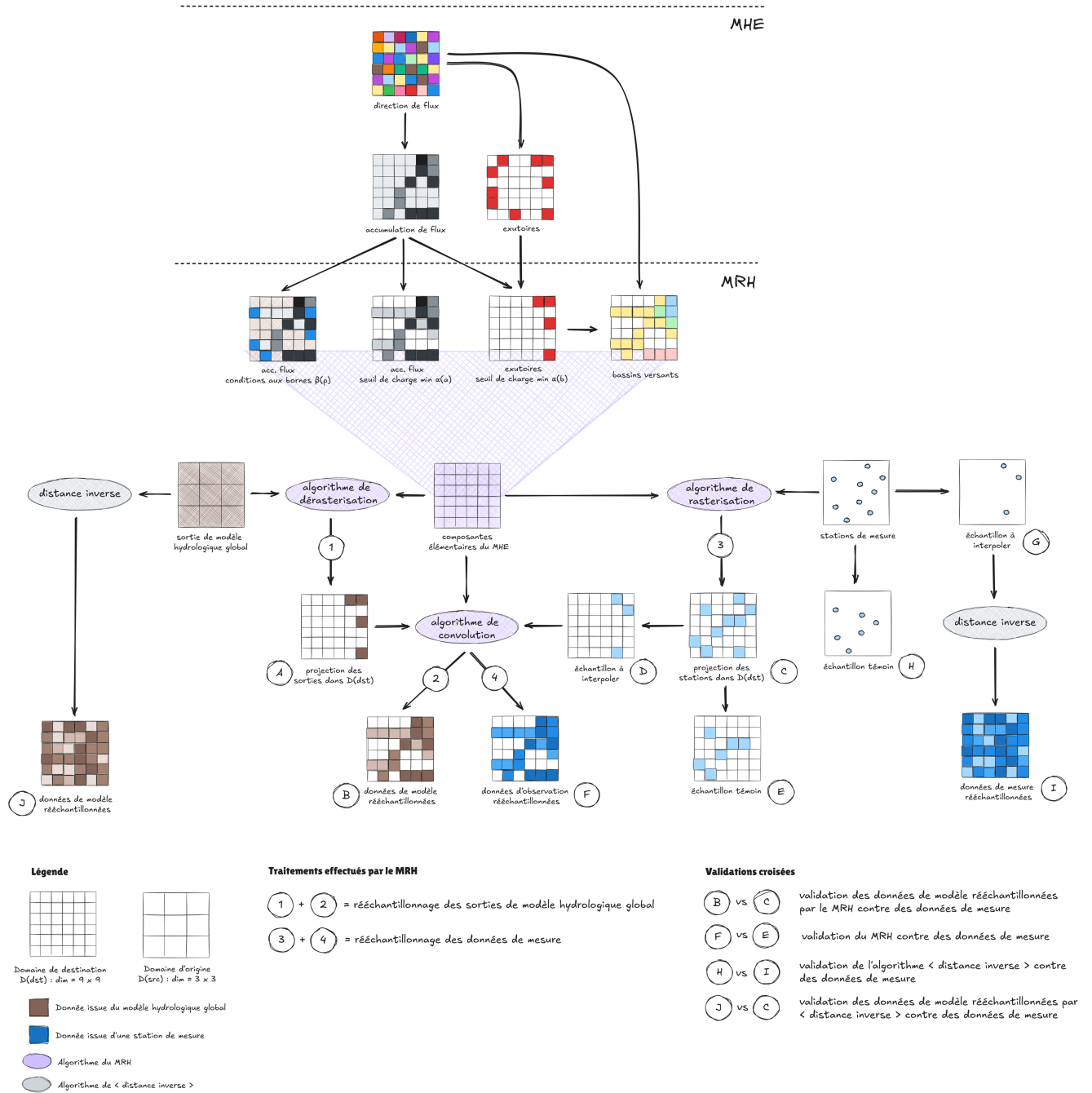


Figure 3.10 – Vue d'ensemble du modèle de rééchantillonnage hydrographique

3.3.2.2 Une méthode de *data assimilation*

La *data assimilation* est un terme qui recouvre un ensemble de méthodes qui visent toutes à coupler des données dites numériques ou modélisées avec des données dites d'observation ou de mesure afin d'améliorer la description du système étudié, c'est-à-dire de réduire les incertitudes entourant la description du système, et plus précisément l'incertitude liée à l'imperfection des données de mesure et à la paramétrisation des données modélisées (DELMOTTE et al. 2011).

Le modèle MRH dans son ensemble peut s'apparenter à une méthode de *data assimilation* puisque son but est de permettre la comparaison pixel-à-pixel entre les sorties d'un modèle hydrologique et des données de mesure, et ceci à travers deux modèles intermédiaires, un "modèle de dérasterisation" qui rééchantillonne le maillage de faible résolution d'un modèle hydrologique global vers un maillage à forte résolution, et un "modèle de rasterisation" qui projette les données de mesure définies par une géométrie `Point` vers le réseau hydrographique modélisé dans le maillage à forte résolution. Ces deux modèles intermédiaires ne sont pas sans faire penser à ce qui est appelé un *observation operator* dans le vocabulaire de la *data assimilation*.

Nous détaillons davantage cet aspect en annexe, où nous formalisons le lien entre le modèle MRH et l'écriture générique d'un modèle de *data assimilation*.

3.3.2.3 Algorithme de convolution hydrographique linéaire du débit

Cette partie ainsi que les deux parties suivantes vont s'attacher à décrire et formaliser le fonctionnement des trois algorithmes constitutifs du modèle MRH.

Principe général

Il s'agit de la méthode de rééchantillonnage mixte physique-statistique à proprement parler. L'enjeu est d'assurer la cohérence hydraulique d'une méthode statistique. En effet, le débit en tout point de l'espace n'est pas estimé à partir d'une transformation pluie-débit (comme dans un modèle hydrologique) mais à partir d'une convolution du débit dit exutoire vers l'amont du réseau hydrographique. Plus simplement, le modèle MRH estime le débit en un point du réseau en lui attribuant une fraction du débit exutoire connu correspondant à la charge hydraulique relative du réseau en ce point. Cette charge hydraulique relative est obtenue par la convolution linéaire d'un transport initial de matière mis en charge à l'amont du réseau, et normalisée sur la charge à l'exutoire.

Le modèle MRH, à travers son algorithme de convolution hydrographique, effectue en quelque sorte une "règle de trois". Il s'agit d'estimer une quantité inconnue à partir de trois autres quantités connues. Plus précisément, on cherche à connaître le débit en tout pixel i d'un réseau hydrographique à partir :

- d'une valeur connue de débit au pixel j du même réseau

- d'une valeur théorique de charge (ou accumulation) au pixel j qu'on notera a_j
- d'une valeur théorique de charge au pixel i qu'on notera a_i

Le débit inconnu $y_{i,t}$ (pour la période t) est ainsi obtenu en appliquant au débit connu $\widehat{x}_{k,t}$ (il s'agit du pixel $x_{j,t}$ de \mathcal{D}_{src} qui a été estimé dans \mathcal{D}_{dst}) la proportion calculée entre la charge théorique respective de ces deux pixels. L'algorithme de convolution fonctionne exclusivement dans le domaine de destination \mathcal{D}_{dst} et donc la conversion $x_{j,t} \in \mathcal{D}_{src} \rightarrow \widehat{x}_{i,t} \in \mathcal{D}_{dst}$ doit déjà avoir été effectuée. La formule de la convolution est alors la suivante :

$$\forall i \in \mathcal{D}_{dst}, k \in \mathcal{D}_{dst}, \quad y_{i,t} = \widehat{x}_{k,t} \cdot \frac{a_i}{a_k}$$

Dans cette formule, on reconnaît χ_i qui est défini comme $\frac{a_i}{a_k}$. Notons que k est fixe et que ce sont les pixels i inconnus qui varient. Notons également que la valeur théorique de charge a ne dépend pas de la période t .

Une remarque sur le pixel k retenu s'impose. Ce dernier correspond en fait à la charge maximale du réseau, et nous en verrons la raison au moment où nous détaillerons l'algorithme de dérasterisation. Pour le moment, nous pouvons souligner la relation suivante :

$$\forall i \in \mathcal{D}_{dst}, k \in \mathcal{D}_{dst}, \quad \chi_i = \frac{a_i}{a_k} = \frac{a_i}{\max_{i \in \mathcal{G}_v} \{a_i\}}$$

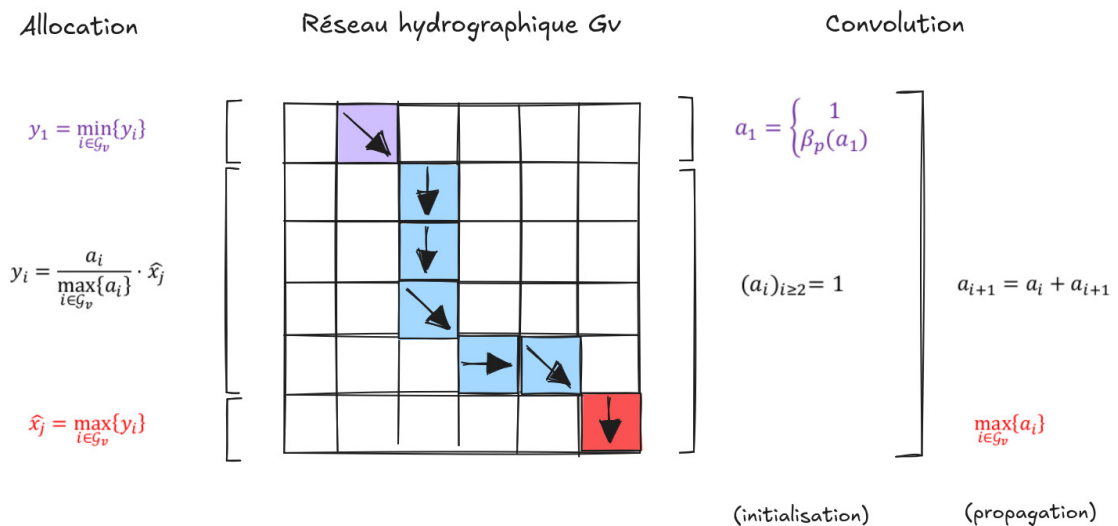


Figure 3.11 – Interpolation du débit par l'algorithme de convolution hydrographique linéaire

Convolution linéaire : emprunt à l'hydrogramme unitaire

La charge (ou accumulation) théorique a en tout pixel du réseau est obtenue selon un principe de convolution linéaire, qui s'appuie sur la théorie de l'hydrogramme unitaire (SHERMAN 1932). Cette théorie fait plusieurs hypothèses :

- l'intensité de la pluie (dans notre cas la charge de matière initiale à transporter) est constante sur l'ensemble du bassin versant.
- infiltration et évapotranspiration sont des paramètres constants sur l'ensemble du bassin versant (dans notre cas, nous les considérerons comme uniformément nuls).

Il y a convolution linéaire du débit car la charge initiale portée par tout pixel s'ajoute de l'amont vers l'aval. Comme dans la théorie de l'hydrogramme unitaire où la pluie initiale est dite unitaire parce qu'elle vaut 1 (dans l'unité que l'on souhaite, e.g. $mm \cdot h^{-1}$), la valeur d'initialisation des pixels du réseau vaut 1, sauf s'ils sont compris dans le vecteur de conditions initiales aux bornes de la zone d'étude β_p . L'existence d'un tel vecteur est nécessaire car la zone d'étude est un système ouvert : le réseau hydrographique reçoit une charge amont exclue de la zone d'étude.

Pour expliquer le principe de convolution linéaire, prenons l'exemple d'un réseau à trois pixels $\{0, 1, 2\}$ rangés de l'amont vers l'aval et sans affluents (i.e. un réseau à un seul cours d'eau) :

- si a_0 n'appartient pas au vecteur de conditions initiales β_p : alors les conditions initiales sont $a_0 = a_1 = a_2 = 1$ et la convolution linéaire donne $a_1 = a_0 + a_1 = 2$ et $a_2 = a_1 + a_2 = 3$
- si a_0 appartient au vecteur de conditions initiales β_p : alors les conditions initiales sont $a_0 = \beta_p(0)$ et $a_1 = a_2 = 1$ et la convolution linéaire donne $a_1 = a_0 + a_1 = \beta_p(0) + 1$ et $a_2 = a_1 + a_2 = \beta_p(0) + 2$

La charge d'un pixel est ainsi toujours égale à sa propre charge initiale + la charge amont. Il s'agit donc d'une somme cumulée.

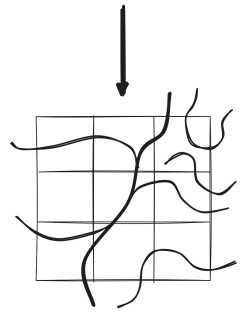
Nous développons en annexes deux remarques sur la nature précise de qui est modélisé par cet algorithme de transport. Une première remarque porte sur la différence la modélisation du ruissellement et la modélisation de l'écoulement : l'algorithme de convolution que nous avons présenté ici modélise l'écoulement et pas le ruissellement, qui sont deux phénomènes que nous proposons de différencier selon que le système est à l'état stable (écoulement) ou dans un état

transitoire (ruissellement). La seconde remarque concerne la différence entre la modélisation d'un transport de masse et la modélisation d'un transport de grandeur hydrométrique (i.e. le débit), où nous insistons sur le fait que s'il y a bien conservation de la masse, il peut ne peut pas y avoir conservation du débit car celui-ci dépend des conditions d'écoulement (section hydraulique, rugosité, vitesse) et pas seulement de la masse transportée.

3.3.2.4 Algorithme de dérasterisation des pixels de débit modélisé

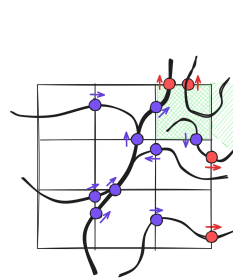
Cet algorithme réalise la première étape du rééchantillonnage de l'image raster de \mathcal{D}_{src} qui consiste à mettre l'image à l'échelle de \mathcal{D}_{dst} . Cette mise à l'échelle consiste plus précisément à attribuer à certains pixels de \mathcal{D}_{dst} les pixels de \mathcal{D}_{src} . À l'issue de cette étape, le réseau hydrographique n'est pas encore intégralement reconstitué ; il a été discrétisé sur certains pixels de \mathcal{D}_{dst} . C'est l'algorithme de convolution linéaire qui permettra d'aboutir à un réseau continu dans \mathcal{D}_{dst} (voir Figure 3.12).

Modèle Hydrographique Élémentaire



~ Accumulation de flux

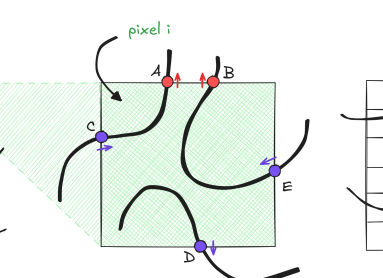
Identification des exutoires



- Exutoire de transfert
- Exutoire final
- Sens d'écoulement

Dérasterisation

Allocation des débits



Zoom sur le pixel i :

Le débit du pixel i peut se répartir sur les nœuds sortants :
 $Q(\text{pixel } i) = Q(A) + Q(B) + Q(D)$

C et E sont des exutoires de transfert entrant :
 $Q(C)$ contribue à $Q(A)$
 $Q(E)$ contribue à $Q(B)$

D est exutoire de transfère sortant :
 il doit être décompté.

Le débit en D est estimé par :
 $Q(D) = \text{accr}(D) \times Q(\text{pixel } i)$
 Où accr désigne l'accumulation relative.

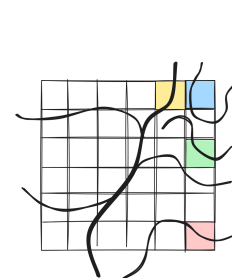
$$\text{accr}(D) = \frac{\text{acc}(D)}{\sum_{A \leq k \leq E} \text{acc}(k)}$$

et acc l'accumulation brute.

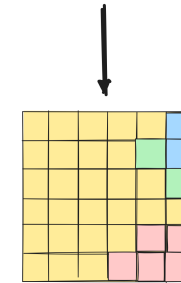
Le débit restant est alors :
 $Q(A+B) = Q(\text{pixel } i) - Q(D)$

Il est enfin alloué à A et B :
 $Q(A) = \text{accr}(A) \times Q(A+B)$
 $Q(B) = \text{accr}(B) \times Q(A+B)$

Image raster finale

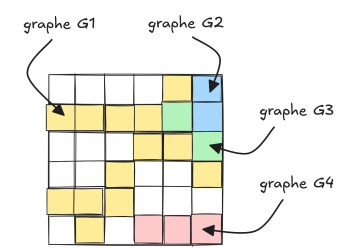


Modèle Hydrographique Élémentaire



Délimitation des sous bassins-
versants à partir des exutoires
finaux

Convolution



X désigne un exutoire final parmi les 4.

$Q(X)$ est alloué au reste des pixels de son bassin versant (ou graphe G_x) en proportion de leur accumulation relative (accr).
 accr est calculée après découpage des bassins versants.

Figure 3.12 – Représentation schématique de l'algorithme de dérasterisation et lien avec l'algorithme de convolution

L'algorithme de dérasterisation n'est pas un algorithme de vectorisation. Son but est de convertir une image raster source comportant des données hydrologiques organisées en réseau (*routed*) dans une image raster destination dont la résolution du pixel est différente (le reste des propriétés `geoTransform` reste inchangé : système de projection, étendue, etc).

La spécificité des valeurs portées par les pixels d'une telle image est qu'elles sont organisées en réseau, c'est-à-dire qu'il y a entre elles une cohérence hydraulique. Les algorithmes de rééchantillonnage traditionnellement utilisés en SIG effectuent également cette première étape de dérasterisation avant l'étape d'interpolation (que nous avons appelé *convolution* dans notre cas), mais sans forcer la dérasterisation par une contrainte hydrographique. La procédure classique de dérasterisation en SIG consiste à convertir le barycentre des pixels de l'image source en pixels dans l'image de destination. Cette attribution s'envisage comme une fonction injective, par laquelle tout pixel de \mathcal{D}_{src} est associé à un unique pixel de \mathcal{D}_{dst} mais où seulement certains pixels de \mathcal{D}_{dst} sont associés à un pixel de \mathcal{D}_{src} . Considérer le centre géométrique d'une surface comme représentatif de la valeur prise par cette surface n'a pas de sens hydrauliquement.

Il nous a ainsi paru plus judicieux de déterminer les pixels à l'intérieur de cette surface qui soient hydrographiquement représentatifs de la valeur de la surface. La proposition que nous faisons considère comme représentatifs les pixels qui correspondent aux exutoires hydrographiques de la surface, c'est-à-dire les pixels par lesquels les cours d'eau s'échappent de la surface. Et il y a en effet plusieurs pixels susceptibles d'être de tels exutoires à l'intérieur de la surface. Notre algorithme n'est ainsi ni injectif (plusieurs pixels de \mathcal{D}_{dst} peuvent être associés à un même pixel de \mathcal{D}_{src}) ni surjectif (il existe des pixels de \mathcal{D}_{dst} qui ne sont associés à aucun pixel de \mathcal{D}_{src}), ce qui le différencie mathématiquement des algorithmes de dérasterisation communs en SIG (pour l'injection du moins). Cette différence mathématique implique, pour les algorithmes conventionnels, une attribution directe de la valeur du pixel source vers le pixel de destination. Pour notre algorithme en revanche, l'attribution n'est pas directe et consiste en deux étapes : une attribution dans l'espace des positions et une attribution dans l'espace des valeurs. Après avoir déterminé quels pixels de destination sont associés à tout pixel source, il faut attribuer (ou répartir) la valeur du pixel source aux différents pixels de destination associés.

Cette attribution dans l'espace des valeurs s'appuie là encore sur le principe de la convolution linéaire du débit le long du réseau hydrographique. Selon ce principe, la valeur maximale de débit

est atteinte au pixel le plus à l'aval du réseau, c'est-à-dire au pixel exutoire. Notre zone d'étude étant un système ouvert (le réseau hydrographique capturé par la zone d'étude reçoit une charge à l'amont et transfère une charge à l'aval qui ne sont pas capturées par la zone d'étude), ce que nous appelons "pixel exutoire" ne correspond pas nécessairement à l'exutoire naturel du bassin versant, mais au pixel où le débit est transféré en dehors de la zone d'étude (où l'eau "s'échappe"). Plusieurs propriétés utiles pour l'attribution des pixels dans l'espace des valeurs découlent de ce principe :

- à tout pixel on peut attribuer une valeur de charge relative dans le réseau (rapport entre sa charge hydraulique et la charge à l'exutoire). Cette charge relative peut constituer une méthode d'attribution dans l'espace des valeurs
- les charges relatives entre les pixels de \mathcal{D}_{dst} n'ont de sens que si les pixels appartiennent au même sous bassin versant (au même graphe)
- enfin, la caractéristique "exutoire" d'un pixel est différenciée selon deux situations :
 - tout pixel de \mathcal{D}_{src} peut être considéré comme une zone d'étude à part entière et comporte à ce titre des pixels exutoires dans le réseau hydrographique sous-jacent de \mathcal{D}_{dst} . Autrement dit, le débit d'un pixel de \mathcal{D}_{src} est transféré aux pixels adjacents et ce transfert se produit par les pixels exutoires superposés dans \mathcal{D}_{dst} . Il s'agit d'un type d'exutoire que nous appellerons "exutoire de transfert" (voir Figure 3.12)
 - à l'échelle de la zone d'étude, certains pixels de \mathcal{D}_{dst} transfèrent l'écoulement en dehors de la zone sans le transférer à d'autres pixels de \mathcal{D}_{src} superposés. Ce type d'exutoire, nous l'appellerons "exutoire final" (voir Figure 3.12)

Pour résumer les différences entre notre algorithme de dérasterisation et ceux communément utilisés en SIG, nous proposons la table suivante :

	espace des positions	espace des valeurs
algorithme SIG	pixel barycentre de \mathcal{D}_{dst} superposé au pixel de \mathcal{D}_{src}	fonction identité $\forall \lambda \in \mathbb{R}, id(\lambda) = \lambda$
algorithme MRH	pixels exutoires de \mathcal{D}_{dst} superposés au pixel de \mathcal{D}_{src}	fonction proportionnelle λ est réparti sur les pixels exutoires au prorata de leur charge relative

TABLE 3.2 – Comparaison des algorithmes de dérasterisation entre le modèle MRH et un modèle SIG quelconque

Pour poursuivre, on rappelle la notation que nous avons adoptée pour les graphes constitutifs

du réseau hydrographique : $\forall i \in \mathcal{D}_{dst}, i \in \mathcal{G}_v$ si i est un pixel du réseau hydrographique \mathcal{G}_v où $v \in \llbracket 1, w \rrbracket$ avec w le nombre d'exutoires. \mathcal{G}_v est ainsi un sous-ensemble de \mathcal{D}_{dst} (\mathcal{G}_v est l'ensemble des pixels i de \mathcal{D}_{dst} considérés comme formant partie d'un cours d'eau actif). On note Γ_w l'ensemble des w graphes.

En reprenant ces notations, on peut décrire l'algorithme de dérasterisation comme une double étape : reprojexion des pixels de \mathcal{D}_{src} vers \mathcal{D}_{dst} et estimation des valeurs de débit dans le nouvel espace \mathcal{D}_{dst} . On rappelle également que a_i désigne la charge hydraulique théorique du pixel i dans \mathcal{D}_{dst} .

Commençons par affiner la définition de \mathcal{G}_v :

$$\forall v \in \llbracket 1, w \rrbracket, \quad \mathcal{G}_v := \left\{ i \in \mathcal{D}_{dst} \mid (a_i \geq \alpha_a) \ \& \ \left(\sum_{i \in \mathcal{G}_v} a_i = a_v \right) \right\}$$

\mathcal{G}_v correspond à l'ensemble des pixels de \mathcal{D}_{dst} qui répondent à deux conditions (l'opérateur logique "et" est désigné par $\&$) : leur charge hydraulique est au moins égale à la charge seuil (qui détermine si un pixel est considéré comme un cours d'eau permanent ou non), et la somme de la charge des pixels de \mathcal{G}_v est égale à la charge du pixel v à l'exutoire de \mathcal{G}_v .

Nous introduisons enfin une notation différenciée pour les exutoires dits "finaux" et les exutoires dits "de transfert". On note Ω_j l'ensemble des pixels $i \in \mathcal{D}_{dst}$ superposés au pixel $j \in \mathcal{D}_{src}$ qui constituent des exutoires finaux. Et on note Φ_j l'ensemble des pixels $i \in \mathcal{D}_{dst}$ superposés au pixel $j \in \mathcal{D}_{src}$ qui constituent des exutoires de transfert.

Nous pouvons à présent formaliser la fonction qui effectue l'association entre \mathcal{D}_{src} et \mathcal{D}_{dst} dans l'espace des valeurs (i.e. l'attribution des valeurs de débits aux pixels reprojétés). Pour toute période t , le débit du pixel $x_{j,t} \in \mathcal{D}_{src}$ peut être décomposé en deux fractions : une fraction de $x_{j,t}$ qui sort de la zone d'étude par les pixels sous-jacents dans \mathcal{D}_{dst} de Ω_j ; et une fraction de $x_{j,t}$ qui est transférée à un autre pixel de la zone d'étude par les pixels sous-jacents dans \mathcal{D}_{dst} de Φ_j . On peut donc écrire $x_{j,t}$ sous la forme suivante :

$$\forall j \in \mathcal{D}_{src}, \quad x_{j,t} = x_{j,t} \cdot (\omega_j + \varphi_j)$$

Les variables ω_j et φ_j représentent respectivement la part du débit qui sort de la zone d'étude et la part qui est transférée à un autre pixel. Ces parts se calculent de la façon suivante : ω_j est

la somme des charges hydrauliques des pixels $i \in \mathcal{D}_{dst}$ superposés à $j \in \mathcal{D}_{src}$ considérés comme des pixels "exutoires finaux" (de Ω_j), et rapportée à la somme des charges hydrauliques de tous les pixels exutoires $i \in \mathcal{D}_{dst}$ superposés à $j \in \mathcal{D}_{src}$; et φ_j correspond au même ratio mais avec comme numérateur les pixels dits "exutoires de transfert". Formellement, on note :

$$\forall j \in \mathcal{D}_{src}, \quad \omega_j = \frac{\sum_{i \in \Omega_j} a_i}{\sum_{i \in \Omega_j} \sum_{i \in \Phi_j} a_i}$$

$$\forall j \in \mathcal{D}_{src}, \quad \varphi_j = \frac{\sum_{i \in \Phi_j} a_i}{\sum_{i \in \Omega_j} \sum_{i \in \Phi_j} a_i}$$

Le lecteur pourra trouver un exemple schématique de calcul dans la Figure 3.12. Enfin, seuls les exutoires finaux de \mathcal{D}_{dst} sont retenus comme pixels à partir desquels l'interpolation sera conduite. La fonction qui permet de leur attribuer une part du débit du pixel superposé dans l'espace des valeurs attribue la valeur du pixel $x_{j,t}$ aux pixels exutoires au prorata de leur charge hydraulique relative.

$$\forall i \in \Omega_j, \quad x_{i,t} = x_{j,t} \cdot \omega_j \cdot \frac{a_i}{\sum_{i \in \Omega_j} a_i}$$

Un dernier raffinement est effectué, qui consiste à ne retenir que les exutoires finaux ayant une charge hydraulique supérieur à une charge seuil, afin de ne pas travailler par la suite sur des bassins versants trop petits). Ainsi l'espace Γ_w est défini par :

$$\forall v \in \Omega_v, \quad \mathcal{G}_v \in \Gamma_w \quad \text{si} \quad a_v \geq \alpha_b$$

Un pixel est susceptible d'être considéré comme l'exutoire d'un graphe si sa charge dépasse la charge seuil α_b .

3.3.2.5 Algorithme de rasterisation des données ponctuelles de débit observé

Cet algorithme a déjà été décrit lorsque nous avons présenté le MHE. Nous le mettons à nouveau ici en avant pour discuter de son influence sur le paramètre δ du modèle.

Pour rappel, cet algorithme permet de projeter la couche vectorielle des stations de mesure (de géométrie `Point`) dans le domaine \mathcal{D}_{dst} . À partir de la position des stations en coordonnées $\langle \mathbf{x}, \mathbf{y} \rangle$, l'algorithme effectue une jointure spatiale par plus proche voisin avec la couche des pixels de \mathcal{D}_{dst} préalablement convertis en couche vectorielle de géométrie `Point` où tout pixel

se voit attribuer la coordonnée $\langle \mathbf{x}, \mathbf{y} \rangle$ de son barycentre.

L'image raster de \mathcal{D}_{dst} utilisée pour la jointure correspond au raster d'accumulation de flux filtré sur les pixels dont la charge est supérieure à la charge seuil α_a . Ces pixels, convertis en coordonnées, sont joints à la couche vectorielle des stations de mesure par l'intermédiaire d'un algorithme glouton. Le principe de fonctionnement est de calculer une matrice des distances entre la couche des coordonnées des pixels et la couche des stations. Dans un ordre arbitraire, l'algorithme prend une station de mesure et lui attribue la coordonnée du pixel le plus proche dans la matrice des distances. Cette coordonnée est alors retirée de la liste des coordonnées disponibles pour les stations suivantes. Il s'agit donc d'un algorithme de type "premier arrivé premier servi" qui permet d'attribuer à toute station une unique coordonnée de pixel, même en cas d'équidistance entre plusieurs stations et une même coordonnée de pixel.

L'inconvénient d'un tel algorithme glouton est que plus le nombre de stations à joindre est important, plus la probabilité d'une jointure avec un voisin qui n'est pas le plus proche à mesure que la liste des coordonnées de pixels disponibles diminue. Cette erreur inhérente à la construction de l'algorithme (incertitude structurelle) est formalisée dans le paramètre δ que nous avons introduit plus haut.

δ peut être envisagé comme un vecteur qui contient la distance euclidienne entre chaque paire de station et de pixel associés. On pourrait envisager d'ajouter à l'algorithme de rasterisation un algorithme de minimisation de δ qui fonctionnerait de la façon suivante : on effectue la jointure spatiale avec l'algorithme glouton autant de fois que de combinaisons dans la matrice des distances, en changeant à chaque itération l'ordre de traitement des stations (la première station à la première itération devient la deuxième station de l'itération suivante, etc), puis on range les distances des paires effectivement associées dans δ après chaque itération. On effectue pour chaque δ la somme de ses distances, et on détermine l'ordre de tirage des stations qui minimise ces sommes. Cette optimisation (minimisation) de δ n'a pas été implémentée numériquement car nous avons préféré fournir une valeur d'erreur globale du modèle MRH plutôt qu'une valeur d'erreur propre à chacun de ses algorithmes constitutifs.

Conclusion

Ce chapitre a détaillé les deux propositions méthodologiques que nous avons formulées pour l'utilisation des données de mesure de débits systématisées dans la mise en œuvre de l'indicateur de (ROCKSTROM et al. 2023). La première proposition méthodologique consiste à construire une méthode AESA à partir de telles données, ce qui distingue la démarche de celle adoptée par (ROCKSTROM et al. 2023) où les données mobilisées peuvent être qualifiées de données de modèle de simulation numérique, en reprenant la typologie de l'échelle de la donnée que nous avons proposée au chapitre 2. La différence qui nous intéresse entre les données de débits mesurés sur site et les sorties de modèles de simulation tient au caractère spatialement explicite des premières et spatialement différencié des secondes.

L'indicateur sur l'eau de surface proposé par (ROCKSTROM et al. 2023) présente également l'intérêt de mettre en tension ces deux types de données, autour de la notion de "variabilité excessive". Si la question de la variabilité du débit dans une perspective d'hydrologie fréquentielle renvoie à certaines métriques de quantification issue du champ des statistiques, dont l'écart-type et le coefficient de variation, qui ne sont que des mesures de l'erreur associée à une valeur normale (ou moyenne) de débit, la variabilité au sens où l'emploie (ROCKSTROM et al. 2023) désigne l'écart entre une valeur de débit en régime perturbé (par les pressions anthropiques) et sa valeur en régime non perturbé. Dans cet indicateur de variabilité, le type de données mobilisées vient ainsi questionner ce qui sert de point de référence à partir duquel évaluer la variation du débit. Nous avons à cet égard distingué entre une référence dite absolue (le régime d'écoulement non perturbé) et une référence dite normale (au sens d'une norme statistique). Les données de mesure de débits quantifiant par construction le régime perturbé, il n'est pas possible de construire une référence absolue mais seulement *normale*. L'indicateur d'excès de variabilité devient ainsi un indicateur d'anomalie de variabilité, mesuré comme l'écart entre la variabilité du débit moyen d'un mois donné sur une chronique et la variabilité moyenne du débit pour tous les mêmes mois de la chronique. Un tel indicateur ne porte donc en soi pas d'information sur l'origine de la pression excessive qu'il quantifie, par exemple celle de savoir si une variabilité anormale tient à un excès de pressions anthropiques ponctuelles, à des aléas naturels ou au forçage radiatif (les aléas naturels pouvant eux-mêmes être forcés). Nous avons donc proposé l'utilisation d'une méthode de régression de Mann-Kendall pour déceler l'impact statistiquement significatif ou non du forçage climatique sur les débits observés et leur variabilité. Cette méthode ne quantifie pas la contribution de chacune des

pressions à la variabilité totale, mais permet du moins de distinguer entre les stations de mesure où l'effet d'une tendance de long terme est significatif de celles où l'effet n'est pas significatif.

Une autre question que nous avons adressée aux données de mesures relève de leur capacité spatialement explicative. La construction de l'indicateur d'anomalie de variabilité a en effet réduit les données de débits à des séries temporelles, tandis que l'un des apports de la démarche de (ROCKSTROM et al. 2023) tient au caractère spatialement différencié des données qu'ils mobilisent, et donc de leurs résultats. Les données de débits issues de stations hydrométriques ne portent pas spontanément l'information sur l'emprise spatiale de la pression anthropique qu'elles mesurent. Aussi avons-nous suggéré de spatialiser les données ponctuelles de débits par le développement de ce que nous avons appelé un modèle hydrographique élémentaire. Une façon hydrologiquement cohérente de spatialiser de telles données est de reconstituer la zone d'apport associée à toute station de mesure (zone d'apport que nous avons appelé bassin versant, bien qu'elle puisse ne pas être hydrologiquement close dans le cas d'un enclassement entre zones d'apport sur un même réseau hydrographique, c'est-à-dire en situation de sous bassins versants enclassés). Cette méthode requiert le développement d'un modèle hydrographique numérique qui, à partir de toute position dans l'espace (la coordonnée spatiale d'une station), reconstitue automatiquement le sous bassin versant correspondant. Avec cette spatialisation des données de débits, nous avons proposé d'analyser l'influence de l'aménagement du paysage sur la variabilité observée aux stations selon une corrélation statistique entre le nombre d'éléments anthropiques sur les bassins versants et la variabilité du débit sur ces bassins versants. Cette démarche permet d'élargir le nombre des pressions anthropiques considérées par rapport au modèle utilisé par (ROCKSTROM et al. 2023), où seulement les barrages et les prélèvements pour l'irrigation sont pris en compte, tandis que la corrélation spatiale que nous avons conduite permet d'intégrer, par exemple, les modes d'occupation du sol.

Cette démarche de spatialisation des données de mesure pour l'analyse de l'influence de l'anthropisation des paysages, ainsi que la démarche d'analyse de significativité de l'influence de la tendance de long terme sur les débits, ne permettent toutefois pas de reproduire l'effet spatialement explicatif d'un modèle. C'est la raison pour laquelle nous avons souhaité proposer une méthode d'assimilation des données de mesure de débit aux données issues d'un modèle de simulation hydrologique global, tel le *Water Balance Model* utilisé dans (ROCKSTROM et al. 2023). La nature fondamentalement hydrologique des données à assimiler nous a conduit à proposer une amélioration des techniques de rééchantillonnage classiquement disponibles chez les

fournisseurs de traitements SIG. Un rééchantillonnage s'avère en effet nécessaire pour effectuer une comparaison directe (c'est-à-dire dans un domaine commun) de données définies ponctuellement dans l'espace (mesures de débits) avec des données définies spatialement par un maillage (modèle hydrologique global). Là où les techniques classiques se basent sur une hypothèse de continuité spatiale des phénomènes et de relation inverse à la distance, nous avons contraint la méthode d'interpolation à respecter le tracé hydrographique (*routing*) que le modèle hydrographique élémentaire nous permet de reconstituer pour toute zone d'étude. La relation inverse à la distance est également contrainte sous la forme d'une relation inverse à la charge hydraulique accumulée le long du réseau. Le modèle de rééchantillonnage hydrographique (MRH) que nous avons développé se structure en trois algorithmes : deux algorithmes sont dédiés à la projection des données depuis leur domaine spatial d'origine vers un domaine spatial cible (un algorithme dit de *rasterisation* pour les données de mesure et un algorithme dit de *dérasterisation* pour les données de modèle), et un algorithme d'interpolation qui reconstitue les valeurs de débits le long du réseau hydrographique modélisé. Ces trois algorithmes peuvent servir une diversité d'applications possibles en hydrologie, comme l'estimation de données manquantes à partir de données de mesure connues. Ici, nous les utilisons pour permettre la comparaison dans un même domaine cible, c'est-à-dire par superposition, des données de mesure avec des données de modèle.

Le chapitre suivant aborde la mise en œuvre numérique des deux propositions méthodologiques ainsi formulées. Il présente ainsi deux niveaux de résultats : les résultats liés à la mise en œuvre numérique en tant que telle (c'est-à-dire les développements logiciels et leurs performances), et les résultats liés à l'évaluation environnementale de l'intégrité des cours d'eau sur le territoire du SCoT du Sud-Loire.

Chapitre 4

Application numérique au territoire du SCoT Sud-Loire

Introduction

Le chapitre 3 a consisté à formuler deux propositions quant à la place des données de mesure dans une évaluation environnementale de l'intégrité des cours sur un territoire basée sur le cadre des frontières du système Terre. Ces deux propositions relèvent, d'une part, d'une méthode d'AESA par des données de mesure, qui a conduit à proposer un indicateur d'anomalie de variabilité et, d'autre part, de leur assimilation aux sorties d'un modèle hydrologique global dans un but de validation du modèle à une échelle locale. Ces deux propositions ont été présentées et formalisées, respectivement comme méthode d'évaluation absolue de la soutenabilité environnementale (MAESA) et comme modèle de rééchantillonnage hydrographique (MRH), mais sont restées pour le moment sans application numérique. Le présent chapitre entend procéder à l'application numérique de ces deux propositions, qui a été formalisée dans le développement de deux logiciels distincts, pour chacune des deux propositions. Ainsi le développement du logiciel **HydroTrendsFrance** constitue l'application numérique de la méthode d'AESA, et le logiciel **relio** celle du modèle MRH.

Dans une première partie (1), nous présenterons les données de mesure mobilisées dans le contexte du cas d'étude sur le territoire du SCoT du Sud-Loire. Cette partie sera l'occasion de revenir sur les grandes bases de données relatives à l'eau disponibles en France, puis de présenter dans le détail les données du réseau Vigicrues, qui seront utilisées dans ce travail d'application numérique. Les parties suivantes consisteront chacune en une mise en œuvre spécifique des don-

nées Vigicrues. La partie (2) en fera usage dans un but de validation du modèle hydrographique élémentaire (MHE) présenté au chapitre précédent. La partie (3) consistera à présenter l'application numérique de la méthode MAESA, c'est-à-dire son déploiement et ses résultats. La partie (4) en fera de même mais vis-à-vis du modèle MRH. Une dernière partie (5) dressera les conclusions générales du travail entrepris et proposera des perspectives de recherche.

4.1 Cas d'application : les données de mesure du réseau Vigicrues sur le territoire du SCoT du Sud-Loire

4.1.1 Identification des données disponibles

Une étape préalable à ce travail de développement logiciel est l'identification de données de mesure appropriées pour l'application numérique des deux propositions. Les données de mesure relatives à l'eau superficielle sur le périmètre de notre territoire d'étude ne sont pas produites par le SCoT du Sud-Loire lui-même, qui ne dispose ni de la compétence eau (réservée aux communes¹) ni n'intervient dans la politique de l'eau (déléguée aux comités de bassins et agences de l'eau depuis 1992, et dont le périmètre d'action est défini par leur bassin hydrographique²). Pour autant, disposer de la compétence eau ou intervenir dans la politique de l'eau ne se traduit pas nécessairement par être producteur de données sur l'eau. La production des données relatives à l'eau de surface en France est une compétence davantage portée par les services techniques de l'État (comme les DREAL, directions régionales de l'environnement, de l'aménagement et du logement) et de ce point de vue, relativement indépendante des périmètres d'action où la donnée produite se destine. Les données relatives à l'eau superficielle peuvent être distinguées selon qu'il s'agit de données portant sur la qualité de l'eau ou sur les quantités d'eau. Les propositions que nous avons développées dans le chapitre 3, de même que l'esprit du cadre des frontières du système Terre, envisagent la question de l'eau de surface dans une perspective essentiellement quantitative, et plus spécifiquement portée sur l'évaluation des débits moyens mensuels.

À notre connaissance, au moins quatre bases de données d'envergure nationale sur les eaux de surface existent : les bases Naïades, Ondes, BNPE et Vigicrues. La base Naïades concentre les données relatives à la qualité des eaux de surface, sur différentes thématiques, comprenant :

1. Voir : <https://www.collectivites-locales.gouv.fr/competences/leau-et-lassainissement>

2. Voir : <https://www.ecologie.gouv.fr/politiques-publiques/gestion-leau-france>

l'hydrobiologie, l'hydromorphologie, la sédimentologie et la structure végétale de la ripisylve, la physicochimie et les températures. Cette base ne dispose pas de données sur les débits et ne nous intéresse donc pas dans le cadre de ce travail. La base Ondes regroupe des données de mesure sur les étiages des petits et moyens cours d'eau. Il s'agit d'observations visuelles réalisées et reportées par les agents de l'OFB pendant l'été. Ces observations sont exclusivement qualitatives (e.g. "Écoulement visible acceptable", "Assec", etc) et non instrumentées ; elles ne nous intéressent donc pas non plus pour ce travail. La BNPE regroupe les données déclarées des organismes assujettis à une redevance au titre de leurs prélèvements en eau, qu'il s'agisse d'eau souterraine ou superficielle. Elle ne renseigne donc pas sur le débit d'écoulement des cours d'eau mais sur les volumes prélevés. Enfin, les données du réseau Vigicrues regroupent les données météorologiques et les données hydrométriques utilisées dans la prévision du risque inondation. Les données hydrométriques sont issues d'un réseau de stations de mesure installées en cours d'eau, mesurant principalement des hauteurs d'eau converties en débits par courbes de tarage, et disponibles en temps réel ou en archives sur des chroniques variables selon l'ancienneté des stations de mesure (CENTRE 2012). Ce sont ces données hydrométriques Vigicrues que nous allons utiliser.

4.1.2 Les données hydrométriques du réseau Vigicrues

Cette section présente les données hydrométriques du réseau Vigicrues que nous allons utiliser pour l'application numérique de nos propositions de recherche.

4.1.2.1 Le réseau Vigicrues

Les données dites Vigicrues sont de plusieurs types : celles qui portent sur la météorologie et celles qui portent sur l'hydrométrie. Vigicrues est le nom du réseau qui opère la vigilance crue en France depuis 2003, après que les graves crues survenues en 1999 et 2000 ont poussé l'État à restructurer les services départementaux d'annonce des crues³. Ce réseau est constitué du service central d'hydrométéorologie et d'appui à la prévision des inondations (Schapi), des 17 services de prévision des crues (SPC) et des 19 unités d'hydrométrie (UH)⁴. Les SPC et les UH sont rattachés aux DREAL bien que leur périmètre d'action ne soit pas défini par les régions

3. Voir : <https://www.centre-val-de-loire.developpement-durable.gouv.fr/le-reseau-vigicrues-fete-ses-20-ans-a4720.html>

4. Voir : https://www.vigicrues.gouv.fr/uploads/bilan_annuel_vigilance.pdf

administratives mais par les bassins hydrographiques. Chacun des six bassins hydrographiques dispose ainsi d'un Schéma Directeur de Prévision des Crues (SDPC). Ces schémas décrivent les modalités d'intervention liées à la gestion du risque, mais règlent également la transmission de l'information entre les différents acteurs compétents et décrivent les dispositifs de mesure existant sur leur bassin. Le bassin hydrographique sur lequel se situe le territoire du SCoT Sud-Loire est principalement le bassin Loire-Bretagne et, dans une moindre mesure, le bassin Rhône Méditerranée-Corse. Dans le SDPC Loire-Bretagne⁵, qui regroupe quatre SPC et est administré par la DREAL Centre, il est indiqué que le bassin dispose d'environ 1400 stations de mesure appartenant à la DGPR (direction générale de la prévention des risques, à laquelle le Schapi est rattaché) qui sont opérées par les SPC et les UH. D'autres données sont mises à disposition du SPC par l'intermédiaire de partenariats locaux avec les gestionnaires d'ouvrages, les syndicats des eaux, les sociétés d'aménagement régional pour les eaux, les bassins versants expérimentaux gérés par des laboratoires de recherche, les collectivités, etc. La Métropole de Saint-Étienne met ainsi à disposition les données issues de son instrumentation sur les rivières du Furan et de l'Ondaine.

Concernant la transmission des données entre services, le Schapi coordonne leur mise à disposition dans une base de données commune, appelée PhyC (Plateforme HYdro Centrale) et accessible via le site HydroPortail ou l'API hubeau. La base PhyC a commencé à être constituée en 2017 comme regroupement des deux bases gérées indépendamment par les SPC et les UH⁶. Les SPC disposent alors d'environ 1800 stations de mesure sur le territoire national, et les UH d'environ 3500, mesurant toutes principalement des hauteurs d'eau converties en débits. L'enjeu de la fusion est alors de disposer d'une base élargie et unifiée, notamment sur le plan des outils logiciels et des modèles hydrologiques employés par chaque service.

4.1.2.2 La production des données hydrométriques Vigicrues

Les données hydrométriques Vigicrues proviennent d'une chaîne de traitements en allers-retours (voir Figure 4.1) opérée par les SPC et les UH, partant de la collecte de la donnée par une instrumentation automatisée en cours d'eau ("donnée brute") jusqu'au stade final de validation-correction et son archivage dans la base PhyC ("donnée validée")⁷.

5. Téléchargeable ici : <https://www.centre-val-de-loire.developpement-durable.gouv.fr/presentation-du-spc-loire-allier-cher-inredre-a3116.html>

6. Voir : <https://professionnels.ofb.fr/fr/node/39>

7. Voir : <https://professionnels.ofb.fr/fr/node/39>

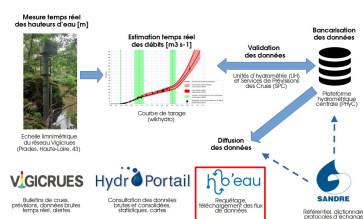


Figure 4.1 – Le parcours de production des données hydrométriques Vigicrues au sein du Système d'Information sur l'Eau (SIE)

À l'étape de la collecte, les données sont généralement récupérées à l'aide d'une échelle limnimétrique (voir Figure 4.2) qui mesure la hauteur d'eau depuis le bas de l'échelle (il s'agit d'une cote). Les débits peuvent être mesurés directement, mais plus rarement, car le matériel est plus onéreux et plus fragile, grâce à deux méthodes au moins : le moulinet hydrométrique qui intègre directement les vitesses ou l'ADCP (profileur acoustique de vitesse par effet Doppler)⁸. Ces données sont ensuite envoyées directement dans la base PhyC avec le statut de "donnée brute".



Figure 4.2 – Échelle limnimétrique du réseau Vigicrues installée sur l'Allier (Prades, Haute-Loire, 2024) (c) Gondran N.

Les données brutes sont ensuite corrigées par les "prévisionnistes" au sein des SPC, qui les déposent dans la PhyC avec le statut de "donnée corrigée". Les données corrigées sont ensuite

8. Voir l'article "Courbe de tarage : extrapolation par modélisation 1D et cartographie par couplage SIG" sur <http://wikhydro.developpement-durable.gouv.fr>

expertisées par les hydromètres des UH (notamment correction des doublons) puis à nouveau déposées dans la PhyC avec le statut "donnée pré-validée". Les données acquièrent finalement le statut "donnée validée" l'année suivant leur dépôt, lorsque les hydromètres effectuent une vérification de la cohérence des données sur l'ensemble de l'année hydrologique. À l'issue de ce parcours, la donnée disponible n'a pas changé de grandeur, il s'agit toujours de la grandeur relevée par l'instrument de mesure en cours d'eau. Ce parcours est avant tout un parcours de validation de la donnée. La conversion des hauteurs d'eau en débits est effectuée selon un autre parcours.

Les hydromètres des UH ont en charge l'opération de jaugeage, qui consiste à convertir les hauteurs d'eau en débits à l'aide de courbes de tarage et courbes de correction, que les hydromètres construisent et entretiennent à partir de relevés de terrain (morphologie, bathymétrie des profils en travers, etc). Une fois les courbes de tarage validées, qui sont déposées dans la PhyC et consultables sur HydroPortail, la conversion des hauteurs d'eau en débits s'effectue automatiquement pour chaque station directement dans la PhyC. Les données de débits ainsi calculées sont déposées avec le statut de "donnée corrigée". Elles subissent donc ensuite le même parcours de validation décrit plus haut que les données de hauteur d'eau.

Les données hydrométriques ainsi produites sont ensuite diffusées selon trois modalités principales :

- via le site internet Vigicrues, où les données sont présentées de façon élaborée sous la forme de bulletin de prévision de crues, et mises à disposition de façon brute pour la consultation mais pour des chroniques d'une durée de quelques semaines seulement ;
- via l'HydroPortail, qui permet essentiellement la consultation des données brutes et consolidées sur toute la durée des chroniques disponibles pour chaque station, mais sans possibilité de les télécharger à la demande pour des chroniques et des stations spécifiques définies par l'utilisateur ;
- via l'API (*Application programming Interface*) hubeau qui permet le requêtage dynamique des données hydrométriques brutes et consolidées selon plusieurs clés (emprise spatiale, code station et couverture temporelle notamment) et le téléchargement des flux de données associés. C'est cette API que nous utiliserons pour la présente application numérique, accessible à partir du *endpoint* <https://hubeau.eaufrance.fr/api/v1/hydrometrie/>.

4.1.2.3 Les incertitudes liées aux données hydrométriques Vigicrues

Si les données sont disponibles avec le statut "validé" dans la PhyC (et accessibles via HydroPortail) cela ne signifie pas pour autant qu'il s'agisse de données "exactes". En effet, le parcours de validation que nous avons décrit plus haut n'effectue pas une correction des incertitudes liées à la production de la donnée hydrométrique, mais une validation de la cohérence sur l'ensemble de la chaîne de production de la donnée. Les incertitudes entourant les valeurs de débits enregistrées dans la PhyC peuvent être distinguées en deux types : les erreurs liées à la mesure des hauteurs d'eau, et les erreurs liées au jaugeage⁹.

Les études de qualité menées par les services de l'État ont montré que l'erreur liée à la mesure des hauteurs d'eau est négligeable devant l'erreur liée à l'estimation du débit (CEREMA 2015c). En effet, selon cette même source, l'erreur moyenne pour la mesure de la hauteur d'eau est comprise entre ± 1 cm en régime normal, et atteint ± 10 cm lors de crues, où l'effet l'écoulement turbulent produit un effet de batillage (vaguelettes qui perturbent la lecture sur l'échelle limnimétrique) et où la montée rapide des eaux (ou la décrue rapide) oblige à considérer une valeur moyenne entre deux lectures. Concernant l'erreur de jaugeage, celle-ci varie selon la technique employée, de $\pm 3\%$ pour les jaugeages au Doppler jusqu'à $\pm 10\%$ pour les jaugeages disposant uniquement des vitesses de surface (moulinet, etc).

L'opération de jaugeage consiste en une campagne de mesures des débits et des hauteurs d'eau sur un même site. Un échantillon de couples (débits, hauteurs) est ainsi obtenu. Une courbe est tracée à partir de cet échantillon par interpolation : il s'agit de la courbe de tarage. La courbe de tarage est également extrapolée pour les valeurs de débits très fortes qui n'ont pas pu être mesurées (crues). Sur le bassin hydrographique Rhône-Méditerranée-Corse par exemple, seulement 9% des stations ont un jaugeage pour de hauteurs d'eau de temps de retour de 10 ans (CEREMA 2015b). Au moins deux grands types d'extrapolation peuvent être distingués, selon la disponibilité des données sur le site : la modélisation hydraulique à partir des équations de Chézy et Manning-Strickler d'une part, ou l'extrapolation logarithmique d'autre part (CEREMA 2015b; CEREMA 2015a).

9. Voir l'article "Incertitudes de la mesure du couple (hauteur – débit) lors du jaugeage" sur <http://wikhydro.developpement-durable.gouv.fr/>

4.2 Validation du modèle hydrographique élémentaire

Cette partie présente la démarche entreprise pour la validation du MHE à partir, notamment, des données Vigicrues, et les résultats obtenus.

4.2.1 Remarque préalable : modélisation réaliste *versus* modélisation constructiviste

Si la question de la "validation" du MHE semble intuitive, entendue comme vérification de la bonne correspondance entre les sorties du modèle et la réalité, c'est-à-dire une validation de la proximité entre le réseau hydrographique réel et le réseau modélisé, la remarque qui suit ajoute un peu d'ambiguïté et de nuances à cette acception immédiate.

La démarche *réaliste* de modélisation entend rendre compte de l'intégralité du réseau hydrographique tel qu'il existe *réellement*, c'est-à-dire, par exemple, représenter tous les tronçons où l'écoulement est permanent, depuis l'exutoire jusqu'aux têtes de bassin. La démarche que nous appelons *constructiviste* en revanche, n'entend pas proposer le reflet fidèle de la réalité mais ambitionne de reconstruire ce qui, dans la réalité, intéresse le modélisateur. C'est dans cette seconde démarche que nous nous situons. En effet, le MHE que nous proposons pourrait viser à reconstruire fidèlement le réseau hydrographique réel selon une valeur appropriée du paramètre qui contrôle le seuil minimum d'accumulation. Mais ce paramètre peut permettre d'optimiser deux autres aspects de la modélisation :

- la position géographique des stations de mesure, c'est-à-dire leur conversion en exutoires hydrographiques (i.e. projeter la position géographique de la station dans l'espace réel vers l'espace hydrographique modélisé)
- la caractérisation du tronçon, c'est-à-dire l'identification du degré d'importance du tronçon dans le réseau (les stations de mesure sur un territoire sont généralement implantées sur les cours d'eau dits "majeurs" ou "importants").

Par ailleurs, la délimitation des bassins versants associés aux stations est indépendante de la représentation exhaustive du réseau hydrographique : le calcul est fondamentalement topographique. La ligne de partage des eaux est en effet obtenue à partir des pentes comparées, et ainsi le bassin versant d'un tronçon amont sera nécessairement compris dans le bassin versant du tronçon aval sans pour autant que ce tronçon aval soit représenté.

Il nous faut alors bien établir ce qui doit faire l'objet d'une validation. Dans cette section, nous proposons d'effectuer une validation très généraliste du MHE en s'intéressant seulement au raster des bassins versants en sortie de modèle. La validation du réseau hydrographique, c'est-à-dire le tracé du chevelu, demande de procéder paramètre par paramètre et porte plus spécifiquement sur, par exemple, la validation de l'emplacement des stations de mesure dans le réseau, la position des exutoires, la charge hydraulique relative le long du chemin hydrographique, etc. Nous aurons l'occasion de revenir en détail sur les paramètres du MHE dans la partie (3) du présent chapitre.

Par ailleurs, si les sorties de modèle appellent spontanément une validation, la validation peut aussi bien porter sur la méthode elle-même. À cet égard, nous n'avons pas développé une méthode propre à ce travail de thèse mais nous nous inscrivons dans la lignée des travaux de (O'CALLAGHAN et al. 1984) et (KARSSENBERG et al. 2010), dont nous reprenons respectivement la méthodologie en trois temps pour la modélisation du réseau hydrographique, et les algorithmes pour le traitement du MNT.

4.2.2 Procédure de validation

Plusieurs techniques existent dans la littérature pour valider le tracé d'un réseau hydrographique modélisé. Leur point commun est de s'appuyer sur les tracés issus de sources externes supposées validées. Il s'agit de considérer un ou plusieurs tracés de référence, dont l'emprise peut ne représenter qu'une partie de l'emprise effectivement étudiée, et de les comparer, selon différentes métriques, au tracé modélisé. Parmi ces métriques, on peut citer celles qui sont utilisées par (VOROSMARTY et al. 2000; FEKETE et al. 2002). Ces deux publications proposent respectivement une validation du réseau hydrographique global STN-30p (simulated topological network 30 arcmin) et une calibration des débits exutoires du modèle hydrologique global WBM (Water Balance Model) sur les données de jaugeage du GRDC (global runoff data center). La publication de (VOROSMARTY et al. 2000) décrit deux métriques pour l'évaluation de la validité du réseau hydrographique modélisé (sa correspondance avec le réseau "réel", du moins avec des réseaux locaux eux-mêmes validés). D'une part, la comparaison de la longueur entre des tronçons de référence et leurs tronçons associés dans le réseau modélisé; d'autre part, la comparaison entre la superficie des bassins versants de référence et leurs bassins associés dans le réseau modélisé. Quant à la publication de (FEKETE et al. 2002), elle propose des métriques

pour la validation puis la calibration de la jointure entre la position réelle des stations et leur position dans le réseau hydrographique modélisé (cette jointure est appelée *coregistration*).

Comme nous le mentionnions plus haut, nous comptons valider le MHE en validant d'abord la délimitation des bassins versants qu'il produit en sortie. Pour cette validation, en s'inspirant de (VOROSMARTY et al. 2000), nous allons utiliser la métrique de l'emprise comparée des bassins versants, entre la couche produite en sortie par le MHE, et une couche issue d'un producteur externe et supposé fiable. Deux démarches sont envisageables : (1) soit nous forçons le MHE à reproduire les bassins versants fournis par un producteur externe, (2) soit il existe une couche fournie par un producteur externe qui reproduise les bassins versants générés par le MHE.

À notre connaissance, les jeux de données existants sur les bassins versants du territoire français sont issus soit de la BD Topage soit de la BD Carthage. La première BD constitue en quelques sortes l'évolution de la seconde en cela qu'elle effectue le passage à grande échelle des entités hydrographiques du système d'information sur l'eau (SIE), dans le but de les rendre compatibles avec les entités cartographiques du référentiel national à grande échelle (RGE) de l'IGN. Cette mise en compatibilité est également une mise en conformité avec les objectifs d'uniformisation des référentiels européens portés par la directive européenne INSPIRE. À sa création en 2017, la BD Topage est ainsi produite de façon à permettre l'identification des entités hydrographiques similaires entre la BD Carthage et la BD Topo. La BD Topage propose un produit appelé "bassins versants topographiques", dont la documentation (OFB et al. 2020) signale qu'ils correspondent "pour l'instant" aux "zones hydrographiques" de la BD Carthage, laquelle propose le découpage du territoire national en "aires hydrographiques", selon cinq niveaux d'agrégation dits gigognes parce qu'emboîtés ("bassin" > "région" > "secteur" > "sous-secteur" > "zone") (IGN 2006). Les bassins versants délimités par ces deux producteurs ne correspondent donc pas aux bassins versants associés à un réseau de stations de mesure (en particulier le réseau Vigicrues).

La démarche (1) est alors à envisager : comment reproduire les bassins versants des BD Topage et Carthage à partir de notre modèle hydrographique ? Il faut pour cela disposer des exutoires associés aux bassins Topage et Carthage, qui seront fournis en entrée à notre modèle. Les deux BD fournissent un produit appelé "nœuds hydrographiques" mais qui s'avère impossible à lier aux bassins versants fournis par les mêmes BD, car les deux produits ne partagent aucun attribut commun pour envisager une jointure attributaire.

Si ces deux BD ne sont pas directement exploitables pour effectuer la validation qui nous intéresse, le BRGM et l'OFB se sont associés en 2014 pour monter un pôle de recherche commun sur l'interopérabilité du SIE entre les différents producteurs nationaux, en particulier les BD Carthage et BD Lisa (relative aux aquifères), appelé pôle INSIDE¹⁰. Cette initiative s'est traduite en 2017 par la création d'un service web de géotraitements portant sur chacune des deux BD Carthage et Lisa. Pour ce qui est de la BD Carthage, les services de géotraitements proposés comportent : une délimitation automatique des bassins versants à partir d'une couche de points fournie en entrée, la reconstruction du parcours amont ou aval à partir d'un point du réseau, la reconstruction du parcours entre deux points fournis en entrée, l'identification du cours d'eau le plus proche d'un point fourni en entrée. Ces services sont accessibles au travers d'une interface graphique¹¹ ou d'une requête directe sur serveur par le protocole de communication WPS¹². Pour ce qui est du service "délimitation des bassins versants" (WatershedWPS), si l'interopérabilité n'est guère mise en exergue dans le service proposé (il s'agit d'une délimitation des bassins versants sans outil de jointure avec d'autres produits du SIE), ce service présente l'avantage de générer les bassins versants qui nous intéressent, c'est-à-dire les bassins versants associés aux stations du réseau Vigicrues. L'interopérabilité proposée ne va pas jusqu'à faire le lien entre ces deux bases (réseau hydrographique Carthage et stations Vigicrues), et la position des stations doit être rentrée manuellement. Ce service permet ainsi d'envisager la validation de notre modèle hydrographique selon la démarche (2).

En pratique, la communication avec le serveur WPS s'avère ardue¹³. Nous n'avons pas eu le choix que d'utiliser l'interface graphique du démonstrateur, pour laquelle nous avons décidé de ne délimiter que le bassin versant couvrant notre zone d'étude, c'est-à-dire en fournissant au démonstrateur la position de la station la plus à l'aval sur la Loire, située en sortie du barrage de Villerest (voir Figure 4.3). On fournit alors à notre modèle cette même position afin de pouvoir comparer la superficie des bassins versants obtenus selon l'un et l'autre des modèles

10. Voir en ligne : <https://www.pole-inside.fr/fr/page/propos>

11. Voir le service GUI : <https://reseau.eaufrance.fr/geotraitements/viewer/bdcarthage> et <https://reseau.eaufrance.fr/geotraitements/viewer/bassinversant>

12. Pour les bassins versants, le service WPS se nomme WatershedWPS et la documentation est accessible via : <https://github.com/INSIDE-information-systems/WatershedWPS>

13. Les tentatives en GET et POST pour une requête `request=Execute` se sont révélées infructueuses, et ceci pour des syntaxes URL et XML, ainsi qu'en utilisant la librairie Python `pyWPS` développée par l'OGR. Notamment, alors que la requête `request=DescribeProcess` est censée retourner la liste des configurations possibles pour la délimitation des bassins versants (c'est-à-dire les différentes configurations paramétriques permettant de convertir le MNT utilisé en raster d'accumulation de flux puis en bassins versants), elle renvoie une liste vide, suggérant que le serveur n'est peut-être plus entretenu.

(MHE et WatershedWPS). Remarquons qu'un traitement préalable doit être effectué sur le bassin versant fourni par le démonstrateur, qui consiste à couper la couche selon l'emprise de notre zone d'étude. En effet, le service INSIDE est proposé sur le périmètre national (France métropolitaine et quelques DROM-TOM), ce qui implique que les bassins versants tracés à partir d'un point remontent nécessairement jusqu'en tête de bassin (pour les bassins circonscrits au territoire national, ce qui est le cas du fleuve Loire). Notre modèle en revanche fonctionne sur un périmètre infranational. Non pas qu'il ne pourrait pas prendre en entrée un périmètre national, mais les capacités de calcul de la machine utilisée contraignent l'utilisation d'un périmètre plus restreint.

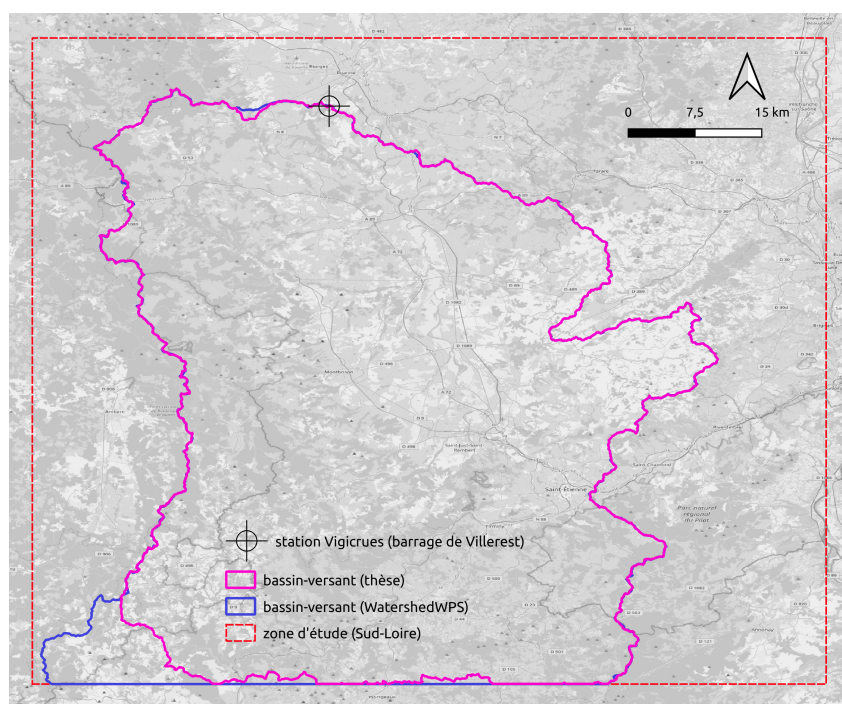


Figure 4.3 – Le bassin versant de la Loire à partir de la station Vigicrues du barrage de Villerest obtenu par le MHE (rose) et par le modèle WatershedWPS (bleu)

Les surfaces des bassins versants calculées à partir des deux modèles sont très proches l'une de l'autre : le bassin versant obtenu à partir du modèle WatershedWPS est 3.5% plus grand que celui obtenu à partir du modèle MHE (4660 km² contre 4500 km²). Ce constat est d'autant plus intéressant que nous ne savons rien de la paramétrisation ni des données d'entrées utilisées par WatershedWPS : on ne connaît pas la résolution ni le producteur du MNT utilisé, on ne connaît pas la valeur des paramètres utilisés pour générer le raster de direction de flux, et on ne connaît pas le seuil minimum d'accumulation ayant servi à tracer le réseau hydrographique. Remarquons également que la documentation de WatershedWPS décrit succinctement le fonctionnement de l'algorithme de projection du point positionné arbitrairement dans l'espace par l'utilisateur

vers le réseau hydrographique modélisé : "le point envoyé en entrée est ramené (*snappé*) sur le réseau hydrographique modélisé, en suivant la plus forte pente jusqu'à retrouver un tronçon". Malgré la différence de construction de cet algorithme et le nôtre, les résultats obtenus sur le bassin versant de la station de Villerest sont très proches (à cette échelle d'étude). Notons que leur algorithme est plus réaliste sur le plan physique que le nôtre puisque, au lieu de ne tenir compte que de la distance entre le point et les pixels du réseau, il intègre une contrainte de pente. En revanche, lors du traitement de plusieurs points en même temps, leur algorithme ne prévoit pas, d'après la description succincte qui en est donnée, la gestion des attributions multiples. Il serait bienvenu toutefois d'effectuer cette validation sur d'autres bassins versants.

Nous allons à présent mobiliser le MHE dans les deux applications numériques (et logicielles) à suivre, `HydroTrendsFrance` et `relio`, qui mettent respectivement en œuvre la méthode AESA et le modèle MRH à partir des données Vigicrues.

4.3 Évaluer l'intégrité des cours d'eau sur un territoire par les données de mesure : le logiciel `HydroTrendsFrance`

`HydroTrendsFrance` est un logiciel libre, sans interface graphique, disponible sur la plateforme `github`¹⁴. Il a été développé dans le cadre de ce travail de thèse afin d'effectuer une évaluation territoriale de l'intégrité des cours d'eau basée sur la version (ROCKSTROM et al. 2023) du cadre des frontières du système Terre. `HydroTrendsFrance` implémente une série de fonctions qui permettent :

- d'extraire les données des stations de mesure du réseau Vigicrues pour un territoire d'intérêt et de les exporter comme tableurs ou comme couches vectorielles cartographiques
- d'effectuer une évaluation de l'intégrité des cours d'eau du territoire basée sur l'esprit de l'indicateur conçu par (ROCKSTROM et al. 2023) en mobilisant un indicateur d'anomalie de variabilité

`HydroTrendsFrance` a été développé dans le langage de programmation Python. Il peut théoriquement traiter toute zone géographique de la France métropolitaine, bien qu'il ait été calibré pour le territoire du Sud-Loire. En conséquence, pour des territoires ayant des

14. Voir en ligne : <https://github.com/Quentin-Dassibat/HydroTrendsFrance>

profils hydrologiques éloignés du profil "plaine" et "moyenne montagne", des défauts hydro-morphologiques peuvent apparaître (notamment pour la délimitation des bassins versants du territoire). HydroTrendsFrance peut théoriquement traiter toute plage temporelle dans la limite de l'étendue temporelle des données mises à disposition par les stations de mesure du réseau Vigicrues.

4.3.1 Description du logiciel HydroTrendsFrance : application numérique de la méthode d'évaluation environnementale

Le logiciel HydroTrendsFrance fonctionne à partir de deux paramètres d'entrées principaux : l'emprise géographique du territoire étudiée (qui peut être fournie sous la forme d'un raster ou d'une couche vectorielle) et une plage temporelle définie par une date de début et une date de fin. Le logiciel fait ensuite appelle à quatre jeux de données principaux pour produire ses résultats en sortie (voir Figure 4.4).

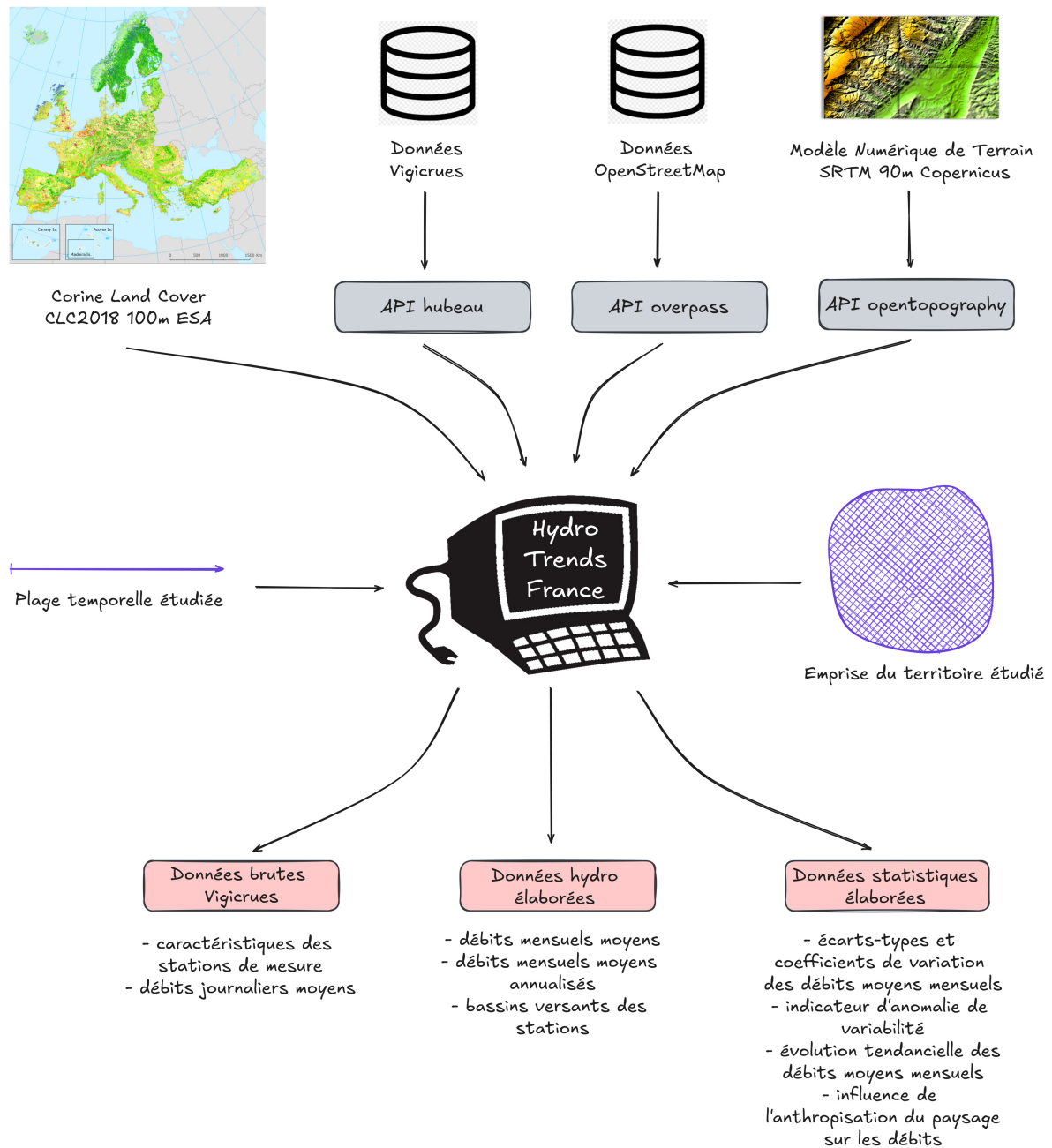


Figure 4.4 – Schéma de fonctionnement du logiciel HydroTrendsFrance

4.3.1.1 Les données fournies en sortie

HydroTrendsFrance propose deux types de sorties. D'une part, une mise à disposition des données Vigicrues récupérées selon les paramètres d'entrée et, d'autre part, une production de données plus élaborées (voir Figure 4.4) :

- données Vigicrues récupérées :
 - La liste des stations de mesure disponibles sur l'emprise géographique fournie en entrée et leurs principales caractéristiques (date d'ouverture, de fermeture, position

géographique)

- Les débits journaliers moyens de toutes les stations qui couvrent la plage temporelle appelée
- données hydrométriques et hydrographiques élaborées par **HydroTrendsFrance** :
 - Les débits mensuels moyens de toutes les stations qui couvrent la plage temporelle appelée
 - Les débits mensuels moyens annualisés¹⁵ de toutes les stations qui couvrent la plage temporelle appelée
 - Les bassins versants associés aux stations de mesure identifiées, en distinguant les bassins versants de toutes les stations identifiées et les bassins versants des seules stations qui couvrent la plage temporelle appelée et renvoient la métrique des débits journaliers moyens (certaines stations Vigicrues ne sont pas jaugées et renvoient seulement des hauteurs d'eau)
- données statistiques pour l'AESA élaborées par **HydroTrendsFrance** :
 - les écarts-types et coefficients de variation associés au débit moyen mensuel moyen de chaque de station de mesure et pour chaque mois
 - l'indicateur d'anomalie de variabilité pour chaque station et chaque mois
 - l'évolution tendancielle sur la période appelée des débits moyens mensuels et de leur variabilité pour chaque station de mesure (amplitude de la tendance et niveau de significativité basés sur une régression de Mann-Kendall)
 - l'influence de l'occupation des sols et de l'anthropisation des paysages sur les débits à l'échelle du territoire étudié, c'est-à-dire toutes stations confondues (coefficient de corrélation de Pearson et niveau de significativité)

Ces données sont disponibles dans deux formats de fichier différents :

- Le format tableur en `.csv`
- Le format vectoriel cartographique en `.gpkg`
 - comme couche de géométrie `Point` : les points correspondent à la position géographique `<x,y>` des stations de mesure

15. Le débit mensuel moyen annualisé est la moyenne sur plusieurs années du débit d'un même mois. Par exemple, la moyenne du débit mensuel de tous les mois de janvier de la période donne le débit mensuel moyen annualisé du mois de janvier.

- comme couche de géométrie **Polygon** : les polygones correspondent à l'emprise des bassins versants associés à chaque station de mesure¹⁶

4.3.1.2 Origine des données d'entrée

Quatre jeux de données d'entrée sont mobilisés par **HydroTrendsFrance** : les données de mesure de débits du réseau Vigicrues bien sûr, mais aussi le modèle numérique de terrain de la zone étudiée qui est utilisé par le MHE pour la spatialisation des données, ainsi que les données *Open Street Map* et Corine Land Cover pour les traitements statistiques relatifs à l'anthropisation des paysages (voir Figure 4.4).

Les données Vigicrues sont requêttables via l'API `hubeau.france` et son *endpoint* `hubeau.eaufrance.fr/api/v1/hydrometrie/`, station par station. Plusieurs agrégations temporelles (appelées "grandeurs hydro" dans le référentiel) sont disponibles en principe : les relevés "temps réel" (pas de temps horaire), les relevés journaliers moyens (débit moyen journalier) et les relevés mensuels moyens (débit moyen mensuel). En pratique et au moment du développement initial du présent protocole numérique de récupération des données (2022), les débits moyens mensuels n'étaient pas disponibles ; les scripts ont donc été construits à partir des débits moyens journaliers. D'autres informations sont disponibles sur les stations, comme la date d'ouverture et, le cas échéant, la date de fermeture, ainsi que la position géographique (coordonnées spatiales $\langle x, y \rangle$ projetées en WGS84).

Les données topographiques servant à délimiter les bassins versants sont récupérées sous la forme d'un MNT (modèle numérique de terrain) qui indique, à une résolution de pixel donnée, l'altitude du pixel normalisé au niveau de la mer. Le MNT correspondant à l'emprise étudiée est récupéré via l'API `opentopography.org` qui est un projet open source de mise à disposition des données fournies par différents producteurs. Nous récupérons dans notre cas le produit SRTM-90 produit par Copernicus (centre de ressources de l'Agence spatiale européenne). Il s'agit d'un MNT au format raster, à une bande, de résolution 90m et projeté en WGS84. L'API requiert l'utilisation d'une clé d'authentification utilisateur mais qui est réutilisable pour toute nouvelle requête.

16. La position géographique de la station de mesure est considérée comme l'exutoire du bassin versant associé, lequel est construit par le modèle hydrographique élémentaire.

Les données ponctuelles d'anthropisation sont récupérées via la librairie OSMnx de Python qui propose une fonction de récupération des données OSM (*Open Street Map*) en effectuant une requête à l'API `overpass-api.de`¹⁷. OSM structure ses données cartographiques en un ensemble de clés (*keys*) et pour chaque clé un ensemble d'entités (*features*). À partir de la documentation OSM Map Features¹⁸ qui décrit les clés et les entités associées, on récupère ainsi les données suivantes :

- la position des stations d'épuration : clé = 'man_made'; entité = 'wastewater_plant'
- la position des lacs de retenue : clé = 'water'; entité = 'reservoir'
- la position des barrages : clé = 'waterway'; entité = 'dam'
- la position des seuils en rivière : clé = 'waterway'; entité = 'weir'
- la position des écluses : clé = 'water'; entité = 'lock'

Notons que les données OSM ont été retenues pour leur facilité d'intégration dans le logiciel HydroTrendsFrance, grâce à l'API flexible `overpass-api.de` qui permet de les requêter. Cependant, des réserves peuvent être émises à leur égard concernant leur fiabilité. En effet, les données OSM sont issues d'une collecte sur des dépôts officiels et d'une cartographie participative accessible et renseignable par tout utilisateur de la communauté. La provenance des données n'est pas toujours indiquée dans les tableaux de données fournis en sortie de requête. Pour ce qui est des éléments que nous avons récupérés ici (les entités ponctuelles d'anthropisation des paysages ayant une influence sur l'écoulement), certaines entités sont attribuées à la source Sandre, qui est le référentiel national des données du Système d'Information sur l'Eau. Le jeu de données sous-jacent dans Sandre s'appelle Obstacles à l'écoulement. Une comparaison systématique entre ce jeu de données et les données OSM que nous avons récupérées gagnerait à être menée pour s'assurer de la robustesse des données mobilisées.

Enfin, le jeu de données utilisé pour l'occupation du sol est le produit Corine Land Cover 2018 (CLC18) de l'Agence Spatiale Européenne. Il s'agit d'une image raster de résolution 100m qui décompose l'utilisation du sol en 44 valeurs de pixel. Le jeu de données est téléchargé pour l'Europe entière puis coupé par intersection avec notre zone d'étude. Une démarche plus fine mais qui n'a pas été implémentée consiste à effectuer une requête sur l'API CLMS de Copernicus directement sur le produit et la zone d'étude souhaitée¹⁹.

17. Voir : https://wiki.openstreetmap.org/wiki/Overpass_API

18. Voir : https://wiki.openstreetmap.org/wiki/Map_features

19. Voir la documentation : https://github.com/eea/clms-api-docs/blob/develop/source/notebooks/clms_portal.ipynb

4.3.2 Résultats

Nous avons utilisé `HydroTrendsFrance` pour extraire les données de trois chroniques différentes sur le territoire du SCoT Sud-Loire :

- 2000-2020 : car il s'agit de la chronique utilisée par (ROCKSTROM et al. 2023). Pour cette chronique, le fichier `log.txt` produit par `HydroTrendsFrance` lors de toute exécution du script, nous renseigne sur le nombre de stations récupérées : 102 stations ont été localisées dans l'emprise géographique demandée, parmi lesquelles 47 satisfont la contrainte temporelle 2000-2020, et parmi celles-ci 42 renvoient la métrique appropriée (débit moyen journalier). Notons que l'emprise géographique demandé a été définie à tâtons en agrandissant progressivement l'enveloppe du périmètre du SCoT Sud-Loire jusqu'à ce que les sous bassins versants (ou zones d'apport) associés à chacune des stations hydrométrique identifiées apporte une couverture hydrographique satisfaisante. En effet, en se limitant aux stations hydrométriques "intra-territoriales", la couverture hydrographique assurée ne représente que les deux tiers de l'emprise territoriale. Nous avons également inclus les stations "extra-territoriales" limitrophes. Ainsi, parmi les 42 stations que nous évoquons plus haut, nous en retenons 13 dites "intra-territoriales" (représentées en jaune dans la Figure 4.5) et 8 dites "extra-territoriales" (représentées en rouge dans la figure 4.5) pour optimiser la couverture hydrographique ;
- 1950-2020 : car il s'agit d'une chronique de 70 ans, durée pour laquelle nous avons constaté que la statistique de Mann-Kendall est généralement significative. Ceci nous permet d'étudier l'effet du forçage climatique sur les débits. Une période plus longue aurait également permis d'assurer cette significativité mais le nombre de stations ouvertes avant 1950 diminue drastiquement. Pour cette période, 5 stations proposent les données demandées.
- 1922-2023 : car il s'agit de la chronique la plus longue sur le territoire pour laquelle 2 stations disposent de données, celle de Bas-en-Basset ouverte en 1918 et celle de Villerest ouverte en 1921. Ceci nous permet d'évaluer l'effet du forçage climatique sur une chronique de plus d'un siècle.

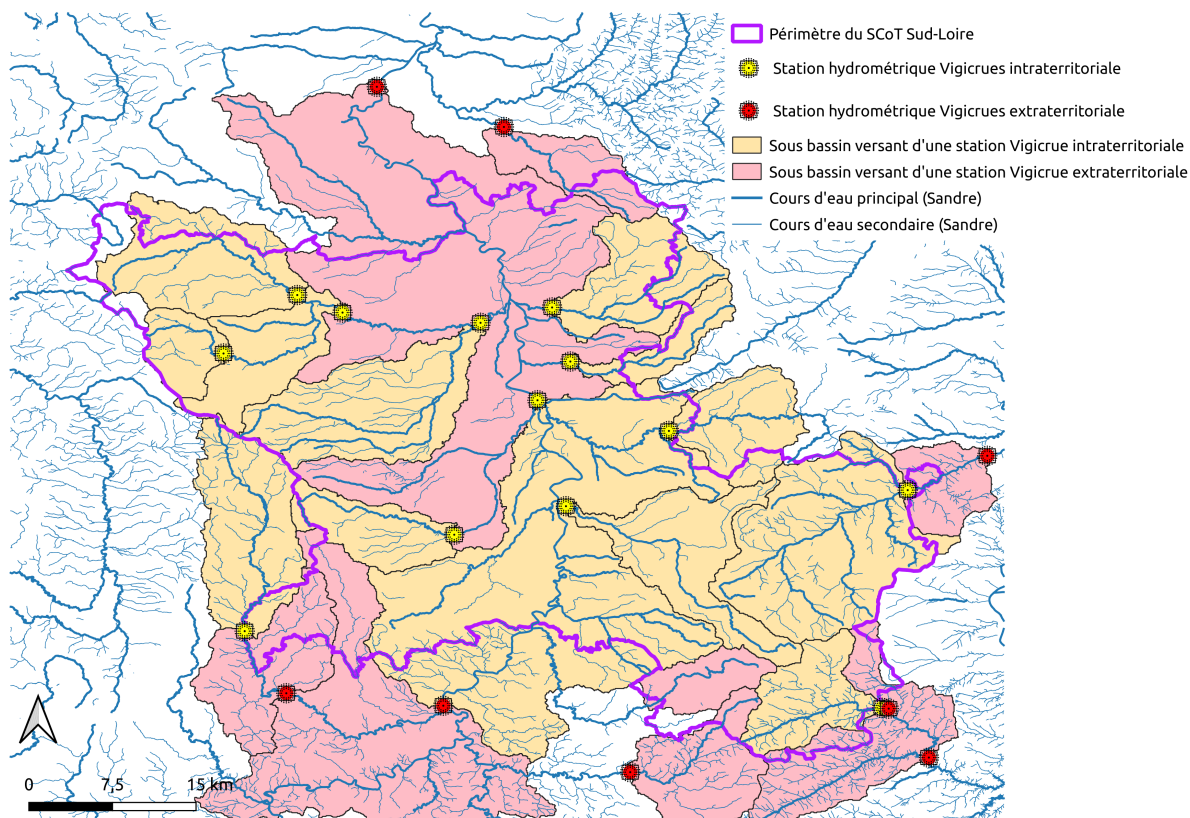


Figure 4.5 – Localisation des stations hydrométriques utilisées pour la chronique 2000-2020.

4.3.2.1 Validité de l'indicateur d'anomalie de variabilité

Remarques sur la notion de variabilité

Pour comprendre les raisons qui nous ont conduit à définir un indicateur d'anomalie de variabilité comme méthode d'évaluation absolue de la soutenabilité environnementale à partir de données de mesure, nous allons commencer par revenir sur la discussion entamée au chapitre 3 concernant cette notion de *variabilité*.

Nous avons discuté dans le chapitre 3 de différentes formulations de la variabilité, qui nous ont conduit à suggérer l'utilisation d'un indicateur d'anomalie de variabilité, c'est-à-dire l'écart d'une variabilité mesurée ponctuellement vis-à-vis d'une variabilité moyenne de long terme (ou variabilité *normale*). Un indicateur simple de variabilité, par exemple le coefficient de variation, n'est en réalité pas directement comparable à un seuil limite (le seuil des $\pm 20\%$ en l'occurrence). La discussion qui suit montre la différence qu'il y a, à partir des résultats de l'application numérique, entre la variabilité entendue au sens statistique (à travers la métrique du coefficient de variation) et le type de variabilité spécifique au cadre conceptuel des limites planétaires.

Le coefficient de variation moyen sur la chronique T pour chaque mois Γ générique de l'année peut se formuler ainsi :

$$\forall i(\text{station}), \forall \Gamma(\text{mois}), cv_{i,\Gamma} = \frac{\sigma_{i,\Gamma}}{\mu_{i,\Gamma}}$$

Littéralement, $cv_{i,\Gamma}$ est un coefficient de variation interannuel : il renseigne par construction sur la variabilité observée entre plusieurs années sur un même mois. $\sigma_{i,\Gamma}$ est en effet l'écart-type entre tous les relevés du mois générique Γ sur les années de la période, et $\mu_{i,\Gamma}$ est le débit moyen du mois Γ sur la même période. L'écart-type est par construction toujours positif et ne fournit donc pas d'information sur le sens de la variabilité (en plus ou moins).

L'interprétation spontanée qu'il est possible d'en faire est que si $cv_{i,\Gamma}$ est supérieur à 20%, on peut considérer que la station i enregistre une variabilité excessive et qu'il y a un risque pour l'intégrité du cours d'eau correspondant.

Sur les 42 stations du territoire qui permettent de conduire l'analyse sur la période 2000-2020 et sur tous les mois de la période, aucune ne respecte $cv_{i,\Gamma} \leq 0.2$. La valeur médiane de $cv_{i,\Gamma}$ est environ 1.2 et la valeur minimum est 0.5. L'enjeu ici est de savoir si l'on peut conclure que toutes les stations dépassent effectivement leur seuil de variabilité limite ou si le couple (indicateur, seuil) ne sont pas appropriés.

La variabilité ainsi mesurée par le coefficient de variation ne traduit pas un écart à une référence environnementale mais un écart à une norme (le module mensuel), et rien n'indique dans la littérature que le débit moyen mensuel (sur une période suffisamment longue) ne puisse pas être considéré comme une référence environnementale. (ROCKSTROM et al. 2023) se basent sur deux articles principaux pour établir leur métrique et le seuil associé : l'article de (POFF et al. 2010) pour la philosophie de l'indicateur et l'article de (RICHTER et al. 2012) pour le seuil.

(POFF et al. 2010) ont conduit une revue de la littérature sur 150 publications relevant de la relation entre l'altération d'un régime hydrologique et l'altération de la biodiversité aquatique. Parmi ces publications, 55 fournissaient les informations suffisantes pour pouvoir calculer un pourcentage d'altération de l'écoulement (*percent change of flow*) et un pourcentage correspondant d'altération de la biodiversité (abondance spécifique et richesse spécifique). En comparant

ainsi un pourcentage de perturbation à un pourcentage de changement, les différentes métriques utilisées dans les différentes publications peuvent être positionnées dans un même graphe. Il est important de souligner néanmoins la diversité des métriques utilisées, puisque c'est bien là l'enjeu qui nous intéresse : de l'altération de quelle grandeur hydrologique parle-t-on ? Dans leur analyse qualitative des publications, Poff et Zimmermann résument les métriques en quatre grandes familles :

- le débit de pointe (*peak flow*), qui comprend : les crues, le débit instantané de pointe, les hautes eaux
- le débit moyen, qui comprend : le débit total sur une période ou le débit moyen sur une période
- le débit de base (*base flow*), qui comprend : le débit basses eaux, le débit d'étiage, la sécheresse
- le débit de court-terme : tout changement de débit observé en moins de 24h.

Poff et Zimmerman précisent également que le calcul du pourcentage d'altération de l'écoulement a été effectué entre la valeur de la métrique "pré-impact" et sa valeur "post-impact". Sans préciser ce qui différencie ces deux états, les auteurs ajoutent qu'ils ont pris en compte dans l'état "pré-impact" les valeurs prises à l'amont des perturbations étudiées. Il n'y a en tout cas pas de nécessité à ce que la valeur "pré-impact", que (ROCKSTROM et al. 2023) appellent "*pristine*", corresponde à un état environnemental de référence moindrement anthropisé.

Poff et Zimmerman proposent trois graphes pour la relation entre l'altération de l'écoulement et l'altération de la biodiversité, correspondant au regroupement de la biodiversité en trois familles taxonomiques. Les auteurs précisent que leur choix de représentation graphique a été de ramener toutes les perturbations à la hausse de l'écoulement estimée à plus de 100% à la valeur de 100% (à l'inverse il n'est physiquement pas possible que les perturbations à la baisse de l'écoulement aient perdu plus de 100% !) : cette procédure n'est d'ailleurs pas sans rappeler la remarque que nous faisons à propos de l'allure de la fonction de distribution des coefficients de variation sur notre territoire, qui tire vers la droite (vers des perturbations de plus de 100%). Les auteurs s'intéressaient à travers ces graphes à la détermination de potentiels effets de seuil. On constate à la lecture des graphes que seuls 2 points (c'est-à-dire 2 publications) sur 55 se situent dans l'intervalle -50%,+50% de perturbation de l'écoulement, et que la majorité des points se situent dans l'intervalle des perturbations à la baisse. Ceci ne signifie pas que les perturbations à la baisse impactent davantage la biodiversité que les perturbations à la hausse,

du moins la revue des auteurs ne permet-elle pas d'avancer une telle conclusion, mais qu'il y a simplement davantage de publications dans leur revue qui ont étudié une perturbation à la baisse. De même, la quasi-absence de points dans l'intervalle $-50\%, +50\%$ de perturbation de l'écoulement ne permet pas de conclure qu'aucune perturbation de la biodiversité n'a lieu pour des perturbations comprises dans cet intervalle, mais simplement qu'une telle pression peut ne pas avoir été documentée dans les publications. Remarquons toutefois que ce n'est peut-être pas pour rien qu'une telle perturbation de l'écoulement n'ait pas été documentée. En effet, sur les 47 stations de notre échantillon pour le Sud-Loire, l'écart-type mensuel n'est jamais inférieur à 55%. Une perturbation de l'écoulement inférieure à 55% (en plus comme en moins, l'écart-type étant symétrique) est donc techniquement invisible sur les stations de mesure pour la période de moyennisation de 20 ans qui a été retenue.

Remarques sur la construction de l'indicateur de variabilité

Qu'en serait-il alors si l'on comparait, pour un mois donné et une station donnée, l'écart-type enregistré sur ce mois et sa valeur normale de long terme ? (par exemple la moyenne de tous les écarts-types de ce même mois et de cette même station sur une période de 20 ans).

Cette formulation compare en réalité encore deux variabilités différentes : la variabilité intramensuelle d'un côté, et de l'autre la variabilité interannuelle d'un même mois. On comprend en effet intuitivement qu'il y a plus de variabilité entre les jours d'un même mois qu'entre le même mois sur plusieurs années, cela tient à l'effet lissant de la moyenne : quand on prend la moyenne sur plusieurs années des moyennes mensuelles, on rajoute une opération de moyennisation (donc de lissage et donc de perte "artificielle" de variabilité statistique) par rapport à la moyenne mensuelle directe sur les débits journaliers.

Une première conclusion importante de ce travail est que la "variabilité" du débit qui intéresse les concepteurs du cadre des limites planétaires (du moins pour la version ROCKSTROM et al. 2023) n'est pas la même que la "variabilité" au sens statistique. Alors que la variabilité au sens des limites planétaires consiste à comparer deux états d'une même métrique, la variabilité au sens statistique, et exprimée par le coefficient de variation, consiste à comparer deux métriques différentes pour le même état (i.e. l'écart-type par rapport à la moyenne pour la même chronique temporelle).

En corollaire, une deuxième conclusion de ce travail en termes de construction d'un indicateur d'AESA pour l'eau de surface, est que l'état environnemental dit de référence doit être exprimé dans la même métrique que l'état environnemental dit perturbé, et que la métrique en soi importe peu.

Cette discussion légitime la proposition de l'indicateur que nous formulons au chapitre 3, celui d'une anomalie de variabilité, où la variabilité enregistrée ponctuellement est comparée à la même variabilité enregistrée sur le temps long. L'indicateur d'anomalie de variabilité est finalement ainsi formulé :

$$\forall i(\text{station}), \forall t(\text{mois}), \Delta cv_{i,t \in \Gamma} = \frac{cv_{i,t} - \overline{cv_{i,t \in \Gamma}}}{\overline{cv_{i,t \in \Gamma}}}$$

Où :

$$\overline{cv_{i,t \in \Gamma}} = \frac{\sum_{t \in \Gamma} cv_{i,t}}{\text{card}(t \in \Gamma)}$$

Il s'agit de calculer l'écart entre la variabilité intramensuelle d'un mois t donné et de la comparer à la moyenne de la variabilité intramensuelle de tous les autres mois t appartenant au même mois générique Γ que lui. Ce calcul est implémenté dans la fonction `plotVariabilityAnomaly` d'`HydroTrendsFrance`, qui permet de mettre en graphique les écarts mensuels de variabilité à la variabilité normale (anomalies) pour chaque station et sur toute la période.

L'anomalie de variabilité des cours d'eau du Sud-Loire

L'implémentation de l'indicateur d'anomalie de variabilité peut prendre plusieurs formes : une forme que l'on peut appeler "désagrégée", où l'anomalie est représentée pour chaque mois de la chronique étudiée et chaque station de mesure (voir Figure 4.6) ; et des formes "agrégées" qui consistent à réduire la chronique à une unique dimension (par exemple, l'anomalie annuelle moyenne ou la fréquence de dépassement de l'anomalie tolérée) ou bien à réduire le territoire à une unique dimension (anomalie moyenne sur toutes les stations) ou bien encore à réduire ces deux dimensions temporelle et spatiale à une unique dimension (moyenne annuelle sur la chronique et les stations). Le graphe de la station de Villerest (station du territoire la plus à l'aval sur la Loire) est reproduit dans la figure ci-dessous.

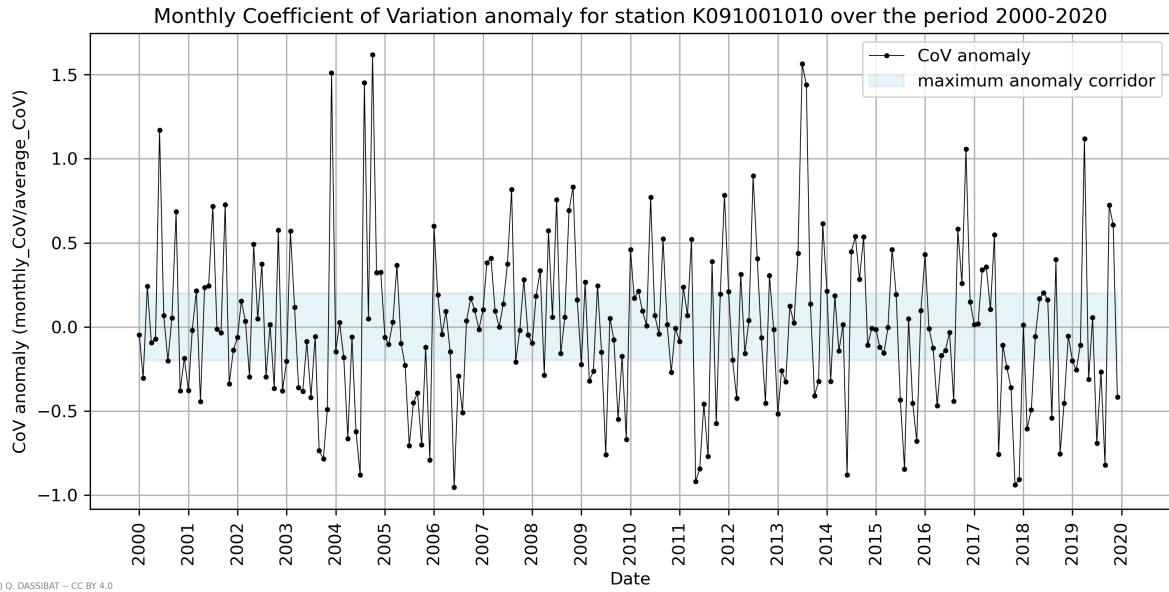


Figure 4.6 – Valeurs mensuelles de l'indicateur d'anomalie de variabilité pour la station la plus à l'aval du territoire du Sud-Loire (station de Villerest, Loire, 2000-2020)

La représentation désagrégée est utile pour visualiser l'indicateur sur une station de mesure à la fois, mais s'avère peu synthétique pour visualiser les résultats sur plusieurs dizaines de stations en même temps (et donc produire une évaluation véritablement *territoriale*).

On peut envisager de réduire la dimension de $\Delta cv_{i,t \in \Gamma}$ (qui peut être grande lorsque le nombre d'années N est grand) de plusieurs façons.

En considérant d'abord la moyenne mensuelle des fréquences de dépassement. Celle-ci s'obtient en calculant, pour une station et un mois d'une année générique, la fréquence des mois sur la chronique où le seuil limite de 20% est dépassé.

$$\Phi_{i,\Gamma} = \frac{\text{card}(t \in \Gamma : |cv_{i,t} - \overline{cv_{i,t,\Gamma}}| \leq 1.2\overline{cv_{i,t,\Gamma}})}{\text{card}(t \in \Gamma)}$$

Où $\text{card}(t \in \Gamma : |cv_{i,t} - \overline{cv_{i,t,\Gamma}}| \leq 1.2\overline{cv_{i,t,\Gamma}})$ désigne pour tout mois générique le nombre de dates dans la chronique où le rapport entre le coefficient de variation à cette date et le coefficient de variation du mois générique associé dépasse 20%, en plus ou en moins. La dimension de $\Phi_{i,\Gamma}$ est alors 12. Numériquement, on attribue la valeur de 1 aux mois t où un tel dépassement est constaté, et une valeur de 0 sinon. Avec le langage Python, cette opération est réalisée grâce à une fonction `lambda` implémentée dans notre tableau de données de type `pandas.DataFrame` grâce à la méthode `pandas.DataFrame.apply(lambda)` :

```
# lambda x: 1 if (x >= percentChange) | (x <= -percentChange) else (np.nan if x
is np.nan else 0)
```

On note que dans la fonction `lambda`, trois valeurs sont en réalité affectées : le 0 et le 1, comme nous le décrivions plus haut, et une valeur vide `np.nan` (il s'agit de l'encodage `numpy` pour les valeurs vides, davantage compatible avec la librairie `pandas` que l'encodage primitif `None` de Python) lorsque les données du mois t ne sont pas disponibles et n'ont pas permis de calculer $cv_{i,t}$. Si l'on avait attribué à ces mois la valeur de 0, on aurait confondu une absence de données avec un non-dépassement de la limite. On effectue ensuite la somme des valeurs 0 et 1 que l'on divise par le nombre de valeurs.

On peut réduire encore davantage la dimension de $\Delta cv_{i,t \in \Gamma}$ à 1, par une moyenne interannuelle pour la station i , soit directement sur la fréquence de dépassement en calculant Φ_i , soit sur la valeur du dépassement en calculant Δcv_i .

$$\Phi_i = \frac{\sum_{1 \leq \Gamma \leq 12} \Phi_{i,\Gamma}}{\text{card}(1 \leq \Gamma \leq 12)}$$

Et :

$$\Delta cv_i = \frac{\sum_{1 \leq \Gamma \leq 12} \Delta cv_{i,t \in \Gamma}}{\text{card}(1 \leq \Gamma \leq 12)}$$

La formulation telle qu'elle de Δcv_i pose toutefois problème car il s'agit de calculer la moyenne des $\Delta cv_{i,t \in \Gamma}$ qui sont centrés sur $\overline{cv_{i,t \in \Gamma}}$. Autrement dit, la moyenne Δcv_i tend nécessairement vers 0, et le calcul de la médiane sur quelques échantillons semble indiquer que la médiane tende elle aussi vers 0 (la distribution des $\Delta cv_{i,t \in \Gamma}$ semble suivre une loi normale). On pourrait en revanche envisager le calcul d'une moyenne à gauche et d'une moyenne à droite de la norme $\overline{cv_{i,t \in \Gamma}}$, mais l'interprétation de cette double moyenne est plus difficile, rendant la lecture du résultat peu compréhensible.

On préfère donc l'indicateur synthétique de la moyenne annuelle des fréquences de dépassement pour chaque station. On peut alors produire une carte pour l'ensemble des bassins versants du territoire (voir Figure 4.7). La génération de cette carte est implémentée dans `HydroTrendsFrance` qui la génère à la volée, mais nous en donnons ici une visualisation retravaillée dans QGIS à partir des couches vectorielles générées par `HydroTrendsFrance`.

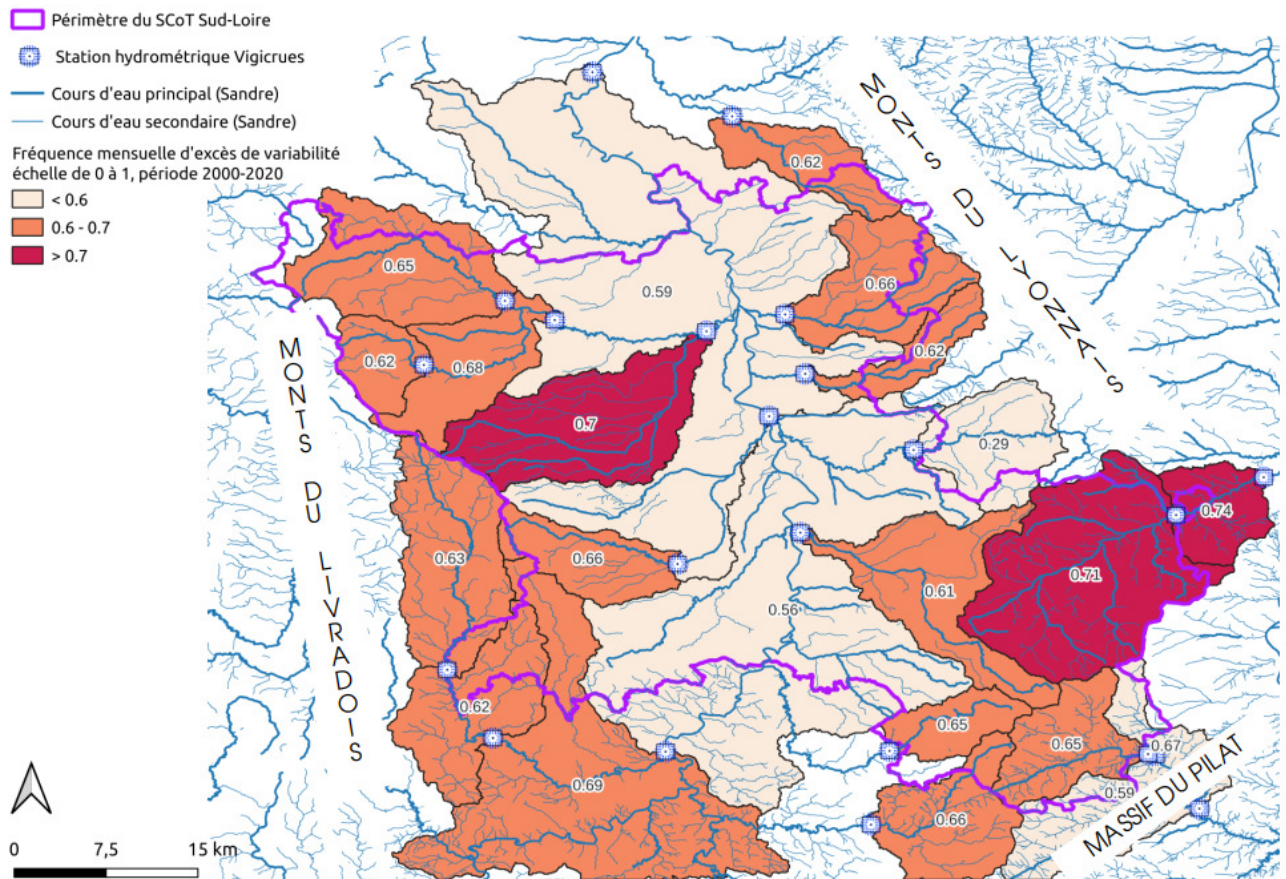


Figure 4.7 – Fréquences mensuelles de dépassement des stations Vigicrues du Sud-Loire (moyenne annuelle 2000-2020, visualisation par les zones d'apport aux stations)

Selon cet indicateur (voir Figure ??), tous les bassins versants du territoire ont une fréquence annuelle moyenne de dépassement oscillant autour de 60% (de 56% à 77%). Ceci s'explique vraisemblablement par le fait qu'ils appartiennent (presque tous) au même bassin versant global (la Loire). La subdivision en sous bassins versants (ou zones d'apports) ne semble pas être discriminante, autrement dit le comportement hydrologique (en termes d'anomalie de variabilité du débit) est relativement homogène entre les sous bassins versants. Ceci est probablement dû à la taille des sous bassins versants, qui est importante, et qui se faisant tend à homogénéiser les caractéristiques anthropiques susceptibles de perturber le débit. Bien que les écarts de dépassements entre les stations soit relativement resserrés, on peut toutefois proposer de distinguer le territoire en 3 zones :

- la plaine de la Loire est associée à une anomalie de variabilité relativement moins importante que le reste du territoire
- les reliefs sont en revanche associés à une anomalie relativement élevée : les contreforts des Monts du Livradois à l'Ouest, les coteaux du massif du Pilat au Sud et les coteaux des Monts du lyonnais à l'Est

- enfin, deux stations hydrométriques se distinguent du reste par leur fréquence mensuelle de dépassement supérieure à 70% : le bassin versant du Vivézy sur la rive gauche de la Loire (70%), et le bassin versant du Gier, au Sud-Est du territoire, qui est intercepté par deux stations hydrométriques (71 et 74% de dépassement).

L'indicateur que nous proposons ici nous semble constituer une alternative fidèle à l'esprit de (ROCKSTROM et al. 2023), tout en ne reposant que sur des données de mesure. De même que leur indicateur, son but est d'alerter sur une variabilité jugée excessive ou anormale du débit d'un cours d'eau. Dans leur indicateur, la normalité est basée sur le résultat d'une modélisation de l'écoulement en régime peu anthropisé ; dans le nôtre, la normalité est à entendre au sens statistique, comme la valeur moyenne sur la chronique d'observation. La normalité est ainsi dans leur indicateur atemporelle, tandis que la nôtre dépend de la longueur de la chronique et de sa position, et est donc associée à une erreur standard.

Nous avons dans une partie précédente étudié l'importance sur la valeur normale (ou moyenne) du débit de ce double effet (ou artefact) statistique, en considérant la mesure d'une erreur standard. Ici, nous pouvons évaluer une plage d'erreur sur la valeur de notre indicateur en comparant la valeur de l'indicateur pour une période de normalisation standard (20 ans, de 2000 à 2020) et une période longue d'environ 100 ans. Cette comparaison n'est possible que pour les stations dont les données sont disponibles pour une telle chronique. Nous allons donc évaluer le biais de notre indicateur dû à son ancrage temporel en nous appuyant sur les stations de Villerest et de Bas-en-Basset. Ces stations ont en plus l'avantage de couvrir le bassin versant de la Loire sur sa quasi-totalité pour le territoire du Sud-Loire.

La valeur de Φ_i à Villerest, au nord du territoire, pour une chronique de 20 ans est de 0.5875 tandis qu'elle est de 0.6514 pour une chronique de 101 ans. A Bas-en-Basset, on passe de 0.6917 sur 20 ans à 0.6964 sur 100 ans. Nous donnons ici les valeurs de Φ_i avec une précision de 4 chiffres après la virgule pour le besoin de la comparaison entre les longueurs de chroniques ; cette précision ne doit pas être prise pour autant comme significative. Pour la station de Villerest, la moyennisation de Φ_i sur une chronique de 20 ans tend à sous-estimer sa valeur d'environ 7 points de pourcentage par rapport à une chronique de 101 ans, tandis que pour celle de Bas-en-Basset, il y a aussi une sous-estimation mais qui est inférieure à 1 point de pourcentage. L'erreur associée à la longueur de la chronique ne semble donc pas si importante. Par précaution, pour tout calcul de Φ_i sur une chronique comprise entre 20 et 100 ans, une

valeur d'erreur de $\pm 10\%$ peut être associée. Cette valeur d'erreur fonctionne par exemple pour la troisième chronique de 70 ans que nous avons testée (1950-2020). Φ_i vaut 0.6333 pour Villerest et 0.7048 pour Bas-en-Basset.

Pour la comparaison, nous pouvons juxtaposer l'indicateur de (ROCKSTROM et al. 2023) avec le nôtre pour le territoire (voir Figure 4.8). Leur indicateur de stress environnemental au seuil de 20% consiste à calculer l'anomalie moyenne de débit pour tous les mois génériques de la chronique 2000-2020, puis à compter sur les 12 mois génériques le nombre de fois où l'anomalie dépasse le seuil de $\pm 20\%$. Numériquement, pour tout carreau i , si l'écart entre le débit moyen interannuel perturbé du mois Γ (qui correspond à ce que nous avons précédemment noté $\mu_{i,\Gamma}$) et le débit moyen interannuel non perturbé du même mois (que l'on peut noter $\mu_{i,\Gamma}^*$) ne dépasse pas le seuil de $\pm 20\%$, alors la valeur 0 est attribuée au mois Γ , sinon la valeur 1. La somme pour un carreau et pour tous les mois génériques Γ est ensuite effectuée, ce qui permet d'obtenir une fréquence mensuelle (qui va de 0 à 12).

La fréquence $\Phi_{i,\Gamma}$ de notre indicateur est différente. Il s'agit du nombre de fois où la variabilité d'un mois réel t dépasse la variabilité normale du mois générique Γ correspondant. La somme des 0 et des 1 porte ainsi sur les mois réels t et pas sur les mois génériques Γ . On pourrait décider d'attribuer la valeur de 1 lorsque $\Phi_{i,\Gamma}$ est différent de 0, mais $\Phi_{i,\Gamma}$ est toujours non nul (il est très peu probable que sur une période de 20 ans jamais aucun mois réel t n'est excédé sa variabilité normale sur la période!). Autrement dit, la conversion directe en fréquence mensuelle donnerait systématiquement la valeur de 12 pour tous les bassins versants i .

Pour la comparaison graphique, nous proposons donc de réaliser une simple opération de bijection des fréquences Φ_i de l'espace $[0,1]$ vers l'espace $[0,12]$. Si $q_{[0,1]}$ est une fréquence de $[0,1]$ alors $q_{[0,12]} = \lfloor 12 \times q_{[0,1]} \rfloor$ où $\lfloor \cdot \rfloor$ désigne la partie entière. Cette opération est quelque peu artificielle, du moins statistique, car la conversion physique entre les métriques n'est pas directement possible. Cela justifie que le chapitre suivant cherchera à développer une méthode de comparaison directe entre les données de modélisation issues du WBM et les données de mesure issu du réseau Vigicrues.

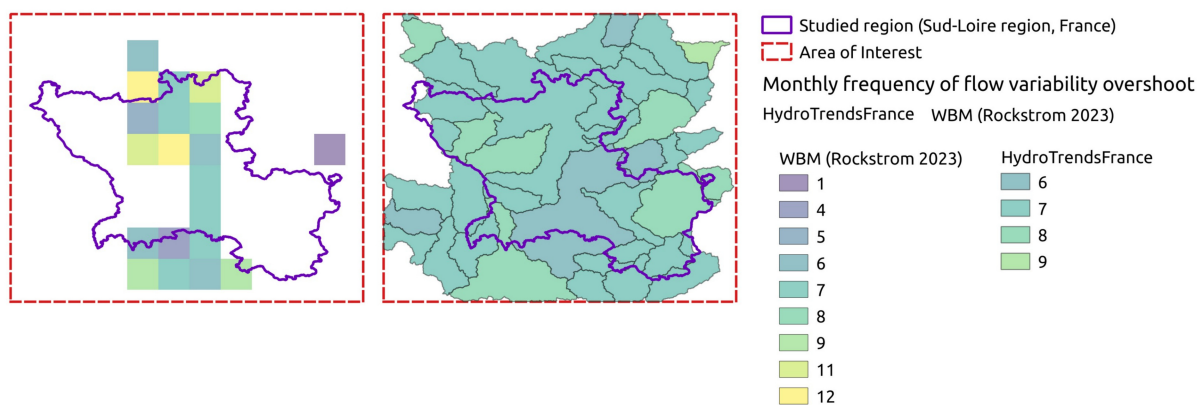


Figure 4.8 – Comparaison des résultats entre l'indicateur de (ROCKSTROM et al. 2023) et l'indicateur d'anomalie de variabilité

On peut noter que l'étendue des valeurs est plus grande avec les données du WBM, tandis que par le protocole d'HydroTrendsFrance, elles sont davantage resserrées. La comparaison spatiale est elle aussi peu évidente car les données du WBM sont spatialisées par un carroyage régulier, tandis que les données d'HydroTrendsFrance sont projetées par sous bassins versants. La comparaison semblerait toutefois indiquer que les situations sont moins contrastées dans la réalité (avec des données de mesure) que dans le modèle WBM, ceci s'expliquant vraisemblablement par l'effet de seuil ou de borne produit par une modélisation carroyée : il n'y a pas de transition douce entre les carreaux, tandis qu'avec des données de mesure réparties sur des bassins versants de taille importante, les effets des pressions anthropiques sont plus diffus.

4.3.2.2 L'effet du changement climatique et de l'anthropisation des paysages

L'approche par la variabilité qui est portée par le cadre des frontières du système Terre présente un intérêt qui va au-delà de la simple production de l'indicateur de fréquence de dépassement que nous discutons précédemment. Cette approche met en effet l'accent sur une métrique hydrologique relativement peu mobilisée dans la politique environnementale de l'eau. Sur le territoire du Sud-Loire, la récente création d'un PTGE (Projet de Territoire pour la Gestion de l'Eau, qui correspond au bassin versant de la Loire en Rhône-Alpes) s'est justement accompagnée d'un regain d'intérêt pour une telle métrique, en proposant, dans l'étude HMUC préliminaire à l'établissement d'un plan de gestion de l'eau, un indicateur dit de "débit biologique", visant précisément à identifier des couloirs maximum de variabilité du débit pour préserver les conditions biologiques des milieux aquatiques. Ce débit biologique est calculé selon la méthode Estimhab, qui consiste à effectuer une série de mesures couplées sur un même site : mesure des débits, de la granulométrie, comptage des espèces, à la fois en régime de basses eaux et de hautes eaux (LRA 2024). Dix-sept sites ont ainsi pu être diagnostiqués selon

cette méthode sur le territoire du PTGE. Notons également que le volet "hydrologie" de l'étude HMUC est censé mettre en avant un concept très proche de celui de "débit non anthropisé" : le "débit désinfluencé" ("Ce volet vise à reconstituer les écoulements tels qu'ils seraient sans les usages anthropiques (prélèvements et rejets). On parle ainsi de « désinfluencer » les débits. Pour ce faire, l'ensemble des prélèvements et points de rejets ont été localisés et quantifiés." ²⁰)

Là où la variabilité au sens des frontières du système Terre permet d'envisager une variation entre un état d'anthropisation jugé moins perturbé et un état jugé actuel, la décomposition de la variabilité au sens statistique offre d'autres perspectives, parmi lesquelles la contribution à la variation observée de la tendance climatique de long terme et de l'anthropisation des paysages (que nous avons associées au type d'incertitude dite "structurelle").

L'effet du changement climatique sur les débits

La contribution du forçage climatique de long terme sur la tendance hydrologique aux stations de mesure du territoire a été évaluée par une régression linéaire de Mann-Kendall qui, par son caractère non paramétrique, permet de s'affranchir des hypothèses quant à la distribution et à l'indépendance des variables aléatoires étudiées, ce qui est nécessaire dans le cas des mesures de débits. Par cette méthode, nous avons cherché à identifier une tendance significative sur deux métriques hydrologiques caractéristiques : d'une part le débit moyen mensuel, d'autre part l'écart-type mensuel. L'hypothèse *a priori* est celle d'une baisse des débits moyens mensuels sur le temps long, et d'une hausse de l'écart-type. D'un point de vue causal, ce sont bien là les deux réactions attendues du cycle de l'eau à l'effet du forçage radiatif : la hausse des températures moyenne de surface augmente le stock d'eau évaporée et donc le stock disponible dans les eaux de surface tend à diminuer ; cette hausse du stock atmosphérique provoque en retour une modification des régimes de pluie, qui se traduit notamment par une fréquence accrue des événements pluvieux extrêmes et donc par un ruissellement fortement variable (IPCC 2023). Si cette chaîne causale est largement documentée, nous n'avons pour autant pas cherché à la vérifier sur le territoire. Nous avons réduit l'analyse à un test d'hypothèse sur la baisse des débits moyens mensuels et la hausse de leur variabilité. La corrélation avec l'évolution des températures sur le territoire ou l'évolution des précipitations n'a pas été menée.

HydroTrendsFrance implémente la régression de Mann-Kendall et le test d'hypothèse dans

20. Voir : <https://sage-loire-rhone-alpes.fr/?p=4086>

sa fonction `hydrofunc.MannKendallStat`. La série temporelle est renseignée sous forme de liste, à savoir dans notre cas une liste des débits mensuels moyens et une liste des écart-types mensuels pour chaque station de mesure. Le test statistique est automatiquement effectué. Deux possibilités sont envisageables : un test bilatéral et un test unilatéral. Pour le premier, il s'agit de différencier le test selon le signe de la métrique Z_{MK} qui, si elle négative, est testée par rapport à la fonction de distribution de la loi normale centrée réduite ; si elle est positive, par rapport à la probabilité réciproque. Ce test permet de tester le sens de la tendance, à la hausse ou à la baisse. Le test unilatéral quant à lui permet plus modestement de tester si la tendance est significativement différente d'une absence de tendance, sans égard au sens de la tendance. Il s'implémente facilement en multipliant par le scalaire 2 la probabilité réciproque de la loi normale centrée réduite. C'est ce dernier test qui est utilisé dans `hydrofunc.MannKendallStat`.

La Figure 4.9 montre le résultat de la régression de Mann-Kendall pour la série temporelle des débits moyens mensuels sur la période 2000-2020. Le choix graphique a été de spatialiser la tendance de long terme observée aux stations par les sous bassins versants associés aux stations. La représentation est la suivante : les bassins versants des stations dont les données sont disponibles pour la période appelée sont délimités par une ligne noire pleine, et les lignes noires en pointillés, plus fines, délimitent quant à elles les bassins versants de chacune des stations du territoire, y compris celles dont les données ne sont pas disponibles. Ce choix a été fait pour rendre compte du fait qu'HydroTrendsFrance a bien pris en compte l'ensemble des stations disponibles, mais que les traitements n'ont pu être effectuées que sur celles dont les données correspondent à la période appelée et à la grandeur hydrométrique requêtée (débits journaliers moyens). Les points noirs présents sur la carte représentent la position des stations qui ont été utilisées pour les traitements (il s'agit en fait de leur position reprojetée dans le réseau hydrographique modélisé pour tracer les bassins versants, et pas de leur position "officielle"). La couleur des bassins versants rend compte du taux de variation tendanciel du débit moyen mensuel, qui s'échelonne sur $[0,1]$ (à multiplier par 100 pour avoir la valeur en pourcentage). On rappelle que ce taux de variation correspond à l'évolution entre l'ordonnée à l'origine (ici, l'année 2000) de la droite de régression de Mann-Kendall et son ordonnée en fin de chronique (ici, l'année 2020). Le texte superposé à la position des stations renseigne sur la p-value associée (qui est considérée comme satisfaisante pour des niveaux de confiance autour de 95%, soit une p-value inférieure à 5%, soit encore 0.05).

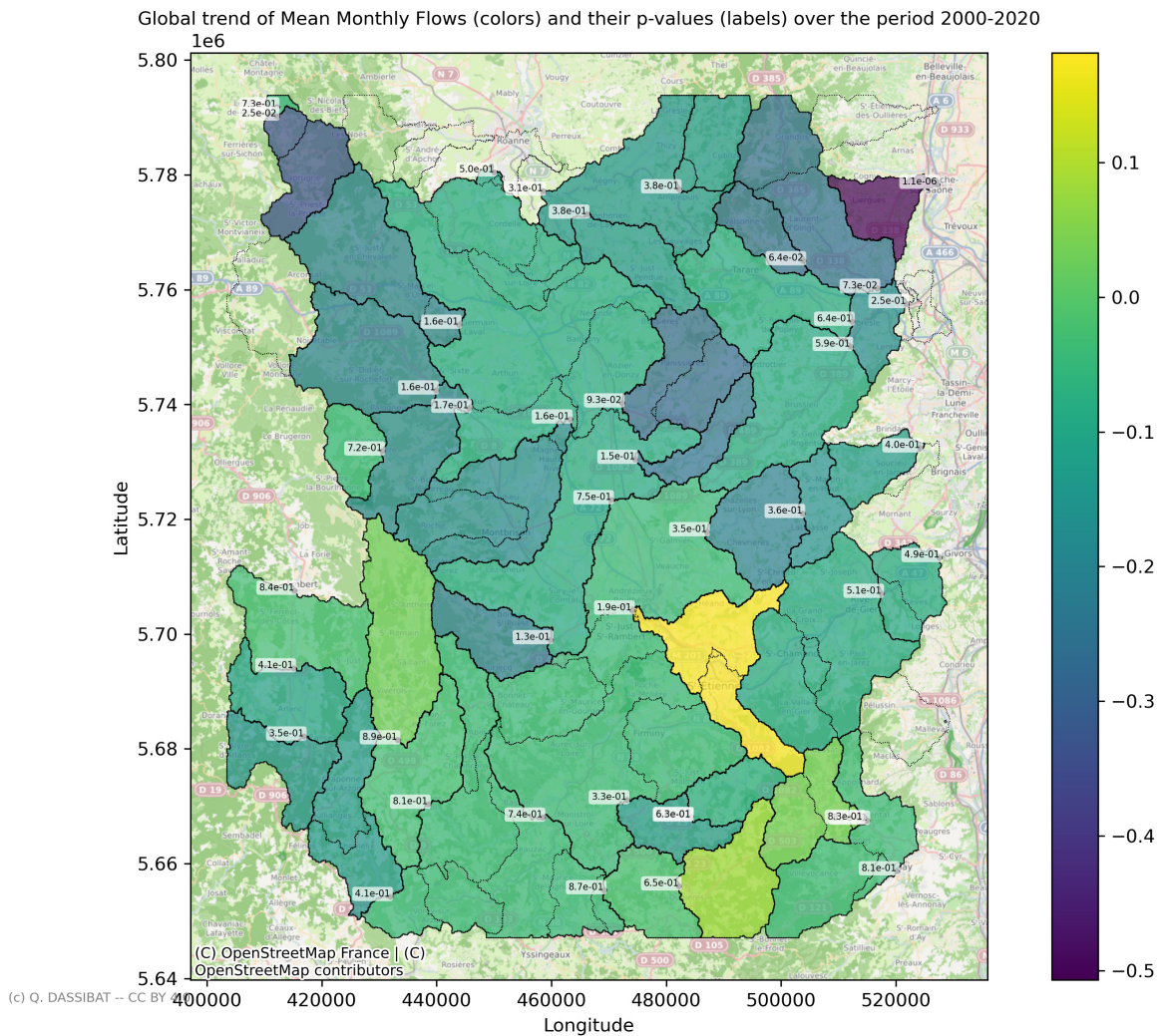


Figure 4.9 – Évolution tendancielle des débits mensuels (Sud-Loire, 2000-2020)

Pour une période de 20 ans, les p-values ne sont globalement pas significatives et, si elles le sont, la tendance observée n'est vraisemblablement pas attribuable au forçage radiatif (du moins pas en majorité) car l'année de référence n'est pas assez éloignée dans le temps pour qu'elle puisse correspondre à un état "avant forçage".

L'effet à long terme du forçage radiatif n'est visible que sur une période longue, typiquement de 100 ans. Avec la chronique 1922-2023, la plus longue disponible sur le territoire du Sud-Loire avec les stations de Villerest et Bas-en-Basset, qui sont d'ailleurs les plus représentatives car elles couvrent la quasi-totalité du bassin versant du territoire (Bas-en-Basset correspondant à la Loire amont sur le territoire, et Villerest à la Loire aval), nous pouvons évaluer la tendance de long terme de l'évolution des débits (voir Figure 4.10). Le bassin versant associé à la station de mesure de Villerest a enregistré une baisse de -20% du débit mensuel moyen sur 100 ans, et cela avec un niveau de confiance très bon, supérieur à 99% (la p-value est de 0.0036). La

tendance est du même ordre pour la station de Bas-en-Basset, où nous estimons la baisse du débit mensuel moyen sur la période à presque -30% avec un niveau de confiance également supérieur à 99% (p-value 0.000004). Ces observations corroborent l'étude HMUC réalisée par le PTGE, où la baisse du débit à Bas-en-Basset depuis 1919 est estimée à -20% (LRA 2022)²¹.

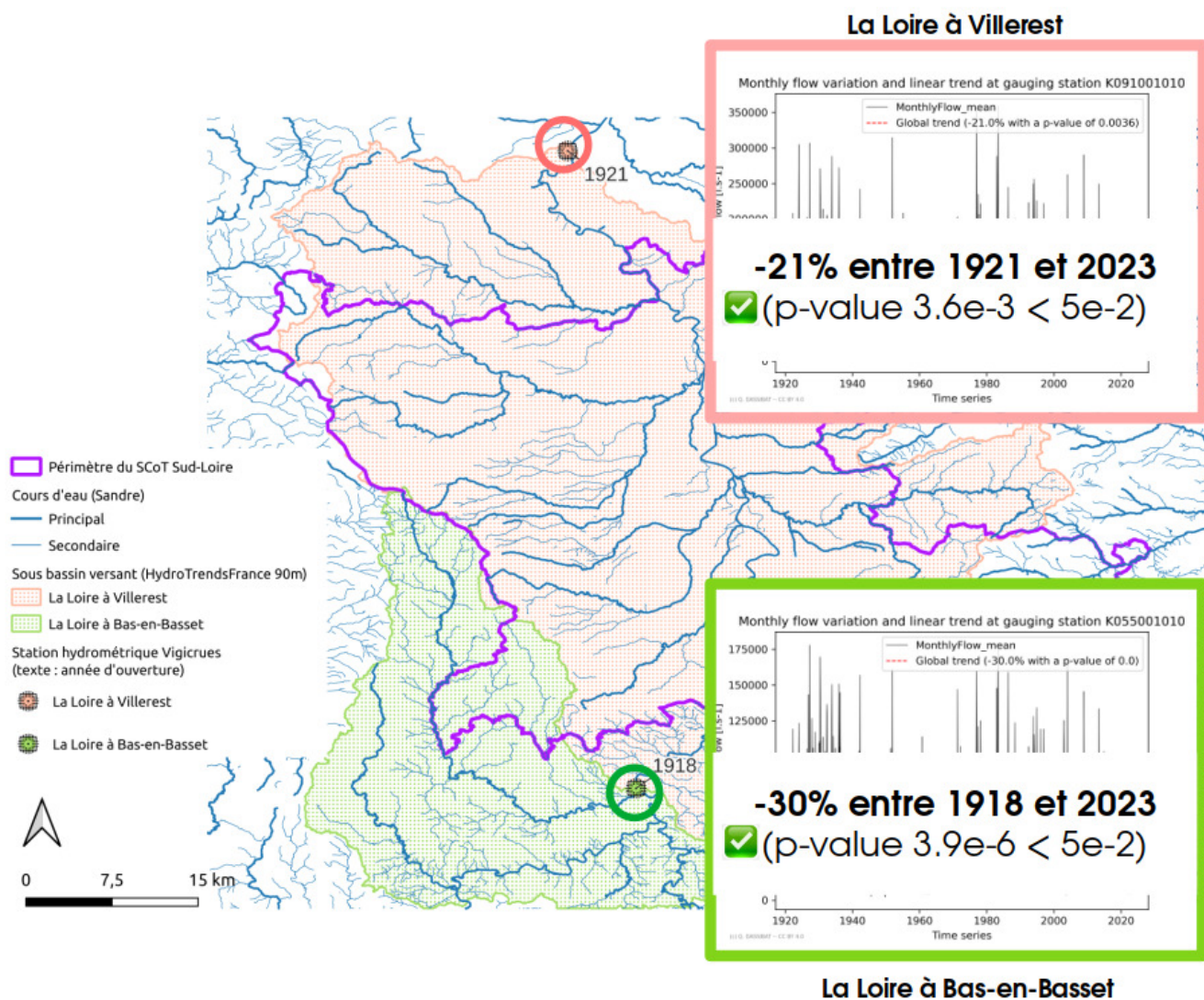


Figure 4.10 – Évolution tendancielle des débits moyen mensuels (station de Villerest : 1921-2023 ; station de Bas-en-Basset : 1918-2023)

Pour ce qui est de la tendance de long terme sur la variabilité, l'effet n'est pas comme attendu à la hausse, mais plutôt à la baisse. Notre test statistique unilatéral ne permet pas d'en être certain pour autant. Pour la station de Villerest, l'écart-type moyen diminuerait tendanciellement de 17%, et de 39% pour celui mesuré à la station de Bas-en-Basset (voir Figure

21. Cet écart de 10 points de pourcentage n'est pas significatif car, comme nous l'avons documenté à propos de l'exactitude des données Vigicrues, les valeurs de débits sont en moyenne incertaine de 10%.

4.11). Cette tendance baissière pourrait s'expliquer par les aménagements de canalisation, d'endiguement et les ouvrage de retenue du fleuve Loire survenus au cours de la chronique (1922-2023). En effet, celle-ci débute à une date antérieure à la mise en service du barrage de Grangent (1957) qui est situé (bien) à l'amont de la station de Villerest. Ce barrage pourrait notamment avoir pour effet d'écarter les crues. Cependant, étant situé à l'aval de la station de Bas-en-Basset, il n'explique pas la baisse de variabilité du débit observée à cette station depuis un siècle. Il faudrait alors s'intéresser à l'histoire des aménagements spécifiques à la Loire amont.

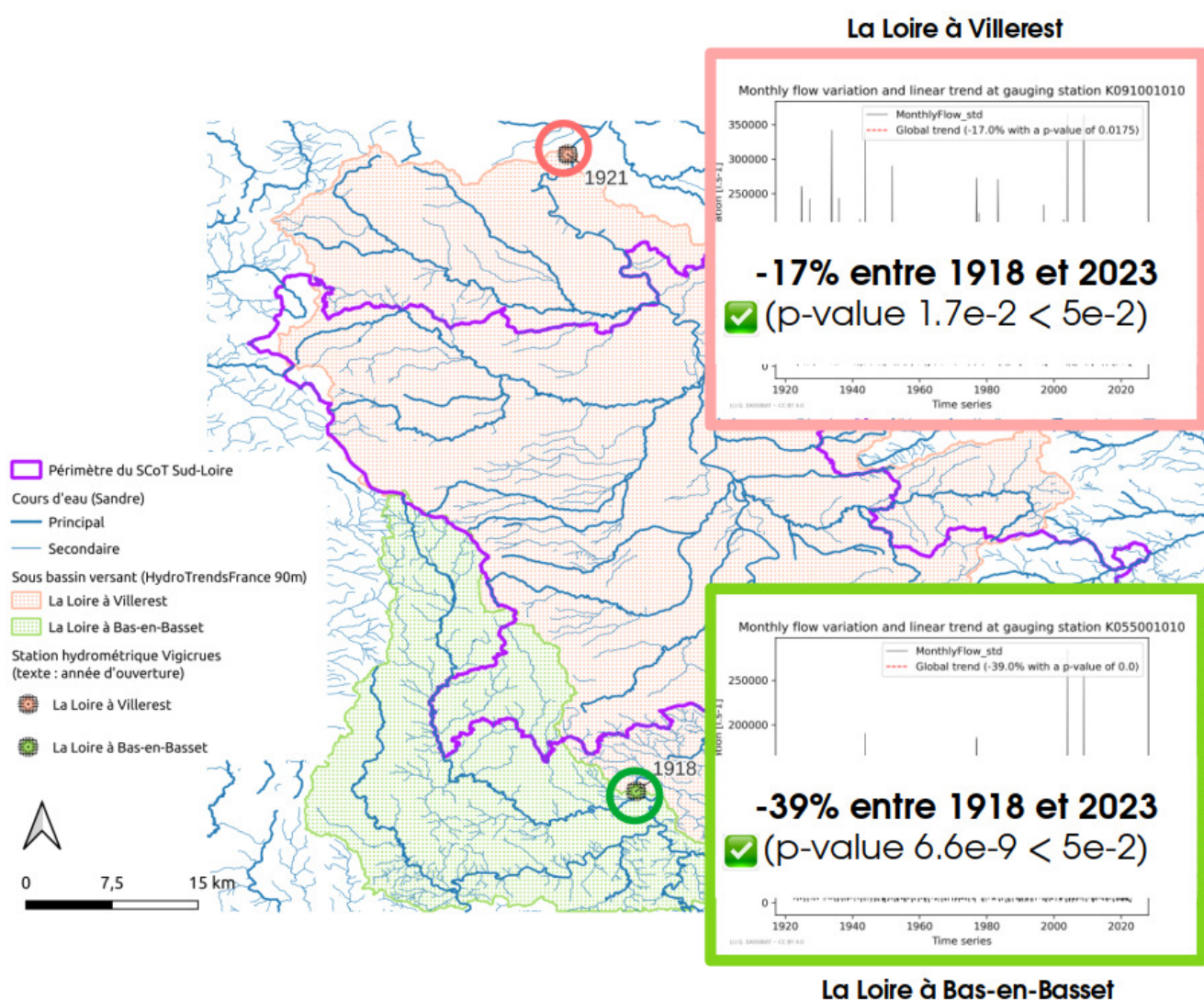


Figure 4.11 – Évolution tendancielle des écarts-types moyens mensuels (station de Villerest : 1921-2023 ; station de Bas-en-Basset : 1918-2023)

L'effet de l'anthropisation des paysages sur les débits

Le contexte d'intervention territoriale dans lequel cette thèse s'inscrit, à savoir le renou-

vement du SCoT du Sud-Loire, légitime le fait de s'intéresser à l'effet de l'aménagement du territoire sur l'intégrité des cours d'eau, c'est-à-dire la corrélation qu'il peut y avoir entre certains éléments anthropiques du paysage et la variabilité des débits.

Nous avons pour cela effectué une corrélation de Pearson ainsi qu'un test statistique unilatéral de p-value entre le coefficient de variation annuel moyen enregistré à chaque station sur une période donnée, et les éléments anthropiques du paysage. C'est ici que la spatialisation des stations de mesure par leur sous bassins versants respectifs prend tout son sens, car ainsi la présence des éléments anthropiques peut être liée au comportement hydrologique d'une station. Remarquons que l'analyse que nous effectuons est statique : si nous partons bien d'une série temporelle de débits, celle-ci est réduite à une seule dimension par le calcul du coefficient de variation annuel moyen sur la période (i.e. la moyenne des coefficients de variation mensuels moyens). Les éléments anthropiques sont également dénombrés de façon statique : l'occupation du sol est observé grâce au millésime 2018 du produit *Corine Land Cover*, et le reste des éléments est récupéré pour l'année 2024 à partir des données *Open Street Map*.

Le coefficient de variation annuel moyen sur la chronique \overline{cv}_i est calculé de la façon suivante :

$$\overline{cv}_i = \frac{\sum_{1 \leq \Gamma \leq 12} cv_{i,\Gamma}}{\text{card}(1 \leq \Gamma \leq 12)}$$

Nous avons décrit dans le chapitre (3) la façon dont les éléments anthropiques du paysage ont été quantifiés de telle sorte à ne retenir qu'une valeur par bassin versant tout en essayant de se prévenir au mieux de l'effet des superficies différentes entre les bassins versants (généralement en exprimant le nombre des éléments anthropiques par unité de surface).

Pour la chronique 2000-2020, nous obtenons ainsi la matrice de corrélation reproduite en Figure 4.12.

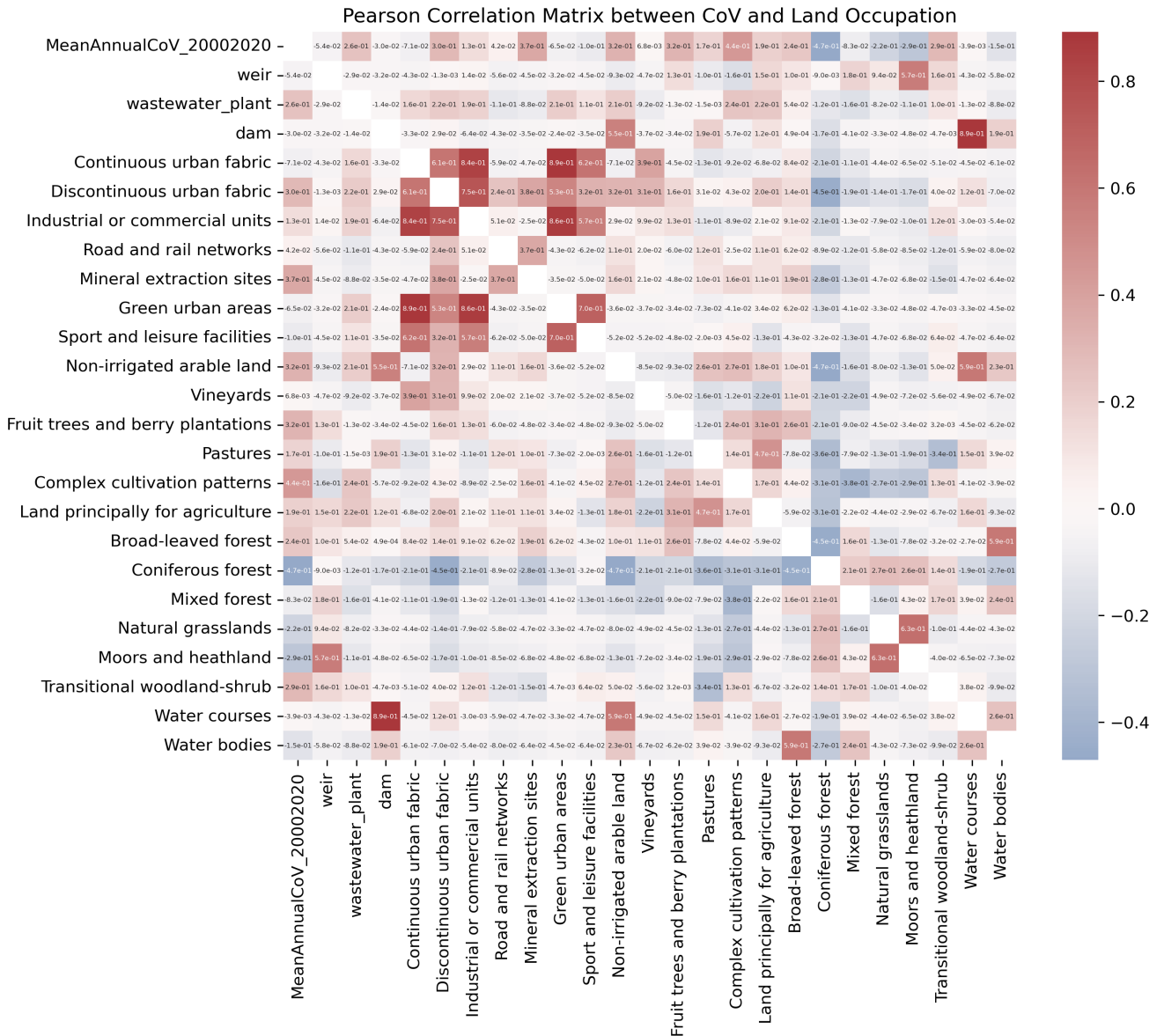


Figure 4.12 – Matrice de corrélation entre le coefficient de variation annuel moyen des débits et les éléments anthropiques du paysage (Sud-Loire, 2000-2020)

La corrélation qui nous intéresse est celle qui met au regard le coefficient de variation est les éléments anthropiques. Le reste des coefficients de corrélation nous importe peu ici. Nous reproduisons l'extraction des corrélations pour la colonne du coefficient de variation des débits dans la Figure 4.13.

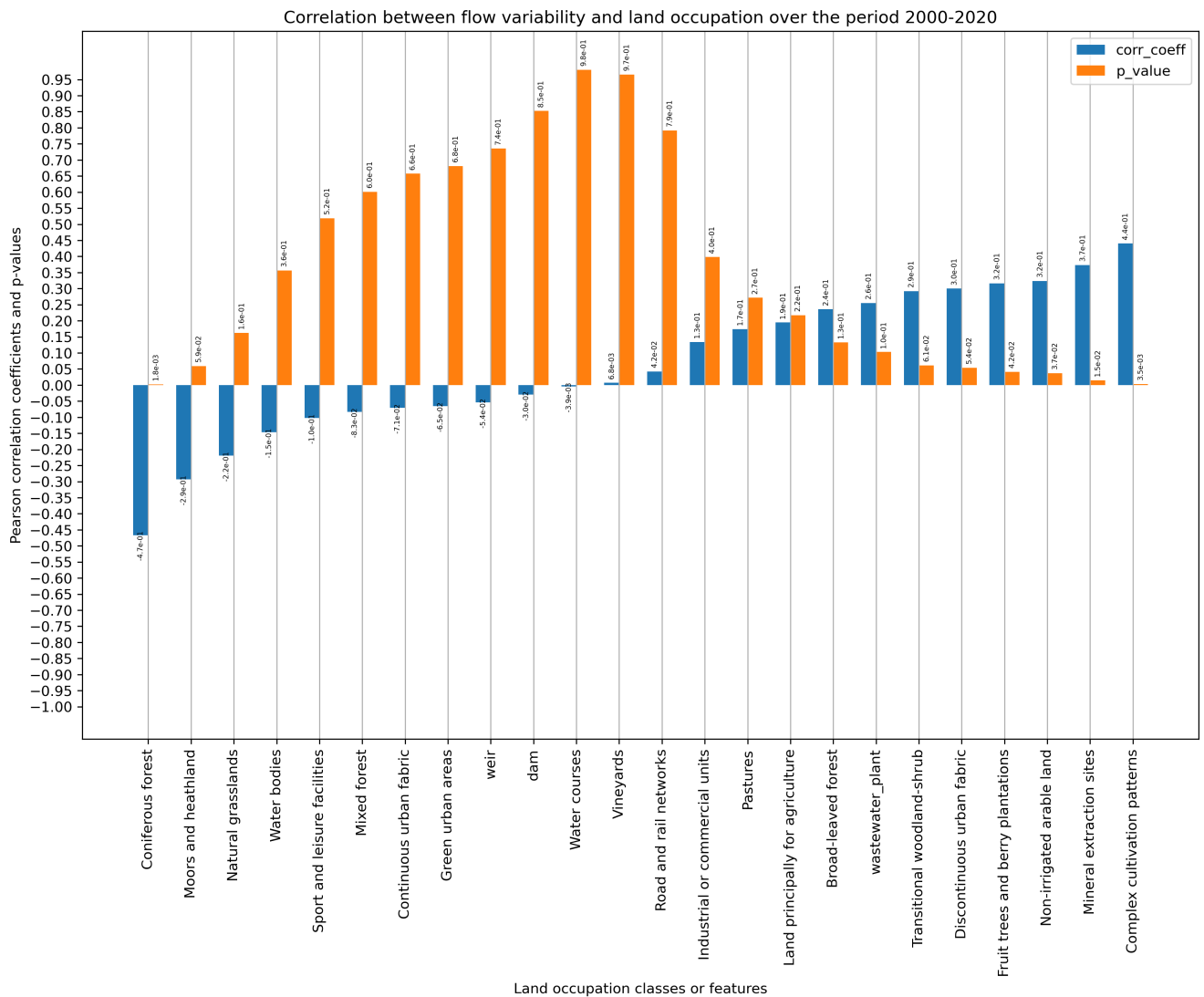


Figure 4.13 – Isolement des variables corrélées à la variabilité des débits (Sud-Loire, 2000-2020)

L’histogramme reproduit en Figure 4.13 permet d’identifier les éléments anthropiques les plus influents sur la variabilité du débit du territoire. Ceux-ci sont classés par ordre croissant d’influence. Lorsque le coefficient de Pearson atteint une valeur d’environ 0.4 (en plus ou en moins), la variable peut être considérée comme influente, et significativement influente lorsque la p-value associée est inférieure à 0.05. Les barrages et les seuils sur le territoire ont un coefficient de corrélation compris entre 0 et -5%, et ne peuvent ainsi être tenu pour responsables de la variabilité enregistrée aux stations, contrairement à l’hypothèse qui préside au choix de modélisation retenu dans (ROCKSTROM et al. 2023). Pour les stations d’épuration, le coefficient de corrélation atteint +26%, ce qui ne permet pas de conclure fermement quant à leur impact sur la variabilité, bien que celui-ci commence à être visible. Il est possible qu’à une échelle plus fine (e.g. des bassins versants de taille plus petite), leur effet soit plus net avec un coefficient de corrélation plus élevé : cette première identification pourrait justifier que l’on s’intéresse

davantage à cet impact.

Quant aux différentes classes d'occupation du sol, on peut distinguer plusieurs groupes :

- les classes d'occupation du sol fortement influentes : il s'agit des classes dont le coefficient de corrélation est supérieur à 40% ou inférieur à -40%. On retrouve les forêts de conifères à -47% ("*coniferous forest*") et les systèmes agricoles mixtes à +44% ("*complex cultivation patterns*", typiquement alternance de cultures en open field et bocage).
- les classes d'occupation du sol potentiellement influentes : il s'agit des classes dont le coefficient de corrélation est compris entre 30 et 40% (en positif ou en négatif). On retrouve ici, du côté positif (i.e. influence de la variabilité du débit à la hausse), à +29% les espaces forestiers clairsemés ("*transitional woodland/shrub*" : typiquement reforestation ou coupes rases), à +30% les espaces urbanisés ("*discontinuous urban fabric*" : typiquement les zones pavillonnaires), à +32% les vergers ("*fruit trees and berry plantations*"), à +32% également les surfaces cultivées non-irriguées ("*non-irrigated arable land*"), à +37% les carrières ("*mineral extraction sites*"); et du côté négatif, on retrouve les landes herbacées d'altitude ("*moors and heathlands*") à -29%.
- les classe d'occupation du sol vraisemblablement peu influentes : il s'agit des classes dont le coefficient de corrélation est compris entre -30% et +30%. On retrouve ici le reste des classes qui n'ont pas été mentionnées précédemment.

Cette analyse spatiale n'a pas la prétention d'être exhaustive ni définitive quant à ses conclusions. Elle permet de souligner en revanche l'influence potentielle, mais qui reste à confirmer par d'autres études plus approfondies, de variables a priori peu considérées. C'est le cas par exemple des carrières, qui pourraient présenter un impact non négligeable sur la variabilité des débits et donc sur l'intégrité des cours d'eau. La corrélation ne permet pas d'ailleurs d'expliquer le phénomène, mais on pourrait songer aux prélèvements d'eau effectués par les activités minières ou encore à l'altération du paysage qui, en réduisant les boisements, renforce le ruissellement lors de fortes intempéries. Il en va de même pour l'exploitation forestière, dont les coupes rases pourraient bien avoir un effet non négligeable à la hausse sur la variabilité des cours d'eau. L'agriculture serait quant à elle bien le premier responsable sur le territoire de la variabilité des débits et donc de l'intégrité quantitative des cours d'eau, davantage que les espaces urbanisés.

Plus généralement, les paramètres choisis par (ROCKSTROM et al. 2023) pour la modélisation du débit dit perturbé ne sont vraisemblablement pas adaptés à la situation du territoire : les

barrages semblent avoir un effet négligeable à cette échelle d’analyse, tandis que l’agriculture a bel et bien un effet significatif, mais le lien de causalité semble moins relever des prélèvements pour l’irrigation que de l’altération des paysages. D’autres paramètres influents ne sont en revanche pas compris dans la modélisation de (ROCKSTROM et al. 2023), comme les sites d’extraction minière ou les espaces forestiers clairsemés résultant d’un mode d’exploitation forestière par coupes rases.

4.4 Assimiler le modèle hydrologique global WBM aux données de débits Vigicrues : le logiciel *relio*

relio est le nom donné au logiciel libre sans interface graphique que nous avons conçu pour implémenter le modèle de rééchantillonnage hydrologique (MRH) formalisé au chapitre (3) et utilisé pour l’assimilation de données de mesures de débits du réseau Vigicrues aux sorties du modèle hydrologique global *Water Balance Model* (WBM). Ce logiciel a été développé spécifiquement pour une application aux données Vigicrues et pour un modèle hydrologique global dont la résolution source est moins précise que la résolution cible.

Cette section présente la mise en œuvre numérique du MRH dans le logiciel *relio*. La mise en œuvre se veut la plus indépendante possible du cas d’étude pour lequel elle a été effectuée, bien que les résultats présentés ci-après portent sur le rééchantillonnage spécifique du modèle WBM et leur assimilation aux données de mesure du réseau Vigicrues.

L’application proposée entend procéder à plusieurs traitements :

- rééchantillonner le raster fourni par le modèle WBM, de résolution $7000 \times 7000\text{m}$ (appartenant au domaine que nous avons appelé \mathcal{D}_{src} au chapitre 3), en un raster de résolution $30 \times 30\text{m}$ (appartenant au domaine \mathcal{D}_{dst})
- rééchantillonner les données ponctuelles de mesure de débits des stations Vigicrues dans le raster de résolution $30 \times 30\text{m}$.

traitements initiaux effectués par le MHE ainsi que les trois algorithmes centraux du modèle MRH. Remarquons que 2 scripts sont dédiés au rééchantillonnage des données Vigicrues et des données du WBM par l’algorithme `invdist` de `gdal`. Ces scripts ne font pas partie du MRH, qui justement a été développé en opposition aux algorithmes continus de rééchantillonnage, mais ont tout de même été intégrés au logiciel `relio` dans le but d’effectuer la validation croisée du MRH.

L’ordre séquentiel d’exécution des scripts est fourni dans un tableau accessible sur le dépôt GitHub de `relio` (voir en ligne : <https://github.com/Quentin-Dassibat/relio>). Nous proposons, en annexe au présent manuscrit, une documentation technique sur le fonctionnement de chacun des algorithmes. L’annexe fournit également une description des algorithmes permettant d’effectuer la paramétrisation du modèle MRH : l’initialisation aux bornes du raster d’accumulation de flux pour le paramètre β_p et le filtrage du raster des exutoires pour le paramètre α_b .

Nous précisons simplement ici l’origine des données que nous manipulons dans `relio`.

Les images raster du WBM définies dans \mathcal{D}_{src} sont issues du *supplementary materials* à l’article de (ROCKSTROM et al. 2023) :

- le raster à 12 bandes des débits mensuels moyens perturbés (DIST)
- le raster à 12 bandes des débits mensuels moyens non-perturbés (PRIST)

Quant aux images raster de \mathcal{D}_{dst} , elles proviennent de plusieurs traitements, qui sont soit externes, soit effectués par le MHE, soit effectués par le MRH :

- les images raster issues de sources externes :
 - un Modèle Numérique de Terrain (MNT), en l’occurrence de résolution 30×30 m. Une résolution faible est importante pour assurer ensuite une meilleure performance de jointure (minimisation de δ) dans l’algorithme de rasterisation. Une résolution de 1×1 m aurait pu être retenue car le produit est disponible pour le territoire d’étude mais la machine avec laquelle les scripts ont été exécutés n’offraient pas une puissance de calcul (en fait surtout capacité de mémoire) suffisante, bien qu’il aurait pu être envisagé d’externaliser les calculs sur serveur distant.
- les images raster issues du modèle hydrographique élémentaire (MHE)
 - un raster de mise en charge initiale : ce raster peut être envisagé comme une pluie unitaire et uniforme s’abattant sur la zone d’étude. Tout pixel du raster reçoit la

- valeur de 1, qui vaut charge initiale avant transport hydraulique de l'amont vers l'aval
- un raster de direction de flux : il s'agit d'un raster qui donne le sens d'écoulement sur chaque pixel en fonction du pendage le plus fort avec ses pixels adjacents
 - un raster d'accumulation de flux brute : ce raster répercute la charge initiale de chaque pixel de l'amont vers l'aval, par convolution linéaire
 - un raster d'accumulation de flux filtrée : ce raster est identique à l'accumulation de flux brute mais ne sont conservés que les pixels dont la charge est supérieure à une charge seuil, que nous avons notée α_a
 - un raster des exutoires bruts : il s'agit d'une restriction du raster d'accumulation de flux correspondant aux pixels définis comme exutoires dans le raster de direction de flux
- les images raster issues de la paramétrisation spécifique au modèle MRH (voir Figure ??) :
 - un raster d'accumulation de flux initialisée aux bornes : ce raster est identique au raster d'accumulation de flux brute mais certains pixels en bordure de la zone d'étude sont initialisés avec une charge qui tient compte de la charge héritée à l'amont de la zone d'étude. Ce raster met en oeuvre le paramètre β_p qui est le vecteur des conditions initiales aux bornes
 - un raster des exutoires filtrés : ce raster est identique au raster des exutoires brutes mais ne sont conservés que les pixels dont la charge est supérieure à une charge seuil, que nous avons notée α_b
 - un raster de bassins versants délimités : ce raster délimite les bassins versants associés au raster des exutoires filtrés
 - images raster en sortie du modèle MRH (voir Figure 4.14)
 - un raster des données WBM rééchantillonnées dans le domaine de destination : ce raster est le résultat des algorithmes de dérasterisation et de convolution du modèle MRH appliqués aux données du WBM
 - un raster des données de mesures Vigicrues reprojctées dans le domaine de destination : ce raster est le résultat de l'algorithme de rasterisation du modèle MRH appliqué aux données Vigicrues
 - un raster des données de mesures Vigicrues rééchantillonnées (reprojctées et interpolées) dans le domaine de destination : ce raster est le résultat des algorithmes de rasterisation et de convolution de modèle MRH appliqués aux données Vigicrues

4.4.2 Mise en œuvre de l’algorithme `invdist` pour la comparaison à `relio`

4.4.2.0.1 Rééchantillonnage des données Vigicrues par `invdist`

Le script `CONTINUOUSresamplingOfStations.py` effectue le rééchantillonnage des valeurs de débits de 10 stations de mesure selon la méthode `invdist`. Nous avons choisis ces stations de mesure parmi l’ensemble des stations de mesure selon un critère de représentativité statistique. En effet, à la façon d’une campagne de mesure qui serait menée sur le territoire, on viserait à obtenir un échantillon représentatif des cours d’eau en présence. Ainsi, pour chaque mois de la période 2000-2020, on récupère une station par décile de distribution des débits : la station correspondant au 10% des débits les plus faibles, puis la stations des 20% des débits les plus faibles, etc.

Cette échantillon est ensuite rééchantillonné vers \mathcal{D}_{dst} selon la méthode `invdist`. Pour cela, nous avons développé une suite de fonctions qui convertissent d’abord l’échantillon des 10 stations en image raster virtuelle (format `.vrt`, compatible avec l’interpolateur natif de `gdal` qui sera utilisé ensuite) puis interpolent ce raster dans \mathcal{D}_{dst} . l’interpolation se fait au moyen de la fonction `relio.interpolate` qui s’appuie sur la fonction `gdal_grid` dont nous donnons ci-après la commande simplifiée pour mieux en saisir les paramètres principaux :

```
gdal_grid -a -a_srs -txe -tye -outsize
```

Le paramètre `-a` introduit le nom de l’algorithme d’interpolation qui sera utilisé. ici, nous utilisons l’algorithme `invdist` de `gdal` bien qu’il en existe d’autres. Cet algorithme est paramétrable, notamment pour ce qui est du rayon de recherche et du facteur de lissage, mais nous ne modifions par les paramètres par défaut. Remarquons que nous avons développé le script `CONTINUOUSresamplingOfStations.py` de telle sorte à permettre de modifier le paramètre `-a` au besoin, ainsi que de modifier les paramètres. Le reste des paramètres permet de spécifier à l’algorithme que l’interpolation doit s’effectuer dans \mathcal{D}_{dst} . Pour cela, la valeur de ces paramètres est héritée du MNT, qui est lui-même défini dans \mathcal{D}_{dst} .

A la suite de cette opération, il reste encore à comparer deux-à-deux les pixels du raster rééchantillonné obtenu en sortie avec les points des stations de mesure. Le premier raster est

pour cela d'abord converti en couche vectorielle de géométrie `Point`. On charge ensuite la couche vectorielle des stations de mesure puis on effectue une jointure spatiale directe par plus proche voisin avec la librairie `geopandas` et sa fonction dédiée `geoDataFrame.sjoin_nearest` dont l'attribution unique (un pixel est attribué à une unique station) est assuré par l'option `how='inner'` qui permet une jointure interne sur la couche des stations. Remarquons que le raster des pixels rééchantillonné est continu : à tout pixel de \mathcal{D}_{dst} a été attribué une valeur, indépendamment du fait que le pixel sous-jacent fasse partie ou non d'un tronçon hydrographique. Il n'y a donc pas besoin d'effectuer l'étape intermédiaire de rasterisation des stations de mesure vers \mathcal{D}_{dst} puisque l'espace est continu en tout point (formellement, pixel).

Enfin, on convertit les débits des stations dans l'unité des débits rééchantillonnés, de la même façon que nous l'avons fait pour l'évaluation du groupe 2.

4.4.2.0.2 Rééchantillonnage des données WBM par `invdist`

Le rééchantillonnage des données WBM par `invdist` est conduit dans le script `CONTINUOUSresamplingOfWBM.py` qui fonctionne à la façon du script que nous présentions dans la section précédente. La différence réside dans les données d'entrée : il ne s'agit plus de rééchantillonner un échantillon des stations de mesure mais directement les pixels du modèle WBM définis dans \mathcal{D}_{src} . Conformément à ce que nous décrivions sur le fonctionnement des algorithmes de rééchantillonnage conventionnels, qui utilisent le barycentre des pixels comme point de départ de l'interpolation, nous commençons par extraire les barycentres des pixels du raster WBM (voir les étapes successives sur la Figure 4.15). Cette couche de points est ensuite convertie en raster virtuel puis interpolée à l'aide de la fonction `relio.interpolate`. De même, nous effectuons ensuite une jointure par plus proche voisin avec la couche des stations de mesure puis nous exprimons dans une unité commune les valeurs de débits.

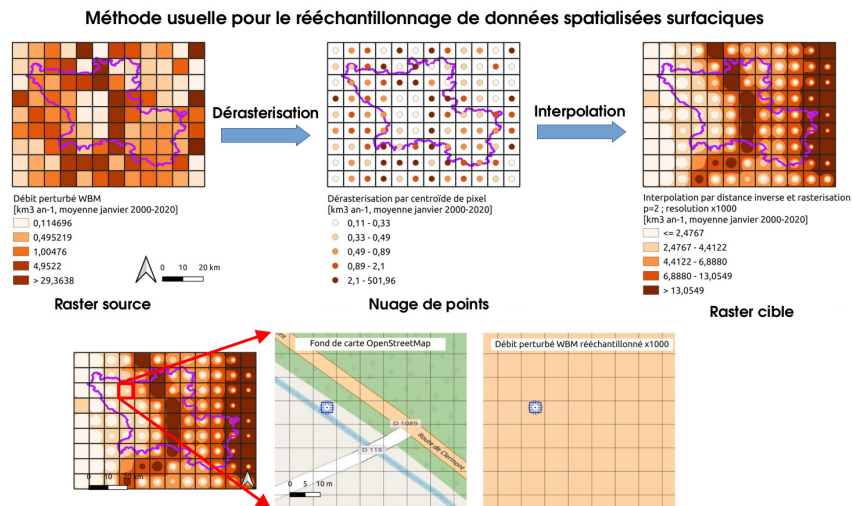


Figure 4.15 – Séquences des traitements effectués pour le rééchantillonnage d'un raster WBM selon l'algorithme de distance inverse.

4.4.3 Résultats

Cette partie présente les résultats liés au rééchantillonnage des sorties du modèle WBM et à la performance du modèle MRH lui-même, en reprenant les groupes de résultats évoqués plus haut. Cette partie vise ainsi à discuter de la possibilité d'un couplage entre les données de mesure sur site du réseau Vigicrues et les données carroyées du modèle WBM.

Cette partie se décompose deux sections. La première présente les résultats relatifs aux performances de **relio**, c'est-à-dire les résultats issus des groupes (1) et (3) décrits ci-dessus. La deuxième section s'intéresse aux résultats du couplage entre données de modèle carroyées et données de mesure ponctuelle, en s'appuyant sur les résultats de la section précédente qui donneront une mesure (partielle) de l'erreur propre à **relio**. Cette section reprendra ainsi les résultats issus des groupes (2) et (4).

4.4.3.1 Performances de **relio**

L'évaluation des performances de **relio** s'opère de deux façons.

Par une évaluation directe d'une part, en utilisant la méthode pour rééchantillonner directement un sous-ensemble des stations Vigicrues dont le résultat sera comparé au reste des stations servant de valeur de référence. Cette vérification ne permet en réalité d'évaluer les performances de **relio** que vis-à-vis d'un seul des trois algorithmes centraux qui le constituent : l'algorithme de convolution hydrographique linéaire. En effet, l'algorithme de

rasterisation des données Vigicrues vers la grille de raster définie dans \mathcal{D}_{dst} est un préalable à l'évaluation directe que nous souhaitons menée. L'erreur que nous constaterons entre les données rééchantillonnées et les données observées hérite de l'erreur propre à cet algorithme, à laquelle s'ajoute l'erreur propre à l'algorithme de convolution par lequel les données sont à proprement parler rééchantillonnées. Autrement dit l'évaluation directe que nous allons mener nous permettra d'obtenir une valeur d'erreur (ou inversement, de performance) indistinctement pour les algorithmes de convolution et de rasterisation. L'algorithme de dérasterisation ne peut en revanche pas être évalué par cette méthode directe puisqu'il n'est pas impliqué dans sa réalisation. C'est la raison pour laquelle, à défaut de pouvoir évaluer l'erreur pour les trois algorithmes d'un coup, nous avons adopté une deuxième méthode, indirecte, par validation croisée.

L'évaluation indirecte, d'autre part, propose une approche alternative à la quantification des erreurs "absolues" de `relio`, qui ne peut pas être conduite car l'algorithme de dérasterisation échappe à la méthode par validation directe. La méthode par validation croisée ne permettra pas d'apporter une valeur d'erreur par rapport à des données réelles ("absolue") mais par rapport au modèle conventionnel en SIG de l'interpolation par distance inverse.

4.4.3.2 Validation directe de `relio`

Commençons par redonner la formalisation mathématique du modèle MRH :

$$y_{i,t} := q(\chi_i, x_{j,t} : \alpha_a, \alpha_b, \beta_p, \delta)$$

Où χ_i désigne la méthode d'interpolation à proprement parler (il s'agit de l'algorithme de convolution hydrographique linéaire), $x_{j,t}$ est la valeur de débit pour le mois t spatialement localisée en j dans l'espace \mathcal{D}_{src} et $y_{i,t}$ est la valeur de débit rééchantillonnée (ou prédite) pour le mois t spatialement localisée en i dans l'espace \mathcal{D}_{dst} .

χ_i ne fonctionne que dans l'espace \mathcal{D}_{dst} . C'est la raison pour laquelle les algorithmes de rasterisation et de dérasterisation ont été développés.

En particulier pour l'algorithme de rasterisation, qui nous intéresse dans cette section, il

effectue le passage de l'espace dans lequel est définie la couche vectorielle des stations (noté \mathcal{E}) vers \mathcal{D}_{dst} :

$$\begin{aligned}\mathcal{E} &\rightarrow \mathcal{D}_{dst} \\ x_{j,t} &\rightarrow x_{i,t}\end{aligned}$$

Contrairement à l'algorithme de dé rasterisation qui modifie la valeur de x lors de la projection de j vers i ($x_{i,t}$ étant alors la valeur estimée $\widehat{x_{i,t}}$), l'algorithme de rasterisation préserve la valeur de x et effectue simplement la projection. La projection dépend en revanche de certains paramètres : α_a et β_p car les pixels i de destination appartiennent au raster d'accumulation de flux filtré au seuil α_a et initialisé aux bornes par β_p ; δ car la jointure entre les pixels i et les points j n'est pas exacte.

Plus concrètement, l'évaluation des performances de **relío** s'effectue à partir d'un échantillon des $(x_{i,t})$ qui vont être interpolés, que l'on peut noter \mathcal{I} , et le reste des $(x_{i,t})$ servira de référence, que l'on peut noter \mathcal{R} . L'évaluation des performances que nous allons conduire dans cette section consiste donc à comparer paire-à-paire les valeurs de \mathcal{R} et les valeurs rééchantillonnées de l'ensemble $(y_{i,t})$

Si nous parlons ici de "performances", l'évaluation portera spécifiquement sur une mesure de l'erreur. L'approche retenue est de comparer les valeurs rééchantillonnées des $(y_{i,t})$ avec les valeurs observées (mais reprojctées) de \mathcal{R} . On se propose pour cela d'employer deux métriques classiquement utilisées pour évaluer les erreurs structurelles des modèles : le biais moyen et le résidu. Les graphiques présentés dans les sections suivantes sont issus du script `PLOT.py`.

En première analyse, on peut déjà représenter le nuage des $(y_{i,t})$ par rapport aux $(x_{i,t})$ de \mathcal{R} (voir Figure 4.16).

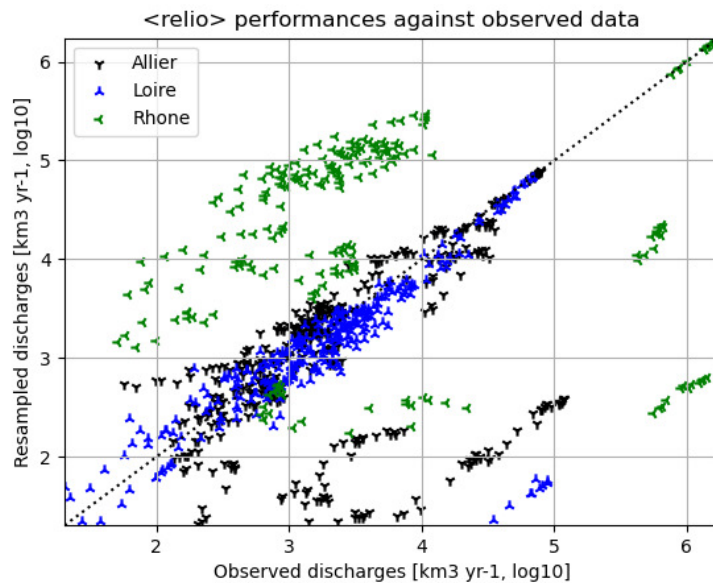


Figure 4.16 – Performances de *relio* par rapport à des données de contrôle

La droite $y = x$ qui est tracée en pointillés permet d'effectuer une première lecture sur le niveau de performance *relio* pour les trois principaux bassins versants de la zone d'étude (la Loire, le Rhône et l'Allier). L'échelle est en \log_{10} pour permettre la visualisation de toutes les valeurs sur un intervalle restreint. On peut remarquer d'une part le sens de variation des nuages de points lorsque l'axe Ox augmente par rapport au sens de variation $y = x$, et d'autre part la position des nuages par rapport à la droite $y = x$. Pour les trois bassins versants, le sens de variation est croissant. Cela signifie que le modèle rend bien compte de l'accumulation de charge entre l'amont et l'aval des bassins versants. En revanche, pour ce qui est d'approcher au plus près la valeur réelle de débit en chaque point, le modèle performe différemment selon les bassins. Le nuage de points du bassin versant de la Loire est ainsi relativement resserré autour de $y = x$, ce qui signifie que les valeurs estimées ne sont globalement ni sous-estimées ni sur-estimées. Le constat vaut également pour le bassin versant de l'Allier bien qu'un nombre important de points se trouve en-dessous de l'axe $y = x$, ce qui signifie que le modèle tend à sous-estimer les débits réels. A l'inverse pour le bassin versant du Rhône, les débits sont sur-estimés par le modèle et très peu de valeurs sont proches de $y = x$, laissant à penser que le modèle performe mal pour ce bassin.

Étude du biais moyen

Le biais moyen (MBE, *Mean Bias Error*) est une métrique qui renseigne sur l'écart relatif entre la valeur prédite par un modèle et la valeur observée correspondante ("relatif", car normé

par la valeur observée).

Le MBE est calculé avec la formule suivante :

$$\forall (i, t : x_{i,t} \in \mathcal{R}), \quad MBE_{i,t} = \frac{y_{i,t} - x_{i,t}}{x_{i,t}}$$

La dimension des $(MBE_{i,t})$ est importante puisqu'il y a une valeur par date et par position. On peut alors en représenter la distribution par la fonction de densité (voir Figure 4.17).

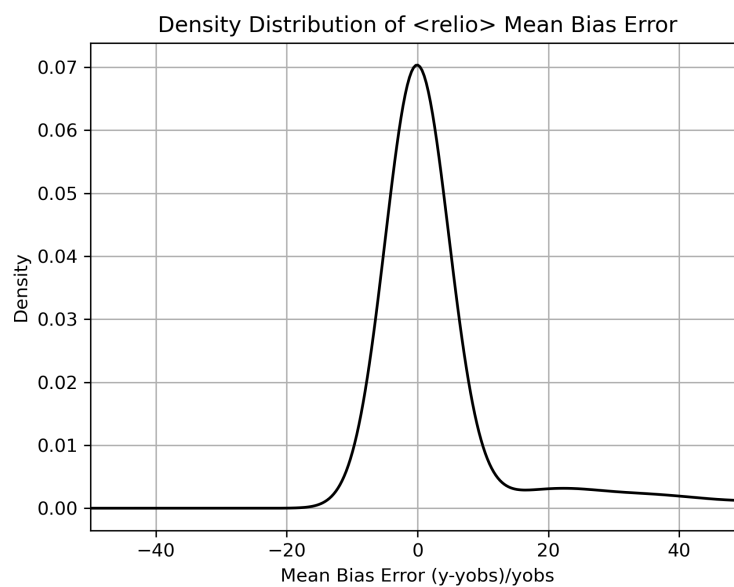


Figure 4.17 – Distribution du MBE de relio

En première observation, la distribution est centrée sur 0, signifiant que le modèle n'est pas biaisé ni du côté de la sous-estimation, ni du côté de la surestimation. Plus précisément, la densité pique est atteinte en $x = -0.03$ pour une probabilité associée de $y = 0.07$.

La valeur moyenne des $MBE_{i,t}$ (notée ϵ) peut être considérée comme une norme d'erreur pour les sorties de **relio**. Elle calculée à l'aide de la formule suivante :

$$\epsilon = \frac{\sum_i \sum_t MBE_{i,t}}{\text{card}_{i,t}(MBE_{i,t})}$$

ϵ vaut environ 1% pour le bassin versant de l'Allier et 19% environ pour la Loire. Le bassin versant du Rhône en revanche présente une erreur très importante ($\epsilon = 2700\%$). Dans la Figure 4.18, est représenté par un trait jaune le MBE médian, qui correspond à la moyenne si la série

est symétrique.

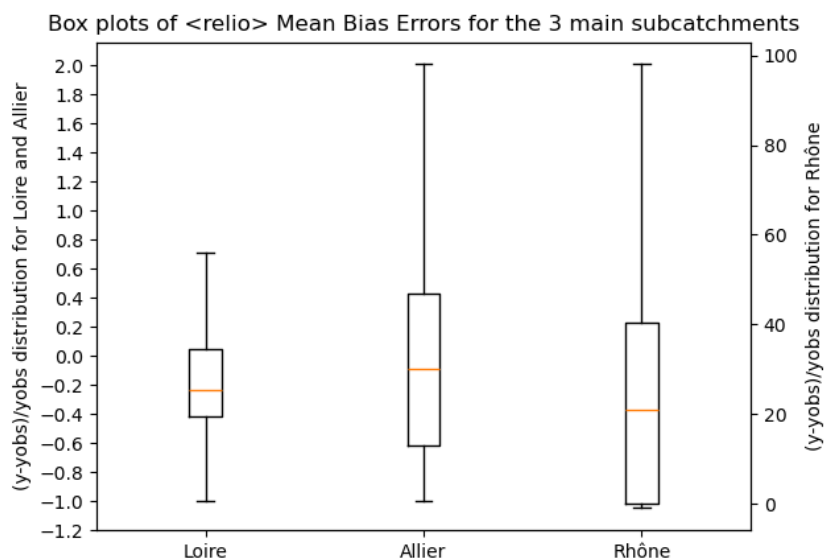


Figure 4.18 – Descripteurs statistiques des distributions du MBE pour les trois principaux bassins versants du territoire

Les MBE médians des bassins versants ne coïncident pas nécessairement avec les MBE moyens. Le bassin versant de la Loire a un MBE médian de -22%, celui de l’Allier de -10% et celui du Rhône de 2000%. Les performances pour les bassins versants de la Loire et de l’Allier sont bonnes en médiane, et il est intéressant de noter que leur MBE médian est négatif, tandis que leur MBE moyen est positif, ce qui signifie que parmi les valeurs extrêmes (points aberrants), les surestimations sont sur-représentées par rapport aux sous-estimations. On préférera donc retenir ces valeurs de MBE médian comme valeur normale d’erreur pour le `relio`.

La ventilation du MBE pour les trois principaux bassins versants de la zone d’étude montre ainsi des résultats très contrastés entre les bassins versants dont la charge transportée à l’amont de la zone d’étude est relativement faible (zone d’étude en tête de bassin), à savoir la Loire et l’Allier, et le bassin versant transportant sur la zone d’étude une charge très importante, à savoir le Rhône. Ce constat témoigne de la forte sensibilité du paramètre β_p dans les performances du modèle.

Étude des résidus

L’étude des résidus permet plus spécifiquement d’étudier le comportement des erreurs de

modèle. On peut distinguer deux types de résidus : le résidu additif et le résidu multiplicatif, tous deux définis à partir du même jeu de paires $(y_{i,t}, x_{i,t})$.

Le résidu additif consiste à estimer la valeur $r_{i,t}$ définie comme $y_{i,t} = x_{i,t} + r_{i,t}$. Il permet d'évaluer si l'erreur peut être corrigée par l'ajout d'un terme. Formellement, le résidu est une variable aléatoire qui dépend du tirage des i et des t . On détermine numériquement la variable aléatoire ϵ en traçant sa fonction de densité (voir Figure ??) selon la formule suivante :

$$\epsilon = (r_{i,t}) = (y_{i,t} - x_{i,t})$$

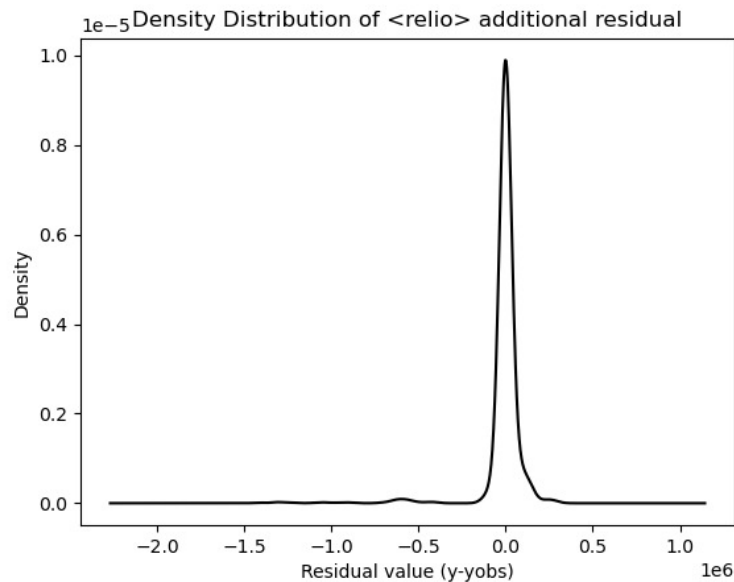


Figure 4.19 – Distribution du résidu additif de `relio` pour tous les mois et toutes les stations

Quant au résidu multiplicatif, il est défini comme la valeur $r_{i,t}$ telle que $y_{i,t} = x_{i,t} \cdot r_{i,t}$. Il permet d'évaluer si l'erreur du modèle peut être corrigée par un facteur. De même on détermine numériquement la variable aléatoire ϵ en traçant sa fonction de densité (voir Figure 4.20), qu'on normalise par les $x_{i,t}$ en ajoutant le terme -1 :

$$\epsilon = (r_{i,t}) = \left(\frac{y_{i,t}}{x_{i,t}} - 1 \right)$$

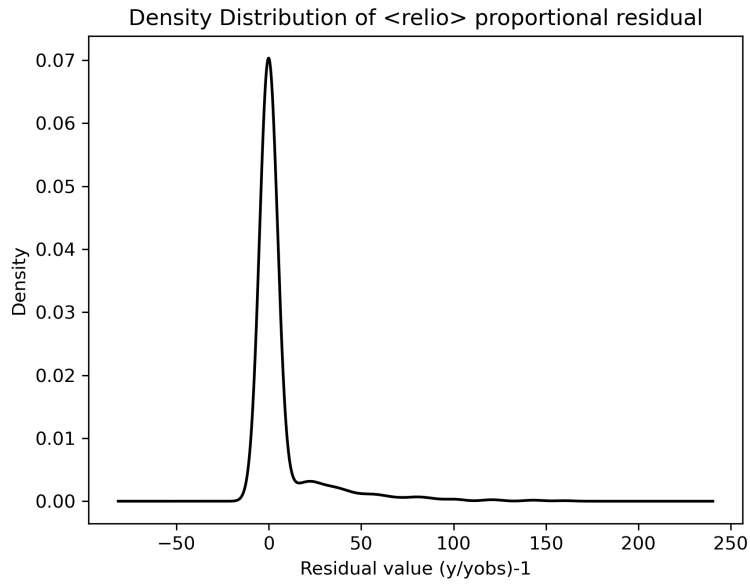


Figure 4.20 – Distribution du résidu multiplicatif de `relio` pour tous les mois et toutes les stations

La comparaison des fonctions de densité entre les résidus additifs et multiplicatifs montre que la probabilité pic des résidus multiplicatifs est très nettement supérieure à celle des résidus additifs (0.07 contre 0.00001). Le comportement de l’erreur de `relio` serait ainsi de type proportionnel. Ainsi dans la formalisation du modèle, le paramètre d’erreur pourrait figurer comme suit :

$$y_{i,t} := r_{i,t} \cdot q(\chi_i, x_{j,t} : \alpha_a, \alpha_b, \beta_p, \delta)$$

4.4.3.3 Validation croisée de `relio`

4.4.3.3.1 Rééchantillonnage de données de mesure

Nous allons employer les mêmes métriques que précédemment et positionner la méthode `invdist` par rapport à `relio`.

En première approche, on peut représenter les nuages de points des deux modèles l’un contre l’autre au regard de la droite $y = x$ de correspondance parfaite entre les débits estimés et les débits observés (voir Figure 4.21).

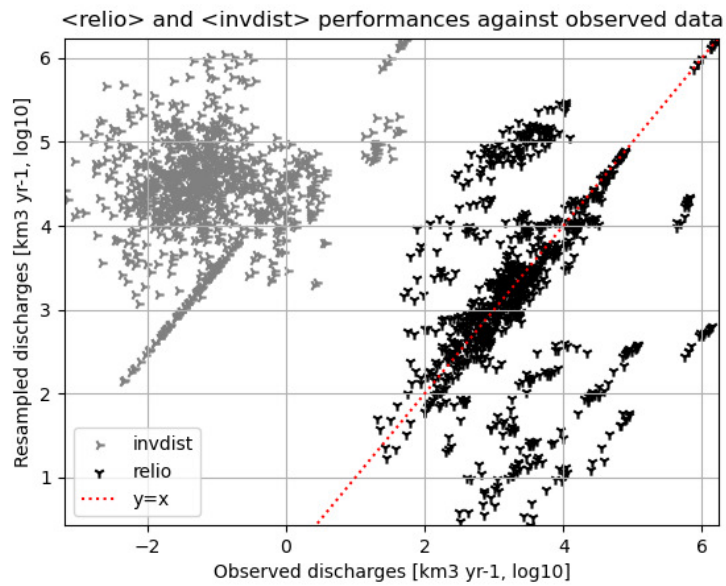


Figure 4.21 – Performances comparées de relio et invdist par rapport à des données de contrôle

On constate que **relio** performe nettement mieux que le modèle **invdist** car son nuage de point est davantage resserré autour de $y = x$. Les deux modèles ont un comportement croissant : plus le débit observé est important, plus le débit estimé l'est aussi, et ceci de façon a priori linéaire pour **relio** et non-linéaire pour **invdist** ($\mathcal{O}y$ semble augmenter plus que proportionnellement par rapport à $\mathcal{O}x$). De plus, **invdist** tend globalement à sur-estimer les valeurs de débit.

Étude du biais moyen

On donne dans la Figure 4.22 la distribution du MBE pour le modèle **invdist**.

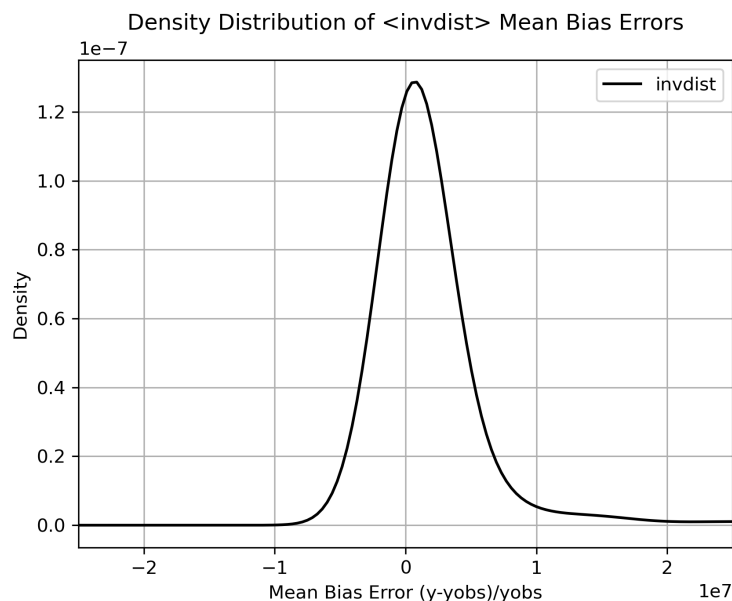


Figure 4.22 – Distribution du MBE de `invdist`

On remarque que le biais moyen de `invdist` n'est pas centré sur 0 mais décalé vers la droite ($x = 7.6e5$), ce qui confirme la tendance du modèle à la surestimation. On observe également la plus grande dispersion des valeurs estimées par `invdist` comparément à `relio` car la densité pique du biais moyen est bien moins forte pour `invdist`, de l'ordre de 0.00000013 contre 0.07 pour `relio`. C'est d'ailleurs pour cette raison que nous n'avons pas représenté les deux distributions dans le même graphe.

Étude des résidus

L'étude des résidus additifs et multiplicatifs comparés entre `relio` et `invdist` confirme les performances accrues de `relio` en matière d'erreur.

Le résidu additif de `relio` présente une distribution davantage resserrée autour d'une valeur pic dont la probabilité est très nettement accrue (voir Figure 4.23). Ce qui est encore davantage le cas avec les résidus multiplicatifs (voir Figure 4.24).

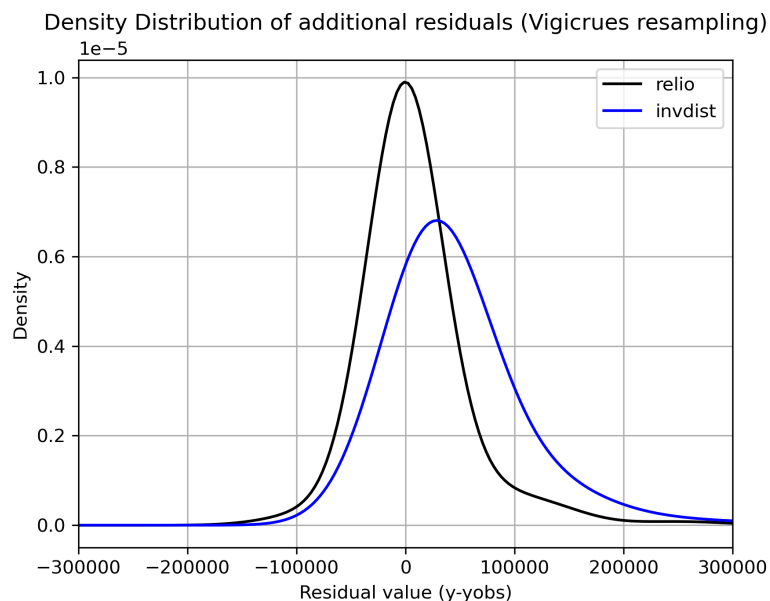


Figure 4.23 – Distributions comparées du résidu additif de relio et invdist

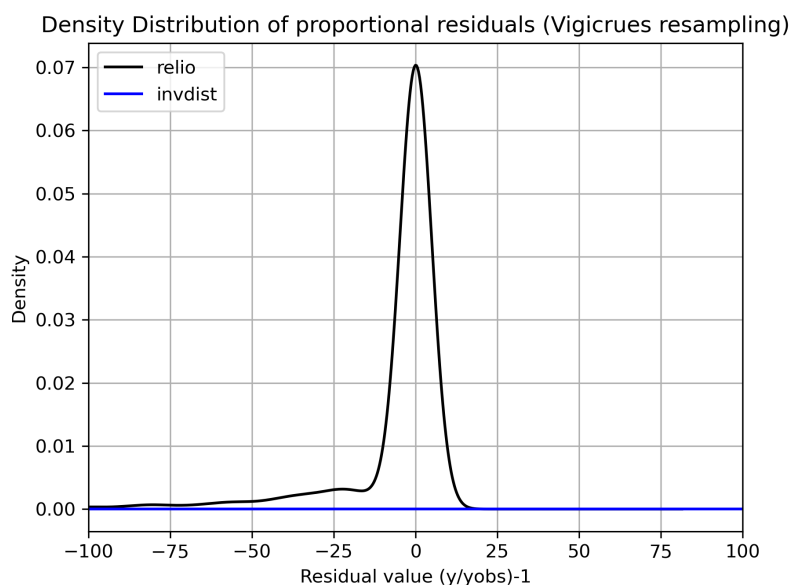


Figure 4.24 – Distributions comparées du résidu multiplicatif de relio et invdist

4.4.3.3.2 Rééchantillonnage des données carroyées du WBM

Nous allons à présent étudier le comportement comparé de `relio` et `invdist` lorsqu'il s'agit de rééchantillonner non plus des données de mesure mais les données modélisées du WBM (voir Figure 4.25), fort des conclusions sur la bonne performance globale de `relio`, bien que différente entre bassins versants (et très mauvaise pour le bassin versant du Rhône,

vraisemblablement à cause d'une mauvaise paramétrisation de β_p).

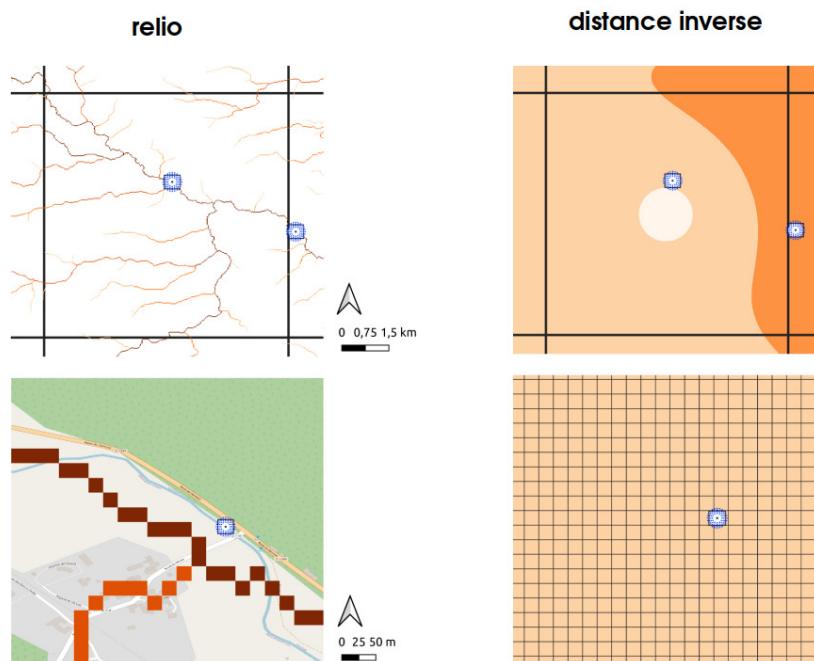


Figure 4.25 – Visualisation du rééchantillonnage des données carroyées WBM selon notre méthode hydrographiquement contrainte (**relio**) et selon une méthode continue (**invdist**)

L'évaluation des performances de **relio** sur le rééchantillonnage des données modélisées porte sur les trois algorithmes de dérasterisation, rasterisation et convolution, contrairement au rééchantillonnage des données de mesure qui ne fait intervenir que les deux derniers algorithmes.

L'algorithme de dérasterisation effectue le passage de l'espace de l'espace source des pixels de débits modélisés par le WBM (\mathcal{D}_{src}) vers \mathcal{D}_{dst} :

$$\begin{aligned} \mathcal{D}_{src} &\rightarrow \mathcal{D}_{dst} \\ x_{j,t} &\rightarrow \widehat{x}_{i,t} \end{aligned} \tag{4.1}$$

Contrairement à l'algorithme de rasterisation qui préserve la valeur de x lors de la projection de j vers i , la dérasterisation conduit à estimer la valeur de x dans l'espace de destination, $x_{i,t}$ étant alors la valeur estimée $\widehat{x}_{i,t}$.

Le modèle MRH, et son implémentation numérique **relio**, qui va être évalué ici s'écrit donc formellement comme :

$$y_{i,t} := q(\chi_i, \widehat{x}_{i,t} : \alpha_a, \alpha_b, \beta_p, \delta)$$

Le raster du modèle WBM qui est rééchantillonné ici correspond aux simulations perturbées (DIST).

En première approche, on peut représenter le nuage de points rééchantillonné par `relio` et `invdist` (voir Figure 4.26).

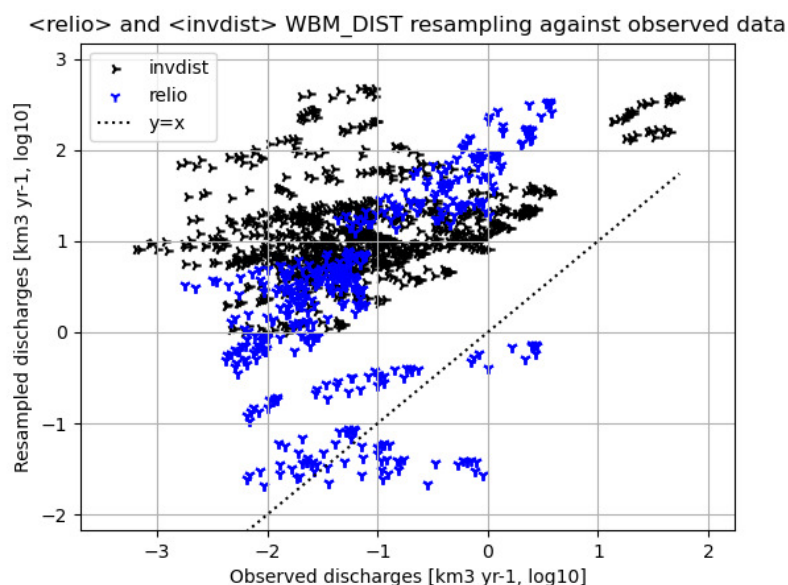


Figure 4.26 – Comparaison des valeurs de débits WBM rééchantillonnées par rapport aux données de mesure Vigicrues, pour `relio` et `invdist`

On remarque une mauvaise performance globale des données estimées par rapport aux données observées. Pour le rééchantillonnage `invdist`, on ne peut guère conclure entre la performance du rééchantillonnage et la performance prédictive du modèle WBM. En revanche, en vue des bonnes performances de `relio`, on peut légitimement questionner les performances du WBM pour cette échelle de territoire.

On a vu en effet dans la section précédente que `relio` est associée à une erreur de -22% pour le bassin versant de la Loire et -10% pour celui de l'Allier (MBE médian). Cette erreur prend en compte les algorithmes de rasterisation et de convolution. L'évaluation des performances pour le rééchantillonnage du WBM intègre en plus l'algorithme de dérasterisation. Si les deux évaluations ne sont ainsi pas parfaitement comparables, on peut toutefois noter que tous ces algorithmes reposent finalement sur le même jeu de paramètres, qui sont au fond les variables influençant l'erreur finale. En particulier, tous reposent sur le raster d'accumulation de flux, qui est le raster central de la méthode `relio`. L'algorithme de dérasterisation fonctionne sur le même principe que l'algorithme de convolution, à savoir attribuer une quantité connue de

débit au prorata d'une charge hydraulique théorique. Il est donc raisonnable de penser que l'erreur propre à l'algorithme de dérasterisation n'est pas bien différente de l'erreur que nous avons évaluée pour les algorithmes de convolution et de rasterisation, bien que nous ne puissions pas la quantifier. En l'occurrence, l'erreur qui apparaît pour le rééchantillonnage des données WBM dépasse de loin l'ordre de grandeur de l'erreur liée au rééchantillonnage des données de mesure, et la différence de performance entre les trois bassins versants est complètement renversée (voir Figure 4.27) : le Rhône devient le bassin versant le plus performant (tout en restant très mauvais), ce qui ne peut guère s'expliquer par l'utilisation de l'algorithme de dérasterisation qui, encore une fois, repose sur le même jeu de paramètres que les deux autres algorithmes. Il est donc raisonnable de penser que la dégradation des performances de rééchantillonnage s'explique moins par le rééchantillonnage lui-même que par une défaillance du modèle WBM à cette échelle d'étude.

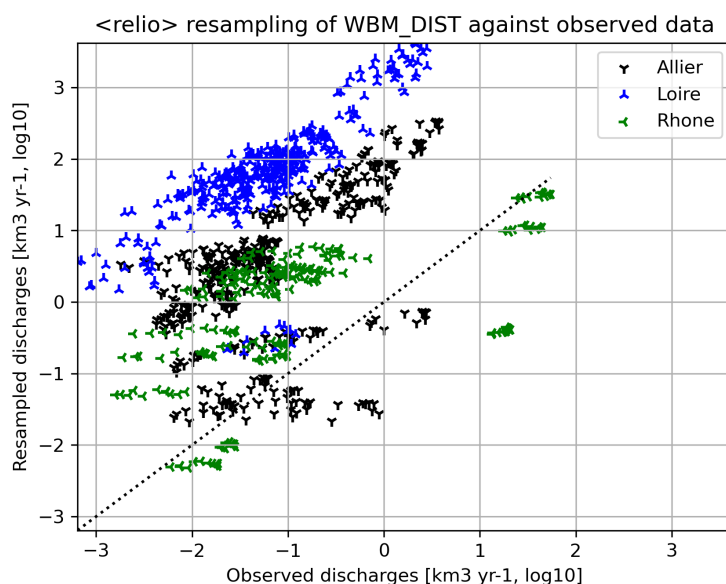


Figure 4.27 – Comparaison des valeurs de débits WBM rééchantillonnées par relio par rapport aux données de mesure Vigicrues

Étude du biais moyen

L'étude du biais moyen confirme l'hypothèse que nous formulons dans la section précédente.

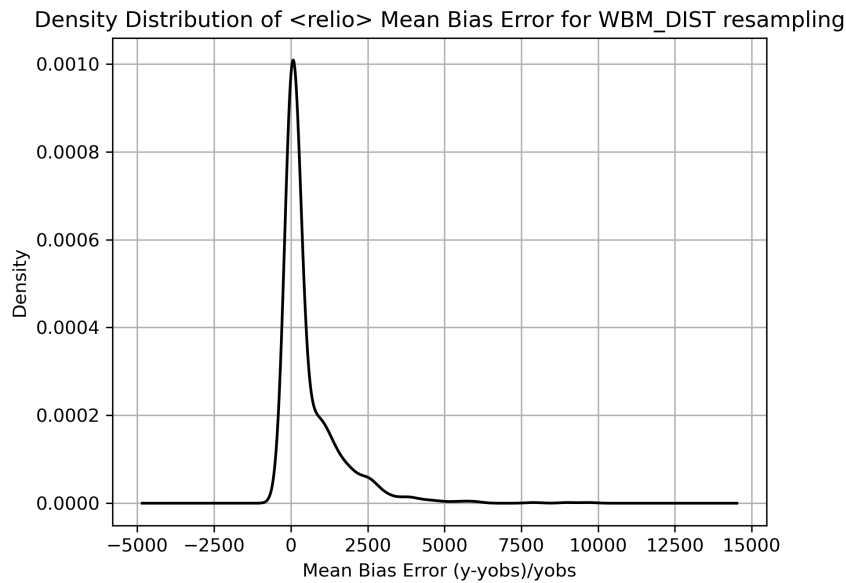


Figure 4.28 – Distribution du MBE pour les valeurs de débits du WBM rééchantillonnées par `relio`

Le pic de densité est atteint en $x = 76.5$ pour une probabilité de $y = 0.001$ (voir Figure 4.28). Rappelons que le biais moyen pour l'évaluation de `relio` sur le rééchantillonnage de données de mesure était de $x = -0.03$ et $y = 0.07$. Il apparaît alors peu probable que l'ajout de l'algorithme de dérasterisation, reposant sur le même jeu de paramètres que les algorithmes de rasterisation et de convolution, ait conduit à un écart de performance de plusieurs ordres de grandeur entre ces deux rééchantillonnages.

Étude des résidus

L'étude des résidus comparés entre `relio` et `invdist` va également dans le sens de ce que nous évoquions dans les sections précédentes. En effet, alors que `relio` présente une erreur moins biaisée et plus resserrée autour d'une valeur pic, à la fois pour les résidus additif et multiplicatif, lorsqu'il s'agit du rééchantillonnage des données WBM, `invdist` présente des performances accrues pour ce qui est des résidus additifs (voir Figure 4.29) et `relio` performe davantage sur les résidus multiplicatifs (voir Figure 4.30).

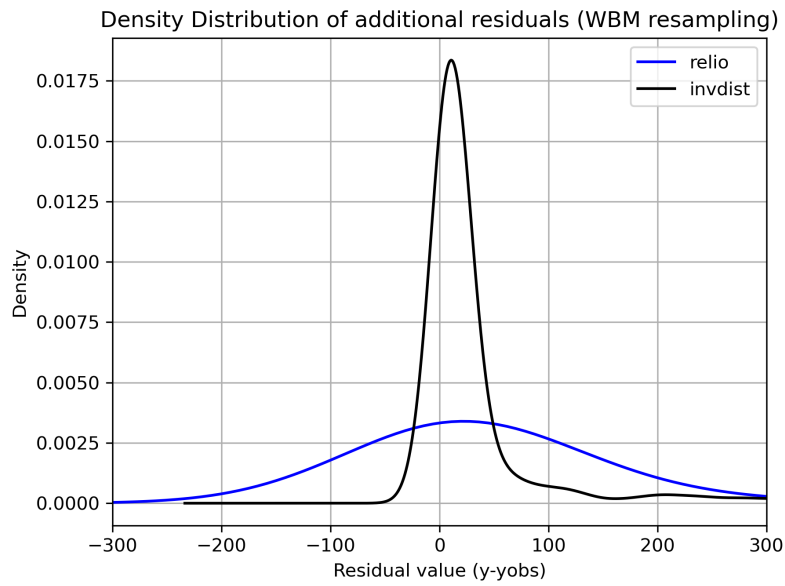


Figure 4.29 – Distribution comparée du résidu additif pour les valeurs de débits du WBM rééchantillonnées par relio et invdist

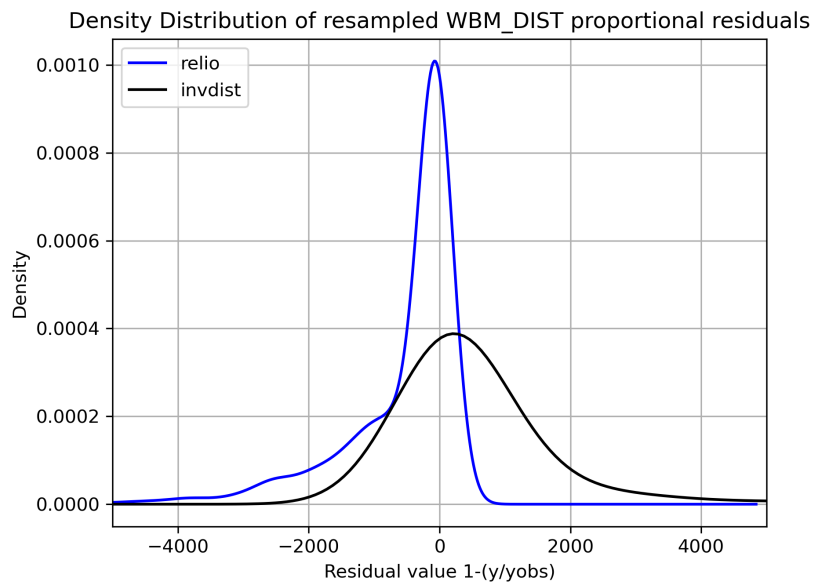


Figure 4.30 – Distribution comparée du résidu multiplicatif pour les valeurs de débits du WBM rééchantillonnées par relio et invdist

Conclusion

Ce chapitre a proposé une mise en œuvre numérique des deux propositions de recherche que nous avons formulées au chapitre 3 concernant la possibilité d'intégrer les données de mesure de débits dans une démarche AESA, en s'appuyant sur le cas d'étude de l'évaluation

environnementale de l'intégrité des cours d'eau sur le territoire du SCoT Sud-Loire par les données du réseau Vigicrues.

Les données du réseau Vigicrues peuvent être caractérisées de spatialement explicites et robustes, et correspondent ainsi au type 1 des données dans la typologie hiérarchique des données que nous avons présentée au chapitre 2. Elles sont spatialement explicites parce qu'il s'agit de mesures de débits sur site effectuées par des stations en cours d'eau (qui enregistrent des hauteurs d'eau converties en débits par courbes de tarage), et elles sont robustes parce que produites de façon systématisée, c'est-à-dire selon un protocole de validation croisée entre les services de prévision des crues (SPC) et les unités d'hydrométrie (UH). L'indicateur sur l'intégrité des cours d'eau proposé par (ROCKSTROM et al. 2023) est calculé, dans leur publication, à partir de données issues d'un modèle hydrologique global (le *Water Balance Model*). Ces données sont carroyées sur un maillage global de résolution environ $7 \text{ km} \times 7 \text{ km}$. Si elles peuvent donc être caractérisées de spatialement différenciées, elles ne sont que partiellement explicites. En effet, le WBM a été validé pour certaines stations de mesure hydrométrique du réseau GDRC (*Global Runoff Data Center*) qui correspondent aux grands bassins versants des masses d'eau continentales du globe (FEKETE et al. 2002). Pour le reste des carreaux qui n'interceptent pas une masse d'eau significative (typiquement lorsque le bassin versant associé au carreau est inférieur à $10\,000 \text{ km}^2$), se pose alors la question de la validité des résultats du WBM. Dans ces conditions, nous avons envisagé deux hypothèses de travail :

- est-il possible, d'une part, de calculer l'indicateur proposé par (ROCKSTROM et al. 2023) en se basant exclusivement sur des données de mesure ? Ceci dans le but de palier le caractère non explicite des données du WBM à l'échelle du territoire que nous étudions ;
- est-il possible, d'autre part, de coupler les données du WBM avec des données de mesure ? Ceci dans le but d'évaluer la validité du WBM sur le territoire et de doter l'indicateur construit à partir de données de mesure d'une référence véritablement absolue (sans pressions anthropiques).

Concernant la première hypothèse de travail, nous avons montré qu'il est en effet possible de construire l'indicateur de (ROCKSTROM et al. 2023) à partir des données Vigicrues au prix d'une inflexion sur le caractère *absolu* d'un tel indicateur. En effet, la particularité d'un indicateur de type "limite planétaire" est de comparer l'état actuel (ou anthropisé) d'un compartiment environnemental avec son état stable sous l'Holocène (ou non anthropisé). Pour ce qui est de l'indicateur utilisé dans (ROCKSTROM et al. 2023) pour les eaux de surface, la définition

de "l'état stable sous l'Holocène" ne correspond pas pour autant à un "état de nature". La morphologie des cours d'eau ainsi que les conditions climatiques utilisées dans les simulations du WBM correspondent en effet à l'état d'anthropisation actuel (2000-2020), considérant qu'il s'agit de variables pour lesquelles il est difficile de revenir à l'état de l'Holocène. En revanche, l'état non anthropisé est modélisé par l'absence de deux sources de pression (ou variables), que sont les prélèvements d'eau pour l'irrigation et la présence de barrages.

Les données Vigicrues ne permettent évidemment pas de remonter à l'époque de l'Holocène, ce qui n'est donc pas requis par l'indicateur proposé dans (ROCKSTROM et al. 2023), qui propose la démarche réaliste de considérer comme référence un état environnemental déjà dévié de l'Holocène. En revanche, les données Vigicrues ne permettent pas d'isoler ce qui, dans l'écoulement qu'elles mesurent, relève d'une pression liée aux barrages, d'une pression liée à l'irrigation, comme de toute autre source de pression. Les pressions mesurées sont en effet diffuses et difficilement attribuables à (ou explicables par) une source de pression spatialement localisée et quantitativement mesurée. Un état de référence environnemental basé sur des données de mesure ne peut donc pas représenter un état "sans pressions anthropiques".

L'indicateur que nous avons proposé est celui d'une anomalie de variabilité, qui décrit littéralement la situation d'une variabilité non normale au sens statistique, c'est-à-dire un écart à la norme de variabilité calculée comme la variabilité moyenne enregistrée par une station de mesure sur la période. En retenant le taux de 20% comme variabilité excessive (RICHTER et al. 2012), nous avons proposé de construire un indicateur de fréquence mensuelle et annuelle de dépassement. Pour la fréquence mensuelle, il s'agit de compter le nombre de fois où un même mois est en situation de variabilité excessive, par exemple, parmi tous les mois de janvier d'une chronique typique (2000-2020), ceux où la station hydrométrique enregistre une variabilité du débit mensuel (calculée comme l'écart-type des débits journaliers moyens) supérieure ou inférieure à 20% de la variabilité moyenne des mois de janvier sur la chronique. Pour ce qui est de la fréquence annuelle de dépassement, il s'agit de calculer la moyenne de la fréquence mensuelle de dépassement sur tous les mois de l'année.

Ainsi pour le Sud-Loire, les stations hydrométriques enregistrent une fréquence annuelle de dépassement comprise entre 56 et 77%, et en moyenne de 60% (moyenne sur toutes les stations). Ce dernier chiffre peut s'interpréter de la façon suivante : en moyenne sur le territoire, la variabilité du débit enregistrée par une station hydrométrique dépasse le corridor de variabilité

raisonnable ($\pm 20\%$) les 2/3 du temps sur la chronique 2000-2020. Le fait que les fréquences de dépassement pour les 42 stations de mesure du territoire soient relativement proches tient sans doute à ce que la majorité d'entre elles appartiennent au même bassin versant global de la Loire, et donc ont en commun la plupart des caractéristiques hydrologiques.

Nous avons souhaité identifier la présence (ou non) d'une tendance de long terme sur cette variabilité globalement excessive, qui pourrait s'expliquer par l'influence du forçage radiatif sur l'évolution de long terme des débits. Nous avons pour cela conduit une régression de Mann-Kendall sur les débits enregistrés aux stations. Pour la chronique 2000-2020, la régression n'est pas significative. Nous avons donc étendu la chronique à 1922-2023 pour les deux stations du territoire disposant d'une telle étendue d'enregistrements : la station de Villerest sur la Loire aval et celle de Bas-en-Basset sur la Loire amont (ces deux stations permettent également de couvrir la quasi-totalité du territoire et sont donc hydrologiquement représentatives). La station de Villerest a enregistré une baisse de 20% du débit mensuel moyen sur 100 ans, et cela avec un niveau de confiance très bon, supérieur à 99%. La tendance est du même ordre pour la station de Bas-en-Basset, où nous estimons la baisse du débit mensuel moyen sur la période à presque 30% avec un niveau de confiance également supérieur à 99%.

S'il y a donc un effet tendanciel certain sur la variabilité, cette régression ne permet pas formellement de distinguer, par exemple, entre l'effet du forçage radiatif et l'anthropisation des paysages. Comme nous l'évoquons plus haut, une évaluation basée sur des données de mesure ne reproduit pas l'effet explicatif d'un modèle spatialement différencié. Nous avons donc proposé une démarche de spatialisation des données hydrométriques ponctuelles, qui consiste à attribuer aux mesures effectuées par une station les sources de pressions situées sur sa zone d'apport (i.e. son bassin versant). Pour cela, nous avons implémenté un modèle hydrographique élémentaire (MHE) dans le logiciel **HydroTrendsFrance**. Son rôle est de reconstituer le bassin versant topographique associé à la position géographique de toute station hydrométrique. L'implémentation numérique du MHE a été validée par comparaison de l'emprise du bassin versant de la station de Villerest obtenue par **HydroTrendsFrance** avec celle obtenue par le service **WatershedWPS** du BRGM et de l'OFB. Le bassin versant obtenu par **WatershedWPS** est ainsi 3.5% plus grand que celui obtenu par **HydroTrendsFrance** (4660 km² contre 4500 km²), et les deux résultats présentent ainsi une bonne proximité.

À partir de cette spatialisation, nous avons calculé une matrice de corrélation de Pearson

entre la variabilité des débits mesurés sur chacun des bassins versants du territoire, et certains éléments du paysage représentatifs capables d’influencer l’écoulement : les barrages, les seuils, les écluses, les stations d’épuration, l’occupation des sols. La matrice de corrélation montre que les barrages et les seuils sur le territoire ne peuvent pas être tenus pour responsables de la variabilité enregistrée aux stations (leur coefficient de corrélation n’est pas significativement différent de zéro), contrairement à l’hypothèse qui préside au choix de modélisation retenu dans (ROCKSTROM et al. 2023). Toutefois, à une échelle plus fine, le coefficient de corrélation pourrait significativement changer : cet effet d’échelle dans l’évaluation d’une pression est connu dans la littérature sous le nom de capacité de charge disponible (*available carrying capacity*) (BJØRN et al. 2019). Pour les stations d’épuration, le coefficient de corrélation atteint +26%, ce qui ne permet pas de conclure fermement quant à leur impact sur la variabilité, bien que celui-ci commence à être visible. Il est possible qu’à une échelle plus fine, leur effet soit plus net avec un coefficient de corrélation plus élevé. Concernant les classes d’occupation du sol, les forêts de conifères semblent jouer un rôle fortement influant à la baisse de la variabilité (-40%) et les systèmes agricoles mixtes à la hausse (+44%). Une perspective de travail serait de mener une analyse diachronique de la corrélation en considérant les variations entre mois de l’année et entre années, ainsi qu’une analyse sur des bassins versants de taille inférieure (ce qui suppose d’interpoler les données Vigicrues, par exemple grâce au modèle de rééchantillonnage hydrologique que nous avons développé).

Concernant la seconde hypothèse de travail, nous avons présenté **relio**, l’application numérique du modèle de rééchantillonnage hydrologique (MRH). Ce modèle permet de rééchantillonner, d’une part, un jeu de données hydrologiques carroyées à faible résolution (typiquement 7000m × 7000m) vers une image raster de haute résolution (typiquement 30m × 30m) et, d’autre part, un jeu de données hydrométriques ponctuelles vers cette même image raster de haute résolution : **relio** permet ainsi la comparaison pixel-à-pixel de deux jeux de données de natures différentes (données de modèle *versus* données de mesure) et de géométrie différente (données carroyées *versus* données ponctuelles). Au cours de ce chapitre, nous avons présenté l’implémentation logicielle du MRH et avons évalué sa performance. Le modèle MRH n’a donc pas fait l’objet d’une application à proprement parler : le cas d’étude du Sud-Loire sur lequel nous avons travaillé nous a permis de discuter des paramètres relatifs à sa calibration (notamment l’initialisation des conditions aux bornes). L’application de **relio** pour l’évaluation environnementale à proprement parler constitue des perspectives de travail.

Pour ce qui est de la validation (ou des performances) de **relio**, rappelons qu'il implémente les trois algorithmes du modèle MRH : l'algorithme de dérasterisation (opération de projection des données carroyées à faible résolution vers l'image raster cible), l'algorithme de rasterisation (opération de projection des données ponctuelles de mesure vers l'image raster cible) et l'algorithme de convolution (opération d'interpolation qui reconstitue les données manquantes le long du réseau hydrographique dans l'image raster cible). Deux démarches de validation de **relio** ont été entreprises : la validation directe, par comparaison des résultats de **relio** à un échantillon témoin de données de mesure, et la validation croisée, par comparaison des résultats de **relio** aux résultats de l'algorithme de distance inverse (**invdist** du fournisseur de traitements **gdal**). Si cette double validation a été entreprise, c'est qu'il n'est pas possible de valider séparément les algorithmes implémentés dans **relio**. De fait, seulement l'algorithme de convolution peut être validé ainsi. L'algorithme de rasterisation est en effet nécessaire pour pouvoir comparer les performances de l'interpolation menée par **relio** à un échantillon témoin de station de mesure, de même que l'algorithme de dérasterisation pour la comparaison des données de modèle. La validation croisée permet alors de comparer les performances de **relio** moins par rapport à des points de contrôle sur site (des stations de mesure) que par rapport aux performances de la méthode la plus usuelle de rééchantillonnage disponible en SIG.

Concernant la validation directe, nous avons distingué les trois bassins versants principaux du territoire : la Loire, le Rhône et l'Allier. Pour le bassin versant de la Loire, la valeur médiane du biais moyen est de -22%, contre -10% pour celui de l'Allier et 2000% pour celui du Rhône. Ces valeurs peuvent être retenues comme les performances **relio** sur son algorithme de convolution. La mauvaise performance sur le bassin du Rhône peut s'expliquer par la difficulté à paramétrer correctement les apports amont de la Saône et du Rhône qui sont interceptés sur la zone d'étude (qui, rappelons-le, a été définie à partir de l'emprise du SCoT Sud-Loire mais qui a ensuite été étendue selon la procédure d'initialisation de la zone d'étude pour minimiser les tronçons hydrographiques entrants). La comparaison des fonctions de densité entre les résidus additifs et multiplicatifs montre que la probabilité pic des résidus multiplicatifs est très nettement supérieure à celle des résidus additifs (0.07 contre 0.00001). Le comportement de l'erreur de **relio** serait ainsi de type proportionnel et une correction par un facteur de proportionnalité distinct pour chaque bassin versant pourrait être envisagée.

Concernant la validation croisée de **relio** par rapport à l'algorithme continu de rééchantillonnage **invdist**, nous avons d'une part conduit une évaluation croisée des deux algorithmes

par rapport à un échantillon de contrôle et, d'autre part, par rapport à leurs performances de rééchantillonnage sur les données du WBM. Pour ce qui est du rééchantillonnage d'un échantillon de données de mesures vis-à-vis d'un échantillon de contrôle, **invdist** tend globalement à sur-estimer les valeurs de débit, car la distribution de son biais moyen de n'est pas centrée sur 0 mais décalée vers la droite ($x = 7.6e5$), ce qui confirme la tendance du modèle à la sur-estimation. On observe également la plus grande dispersion des valeurs estimées par **invdist** comparément à **relio** car la densité pique du biais moyen est bien moins forte pour **invdist**, de l'ordre de 0.00000013 contre 0.07 pour **relio**. Pour ce qui est du rééchantillonnage des données du WBM, l'évaluation des performances de **relio** intègre alors également celles de l'algorithme de dérasterisation. S'il n'est pas possible d'évaluer sa performance isolément, l'algorithme de dérasterisation fonctionne sur le même principe que l'algorithme de convolution, à savoir attribuer une quantité connue de débit au prorata d'une charge hydraulique théorique. Il est donc raisonnable de penser que l'erreur propre à l'algorithme de dérasterisation n'est pas bien différente de l'erreur que nous avons évaluée précédemment, bien que nous ne puissions pas la quantifier. L'erreur qui apparaît pour le rééchantillonnage des données WBM dépasse de loin l'ordre de grandeur de l'erreur liée au rééchantillonnage des données de mesure et **relio** performe alors aussi mal que **invdist**. Il est donc raisonnable de penser que la dégradation significative des performances de **relio** s'explique moins par une défaillance du modèle WBM à cette échelle d'étude.

Conclusion générale

Ce travail de thèse de Doctorat a été orienté par l'hypothèse directrice que, dans l'application du cadre des limites planétaires sur un territoire, il ne se joue pas seulement une question d'échelle spatiale (de l'échelle globale vers l'échelle locale) mais aussi une question de qualité de la donnée environnementale. Cette hypothèse s'inscrit dans le contexte disciplinaire de l'évaluation absolue de la soutenabilité environnementale (AESA). Dans un tel contexte, le cadre des limites planétaires est considéré comme une méthode de quantification des pressions anthropiques sur l'environnement. Nous avons souligné au chapitre 1, que le cadre des limites planétaires n'est pas tenu à ce rôle de quantification des pressions, et qu'il peut s'envisager comme un cadre réflexif ou prospectif dont la particularité est, d'une part, de décentrer la focale des politiques environnementales de la seule prise en compte du carbone et, d'autre part de penser les pressions anthropiques de façon systémique, c'est-à-dire de considérer les interactions entre compartiments environnementaux d'une même politique environnementale.

Une réflexion sur la qualité de la donnée environnementale dans un contexte d'AESA

S'intéresser à la qualité de la donnée environnementale n'est évidemment pas une question nouvelle. Dans le champ de l'AESA, nous avons souligné au chapitre 2 les discussions dont elle a fait l'objet, notamment dans (BJØRN et al. 2019), pour qui la collecte des données et l'évaluation de leur robustesse doivent être considérées comme une phase à part entière de la méthode d'évaluation absolue ; ou encore dans (ZIPPER et al. 2020) qui traitent plus spécifiquement de l'intégration du cycle de l'eau au cadre des limites planétaires et qui identifient la qualité de la donnée, en particulier sa validation aux échelles auxquelles elle est mobilisée, comme un enjeu scientifique majeur pour le champ. Si la question de la donnée est soulevée, peu d'auteurs ont en revanche proposé une méthodologie ou un cadre analytique pour l'évaluation de la qualité de la donnée environnementale mobilisée dans une évaluation absolue. La question de la qualité de la donnée nous semble avoir été davantage documentée sous l'angle de la validation des

résultats de modèles de simulation utilisés pour la quantification des limites planétaires dans le cadre global. Dans cette perspective, l'évaluation de la qualité de la donnée est réduite à un traitement des erreurs de modèles.

Le chapitre 2, en explorant les jeux de données mobilisés dans le champ des AESA relatives à l'eau de surface, a mis en avant les différentes façons de traiter les erreurs de modèles dans cette littérature. Une première distinction peut être faite entre une démarche de validation et une démarche de calibration d'un modèle. La validation consiste à comparer les résultats obtenus par un modèle à des résultats provenant de sources externes et jugés eux-mêmes validés. La calibration pour sa part consiste à optimiser les valeurs des paramètres en entrées de modèles pour minimiser l'erreur des résultats en sorties. La calibration est une démarche que nous n'avons pas retrouvée dans notre littérature. Elle s'apparente aux techniques de *data assimilation* qui sont largement mobilisées en météorologie et océanographie (GHORBANIDEHNO et al. 2020), mais peu en hydrologie (BARBAROSSA et al. 2018). La validation en revanche est une démarche fréquemment mobilisée dans notre littérature, dans laquelle nous avons pu identifier au moins deux types d'approches.

D'une part, et pour prendre un exemple, (ROCKSTROM et al. 2023) ont recours aux simulations du *Water Balance Model*, un modèle hydrologique global, pour la quantification de la frontière du système terre relative aux eaux de surface. Ce modèle, développé par (VOROSMARTY et al. 2000; WISSER et al. 2010) a été validé par (FEKETE et al. 2002). L'approche retenue pour la validation a consisté à comparer les valeurs de ruissellement obtenues par le modèle avec les valeurs de ruissellement enregistrées par un réseau de stations hydrométriques installées en cours d'eau (*Global Runoff Data Center*). Alors que le modèle fournit une valeur de ruissellement pour chaque carreau de résolution $7 \text{ km} \times 7 \text{ km}$ env. sur un maillage qui couvre les surfaces émergées du globe, les stations hydrométriques utilisées pour la validation ne correspondent qu'à une petite fraction de ces carreaux. Nous disons alors qu'il s'agit d'une *validation partielle du modèle sur son domaine de valeurs par des données d'observation*.

D'autre part, une approche alternative consiste à valider les résultats de modèle par la méthode des ensembles. Cette technique a notamment été mobilisée dans les publications qui entendent proposer une reformulation de la limite planétaire relative à l'eau (ROCKSTRÖM et al. 2009b; PASTOR et al. 2014; PORKKA et al. 2024). Il s'agit de produire un même résultat

par un ensemble de modèles différents mais instanciés avec les mêmes données d'entrée, et de considérer que l'étendue des résultats entre les différents modèles constitue une plage d'erreur pour la valeur souhaitée (TELTEU et al. 2021). Nous disons alors qu'il s'agit d'une *validation du modèle par intervalle de confiance*.

Le travail de (PASTOR et al. 2020) propose une montée en généralité sur la façon de traiter les incertitudes dans ce type de modèles de simulation, qui débouche sur un cadre analytique que nous avons repris au chapitre 3 pour l'identification des contributeurs statistiques à la variabilité des débits en cours d'eau. Ces travaux réduisent toutefois l'ensemble des types de données mobilisables dans une AESA aux seules données issues de modèles de simulation numérique. Le travail de (CRENNA et al. 2019) propose précisément un cadre analytique plus englobant, en introduisant une typologie hiérarchique des sources de données. Cette typologie présente l'intérêt de classer les sources de données entre elles, et donc de définir une priorité dans les données mobilisées pour l'évaluation environnementale, mais aussi de définir des critères de robustesse entre sources d'une même classe.

Une réflexion sur la question des échelles spatiales dans un contexte d'AESA

La question des échelles spatiales, entre le global et le local, a également suscité de nombreuses discussions, à la fois dans le champ de l'AESA et dans celui des sciences du système Terre. Dans le champ de l'AESA, cette question est généralement résumée par la distinction entre une approche méthodologique dite *bottom-up* et une approche dite *top-down* pour la quantification de la capacité de charge (BAI et al. 2024). La première consiste à définir une capacité de charge à l'échelle du système étudié et reconnaît ainsi l'existence de seuils environnementaux sous-globaux. Certains auteurs parlent ainsi de *regional safe operating space* (DEARING et al. 2014). Pour la seconde approche en revanche, il s'agit d'effectuer une allocation de la capacité de charge globale qui soit représentative et juste pour le système étudié (HJALSTED et al. 2021). Une littérature très importante s'est précisément concentrée sur cette question du principe d'allocation (RYBERG et al. 2020). Quant au champ des sciences du système Terre, nous avons montré au chapitre 2 comment l'idée d'une définition locale de la limite planétaire relative à l'eau s'est progressivement imposée, notamment sous l'impulsion des travaux de (DEARING et al. 2014; GERTEN et al. 2013; GLEESON et al. 2020b). Dans la révision du cadre proposée en 2015 (STEFFEN et al. 2015b), la possibilité d'une variable de contrôle localement définie pour l'eau de surface (les prélèvements anthropiques ne doivent

pas excéder une certaine fraction de l'eau bleue disponible localement afin de maintenir le fonctionnement des écosystèmes aquatiques) est envisagée aux côtés de la variable de contrôle globale (le volume annuel d'eau douce appropriable par les activités humaines). La variable de contrôle "locale" n'est toutefois envisagée qu'à l'échelle des grands bassins versants. Si les auteurs reconnaissent qu'il peut exister localement des seuils critiques pour l'eau de surface, tous ces seuils locaux ne représentent pas la même menace pour la stabilité du système Terre, et les grands bassins versants constituent ainsi l'échelle à partir de laquelle le dépassement de la limite localement définie sur l'eau de surface devient critique pour le système lui-même (GLEESON et al. 2020a) : une vision qui renvoie à celle des points de bascule (LENTON et al. 2008). C'est véritablement à partir de la mise à jour proposée par (ROCKSTROM et al. 2023) en 2023, que la variable de contrôle retenue pour l'eau de surface se définit fondamentalement à l'échelle locale, et c'est la somme des transgressions des limites locales qui conduit à la transgression de la limite globale. En effet, la déstabilisation des écosystèmes aquatiques, provoquée par une altération du débit des cours d'eau supérieure à $\pm 20\%$ du débit moyen pré-industriel, est jugée d'égale importance quel que soit l'endroit sur la planète.

On pourrait alors envisager de rapprocher la façon de définir une *approche locale* dans l'un et l'autre des deux champs scientifiques : la variable de contrôle proposée par (ROCKSTROM et al. 2023) semble en effet correspondre à une approche *bottom-up* en AESA puisqu'elle consiste à reconnaître l'existence d'une capacité de charge locale. En réalité, ce que désigne le terme *local* ne se rapporte pas au même type d'*échelle* : l'échelle de la variable de contrôle pour les sciences du système Terre et l'échelle du système étudié pour l'AESA. La conclusion du chapitre 2 nous a ainsi amené à souligner le rôle de la donnée environnementale dans la correspondance entre ces deux échelles spatiales. L'enjeu est effet de parvenir à une quantification environnementale à chacune de ces deux échelles, celle du système étudié et celle de la capacité de charge. Quand bien même ces deux échelles se correspondraient-elles conceptuellement, leur mise en chiffre fait nécessairement intervenir des données environnementales qui, elles aussi, sont associées à une échelle spatiale spécifique.

Apprécier l'échelle spatiale par l'échelle de la donnée : un cadre analytique pour la caractérisation d'une démarche *bottom-up* en AESA

Bien que la question de la qualité de la donnée environnementale ait déjà été partiellement abordée dans notre littérature, de même que celle des échelles spatiales, nous avons souhaité

confronter ces deux questions en posant d'une part la question de l'échelle spatiale en termes de qualité de la donnée environnementale et, inversement, la question de la qualité de la donnée environnementale en termes d'échelle spatiale.

Nous avons ainsi suggéré de distinguer entre plusieurs types d'échelles spatiales : l'*échelle de l'étude* d'une part et ce que nous avons appelé l'*échelle de la donnée* d'autre part. Au regard de la littérature que nous avons étudiée, l'échelle de l'étude nous paraît recouvrir deux dimensions :

- l'échelle spatiale du système étudié : c'est l'échelle à laquelle se définit géographiquement le système que l'on étudie, qui correspond à l'étape de définition du périmètre géographique dans une évaluation environnementale, où le système est spatialement borné. Par exemple, pour une évaluation de l'intégrité des cours sur le territoire du SCoT Sud-Loire, l'échelle du système correspond aux frontières administratives du Sud-Loire.
- la granularité spatiale du système étudié : c'est l'échelle à laquelle on se propose d'effectuer la quantification environnementale du système étudié, qui peut être plus fine, identique ou plus large que l'échelle du système étudié. Par exemple, on peut décider de cartographier le SCoT Sud-Loire par un maillage régulier d'une résolution de $1\text{km} \times 1\text{km}$ pour évaluer l'intégrité des cours d'eau sur chaque carreau, ou bien peut-on décider d'instrumenter d'un dispositif de mesure du débit le cours d'eau le plus important du territoire.

De même, nous définissons l'échelle de la donnée par deux dimensions :

- la granularité de la donnée : il s'agit de la résolution spatiale de chaque jeu de données qui sert à quantifier la variable de contrôle et/ou la capacité de charge. Par exemple, on décide de récupérer les résultats du *Water Balance Model*, qui sont carroyés sur un maillage de résolution $7\text{ km} \times 7\text{km}$, pour quantifier la capacité de charge (débit non-anthropisé des cours d'eau), et un jeu de données hydrométriques donnant le débit actuel du territoire sur un maillage de résolution $1\text{km} \times 1\text{km}$ pour quantifier la variable de contrôle (débit anthropisé).
- la source hiérarchique de la donnée : il s'agit de la typologie hiérarchique des données que nous avons reprise de (CRENNA et al. 2019) et modifiée pour tenir compte de la particularité des jeux de données mobilisés dans la littérature sur la limite planétaire relative à l'eau de surface. Cette typologie modifiée classe hiérarchiquement les sources de données selon qu'elles sont plus ou moins spatialement explicites (ce qui dépend de leur validation par rapport à des données d'observation) et plus ou moins robustes (ce qui dépend de la validation de la méthode par laquelle elles ont été produites).

En s'appuyant sur ces distinctions, on pourrait alors répondre à la question de savoir si le SCoT Sud-Loire est une échelle spatiale adaptée au cadre des limites planétaires. Nous répondons plus généralement que la question n'est pas de savoir si l'échelle du système correspond à l'échelle à laquelle on peut définir une limite planétaire, mais si l'enchâssement des échelles que nous venons de définir est cohérent au cours de la méthode d'AESA. Ainsi, une AESA conduite sur un territoire n'est pas seulement *bottom-up* parce qu'on aurait défini une capacité de charge à l'échelle de l'étude (BAI et al. 2024), mais parce qu'il y a une cohérence méthodologique entre l'échelle de l'étude (échelle du système étudié et granularité spatiale du système étudié) et l'échelle de la donnée (granularité spatiale des jeux de données et source hiérarchique des jeux de données). Nous avons ainsi suggéré de parler d'un *gradient d'approche bottom-up* pour rendre compte de ces multiples agencements entre échelles.

Retour sur les questions de recherche

En nous plaçant dans le cadre analytique que nous avons développé, nous avons formulé deux questions de recherche :

1. Comment effectuer l'évaluation environnementale de l'intégrité des cours d'eau sur un territoire au regard de l'indicateur proposé par (ROCKSTROM et al. 2023) à partir de mesures de débits sur site systématisées ? Cette proposition étudie la faisabilité d'un indicateur AESA qui assure la cohérence méthodologique entre l'échelle de l'étude et l'échelle de la donnée à partir des données classées comme les plus spatialement explicites et robustes.
2. Comment coupler des données issues de mesures de débits sur site systématisées avec des données issues des sorties d'un modèle hydrologique global ? Autrement dit, comment coupler des données spatialement explicites avec des données non spatialement explicites mais spatialement différenciées ? Cette proposition étudie la faisabilité d'une assimilation entre sources de données différentes et étudie ainsi la possibilité de combiner des degrés différents d'approche *bottom-up*.

Le chapitre 3 a détaillé la méthode associée à chacune de ces deux questions de recherche, et le chapitre 4 a exposé l'application numérique qui en a été faite sur le territoire du SCoT Sud-Loire en mobilisant les données de mesure de débits du réseau national Vigicrues et les données de modèle du *Water Balance Model* (WBM).

Pour la question de recherche (1), nous avons discuté des enjeux méthodologiques que pose la définition d’une variable de contrôle basée sur la variabilité du débit à partir de données de mesure sur site. Le type de données mobilisées vient en effet questionner ce qui sert de point de référence à partir duquel évaluer la variation du débit. Nous avons à cet égard distingué entre une référence dite *absolue* (le régime d’écoulement non perturbé) et une référence dite *normale* (au sens d’une norme statistique). Les données de mesure de débits sur site quantifiant par construction le régime perturbé, il n’est pas possible de construire une référence absolue mais seulement *normale*.

Dans leur publication, (ROCKSTROM et al. 2023) proposent de définir le débit de référence en tout point d’un cours d’eau (en fait, sur tout carreau de $7 \text{ km} \times 7 \text{ km}$) d’après les résultats d’un modèle hydrologique global (le WBM) qui simule un écoulement de l’eau tel qu’il serait vraisemblablement sans les perturbations anthropiques liées à l’irrigation et aux barrages. En calculant la moyenne, pour chaque carreau, des débits mensuels moyens sur la durée de la simulation (20 ans, de 2000 à 2020), les auteurs établissent un débit de référence qui est dit absolu car censé représenter un débit préservé des activités humaines. En revanche, lorsque l’on dispose d’une série temporelle de valeurs de débits mesurées par une station hydrométrique (typiquement, en France, les débits mensuels moyens d’une station du réseau Vigicrues), la moyenne des valeurs de la série ne donne qu’une référence que l’on peut qualifier de *normale* sur la période. On peut dans les deux cas (référence absolue et référence normale) calculer l’écart du débit moyen d’un mois donné (par exemple le débit moyen du mois de janvier 2001) avec sa valeur de référence sur la période (le débit moyen de tous les mois de janvier de la période 2000-2020). L’écart par rapport au débit dit absolu permet de définir un indicateur d’*excès de variabilité*, tandis que l’écart par rapport au débit normal donne un indicateur d’*anomalie de variabilité*. On peut en revanche conserver dans les deux cas une même valeur de seuil de variabilité : si l’écart dépasse le couloir des $\pm 20\%$, alors il y a dépassement de la limite.

Nous avons implémenté cet indicateur d’anomalie de variabilité dans le logiciel `HydroTrendsFrance`, disponible sous licence libre²², qui fonctionne à partir des données du réseau Vigicrues. Ainsi, pour le SCoT Sud-Loire, les stations hydrométriques enregistrent une fréquence annuelle de dépassement comprise entre 56 et 77%, et en moyenne de 60% (moyenne sur toutes les stations) sur la période 2000-2020. Cela signifie que sur une année, les cours d’eau du territoire sont en moyenne anormalement perturbés deux mois sur trois (de un

22. Voir le dépôt GitHub : <https://github.com/Quentin-Dassibat/HydroTrendsFrance>

mois sur deux jusqu'à trois mois sur quatre).

Un tel indicateur d'anomalie de variabilité ne porte en revanche pas d'information sur l'origine de la pression excessive qu'il quantifie. Nous avons notamment souhaité savoir si une tendance de long terme pouvait interférer à la hausse ou à la baisse sur la variabilité excessive observée (une interférence qui pourrait notamment s'expliquer par l'effet du forçage climatique). Nous avons donc proposé l'utilisation d'une méthode de régression de Mann-Kendall pour détecter une éventuelle tendance de long terme statistiquement significative sur les débits observés. Cette méthode a ainsi été implémentée dans **HydroTrendsFrance**. Les deux stations les plus hydrologiquement représentatives du Sud-Loire, celle de Villerest et celle de Bas-en-Basset, ont ainsi enregistré une baisse respective de 20 et 30% du débit mensuel moyen sur 100 ans (1922-2023), et cela avec un niveau de confiance supérieur à 99%.

Une autre question que nous avons adressée aux données de mesure relève de leur capacité spatialement explicative. Les données de débits issues de stations hydrométriques ne portent en effet pas spontanément l'information sur l'emprise spatiale de la pression anthropique qu'elles mesurent. Aussi avons-nous suggéré de spatialiser les données ponctuelles de débits par le développement de ce que nous avons appelé un modèle hydrographique élémentaire, que nous avons implémenté dans **HydroTrendsFrance**. Avec cette spatialisation des données de débits, nous avons proposé d'analyser l'influence de l'aménagement du territoire sur la variabilité observée aux stations. Le résultat de la corrélation sur le SCoT du Sud-Loire montre, par exemple, que les forêts de conifères²³ semblent jouer un rôle fortement influant à la baisse de la variabilité (donc un rôle de stabilisation de l'écoulement, par exemple en écrétant les événements extrêmes), tandis que les champs agricoles mixtes (champs ouverts et en bocage) seraient responsables d'une hausse significative de la variabilité (donc d'un effet perturbateur sur la stabilité des débits).

Pour ce qui est de la seconde question de recherche (2), nous avons développé un modèle d'assimilation des données hydrologiques qui permet de coupler les données issues d'un modèle hydrologique global avec les données issues d'un réseau de stations de mesure. Ce modèle permet de rééchantillonner chacun de ces deux types de données hydrologiques : d'une part, en rééchantillonnant un jeu de données hydrologiques carroyées à faible résolution (typiquement 7000m × 7000m) vers une image raster de haute résolution (typiquement 30m × 30m) et,

23. Et vraisemblablement d'autres types de boisements, mais ce type de forêt est surreprésenté sur le SCoT Sud-Loire.

d'autre part, un jeu de données hydrométriques ponctuelles vers cette même image raster de haute résolution. Ce modèle de rééchantillonnage permet ainsi la comparaison pixel-à-pixel de deux jeux de données hydrologiques de nature différente (données de modèle *versus* données de mesure) et de géométrie différente (données carroyées *versus* données ponctuelles). Le développement d'un modèle de rééchantillonnage spécifique aux données hydrologiques s'est en effet avéré nécessaire car les techniques de rééchantillonnage classiquement disponibles chez les principaux fournisseurs de traitements SIG reposent sur une hypothèse de continuité spatiale des phénomènes, ce qui n'est pas le cas des données hydrologiques qui sont spatialement contraintes par le tracé du réseau hydrographique.

Nous avons développé numériquement ce modèle sous le nom de **relio**, qui est disponible en licence libre²⁴. Nous avons utilisé **relio** pour le rééchantillonnage des données issues du *Water Balance Model* et des données issues des stations de mesure du réseau Vigicrues. Le modèle **relio** a été validé par rapport à des données de mesure de débits et selon une approche croisée par comparaison avec les valeurs de débits obtenues selon un algorithme classique de rééchantillonnage continu, à savoir l'algorithme de distance inverse de la librairie **gdal**. L'erreur des sorties du modèle **relio** est différente selon le bassin versant considéré sur le territoire du SCoT Sud-Loire²⁵. Ainsi, sur le bassin versant de la Loire, nous avons retenu une erreur de -22%, qui est l'écart relatif entre les débits rééchantillonnés et les débits mesurés par les stations. Pour le bassin versant du Rhône en revanche, l'erreur atteint un facteur 20, ce qui témoigne de la difficulté à initialiser les conditions aux bornes du modèle lorsque celui-ci fonctionne sur des bassins hydrographiques non clos (système ouvert). À titre de comparaison, l'erreur associée à l'algorithme de distance inverse atteint un facteur $7.6 \cdot 10^5$ en moyenne sur tous les bassins versants de la zone d'étude. D'autre part, le rééchantillonnage des données du WBM par **relio** a permis leur comparaison avec les données d'observation du réseau Vigicrues. Sur le bassin versant de la Loire, où les performances de **relio** se sont révélées satisfaisantes, le biais moyen observé entre les données WBM rééchantillonnées et les données de mesure atteint un facteur 77. cela signifie, malgré l'erreur propre à **relio** mais qui négligeable devant l'erreur des données WBM rééchantillonnées, que le WBM n'est effectivement pas valide sur l'ensemble des carreaux de son maillage, comme l'avait documenté (FEKETE et al. 2002)

24. Voir le dépôt GitHub : <https://github.com/Quentin-Dassibat/relio>

25. Remarquons que pour l'amélioration des performances du modèle (meilleure paramétrisation des conditions aux bornes), nous avons étendu la zone d'étude au-delà des frontières administratives du SCoT Sud-Loire. Nous avons intégré une portion du bassin versant du Rhône à l'est du territoire par les monts du lyonnais, une portion du bassin versant de l'Allier à l'ouest par les monts du Forez, et une portion du bassin de la Loire amont sur le département de la Haute-Loire au sud du territoire

pour des bassins versants de taille inférieure à 10^5 km^2 (comme le territoire du SCoT Sud-Loire).

Perspectives de recherche

Le premier apport de ce travail porte sur le plan analytique, c'est-à-dire dans le développement d'un cadre analytique pour la caractérisation d'une AESA bottom-up. À cet égard, nous avons suggéré l'existence d'un *gradient d'approche bottom-up* en opposition à la dichotomie trop nette entre approche *bottom-up* et approche *top-down*. Le cas d'étude à partir duquel nous avons construit ce cadre analytique a réduit l'analyse sur plusieurs aspects : (i) il a seulement porté que sur le compartiment environnemental de l'eau de surface ; (ii) il n'a porté que sur un seul de territoire ; (iii) il n'a porté que sur la version (ROCKSTROM et al. 2023) du cadre des limites planétaires.

(i) Nous avons retenu l'eau de surface du fait des préoccupations fortes liées à l'enjeu de la ressource en eau parmi les acteurs du SCoT Sud-Loire, mais aussi en raison de la disponibilité d'un système d'information sur l'eau dont les données sont à la fois robustes et spatialement explicites. Répéter l'exercice sur d'autres limites planétaires, c'est-à-dire sur d'autres compartiments environnementaux, implique ainsi de disposer de telles données. Sur le territoire du Sud-Loire, une récente étude menée par le SAGE Loire en Rhône-Alpes a par exemple permis de conduire une étude quantitative sur l'état des aquifères de la rive droite de la plaine du Forez (PAULET et al. 2023), ce qui ouvre la voie à l'étude du compartiment des eaux souterraines au regard du cadre des limites planétaires, dans la double perspective où nous l'avons conduite pour les eaux de surface : la construction d'une méthode AESA basée sur des données de mesure sur site robustes, et le couplage de telles données avec des sorties de modèles hydrogéologique. De même, l'étude HMUC en cours (Hydrologie Milieu Usage Climat) qui est conduite sur le territoire par le récent Projet de territoire pour la gestion de l'eau (PTGE), a produit des données croisées de débits et de comptage d'espèces sur différents cours d'eau du territoire²⁶, ouvrant alors la voie à une répétition de l'étude sur les eaux de surface que nous avons menée mais avec un autre jeu de données de mesure, ou bien à l'étude de la limite planétaire relative à la biodiversité par le prisme de la biodiversité aquatique.

(ii) Le territoire du SCoT Sud-Loire sur lequel nous avons bâti notre cas d'étude ne représente qu'un type particulier de système sous-global, c'est-à-dire un système territoire, que

26. Voir <https://sage-loire-rhone-alpes.fr/?p=4086>

nous avons qualifié au chapitre 1 de multifonctionnel et de multi-zone. Il serait souhaitable de confronter le cadre analytique à d'autres types de systèmes sous-globaux, de type unifonctionnel ou de type site, mais aussi à d'autres territoires, par exemple des territoires côtiers ou de montagne, ou même d'autres territoires de plaine dont les paysages, l'occupation du sol et les activités anthropiques diffèrent d'avec le SCoT Sud-Loire. Cette perspective de recherche s'inscrit dans les perspectives propres au champ de l'évaluation environnementale, qui pourrait voir un intérêt non seulement à confronter le cadre analytique que nous proposons à d'autres types de systèmes, mais aussi à considérer le système dans une approche non-territoriale (au sens de l'évaluation environnementale), c'est-à-dire dans une approche consommation : nous aurions en effet pu nous intéresser aux pressions sur l'eau de surface causées par le SCoT Sud-Loire sur d'autres territoires en raison de la consommation et des modes de vie de sa population résidente.

(iii) Nous n'avons abordé la question de l'eau de surface qu'au prisme de la variable de contrôle proposée par (ROCKSTROM et al. 2023). Ce choix tenait à la fois au fait qu'il s'agissait de la version la plus actuelle du cadre des limites planétaires au moment où nous avons entrepris nos recherches, ainsi qu'à l'échelle fondamentalement "locale" à laquelle se quantifie la variable de contrôle, nous offrant alors la possibilité d'une réflexion sur la façon dont se traduit une échelle locale dans le cadre des limites planétaires à une échelle locale dans une évaluation AESA (ce dont nous avons discuté plus haut). La formulation de la variable de contrôle pour l'eau de surface dans (RICHARDSON et al. 2023) et, plus encore dans la dernière actualisation de (PORKKA et al. 2024), qui ne se basent plus directement sur une valeur normale de débit mais sur une valeur normale d'anomalie de variabilité, questionnerait à nouveaux frais la construction d'un indicateur exclusivement basé sur des données de mesures de débits.

Le second apport de ce travail porte sur le plan quantitatif. Les deux questions de recherche que nous avons posées portent en effet les aspects quantitatifs de l'AESA, l'une en s'intéressant à la construction d'une variable de contrôle adaptée à des données de mesure, l'autre en s'intéressant au couplage entre différentes sources hiérarchiques de données.

Pour ce qui est de la première démarche de quantification, nous avons souligné l'absence de pouvoir explicatif, ou causal, d'une mesure de variabilité du débit. Entendue comme erreur par rapport à une valeur normale, cette variabilité peut être décomposée selon la typologie des incertitudes (PASTOR et al. 2020). Nous n'avons étudié dans ce travail que deux types

d'incertitude, ou erreur : l'incertitude épistémique liée à une tendance de long terme dans les séries temporelles, et l'incertitude paramétrique liée à la longueur de la plage de moyennisation. Une perspective de travail consiste alors à identifier de façon systématique les différents contributeurs à la variabilité des débits selon cette typologie des incertitudes. Nous avons suggéré au chapitre 3 plusieurs approches méthodologiques pour mener une telle identification, sans pour autant les mettre en œuvre numériquement. Par ailleurs, la capacité explicative des données de mesure peut être améliorée par leur spatialisation. Nous avons à cet égard proposé de conduire une analyse de corrélation entre l'anthropisation des paysages et la variabilité des débits. Cette analyse nécessite d'être comparée à d'autres études ayant déjà posé la question (BELLETTI et al. 2020), menée de façon diachronique et à une échelle de bassin versant plus fine (par exemple en interpolant les valeurs de débits connues le long du réseau hydrographique par le modèle `relio`).

Pour ce qui est de la seconde démarche de quantification, liée au couplage entre des données de mesure et des données de modèle hydrologique global, une perspective de recherche consiste à minimiser les erreurs du modèle `relio` en minimisant les erreurs paramétriques liées à la bonne instanciation des conditions aux bornes de l'étude. La démarche la plus prometteuse à cet égard nous semble résider dans une extension de la zone d'étude de telle sorte à ce que tous les bassins soient hydrologiquement clos (pour supprimer le paramètre lié à l'estimation de la charge aval). L'extension de la zone d'étude requiert alors une puissance de calcul suffisante pour laquelle nous suggérons de repenser les scripts en vue de les exécuter sur un serveur distant. Pour la calibration du reste des paramètres de `relio`, notamment les erreurs de projection des stations de mesure dans le maillage cible ou encore la définition d'un seuil d'accumulation minimum pour la représentation d'un régime d'écoulement permanent, une méthode de data assimilation par machine learning pourrait être envisagée, en améliorant la prédiction du modèle par la prise en compte de données d'entrée élargies : surfaces de bassin amont, altitude, pente, précipitations, température de surface, évapotranspiration potentielle (BARBAROSSA et al. 2018).

Bibliographie

- AHN, Sangtae et Jeffrey A. FESSLER (2003). *Standard Errors of Mean, Variance, and Standard Deviation Estimators*. Rapp. tech. Report dated July 24, 2003. EECS Department, The University of Michigan.
- ALCAMO, Joseph, Petra DÖLL, Thomas HENRICHS, Frank KASPAR, Bernhard LEHNER, Thomas RÖSCH et Stefan SIEBERT (2003). “Development and testing of the WaterGAP 2 global model of water use and availability”. In : *Hydrological Sciences Journal* 48.3. Publisher : Taylor & Francis _eprint : <https://doi.org/10.1623/hysj.48.3.317.45290>, p. 317-337. ISSN : 0262-6667. DOI : 10.1623/hysj.48.3.317.45290.
- ALDAYA, Maite M, Ashok K CHAPAGAIN, Arjen Y HOEKSTRA et Mesfin M MEKONNEN (2012). *The water footprint assessment manual : Setting the global standard*. Routledge.
- ALLAN, John A (1998). “Virtual water : a strategic resource”. In : *Ground water* 36.4, p. 545-547.
- ALLAN, John Anthony et al. (1993). “Fortunately there are substitutes for water otherwise our hydro-political futures would be impossible”. In : *Priorities for water resources allocation and management* 13.4, p. 26.
- ALTMAN, Douglas G et J Martin BLAND (2005). “Standard deviations and standard errors”. In : *BMJ : British Medical Journal* 331.7521, p. 903. ISSN : 0959-8138.
- ANDRÉ, Pierre, Claude-E. DELISLE, Jean-Pierre REVÉRET et A. SENE (1999). *L'évaluation des impacts sur l'environnement : Processus, acteurs et pratique*. Presse Internationales Polytechnique, p. 416. DOI : 10.4000/vertigo.5098.
- ARMSTRONG MCKAY, David I., Arie STAAL, Jesse F. ABRAMS, Ricarda WINKELMANN, Boris SAKSCHEWSKI, Sina LORIANI, Ingo FETZER, Sarah E. CORNELL, Johan ROCKSTRÖM et Timothy M. LENTON (2022). “Exceeding 1.5°C global warming

- could trigger multiple climate tipping points”. In : *Science* 377.6611. Publisher : American Association for the Advancement of Science, eabn7950. DOI : 10.1126/science.abn7950.
- ARTHINGTON, Angela H, Anik BHADURI, Stuart E BUNN, Sue E JACKSON, Rebecca E THARME, Dave TICKNER, Bill YOUNG, Mike ACREMAN, Natalie BAKER, Samantha CAPON et al. (2018). “The Brisbane declaration and global action agenda on environmental flows (2018)”. In : *Frontiers in Environmental Science* 6, p. 45.
- BAI, Junhong, Yanan GUAN, Peipei LIU, Ling ZHANG, Baoshan CUI, Xiaowen LI et Xinhui LIU (2020). “Assessing the safe operating space of aquatic macrophyte biomass to control the terrestrialization of a grass-type shallow lake in China”. In : *Journal of Environmental Management* 266, p. 110479. ISSN : 0301-4797. DOI : 10.1016/j.jenvman.2020.110479.
- BAI, Xuemei, Syezlin HASAN, Lauren Seaby ANDERSEN, Anders BJØRN, Şiir KILKIŞ, Daniel OSPINA, Jianguo LIU, Sarah E CORNELL, Oscar SABAG MUÑOZ, Ariane de BREMOND et al. (2024). “Translating Earth system boundaries for cities and businesses”. In : *Nature Sustainability* 7.2, p. 108-119.
- BARBAROSSA, Valerio, Mark A. J. HUIJBREGTS, Arthur H. W. BEUSEN, Hylke E. BECK, Henry KING et Aafke M. SCHIPPER (2018). “FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015”. In : *Scientific Data* 5.1. Publisher : Nature Publishing Group, p. 180052. ISSN : 2052-4463. DOI : 10.1038/sdata.2018.52.
- BARNOSELL, Irene et Carlos POZO (2024). “The impacts of the European chemical industry on the planetary boundaries”. In : *Sustainable Production and Consumption* 44, p. 188-207. ISSN : 2352-5509. DOI : 10.1016/j.spc.2023.12.006.
- BELLETTI, Barbara, Carlos GARCIA DE LEANIZ, Joshua JONES, Simone BIZZI, Luca BÖRGER, Gilles SEGURA, Andrea CASTELLETTI, Wouter VAN DE BUND, Kim AARESTRUP, James BARRY et al. (2020). “More than one million barriers fragment Europe’s rivers”. In : *Nature* 588.7838, p. 436-441.
- BJØRN, Anders, Chanjief CHANDRAKUMAR, Anne-Marie BOULAY, Gabor DOKA, Kai FANG, Natacha GONDRAN, Michael Zwicky HAUSCHILD, Annemarie KERKHOF, Henry KING, Manuele MARGNI, Sarah MCLAREN, Carina MUELLER, Miłołaj OWSIANIAK, Greg PETERS, Sandra ROOS, Serenella SALA, Gustav SANDIN, Sarah

- SIM, Marcial VARGAS-GONZALEZ et Morten RYBERG (2020a). “Review of life-cycle based methods for absolute environmental sustainability assessment and their applications”. In : *Environmental Research Letters* 15.8. Publisher : IOP Publishing, p. 083001. ISSN : 1748-9326. DOI : 10.1088/1748-9326/ab89d7.
- BJØRN, Anders, Miriam DIAMOND, Mikołaj OWSIANIAK, Benoît VERZAT et Michael Zwicky HAUSCHILD (2015a). “Strengthening the Link between Life Cycle Assessment and Indicators for Absolute Sustainability To Support Development within Planetary Boundaries”. In : *Environmental Science & Technology* 49.11. Publisher : American Chemical Society, p. 6370-6371. ISSN : 0013-936X. DOI : 10.1021/acs.est.5b02106.
- BJØRN, Anders et Michael Z. HAUSCHILD (2013). “Absolute versus Relative Environmental Sustainability : What can the Cradle-to-Cradle and Eco-efficiency Concepts Learn from Each Other?” In : *Journal of Industrial Ecology* 17.2, p. 321-332. ISSN : 10881980. DOI : 10.1111/j.1530-9290.2012.00520.x.
- BJØRN, Anders et Michael Zwicky HAUSCHILD (2015b). “Introducing carrying capacity-based normalisation in LCA : framework and development of references at midpoint level”. In : *The International Journal of Life Cycle Assessment* 20.7, p. 1005-1018. ISSN : 0948-3349, 1614-7502. DOI : 10.1007/s11367-015-0899-2.
- BJØRN, Anders, Katherine RICHARDSON et Michael Zwicky HAUSCHILD (2019). “A Framework for Development and Communication of Absolute Environmental Sustainability Assessment Methods”. In : *Journal of Industrial Ecology* 23.4. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jiec.12820>, p. 838-854. ISSN : 1530-9290. DOI : 10.1111/jiec.12820.
- BJØRN, Anders, Sarah SIM, Anne-Marie BOULAY, Henry KING, Julie CLAVREUL, Wan Yee LAM, Valerio BARBAROSSA, Cécile BULLE et Manuele MARGNI (2020b). “A planetary boundary-based method for freshwater use in life cycle assessment : Development and application to a tomato production case study”. In : *Ecological Indicators* 110, p. 105865. ISSN : 1470-160X. DOI : 10.1016/j.ecolind.2019.105865.
- BOGARDI, Janos J, Balázs M FEKETE et Charles J VÖRÖSMARTY (2013). “Planetary boundaries revisited : a view through the ‘water lens’”. In : *Current Opinion in Environmental Sustainability* 5.6, p. 581-589.
- BORUCKE, Michael, David MOORE, Gemma CRANSTON, Kyle GRACEY, Katsunori IHA, Joy LARSON, Elias LAZARUS, Juan Carlos MORALES, Mathis WACKERNAGEL et

- Alessandro GALLI (2013). “Accounting for demand and supply of the biosphere’s regenerative capacity : The National Footprint Accounts’ underlying methodology and framework”. In : *Ecological indicators* 24, p. 518-533.
- BOULAY, Anne-Marie, Jane BARE, Lorenzo BENINI, Markus BERGER, Michael J. LATHUILLIÈRE, Alessandro MANZARDO, Manuele MARGNI, Masaharu MOTOSHITA, Montserrat NÚÑEZ, Amandine Valerie PASTOR, Bradley RIDOUTT, Taikan OKI, Sebastien WORBE et Stephan PFISTER (2018). “The WULCA consensus characterization model for water scarcity footprints : assessing impacts of water consumption based on available water remaining (AWARE)”. In : *The International Journal of Life Cycle Assessment* 23.2, p. 368-378. ISSN : 0948-3349, 1614-7502. DOI : 10.1007/s11367-017-1333-8.
- BOUTAUD, Aurélien et Natacha GONDRAN (2020). *Les limites planétaires*. La Découverte.
- BUNSEN, Jonas, Markus BERGER et Matthias FINKBEINER (2021). “Planetary boundaries for water – A review”. In : *Ecological Indicators* 121, p. 107022. ISSN : 1470-160X. DOI : 10.1016/j.ecolind.2020.107022.
- CAESAR, L, B SAKSCHEWSKI, L ANDERSEN et T BERINGER (2024). *Planetary Health Check Report 2024*. Potsdam, Germany : Potsdam Institute for Climate Impact Research, p. 96.
- CENTRE, DREAL (2012). *Schéma directeur de prévision des crues - bassin Loire-Bretagne et bassins Charente et Seudre*.
- CEREMA (2015a). *Courbe de tarage : extrapolation par modélisation 1D et cartographie par couplage SIG*. In : *Wikhydro*.
- (2015b). *Incertitudes de la courbe de tarage*. In : *Wikhydro*.
- (2015c). *Incertitudes de la mesure du couple (hauteur – débit) lors du jaugeage*. In : *Wikhydro*.
- CHANDRAKUMAR, Chanjief et Sarah J. MCLAREN (2018). “Towards a comprehensive absolute sustainability assessment method for effective Earth system governance : Defining key environmental indicators using an enhanced-DPSIR framework”. In : *Ecological Indicators* 90, p. 577-583. ISSN : 1470-160X. DOI : 10.1016/j.ecolind.2018.03.063.
- CHAPAGAIN, A et A HOEKSTRA (2004). *Water footprints of nations*. UNESCO-IHE Institute for Water Education.

- CHILÈS, Jean-Paul et Nicolas DESASSIS (2018). “Fifty years of kriging”. In : *Handbook of mathematical geosciences : Fifty years of IAMG*, p. 589-612.
- CHIPPERFIELD, Martyn P., Ryan HOSSAINI, Stephen A. MONTZKA, Stefan REIMANN, David SHERRY et Susann TEGTMEIER (2020). “Renewed and emerging concerns over the production and emission of ozone-depleting substances”. In : *Nature Reviews Earth & Environment* 1.5, p. 251-263. ISSN : 2662-138X. DOI : 10.1038/s43017-020-0048-8.
- CORSATEA, Teodora Diana, Soeren LINDNER, Iñaki ARTO, Maria Victoria ROMÁN, José Manuel RUEDA-CANTUCHE, A VELÁZQUEZ AFONSO, Antonio F AMORES et Frederik NEUWAHL (2000). “World input-output database environmental accounts”. In : *Update* 2016, p. 54.
- CRENNA, Eleonora, Michela SECCHI, Lorenzo BENINI et Serenella SALA (2019). “Global environmental impacts : data sources and methodological choices for calculating normalization factors for LCA”. In : *The International Journal of Life Cycle Assessment* 24.10, p. 1851-1877. ISSN : 1614-7502. DOI : 10.1007/s11367-019-01604-y.
- CRUTZEN, Paul (2002). “Geology of mankind | Nature”. In : *Nature*.
- CRUTZEN, Paul J. et Eugene F. STOERMER (2000). “The ‘Anthropocene’ (2000)”. In : *Paul J. Crutzen and the Anthropocene : A New Epoch in Earth’s History*. Sous la dir. de Susanne BENNER, Gregor LAX, Paul J. CRUTZEN, Ulrich PÖSCHL, Jos LELIEVELD et Hans Günter BRAUCH. *The Anthropocene : Politik—Economics—Society—Science*. Cham : Springer International Publishing, p. 19-21. ISBN : 978-3-030-82202-6. DOI : 10.1007/978-3-030-82202-6_2.
- DAHAN, Amy (2007). “Introduction : Modèles et fabrications du futur : du débat sur la croissance au débat climatique et retour”. In : *Les modèles du futur*. Section : Géographie. La Découverte, p. 7-18. ISBN : 978-2-7071-5013-4. DOI : 10.3917/dec.dahan.2007.01.0007.
- DASSIBAT, Quentin (2022). “Des limites planétaires aux limites du Sud-Loire”. In : *Anthropocène 2050*.
- DASSIBAT, Quentin et Natacha GONDRAN (2024). “Limites planétaires et évaluation environnementale des organisations”. In : *Techniques de l’Ingénieur*. Article de référence. DOI : <https://doi.org/10.51257/a-v1-ag504>.

- DE FRAITURE, C, de D MOLDEN, U AMARASINGHE et I MAKIN (2001). “PODIUM : Projecting water supply and demand for food production in 2025”. In : *Physics and Chemistry of the Earth, Part B : Hydrology, Oceans and Atmosphere* 26.11-12, p. 869-876.
- DEARING, John A., Rong WANG, Ke ZHANG, James G. DYKE, Helmut HABERL, Md. Sarwar HOSSAIN, Peter G. LANGDON, Timothy M. LENTON, Kate RAWORTH, Sally BROWN, Jacob CARSTENSEN, Megan J. COLE, Sarah E. CORNELL, Terence P. DAWSON, C. Patrick DONCASTER, Felix EIGENBROD, Martina FLÖRKE, Elizabeth JEFFERS, Anson W. MACKAY, Björn NYKVIST et Guy M. POPPY (2014). “Safe and just operating spaces for regional social-ecological systems”. In : *Global Environmental Change* 28, p. 227-238. ISSN : 0959-3780. DOI : 10.1016/j.gloenvcha.2014.06.012.
- DEGIORGI, François et Pierre-Marie BADOT (2020). “Étude de l'état de santé des rivières karstiques en relation avec les pressions anthropiques sur leurs bassins versants”. In : *Bilan Synthétique des Opérations Réalisées et des Recherches et Analyses Effectuées et Disponibles. Université de Franche-Comté*, p. 47.
- DELMOTTE, B, S RICCI, M ROCHOUX et O THUAL (2011). *Introduction to data assimilation. Application to the calibration of a model parameter on the 1D diffusion equation*.
- DEREPASKO, Diana, Joseph HA GUILLAUME, Avril C HORNE et Martin VOLK (2021). “Considering scale within optimization procedures for water management decisions : Balancing environmental flows and human needs”. In : *Environmental Modelling & Software* 139, p. 104991.
- DESMOITIER, Nicolas, Mathilde KOLENDA, Karen Holm OLSEN et Morten Walbech RYBERG (2023). “Methods for assessing social impacts of policies in relation to absolute boundaries”. In : *Environmental Impact Assessment Review* 101, p. 107098. ISSN : 0195-9255. DOI : 10.1016/j.eiar.2023.107098.
- DESROSIÈRES, Alain (2008). “Chapitre 1. La statistique, outil de gouvernement et outil de preuve. Introduction”. In : *Pour une sociologie historique de la quantification*. Presses des Mines. DOI : 10.4000/books.pressesmines.909.
- DIETZENBACHER, Erik, Bart LOS, Robert STEHRER, Marcel TIMMER et Gaaitzen DE VRIES (2013). “The construction of world input-output tables in the WIOD project”.

- In : *Economic Systems Research* 25.1, p. 71-98. ISSN : 0953-5314, 1469-5758. DOI : 10.1080/09535314.2012.761180.
- DUMERGUES, Laurent (2020). “Gaz à effet de serre : enjeux de réduction et méthodes de comptabilisation”. In : *Techniques de l’Ingénieur*. Article de référence.
- EHRlich, Paul R. et John P. HOLDREN (1971). “Impact of Population Growth : Complacency concerning this component of man’s predicament is unjustified and counter-productive.” In : *Science* 171.3977, p. 1212-1217. ISSN : 0036-8075, 1095-9203. DOI : 10.1126/science.171.3977.1212.
- EPURES (2021). *Quelle participation du Sud-Loire à l’atteinte des limites planétaires ?* Agence d’urbanisme de la région stéphanoise.
- FALKENMARK, M (1995). “Land-water linkages : a synopsis”. In : *FAO land and water bulletin* 1, p. 15-17.
- FALKENMARK, Malin et Johan ROCKSTRÖM (2004). *Balancing water for humans and nature : the new approach in ecohydrology*. Earthscan.
- FANNING, Andrew L, Daniel W O’NEILL, Jason HICKEL et Nicolas ROUX (2022). “The social shortfall and ecological overshoot of nations”. In : *Nature sustainability* 5.1, p. 26-36.
- FEKETE, Balázs M., Charles J. VÖRÖSMARTY et Wolfgang GRABS (2002). “High-resolution fields of global runoff combining observed river discharge and simulated water balances”. In : *Global Biogeochemical Cycles* 16.3. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/1999GB001254>, p. 15–1–15-10. ISSN : 1944-9224. DOI : 10.1029/1999GB001254.
- FLÖRKE, Martina, Ellen KYNAST, Ilona BÄRLUND, Stephanie EISNER, Florian WIMMER et Joseph ALCAMO (2013). “Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development : A global simulation study”. In : *Global Environmental Change* 23.1, p. 144-156. ISSN : 0959-3780. DOI : 10.1016/j.gloenvcha.2012.10.018.
- FNAU (2024). *Les limites planétaires : Une nouvelle boussole pour l’habitabilité de la Terre*. 15^e éd. Points Fnau - Alternatives. Forthcoming. Paris, France : Gallimard. ISBN : 978-2-07-299998-0.
- FORRESTER, Jay W. (1971). *World Dynamics*. Cambridge, MA : Wright-Allen Press. ISBN : 978-0-262-56065-9.

- FREYDANK, Katharina et Stefan SIEBERT (2008). *Towards mapping the extent of irrigation in the last century : time series of irrigated area per country*. Rapp. tech. Univ., Inst. of Physical Geography.
- FRIELER, Katja, Stefan LANGE, Franziska PIONTEK, Christopher PO REYER, Jacob SCHEWE, Lila WARSZAWSKI, Fang ZHAO, Louise CHINI, Sebastien DENVIL, Kerry EMANUEL et al. (2017). “Assessing the impacts of 1.5 C global warming–simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b)”. In : *Geoscientific Model Development* 10.12, p. 4321-4345.
- GEMENNE, François (2021). “3. Migrations et déplacements de populations”. In : *Géopolitique du climat*. T. 3e éd. Objectif Monde. Paris : Armand Colin, p. 63-79. ISBN : 978-2-200-63242-7.
- GENTY, Aurélien, Inaki ARTO et Frederik NEUWAHL (2012). *Final Database of Environmental Satellite Accounts : Technical Report on Their Compilation*. IPTS, p. 69.
- GERTEN, Dieter, Vera HECK, Jonas JÄGERMEYR, Benjamin Leon BODIRSKY, Ingo FETZER, Mika JALAVA, Matti KUMMU, Wolfgang LUCHT, Johan ROCKSTRÖM, Sibyll SCHAPHOFF et al. (2020). “Feeding ten billion people is possible within four terrestrial planetary boundaries”. In : *Nature Sustainability* 3.3, p. 200-208.
- GERTEN, Dieter, Holger HOFF, Johan ROCKSTRÖM, Jonas JÄGERMEYR, Matti KUMMU et Amandine V PASTOR (2013). “Towards a revised planetary boundary for consumptive freshwater use : role of environmental flow requirements”. In : *Current Opinion in Environmental Sustainability*. Aquatic and marine systems 5.6, p. 551-558. ISSN : 1877-3435. DOI : 10.1016/j.cosust.2013.11.001.
- GHORBANIDEHNO, Hojat, Amalia KOKKINAKI, Jonghyun LEE et Eric DARVE (2020). “Recent developments in fast and scalable inverse modeling and data assimilation methods in hydrology”. In : *Journal of Hydrology* 591, p. 125266.
- GILBERT, Richard O (1987). *Statistical methods for environmental pollution monitoring*. John Wiley & Sons.
- GLEESON, Tom, Lan WANG-ERLANDSSON, Miina PORKKA, Samuel C ZIPPER, Fernando JARAMILLO, Dieter GERTEN, Ingo FETZER, Sarah E CORNELL, Luigi PIEMONTESE, Line J GORDON et al. (2020a). “Illuminating water cycle modifications and Earth system resilience in the Anthropocene”. In : *Water Resources Research* 56.4, e2019WR024957.

- GLEESON, Tom, Lan WANG-ERLANDSSON, Samuel C. ZIPPER, Miina PORKKA, Fernando JARAMILLO, Dieter GERTEN, Ingo FETZER, Sarah E. CORNELL, Luigi PIEMONTESE, Line J. GORDON, Johan ROCKSTRÖM, Taikan OKI, Murugesu SIVAPALAN, Yoshihide WADA, Kate A. BRAUMAN, Martina FLÖRKE, Marc F. P. BIERKENS, Bernhard LEHNER, Patrick KEYS, Matti KUMMU, Thorsten WAGENER, Simon DADSON, Tara J. TROY, Will STEFFEN, Malin FALKENMARK et James S. FAMIGLIETTI (2020b). “The Water Planetary Boundary : Interrogation and Revision”. In : *One Earth* 2.3, p. 223-234. ISSN : 2590-3322. DOI : 10.1016/j.oneear.2020.02.009.
- GLEICK, Peter H., Gary H. WOLFF, Heather COOLEY, Meena PALANIAPPAN, Andrea SAMULON, Emily LEE, Jason MORRISON et David KATZ (2006). *The World's Water, Volume 5*. The Pacific Institute Island Press, p. 392.
- GLOBAL FOOTPRINT NETWORK (2023). *National Footprint and Biocapacity Accounts, 2023 Edition*. Oakland, CA, USA : Global Footprint Network.
- GONDRAN, Natacha (2012). “The ecological footprint as a follow-up tool for an administration : Application for the Vanoise National Park Ecological Indicator.” In : *Ecological Indicators*.
- GONDRAN, Natacha, Quentin DASSIBAT et Maud MARSAUCHE (2023). “Quelles utilisations du cadre des limites planétaires à l'échelle locale ?” In : *La France face aux neuf limites planétaires*. Théma. Commissariat général au développement durable, p. 61-65.
- GORDON, W. S., J. S. FAMIGLIETTI, N. L. FOWLER, T. G. F. KITTEL et K. A. HIBBARD (2004). “Validation of simulated runoff from six terrestrial ecosystem models : results from vemap”. In : *Ecological Applications* 14.2. Publisher : John Wiley & Sons, Ltd, p. 527-545. ISSN : 1051-0761. DOI : 10.1890/02-5287.
- GREEN, Andy J, Paloma ALCORLO, Edwin THM PEETERS, Edward P MORRIS, José L ESPINAR, Miguel Angel BRAVO-UTRERA, Javier BUSTAMANTE, Ricardo DÍAZ-DELGADO, Albert A KOELMANS, Rafael MATEO, Wolf M MOOIJ, Miguel RODRÍGUEZ-RODRÍGUEZ, Egbert H van NES et Marten SCHEFFER (2017). “Creating a safe operating space for wetlands in a changing climate”. In : *Frontiers in Ecology and the Environment* 15.2. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/fee.1459>, p. 99-107. ISSN : 1540-9309. DOI : 10.1002/fee.1459.

- GREEN, Pamela A. (2023). *Safe and Just Earth Systems Boundaries for Surface Water : Hydrologic Alteration of Environmental Flows*. Version v0.2.0. DOI : 10.5281/zenodo.7674802.
- GRENOBLE, Ville de (2023). *Grenoble en transition : rapport annuel sur la situation en matière de développement durable*. Ville de Grenoble, p. 122.
- GUPTA, Joyeeta, Diana LIVERMAN, Klaudia PRODANI, Paulina ALDUNCE, Xuemei BAI, Wendy BROADGATE, Daniel CIOBANU, Lauren GIFFORD, Chris GORDON, Margot HURLBERT, Cristina Y. A. INOUE, Lisa JACOBSON, Norichika KANIE, Steven J. LADE, Timothy M. LENTON, David OBURA, Chukwumerije OKEREKE, Ilona M. OTTO, Laura PEREIRA, Johan ROCKSTRÖM, Joeri SCHOLTENS, Juan ROCHA, Ben STEWART-KOSTER, J. DAVID TÀBARA, Crelis RAMMELT et Peter H. VERBURG (2023). “Earth system justice needed to identify and live within Earth system boundaries”. In : *Nature Sustainability* 6.6. Number : 6 Publisher : Nature Publishing Group, p. 630-638. ISSN : 2398-9629. DOI : 10.1038/s41893-023-01064-1.
- HALLMANN, Caspar A., Martin SORG, Eelke JONGEJANS, Henk SIEPEL, Nick HOFLAND, Heinz SCHWAN, Werner STENMANS, Andreas MÜLLER, Hubert SUMSER, Thomas HÖRREN, Dave GOULSON et Hans de KROON (2017). “More than 75 percent decline over 27 years in total flying insect biomass in protected areas”. In : *PLOS ONE* 12.10. Publisher : Public Library of Science, e0185809. ISSN : 1932-6203. DOI : 10.1371/journal.pone.0185809.
- HAUSCHILD, Michael Z., Sami KARA et Inge RØPKE (2020). “Absolute sustainability : Challenges to life cycle engineering”. In : *CIRP Annals* 69.2, p. 533-553. ISSN : 0007-8506. DOI : 10.1016/j.cirp.2020.05.004.
- HÄYHÄ, Tiina, Paul L LUCAS, Detlef P van VUUREN, Sarah E CORNELL et Holger HOFF (2016). “From Planetary Boundaries to national fair shares of the global safe operating space—How can the scales be bridged?” In : *Global Environmental Change* 40, p. 60-72.
- HEIDE, Mia, Michael Z. HAUSCHILD et Morten RYBERG (2023). “Reflecting the importance of human needs fulfilment in absolute sustainability assessments : Development of a sharing principle”. In : *Journal of Industrial Ecology* 27.4. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jiec.13405>, p. 1151-1164. ISSN : 1530-9290. DOI : 10.1111/jiec.13405.

- HEIJUNGS, Reinout, Arjan de KONING et Jeroen B. GUINÉE (2014). “Maximizing affluence within the planetary boundaries”. In : *The International Journal of Life Cycle Assessment* 19.6, p. 1331-1335. ISSN : 1614-7502. DOI : 10.1007/s11367-014-0729-y.
- HJALSTED, Anjila Wegge, Alexis LAURENT, Martin Marchman ANDERSEN, Karen Holm OLSEN, Morten RYBERG et Michael HAUSCHILD (2021). “Sharing the safe operating space : Exploring ethical allocation principles to operationalize the planetary boundaries and assess absolute sustainability at individual and industrial sector levels”. In : *Journal of Industrial Ecology* 25.1, p. 6-19.
- HOEKSTRA, A et P HUNG (2002). *Virtual Water Trade : A Quantification of Virtual Water Flows between Nations in Relation to International Crop Trade*. UNESCO-IHE Institute for Water Education.
- HOEKSTRA, A et M MEKONNEN (2010a). *The green, blue, grey water footprint of farm animals and animals products*. UNESCO-IHE Institute for Water Education.
- (2010b). *The green, blue and grey water footprint of crops and derived crop products*. UNESCO-IHE Institute for Water Education.
- (2011a). *Global water scarcity : the monthly blue water footprint compared to blue water availability for the world’s major river basins*. UNESCO-IHE Institute for Water Education.
- (2011b). *National Water Footprint accounts : the green, blue, grey water footprint of production and consumption*. UNESCO-IHE Institute for Water Education.
- HUANG, Zhongwei, Mohamad HEJAZI, Xinya LI, Qiuhong TANG, Chris VERNON, Guoyong LENG, Yaling LIU, Petra DÖLL, Stephanie EISNER, Dieter GERTEN et al. (2018). “Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns”. In : *Hydrology and Earth System Sciences* 22.4, p. 2117-2133.
- IGN (2006). *BD CARTHAGE ® version 3 - Descriptif de contenu*.
- IPCC (2018). *Global Warming of 1.5°C : An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. Geneva, Switzerland.
- (2021). *Sixth Assessment Report AR6 - Physical Basis*. IPCC.

- IPCC (2022). *Sixth Assessment Report AR6 - Impacts, Adaptation and Vulnerability*.
- (2023). “Technical Summary”. In : *Climate Change 2021 – The Physical Science Basis : Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Sous la dir. d’INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). Cambridge : Cambridge University Press, p. 35-144. ISBN : 978-1-00-915788-9. DOI : 10.1017/9781009157896.002.
- IUCN (2020). *IUCN Red List 2017-2020 Report*.
- KARSENBERG, Derek, Oliver SCHMITZ, Peter SALAMON, Kor DE JONG et Marc FP BIERKENS (2010). “A software framework for construction of process-based stochastic spatio-temporal models and data assimilation”. In : *Environmental Modelling & Software* 25.4, p. 489-502.
- KENDALL, Maurice George (1948). *Rank correlation methods*. Griffin.
- LAVISSE, Téo, Rémy PANARIELLO, Fabien PERDU, Sébastien ROLÈRE et Peggy ZWOLINSKI (2024). “Developing a Methodological Framework for Assessing Absolute Sustainability in Battery Upscaling within Planetary Boundaries”. In : *Procedia CIRP*. 31st CIRP Conference on Life Cycle Engineering 122, p. 766-771. ISSN : 2212-8271. DOI : 10.1016/j.procir.2024.01.106.
- LEHNER, Bernhard (2012). *Derivation of watershed boundaries for GRDC gauging stations based on the HydroSHEDS drainage network*. Koblenz, Germany : GRDC - Global Runoff Data Center, p. 18.
- LENTON, Timothy M., Hermann HELD, Elmar KRIEGLER, Jim W. HALL, Wolfgang LUCHT, Stefan RAHMSTORF et Hans Joachim SCHELLNHUBER (2008). “Tipping elements in the Earth’s climate system”. In : *Proceedings of the National Academy of Sciences* 105.6. Publisher : Proceedings of the National Academy of Sciences, p. 1786-1793. DOI : 10.1073/pnas.0705414105.
- LENZEN, Manfred, Arne GESCHKE, Muhammad Daaniyall ABD RAHMAN, Yanyan XIAO, Jacob FRY, Rachel REYES, Erik DIETZENBACHER, Satoshi INOMATA, Kei-chiro KANEMOTO, Bart LOS, Daniel MORAN, Hagen SCHULTE IN DEN BÄUMEN, Arnold TUKKER, Terrie WALMSLEY, Thomas WIEDMANN, Richard WOOD et Norihiko YAMANO (2017). “The Global MRIO Lab – charting the world economy”. In : *Economic Systems Research* 29.2. Publisher : Routledge _eprint :

<https://doi.org/10.1080/09535314.2017.1301887>, p. 158-186. ISSN : 0953-5314. DOI : 10.1080/09535314.2017.1301887.

- LENZEN, Manfred, Keiichiro KANEMOTO, Daniel MORAN et Arne GESCHKE (2012). “Mapping the Structure of the World Economy”. In : *Environmental Science & Technology* 46.15, p. 8374-8381. ISSN : 0013-936X, 1520-5851. DOI : 10.1021/es300171x.
- LI, Mo, Thomas WIEDMANN, Kai FANG et Michalis HADJIKAKOU (2021). “The role of planetary boundaries in assessing absolute environmental sustainability across scales”. In : *Environment International* 152, p. 106475. ISSN : 0160-4120. DOI : 10.1016/j.envint.2021.106475.
- LI, Mo, Thomas WIEDMANN, Junguo LIU, Yafei WANG, Yuanchao HU, Zongyong ZHANG et Michalis HADJIKAKOU (2020). “Exploring consumption-based planetary boundary indicators : An absolute water footprinting assessment of Chinese provinces and cities”. In : *Water Research* 184, p. 116163. ISSN : 0043-1354. DOI : 10.1016/j.watres.2020.116163.
- LIU, Xingcai, Wenfeng LIU, Liu LIU, Qihong TANG, Junguo LIU et Hong YANG (2021). “Environmental flow requirements largely reshape global surface water scarcity assessment”. In : *Environmental Research Letters* 16.10, p. 104029.
- LOISEAU, Eléonore, Lynda AISSANI, Samuel LE FÉON, Faustine LAURENT, Juliette CERCEAU, Serenella SALA et Philippe ROUX (2018). “Territorial Life Cycle Assessment (LCA) : What exactly is it about ? A proposal towards using a common terminology and a research agenda”. In : *Journal of Cleaner Production* 176, p. 474-485.
- LOVELOCK, J. E. (1972). “Gaia as seen through the atmosphere”. In : *Atmospheric Environment (1967)* 6.8, p. 579-580. ISSN : 0004-6981. DOI : 10.1016/0004-6981(72)90076-5.
- LOVELOCK, James E et Lynn MARGULIS (1974). “Atmospheric homeostasis by and for the biosphere : the Gaia hypothesis”. In : *Tellus* 26.1-2, p. 2-10.
- LRA, SAGE (2022). *Plaquette de présentation du PTGE LRA*.
- (2024). “Commission Locale de l’Eau - Présentation des résultats volet "milieu" étude HMUC”. CLE. <https://sage-loire-rhone-alpes.fr/?p=4086>.
- LUSSAULT, Michel et Valérie DISDIER (2021). *Néolithique Anthropocène. Dialogue autour des 12000 dernières années*. Editions 205. ISBN : 978-2-919380-40-4.

- M. ALGUNAIBET, Ibrahim, Carlos POZO, Ángel GALÁN-MARTÍN, Mark A. J. HUIJBREGTS, Niall Mac DOWELL et Gonzalo GUILLÉN-GOSÁLBEZ (2019). “Powering sustainable development within planetary boundaries”. In : *Energy & Environmental Science* 12.6. Publisher : Royal Society of Chemistry, p. 1890-1900. DOI : 10.1039/C8EE03423K.
- MARGAT, Jean, Karen FRENKEN et Jean-Luc FAURÈS (2005). *Key Water Resources Statistics in AQUASTAT*. Vienna : FAO’s Global Information System on Water et Agriculture, p. 13.
- MCLAUGHLIN, Dennis (2002). “An integrated approach to hydrologic data assimilation : interpolation, smoothing, and filtering”. In : *Advances in Water Resources* 25.8-12, p. 1275-1286.
- MEADOWS, Donella H, Dennis L MEADOWS, Jørgen RANDERS et William W BEHRENS III (1972). *The limits to growth-club of rome*. Club of Rome.
- MILLENNIUM ECOSYSTEM ASSESSMENT (2005). *Millennium Ecosystem Assessment Report*.
- MOLDAN, Bedřich, Svatava JANOUŠKOVÁ et Tomáš HÁK (2012). “How to understand and measure environmental sustainability : Indicators and targets”. In : *Ecological Indicators*. Indicators of environmental sustainability : From concept to applications 17, p. 4-13. ISSN : 1470-160X. DOI : 10.1016/j.ecolind.2011.04.033.
- MONFREDA, Chad, Navin RAMANKUTTY et Jonathan A FOLEY (2008). “Farming the planet : 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000”. In : *Global biogeochemical cycles* 22.1.
- MTES (2019). *L’état de l’environnement en France - Édition 2019*. Paris, France : Ministère de la Transition écologique et solidaire.
- (2023). *La France face aux neuf limites planétaires*. Ministère de la Transition Ecologique et de la Cohésion des Territoires.
- NATIONS, United (2015). *Transforming our World : The 2030 Agenda for Sustainable Development A/RES/70/1*. United Nations.
- O’CALLAGHAN, John F et David M MARK (1984). “The extraction of drainage networks from digital elevation data”. In : *Computer vision, graphics, and image processing* 28.3, p. 323-344.

- OFB et IGN (2020). *Document de présentation - Description du référentiel hydrographique (BD TOPAGE® version 1)*.
- PASTOR, A. V., F. LUDWIG, H. BIEMANS, H. HOFF et P. KABAT (2014). “Accounting for environmental flow requirements in global water assessments”. In : *Hydrology and Earth System Sciences* 18.12, p. 5041-5059. DOI : 10.5194/hess-18-5041-2014.
- PASTOR, A.V., D.C.S. VIEIRA, F.H. SOUDIJN et O.Y. EDELENBOSCH (2020). “How uncertainties are tackled in multi-disciplinary science? A review of integrated assessments under global change”. In : *CATENA* 186, p. 104305. ISSN : 03418162. DOI : 10.1016/j.catena.2019.104305.
- PAULET, F. Crastes de et S. LECONTE (2023). *Synthèse hydrogéologique concernant la rive droite de la plaine du Forez : Rapport final*. Rapp. tech. BRGM/RP-72380-FR. Version Version 2. Étude réalisée dans le cadre des opérations de service public du BRGM. BRGM.
- PERSSON, Linn, Bethanie M CARNEY ALMROTH, Christopher D COLLINS, Sarah CORNELL, Cynthia A DE WIT, Miriam L DIAMOND, Peter FANTKE, Martin HASSELLÖV, Matthew MACLEOD, Morten W RYBERG et al. (2022). “Outside the safe operating space of the planetary boundary for novel entities”. In : *Environmental science & technology* 56.3, p. 1510-1521.
- PFISTER, Stephan et Peter BAYER (2014). “Monthly water stress : spatially and temporally explicit consumptive water footprint of global crop production”. In : *Journal of Cleaner Production*. Towards eco-efficient agriculture and food systems : Selected papers from the Life Cycle Assessment (LCA) Food Conference, 2012, in Saint Malo, France 73, p. 52-62. ISSN : 0959-6526. DOI : 10.1016/j.jclepro.2013.11.031.
- PFISTER, Stephan, Annette KOEHLER et Stefanie HELLWEG (2009). “Assessing the Environmental Impacts of Freshwater Consumption in LCA”. In : *Environmental Science & Technology* 43.11. Publisher : American Chemical Society, p. 4098-4104. ISSN : 0013-936X. DOI : 10.1021/es802423e.
- PFISTER, Stephan, Samuel VIONNET, Tereza LEVOVA et Sebastien HUMBERT (2016). “Ecoinvent 3 : assessing water use in LCA and facilitating water footprinting”. In : *The International Journal of Life Cycle Assessment* 21.9, p. 1349-1360. ISSN : 1614-7502. DOI : 10.1007/s11367-015-0937-0.

- POFF, N Leroy et Julie KH ZIMMERMAN (2010). “Ecological responses to altered flow regimes : a literature review to inform the science and management of environmental flows”. In : *Freshwater biology* 55.1, p. 194-205.
- PORKKA, Miina, Vili VIRKKI, Lan WANG-ERLANDSSON, Dieter GERTEN, Tom GLEESON, Chinchu MOHAN, Ingo FETZER, Fernando JARAMILLO, Arie STAAL, Sofie te WIERIK et al. (2024). “Notable shifts beyond pre-industrial streamflow and soil moisture conditions transgress the planetary boundary for freshwater change”. In : *Nature Water* 2.3, p. 262-273.
- RAMÍREZ, Francisco, Maria Grazia PENNINO, Marta ALBO-PUIGSERVER, Jeroen STEENBEEK, Jose M. BELLIDO et Marta COLL (2021). “SOS small pelagics : A safe operating space for small pelagic fish in the western Mediterranean Sea”. In : *Science of The Total Environment* 756, p. 144002. ISSN : 0048-9697. DOI : 10.1016/j.scitotenv.2020.144002.
- RAWORTH, Kate (2012). *A safe and just space for humanity*. Oxfam Discussion Papers. Oxfam.
- (2017). *Doughnut Economics : Seven Ways to Think Like a 21st-Century Economist*. Random House. ISBN : 1-84794-137-0.
- RAWORTH, Kate, Olya KRESTYANINOVA, Fredrik ERIKSSON, Lucy FEIBUSCH, Carlota SANZ, Janine BENYUS, Jamie DWYER, Nicole Hagerman MILLER, Annerieke DOUMA, Inge ter LAAK, Nicolas RASPAIL, Liesl EHLERS et Julia LIPTON (2020). *The Amsterdam City Doughnut : A Tool for Transformative Action*. Rapp. tech. Commissioned by the City of Amsterdam. Doughnut Economics Action Lab (DEAL).
- REES, William E (1992). “Ecological footprints and appropriated carrying capacity : what urban economics leaves out”. In : *Environment and urbanization* 4.2, p. 121-130.
- REICHLER, Rolf H (2008). “Data assimilation methods in the Earth sciences”. In : *Advances in water resources* 31.11, p. 1411-1418.
- RICHARD, Jacques (2012). *Comptabilité et Développement Durable*. Paris, France : Economica. ISBN : 978-2-7178-6242-1.
- RICHARDSON, Katherine, Will STEFFEN, Wolfgang LUCHT, Jørgen BENDTSEN, Sarah E. CORNELL, Jonathan F. DONGES, Markus DRÜKE, Ingo FETZER, Govindasamy BALA, Werner von BLOH, Georg FEULNER, Stephanie FIEDLER, Dieter GERTEN, Tom GLEESON, Matthias HOFMANN, Willem HUISKAMP, Matti KUMMU,

- Chinchu MOHAN, David NOGUÉS-BRAVO, Stefan PETRI, Miina PORKKA, Stefan RAHMSTORF, Sibyll SCHAPHOFF, Kirsten THONICKE, Arne TOBIAN, Vili VIRKKI, Lan WANG-ERLANDSSON, Lisa WEBER et Johan ROCKSTRÖM (2023). “Earth beyond six of nine planetary boundaries”. In : *Science Advances* 9.37. Publisher : American Association for the Advancement of Science, eadh2458. DOI : 10.1126/sciadv.adh2458.
- RICHTER, Brian D, MM DAVIS, Colin APSE et Christopher KONRAD (2012). “A presumptive standard for environmental flow protection”. In : *River Research and Applications* 28.8, p. 1312-1321.
- ROCKSTROM, Johan, Joyeeta GUPTA, Dahe QIN, Steven LADE, Jesse ABRAMS, Lauren ANDERSEN, David ARMSTRONG, Xuemei BAI, Govindasamy BALA et Stuart BUNN (2023). “Safe and just Earth system boundaries”. In : *Nature* 619.7968, p. 102-111.
- ROCKSTRÖM, Johan (1997). “On-farm agrohydrological analysis of the Sahelian yield crisis : rainfall partitioning, soil nutrients and water use efficiency of pearl millet”. Härtill sex uppsatser. Thèse de doct. Stockholm University, Department of Systems Ecology, p. 62. ISBN : 91-7153-654-X.
- ROCKSTRÖM, Johan, Will STEFFEN, Kevin NOONE, Åsa PERSSON, F. Stuart CHAPIN, Eric F. LAMBIN, Timothy M. LENTON, Marten SCHEFFER, Carl FOLKE, Hans Joachim SCHELLNHUBER, Björn NYKVIST, Cynthia A. de WIT, Terry HUGHES, Sander van der LEEUW, Henning RODHE, Sverker SÖRLIN, Peter K. SNYDER, Robert COSTANZA, Uno SVEDIN, Malin FALKENMARK, Louise KARLBERG, Robert W. CORELL, Victoria J. FABRY, James HANSEN, Brian WALKER, Diana LIVERMAN, Katherine RICHARDSON, Paul CRUTZEN et Jonathan A. FOLEY (2009a). “A safe operating space for humanity”. In : *Nature* 461.7263. Publisher : Nature Publishing Group, p. 472-475. ISSN : 1476-4687. DOI : 10.1038/461472a.
- ROCKSTRÖM, Johan, Will STEFFEN, Kevin NOONE, Åsa PERSSON, F. Stuart III CHAPIN, Eric LAMBIN, Timothy LENTON, Marten SCHEFFER, Carl FOLKE, Hans Joachim SCHELLNHUBER, Björn NYKVIST, Cynthia de WIT, Terry HUGHES, Sander van der LEEUW, Henning RODHE, Sverker SÖRLIN, Peter SNYDER, Robert COSTANZA, Uno SVEDIN, Malin FALKENMARK, Louise KARLBERG, Robert CORELL, Victoria FABRY, James HANSEN, Brian WALKER, Diana LIVERMAN, Katherine RICHARDSON, Paul CRUTZEN et Jonathan FOLEY (2009b). “Planetary

- Boundaries : Exploring the Safe Operating Space for Humanity”. In : *Ecology and Society* 14.2. Publisher : The Resilience Alliance. ISSN : 1708-3087. DOI : 10.5751/ES-03180-140232.
- ROCKSTRÖM, Johan, Lan WANG-ERLANDSSON, Carl FOLKE, Dieter GERTEN, Line J. GORDON et Patrick W. KEYS (2024). “Malin Falkenmark : Water pioneer who coined the notion of water crowding and coloured the water cycle”. In : *Ambio* 53.5, p. 657-663. ISSN : 1654-7209. DOI : 10.1007/s13280-024-01989-7.
- RODRIGUES, Jérémy, Natacha GONDRAN, Adrien BEZIAT et Valérie LAFOREST (2021). “Application of the absolute environmental sustainability assessment framework to multifunctional systems – The case of municipal solid waste management”. In : *Journal of Cleaner Production* 322, p. 129034. ISSN : 0959-6526. DOI : 10.1016/j.jclepro.2021.129034.
- ROST, Stefanie, Dieter GERTEN, Alberte BONDEAU, Wolfgang LUCHT, Janine ROHWER et Sibyll SCHAPHOFF (2008). “Agricultural green and blue water consumption and its influence on the global water system”. In : *Water Resources Research* 44.9.
- RYBERG, Morten W., Martin Marchman ANDERSEN, Mikołaj OWSIANIAK et Michael Z. HAUSCHILD (2020). “Downscaling the planetary boundaries in absolute environmental sustainability assessments – A review”. In : *Journal of Cleaner Production* 276, p. 123287. ISSN : 0959-6526. DOI : 10.1016/j.jclepro.2020.123287.
- RYBERG, Morten W., Mikołaj OWSIANIAK, Katherine RICHARDSON et Michael Z. HAUSCHILD (2018). “Development of a life-cycle impact assessment methodology linked to the Planetary Boundaries framework”. In : *Ecological Indicators* 88, p. 250-262. ISSN : 1470-160X. DOI : 10.1016/j.ecolind.2017.12.065.
- SALA, Serenella, Biagio CIUFFO et Peter NIJKAMP (2015). “A systemic framework for sustainability assessment”. In : *Ecological economics* 119, p. 314-325.
- SALA, Serenella, Eleonora CRENNNA, Michela SECCHI et Esther SANYÉ-MENGUAL (2020). “Environmental sustainability of European production and consumption assessed against planetary boundaries”. In : *Journal of Environmental Management* 269, p. 110686. ISSN : 0301-4797. DOI : 10.1016/j.jenvman.2020.110686.
- SEN, Pranab Kumar (1968). “Estimates of the regression coefficient based on Kendall’s tau”. In : *Journal of the American statistical association* 63.324, p. 1379-1389.

- SHEPARD, Donald (1968). “A two-dimensional interpolation function for irregularly-spaced data”. In : *Proceedings of the 1968 23rd ACM national conference*, p. 517-524.
- SHERMAN, LeRoy K (1932). “The relation of hydrographs of runoff to size and character of drainage-basins”. In : *Eos, Transactions American Geophysical Union* 13.1, p. 332-339.
- SHIKLOMANOV, Igor A (2000). “World water resources and water use : Present assessment and outlook for 2025”. In : *World water scenarios analyses* 396.
- SIEBERT, Stefan et Petra DÖLL (2010). “Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation”. In : *Journal of Hydrology* 384.3-4, p. 198-217.
- SILVA, Samara Fernanda da, Diego Cezar dos SANTOS ARAÚJO, Jussara FREIRE DE SOUZA VIANA, Andrea SOUSA FONTES, Yvonilde DANTAS PINTO MEDEIROS et Suzana Maria GICO LIMA MONTENEGRO (2024). “Analysis of the correlation between land use and surface runoff in a Brazilian savanna basin”. In : *Journal of South American Earth Sciences* 133, p. 104724. ISSN : 0895-9811. DOI : 10.1016/j.jsames.2023.104724.
- SØRUP, Hjalte J. D., Sarah BRUDLER, Berit GODSKESEN, Yan DONG, Sara M. LERER, Martin RYGAARD et Karsten ARNBJERG-NIELSEN (2020). “Urban water management : Can UN SDG 6 be met within the Planetary Boundaries?” In : *Environmental Science & Policy* 106, p. 36-39. ISSN : 1462-9011. DOI : 10.1016/j.envsci.2020.01.015.
- STADLER, Konstantin, Richard WOOD, Tatyana BULAVSKAYA, Carl-Johan SÖDERSTEN, Moana SIMAS, Sarah SCHMIDT, Arkaitz USUBIAGA, José ACOSTA-FERNÁNDEZ, Jeroen KUENEN, Martin BRUCKNER, Stefan GILJUM, Stephan LUTTER, Stefano MERCIAI, Jannick H. SCHMIDT, Michaela C. THEURL, Christoph PLUTZAR, Thomas KASTNER, Nina EISENMENGER, Karl-Heinz ERB, Arjan de KONING et Arnold TUKKER (2018). “EXIOBASE 3 : Developing a Time Series of Detailed Environmentally Extended Multi-Regional Input-Output Tables”. In : *Journal of Industrial Ecology* 22.3. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jiec.12715>, p. 502-515. ISSN : 1530-9290. DOI : 10.1111/jiec.12715.

- STATISTICS, Oxford Dictionary of (2008). *Proxy Variable*. Published online : 2008, Current Online Version : 2008. Oxford University Press. ISBN : 9780199541454. eprint : 9780191726866.
- STEFFEN, Will, Wendy BROADGATE, Lisa DEUTSCH, Owen GAFFNEY et Cornelia LUDWIG (2015a). “The trajectory of the Anthropocene : The Great Acceleration”. In : *The Anthropocene Review* 2.1. Publisher : SAGE Publications, p. 81-98. ISSN : 2053-0196. DOI : 10.1177/2053019614564785.
- STEFFEN, Will, Katherine RICHARDSON, Johan ROCKSTRÖM, Sarah CORNELL, Ingo FETZER, Elena BENNETT, Reinette BIGGS, Stephen CARPENTER, Wim VRIES, Cynthia de WIT, Carl FOLKE, Dieter GERTEN, Jens HEINKE, Linn PERSSON, Veerabhadran RAMANATHAN, Belinda REYERS et Sverker SÖRLIN (2015b). “Planetary Boundaries : Guiding Human Development on a Changing Planet”. In : *Science*. DOI : 10.1126/science.1259855.
- STEFFEN, Will, Katherine RICHARDSON, Johan ROCKSTRÖM, Hans Joachim SCHELLNHUBER, Opha Pauline DUBE, Sébastien DUTREUIL, Timothy M LENTON et Jane LUBCHENCO (2020a). “The emergence and evolution of Earth System Science”. In : *Nature Reviews Earth & Environment* 1.1, p. 54-63.
- (2020b). “The emergence and evolution of Earth System Science”. In : *Nature Reviews Earth & Environment* 1.1, p. 54-63. ISSN : 2662-138X. DOI : 10.1038/s43017-019-0005-6.
- STEFFEN, Will, Johan ROCKSTROM, Katherine RICHARDSON, Timothy LENTON, Carl FOLKE, Diana LIVERMAN, Colin SUMMERHAYES, Anthony BARNOSKY, Sarah CORNELL et Michel CRUCIFIX (2018). “Trajectories of the Earth System in the Anthropocene”. In : *Proceedings of the National Academy of Sciences* 115.33, p. 8252-8259.
- STEFFEN, Will, Angelina SANDERSON, Peter TYSON, Jill JÄGER, Pamela MATSON, Berrien MOORE, Frank OLDFIELD, Katherine RICHARDSON, H. John SCHELLNHUBER, B. L. TURNER et Robert J. WASSON (2005). *Global Change and the Earth System : A Planet Under Pressure*. Global Change — The IGBP Series. Berlin, Heidelberg : Springer. ISBN : 978-3-540-26594-8 978-3-540-26607-5. DOI : 10.1007/b137870.
- STEWART-KOSTER, Ben, Stuart E. BUNN, Pamela GREEN, Christopher NDEHEDEHE, Lauren S. ANDERSEN, David I. ARMSTRONG MCKAY, Xuemei BAI, Fabrice

- DECLERCK, Kristie L. EBI, Christopher GORDON, Joyeeta GUPTA, Syezlin HASAN, Lisa JACOBSON, Steven J. LADE, Diana LIVERMAN, Sina LORIANI, Awaz MOHAMED, Nebojsa NAKICENOVIC, David OBURA, Dahe QIN, Crellis RAMMELT, Juan C. ROCHA, Johan ROCKSTRÖM, Peter H. VERBURG et Caroline ZIMM (2024). “Living within the safe and just Earth system boundaries for blue water”. In : *Nature Sustainability* 7.1. Publisher : Nature Publishing Group, p. 53-63. ISSN : 2398-9629. DOI : 10.1038/s41893-023-01247-w.
- STRAHLER, Arthur N (1957). “Quantitative analysis of watershed geomorphology”. In : *Eos, Transactions American Geophysical Union* 38.6, p. 913-920.
- SYMPOSIUM, 10th International River (2007). *The Brisbane Declaration*.
- SYNDICAT MIXTE DU SUD-LOIRE (2022). *Plaquette de présentation de la démarche du Plan Paysage du Sud-Loire*. Brochure. Réalisation accompagnée par les bureaux d’études Fabriques et Caudex.
- SZABÓ, Zoltán, Krisztina BUCZKÓ, Aritina HALIUC, Ilona PÁL, János L. KORPONAI, Róbert-Csaba BEGY, Daniel VERES, Tomi P. LUOTO, Andreea R. ZSIGMOND et Enikő K. MAGYARI (2020). “Ecosystem shift of a mountain lake under climate and human pressure : A move out from the safe operating space”. In : *Science of The Total Environment* 743, p. 140584. ISSN : 0048-9697. DOI : 10.1016/j.scitotenv.2020.140584.
- TAN, Nairong, Xiaolei WANG, Hao WANG, Zhuoqun GAO, Xiaoying CHANG et Tao MA (2022). “Downscaling of planetary boundaries and sustainability management : A nexus analysis of water, land and major functions at the national-provincial level”. In : *Sustainable Horizons* 3, p. 100028. ISSN : 2772-7378. DOI : 10.1016/j.horiz.2022.100028.
- TARBOTON, David G, Rafael L BRAS et Ignacio RODRIGUEZ-ITURBE (1991). “On the extraction of channel networks from digital elevation data”. In : *Hydrological processes* 5.1, p. 81-100.
- TELTEU, Camelia-Eliza, Hannes MÜLLER SCHMIED, Wim THIERY, Guoyong LENG, Peter BUREK, Xingcai LIU, Julien Eric Stanislas BOULANGE, Lauren Seaby ANDERSEN, Manolis GRILLAKIS, Simon Newland GOSLING et al. (2021). “Understanding each other’s models : an introduction and a standard representation of 16 global water mo-

- dels to support intercomparison, improvement, and communication”. In : *Geoscientific Model Development* 14.6, p. 3843-3878.
- TOMLIN, CD et JK BERRY (1979). “Mathematical structure for cartographic modeling in environmental analysis”. In : *Proceedings of the American Congress on Surveying and Mapping annual meeting*.
- TOWNSEND, Peter (1987). “Deprivation”. In : *Journal of Social Policy* 16.2. Publisher : Cambridge University Press, p. 125-146. ISSN : 1469-7823, 0047-2794. DOI : 10.1017/S0047279400020341.
- UNITED NATIONS (1997). *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. Conference of the Parties, Third Session, Kyoto. United Nations Framework Convention on Climate Change.
- (2015). *Paris Agreement*. Conference of the Parties, Twenty-first Session, Paris. United Nations Framework Convention on Climate Change.
- UPPALA, Sakari M, PW KÅLLBERG, Adrian J SIMMONS, U ANDRAE, V Da Costa BECHTOLD, M FIORINO, JK GIBSON, J HASELER, A HERNANDEZ, GA KELLY et al. (2005). “The ERA-40 re-analysis”. In : *Quarterly Journal of the Royal Meteorological Society : A journal of the atmospheric sciences, applied meteorology and physical oceanography* 131.612, p. 2961-3012.
- VANHAM, Davy, Mesfin M. MEKONNEN et Ashok K. CHAPAGAIN (2020). “Arjen Y. Hoekstra 1967–2019”. In : *Nature Sustainability* 3.2. Publisher : Nature Publishing Group, p. 80-80. ISSN : 2398-9629. DOI : 10.1038/s41893-019-0468-y.
- VARGAS-GONZALEZ, Marcial, François WITTE, Patricia MARTZ, Laurent GILBERT, Sébastien HUMBERT, Olivier JOLLIET, Rosalie van ZELM et Jacques L’HARIDON (2019). “Operational Life Cycle Impact Assessment weighting factors based on Planetary Boundaries : Applied to cosmetic products”. In : *Ecological Indicators* 107, p. 105498. ISSN : 1470-160X. DOI : 10.1016/j.ecolind.2019.105498.
- VEA, Eldbjørg Blikra, Morten RYBERG, Katherine RICHARDSON et Michael Z. HAUSCHILD (2020). “Framework to define environmental sustainability boundaries and a review of current approaches”. In : *Environmental Research Letters* 15.10. Publisher : IOP Publishing, p. 103003. ISSN : 1748-9326. DOI : 10.1088/1748-9326/abac77.

- VILLE DE GRENOBLE (2022). *Rapport annuel sur le développement durable 2022 : Grenoble en transitions*. Grenoble, France : Ville de Grenoble.
- VIONNET, S, L LESSARD, T LEVOVA et S HUMBERT (2012). *Quantis water database—technical report*. Lausanne : Quantis International, p. 89.
- VOROSMARTY, Charles J, Pamela GREEN, Joseph SALISBURY et Richard B LAMMERS (2000). “Global water resources : vulnerability from climate change and population growth”. In : *science* 289.5477, p. 284-288.
- VYSNA, V., J. MAES, J.E. PETERSEN, A. LA NOTTE, S. VALLECILLO, N. AIZPURUA, E. IVITS et A. TELLER (2021). *Accounting for ecosystems and their services in the European Union (INCA) : Final report from phase II of the INCA project aiming to develop a pilot for an integrated system of ecosystem accounts for the EU*. Statistical report. Luxembourg : Publications Office of the European Union.
- WANG-ERLANDSSON, Lan, Arne TOBIAN, Ruud J VAN DER ENT, Ingo FETZER, Sofie te WIERIK, Miina PORKKA, Arie STAAL, Fernando JARAMILLO, Heindriken DAHLMANN, Chandrakant SINGH et al. (2022). “A planetary boundary for green water”. In : *Nature Reviews Earth & Environment* 3.6, p. 380-392.
- WCED (1987). *Our Common Future*. Oxford, UK : Oxford University Press.
- WERNET, Gregor, Christian BAUER, Bernhard STEUBING, Jürgen REINHARD, Emilia MORENO-RUIZ et Bo WEIDEMA (2016). “The ecoinvent database version 3 (part I) : overview and methodology”. In : *The International Journal of Life Cycle Assessment* 21.9, p. 1218-1230. ISSN : 1614-7502. DOI : 10.1007/s11367-016-1087-8.
- WISSER, D., B. M. FEKETE, C. J. VÖRÖSMARTY et A. H. SCHUMANN (2010). “Reconstructing 20th century global hydrography : a contribution to the Global Terrestrial Network- Hydrology (GTN-H)”. In : *Hydrology and Earth System Sciences* 14.1, p. 1-24. DOI : 10.5194/hess-14-1-2010.
- WWF (2018). *Living Planet Report 2018*.
- XUE, Ying et Bhavik R. BAKSHI (2022). “Metrics for a nature-positive world : A multiscale approach for absolute environmental sustainability assessment”. In : *Science of The Total Environment* 846, p. 157373. ISSN : 0048-9697. DOI : 10.1016/j.scitotenv.2022.157373.

- YANG, Yi et Reinout HEIJUNGS (2017). “A generalized computational structure for regional life-cycle assessment”. In : *The International Journal of Life Cycle Assessment* 22.2, p. 213-221. ISSN : 1614-7502. DOI : 10.1007/s11367-016-1155-0.
- YANG, Yi, Wesley W. INGWERSEN et David E. MEYER (2018). “Exploring the relevance of spatial scale to life cycle inventory results using environmentally-extended input-output models of the United States”. In : *Environmental Modelling & Software* 99, p. 52-57. ISSN : 1364-8152. DOI : 10.1016/j.envsoft.2017.09.017.
- YOUNG, Derek Scott (2018). *Handbook of regression methods*. Chapman et Hall/CRC.
- ZIPPER, Samuel, Fernando JARAMILLO, Lan WANG-ERLANDSSON, Sarah CORNELL, Tom GLEESON, Miina PORKKA, Tiina HÄYHÄ, Anne-Sophie CRÉPIN, Ingo FETZER, Dieter GERTEN, Holger HOFF, Nathaniel MATTHEWS, Constanza RICAURTE-VILLOTA, Matti KUMMU, Yoshihide WADA et Line GORDON (2020). “Integrating the Water Planetary Boundary With Water Management From Local to Global Scales”. In : *Earth's Future* 8.2. _eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019EF001377>, e2019EF001377. ISSN : 2328-4277. DOI : 10.1029/2019EF001377.

Annexes

1 Annexes au chapitre 3

1.1 Annexe à la section 4.2.2

À partir de cette formalisation élémentaire, nous allons montrer que le MRH appartient à la famille des modèles de *data assimilation*.

Ce terme recouvre un ensemble de méthodes qui visent toutes à coupler des données dites numériques ou modélisées avec des données dites d'observation ou de mesure afin d'améliorer la description du système étudié, c'est-à-dire de réduire les incertitudes entourant la description du système, et plus précisément l'incertitude liée à l'imperfection des données de mesure et à la paramétrisation des données modélisées (DELMOTTE et al. 2011). La *data assimilation* s'inscrit plus généralement dans le champ de la "modélisation inverse" (*inverse modeling*, par opposition à *forward modeling*) en cela que la démarche ne va plus "des paramètres vers les sorties de modèles" mais "des sorties de modèle vers l'optimisation des paramètres d'entrée" (GHORBANIDEHNO et al. 2020). Les modèles hydrologiques globaux que nous avons déjà eu l'occasion de présenter n'ont pas recours à la *data assimilation*, ce qui s'explique moins par le fait que, ayant un domaine d'étude global, la question des bornes du système ne se poserait pas²⁷, que par une "philosophie" différente. Comme nous avons pu le voir, la calibration de ces modèles sur des données de mesure n'est pas l'enjeu central de leur construction. Le *Water Balance Model* par exemple n'est calibré que sur un nombre de pixels très restreint comparé à l'ensemble des pixels de son maillage, qui correspondent à la position de quelques centaines de stations de mesure (FEKETE et al. 2002). La calibration du WBM intervient d'ailleurs à la fin, comme validation et pas tellement comme correction interne du modèle lui-même. Ces modèles

27. elle se pose en réalité mais moins trivialement : les bornes dans un système global sont par exemple les bornes entre les sous-systèmes constitutifs du système global, typiquement les échanges entre les compartiments neige, humidité et ruissellement dans un modèle hydrologique, ou encore les échanges physiques aux bornes entre les pixels élémentaires du maillage

ne sont en effet pas orientés vers la donnée (*data-driven*) du fait d'une production et d'un accès historiquement limités à la donnée, ce qui positionnement épistémologiquement le champ de l'hydrologie comme une "science des données limitées" (*data-limited science*).

For the most part, hydrology has been a data-limited science. Most of the fluxes in the hydrologic cycle are difficult to measure directly, especially over extensive areas. Even readily measured hydrologic variables are typically available only at isolated locations or times. This situation may change dramatically over the next few decades with the increased availability of remote sensing data and automated ground-based sensors. [...] If hydrologists are to make best use of emerging remote and ground-based data sources they will need to devote more attention to intelligent methods for processing and interpreting large amounts of information. Traditional manual approaches will simply not be able to cope with the volume of data generated by new sensors. Even computer-based estimation procedures may be overwhelmed. Meteorologists and oceanographers have addressed these challenges by developing a set of so-called data assimilation techniques. (MCLAUGHLIN 2002)

De fait, nous ne sommes pas parvenus à trouver d'articles qui proposent de coupler un modèle hydrologique (global) avec des techniques de *data assimilation*, pourtant très utilisées dans les champs voisins de la météorologie et de l'océanographie (qui sont d'ailleurs à l'origine de ces techniques statistiques), ainsi qu'en hydrogéologie, comme le remarquent dans leur revue de la littérature sur les applications de la *data assimilation* en hydrologie (GHORBANIDEHNO et al. 2020) : "*monitoring terrestrial hydrological processes like snow cover, precipitation and drought, and characterization of aquifers, [...] tracking non-aqueous fluids or solutes*".

Le modèle MRH dans son ensemble peut s'apparenter à une méthode de *data assimilation* puisque son but est de permettre la comparaison pixel-à-pixel entre les sorties d'un modèle hydrologique et des données de mesure, et ceci à travers deux modèles intermédiaires, un "modèle de dérasterisation" qui rééchantillonne le maillage de faible résolution d'un modèle hydrologique global vers un maillage à forte résolution, et un "modèle de rasterisation" qui projette les données de mesure définies par une géométrie `Point` vers le réseau hydrographique modélisé dans le maillage à forte résolution. Ces deux modèles intermédiaires ne sont pas sans faire penser à ce qui est appelé un *observation operator* dans le vocabulaire de la *data assimilation*.

Nous détaillons davantage cet aspect en annexe, où nous formalisons le lien entre le modèle MRH et l'écriture générique d'un modèle de *data assimilation*.

As the background [first guess of a set of parameters] and observational information need to be combined in the course of data assimilation, an operator H mapping a parameter from the control space onto the observation space is required. This observation operator is generally non-linear as a composition of a model integration M (from model parameters to state variables) and of an interpolation process (from grid points to observational points). (DELMOTTE et al. 2011).

Dans son article, (MCLAUGHLIN 2002) propose de distinguer entre trois types de méthodes de *data assimilation* dans des applications à l'hydrologie :

[...] it is helpful to distinguish the three variations on the general data assimilation problem : interpolation, smoothing, and filtering. Interpolation problems frequently arise when we wish to construct a static map of a spatially distributed variable (e.g. instantaneous precipitation) from scattered point measurements. Smoothing (or reanalysis) problems arise when we wish to reconstruct historical conditions (e.g. climate fluctuations over the last few decades) from an existing data set. Filtering problems arise in real-time control applications or when we wish to initialize models for forecasting purposes. We distinguish these three data assimilation problems because they serve different objectives and require somewhat different approaches. In the following sections we consider each problem in more detail and discuss possible solutions. (MCLAUGHLIN 2002)

Le modèle MRH peut ainsi être caractérisé de "problème d'interpolation".

Quant aux méthodes numériques disponibles pour optimiser les paramètres d'un modèle, de très nombreuses méthodes existent et ont déjà été employées dans le champ de l'hydrologie, comme le soulignent (GHORBANIDEHNO et al. 2020) qui proposent de distinguer les méthodes numériques qu'ils appellent *déterministes*, des méthodes *stochastiques* : "*Deterministic inversion methods (also referred to as parameter estimation or model calibration) use gradient-based techniques to obtain estimates that minimize the misfit between observations and model predictions. [...] Stochastic methods also obtain estimates that minimize the misfit between observations*

and model predictions, but conceptualize the unknowns and data as random variables with probability distributions that represent their uncertainty." S'il peut exister mathématiquement une solution optimale pour l'optimisation du vecteur des paramètres, les dimensions manipulées, souvent très grandes, obligent à l'utilisation de techniques dites variationnelles (*variational data assimilation*) (REICHLÉ 2008).

Sans présenter ici dans le détail les différentes méthodes d'assimilation, qui sont mathématiquement non-triviales, nous pouvons toutefois poser quelques éléments de vocabulaire afin d'illustrer la proximité conceptuelle entre la *data assimilation* et le modèle MRH. (DELMOTTE et al. 2011) et (REICHLÉ 2008) proposent par exemple une formulation élémentaire et générale pour un modèle de *data assimilation*, dont nous reprenons la plupart des notations.

On peut commencer par poser de façon très générale le modèle $Y = M(X)$ où Y est le vecteur des sorties du modèle M compte tenu du vecteur des entrées X (paramètres du modèle). Il existe parmi les différentes valeurs possibles de X une solution X^t qui correspond à la réalité, appelée *true vector*. On note X^b la configuration initiale de X , appelée *background vector*, qui pourrait correspondre à une première estimation non optimale de X et X^b correspond à l'espace de contrôle (*control space*). On note aussi X^a la configuration optimale de X , appelée *analysis vector*. X^a est ainsi plus proche de X^t que ne l'est X^b . On note également Y^0 le vecteur des données de mesure correspondant aux sorties du modèle et Y^0 correspond à l'espace des observations (*observation space*). Le lien entre l'espace des observations Y^0 et l'espace de contrôle n'est pas direct, il faut un opérateur d'observation (*observation operator*), noté H et défini non-linéairement comme composition entre le modèle M et une interpolation S (qui projette typiquement les données de mesure de façon ponctuelle dans le maillage du modèle). Ainsi : $H(X) = S(M(X))$. La démarche consiste ensuite à reconnaître l'existence d'incertitudes entourant les valeurs de X^b et Y^0 , respectivement notées ϵ^b et ϵ^0 , et qui sont inconnues mais estimées par $\epsilon^b \sim \mathcal{N}(0, B)$ et $\epsilon^0 \sim \mathcal{N}(0, R)$ où B (resp. R) est la matrice de covariance des erreurs de X^b (resp. Y^0). On peut alors poser le problème inverse : en connaissant X^b et Y^0 , on cherche à approximer la meilleur valeur de X^t telle que :

$$\begin{cases} Y^0 = H(X^t) + \epsilon^0 \\ X^b = X^t + \epsilon^b \end{cases}$$

La résolution de ce système implique ensuite différents types d'algorithmes dont on peut trouver une description générique dans, notamment, (DELMOTTE et al. 2011; MCLAUGHLIN 2002;

REICHLÉ 2008).

En reprenant les notations et le vocabulaire introduits, on comprend que le modèle MRH dans son ensemble est conceptuellement rattaché à la famille des méthodes de *data assimilation*.

Par exemple, on pourrait souhaiter optimiser seulement la valeur du paramètre β_p dans le modèle MRH. Rappelons que l'indice p désigne le $p^{\text{ième}}$ quantile de la fonction de répartition des charges hydrauliques entrant sur le domaine d'étude par tout pixel situé en bordure du domaine. L'introduction de p permet de réduire le nombre de pixels contenus dans β drastiquement car la fonction de répartition suit une loi log-normale, et donc d'alléger le travail manuel d'identification que nous avons déjà décrit. Ainsi, β dénote la valeur théorique des conditions aux bornes, qui correspond à X^t dans les notations précédemment introduites, et β_p est une réalisation "en première instance" de β , qui correspond au vecteur X^b . Dans notre cas, la valeur de Y^0 n'est pas évidente à déterminer car notre terrain d'étude ne dispose pas d'une instrumentation qui permette de déterminer la charge hydraulique au voisinage des pixels en bordure du domaine d'étude. Quant à H , il s'agit du modèle de rasterisation des données ponctuelles d'observation vers le maillage du réseau hydrographique modélisé.

Plus généralement, le vecteur X^a désigne la réalisation optimale de l'ensemble des paramètres du modèle MRH, en posant $X^b = (\alpha_a, \alpha_b, \beta_p, \delta)$. On imagine alors bien que la résolution de X^a n'est pas triviale numériquement.

1.2 Annexe à la section 4.2.3

Deux remarques sur la nature précise de qui est modélisé par l'algorithme de convolution hydrographique linéaire dans le modèle MRH.

Première remarque : modéliser le ruissellement *versus* modéliser l'écoulement

Les pixels de raster générés par l'algorithme de convolution, c'est-à-dire par l'opération d'interpolation du modèle MRH, ne représentent pas un ruissellement mais un écoulement. L'écoulement peut être défini comme un ruissellement permanent, autrement dit comme la fraction résiduelle du ruissellement assurant l'état stable du système physique modélisé. La plupart des modèles hydrologiques représentent le ruissellement, en calculant de façon discrète

(i.e. pour chaque pixel du domaine et pour chaque pas de temps) l'équilibre massique de l'eau, en distinguant généralement trois composantes : l'évapotranspiration, l'infiltration et le ruissellement (WISSER et al. 2010), qui s'équilibrent à chaque pas de temps suite à une injection initiale (e.g. une pluie). Lorsque l'hystérèse liée à l'injection initiale cesse (e.g. le temps de concentration lié à la pluie est échu), si un ruissellement a toujours lieu sur le pixel, comme le temps est alors statique, il s'agit d'un ruissellement stable, appelé écoulement (TARBOTON et al. 1991).

Nous aborderons dans la remarque (3) qui suit les problèmes que pose le transfert discret de l'écoulement d'un pixel à un autre, qui n'est évidemment pas quelque chose de statique. Néanmoins, ces deux notions, ruissellement et écoulement, peuvent assez simplement être distinguées selon ce critère dynamique/statique : l'écoulement est le résultat d'un ruissellement en excès une fois l'impulsion initiale passée. Le modèle MRH ne représente pas le ruissellement mais le transfert, ou convolution, de l'écoulement le long d'un réseau hydrographique. C'est pour cette raison que nous n'effectuons pas de balance hydrologique. Nous considérons que l'état stable est déjà atteint et étudions seulement le transfert de la masse de façon statique. Ceci appelle une remarque supplémentaire.

Seconde remarque : modéliser la masse *versus* modéliser une grandeur hydrologique

Cette remarque est sans doute la plus importante car elle explique une part de l'erreur d'estimation des débits du modèle MRH. La masse est une grandeur parfaitement statique puisqu'elle n'est pas définie par rapport au temps, il s'agit d'une quantité d'éléments (e.g. des mols de molécules H_2O). Pour autant, l'écoulement dans un cours d'eau n'est jamais exprimé en masse, mais selon différentes grandeurs hydrologiques, dont certaines sont définies par rapport au temps, comme le débit, généralement utilisée pour quantifier l'écoulement, et d'autres sont indépendantes du temps, comme la hauteur d'eau, généralement utilisée pour quantifier le ruissellement. Pour ajouter en complexité, certaines publications calculent le ruissellement à l'exutoire d'un bassin versant à partir du débit à l'exutoire, en rapportant ce débit à la superficie du bassin versant (BJØRN et al. 2020b). Lorsque l'on étudie la propagation de l'écoulement, il peut y avoir conservation de la masse mais pas conservation de la grandeur hydrologique retenue, par exemple le débit. Mathématiquement, si i désigne un pixel de réseau hydrographique et $i + 1$ le pixel contigu à l'aval, alors l'écoulement calculé en $i + 1$ est la somme de son propre écoulement (état stable de $i + 1$) et de l'écoulement hérité de l'amont (état stable de i). En notant m est la

masse et q le débit, le problème peut alors se poser ainsi :

$$\begin{cases} m_{i+1} = m_{i+1} + m_i \\ q_{i+1} \neq q_{i+1} + q_i \end{cases}$$

La définition élémentaire du débit est une section d'écoulement s [m^2] multipliée par une vitesse d'écoulement v [$m \cdot s^{-1}$], i.e. $\forall i, q_i = s_i \times v_i$ et q_i est alors en [$m^3 \cdot s^{-1}$]. Pour que le débit soit préservé linéairement entre les pixels i et $i + 1$, il faut que s et v demeurent constantes. Ceci implique, toutes choses égales par ailleurs (i.e. l'évapotranspiration et l'infiltration spécifique des pixels i et $i + 1$ sont identiques) :

$$(q_{i+1} = q_{i+1} + q_i) \Leftrightarrow (s_i \times v_i = s_{i+1} \times v_{i+1})$$

Le modèle MRH repose sur cette hypothèse de constance des conditions morphologiques : l'écoulement se propage sur des pixels morphologiquement identiques. La vitesse d'écoulement en régime non turbulent peut s'obtenir à partir des équations de Chézy et Manning :

$$\begin{cases} \text{Chézy eq.} & v = C\sqrt{Ra} \\ \text{Manning eq.} & C = \frac{1}{n}R^{1/6} \end{cases}$$

Où R est le rayon hydraulique (e.g. la hauteur d'eau) typiquement en [m], a est la pente vers l'aval typiquement en [$m \cdot m^{-1}$] et C est le coefficient de Manning en [$m^{1/2} \cdot s^{-1}$] avec $1/n$ qui désigne la rugosité des parois. À l'aide de ces équations, nous pouvons isoler les conditions (ou variables) qui réalisent l'équivalence que nous avons posée plus haut : pour que le débit soit linéairement propagé, il faut que la section d'écoulement s , la hauteur d'eau R , la pente a et la rugosité du lit $1/n$ soient constantes entre les pixels i et $i + 1$.

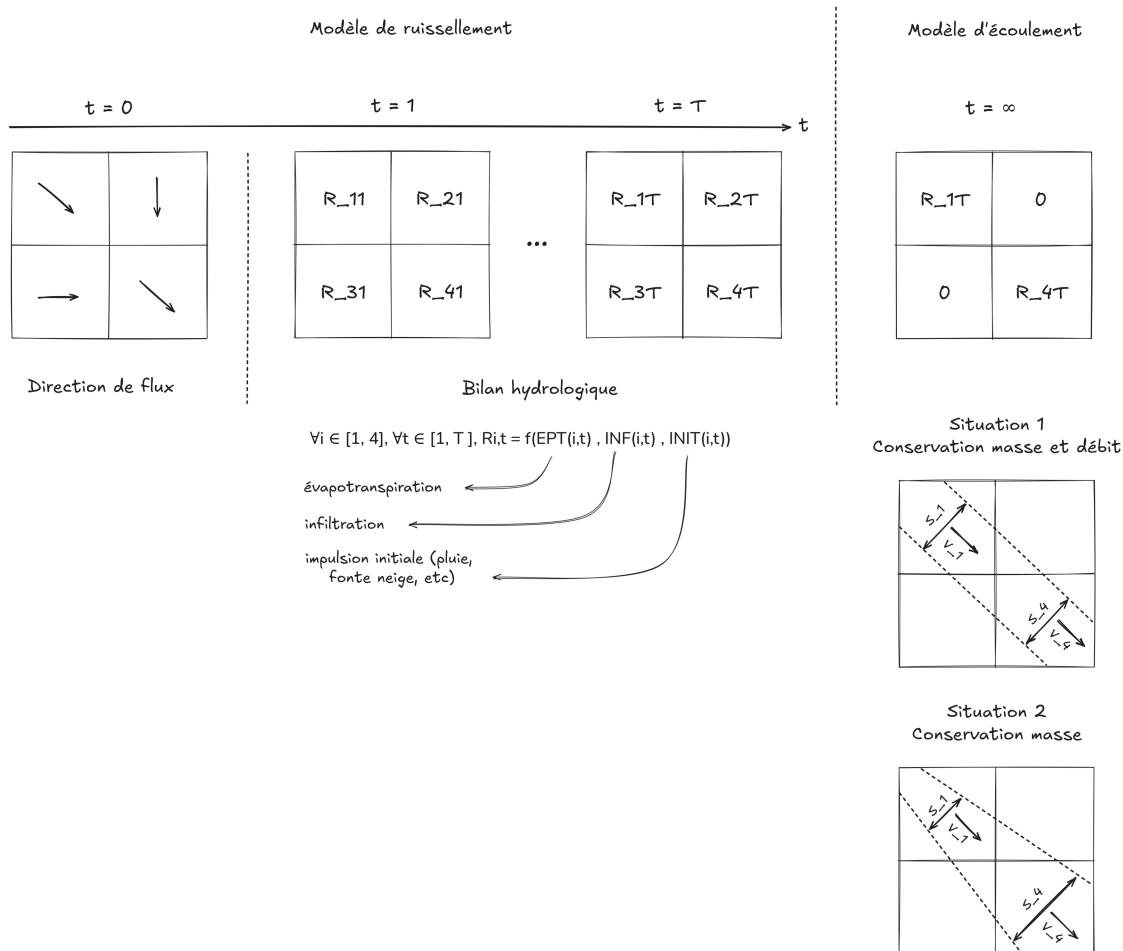


Figure A.1 – Représentation schématique du lien entre un modèle de ruissellement et un modèle d'écoulement

2 Annexes au chapitre 4

2.1 Annexe à la section 4.2

HydroTrendsFrance : preuve de concept

HydroTrendsFrance a été initialement développé à partir du territoire du Sud-Loire (42), qui correspond au regroupement de quatre EPCI, mais a été testé sur deux autres territoires ne correspondant chacun qu'à un seul EPCI : la Métropole du Grand Lyon qui présente le point commun avec le Sud-Loire d'être un territoire enclavé sans accès maritime et la communauté d'agglomération du Cotentin (autour de Cherbourg-en-Cotentin) qui est un territoire littoral.

L'un des enjeux qui a motivé la structuration d'HydroTrendsFrance sous sa forme actuelle était de rendre accessible les données hydrométriques de débits et de variabilité à des collectivi-

tés territoriales dont la politique environnementale de l'eau n'est pas la compétence, avec l'idée que la connaissance quantitative des milieux est déjà une première étape vers des indicateurs de type "limites planétaires". Ce premier pas vers la territorialisation du cadre, l'accès aux données et à la connaissance des processus hydrologiques territoriaux, a justifié de nombreux choix de développement logiciel : la structure globale des répertoires de fichier, la nomenclature des fichiers adaptables à l'emprise et à la chronique souhaitées, le paramétrage souple des modules pour exécuter différentes fonctionnalités d'HydroTrendsFrance le plus indépendamment et rapidement possible. Pour valider la possibilité de généraliser l'utilisation d'HydroTrendsFrance à d'autres territoires que le Sud-Loire et d'autres chroniques que la période 2000-2020, nous avons souhaité vérifier :

- l'interopérabilité : l'architecture des scripts doit être indépendante de l'environnement d'exécution en local sur la machine que nous avons utilisée pour le développement
- les exceptions : les exceptions levées par les données disponibles (ou, justement, non disponibles) sur d'autres territoires ne doit pas obérer la bonne exécution des fonctionnalités d'HydroTrendsFrance.
- la lisibilité : le paramétrage du fichier `main.py` doit être suffisamment clair et compréhensible pour une utilisation par un-e autre développeur-euse, et de même pour les fichiers générés en sortie de scripts.

L'interopérabilité

Grâce au travail attentif de Duli Raschid, élève-ingénieur de l'école des Mines de Saint-Etienne et en stage à la Métropole du Grand Lyon entre mars et août 2024 sur le sujet de l'application des limites planétaires au territoire du Grand Lyon, l'interopérabilité des scripts a pu être testée, non seulement entre deux environnements de travail Python différents, mais aussi entre deux systèmes d'exploitation différents (système Unix avec OS Ubuntu et système Windows avec OS Windows 10). Son travail a notamment permis d'harmoniser l'écriture des chemins d'accès aux fichiers.

Son travail a permis de mettre en œuvre une méthode de calcul pour la limite sur l'eau formulée dans la version 2015 du cadre, en couplant les données de mesure Vigicrues spatialisées à l'échelle des bassins versants (HydroTrendsFrance) et les données spatialisées de la BNPE pour les prélèvements en eau enregistrés sur ces mêmes bassins versants.

Les exceptions

Cet aspect est de loin le plus difficile à gérer car à chaque nouveau territoire et chronique appelés, la réponse du serveur hubeau à la requête varie et en conséquences les tableaux de données à traiter. L’affichage des graphiques et le calcul des ratios est très sensible aux valeurs `null` qui conduisent à des divisions par 0 et donc à des erreurs. Malgré un grand nombre d’appels conditionnels ou de filtrage des tableaux sur les données non-`null`, des exceptions peuvent malgré tout se produire.

Une autre source d’exception importante réside dans les formats de fichiers. Il n’est ainsi pas encore possible de travailler à partir de fichiers vectoriels autres qu’au format `.gpkg` bien que les bibliothèques `gdal` et `geopandas` soient compatibles avec d’autres formats vectoriels. Pour les fichiers shapefile par exemple, certains scripts travaillent avec des noms de colonne qui dépassent la longueur maximum autorisée par le format shapefile, ce qui pose de nombreux problèmes en chaîne lorsque les fonctions font référence à des noms de colonnes tronqués par le format shapefile.

Du même type, des exceptions liées au système de projection peuvent apparaître et il n’est pas encore possible de travailler avec un SRC autre que l’EPSG 4326 bien que cela soit présenter comme un paramètre.

La lisibilité

La structure des scripts vise à réduire leur utilisation au paramétrage initial dont tous les paramètres sont listés en tête du script `main.py`. Un tableur au format `.ods` est fourni sur le dépôt github d’`HydroTrendsFrance` afin de documenter les fichiers de sortie générés par les scripts, en distinguant les produits finis bruts (tableaux de données élaborées disponibles dans les formats `.csv` et `.gpkg`), les produits finis mis en page (sous forme de graphes ou de cartes, au format `.png`), les produits intermédiaires (tableaux de données brutes ou peu élaborées) et les fichiers temporaires (ayant servi dans différents traitements et pouvant être utilisés à des fins de raffinement).

Quelques remarques sur la comparaison des résultats pour différents territoires

Les cartes suivantes représentent l’indicateur d’anomalie de variabilité que nous avons

détaillé plus haut, pour les trois territoires du Sud-Loire, du Grand Lyon et du Cotentin.

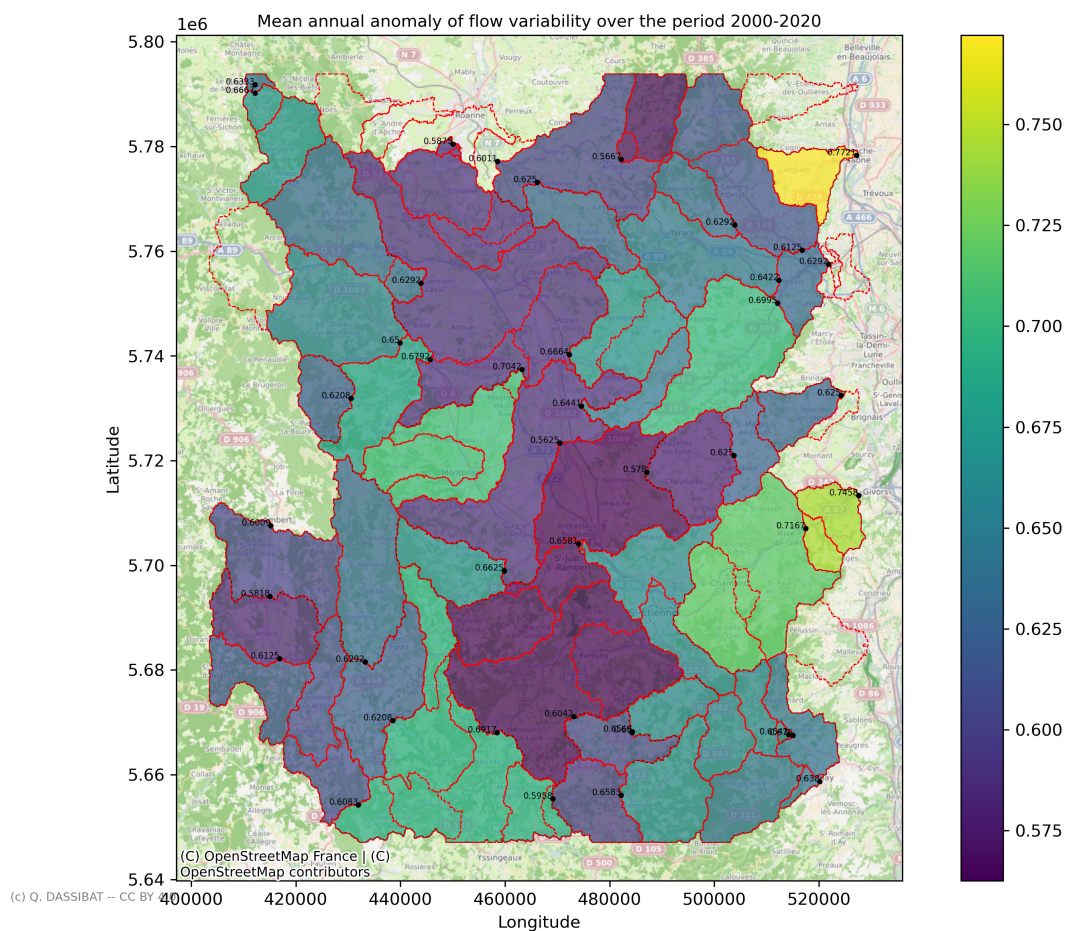


Figure A.2 – Anomalie de variabilité sur le territoire du SCoT du Sud-Loire

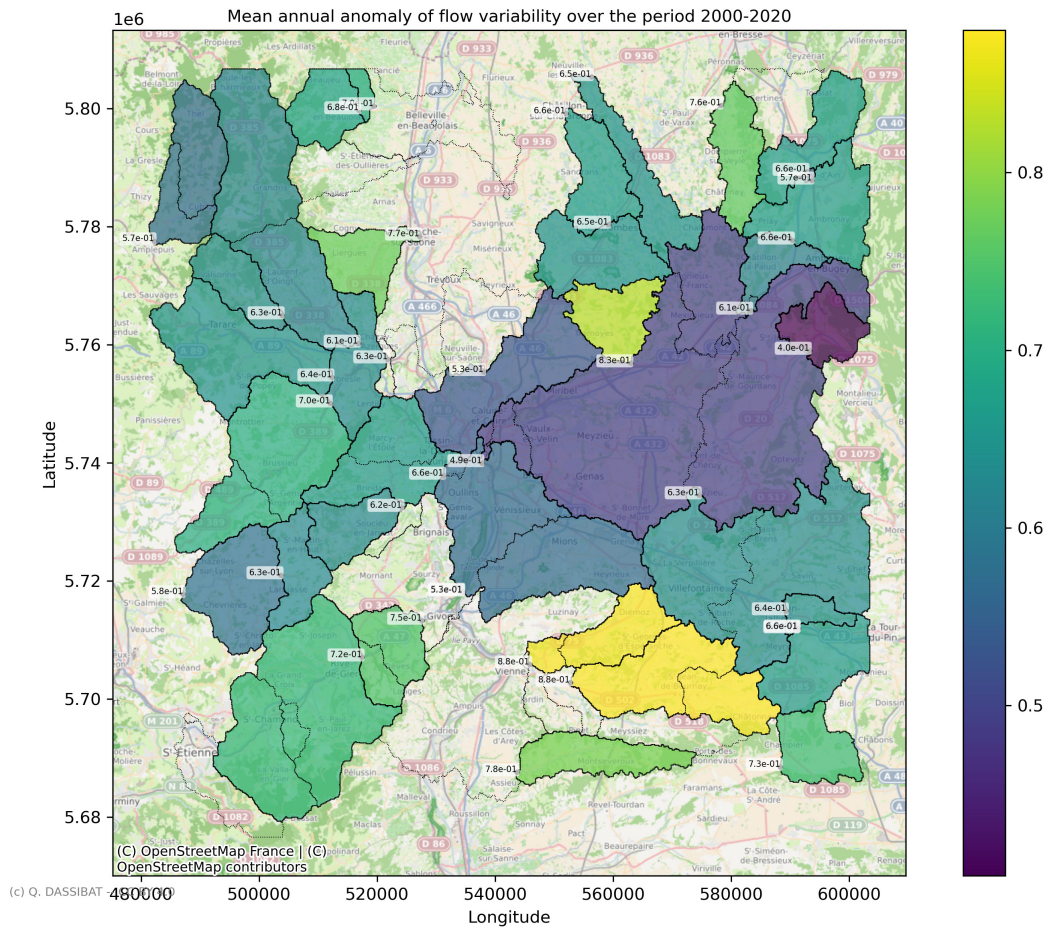


Figure A.3 – Anomalie de variabilité sur le territoire de la Métropole du Grand Lyon

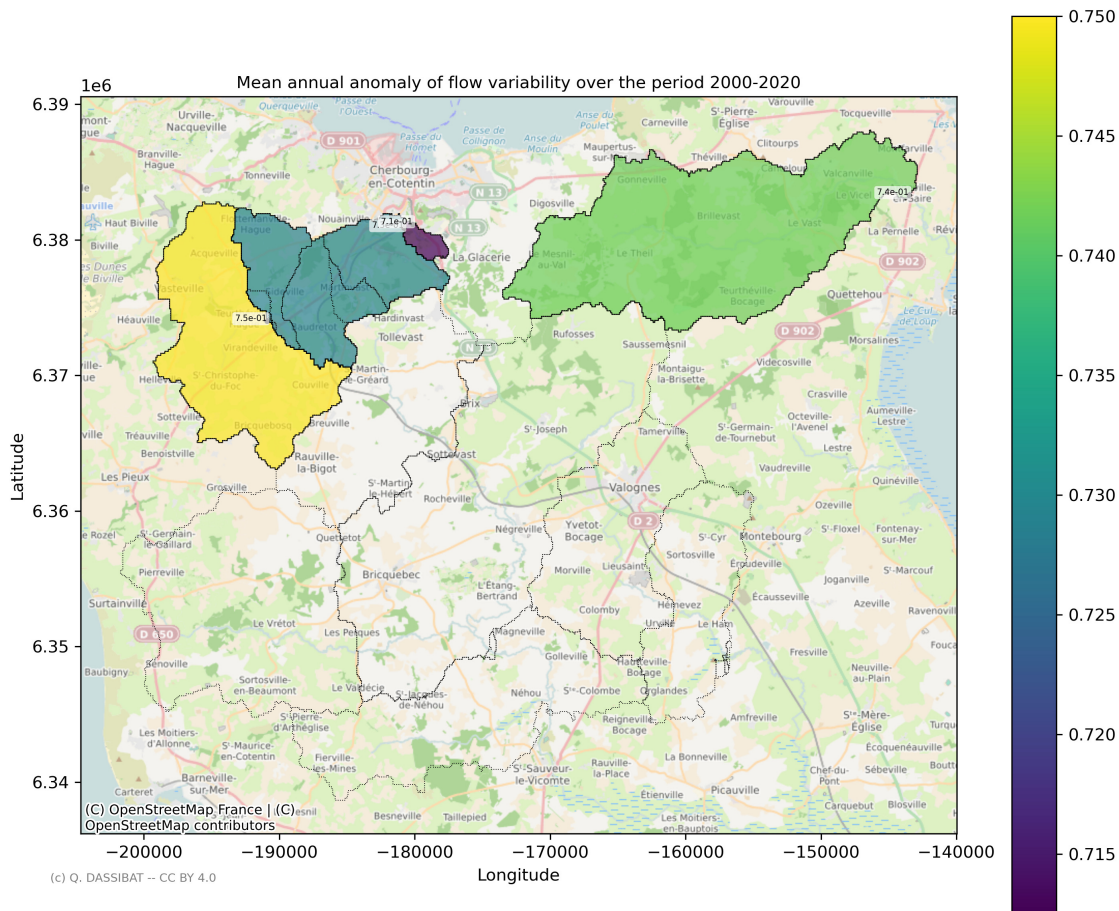


Figure A.4 – Anomalie de variabilité sur le territoire de la communauté d’agglomération du Cotentin

On remarque qu’avec un seuil maximum d’anomalie de variabilité à $\pm 20\%$, presque aucun bassin versant d’aucun des trois territoires ne respecte sa limite. Ce qui pose deux questions : ou bien effectivement l’intégrité des cours d’eau est-elle menacée sur ces trois territoires, ou bien le seuil de $\pm 20\%$ est-il trop strict lorsque le calcul est effectué sur la base de données de mesure ?

2.2 Annexe à la section 5.1

Cette annexe décrit le fonctionnement des algorithmes du MRH mis en œuvre dans *relio* et leur agencement dans les différents scripts constitutifs du logiciel.

L’annexe apporte également une description des algorithmes permettant d’effectuer la paramétrisation du modèle MRH : l’initialisation aux bornes du raster d’accumulation de flux pour le paramètre β_p et le filtrage du raster des exutoires pour le paramètre α_b .

2.2.1 Algorithmes pour les traitements initiaux

Ce sont les rasters (1) et (2) qui nous intéressent ici. Le script `MODELinputs.py` effectue les traitements initiaux sur ces rasters et définit l'ensemble des rasters compris dans \mathcal{D}_{src} . Deux traitements sont effectués : les multi-bandes sont éclatées au profit de 12 rasters à 1 bande (ce choix n'est pas technique mais permet de s'y retrouver plus simplement dans la suite des traitements), puis ces 12 rasters sont coupés à la dimension de la zone d'étude (afin d'alléger les temps de calcul). Remarquons que l'opération de découpage est effectuée par la fonction `relio.clip` qui implémente la commande `gdalwarp` dans laquelle l'utilisation d'une méthode de rééchantillonnage est nécessaire. En effet, l'algorithme impose l'emprise spatiale du raster en sortie à celle de la zone d'étude (l'emprise spatiale étant définie comme les deux paires de coordonnées au sud-ouest et au nord-est : `<xmin,ymin>` et `<xmax,ymax>`). En forçant ce paramètre, il devient nécessaire d'ajuster la taille du pixel dans le raster en sortie, qui ne peut pas être conservée à sa taille en entrée. Le rééchantillonnage se fait alors par plus proche voisin (paramètre `-r near`) qui est une méthode embarquée dans la commande `gdalwarp`. A l'issue de ces opérations, on a par exemple le raster `WBM_TerraClimate_OUTPUTQ_DIST_mLTM_-2000-2020_band1_clip.tif` qui correspond au raster de débits perturbés (DIST) pour le mois de janvier (`band1`) et coupé aux dimensions de la zone d'étude (`clip`).

Enfin, le script `GAUGINGdata.py` permet de préparer les données de mesure requêtes sur l'emprise territoriale appelée. Les débits moyens journaliers retournés par l'API `hubeau.france.fr` sont convertis en débits mensuels moyens interannuels, conformément à la métrique utilisée dans le modèles des frontières du système Terre. Ces traitements ont déjà été largement documentés dans le chapitre 4 de ce manuscrit.

La description des algorithmes du MHE a déjà été effectuée dans le chapitre (3) pour la génération des images suivantes :

- raster de direction de flux
- raster d'accumulation de flux
- raster des bassins versants
- raster des données d'observatin (rasterisées)

`relio` apporte un raffinement dans la génération du raster d'accumulation de flux et partant, dans l'identification des exutoires "significatifs" de la zone d'étude.

Le raster de direction de flux

Cette algorithmme ayant déjà décrit, nous donnons simplement ici quelque informations concernant son agacement dans l'architecture des scripts de `relio`. Cet algorithmme convertit le MNT (de résolution 30×30m en un raster de direction de flux. Cette conversion, réalisée dans le script `INITIALIZEaccuflux.py`, fonctionne en deux étapes successives : le MNT est d'abord converti au format PCRaster à l'aide de la fonction `relio.convert_to_pcraster` qui implémente la fonction `gdalwarp`, puis transformé en raster de direction de flux avec la fonction `relio.create_flowdirection` qui implémente la fonction `pcraster.lddcreate`.

Raffinement du raster d'accumulation de flux

Nous allons d'abord décrire l'opération de base qui consiste à créer le raster d'accumulation de flux. Nous décrirons ensuite les raffinements apportés à ce raster, à savoir la prise en compte des conditions aux bornes de la zone d'étude et le filtrage à un seuil de charge donné.

Pour la mise en charge initiale, nous avons créé un raster dont tous les pixels valent 1, pour simuler une pluie unitaire uniforme. En sortie de `relio.create_accuflux`, tous les pixels valent donc au moins 1. L'enjeu est alors de savoir à partir de quel niveau d'accumulation (ou charge cumulée, ou tout simplement charge), on peut considérer un pixel comme formant partie d'un cours d'eau actif. Ce seuil de charge minimum est désignée par le paramètre α_a . Le script `HYDROGRAPHICinputs.py` permet d'itérer sur la valeur de α_a en modifiant le paramètre `epsilonAccuflux` et en récupérant itérativement la couche `DEMx_SRTM30_wgs84_pcraster_-accuflux_sup{str(epsilonAccuflux)}.tif` générée en sortie.

Nous avons commencé par décrire ce raffinement du raster d'accumulation de flux en premier car il s'agit du plus évident, ou intuitif. En pratique, il est plus simple de commencer numériquement par le raffinement à suivre : la prise en compte des conditions aux bornes. Cette étape est indispensable pour le bon fonctionnement de la méthode `relio` dont le principe est d'allouer un débit connu en un point du réseau au reste du réseau par rapport de proportionnalité entre la charge théorique en ce point connu et tout autre point du réseau. Il faut donc que les charges hydrauliques relatives du réseau rendent compte au plus proche de la réalité hydrologique de la zone d'étude. Un facteur particulièrement influant pour la prise en

compte réaliste des charges relatives du réseau réside dans les conditions aux bornes de la zone d'étude. En effet, la zone d'étude étant un système ouvert et arbitrairement positionnée (i.e. positionnée par un découpage administratif), elle ne peut nécessairement pas correspondre au découpage naturel des bassins versants. De là vient que les bassins versants de la zone d'étude sont des sous-ensembles des bassins versants naturels, en particulier pour ceux qui se situent en bordure de la zone d'étude. Deux situations peuvent alors apparaître pour de tels bassins versants "coupés" : soit l'exutoire se situe à l'intérieur de la zone d'étude et la tête de bassin se situe en dehors, soit à l'inverse l'exutoire se situe à l'intérieur de la zone d'étude et la tête de bassin à l'extérieur, soit une situation intermédiaire entre ces deux situations. La situation la moins favorable pour notre estimation des charges relatives du réseau est lorsque la tête de bassin se situe en dehors de la zone d'étude. En effet, le raster de charge initiale agit comme si tout pixel était une tête de bassin. Or certains pixels de la zone d'étude sont susceptibles d'hériter d'une charge à l'amont. La situation où l'exutoire naturel ne se trouve pas dans la zone d'étude n'est pas dérangeante pour notre étude, car l'exutoire "fictif" qui se trouvera en bordure de zone d'étude représentera bien la charge amont accumulée jusqu'à lui.

Déterminer la charge héritée de l'amont n'est numériquement pas chose facile. Une méthode brutale voudrait que l'on élargisse la zone d'étude jusqu'à ce qu'elle englobe les bassins versants naturels. Pour notre cas d'étude en Sud-Loire, la zone d'étude que nous avons fixé comparant le bassin versant du Rhône, de la Saône, de l'Allier et de la Loire. On imagine bien dès lors la très grande de taille des jeux de données à manipuler pour une résolution de 30×30m, qui ne peuvent être traités que sur serveur de calcul. En outre, plus la zone d'étude est élargie, moins le raster de charge unitaire et uniforme devient pertinent (réaliste) car plus l'hypothèse d'uniformité s'affaiblit. Une méthode plus fine, et que nous avons adoptée, consiste à déterminer itérativement une zone d'étude (ou enveloppe) optimale, c'est-à-dire à élargir progressivement l'enveloppe à partir d'une taille initiale jusqu'à ce que le nombre de cours d'eau qui intersectent de l'amont la zone d'étude soit jugé raisonnable pour envisager une correction manuelle de la charge initiale des pixels auxquels ils se branchent dans le réseau.

Le script `SETenvelope.py` est dédié à cette opération initiale par laquelle l'enveloppe optimale est déterminée. L'idée est de fournir une enveloppe en entrée (ou emprise) et d'effectuer une différence symétrique avec la couche vectorielle du réseau hydrographique de la BD Topage de géométrie `Linestring` à l'aide de la librairie `geopandas`. On obtient ainsi la couche vectorielle des tronçons hydrographiques qui intersectent la zone d'étude, mais seulement

leur portion amont ou aval. Le script calcule ensuite la longueur de ces tronçons amont ou aval (en mètres) avec la méthode `length` de `geopandas` puis génère la fonction de distribution des longueurs de tronçons et enfin propose d'exporter la couche vectorielle avec un seuil de coupure dans la fonction de distribution. Par exemple, si le seuil de coupure (qui correspond au paramètre `percentile` dans le script) est fixé à 90%, cela signifie que la couche vectorielle des tronçons amont ou aval à la zone d'étude ne comportera que tronçons qui, en cumulé, représentent 90% de la longueur totale de tous les tronçons. Nous avons en effet observé que la distribution des longueurs suit approximativement une loi log : un petit nombre de tronçons sont responsable d'un grand pourcentage de la longueur amont et aval. Ne conserver que ces quelques tronçons permet ainsi de s'économiser la tâche manuelle et fastidieuse de correction des charges initiales aux pixels d'embranchement.

Pourquoi s'intéresser à la longueur des tronçons amont et aval à la zone d'étude pour déterminer les charges initiales aux bornes? L'idée est d'utiliser la longueur d'un tronçon comme un proxy de sa charge hydraulique. Si l'on divise en effet sa longueur (exprimée en mètres) par la résolution 30m des pixels, on obtient un "équivalent-nombre-de-pixels" qui est très proche de la façon dont fonctionne l'accumulation de flux, où l'initialisation de la charge à 1 et la convolution pixel-par-pixel revient également à dénombrer les pixels le long du réseau. Cette technique d'estimation du débit amont est toutefois à considérer comme une première tentative et demeure largement imparfaite. En effet, elle ne prend en compte que le tronçon principal qui arrive ou part du territoire, et pas ses affluents en dehors de la zone d'étude, ce qui est problématique pour le cas des cours d'eau importants arrivant sur la zone d'étude, dont la charge a nécessairement été augmentée de celles de ses affluents à l'amont du territoire. La méthode tend ainsi à sous-estimer les charges héritées de l'amont. Enfin, la méthode ne permet pas de déterminer le sens d'écoulement des tronçons : s'agit-il de tronçons amont ou aval à la zone d'étude? Seuls nous intéressent les cours d'eau amont (qui apportent une charge à la zone d'étude), et seule la connaissance du territoire par l'utilisateur permet de les identifier.

Nous pouvons en profiter pour préciser la formalisation mathématique du paramètre β_p , qui désigne précisément le proxy "équivalent-nombre-de-pixels" que nous venons de décrire.

β_p est en effet le vecteur des charges initiales aux bornes de la zone d'étude. Plus précisément, il contient les valeurs de charge des n pixels de \mathcal{D}_{dst} qui héritent d'une charge à l'amont, $\beta_p := (a_k)_{1 \leq k \leq n}$. Comme nous l'avons précisé, les a_k ne sont pas connus mais estimés par le proxy

"équivalent-nombre-de-pixels". On note l_k la longueur du tronçon à l'amont du pixel k et r la résolution d'un pixel de \mathcal{D}_{dst} . On calcule alors les a_k de la façon suivante :

$$\forall 1 \leq k \leq n, \quad \widehat{a}_k = \frac{l_k}{r}$$

On peut noter L la variable aléatoire des $(l_k)_{1 \leq k \leq n}$. Avec ces notations, l'indice p de β_p correspond alors au $p^{\text{ième}}$ quantile de la fonction de distribution de L , à savoir : $\mathcal{P}(L \leq \gamma) = p$ où γ (gamma) désigne la longueur amont cumulée telle que p soit réalisé. Si les \widehat{a}_k sont rangés par ordre croissant, numériquement on cherche le premier k tel que p soit réalisé.

$$k? \quad | \quad \mathcal{P} \left(L \leq \sum_{i=1}^k \widehat{a}_i \right) = p \quad \text{où} \quad \forall 1 \leq i \leq k, \quad \widehat{a}_i \leq \widehat{a}_{i+1}$$

Une fois la taille optimale (du moins satisfaisante) de l'enveloppe déterminée, il s'agit de modifier manuellement la valeur des pixels qui se situent à l'embranchement avec les tronçons amont dans le raster de charge initiale. Ces pixels se verront attribuer la valeur "équivalent-nombre-de-pixels" qui a été calculée précédemment, tandis que les pixels non concernés par un embranchement amont en bordure de zone d'étude conserveront une charge unitaire. Deux scripts sont dédiés à cette opération : `INITIALIZEaccuflux.py` qui doit être exécuter en premier et `INITIALIZEmaterial.py` qui est exécuter en second.

Le script `INITIALIZEaccuflux.py` permet de générer un raster d'accumulation de flux à partir d'une charge initiale unitaire et uniforme. Le script renvoie également une copie du raster d'accumulation de flux mais à deux bandes : une bande qui donne les valeurs de charge accumulée et une bande qui donne un identifiant de pixel (nombre entier allant de 1 à la dimension du raster). Ce raster à deux bandes est ensuite affiché dans une interface graphique SIG et superposé à la couche vectorielle des tronçons amont et aval générée précédemment. Il s'agit alors d'effectuer la correspondance entre les tronçons amont et leur pixel d'embranchement. Dans le script `INITIALIZEmaterial.py`, on vient instancier un vecteur d'initialisation aux bornes où l'on range pour chaque pixel d'embranchement, l'identifiant du pixel et la valeur "équivalent-nombre-de-pixels" correspondante dans la couche vectorielle. L'exécution de ce second script renvoie un nouveau raster de charge initiale paramétré aux bornes.

Enfin, nous revenons à la première étape que nous décrivions (qu'il était plus intuitif de présenter en premier bien qu'elle n'intervienne qu'en dernier dans les traitements). Le script `HYDROGRAPHICinputs.py` peut être exécuter car il s'occupe de générer le raster d'accumulation

de flux à partir du raster modifié de charge initiale. Il s'agit alors d'itérer sur les valeurs de α , comme nous décrivions plus haut, afin de déterminer visuellement la valeur la plus satisfaisante, α_a . Remarquons que le script `SETaccufluxThreshold.py` a pour but d'accompagner la détermination optimale de α_a en minimisant l'écart de densité entre les pixels non-NoData du raster d'accumulation de flux et les pixels d'une couche vectorielle du réseau hydrographique Topage préalablement converti en image raster.

Raffinement du raster des exutoires

Ce raster est un raster intermédiaire entre le raster d'accumulation de flux et le raster des bassins versants.

Un raffinement supplémentaire consiste à introduire un filtre qui ne retient que les pixels exutoires ayant accumulé une charge supérieure à un certain seuil. Ce seuil permet de réduire considérablement les temps de calculs. En effet, de nombreux cours d'eau de longueur très faible sortent de la zone d'étude et sont à ce titre considérés comme des exutoires. Ceci conduit à multiplier le nombre des bassins versants, notamment par de tous petits bassins en bordure d'image, ce qui ne pose en soi pas de problème pour les traitements suivants mais conduit à multiplier le nombre des traitements car la suite des opérations consistera à traiter chaque bassin versant au cas par cas. Or le fonctionnement des traitements par images raster implique de toujours garder la même dimension d'image (pour que les superpositions spatiales soient exactes) : même pour de tous petits bassins versants, la taille totale de la zone d'étude est préservée, ce qui conduit à traiter des images quasiment vides de données alors que le temps de traitement d'une image dépend du nombre de pixels à lire et pas de l'information portée par les pixels. Ainsi donc un seuil d'accumulation α_b est fixé itérativement.

L'opération d'extraction des pixels exutoires est effectuée dans le script `OUTLETSlocation.py` et commence par extraire du raster \mathcal{D}_{src} les pixels en bordure d'image puis, pour chacun de ces pixels, les pixels superposés de \mathcal{D}_{dst} en bordure sont extraits. Cette deuxième opération est effectuée à l'aide de la fonction `rel.io.extract_edge_outlets` qui, malgré son nom, élargit la recherche des exutoires au voisinage des pixels en bordure car il existe un bruit numérique par lequel certains pixels sont bien encodés comme des exutoires dans le raster de direction de flux sans se situer exactement sur la bordure de l'image, mais à quelques pixels près. Le raster d'accumulation de flux filtré est ensuite également coupé à la dimension de chaque pixel de \mathcal{D}_{src}

en bordure (afin de réduire les temps de calcul) puis la fonction `relio.significant_outlets` applique le filtre α_b sur les exutoires identifiés.

Précisons davantage le fonctionnement de `OUTLETSlocation.py` qui effectue une série d'opérations décisives dans l'ensemble de la méthode `relio`. C'est en effet à ce stade que sont distingués, pour chaque pixel en bordure de \mathcal{D}_{src} , les pixels sous-jacents de \mathcal{D}_{dst} correspondant aux exutoires finaux et aux exutoires de transfère. La fonction `relio.significant_outlets` opère à l'intérieur d'un pixel de \mathcal{D}_{src} , qu'un peut noter j . Elle passe en revue tous les pixels i sous-jacent de \mathcal{D}_{dst} (raster de direction de flux) situés en bordure. Pour chaque arrête de j , la fonction cherche les directions d'écoulement des pixels i situés l'arrête telles qu'elles dirigent l'écoulement en dehors de j . Par exemple, pour l'arrête Nord de j , seules les directions d'écoulement Nord, Nord-Ouest, Nord-Est conduisent le flux à "s'échapper" de j . Dans l'encodage D8 du raster de direction de flux, ces trois directions données en exemple correspondent aux valeurs 7, 8 et 9. A ces valeurs, nous ajoutons la valeur 5 qui correspond à l'encodage des dépression effectué par la fonction `pcraster.lddcreate` et c'est pour cette valeur que nous élargissons la recherche au voisinage des pixels i en bordure j (au voisinage "interne", c'est-à-dire toujours en restant bien à l'intérieur de j). La fonction `relio.significant_outlets` opère de façon analogue pour les directions Sud, Est et Ouest d'écoulement. Elle crée un raster en sortie dont la dimension est égale à taille du pixel j et la résolution est égale à la taille des pixels i . Pour des questions pratiques, nous exportons ce raster comme une couche vectorielle de géométrie `Point`, ce qui permettra dans un second temps, une fois tous les pixels j en bordure traités, d'invoquer la fonction `pandas.concat` pour concaténer les couches de sortie et les traiter plus facilement comme objets `DataFrames`. A l'issue de cette étape, tous les pixels (désormais "entités" de couche vectorielle) sont passés en revue et labellisés `True` ou `False` selon qu'ils correspondent respectivement à des exutoires finaux ou de transfère. Plus précisément, si leur code de direction vaut 5, ils sont considérés comme des exutoires finaux, sinon comme des exutoires de transfère. La valeur correspondant à chacun de ces pixels dans le raster d'accumulation de flux filtré est ensuite récupérée et ajoutée au `DataFrame` (par jointure attributaire d'identifiant de pixel, ce que permet le travail sur des couches vectorielles plutôt que sur des images raster, et la fonction `relio.extract_cellValues` assure la conversion d'un raster vers une couche vectorielle).

Remarquons enfin que le raster de \mathcal{D}_{src} à partir duquel sont extraits les pixels en bordure a arbitrairement été choisi comme le raster des débits perturbés pour le mois de janvier (`WBM_TerraClimate_OUTPUTQ_DIST_mLTM_2000-2020_band1_clip.tif`). Ce choix n'a aucune

importance puisque les 24 rasters de \mathcal{D}_{src} partagent les mêmes propriétés spatiales (résolution, dimension, projection) et qu'on ne s'intéresse ici pas encore aux valeurs des pixels mais simplement à leurs propriétés spatiales.

Le raster des bassins versants

La délimitation des bassins versants est essentielle dans la méthode `relio` car celle-ci attribue un débit connu à l'exutoire le long du chemin hydrographique au prorata de la charge hydraulique théorique relative de tout point du réseau par rapport à l'exutoire. Or on ne peut comparer des charges hydrauliques qu'entre deux points (plutôt pixels) d'un même bassin versant car la charge à l'exutoire est fonction de l'accumulation des charges à l'amont et est indépendante des charges observées sur d'autres bassins versants.

Le script `SUBCATCHMENTSdelineation.py` effectue la délimitation des bassins versants. Nous utilisons la fonction `pcraster.catchment` qui opère à partir d'un raster de direction de flux et d'une liste d'exutoires avec leurs coordonnées $\langle x, y \rangle$ contenus dans un fichier `.txt`. La fonction `relio.create_subcatchments` implémente cette fonction `PCRaster` et se charge automatiquement de générer le fichier `.txt` à partir d'un `DataFrame`. Ce `DataFrame` est précisément issu du traitement précédent, où les exutoires ont été identifiés et distingués en exutoires finaux et de transfère. Pour la délimitation des bassins versants, ce sont seulement les exutoires finaux de la zone d'étude qui nous intéressent.

2.2.2 Algorithme de dérasterisation des pixels de débit modélisé

Il s'agit de l'un des trois algorithmes centraux de la méthode `relio`. Son but est de convertir les pixels d'un raster de faible résolution (typiquement $7000\text{m} \times 7000\text{m}$) vers une grille raster à forte résolution (typiquement $30\text{m} \times 30\text{m}$). Pour cela, nous avons déjà décrit dans le chapitre (3) le principe de fonctionnement en deux temps : dans l'espace des positions et dans l'espace des valeurs. Dans le premier espace, le passage du raster d'origine à faible résolution vers le raster de destination à forte résolution consiste à localiser les pixels du domaine de destination qui constituent les exutoires sous-jacents des pixels d'origine en bordure de zone d'étude. Cette première étape est en réalité effectuée par le script `OUTLETSlocation.py` au moment où le raster des exutoires est généré. La description de ce script aurait pu être effectuée ici car elle est partie intégrante de l'algorithme de dérasterisation. Cependant, la localisation des exutoires

conditionne la délimitation des bassins versants, et trouvait donc aussi sa place dans la section des traitements initiaux.

Pour ce qui est de l'espace des valeurs, le passage de \mathcal{D}_{src} à \mathcal{D}_{dst} consiste à attribuer (ou allouer) la valeur des pixels de \mathcal{D}_{src} vers les pixels exutoires de \mathcal{D}_{dst} au prorata de leur charge hydraulique théorique relative. On rappelle que la méthode de rééchantillonnage **rel**io consiste à "éclater" le raster initial à faible résolution vers les exutoires sous-jacents dans le raster à forte résolution, puis à convoluer le débit des exutoires le long de leur réseau hydrographique respectif. Il faut donc d'abord connaître la valeur du débit qui est attribué aux pixels exutoires sous-jacents. Deux cas de figure ont déjà été distingués et sont rappelés ici : pour tout pixel en bordure du raster d'origine, soit l'écoulement sort de la zone d'étude en sortant du pixel, soit l'écoulement rejoint un autre pixel en sortant du pixel. C'est la première raison pour laquelle la valeur d'un pixel dans le raster d'origine ne peut pas être attribuée directement à un unique pixel sous-jacent dans le raster de destination : il faut retrancher à cette valeur la part qui est transférée à un autre pixel (pour éviter les doubles comptabilités). La deuxième raison qui empêche cette attribution directe et unique est que, même une fois décomptée la part du débit qui est transférée à un pixel adjacent (par les exutoires sous-jacents dits "de transfère"), les points de sortie de la zone d'étude sous-jacents à un pixel du raster d'origine sont nécessairement plusieurs. Il faut donc également répartir le débit sortant sur ces différents pixels sous-jacents (exutoires sous-jacents dits "finaux").

L'allocation des valeurs des pixels du raster d'origine est réalisée en deux étapes :

1. décompte de la fraction "transférée" du débit dans le débit total du pixel
2. répartition du reste du débit sur les différents exutoires "finaux" sous-jacents

Ces deux étapes reposent sur un même principe d'allocation, celui d'un prorata à la charge hydraulique théorique relative. C'est le script `OUTLETSdischarge.py` qui les met en oeuvre. Celui-ci commence par associer aux pixels du raster des exutoires (importé sous la forme vectorielle d'un `DataFrame`) la valeur des pixels correspondant dans le raster d'accumulation de flux. Le script fonctionne ensuite pixel d'origine par pixel d'origine. Pour chacun, l'étape (1) est d'abord effectuée en ne retenant que les pixels du raster des exutoires correspondant à l'identifiant du pixel (cet identifiant a été ajouté pendant la génération initiale du raster des exutoires). Pour les pixels retenus (autrement dit ceux qui sont superposés au pixel du raster d'origine), la somme des valeurs d'accumulation de flux sur ces pixels est effectuée : il s'agit de

l'accumulation totale aux exutoires du pixel d'origine. Un deuxième filtre est ensuite appliqué sur le type d'exutoire : ne sont retenus que les pixels qui ne sont pas labellisés comme exutoires finaux. La somme des valeurs d'accumulation de flux sur ces pixels est alors effectuée : il s'agit de l'accumulation totale aux exutoire de transfère du pixel d'origine. Le ratio entre ces deux sommes (accumulation aux exutoires de transfère sur accumulation à tous les exutoires) donne la part du débit du pixel d'origine qui est transférée aux pixels adjacents.

Pour l'étape (2), on ne considère plus alors la valeur des pixels d'origine mais leur valeur décomptée des débits de transfère calculés à l'étape (1). On appelle cette nouvelle valeur de débit "débit exutoire". Il ne reste alors plus qu'à répartir ce débit exutoire sur les pixels exutoires finaux sous-jacents. Pixel d'origine par pixel d'origine là encore, on filtre le `DataFrame` du raster des exutoires sur les exutoires "finaux" puis on calcule le ratio pour chacun d'eau entre leur valeur d'accumulation de flux et la somme de ces valeurs. Pour chacun de ces pixels sous-jacents, on multiplie finalement leur ratio de charge théorique relative par le débit exutoire du pixel d'origine superposé.

Remarquons enfin que cet algorithme fonctionne à partir d'un unique raster de \mathcal{D}_{src} , qui a arbitrairement été choisi comme le raster des débits perturbés pour le mois de janvier (`WBM_TerraClimate_OUTPUTQ_DIST_mLTM_2000-2020_band1_clip.tif`). Ce choix n'a aucune importance puisque les 24 rasters de \mathcal{D}_{src} partagent les mêmes propriétés spatiales (résolution, dimension, projection) et qu'on ne s'intéresse ici pas encore aux valeurs des pixels mais simplement à leurs propriétés spatiales.

2.2.3 Algorithme de convolution linéaire du débit

Cet algorithme succède logiquement à l'algorithme de dérasterisation dont nous décrivions la mise en oeuvre numérique ci-dessus. En effet, une fois que les pixels du raster d'origine ont été "éclatés" sur la grille du raster de destination, il reste à répercuter la valeur le long du réseau hydrographique. `relio` propose d'effectuer cette remontée du réseau hydrographique, pixel par pixel, selon un principe de convolution linéaire. L'hypothèse fondatrice est que tout pixel transfère vers l'aval la charge hydraulique qu'il reçoit de l'amont et sa charge d'initialisation. Le pixel à l'exutoire du réseau a donc accumulé la charge de tous les pixels à l'amont. Si l'on initialise le réseau avec une charge hydraulique unitaire (ou théorique), on peut donc exprimer la part (ou contribution) de chaque pixel dans la décharge du réseau à l'exutoire. C'est ce que nous avons appelé charge hydraulique théorique relative. Cette charge relative n'a de sens

qu'entre les pixels d'un même bassin versant.

Le script `ALLOCATEmodelDischarge.py` effectue une telle convolution. Il commence par couper le raster d'accumulation de flux (filtré au seuil α_a) à l'emprise de chaque bassin versant contenu dans le raster des bassins versants, puis exporte ce raster coupé en couche vectorielle de géométrie `Point`. Le raster des pixels dérasterisés (étape précédente) est également chargé en format couche vectorielle de géométrie `Point`, et coupé à l'emprise de chaque bassin versant. Nous avons évoqué dans la section précédente le fait qu'un même pixel du raster d'origine pouvait avoir plusieurs exutoires "finaux" sous-jacents. L'étape de délimitation des bassins versants permet mécaniquement de ne plus être confronté à ce problème car pour tout exutoire final, un bassin versant est créé. Lorsque le traitement se fait bassin versant par bassin versant, l'exutoire est unique. Ainsi pour chacun, la charge théorique à l'exutoire et la valeur de débit alloué à l'exutoire sont récupérées. La valeur de charge théorique relative de chaque pixel (en fait maintenant, entité vectorielle `Point`) est ensuite calculée par rapport à la charge théorique de l'exutoire. La ratio de charge relative peut alors être appliqué à la valeur de débit alloué à l'exutoire. Le `DataFrame` en sortie est finalement exporté comme raster.

Le script `ALLOCATEobservedDischarge.py` est l'analogue de celui que nous venons de décrire mais s'applique aux débits observés des stations préalablement rasterisés. Son fonctionnement est identique à ceci près qu'il ne prend pas en entrée la couche dérasterisée des pixels de débit modélisé, mais directement la couche des débits observés (préalablement projetés dans la grille raster de \mathcal{D}_{dst} , cf. section suivante). Le script considère alors, pour chaque bassin versant, les pixels ayant la plus grosse charge hydraulique et lit la valeur de débit observé qui leur est associée. Cette valeur de débit est ensuite convoluée le long du réseau au prorata de la charge hydraulique relative de chaque pixel par rapport à l'exutoire. Remarquons dans ce cas que le pixel sur lequel se situe une station de mesure fait office d'exutoire au bassin versant : les bassins versants ne sont pas délimités par rapport aux stations, mais comprennent de fait les bassins des stations car leur délimitation a été effectuée par rapport à l'emprise de la zone d'étude, qui elle-même contient nécessairement toutes les stations utilisées.

Enfin, notons que cet algorithme doit être exécuté autant de fois que de rasters dans \mathcal{D}_{src} , c'est-à-dire 12 fois par type de modèle (`DIST` et `PRIST`). En effet, les valeurs de débits changent pour chaque raster, bien que la charge relative théorique soit constante.

2.2.4 Algorithme de rasterisation des données ponctuelles de débit observé

Nous avons évoqué ci-dessus le script `ALLOCATEobservedDischarge.py` comme un analogue du script `ALLOCATEmodelDischarge.py` alors qu'il y en réalité une différence initiale et fondamentale entre la convolution des données de modèle et celle des données de mesure. Les données de modèle sur lesquelles nous effectuons la convolution sont en effet issues d'un travail préalable de dérasterisation (ou éclatement) des données de \mathcal{D}_{src} . Une fois ce travail effectué, les données de modèle sont définies dans l'espace \mathcal{D}_{dst} , qui est le même espace où est défini le support de la convolution, à savoir le raster d'accumulation de flux. Une préparation des données de mesure est donc également nécessaire pour qu'elles s'inscrivent dans le même espace \mathcal{D}_{dst} . Les données de mesure sont effet exogènes à la modélisation hydrographique que nous avons effectuée dans \mathcal{D}_{dst} . Il est donc très peu probable (en réalité presque impossible) que la position d'une station du réseau Vigicrues intersecte le pixel correspondant dans le réseau hydrographique modélisé. Les données de mesure doivent ainsi être rasterisée en forçant l'opération de carroyage à les associer à un pixel contributif du réseau hydrographique (i.e. un pixel non-NoData du raster d'accumulation de flux filtré au seuil α_a).

C'est à cet effet qu'est prévu le script `JOINaccuflux2stations.py` et en particulier l'appel qu'il fait à la fonction `relio.join_points_to_pixels` que nous avons déjà eu l'occasion de décrire dans le chapitre 4 de ce manuscrit. Cette fonction implémente un algorithme glouton de jointure spatiale que nous avons déjà eu l'occasion de décrire dans la partie concernant le MHE au chapitre (3). La vérification de la jointure s'effectue manuellement après cette première opération automatique, dans un logiciel de SIG à interface graphique, en superposant le raster d'accumulation de flux et la couche vectorielle de géométrie `Point` générée par le script `JOINaccuflux2stations.py` en sortie. Si une station enregistre un débit qui ne correspond vraisemblablement pas à la valeur d'accumulation à laquelle elle a été associée, alors la position du point est décalée en superposition d'un pixel proche et vraisemblable dans le raster d'accumulation de flux. Notons que cette correction manuelle est menée à partir des valeurs de débits prises arbitrairement pour le mois de janvier.

Cette démarche de jointure en deux temps entre la position des stations et le réseau hydrographique modélisé est similaire à celle de (LEHNER 2012) où l'enjeu est de joindre les stations du GRDC (Global Runoff Data Center) avec le réseau hydrographique HydroSHEDS de résolution 15 arcsec. Les deux temps de la procédure sont :

1. Une jointure automatisée. Un rayon de recherche de 5km autour de chaque station déter-

minent les pixels à partir desquels sont délimités les bassins versants potentiels de chaque station. Leur surface est calculée et comparées à la valeur d'emprise du bassin amont de la station dans les données GDRC. Les pixels dont le bassin versant présente un écart de plus de 50% avec la valeur connue ne sont pas retenus. La distance euclidienne entre la station et les pixels dans le rayon de recherche sont également calculées. Le pixel retenu est celui dont la surface est la plus proche de la surface connue *et* qui est le moins distant de la station.

2. Une jointure manuelle : pour les stations n'ayant aucun pixel dont la surface du bassin versant est moins de 50% différente de la surface réelle, ou les stations qui n'ont pas de données de surface de bassin amont dans la base GDRC, ces stations sont attribuées manuellement à un pixel.

Cette démarche que le GDRC a implémenté est reprise dans (BARBAROSSA et al. 2018) mais à partir du réseau hydrographique HydroSHEDS de résolution 30 arcmin.

C'est cette correction manuelle d'erreur de jointure que le paramètre δ du modèle MRH désigne. Nous ne donnons de formalisation plus poussée de ce paramètre mais nous le reportons comme tel afin d'en souligner l'existence et, très certainement, la forte influence sur les performances du modèle. Une analyse de sensibilité entre différents jeux de δ pourrait à cet égard être menée.

2.2.5 Algorithmes pour les traitements finaux

Cinq scripts ont été spécifiquement développés dans le but de mettre en forme les résultats obtenus par les algorithmes principaux.

Nous rappelons ici les 4 groupes de résultats recherchés :

1. les données de mesure Vigicrues rééchantillonnées par **relio** *versus* les données Vigicrues non-rééchantillonnées
2. les données du modèle WBM rééchantillonnées par **relio** *versus* les données de mesure Vigicrues non-rééchantillonnées
3. les données de mesure Vigicrues rééchantillonnées par **invdist** *versus* les données Vigicrues non-rééchantillonnées

4. les données du modèle WBM rééchantillonnées par `invdist` *versus* les données Vigicrues non-rééchantillonnées

Traitements pour les résultats du groupe 1

Le groupe (1) vise à évaluer les performances de la méthode `relio` en considérant un sous-ensemble des données Vigicrues qui seront rééchantillonnées selon la méthode et comparées au reste de l'ensemble des données Vigicrues intouchées. Si la méthode `relio` se décompose en trois algorithmes centraux (dérasterisation, rasterisation et convolution), l'évaluation du groupe (1) porte spécifiquement sur l'algorithme de convolution. En effet, pour procéder à cette évaluation, les données Vigicrues doivent d'abord avoir été rasterisées, c'est-à-dire exprimées dans l'espace \mathcal{D}_{dst} . Quant à l'algorithme de dérasterisation, il ne s'applique qu'aux données de modèle initialement carroyées, ce qui n'est pas le cas des données Vigicrues.

Le script `PERFORMANCEmetrics.py` a été développé dans le but d'évaluer le groupe (1). Au préalable, les scripts `JOINaccuflux2stations.py` et `ALLOCATEobservedDischarge.py` doivent avoir été exécutés. Pour rappel, ils permettent respectivement d'associer à toute station définie dans la couche vectorielle de géométrie `Point` un pixel du raster d'accumulation de flux, et d'effectuer la convolution bassin versant par bassin versant à partir des stations ayant la plus grande valeur de débit. C'est au cours de l'exécution de ce deuxième script que l'échantillon des stations Vigicrues qui serviront de support à la convolution sont sélectionnées. Il faut s'imaginer qu'à l'issue de ce script, lorsque que l'on convertit en couche vectorielle l'image raster en sortie, on obtient un tableau de données dont la profondeur (nombre de lignes) est égale au nombre de bassins versants contenus dans le raster des bassins versants et dont les colonnes sont la valeur d'accumulation de flux et le débit observé. Autrement, dit une et une seule station d'observation est retenue pour chaque bassin versant. Le reste des stations de mesure servira de point de contrôle. Pour la comparaison pixel-à-pixel, entre ces stations intouchées et les pixels rééchantillonnées, il faut là aussi se figurer que le script `ALLOCATEobservedDischarge.py` conserve pour chaque station les valeurs observées de débits mensuels et calcule les valeurs convoluées de débits mensuels, qui sont deux attributs distincts pour une même entité dans la couche vectorielle en sortie.

Le script `PERFORMANCEmetrics.py` récupère alors cette couche vectorielle et effectue le calcul des métriques de performance élémentaires (que nous décrivons dans la section "Résultats") et

effectue un export des résultats bassin versant par bassin versant.

Traitements pour les résultats du groupe 2

L'évaluation du groupe (2) vise à comparer les données prédites par le modèle WBM avec les données de mesure Vigicrues. Cette évaluation est le résultat principal pour lequel la méthode `relio` a été développée : permettre la comparaison entre des données de modèle carroyées et des données de mesure ponctuelles. L'évaluation du groupe (2) ne vise donc pas les performances propres à `relio`, qui sont évaluées avec les groupes (1) et (3).

Les scripts `MERGEresample2observed.py` et `PERFORMANCEmetrics2.py` ont été développés dans le but de procéder à l'évaluation du groupe (2).

Le script `MERGEresample2observed.py` doit être exécuté en premier. Ce script a le rôle décisif de fusionner le raster des pixels rééchantillonnés de débits modélisés avec le raster des stations de mesure, qui ont en effet été générés indépendamment l'un de l'autre. Il commence par fusionner les rasters de débits rééchantillonnés (il y en a un par bassin versant) à l'aide de la fonction `relio.merge_sameSizedRasters` que nous avons développée pour "coller" bout-à-bout des rasters contigus afin de former un unique raster qui les inclut tous. Cette opération est nécessaire car la rasterisation des stations de mesure a été effectuée avec le raster d'accumulation de flux dont la dimension est égale à celle de la zone d'étude. En effet, pour fusionner les stations avec les débits rééchantillonnés, nous allons fonctionner par plus proche voisin, ce qui suppose de fusionner des rasters de même dimension. La fusion par plus proche voisin recourt à la fonction `relio.join_points_to_pixels` que nous avons déjà eu l'occasion de présenter. A l'issue de cette opération, chaque station de mesure rasterisée est associée à son plus proche voisin dans le raster des débits rééchantillonnés (en fait, son "plus proche voisin" est le pixel en superposition car les deux rasters sont définis dans \mathcal{D}_{dst} et donc se superposent parfaitement). La couche vectorielle obtenue en sortie est alors découpée par bassins versants pour permettre une analyse différenciée des résultats. Une dernière étape consiste à convertir les valeurs de débits observés dans l'unité des valeurs de débits modélisés, . La conversion des $l.s^{-1}$ vers des $km^3.an^{-1}$ est effectuée avec le facteur de conversion $c = 1e^{-3} \times 1e^{-9} \times 60 \times 60 \times 24 \times 365$ qui s'implémente dans la formule : $y_{i,t} [km^3.an^{-1}] = y_{i,t} [l.s^{-1}] \times c$

Traitements pour les résultats du groupe 3

L'évaluation du groupe (3) permet la validation croisée de `relio`. En effet, comme nous l'avons évoqué plus haut, le groupe (1) permet seulement de valider l'algorithme de convolution linéaire. Les algorithmes de rasterisation et dérasterisation en revanche ne font pas l'objet d'une validation spécifique. A défaut, il est intéressant de comparer `relio` à une autre méthode de rééchantillonnage plus conventionnelle, en l'occurrence la méthode de la distance inverse, que nous avons appelé `invdist`. La comparaison deux-à-deux des sorties de `relio` et `invdist` n'est pas particulièrement utile car elle renseigne seulement sur les écarts de prédiction entre les deux modèles, et pas par rapport à une référence jugée fiable. En revanche, la comparaison successive de `relio` avec un échantillon de données de référence, puis de `invdist` avec un échantillon de référence permet d'effectuer une comparaison croisée des modèles vis-à-vis d'une référence exogène.

L'évaluation du groupe (3) est effectuée dans le script `CONTINUOUSresamplingOfStations.py` qui effectue le rééchantillonnage des valeurs de débits de 10 stations de mesure selon la méthode `invdist`. Nous avons choisis ces stations de mesure parmi l'ensemble des stations de mesure selon un critère de représentativité statistique. En effet, à la façon d'une campagne de mesure qui serait menée sur le territoire, on viserait à obtenir un échantillon représentatif des cours d'eau en présence. Ainsi, Pour chaque mois de la période 2000-2020, on récupère une station par décile de distribution des débits : la station correspondant au 10% des débits les plus faibles, puis la stations des 20% des débits les plus faibles, etc.

Cette échantillon est ensuite rééchantillonné vers \mathcal{D}_{dst} selon la méthode `invdist`. Pour cela, nous avons développé une suite de fonctions qui convertissent d'abord l'échantillon des 10 stations en image raster virtuelle (format `.vrt`, compatible avec l'interpolateur natif de `gdal` qui sera utilisé ensuite) puis interpolent ce raster dans \mathcal{D}_{dst} . l'interpolation se fait au moyen de la fonction `relio.interpolate` qui s'appuie sur la fonction `gdal_grid` dont nous donnons ci-après la commande simplifiée pour mieux en saisir les paramètres principaux :

```
gdal_grid -a -a_srs -txe -tye -outsize
```

Le paramètre `-a` introduit le nom de l'algorithme d'interpolation qui sera utilisé. ici, nous utilisons l'algorithme `invdist` de `gdal` bien qu'il en existe d'autres. Cet algorithme est paramétrable, notamment pour ce qui est du rayon de recherche et du facteur de lissage, mais nous ne modifions par les paramètres par défaut. Remarquons que nous avons développé

le script `CONTINUOUSresamplingOfStations.py` de telle sorte à permettre de modifier le paramètre `-a` au besoin, ainsi que de modifier les paramètres. Le reste des paramètres permet de spécifier à l’algorithme que l’interpolation doit s’effectuer dans \mathcal{D}_{dst} . Pour cela, la valeur de ces paramètres est héritée du MNT, qui est lui-même défini dans \mathcal{D}_{dst} .

A la suite de cette opération, il reste encore à comparer deux-à-deux les pixels du raster rééchantillonné obtenu en sortie avec les points des stations de mesure. Le premier raster est pour cela d’abord converti en couche vectorielle de géométrie `Point`. On charge ensuite la couche vectorielle des stations de mesure puis on effectue une jointure spatiale directe par plus proche voisin avec la librairie `geopandas` et sa fonction dédiée `geoDataFrame.sjoin_nearest` dont l’attribution unique (un pixel est attribué à une unique station) est assuré par l’option `how='inner'` qui permet une jointure interne sur la couche des stations. Remarquons que le raster des pixels rééchantillonné est continu : à tout pixel de \mathcal{D}_{dst} a été attribué une valeur, indépendamment du fait que le pixel sous-jacent fasse partie ou non d’un tronçon hydrographique. Il n’y a donc pas besoin d’effectuer l’étape intermédiaire de rastérisation des stations de mesure vers \mathcal{D}_{dst} puisque l’espace est continu en tout point (formellement, pixel).

Enfin, on convertit les débits des stations dans l’unité des débits rééchantillonnés, de la même façon que nous l’avons fait pour l’évaluation du groupe 2.

Traitements pour les résultats du groupe 4

L’évaluation du groupe (4) a avant tout un but indicatif puisqu’elle consiste à rééchantillonner les débits du modèle WBM par une méthode de rééchantillonnage conventionnelle et continue, en l’occurrence `invdist`. En effet, la comparaison directe des débits WBM rééchantillonnés par `relio` et `invdist` est assez faible en conclusion car sans référence externe il est difficile d’évaluer les performances comparées des deux méthodes.

L’évaluation du groupe (4) est conduite dans le script `CONTINUOUSresamplingOfWBM.py` qui fonctionne à la façon du script que nous présentions précédemment. La différence réside dans les données d’entrée : il s’agit plus de rééchantillonner un échantillon des stations de mesure mais directement les pixels du modèle WBM définis dans \mathcal{D}_{src} . Conformément à ce que nous décrivions sur le fonctionnement des algorithmes de rééchantillonnage conventionnels, qui utilise le barycentre des pixels comme point de départ de l’interpolation, nous commençons par extraire

les barycentres des pixels du raster WBM. Cette couche de points est ensuite convertie en raster virtuel puis interpolée à l'aide de la fonction `relio.interpolate` dont nous gardons les paramètres inchangés par rapport au groupe (3). De même, nous effectuons ensuite une jointure par plus proche voisin avec la couche des stations de mesure puis nous exprimons dans une unité commune les valeurs de débits.

2.2.6 Paramétrisation du raster d'accumulation de flux : initialisation aux bornes

Nous allons d'abord décrire l'opération de base qui consiste à créer le raster d'accumulation de flux. Nous décrirons ensuite les raffinements apportés à ce raster, à savoir la prise en compte des conditions aux bornes de la zone d'étude et le filtrage à un seuil de charge donné.

Pour la mise en charge initiale, nous avons créé un raster dont tous les pixels valent 1, pour simuler une pluie unitaire uniforme. En sortie de `relio.create_accuflux`, tous les pixels valent donc au moins 1. L'enjeu est alors de savoir à partir de quel niveau d'accumulation (ou charge cumulée, ou tout simplement charge), on peut considérer un pixel comme formant partie d'un cours d'eau actif. Ce seuil de charge minimum est désignée par le paramètre α_a . Le script `HYDROGRAPHICinputs.py` permet d'itérer sur la valeur de α_a en modifiant le paramètre `epsilonAccuflux` et en récupérant itérativement la couche `DEMx_SRTM30_wgs84_pcraster_-accuflux_sup{str(epsilonAccuflux)}.tif` générée en sortie.

Nous avons commencé par décrire ce raffinement du raster d'accumulation de flux en premier car il s'agit du plus évident, ou intuitif. En pratique, il est plus simple de commencer numériquement par le raffinement à suivre : la prise en compte des conditions aux bornes. Cette étape est indispensable pour le bon fonctionnement de la méthode `relio` dont le principe est d'allouer un débit connu en un point du réseau au reste du réseau par rapport de proportionnalité entre la charge théorique en ce point connu et tout autre point du réseau. Il faut donc que les charges hydrauliques relatives du réseau rendent compte au plus proche de la réalité hydrologique de la zone d'étude. Un facteur particulièrement influant pour la prise en compte réaliste des charges relatives du réseau réside dans les conditions aux bornes de la zone d'étude. En effet, la zone d'étude étant un système ouvert et arbitrairement positionnée (i.e. positionnée par un découpage administratif), elle ne peut nécessairement pas correspondre au découpage naturel des bassins versants. De là vient que les bassins versants de la zone d'étude sont des sous-ensembles des bassins versants naturels, en particulier pour ceux qui se situent

en bordure de la zone d'étude. Deux situations peuvent alors apparaître pour de tels bassins versants "coupés" : soit l'exutoire se situe à l'intérieur de la zone d'étude et la tête de bassin se situe en dehors, soit à l'inverse l'exutoire se situe à l'intérieur de la zone d'étude et la tête de bassin à l'extérieur, soit une situation intermédiaire entre ces deux situations. La situation la moins favorable pour notre estimation des charges relatives du réseau est lorsque la tête de bassin se situe en dehors de la zone d'étude. En effet, le raster de charge initiale agit comme si tout pixel était une tête de bassin. Or certains pixels de la zone d'étude sont susceptibles d'hériter d'une charge à l'amont. La situation où l'exutoire naturel ne se trouve pas dans la zone d'étude n'est pas dérangeante pour notre étude, car l'exutoire "fictif" qui se trouvera en bordure de zone d'étude représentera bien la charge amont accumulée jusqu'à lui.

Déterminer la charge héritée de l'amont n'est numériquement pas chose facile. Une méthode brutale voudrait que l'on élargisse la zone d'étude jusqu'à ce qu'elle englobe les bassins versants naturels. Pour notre cas d'étude en Sud-Loire, la zone d'étude que nous avons fixé comparant le bassin versant du Rhône, de la Saône, de l'Allier et de la Loire. On imagine bien dès lors la très grande de taille des jeux de données à manipuler pour une résolution de 30×30m, qui ne peuvent être traités que sur serveur de calcul. En outre, plus la zone d'étude est élargie, moins le raster de charge unitaire et uniforme devient pertinent (réaliste) car plus l'hypothèse d'uniformité s'affaiblit. Une méthode plus fine, et que nous avons adoptée, consiste à déterminer itérativement une zone d'étude (ou enveloppe) optimale, c'est-à-dire à élargir progressivement l'enveloppe à partir d'une taille initiale jusqu'à ce que le nombre de cours d'eau qui intersectent de l'amont la zone d'étude soit jugé raisonnable pour envisager une correction manuelle de la charge initiale des pixels auxquels ils se branchent dans le réseau.

Le script `SETenvelope.py` est dédié à cette opération initiale par laquelle l'enveloppe optimale est déterminée. L'idée est de fournir une enveloppe en entrée (ou emprise) et d'effectuer une différence symétrique avec la couche vectorielle du réseau hydrographique de la BD Topage de géométrie `Linestring` à l'aide de la librairie `geopandas`. On obtient ainsi la couche vectorielle des tronçons hydrographiques qui intersectent la zone d'étude, mais seulement leur portion amont ou aval. Le script calcule ensuite la longueur de ces tronçons amont ou aval (en mètres) avec la méthode `length` de `geopandas` puis génère la fonction de distribution des longueurs de tronçons et enfin propose d'exporter la couche vectorielle avec un seuil de coupure dans la fonction de distribution. Par exemple, si le seuil de coupure (qui correspond au paramètre `percentile` dans le script) est fixé à 90%, cela signifie que la couche vectorielle

des tronçons amont ou aval à la zone d'étude ne comportera que tronçons qui, en cumulé, représentent 90% de la longueur totale de tous les tronçons. Nous avons en effet observé que la distribution des longueurs suit approximativement une loi log : un petit nombre de tronçons sont responsable d'un grand pourcentage de la longueur amont et aval. Ne conserver que ces quelques tronçons permet ainsi de s'économiser la tâche manuelle et fastidieuse de correction des charges initiales aux pixels d'embranchement.

Pourquoi s'intéresser à la longueur des tronçons amont et aval à la zone d'étude pour déterminer les charges initiales aux bornes? L'idée est d'utiliser la longueur d'un tronçon comme un proxy de sa charge hydraulique. Si l'on divise en effet sa longueur (exprimée en mètres) par la résolution 30m des pixels, on obtient un "équivalent-nombre-de-pixels" qui est très proche de la façon dont fonctionne l'accumulation de flux, où l'initialisation de la charge à 1 et la convolution pixel-par-pixel revient également à dénombrer les pixels le long du réseau. Cette technique d'estimation du débit amont est toutefois à considérer comme une première tentative et demeure largement imparfaite. En effet, elle ne prend en compte que le tronçon principal qui arrive ou part du territoire, et pas ses affluents en dehors de la zone d'étude, ce qui est problématique pour le cas des cours d'eau importants arrivant sur la zone d'étude, dont la charge a nécessairement été augmentée de celles de ses affluents à l'amont du territoire. La méthode tend ainsi à sous-estimer les charges héritées de l'amont. Enfin, la méthode ne permet pas de déterminer le sens d'écoulement des tronçons : s'agit-il de tronçons amont ou aval à la zone d'étude? Seuls nous intéressent les cours d'eau amont (qui apportent une charge à la zone d'étude), et seule la connaissance du territoire par l'utilisateur permet de les identifier.

Nous pouvons en profiter pour préciser la formalisation mathématique du paramètre β_p , qui désigne précisément le proxy "équivalent-nombre-de-pixels" que nous venons de décrire.

β_p est en effet le vecteur des charges initiales aux bornes de la zone d'étude. Plus précisément, il contient les valeurs de charge des n pixels de \mathcal{D}_{dst} qui héritent d'une charge à l'amont, $\beta_p := (a_k)_{1 \leq k \leq n}$. Comme nous l'avons précisé, les a_k ne sont pas connus mais estimés par le proxy "équivalent-nombre-de-pixels". On note l_k la longueur du tronçon à l'amont du pixel k et r la résolution d'un pixel de \mathcal{D}_{dst} . On calcule alors les a_k de la façon suivante :

$$\forall 1 \leq k \leq n, \quad \hat{a}_k = \frac{l_k}{r}$$

On peut noter L la variable aléatoire des $(l_k)_{1 \leq k \leq n}$. Avec ces notations, l'indice p de β_p

correspond alors au $p^{\text{ième}}$ quantile de la fonction de distribution de L , à savoir : $\mathcal{P}(L \leq \gamma) = p$ où γ (gamma) désigne la longueur amont cumulée telle que p soit réalisé. Si les \widehat{a}_k sont rangés par ordre croissant, numériquement on cherche le premier k tel que p soit réalisé.

$$k? \quad | \quad \mathcal{P} \left(L \leq \sum_{i=1}^k \widehat{a}_i \right) = p \quad \text{où} \quad \forall 1 \leq i \leq k, \quad \widehat{a}_i \leq \widehat{a}_{i+1}$$

Une fois la taille optimale (du moins satisfaisante) de l'enveloppe déterminée, il s'agit de modifier manuellement la valeur des pixels qui se situent à l'embranchement avec les tronçons amont dans le raster de charge initiale. Ces pixels se verront attribuer la valeur "équivalent-nombre-de-pixels" qui a été calculée précédemment, tandis que les pixels non concernés par un embranchement amont en bordure de zone d'étude conserveront une charge unitaire. Deux scripts sont dédiés à cette opération : `INITIALIZEaccuflux.py` qui doit être exécuter en premier et `INITIALIZEmaterial.py` qui est exécuter en second.

Le script `INITIALIZEaccuflux.py` permet de générer un raster d'accumulation de flux à partir d'une charge initiale unitaire et uniforme. Le script renvoie également une copie du raster d'accumulation de flux mais à deux bandes : une bande qui donne les valeurs de charge accumulée et une bande qui donne un identifiant de pixel (nombre entier allant de 1 à la dimension du raster). Ce raster à deux bandes est ensuite affiché dans une interface graphique SIG et superposé à la couche vectorielle des tronçons amont et aval générée précédemment. Il s'agit alors d'effectuer la correspondance entre les tronçons amont et leur pixel d'embranchement. Dans le script `INITIALIZEmaterial.py`, on vient instancier un vecteur d'initialisation aux bornes où l'on range pour chaque pixel d'embranchement, l'identifiant du pixel et la valeur "équivalent-nombre-de-pixels" correspondante dans la couche vectorielle. L'exécution de ce second script renvoie un nouveau raster de charge initiale paramétré aux bornes.

Enfin, nous revenons à la première étape que nous décrivions (qu'il était plus intuitif de présenter en premier bien qu'elle n'intervienne qu'en dernier dans les traitements). Le script `HYDROGRAPHICinputs.py` peut être exécuter car il s'occupe de générer le raster d'accumulation de flux à partir du raster modifié de charge initiale. Il s'agit alors d'itérer sur les valeurs de α , comme nous décrivions plus haut, afin de déterminer visuellement la valeur la plus satisfaisante, α_a . Remarquons que le script `SETaccufluxThreshold.py` a pour but d'accompagner la détermination optimale de α_a en minimisant l'écart de densité entre les pixels non-NoData du raster d'accumulation de flux et les pixels d'une couche vectorielle du réseau hydrographique Topage préalablement converti en image raster.

2.2.7 Paramétrisation du raster des exutoires : seuil de charge

Un raffinement apporté par le MRH au raster des exutoires fourni par le MHE consiste à introduire un filtre qui ne retient que les pixels exutoires ayant accumulé une charge supérieure à un certain seuil. Ce seuil permet de réduire considérablement les temps de calculs. En effet, de nombreux cours d'eau de longueur très faible sortent de la zone d'étude et sont à ce titre considérés comme des exutoires. Ceci conduit à multiplier le nombre des bassins versants, notamment par de tous petits bassins en bordure d'image, ce qui ne pose en soi pas de problème pour les traitements suivants mais conduit à multiplier le nombre des traitements car la suite des opérations consistera à traiter chaque bassin versant au cas par cas. Or le fonctionnement des traitements par images raster implique de toujours garder la même dimension d'image (pour que les superpositions spatiales soient exactes) : même pour de tous petits bassins versants, la taille totale de la zone d'étude est préservée, ce qui conduit à traiter des images quasiment vides de données alors que le temps de traitement d'une image dépend du nombre de pixels à lire et pas de l'information portée par les pixels. Ainsi donc un seuil d'accumulation α_b est fixé itérativement.

L'opération d'extraction des pixels exutoires est effectuée dans le script `OUTLETSlocation.py` et commence par extraire du raster \mathcal{D}_{src} les pixels en bordure d'image puis, pour chacun de ces pixels, les pixels superposés de \mathcal{D}_{dst} en bordure sont extraits. Cette deuxième opération est effectuée à l'aide de la fonction `relio.extract_edge_outlets` qui, malgré son nom, élargit la recherche des exutoires au voisinage des pixels en bordure car il existe un bruit numérique par lequel certains pixels sont bien encodés comme des exutoires dans le raster de direction de flux sans se situer exactement sur la bordure de l'image, mais à quelques pixels près. Le raster d'accumulation de flux filtré est ensuite également coupé à la dimension de chaque pixel de \mathcal{D}_{src} en bordure (afin de réduire les temps de calcul) puis la fonction `relio.significant_outlets` applique le filtre α_b sur les exutoires identifiés.

Précisons davantage le fonctionnement de `OUTLETSlocation.py` qui effectue une série d'opérations décisives dans l'ensemble de la méthode `relio`. C'est en effet à ce stade que sont distingués, pour chaque pixel en bordure de \mathcal{D}_{src} , les pixels sous-jacents de \mathcal{D}_{dst} correspondant aux exutoires finaux et aux exutoires de transfère. La fonction `relio.significant_outlets` opère à l'intérieur d'un pixel de \mathcal{D}_{src} , qu'un peut noter j . Elle passe en revue tous les pixels i

sous-jacent de \mathcal{D}_{dst} (raster de direction de flux) situés en bordure. Pour chaque arrête de j , la fonction cherche les directions d'écoulement des pixels i situés l'arrête telles qu'elles dirigent l'écoulement en dehors de j . Par exemple, pour l'arrête Nord de j , seules les directions d'écoulement Nord, Nord-Ouest, Nord-Est conduisent le flux à "s'échapper" de j . Dans l'encodage D8 du raster de direction de flux, ces trois directions données en exemple correspondent aux valeurs 7, 8 et 9. A ces valeurs, nous ajoutons la valeur 5 qui correspond à l'encodage des dépression effectué par la fonction `pcraster.lddcreate` et c'est pour cette valeur que nous élargissons la recherche au voisinage des pixels i en bordure j (au voisinage "interne", c'est-à-dire toujours en restant bien à l'intérieur de j). La fonction `relio.significant_outlets` opère de façon analogue pour les directions Sud, Est et Ouest d'écoulement. Elle crée un raster en sortie dont la dimension est égale à taille du pixel j et la résolution est égale à la taille des pixels i . Pour des questions pratiques, nous exportons ce raster comme une couche vectorielle de géométrie `Point`, ce qui permettra dans un second temps, une fois tous les pixels j en bordure traités, d'invoquer la fonction `pandas.concat` pour concaténer les couches de sortie et les traiter plus facilement comme objets `DataFrames`. A l'issue de cette étape, tous les pixels (désormais "entités" de couche vectorielle) sont passés en revue et labellisés `True` ou `False` selon qu'ils correspondent respectivement à des exutoires finaux ou de transfère. Plus précisément, si leur code de direction vaut 5, ils sont considérés comme des exutoires finaux, sinon comme des exutoires de transfère. La valeur correspondant à chacun de ces pixels dans le raster d'accumulation de flux filtré est ensuite récupérée et ajoutée au `DataFrame` (par jointure attributaire d'identifiant de pixel, ce qui permet le travail sur des couches vectorielles plutôt que sur des images raster, et la fonction `relio.extract_cellValues` assure la conversion d'un raster vers une couche vectorielle).

Remarquons enfin que le raster de \mathcal{D}_{src} à partir duquel sont extraits les pixels en bordure a arbitrairement été choisi comme le raster des débits perturbés pour le mois de janvier (`WBM_TerraClimate_OUTPUTQ_DIST_mLTM_2000-2020_band1_clip.tif`). Ce choix n'a aucune importance puisque les 24 rasters de \mathcal{D}_{src} partage les mêmes propriétés spatiales (résolution, dimension, projection) et qu'on ne s'intéresse ici pas encore aux valeurs des pixels mais simplement à leurs propriétés spatiales.

Résumé

Le cadre des limites planétaires alerte sur la déstabilisation irréversible des grands équilibres climatiques du système Terre causée par les activités humaines. En définissant les conditions d'habitabilité de la Terre, il permet de caractériser à nouveaux frais la soutenabilité des activités humaines à une échelle locale à travers un ensemble de méthodes dites de soutenabilité environnementale absolue (AESA). Ce travail de thèse de Doctorat étudie la possibilité d'une utilisation du cadre des limites planétaires comme outil chiffré d'évaluation des pressions anthropiques générées par un territoire sur l'environnement. Cette question de la validité de la descente d'échelle du cadre, du système Terre vers un territoire, est abordée par le prisme de la fiabilité des données environnementales. En se concentrant sur la limite planétaire relative à l'eau de surface et en prenant comme cas d'étude le Schéma de Cohérence Territoriale (ScoT) du Sud-Loire, ce travail commence par proposer une approche systématique pour la caractérisation et la hiérarchisation des données environnementales liées à l'eau dans un contexte d'AESA. Le travail développe ensuite deux méthodes d'AESA s'appuyant sur les mesures systématisées de débits en cours d'eau du réseau national Vigicrues. Une première méthode tient au développement du logiciel **HydroTrendsFrance** qui permet d'évaluer l'intégrité des cours d'eau sur le territoire du SCoT Sud-Loire au regard de la limite planétaire. La seconde méthode relève du développement de **relio**, un modèle numérique de *data assimilation* permettant le couplage entre les données d'un modèle hydrologique global et les données Vigicrues.

Mots clés : limites planétaires, évaluation environnementale, évaluation absolue de la soutenabilité environnementale, système d'information sur l'eau, données environnementales, géomatique

Abstract

The planetary boundaries framework warns of the irreversible destabilization of Earth's major climatic balances caused by human activities. By defining the conditions for Earth's habitability, it provides a new perspective for characterizing the sustainability of human activities at a local scale through a set of methods known as absolute environmental sustainability assessment (AESA). This PhD thesis examines the potential use of the planetary boundaries framework as a quantitative tool for assessing the anthropogenic pressures exerted by a local area on the environment. The validity of scaling down the framework—from the Earth system to a specific territory—is explored through the lens of environmental data reliability. Focusing on the freshwater planetary boundary, and using the Sud-Loire region in France as a case study, this work begins by proposing a systematic approach to characterize and prioritize environmental data related to water within an AESA context. It then develops two AESA methods based on standardized measurements of river flows from the Vigicrues national monitoring network. The first method involves the development of the software `HydroTrendsFrance`, which evaluates the integrity of watercourses in the Sud-Loire region against the corresponding freshwater planetary boundary. The second method consists on the development of `relio`, a numerical data assimilation model that enables coupling data from a global hydrological model and Vigicrues observations.

Key words : planetary boundaries, environmental assessment, absolute environmental sustainability assessment, water information system, environmental data, geomatics