



Un modèle spatio-temporel hybride de SIG temporel : application à la géomorphologie marine

Younes Hamdani

► To cite this version:

Younes Hamdani. Un modèle spatio-temporel hybride de SIG temporel : application à la géomorphologie marine. Informatique [cs]. Université de Bretagne Occidentale (UBO), 2020. Français. NNT : . tel-03403473v1

HAL Id: tel-03403473

<https://hal.science/tel-03403473v1>

Submitted on 12 Feb 2021 (v1), last revised 26 Oct 2021 (v3)

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.

THESE DE DOCTORAT DE

L'UNIVERSITE
DE BRETAGNE OCCIDENTALE

ECOLE DOCTORALE N° 598
Sciences de la Mer et du littoral
Spécialité : *Géomatique*

Par

Younes HAMDANI

Un modèle spatio-temporel hybride de SIG temporel : application à la géomorphologie marine

Thèse présentée et soutenue à Brest, le **26/11/2020**
Unité de recherche : **Institut de Recherche de l'école navale (IReNav)**

Rapporteurs avant soutenance :

Marlène VILLANOVA-
OLIVER

Maître de conférences HDR
Université Grenoble Alpes

Rodolphe DEVILLERS

Directeur de Recherche HDR
IRD, Montpellier

Composition du Jury :

Rapporteurs :
Marlène VILLANOVA-OLIVER

Maître de conférences HDR
Université Grenoble Alpes

Rodolphe DEVILLERS

Directeur de Recherche HDR
IRD, Montpellier

Examineurs
Thérèse LIBOUREL

Professeur Émérite des Universités
Université de Montpellier

Géraldine DEL MONDO

Maître de conférences
INSA Rouen

Laurent NANA

Professeur des Universités
Université de Bretagne Occidentale

Directeur de thèse :
Christophe CLARAMUNT

Professeur des Universités
Ecole Navale

Encadrant de thèse :
Rémy THIBAUD

Enseignant-Chercheur
ENSTA Bretagne

Remerciements

La thèse elle-même est une expérience à une échelle temporelle plus réduite qui symbolise le parcours de la vie dont la fin principale est l'amélioration de soi et la contribution collective. Entre ordre et chaos, certitude et doute, satisfaction et frustration, victoire et défaite, le chemin passe de l'un à l'autre de manière régulière. Ces phénomènes que nous percevons comme opposés peuvent être considérés sous un autre angle comme complémentaires et ce passage de l'un à l'autre crée une certaine dynamique et constitue une sorte d'enrichissement et de guide dans notre cheminement vers le sens et la perfection. Dans cette voie, nous devons continuellement apprendre et retracer un chemin de progression clair pour atteindre le but, en gardant toujours la foi et l'espoir. Je dis donc un énorme MERCI à tous ceux qui m'ont aidé d'une manière ou d'une autre à mieux apprendre, à voir plus clair et à progresser grâce à leur qualité et leur soutien.

Tout d'abord, je tiens à remercier chaleureusement et sincèrement mon directeur de thèse, Monsieur Christophe Claramunt, pour sa patience, son interactivité et son intelligence visionnaire dans ce processus d'élaboration de la thèse. Je voudrais également lui dire combien ses encouragements permanents, sa confiance en moi et ses conseils pertinents m'ont été d'un grand soutien et que ses relectures scientifiquement rigoureuses m'a permis de progresser en termes de qualité de rédaction et de méthodologie scientifique. J'ai beaucoup apprécié son soutien dans les moments difficiles de la thèse, son attention particulière aux doctorants et son ouverture d'esprit qui a contribué à maintenir une bonne ambiance dans le laboratoire de recherche. Donc un GRAND merci.

Je tiens à remercier avec la même chaleur et la même vivacité mon encadrant de thèse, Monsieur Rémy Thibaud, pour tout le travail et l'effort d'accompagnement et d'encadrement qu'il a investi et surtout pour les innombrables heures de discussions intéressantes liées à mon travail. Je tiens également à lui dire combien sa proactivité, son implication, son

esprit critique et sa rigueur scientifique ont énormément contribué au développement de ce travail de thèse. J'ai vraiment apprécié son énergie positive, son écoute et sa compréhension tout au long de mes travaux de recherche. Donc aussi un GRAND merci.

Mes remerciements vont également à tous les membres du jury pour l'honneur qu'ils m'ont fait en acceptant d'évaluer mon travail. Je tiens donc à remercier Madame Marlène VILLANOVA-OLIVER, en sa qualité de rapporteur, pour sa relecture très attentive et minutieuse de mon manuscrit de thèse ainsi que pour ses remarques et suggestions très pertinentes et constructives. Je remercie également chaleureusement Monsieur Rodolphe DEVILLERS, en sa qualité de rapporteur, pour avoir consacré du temps à la lecture du manuscrit ainsi que pour ses précieuses remarques et suggestions. Je tiens également à remercier sincèrement Madame Géraldine DEL MONDO, en sa qualité d'examinatrice, qui a pu accompagner l'avancement de ce sujet de thèse en tant que membre de ma comité de thèse, pour tout son intérêt et son implication ainsi que pour ses conseils, remarques et suggestions constructives et pertinentes. Je voudrais également exprimer ma gratitude à Madame Thérèse LIBOUREL et Monsieur Laurent NANA en leur qualité d'examineurs qui ont montré de l'intérêt pour mon travail, par leur lecture, leur écoute et leur présence à la soutenance. Je tiens également à remercier Monsieur Pierre HALLOT en sa qualité de membre du comité de thèse pour le temps consacré à l'évaluation de mon travail ainsi que pour ses remarquables commentaires et suggestions.

Je tiens à exprimer ma sincère gratitude à l'équipe exceptionnelle du laboratoire IRENav de l'Ecole Navale. Je tiens à remercier tout particulièrement Monsieur Abdel Boudraa, Monsieur Eric Saux et Monsieur Joseph Poupin de leurs précieux aides et conseils. Ils ont chacun rendu mon expérience à l'IRENav plus positive, plus productive et plus agréable.

Ma gratitude va également à tout le personnel de l'École Navale qui travaille afin de fournir les conditions nécessaires et favorables pour finaliser les différents projets. Je remercie également le personnel du secrétariat de l'EDSML qui accompagne les doctorants dans leurs démarches administratives, en particulier Madame Elisabeth Bondu et Madame Elodie Benon pour leur gentillesse, leur patience et leur disponibilité.

Je tiens à remercier tous mes collègues doctorants et post-doctorants, avec qui j'ai partagé de bons moments et qui se reconnaîtront, pour leur bonne humeur quotidienne et leurs encouragements. Une reconnaissance particulière à mon collègue et ami Cyril Carré pour ses multiples relectures

de mon manuscrit de thèse, dont les précieuses remarques, conseils en matière scientifique et technique m'ont beaucoup aidé. Je le remercie tout particulièrement sur le plan humain pour son sourire, ses encouragements et son soutien moral tout au long de ma thèse.

Je tiens à remercier mes amis qui m'ont aidé et m'ont beaucoup encouragé durant ce projet. Je pense à Salah Elani (le grand frère et l'homme de la situation), Hadj bay Ahmed (le penseur exceptionnel), Joseph moubogha (l'artiste réjouissant), Asma Rejeb (la flamme éclatante d'énergie positive), Karim karrout (esprit d'aventure et recherche de la vérité), Sirine Laajili (sérénité et gentillesse), Brahim Maylal (l'humoriste rationnel), Hicham Amar (collaborateur dévoué), Abdelatif El Hajbi (le grand diplomate), Bilal Ziane (Big boss), Fatiha Mohammed Aarab (générosité et gentillesse), Wissame Laddada (la classe!), Loic salamon (notre grand frère attentionné), Ludovic Moncla (scientifique brillant et joyeux), Fujino Iwao (sagesse et ouverture d'esprit), José Duarte (soldat de la science), Pierre Leffondré (chercheur hors standard), Pierre Mouneyrou (le pédagogue progressiste), Dan Feldman (Philosophe humaniste).

Un immense et spéciale remerciement à ma mère à qui je dédie spécialement cette thèse. Son amour et son soutien inconditionnel, sa bonne humeur et sa clairvoyance était et reste toujours pour moi une motivation et énergie motrice énorme et quoique je fasse ou quoique j'écris je ne saurai jamais comment la remercier autant. De même, je serais éternellement reconnaissant à mon père qui m'a toujours soutenue et qui a réussi avec succès à me transmettre l'amour de la science et de la connaissance et qui restera pour moi un père exemplaire. Je remercie aussi mon très cher frère Ismail dont son encouragement, ses réflexions et ses conseils m'ont beaucoup aidé et soutenu durant cette thèse. Sans oublier bien évidemment, je ne remercierai jamais assez mon oncle Mustapha Helaline pour ses nombreux conseils et et pour son partage d'expérience et sa bonne humeur. Aussi, je remercie tous les membres de ma grande famille, particulièrement je pense à : Rabiia, Simohammed, Gasem, Najib pour leur amour et soutien.

Enfin, je tiens à remercier chaleureusement Monsieur Antony Galton, le brillant chercheur dont les travaux de recherche ont été pour moi une grande base et une grande source d'inspiration dont j'ai principalement fondé et construit les contributions de cette recherche. Je voudrais également remercier Monsieur Mohammed Shahrour, l'intellectuel et l'ingénieux penseur dont les écrits philosophiques m'ont également inspiré et, surtout, ont aiguisé mon sens méthodologique et mes qualités de raisonnement.

Table des matières

Remerciements	v
Liste des figures	ix
Liste des tableaux	xii
1 Introduction	1
1.1 Contexte	1
1.2 Problématique et motivation	3
1.3 Questionnements et objectifs scientifiques	5
1.4 Structure de la thèse	7
2 Etat de l'Art	9
2.1 Représentation de l'espace	9
2.1.1 Représentation discrète : modèle objet	11
2.1.2 Représentation continue : modèle champ	13
2.2 Modèles de SIG	14
2.2.1 Modèle vecteur	14
2.2.2 Modèle Tessellation	15
2.2.3 Modèle SIG pour représenter un champ continu	17
2.3 Représentation hybride de l'espace géographique	18
2.4 Représentation du temps	21
2.4.1 Caractérisation de la nature du temps :	21
2.4.2 Dichotomie entre objet temporel et champ temporel	23
2.5 Notions temporelles de champ et d'objet	24
2.5.1 Intégration du temps dans la notion de champ	24
2.5.2 Intégration du temps dans la vision objet	25
2.6 Événement et Processus	27
2.7 Espace et Temps en SIG	28
2.8 Modèles de données spatiaux	30
2.8.1 Modélisation spatiale	30

2.8.2	Modèle spatial de base de données	31
2.9	Vers des modèles spatio-temporels	33
2.9.1	Three Domain Model	33
2.9.2	Modèles basés sur l'approche continue	34
2.9.3	Modèles orientés événement et processus	37
2.9.4	Modèles spatio-temporels basés sur un langage concep- tuel	39
2.9.5	Modèles basés sur une approche de graphe :	41
2.9.6	Approche "Objet mobile"	43
2.9.7	Analyse et critique globale des modèles	44
2.10	Vers des modèles de données hybrides	46
2.10.1	Approche basée sur l'intégration par primitive	46
2.10.2	Approche basée sur une intégration à travers un concept abstrait	51
2.10.3	Approche basée sur l'association des visions champ et objets	54
2.11	Évaluation et analyse des modèles hybrides	54
3	Vers une représentation hybride de l'espace géographique	59
3.1	Concepts spatio-temporels principaux	60
3.2	Définition d'un champ spatio-temporel	61
3.3	Définition d'un objet spatio-temporel	62
3.3.1	Continuant versus occurrent	62
3.3.2	Relations spatiales et spatio-temporelles	63
3.4	Vision hybride du temps	65
3.4.1	Objet-champ temporel	66
3.4.2	Champ-objet temporel	66
3.5	Vers une représentation spatio-temporelle Hybride	67
3.5.1	Association des deux visions de l'espace <i>field</i> et objet à travers le temps	67
3.5.2	Vision Hybride : <i>Field</i> vers objet	68
3.5.3	Vision Hybride : Objet vers <i>field</i>	70
3.5.4	Vision Hybride : Évolution du <i>field</i> interne d'un objet	75
3.6	Schéma conceptuel pour des requêtes spatio-temporelles hy- brides.	80
3.7	Conclusion et synthèse	82
4	Vers un modèle spatio-temporel hybride	85
4.1	Modèle de données hybride : vue globale	86
4.2	Modèle Conceptuel	87
4.2.1	Modèle spatio-temporel basé sur la vision de champ	89

4.2.2	Modèle spatio-temporel basé sur la vision objet	91
4.2.3	Modèle conceptuel hybride	92
4.3	Modèle logique pour représenter l'approche hybride	94
4.3.1	Types de données pour le modèle hybride	94
4.3.2	Fonctions hybrides	96
4.3.3	Opération composées temporelles, spatiales et spatio-temporelles	100
4.4	Modèle d'implémentation	102
4.4.1	Architecture et plateforme	103
4.4.2	Schémas de base de données objet-relationnel hybride	105
4.5	Conclusion et synthèse	110
5	Experimentation et Application	112
5.1	Méthodologie générale de l'expérimentation	113
5.2	Contexte et zone d'étude	114
5.3	Aperçu de l'ensemble des données	115
5.4	Initialisation du modèle de données	116
5.5	Expression et visualisation des requêtes hybrides	117
5.5.1	Identification des changements champ et objet	118
5.5.2	Analyse des capacités d'interrogation des modèles de données hybrides	119
5.6	Évaluation et Analyse des performances du modèle de données hybrides	128
5.6.1	Préparation de l'expérimentation	131
5.7	Conclusion et synthèse	135
	Conclusion générale	136
	Synthèse de la contribution	136
	Discussion et analyse critique des réponses apportées aux questionnements scientifiques	140
	Extensions de recherche et perspectives applicatives	142
	Bibliographie	145

Liste des figures

1.1	Positionnement de la thèse par rapport à l'intersection de trois domaines scientifiques principaux	6
2.1	Illustration de la relation entre les notions de domaine, codomaine et image de la fonction f qui représente la notion de champ.(Worboys and Duckham, 2004)	14
2.2	Différentes stratégies pour la représentation numérique d'un champ (Longley et al., 2015)	18
2.3	Algèbre pour les relations temporelles entre deux intervalles (Allen, 1983)	22
2.4	Shéma conceptuel du three-domain-model (Yuan, 1994)	34
2.5	La structure du modèle ESTDM est constituée d'un en-tête (header), d'une carte initiale (base map) et d'une liste d'événements (event list), l'événement étant lui-même formé de composants (components)	35
2.6	Diagramme conceptuel montrant la relation de hiérarchie entre les trois classes d'objets : ST object class, ST Dataset Class et ST-Element Class	37
2.7	Le modèle GEM : objets, événements et leurs interactions (Worboys and Hornsby, 2004)	38
2.8	Les trois types de comportements pour un objet temporel selon le modèle de Renolen (Renolen, 1997)	41
2.9	Modèle logique montrant comment les données se sont organisées autour de state, processus et événement Yuan (2001)	47
2.10	Modèle de données logique pour représenter les concepts de zone, processus et événement (McIntosh and Yuan, 2005)	48
2.11	Structure de données du modèle EDGIS (Pultar et al., 2010)	50
2.12	Modèle de données conceptuel pour l'intégration de l'approche hybride (De Bakker, 2017)	53

3.1	Figure montrant l'association d'un continuant c avec une collection de localisations spatiales L_i	68
3.2	Différence entre la fonction Ct_{fh} (a) et Ct_{fs} (b)	72
3.3	Figure illustrant le suivi de la variation interne d'un objet en l'associant avec une partition de l'espace champ qui décrit la surface interne de l'objet à chaque temps.	77
3.4	Dérivation des fonctions hybrides issue de l'association bidirectionnelles entre des objets et des champs à travers le temps	83
4.1	Intégration de la représentation hybride selon un processus général de modélisation des données	87
4.2	Modèle conceptuel représentant la vision field.	90
4.3	Modèle conceptuel représentant la vision objet.	92
4.4	Modèle conceptuel hybride représentant l'association entre les deux visions field et objet.	93
4.5	L'ensemble des fonctions spatio-temporelles ordinaires et hybrides associées au deux types de données Ctn et Ocr	95
4.6	Structure des type de données $LocSet$	97
4.7	Implémentation du modèle de données hybride avec une architecture basée sur PostgreSQL	104
4.8	Instanciation du modèle champ	106
4.9	Instanciation du modèle objet	107
4.10	Instantiation de la struture de donnée de type $LocSet$ au sein de la table $CState$	107
4.11	Schéma générique de la base de données permettant d'implémenter le modèle hybride	108
5.1	Méthodologie d'expérimentation du modèle hybride	114
5.2	zone d'étude	115
5.3	Données de test	116
5.4	Les quatre dunes les plus déplacées où les fonds marins ont subi une érosion supérieure à 0,7 m (localisées avec des points rouges)	121
5.5	Pour les localisations où l'érosion est supérieure à 1 m (situés avec des points rouges), identification les deux paires de dunes qui se sont le plus rapprochées (i.e., la dune E avec la dune L et la dune A avec la dune L). Les flèches indiquent les directions de déplacement.	123
5.6	Localisation et amplitude des variations de profondeur du fond marin entre t_1 et t_3 , associées à la présence de la dune dont la superficie a le plus évoluée.	124

5.7	Localisation et amplitude des variations de profondeur du fond marin entre t_1 et t_3 , associées à la présence de la paire de dunes ayant le taux d'éloignement le plus élevé.	125
5.8	variation de profondeur et vecteur de déplacement à la surface de la dune dont le taux de déplacement est le plus important entre t_1 et t_2	127
5.9	Carte illustrant la variation de la profondeur du fond marin au sein des emprises spatiale de la paire des dunes ayant eu le taux de rapprochement le plus élevé entre deux instants t_1 et t_3 .129	
5.10	Évolution de la taille du stockage de la table <i>FieldState</i> en fonction de l'attribut <i>field_value</i>	130
5.11	Graphes mettant en évidence l'évolution du temps de calcul en fonction du nombre de localisations spatio-temporelles des requêtes qui utilisent la fonction Ft_c	133
5.12	Graphiques indiquant l'évolution du temps de calcul en fonction du nombre de localisations spatio-temporelles des requêtes hybrides utilisant la fonction Ct_{fh}	133

Liste des tableaux

2.1	Tableau comparatif évaluant les modèles hybrides selon les critères sémantiques	56
2.2	Tableau comparatif évaluant les modèles hybrides selon les critères non fonctionnels	57
3.1	Schéma conceptuel montrant comment chaque fonction hybride pourrait être utilisée pour exprimer des requêtes hybrides.	81

Abbreviations

-
ADT : Abstract Data Type
API : Application Programming Interface
CRI : Cadre de Référence Interne
CRM : Cadre de Référence Mobile
HIQ : Hybrid Interaction Query
IDW : Inverse Distance weighted
MBD : Modèle de Base de Données
MDG : Modèle de Données Géographique
SGBD : Système de Gestion de Base de Données
SIG : Système d'Information Géographique
STP : Spatio-Temporal Point
UML : Unified Modeling Language

Chapitre 1

Introduction

1.1 Contexte

Grâce au développement permanent de la technologie des capteurs géo-localisés et des infrastructures d’acquisition des données géo-localisées, les chercheurs en sciences environnementales et urbaines disposent de nouvelles possibilités pour étudier les phénomènes géographiques à différentes échelles temporelles et spatiales. En effet, le large éventail de données géographiques désormais disponibles offre de multiples possibilités d’exploration des données pour une meilleure compréhension de la nature de la dynamique des processus environnementaux sous-jacents. Face à de telles opportunités, des défis majeurs se manifestent et soulèvent en particulier des questions liées à la conceptualisation et à la représentation informatique de l’espace-temps. Ces interrogations sont un sujet d’étude majeur des Sciences de l’Information Géographique, celles-ci étant définies comme un domaine de recherche qui tend à redéfinir les concepts géographiques et leur utilisation et mise en oeuvre à partir des systèmes d’information géographique ([Mark et al., 2000](#)).

Parmi les thèmes de recherche qui ont bénéficié d’une attention particulière des chercheurs, l’intégration de la dimension temporelle au sein des systèmes d’information géographique est l’un des exemples les plus importants. Cette intégration du temps dans les SIGs a constitué un sujet de recherche depuis la fin des années 1980 et est encore en plein développement. Cela a débuté avec ([Peuquet, 1984](#)) qui, en présentant une taxonomie et une étude approfondie des modèles de données spatiales, a clairement mentionné l’importance du temps dans les systèmes d’informations géographiques et la nécessité d’une meilleure intégration de la dimension temporelle. Toutefois, il nous faut attendre les travaux de ([Langran and Chrisman, 1988](#)) pour in-

roduire, pour une première fois, quelques concepts fondamentaux d'un SIG temporel. Parallèlement d'autres travaux ont été initiés afin d'intégrer le temps au sein des bases de données spatiales ([Armstrong, 1988](#)) ce qui, de fait, a constitué un domaine connexe de recherche orienté vers les bases de données spatio-temporelles. Le rôle des bases de données est essentiel du fait que ces dernières permettent de stocker, de gérer et de manipuler les données géographiques dans l'espace et le temps.

La conception en amont d'un modèle de données spatio-temporelles doit favoriser une meilleure représentation et donc compréhension des phénomènes géographiques dynamiques. Un modèle de données spatio-temporelles doit permettre de représenter divers types de changements à partir desquels un phénomène est observé et caractérisé. Un tel modèle devrait constituer le noyau d'un SIG temporel, permettant ainsi un stockage et une intégration efficace des données, et facilitant les processus d'analyse spatio-temporelle.

Un autre thème de recherche souvent abordé concerne le mode de représentation de l'information spatiale et leur abstraction au sein d'un système d'information. Les deux abstractions les plus prédominantes en sciences de l'information géographique sont les notions de champ et d'objet. L'approche basée sur les champs conçoit l'espace comme un ensemble infini de localisations dont chacune est associée à une ou plusieurs valeurs caractérisant une propriété (température, albédo, radiation solaire, etc.), tandis que la seconde conçoit l'espace comme un ensemble d'entités distinctes ayant chacune une identité, une étendue spatiale, des attributs et des relations avec d'autres entités ([Goodchild, 1989](#); [Coulélis, 1992](#); [Lardon et al., 1999](#)). Ces deux abstractions sont des conceptualisations génériques permettant de capturer et de modéliser la nature de différents types de phénomènes géographiques pour différents domaines. Le problème principal posé par cette dichotomie d'approches, selon ([Worboys, 1994](#)) et ([Galton, 2003](#)), est relatif au fait que ces deux conceptualisations ont souvent été employées séparément, sans orientation claire pour en privilégier l'une ou l'autre.

Il est vrai qu'il existe un nombre important d'approches de conceptualisation de phénomènes géographiques pouvant être appréhendées par une seule vision, soit par le modèle objet (par exemple la gestion des parcelles cadastrales, la structure du réseau routier, les trajectoires des piétons, etc.), soit par le modèle de champ (par exemple la distribution spatiale de la température, l'indice de végétation et la concentration de CO₂, etc.). Cependant, plusieurs phénomènes méritent néanmoins d'être approchés avec l'utilisation des deux modèles afin de les représenter plus complètement. C'est le cas en particulier lorsque la définition de l'objet est contrainte par la nature d'un champ de valeurs (une montagne, une dune, une zone de température), ou par un champ de valeurs lié à la présence d'un objet (température liée à la présence d'un

nuage, vitesse de vent liée à la présence d'un ouragan). Il apparaît de fait que ces deux approches et conceptualisations de l'espace ne sont donc pas mutuellement exclusives mais plutôt complémentaires (Worboys, 1994). Ramenée à la problématique de la représentation de la dimension temporelle et de la dynamique des espaces géographiques, l'abstraction de l'espace en champs et/ou objets, introduit de nombreuses questions de recherche sur l'identité, la filiation, et l'évolution des relations spatiales au court du temps (Hornsby and Egenhofer, 2000; Del Mondo et al., 2010). Les caractéristiques géographiques qui composent ces phénomènes peuvent être perçues comme des continnants (i.e., des objets dont l'existence dure dans le temps) ou comme des occur-rents (i.e., des objets dont l'existence est limitée dans le temps) (Galton, 2004; Grenon and Smith, 2004).

Parvenir à une conceptualisation qui combine le modèle de champ avec celui de l'objet devrait permettre d'élargir le processus d'abstraction et de généralisation de l'information à représenter et ainsi de pouvoir aborder des phénomènes géographiques souvent identifiés comme des phénomènes complexes (Yuan, 2001), distribués (McIntosh and Yuan, 2005) ou multi-aspects (Galton, 2004).

1.2 Problématique et motivation

Le défi de cette thèse consiste à élaborer un modèle générique permettant de gérer la nature des différents phénomènes géographiques. Cette perspective est déjà identifiée comme l'une des majeures directions de recherche encore à aborder dans le domaine des sciences de l'information géographique (Claramunt and Stewart, 2015). Cette problématique est articulée autour de la question suivante : dans quelle mesure les concepts, modèles et cadres de raisonnement spatio-temporels développés à ce stade sont-ils généralisables à tous les domaines d'applications des sciences de l'information géographique ?

La réponse à cette question dépend principalement du niveau de généralisation des abstractions conceptuelles selon lesquelles l'espace est représenté et le cadre applicatif considéré.

En reconnaissant que la réalité est plus complexe que les simples représentations de l'espace basées sur cette dichotomie objet/champ, plusieurs recherches se sont orientées vers une vision hybride pouvant permettre une meilleure association conceptuelle entre ces deux modèles de l'espace. Une première notion clé proposée est celle d'un objet-champ défini comme un champ dans lequel chaque position est associée non seulement à une valeur, mais également à un type d'objet géographique sous-jacent (Galton, 2001a).

Un deuxième concept clé, et toujours dual, est celui de champ-objet, défini par (Yuan, 2001; Goodchild and Yuan, 2007) comme un objet géographique avec une hétérogénéité interne conceptualisée comme un champ. Bien que les visions de champ et d'objet étaient bien appréhendées par ces modèles, l'ambition principale de ces travaux n'est pas la recherche d'une conceptualisation hybride de l'espace, mais plutôt l'identification formelle de correspondances conceptuelles entre les deux approches. De plus, ces modèles sont essentiellement statiques et, jusqu'à présent, la dimension temporelle n'a été que partiellement explorée.

Afin de prendre en compte une approche hybride intégrant le temps, une série de travaux successifs ont abordé l'intérêt de combiner la vision du champ et celle de l'objet dans un même modèle spatio-temporel. Le plus souvent des implémentations préliminaires illustrent une partie du potentiel d'une telle représentation au niveau conceptuel. Cependant, pour la plupart des modèles proposés jusqu'à ce jour, il manque toujours une base théorique formelle établissant les concepts fondamentaux d'une telle approche hybride. En effet, les modèles actuellement proposés se sont concentrés sur un seul aspect de l'approche hybride, à savoir la représentation de la structure interne de type champ à l'intérieur des emprises spatio-temporelles des objets. Cet aspect est loin d'être suffisant pour mettre en relief tout le potentiel que l'approche hybride pourrait offrir aux scientifiques notamment lorsqu'elle est étendue avec la dimension temporelle. Ne pas le faire réduit la capacité de manipuler et d'interroger des modèles de données spatio-temporelles qualifiés d'hybrides face à des requêtes explorant différents aspects de l'interaction entre les deux représentations de champ et d'objet. Pour illustrer nos propos, nous proposons trois exemples de requêtes spatio-temporelles qualifiées d'hybrides :

- S'interroger sur le possible impact des conditions atmosphériques (i.e., la dynamique du champ) sur la résilience d'une réserve biologique (i.e., l'évolution des objets) au cours d'un phénomène d'ouragan (i.e., un événement). Dans ce cas la résilience de la réserve biologique se mesure en caractérisant le changement de son emprise spatiale.
- S'interroger sur les possibles conséquences de l'évolution d'un champ de dunes marines (i.e., la dynamique des objets) sur la variation temporelle de la profondeur locale du fond marin (i.e., la dynamique du champ).
- S'interroger sur la variation spatio-temporelle de la pente (i.e., la dynamique du champ) au sein de l'emprise spatiale d'une dune

marine (i.e., la dynamique des objets) lors de ses déplacements.

Dans l'ensemble, ces exemples mettent l'accent sur les interactions possibles entre les dynamiques de type champ et celles de type objet. Ces requêtes spatio-temporelles montrent également la manière dont les concepts de continuant et celui d'occurrent devraient être pris en compte de manière implicite, notamment pour la manipulation d'objet dans le temps.

La problématique de cette thèse consiste à élaborer un modèle spatio-temporel qui pourrait transcender ces dichotomies, c'est à dire à la fois les prendre en compte par la modélisation mais aussi considérer l'ensemble des interconnexions qu'elles impliquent. Une telle approche devrait permettre de fournir une représentation duale de l'évolution de divers phénomènes géographiques complexes.

1.3 Questionnements et objectifs scientifiques

La problématique à laquelle nous sommes confrontés nous amène à poser les questions scientifiques suivantes :

- Comment formaliser une représentation hybride de l'espace géographique et quelle est son apport pour une modélisation spatio-temporelle au niveau conceptuel ?
- Comment et sous quelle forme logique peut-on appréhender et intégrer une telle approche hybride dans un SIG temporel opérationnel et quels seront les modes et langages de manipulation qui en tireront tout le bénéfice ?
- Quelle est la contribution concrète d'une telle démarche intégrée de conceptualisation, de représentation et de traitement, à l'amélioration des outils et méthodes numériques pour la compréhension des phénomènes géographiques complexes par les sciences environnementale et urbaines ?

Les objectifs concrets de cette thèse peuvent être reformulés par les objectifs suivants :

- Élaborer un modèle spatio-temporel qui puisse bénéficier simul-

tanément des avantages respectifs des approches continues et discrètes de l'espace et du temps. De plus, ce modèle doit être établi à partir d'une formalisation duale/hybride ;

- Concevoir ce modèle avec des choix logiques performants en conservant toute sa sémantique et son pouvoir d'expression ;
- Implémenter le modèle avec les fonctions et opérateurs qui en dépendent ;
- Montrer les avantages d'un tel modèle pour une application concernant un phénomène géographique complexe, illustrant ses capacités de représentation et de traitement et d'analyse : observation de la dynamique des fonds marins littoraux et des dunes sous-marines associées.

Au niveau du positionnement, notre recherche est orientée vers la représentation, la gestion et la manipulation des données spatio-temporelles. Les domaines de recherche concernés sont présentés dans la figure 1.1.

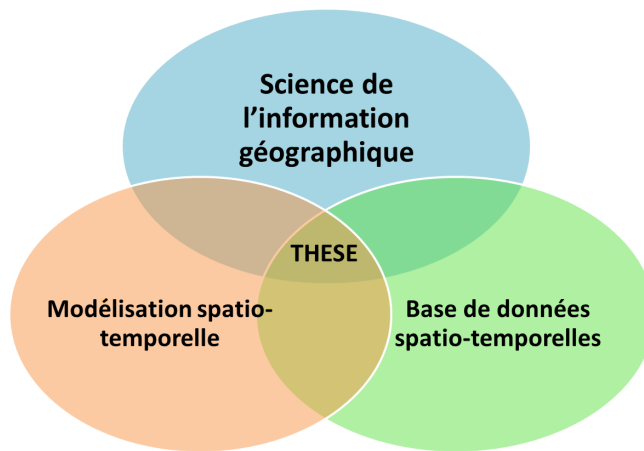


FIGURE 1.1 – Positionnement de la thèse par rapport à l'intersection de trois domaines scientifiques principaux

Les sciences de l'information géographique apportent des concepts et des mécanismes qui s'appuient sur les systèmes d'information géographique (SIG) comme outils opérationnels. Ces sciences de l'information géographique ont pour objectif de faire progresser notre compréhension des phénomènes géographiques en améliorant le cadre théorique, méthodologique et technologique. Quant au domaine de la modélisation spatio-temporelle, il nous four-

nira les abstractions spatiales et temporelles que nous pourrions utiliser pour construire notre modèle formel. Enfin, nous faisons appel au domaine des bases de données spatio-temporelles dans la mesure où notre modèle hybride sera dédié à la gestion, la manipulation des données spatio-temporelles et la catégorisation et l'identification de fonctions d'interrogation et d'analyse.

1.4 Structure de la thèse

Le document de thèse est structuré selon les quatre chapitres suivants :

- Chapitre 2 : ce chapitre présente un état de l'art dans le domaine des sciences de l'information géographique à travers les concepts de champ, d'objet et de temps. Un ensemble de modèles spatio-temporels est présenté, avec une comparaison plus directement orientée vers les modèles spatio-temporels hybrides.
- Chapitre 3 : ce chapitre procède à la formalisation d'une représentation hybride de l'espace géographique prenant en compte l'intégration de la dimension temporelle. Cette démarche de modélisation est illustrée par des exemples orientés vers des phénomènes d'évolution des dunes dans un environnement côtier.
- Chapitre 4 : ce chapitre élabore un modèle de données de SIG temporel qui offre une représentation hybride de l'espace basée sur les notions d'objet et de champ en s'affranchissant d'une dichotomie partitionnant les notions de continuant et d'occurrent. Il présente essentiellement les types de données, les fonctions hybrides de manipulation et des opérations d'analyse spatio-temporelle nécessaires pour la construction d'un langage de manipulation des données hybrides. Ce modèle de données est mis en œuvre à partir d'une architecture basée essentiellement sur le système de base de données objet relationnel libre PostgreSQL et son extension spatiale PostGIS.
- Chapitre 5 : ce chapitre présente une illustration et implémentation d'un modèle de données spatio-temporel hybride, en permettant l'expression d'une série de fonctions hybrides à partir de requêtes SQL intégrées qui montrent le potentiel de l'approche globale de modélisation et de manipulation des données. L'accent est mis sur les requêtes spatio-temporelles décrivant une bonne interaction entre la dynamique de l'espace appréhendé selon une vision champ

et celle d'une perception objet afin de globalement étudier un phénomène dynamique de géomorphologie marine. Une série de tests de performance et d'évolution du système est finalement présentée et montre comment le modèle de données permet, dans son ensemble, l'exécution de différentes requêtes spatio-temporelles hybrides.

Nous concluons l'ensemble de la thèse et soulignons l'apport, les contributions et les limites de ce travail de recherche. Enfin, nous présenterons les perspectives de recherche offertes par notre modèle et son approche hybride pour la modélisation d'un espace géographique.

Chapitre 2

Etat de l'Art

2.1 Représentation de l'espace

"Space and time are the framework within which the mind is constrained to construct its experience of reality."

Immanuel Kant

Du point de vue des sciences de l'information géographique l'espace peut être considéré comme un référentiel ou un conteneur dans lequel les objets existent et les événements se déroulent ([Worboys and Duckham, 2004](#)). ([Chrisman, 1978](#)) souligne que l'espace peut être dual en terme de représentation, car ce dernier peut être perçu soit sous l'angle d'une vision absolue, en le considérant comme un ensemble infini de points stationnaires possédant des propriétés que l'on peut mesurer (ex : température, albédo, humidité du sol), soit sous l'angle d'une vision relative constituée d'objets en relation ayant des propriétés spatiales. Cette vision distinctive de l'espace, a un impact fondamental dans les démarches de conceptualisation de l'espace géographique. Celui-ci peut être plus précisément défini selon ([Worboys and Duckham, 2004](#)) comme 1) une collection de localisations (i.e., points localisés dans un espace géographique terrestre) où des phénomènes sont mesurables ou 2) un ensemble d'entités géographiques référencées comme une région terrestre.

Au-delà de la dualité de ces représentations, il apparaît intuitivement qu'il pourrait être pertinent de concevoir une forme conceptuelle atomique unifiée, à partir de laquelle ces deux définitions duales de l'espace pourraient être spécialisées. Dans les travaux de ([Goodchild, 2001](#); [Goodchild and Yuan,](#)

2007), cette forme et abstraction est conceptualisée à partir de la notion de *geo-atom*, définie comme un concept localisé dans l'espace-temps possédant différentes propriétés que l'on peut mesurer. Ce concept constitue une sorte d'abstraction de haut niveau de l'espace géographique.

Sans considérer la composante temporelle, l'espace géographique S peut être représenté comme un sous-ensemble d'un espace Euclidien à trois dimensions ($S \subset \mathbb{R}^3$). Si la variable temporelle est considérée, il faudra logiquement une dimension additionnelle. Cependant, comme l'indiquent (Worboys and Duckham, 2004) et (Liu et al., 2008), l'espace géographique tel qu'il est conceptualisé actuellement est souvent réduit à une bi-dimensionnalité, pour des raisons de facilité de représentation et de gestion dans de nombreuses applications.

Plus concrètement, la dimension spatiale de l'espace géographique, selon (Goodchild and Yuan, 2007), comprend clairement la surface de la Terre, de toutes ses composantes terrestres, aquatiques ainsi que sa composante atmosphérique, de manière à inclure les domaines de l'hydrogéologie, de la géologie, de l'océanographie et des sciences de l'atmosphère. La dimension temporelle, elle, nous renvoie dans le passé, ou nous projette vers le futur pour des applications de planification ou de prédiction.

Afin d'assurer la représentation et la gestion pratique des phénomènes géographiques au sein des modèles de SIG, ces derniers doivent définir des échelles de représentation liées au niveau de détail représenté par les données (Longley et al., 2005). Ainsi, pour les SIGs actuels et dans la plupart des applications, le seuil le plus fin de représentation spatiale (i.e., résolution spatiale) n'est que rarement inférieur à quelques centimètres (Goodchild and Yuan, 2007). Cela n'exclut pas la possibilité d'aller à une échelle beaucoup plus fine, toutefois cela nécessitera d'adopter ou de changer certains paradigmes spatiaux, parmi lesquels on pourra, par exemple, faire face à la nécessité de définir de nouveaux référentiels.

Du point de vue des sciences de l'information géographique, deux modèles de représentation distincts s'imposent dans les applications de SIG selon les caractéristiques et la problématique traitée (Couclelis, 1992; Peuquet, 1994; Galton, 2004). Ces deux modèles sont, d'une part, le modèle objet qui représente les caractéristiques discrètes du domaine géographique et, d'autre part, le modèle champ qui représente les caractéristiques continues de ce même domaine géographique. Bien que dans la plupart des cas le contexte de l'application permette de faire un choix entre les deux conceptualisations. Dans certaines situations, la distinction n'est pas aussi claire et les modèles objet et champ sont tous deux nécessaires notamment pour développer des analyses spatiales (Galton, 2001a; Cova, 2002; Galton, 2003; Cova, 2016). C'est le cas par exemple pour la représentation des conditions

météorologiques ou de la dynamique de la morphologie des terrains naturels.

2.1.1 Représentation discrète : modèle objet

(Couclelis, 1992) définit la notion objet d'un espace géographique comme un référentiel peuplé par des entités ou des objets ayant des caractéristiques spatiales. Nous pourrions donner une image de cette conceptualisation comme celle d'un conteneur peuplé d'objets géographiques, chacun ayant une identité, une emprise spatiale avec des limites plus ou moins bien définies, des attributs, des relations avec son environnement et un comportement. Cette vue est très compatible avec le langage naturel, car la réalité géographique est généralement décrite à l'aide d'objets. Cette approche matérialise des entités que l'on observe dans la nature et qui constituent des traits ou des formes du terrain, telles que les montagnes, les rivières et les lacs ainsi que des entités liées aux activités artificielles de l'homme, telles que les routes, les bâtiments et les villes. L'observateur humain a toujours une tendance inhérente à attribuer une identité et une composante spatiale aux caractéristiques géographiques. L'exemple suivant : "un front froid (objet) traverse le désert (objet) et se dirige vers les montagnes (objet)" offre une description d'un phénomène météorologique, et illustre bien cette tendance (Longley et al., 2005).

La composante spatiale des objets géographiques peut être caractérisée et formalisée dans les SIGs par sa dimension topologique, c'est-à-dire une collection connectée de points formant des entités linéaires (une dimension), des entités surfaciques (deux dimensions) ou des entités volumiques (trois dimensions) (Galton, 2001a; Liu et al., 2008). Dans la réalité, tous les objets qui sont perceptibles par l'homme sont en trois dimensions, et leur représentation en deux dimensions ne constitue qu'une approximation. (Del Mondo, 2011) désigne cette composante approximation sous le nom d'empreinte spatiale, qu'elle définit comme la partie de l'espace qui lui est directement associée, et caractérisée en premier lieu par sa forme géométrique ainsi que par son positionnement dans l'espace basé sur un système de coordonnées (géographique, Lambert, UTM, etc.), et en second lieu par l'ensemble des relations spatiales qu'elle peut former avec d'autres entités. Ces relations sont classifiées selon trois types : relations topologiques (Egenhofer et al., 1989; Egenhofer and Herring, 1990), métriques (Pullar, 1988) ou d'orientation (Frank, 1996; Zimmermann and Freksa, 1996).

L'empreinte spatiale d'un objet donné pourrait être liée à des attributs spatialement intensifs ou bien à des attributs spatialement extensifs dont une certaine propagation se reflète potentiellement dans la division ou la fusion des objets (Zeiler, 1999). Prenons l'exemple d'une forme urbaine, un quartier

défini comme un objet ayant une empreinte spatiale. Lorsque ce quartier est subdivisé en deux unités plus petites, la valeur d'un attribut spatialement extensif comme la superficie doit être divisée et répartie entre ces deux unités, de telle sorte que la somme des deux valeurs de cet attribut liées aux deux sous-parties séparées du quartier, soit égale à la valeur totale de la superficie de l'ensemble, tandis qu'une valeur d'un attribut spatialement intensif peut persister et s'appliquer aux deux sous-parties du quartier. Un exemple classique est celui de la densité totale de population relative à chaque quartier qui pourra être variable dans les deux unités subdivisées, de telles contraintes posent la problématique de relation et de la propagation possible des attributs relationnels liés à des objets spatiaux (Claramunt and Mainguenaud, 1995).

Parmi les caractéristiques les plus importantes d'une empreinte spatiale, ce sont ses limites ou frontières qui définissent implicitement la portion de l'espace occupée. Selon (Smith and Varzi, 1997, 2000), la limite d'une emprise spatiale pourrait être *fiat* ou *bona fide*. Les limites dites *fiat* sont introduites par la perception humaine, ou autrement dit, elles sont des produits artificiels d'actes humains cognitifs. L'exemple le plus classique concerne les divisions administratives comme les secteurs, les quartiers, etc. D'un autre côté, les limites dites *bona fide* concernent les limites qui se sont fondées sur des caractéristiques de la réalité physique et qui reflètent une certaine forme de cohésion interne ou d'homogénéité : des bâtiments, un lac, etc. (Smith and Varzi, 1997, 2000).

Selon (Goodchild and Yuan, 2007), dans le cas des limites dites *bona fide*, les règles définissant l'appartenance à un ensemble de points constitutifs de l'emprise spatiale peuvent être complexes et vagues, comme par exemple à travers les règles définissant l'appartenance à des caractéristiques topographiques (ex : montagne, vallée) ou des caractéristiques météorologiques (ex : tempête violente, altitude atmosphérique). (Galton, 2004) souligne que cette distinction pourrait être remise en question par certains phénomènes ou cas problématiques. Dans ce sens, il cite l'exemple des limites spatiales d'une métropole qui agissent à la fois comme une démarcation humaine artificielle, mais en même temps comme l'appartenance à un ensemble de points constitutifs de son emprise spatiale, mais qui restent aussi vagues et dépendantes de plusieurs paramètres administratifs, sociologiques ou physiques.

Un dernier aspect qui concerne les limites spatiales d'un objet est dans certains cas son caractère indéterminé qui fait appel à la notion de flou ou d'imprécision. Selon (Burrough and Frank, 1996), cette notion peut être modélisée par une fonction d'appartenance partielle. Pour illustrer cette fonction, considérons un objet o et son ensemble de points constitutifs P_i . La fonction d'appartenance f_a consiste à associer à chaque point $p \in P_i$ une va-

leur $f_a(p)$ entre 0 et 1 représentant le degré d'appartenance à l'objet o . Dans le cas où la valeur de $f_a(p)$ est binaire, c'est-à-dire soit 1 ou 0, les limites de l'objet seront fermement déterminées.

2.1.2 Représentation continue : modèle champ

Dans un espace géographique, un nombre potentiellement infini de propriétés peut être identifié dans chaque position (température, pression, humidité, etc.). La relation ou l'association entre une position et sa valeur est représentée sous la forme d'une fonction qu'on appelle champ. Un champ peut être défini comme un ensemble d'observations dont chacune a été mesurée et associée à une composante géo-référencée.

(Galton, 2001a) définit formellement la notion de champ comme une fonction $f: L \rightarrow V$ qui relie l'ensemble L des points dans l'espace géographique avec l'ensemble V qui constitue les valeurs caractéristiques du champ. L'ensemble L constitue le domaine du champ (i.e., $dom(f)$) alors que l'ensemble V dénote le codomaine du f (i.e., $codom(f)$). Le sous-ensemble des valeurs v appartenant au codomaine V étant associés aux éléments du domaine L est appelé image de la fonction f dénotée comme $img(f)$. La figure 2.1 montre la relation entre les différents éléments d'un champ.

Si le résultat de l'application de la fonction f à l'élément l de L est v alors on écrit $v = f(l)$ ou $f: l \mapsto v$. L'ensemble des valeurs de l'image du champ f donne généralement des valeurs mesurées, calculées ou classifiées d'un ou des attributs (ex. la température, densité de population, occupation des sols, etc.). Ces valeurs peuvent être continues ou discrètes, ordonnées ou non ordonnées, finies ou infinies, numériques ou symboliques. Nous pouvons aussi faire face à un type spécial de champ qui est le champ Booléen dont l'ensemble de ses valeurs $v \in \{true, false\}$ sont binaires. Ces valeurs peuvent être aussi de type scalaire, vecteur (ex : la vitesse du courant marin ou du vent) ou tenseur (ex : stress hydrique, conductivité hydraulique dans les formations poreuses anisotropes).

Le domaine L du champ est considéré comme un ensemble infini, avec des éléments tri-dimensionnels si l'on considère les trois axes de l'espace, ou de quatre dimensions si l'on considère le temps $dim(dom(f)) = \mathbb{R}^4$. Ainsi, cette prise en compte de l'infini rend la notion de champ appropriée pour représenter les phénomènes continus dans un espace géographique. Cependant l'utilisation du champ se heurte aux limites de notre capacité à acquérir, enregistrer et traiter des données. Cela oblige à utiliser, dans la pratique, un domaine spatial discret constituant, soit une partition de l'espace en unités conjointes regroupant (ou pas) plusieurs éléments du champ d'origine, soit un échantillonnage plus ou moins fin de certains éléments du champ d'origine.

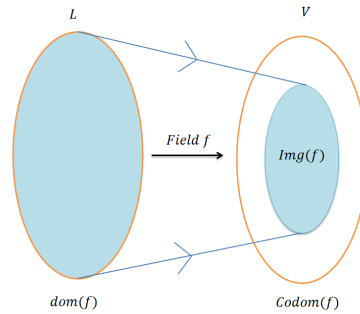


FIGURE 2.1 – Illustration de la relation entre les notions de domaine, codomaine et image de la fonction f qui représente la notion de champ. (Worboys and Duckham, 2004)

2.2 Modèles de SIG

Bien que les deux modèles champ et objet correspondent à des outils conceptuels représentant et modélisant les phénomènes géographiques, leur représentation numérique dans des systèmes logiciels opérationnels impose la prise en compte de plusieurs contraintes (Cova, 2016). Ces dernières sont principalement fixées par les limites de nos représentations, notamment celles posées par les informations dites utiles et celles imposées par les limites de nos ordinateurs et systèmes logiciels. Ces limites imposent de passer par des représentations intermédiaires, et en particulier des modèles logiques et physiques de représentation de ces données qui prennent en compte le caractère fini et quantifié de la représentation numérique (Longley et al., 2005). Dans ce sens, nous distinguons trois modèles ou stratégies d'implémentation qui vont être présentés dans les sections suivantes : le modèle vecteur, le modèle basé sur une tessellation de l'espace et le modèle basé sur l'interpolation.

2.2.1 Modèle vecteur

Le modèle vecteur consiste à donner une représentation géométrique des éléments abstraits ou concrets du domaine géographique (Worboys and Duckham, 2004). Il se base surtout sur les trois primitives géométriques principales que sont le point, la polyligne, et le polygone :

- **Point** : c'est une entité de dimension nulle qui contient un couple de coordonnées (paire (x, y) dans un espace bidimensionnel voire un triplet avec le composant z dans un espace tridimensionnel). Le tuple de coordonnées associé à un point p_1 mesure la distance de ce

point par rapport à une origine dans la direction de chaque axe de référence. La collection de tous ces points constitue le plan cartésien, souvent dénoté comme $\mathbb{R}^2$ ou $\mathbb{R}^3$. Les points sont généralement utilisés pour modéliser des caractéristiques singulières et discrètes (des poteaux électriques, des localisations d'échantillons, des positions de mobiles, etc.) ou des abstractions ponctuelles de surfaces (bâtiments, agglomération, etc.) (Zeiler, 1999).

- **Polyligne** : Une polyligne est définie comme une séquence finie de segments $[m_1, m_2, \dots, m_k]$ de telle sorte que chaque point d'extrémité d'un segment p_i est partagé exclusivement par deux segments m_{i-1} et m_i , sauf éventuellement deux points terminaux représentant les extrémités de la polyligne (i.e., p_1 et p_n) (Worboys and Duckham, 2004). Les polygones sont utilisées pour représenter des entités linéaires telles que les routes, les ruisseaux, les failles, les limites, etc. Elles sont caractérisées principalement par la propriété de longueur.
- **Polygone** : Un polygone est une région du plan délimitée par une polyligne fermée, qui constitue ses limites. Ces limites sont représentées par une série de points ou de sommets $[p_1, p_2, \dots, p_{k+1}]$ reliés par une collection ordonnée de segments de lignes $[m_1, m_2, \dots, m_k]$. La condition est que le premier point p_1 sur le premier segment de ligne m_1 est identique au dernier point p_{k+1} sur le dernier segment de ligne m_k (Worboys and Duckham, 2004). Les polygones sont utilisés pour représenter des caractéristiques telles que les surfaces d'une agglomération, les formations géologiques, les lacs, les communautés végétales, etc. Les polygones ont les propriétés de surface et de périmètre.

2.2.2 Modèle Tessellation

Une tessellation consiste à diviser l'espace en cellules contiguës, dont les limites entre deux cellules sont déterministes (i.e., tranchantes). Selon la forme des cellules, la tessellation peut être classée en : (1) tessellation irrégulière ou (2) tessellation régulière (Goodchild and Yuan, 2007; Liu et al., 2008).

Une tessellation irrégulière est caractérisée par l'attribution d'une forme géométrique variable pour toutes les cellules composant l'espace. L'exemple typique qui tombe dans cette catégorie est le diagramme de Voronoï qui consiste à segmenter le plan en un ensemble de polygones de sorte que

chaque polygone convexe contient la région du plan la plus proche de chacun des points spatialement distribués (Gold and Condal, 1995). La figure 2.2-D illustre le cas où la limite entre les tesselles n'est pas liée à une forme géométrique précise mais à la frontière entre des espaces aux propriétés différentes.

Une tessellation régulière est caractérisée par l'attribution d'une forme géométrique identique et fixe en terme de taille pour toutes les cellules composant l'espace. Dans le cas où cette forme est rectangulaire et régulière, ce sera le modèle raster (Figure 2.2-C). Plus précisément, le modèle raster consiste en un assemblage ou une grille de cellules, souvent appelées pixels dont chacune de ces cellules est associée à une paire de coordonnées (x, y) , ainsi qu'une valeur d'une variable (ex. température).

Les modèles discutés (i.e., voronoï et raster) sont représentés en un espace de deux dimensions, cependant ils peuvent être étendus à des espaces de dimension supérieure. Par exemple, l'unité de base de la tessellation régulière tridimensionnelle est un voxel, tandis que des diagrammes de Voronoï tridimensionnels peuvent également être générés (Ledoux and Gold, 2004).

Bien qu'il soit courant d'associer une vue de champ à un modèle de données de tessellation (plus particulièrement le mode raster en raison de sa structure de données matricielles facilitant le stockage ainsi que la programmation d'algorithmes pour l'analyse spatiale.) et une vue d'objet à un modèle de données vecteur, les champs peuvent être également représentés à l'aide du modèle vecteur et les objets peuvent être représentés à l'aide du modèle raster (Longley et al., 2005). Il existe de nombreux exemples de représentation d'un champ à l'aide du modèle vecteur, comme dans une carte de classification des sols où chaque emplacement est associé à une classe, celle-ci étant définie sous forme d'un polygone (Goodchild and Yuan, 2007). Un exemple d'utilisation d'un raster pour représenter des objets correspond au modèle d'appartenance floue où à chaque cellule du raster est défini le degré d'appartenance de cette cellule à un objet, en utilisant une échelle de 0 à 1 (Cova, 2016). Généralement, nous pouvons dire que les modèles champ et objet sont indépendants de ces deux stratégies de représentations numériques.

Il apparaît également qu'une représentation numérique d'un champ par les modèles vecteurs et tessellations est complexe du point de vue computationnel et surtout représente l'espace comme un ensemble de composants géométriques (ponctuels ou polygonaux) ayant des propriétés uniques. L'image la plus connue montre les flancs d'une montagne non pas comme une pente plutôt régulière mais comme une succession de terrasses horizontales. De plus, pour une représentation d'un champ continu par le mode de tessellation, on perd tous les détails sur les variations spatiales au sein des tesselles, car généralement une seule valeur est attribuée pour chaque tesselle.

Les SIG implémentent d'autres stratégies qui se basent essentiellement sur les notions d'échantillonnage et d'interpolation. Cette dernière (i.e., interpolation) est définie comme le moyen d'estimer la valeur du champ pour des localisations où la valeur de la variable étudiée n'est pas enregistrée ([Goodchild, 2003](#)).

2.2.3 Modèle SIG pour représenter un champ continu

Pour pouvoir représenter la variation continue d'une variable dans le domaine spatial d'un champ, une deuxième méthode différente de la tessellation peut être adoptée. Il s'agit de sélectionner un ensemble fini de géométries de type points ou polygones dont les attributs sont connus, constituant la référence spatiale d'un champ, et de déterminer pour les autres points de l'espace l'attribut recherché par interpolation. Ces géométries (i.e., la référence spatiale) n'ont pas la vocation de représenter des entités spatiales dans la réalité, et si elles sont utilisées c'est seulement pour un objectif de représentation et de discrétisation du domaine géographique du champ ([Goodchild and Yuan, 2007](#); [Liu et al., 2008](#)).

Les images A, B, E et F de la Figure 2.2 montrent quatre types de référence spatiale. A) correspond à un échantillonnage sur un ensemble de points régulièrement espacés; B) montre des échantillons sur un ensemble de points distribués irrégulièrement; E) montre un échantillonnage sur un ensemble de points distribués irrégulièrement et structuré sous la forme d'un modèle triangulaire irrégulier (TIN) ([Franklin, 1973](#)). Dans ce dernier cas, les valeurs des points du champ sont estimées par interpolation spatiale polynomiale linéaire ou cubique en utilisant les valeurs des sommets du triangle ([Mitas and Mitasova, 1999](#)); enfin, F) représente un échantillonnage correspondant aux isohypses (courbes de niveau)

La disposition et le choix d'une méthode d'interpolation dans ce type de processus sont importants pour obtenir une information plus fiable. Par exemple la méthode d'inverse de distance pondéré (IDW) est basée sur le postulat que la valeur à un point non échantillonné peut être approximée par une moyenne pondérée de valeurs à des points qui sont spatialement proches ([Burrough, 1986](#); [Watson, 1992](#)). Selon ([Goodchild, 2004](#)) la majorité des méthodes d'interpolation repose sur la notion de distance. Ce lien est issu de la première loi de Tobler, dite la première loi de la géographie, qui stipule que "tout est lié à tout, mais les choses proches sont plus liées que les choses éloignées" ([Tobler, 1970](#)). Selon ([Goodchild, 2004](#)) sans cette loi et principe, toute interpolation spatiale ne serait pas fiable.

La différence entre les modèles basés sur l'interpolation et ceux basés

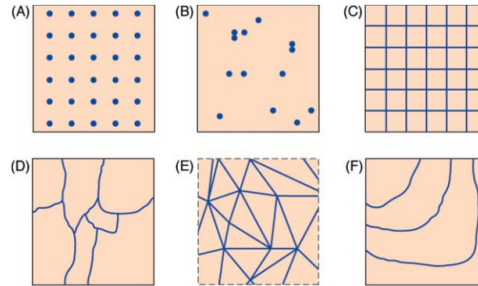


FIGURE 2.2 – Différentes stratégies pour la représentation numérique d'un champ (Longley et al., 2015)

sur la tessellation peut être résumée de la façon suivante : dans le cas d'une tessellation, pour déterminer la valeur d'un champ en un point x il faut identifier le tesselle qui contient x , et dans le cas d'une interpolation, il faut déterminer les échantillons de la référence spatiale à utiliser et une méthode d'interpolation qui doit être appliquée (Liu et al., 2008).

2.3 Représentation hybride de l'espace géographique

Selon (Galton, 2003), les modèles conceptuels champ et objet ont généralement été présentés par les sciences de l'information géographique, comme les pôles d'une dichotomie nette, avec une interrogation fréquente portant sur le choix de l'un de ces deux types de modèle. Il existe un nombre important de thématiques liées aux phénomènes géographiques qui peuvent être gérés par une seule vision, soit à travers le modèle objet (ex. gestion des parcelles, structure de réseau routier, trajectoires piétonnes, etc.), soit par un modèle champ (ex. distribution spatiale de la température, de l'indice de végétation et de la concentration de CO₂, etc.). Néanmoins, plusieurs phénomènes nécessitent l'emploi des deux modèles à la fois, afin de pouvoir les capturer convenablement. C'est le cas par exemple de la dynamique des incendies de forêt, des dunes côtières ou des ouragans (Claramunt et al., 1993; Yuan, 2001; Galton, 2003). Ces phénomènes contiennent des caractéristiques à la fois de type objet et de type champ et sont considérés comme complexes (Yuan, 2001) ou distribués (McIntosh and Yuan, 2005) ou multi-aspects (Galton, 2004). Prenons l'exemple de la dynamique des dunes côtières. Cette dernière peut être modélisée par une approche objet où une dune possède

des limites bien définies extraites de sa géomorphologie et de sa structure interne, une identité, et des propriétés (ex. hauteur maximale, sinuosité, direction de déplacement, etc.) (Thibaud et al., 2013). La même thématique peut être vue comme la dynamique d'une surface continue d'élévation où les dunes sont représentées par des variations du relief. Un autre exemple est celui d'un phénomène d'ouragan. En tant que champ, ce phénomène peut être capturé par une ou plusieurs propriétés continues dans l'espace (ex. pression barométrique, vitesse du vent, couverture nuageuse) sur une région donnée. D'un autre côté, en tant qu'objet, un ouragan sera représenté par son extension géométrique, une identité et des attributs (ex. la vitesse maximale du vent soutenu, la pression oculaire, la vitesse et la direction du mouvement).

Ces exemples montrent que cette dichotomie reste réductrice et insuffisante pour la compréhension d'un ensemble des processus géographiques pouvant décrire des phénomènes, d'autant plus que les deux conceptualisations *objet* et *champ* sont souvent disponibles conceptuellement. Cela nous amène à reconnaître que les deux modèles conceptuels *objet* et *champ* ne sont donc pas mutuellement exclusifs (Worboys, 1994). Selon (Galton, 2003) il devrait être possible de proposer des formes intermédiaires de représentation entre les deux modèles, et de mettre en place des opérations permettant d'enrichir l'information par des méthodes reliant objets et champs. (Goodchild and Yuan, 2007) ont qualifié ces méthodes intermédiaires en trois catégories :

- Méthodes d'estimation de la densité, qui créent une surface continue de densité (i.e., champ) à partir d'une collection d'objets discrets. A chaque position de l'espace, en appliquant par exemple un cercle d'expansion, on détermine le nombre d'objets contenus dans ce cercle divisé par sa surface ;
- Méthodes d'extraction d'objets qui identifient des objets discrets à partir d'analyse d'image, d'outils de segmentation et de reconnaissance de formes ;
- Méthodes d'identification de points et de lignes remarquables correspondant aux caractéristiques géométriques saillantes de la surface (ex. pics, creux, fossés, crêtes, etc.).

Ces méthodes peuvent être complétées par des notions connexes telles que l'approche floue, dans laquelle à chaque position est définie une fonction dont la valeur correspond au degré d'appartenance de cette position à un objet. Les valeurs possibles varient entre 0 (i.e., pas d'appartenance à l'objet) et 1 (i.e., appartenance complète dans l'objet) (Cova, 2016).

Comme précédemment mentionné, en reconnaissant que la réalité est plus complexe que les simples représentations de l'espace basées sur la dichotomie objet/champ, plusieurs recherches se sont orientées vers une vision hybride permettant une association conceptuelle entre ces deux modèles de l'espace, à partir de laquelle un ensemble de concepts clés ont été proposés. La première notion clé consiste en la notion d'**objet-champ** qui a été définie comme un champ dans lequel chaque position est associée non seulement à une valeur, mais également à un type d'objet géographique (Galton, 2001a). En conséquence, l'espace est "matérialisé" en régions qui sont homogènes du point de vue des propriétés du champ, et où tout ou partie de ces régions peuvent être identifiées comme des objets (Galton, 2003). Ainsi, l'association entre le domaine de champ et l'objet est conservée. Ce concept est formalisé comme une paire ordonnée (f, t_y) tel que t_y constitue un type d'objet (parcelle, nuage, montagne, etc.) et f est une fonction qui relie chaque position du domaine géographique à un et un seul objet o de type t_y . La condition est que l'ensemble des points p liés à l'objet o par la fonction f doivent forcément correspondre à l'emprise spatiale de o : $Loc(o)$.

(Cova, 2002) généralise davantage en définissant un objet-champ comme une relation binaire mathématique R_F qui associe un point avec un ou plusieurs objets. (Galton, 2001a) ajoute que l'objet associé à une position spécifique n'inclue pas obligatoirement cette position. Il donne l'exemple du champ de visibilité spatiale appelé "Viewshed", où chaque position d'observation est liée à un objet distant, en l'occurrence, la portion de l'espace géographique visible depuis cette position.

La seconde notion clé consiste en la notion de **champ-objet** définie par (Yuan, 2001; Goodchild and Yuan, 2007). Cette notion correspond à la conceptualisation de l'objet géographique et de sa surface interne ; celle-ci considérée, comme un champ. Par exemple, une dune côtière est définie en tant d'objet par son emprise géométrique, et en tant que champ par la variation d'altitude et de pentes qui y est associée. Selon les auteurs, l'évolution de la structure interne d'un champ-objet peut caractériser la dynamique. Ils prennent l'exemple d'une tempête convective où la variation interne de la vitesse des vents, d'intensité des précipitations, les variations de température et pression caractérisent à la fois la dynamique de la tempête et la façon avec laquelle elle peut évoluer spécialement (i.e., se déplacer et/ou changer de dimension) dans certaines conditions atmosphériques.

Le fait de pouvoir extraire des concepts permettant la combinaison des deux visions de l'espace a amené (Galton, 2003) à la conclusion suivante : *"les deux formes de conceptualisation champ et objet ne sont pas des systèmes monolithiques séparés, mais ils sont intimement liés par un réseau complexe*

d'interconnexions. Certains d'entre eux sont bien compris, mais d'autres ont à peine commencé à être explorés, même dans le cas purement statique". Ainsi Galton a signalé qu'il reste un effort de conceptualisation à réaliser qui consisterait à trier ce réseau d'interconnexions, et à concevoir des formes appropriées de représentation et de manipulation pour rendre compte de ces types de phénomènes complexes.

2.4 Représentation du temps

La nature du temps reste un sujet de préoccupation de longue date qui suscite de nombreuses discussions et débats dans de nombreuses traditions intellectuelles et philosophiques (Galton, 2016). Généralement, et à partir des réflexions menées par les philosophes et les scientifiques pour comprendre la nature du temps, deux principaux concepts sur le temps ont émergé. Premièrement, le temps physique qui est objectif et existe en dehors de la conscience humaine, et consiste en une véritable entité indépendante de la conscience faisant partie du monde réel. Deuxièmement, le temps psychologique qui est subjectif et dont l'existence dépend de la conscience humaine (Whitrow, 1989). Du point de vue ontologique et selon (Baumann et al., 2012), le temps recèle les deux aspects. D'une part, l'homme perçoit le temps en relation avec les entités matérielles à travers des effets de durée, de persistance, de déroulement, de non-simultanéité, d'ordre, de passé, présent et futur, ainsi le changement et de passage du temps. Ces effets pour (Baumann et al., 2012) sont considérés comme des phénomènes dépendants de la conscience. D'autre part, cela suppose que ces phénomènes temporels sont générés par les différentes dispositions et configurations que les entités matérielles possèdent. Ces différentes dispositions constituent la notion de temporalité.

En ce qui concerne les sciences de l'information géographique, le temps a souvent, mais pas toujours, été associé à l'espace, soit comme une extension de l'espace, soit par analogie avec lui (Galton, 2001b, 2004). La façon de modéliser, représenter et enregistrer le temps représente encore un défi fondamental pour les SIG temporels (Peuquet, 1994; Thériault and Claramunt, 1999; Siabato et al., 2018).

2.4.1 Caractérisation de la nature du temps :

La notion de temps peut s'appréhender et se mesurer par ensemble de valeurs temporelles qui localisent soit une mesure ponctuelle du temps soit un intervalle défini par deux instants, une pour le début du phénomène et

une seconde pour sa fin (Galton, 1997). Nous pouvons attribuer plusieurs caractéristiques permettant de décrire plus précisément le temps. Celles-ci sont considérées également comme des facteurs qu'un modèle de donnée de SIG doit prendre en considération pour une intégration réussie du temps :

Ordre du temps : c'est la structure qui permet de modéliser temporellement un phénomène. Le temps peut être conceptualisé avec une dimension linéaire, cyclique ou parallèle. La dimension linéaire représente la progression, elle implique le déroulement consécutif des événements sur un axe orienté. La forme cyclique représente la continuation et l'uniformité, elle conceptualise le fait qu'un temps peut avoir un caractère récurrent (Wachowicz and Healey, 1994). La dimension parallèle permet de combiner des séquences d'événements différents se déroulant simultanément, ou pas. Ces séquences peuvent être issues d'événements communs ou pas, et offrent la possibilité de prévoir les scénarios futurs ou de reconstruire les événements du passé à partir du présent.

Relation temporelle : selon (Cole and King, 1968), "Le temps est essentiellement considéré comme un ensemble de relations temporelles entre des événements et des objets". Par conséquent, l'analyse temporelle des phénomènes nécessite de formaliser les relations que l'on peut observer entre les états temporels. Les premiers travaux qui ont été réalisés dans ce sens étaient ceux de (Allen, 1983). Cette approche a été étendue et généralisée par une algèbre qualitative permettant de manipuler des intervalles convexes (Figure 2.3).

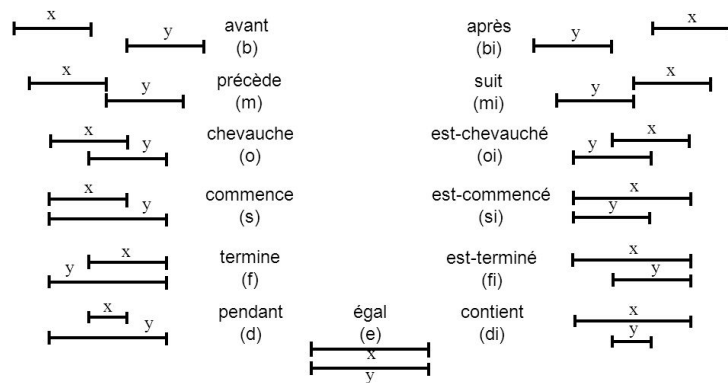


FIGURE 2.3 – Algèbre pour les relations temporelles entre deux intervalles (Allen, 1983)

Granularité temporelle : ce concept correspond au niveau de détail temporel utilisé pour observer et analyser un phénomène. En d'autres termes,

il définit la plus petite unité temporelle nous permettant de représenter le maximum des détails, en termes de changement dans un jeu de données. Un questionnement lié est celui par exemple de savoir si nous pouvons passer d'une représentation par jour à une autre par heure ou par seconde, etc., c'est-à-dire, s'il est possible en changeant l'unité de temps, d'avoir une représentation des données correspondant à cette unité choisie. Prenons l'exemple des types de données dans un système de gestion de base de données pour représenter le temps ; un type de données date, a une granularité d'un jour tandis qu'un horodatage peut avoir une granularité d'une seconde.

Densité temporelle : Une autre sémantique temporelle qui se présente comme un élément important à définir est celle de la densité temporelle. Celle-ci correspond à la fréquence des changements affectant un objet ou l'un de ses attributs. (Renolen, 1997) dans son modèle distingue différents types de changements, par exemple : changement avec continuité, état stable interrompu soudainement par un événement, changement discret, etc. Cette sémantique dépend essentiellement de la fréquence des changements qui affecte un phénomène, et de la répartition de ces changements sur l'axe du temps.

Temps transactionnel et temps valide : Cette notion a été introduite par (Snodgrass and Ahn, 1985). Le temps transactionnel correspond au temps d'enregistrement d'un événement dans une base de données. Le temps valide correspond au temps réel quand l'événement s'est effectivement produit. Les deux temps peuvent être évidemment différents. Le modèle prenant en compte les deux s'appelle bi-temporel (Snodgrass and Ahn, 1985).

2.4.2 Dichotomie entre objet temporel et champ temporel

(Galton, 2003) note qu'il existe au niveau conceptuel une sorte de dichotomie de type objet/champ qui se pose aussi bien pour le temps que pour l'espace. Selon lui, l'analogie temporelle avec les objets correspond aux événements et aux processus de toutes sortes.

L'objet temporel correspond à un événement relativement délimité dans l'espace et le temps. C'est le cas par exemple de la déclaration d'indépendance des États-Unis en 4 juillet 1776 signée au Pennsylvania State House à Philadelphie. Mais il existe également de nombreux événements dont l'étendue temporelle est plus difficilement définie : l'exemple de la pandémie espagnole en 1920 ou de l'érosion côtière dans la plage Porz Milin près de Brest. Tous ces événements, à la différence des objets spatiaux n'existent pas dans leur

intégralité à un moment donné.

L'approche champ temporel conçoit le temps comme un ensemble de plusieurs instants ou intervalles. Il définit un domaine d'une fonction f , dont celle-ci renvoie pour chaque instant ou intervalle temporel une valeur V exprimé comme $f: T \rightarrow V$. La valeur d'un champ temporel pourrait correspondre à celle d'un attribut (ex. température, population) ou à un événement donné.

2.5 Notions temporelles de champ et d'objet

Afin d'intégrer la dimension temporelle au niveau d'une représentation spatiale d'un phénomène géographique, il est nécessaire de faire la différence entre l'évolution d'un objet et l'évolution d'un champ, toutes deux reflétant les changements dans l'espace au fil du temps.

L'évolution d'un champ correspond à des variations en fonction du temps des valeurs liées à l'ensemble des emplacements spatiaux qui composent le domaine du champ. L'évolution des objets matérialise les changements de la géométrie, de la forme, de la sémantique des objets (i.e., des attributs non spatiaux) ainsi que des relations entre les objets voisins (ex. relations topologiques, relations directionnelles).

Afin de pouvoir mieux capturer cette dualité, il est nécessaire d'intégrer le temps dans les notions de champ et d'objet. L'objectif consiste à en déduire différentes options pour une représentation de la spatio-temporalité. Pour formaliser chaque cas nous considérons L et T comme les domaines spatial et temporel, respectivement, soit V un ensemble de valeurs à attribuer aux différentes entités géographiques.

2.5.1 Intégration du temps dans la notion de champ

Intégrer la notion du temps au sein de la vision champ de l'espace génère trois principales configurations :

- Cas 1) consiste à extraire ou dériver un champ spatial d'un champ temporel. Nous pouvons formaliser cette notion par la fonction $f_1: T \rightarrow (L \rightarrow V)$. Cette forme de représentation est appelée par (Galton, 2004), une forme "snapshot" et par (Wang et al., 2016) point de contrôle temporel (i.e., temporal checkpoint). Dans cette formalisation, chaque instant est associé à un état spatial qui dans ce

cas représente une variable spatialement distribuée.

- Cas 2) est l'inverse de la précédente formalisation où il s'agit de dériver un champ temporel d'un champ spatial. Nous pouvons formaliser cette notion par la fonction $f_2: L \rightarrow (T \rightarrow V)$. Cette forme de représentation est appelée par (Galton, 2004), une forme histoire et par (Wang et al., 2016) point de contrôle spatial (i.e., spatial checkpoint). Dans cette formalisation, chaque position du domaine géographique S est associé à un historique de valeurs sous forme d'un ensemble de paires temps-valeur.
- Cas 3) les deux formes peuvent être représentées par cette fonction $f_3: L \times T \rightarrow V$. Il s'agit d'associer la référence temporelle à chaque position du domaine L , ce qui crée la notion de points spatio-temporels avec pour chacune, une valeur qui lui est associée. Cette forme est nommée état (i.e., State).

Cette vision spatio-temporelle de la notion de champ est considérée comme équivalente à la vision Eulérienne (Wang et al., 2016). Elle consiste à observer la variation d'un phénomène en fixant des positions du domaine géographique et en contrôlant le temps. Un modèle général du mouvement peut être construit à partir de cette structure.

2.5.2 Intégration du temps dans la vision objet

Avant d'effectuer une analogie à la notion de champ précédemment étudiée, il est nécessaire d'évoquer l'importance de la notion d'identité. Cette notion a été évoquée et traitée par (Homsby and Egenhofer, 1997; Hornsby and Egenhofer, 2000), en la considérant comme un moyen fondamental qui permet d'identifier de façon unique chaque entité dans l'espace, et de pouvoir suivre sa dynamique et évolution. Cette identification permet de caractériser les changements qui sont opérés sur une ou plusieurs entités spatiales. Afin de formaliser l'intégration du temps dans la vision objet, nous allons considérer un ensemble O des entités spatiales discrètes ayant chacune une identité, empreinte spatiale ainsi que des attributs, dont leurs valeurs, appartiennent à l'ensemble V . Nous distinguons trois cas :

- Cas 1) consiste à extraire ou dériver un objet spatial d'un champ temporel. Nous pouvons formaliser cette notion par la fonction $f_1: T \rightarrow (O \rightarrow V)$. Cette forme de représentation est appelée par (Galton, 2004), une forme "configuration" et par (Wang et al., 2016)

”inventaires dynamiques” (i.e., Dynamic Inventories). Dans cette forme chaque instant est associé à un état spatial qui dans ce cas représente une configuration de l’objet o . Elle est matérialisée par une paire objet-valeur (o, v) .

- Cas 2) ce cas constitue l’inverse de la première où il s’agit de dériver un champ temporel des objets spatiaux. On peut formaliser cette notion par la fonction $f_2: O \rightarrow (T \rightarrow V)$. Dans cette formalisation, chaque objet o est associé à un historique de ses valeurs sous la forme d’un ensemble de paires temps-valeurs. Dans le cas où la valeur comprendrait l’empreinte spatiale de l’objet, on parle d’une forme trajectoire (Wang et al., 2016). Nous pouvons imaginer l’exemple d’une parcelle, associée à ses propriétaires successifs, ainsi qu’aux changements de frontières produits durant son existence.
- Cas 3) les deux formes peuvent être représentées par cette fonction $f_3: O \times T \rightarrow V$. Il s’agit ici d’associer la référence temporelle à chaque objet de l’ensemble O ce qui crée la notion d’objets spatio-temporels avec pour chacun une valeur est associée.

Cette vision temporelle de la représentation objet de l’espace est nommée par (Wang et al., 2016) la vision Lagrangienne. Cette vision permet de fixer l’identité d’un objet, de contrôler le temps et d’observer le changement des valeurs ainsi que les possibles évolutions spatiales comme les déplacements.

Nous pouvons nommer les objets ayant cette forme de représentation des ”**continnants**”. Ces derniers sont définis comme des objets qui perdurent dans le temps, subissent des changements et conservent leurs identités (Galton, 2004; Plumejeaud et al., 2009). Ces changements reflètent une évolution de leur géométrie, de leurs propriétés et de leurs relations avec leurs objets voisins (topologiques, spatiaux ou métriques). Ils ont une histoire et existent dans leur ensemble à chaque instant de leur existence. Ces objets ont un historique à travers les instants de leur existence.

La notion de continnant implique de prendre en considération la notion de changement des relations spatiales entre objets ainsi que les relations spatio-temporelles (Sriti et al., 2005; Del Mondo, 2011). Les relations spatiales sont des relations entre entités présentes en un même temps, alors que les relations spatio-temporelles sont des relations entre entités identifiées à des temps consécutifs, similaires ou postérieurs.

Néanmoins, l’intégration du temps dans la vision objet ne s’arrête pas ici, car selon (Grenon and Smith, 2004; Galton, 2004), il faut différencier à ce stade l’approche dite endurante (i.e., continuant) et l’approche dite per-

durante (i.e., *occurrent*). Dans l'approche *endurante*, l'espace est composé d'éléments qui durent dans le temps et subissent des changements. L'approche *perdurante* considère l'espace comme un ensemble d'objets perdurant (i.e., *occurrent*) définis comme des entités étendues dans le temps (ex. *tornado*). Leur existence s'étend sur une période donnée et entière, du début à la fin, et elles n'existent pas, dans leur complétude, à un seul moment donné. Contrairement à un *continuant*, ils ne peuvent pas subir de changement. Selon (Galton, 2004), cette notion de perdurant appartient à la vision de l'espace à quatre dimensions, dont le temps constitue la quatrième dimension permettant de repérer un objet dans l'espace-temps. Cette vision permet de se détacher totalement de la notion d'identité dans la mesure où un changement ne peut caractériser un perdurant (i.e., *occurrents*). Celui-ci est représenté entièrement ou par une emprise spatio-temporelle durant toute son existence. Cette notion est étroitement liée à la notion d'événement et de processus ce qui nous amènera à la prochaine section.

2.6 Événement et Processus

Les scientifiques ont souligné le rôle important que ces deux concepts d'événement et de processus peuvent jouer pour capturer et améliorer notre compréhension de la dynamique des phénomènes géographiques (Claramunt and Thériault, 1995; Hornsby and Yuan, 2008). Cependant au sein de la communauté des sciences de l'information géographique, il n'y a pas actuellement de consensus sur la définition de ces deux concepts ainsi que sur les différences entre les deux. Selon (Siabato et al., 2018), nous ne savons pas exactement "quand l'un commence et quand l'autre cesse". Il n'y a pas de relations uniques ou exclusives entre eux, car leurs natures changent en fonction des phénomènes modélisés (Carré, 2020). En général, nous pouvons discerner quatre grandes tendances en termes de définitions stabilisées de ces deux concepts :

- selon (Allen and Ferguson, 1994), les événements sont des occurrences discrètes d'un phénomène avec une emprise temporelle, un début et une fin. Ces événements sont constitués d'un ensemble de processus, dont chacun correspond à des changements (ex. *marcher*, *pleuvoir*).
- la définition de (Claramunt and Thériault, 1995; Claramunt et al., 1997) considère un processus comme l'agrégation d'un ensemble de changements étant liés ou produits simultanément; alors qu'un événement peut être représenté comme un ensemble de processus

qui modifient des entités. Ils ajoutent que comme l'observation ou la détection des événements est dépendante de la granularité temporelle, il n'est pas possible de définir une frontière ferme entre ces deux concepts.

- une autre vision proche est traduite par la définition de (Yuan, 2001) qui considère un processus comme une agrégation d'états temporels successifs alors qu'un événement n'est qu'une agrégation spatio-temporelle de processus. L'ensemble des processus liés à un événement sont appelés processus composites. (McIntosh and Yuan, 2005) propose un modèle qui suit la même logique en précisant que les processus peuvent être liés entre eux par des relations spatio-temporelles, plus particulièrement topologiques, et qui matérialisent des événements.
- (Galton and Worboys, 2005) définissent un événement comme un épisode terminé de l'histoire; il apparaît à un instant dans le temps et peut par la suite apparaître à un autre moment dans le temps comme un événement ultérieur, mais différent du précédent. La différence majeure entre un événement, un objet et un processus est que ces deux derniers peuvent subir des changements et que ces changements peuvent être décrits comme des états multiples, alors qu'un événement ne subit pas de changements. Vu qu'un événement n'est qu'un état spécifique (instantanée ou avec une durée), cela sous-entend qu'un processus pourrait être composé de plusieurs événements, ce qui va à l'encontre des deux définitions précédentes.

Malgré ces différences en termes de définitions, plusieurs constats peuvent être formulés. (Yuan, 2001) et (Galton and Worboys, 2005) conviennent qu'un processus implique différents états. Le processus et l'événement sont considérés comme des entités ayant leurs propres attributs et possèdent généralement des relations internes avec les objets géographiques (Claramunt and Theriault, 1996; Worboys and Hornsby, 2004; Hornsby and Yuan, 2008).

2.7 Espace et Temps en SIG

Un système d'information géographique (SIG) peut être défini comme un système d'information qui vise à utiliser la technologie informatique pour permettre la saisie, la modélisation, la manipulation, la récupération, l'analyse et la présentation de données géographiquement référencées

([Worboys and Duckham, 2004](#)). Nous pouvons distinguer trois principales composantes constituant le cœur d'un SIG. 1) La composante modèle de données qui permet de stocker, gérer manipuler et requêter l'ensemble des données géographiques. Cette composante est chargée de gérer les données de type vecteur et celles de type raster. 2) La composante de traitement et d'analyse qui offre à l'utilisateur une panoplie de fonctions de traitements et d'outils d'analyse pour pouvoir exploiter des données géographiques et en dériver des informations utiles et disponibles pour l'interprétation. 3) Enfin, la composante de visualisation qui distingue les SIGs des autres systèmes d'information. Cette dernière composante est capable de restituer le résultat des traitements des données géographiques en se basant sur le paradigme de la carte géographique en 2 ou 3 dimensions. Les données sont affichées dans une carte numérique avec généralement le principe de "layering" ou superposition de différentes couches en associant chaque type de donnée spatiale à une couche.

Un SIG temporel nécessite l'intégration de la dimension temporelle dans chacune de ces trois composantes. Selon ([Langran and Chrisman, 1988](#)), cette intégration doit permettre à un SIG temporel non seulement de posséder des fonctions traditionnelles de traitement d'informations spatiales, mais en plus offrir de l'information spatio-temporelle qui sera destinée à :

- Expérimenter l'applicabilité de théories et hypothèses scientifiques portant sur des phénomènes spatio-temporels.
- Analyser la dynamique d'un phénomène ainsi que les mécanismes qui l'accompagnent (événements, processus, causalités).
- Aider à la prise de décisions d'aménagement et de planification.

Selon ([Wachowicz and Healey, 1994](#)), le but d'un SIG temporel est de comprendre les changements et leurs effets plutôt que simplement les reproduire par une visualisation sous forme d'instantanés (i.e., snapshots).

La recherche tendant à construire un SIG temporel traite des modèles intégrant la dimension temporelle dans la composante spatiale du modèle de données ([Peuquet, 2002](#)). Ces modèles sont appelés des modèles de données spatio-temporelles. Ces modèles de données spatio-temporelles peuvent logiquement aider à améliorer la gestion des données spatio-temporelles décrivant la dynamique des phénomènes géographiques. Dans cette section, nous nous intéressons à la description des processus de modélisation spatio-temporelle

et faisons un état de l'art des différents modèles spatio-temporels actuellement proposés.

2.8 Modèles de données spatiaux

La composante "base de données" représente un aspect essentiel d'un SIG temporel qui doit permettre d'intégrer des données spatio-temporelles et d'intégrer les concepts temporels précédemment décrits. Un tel modèle constituera le noyau d'un SIG temporel permettant de traiter des données de SIG dans toutes ses dimensions spatiales et temporelles. Nous distinguons le processus de modélisation encapsulé par un modèle de données générique (exemple du modèle relationnel, orienté objet, orienté graphe, etc.).

2.8.1 Modélisation spatiale

Un modèle de données décrit comment la réalité doit être représentée pour une application donnée. Plus précisément, il définit les entités logiques, les attributs, les contraintes et les relations qui composent la structure d'une base de données (Date, 2011). Une étape importante de la modélisation des données consiste à sélectionner les phénomènes réels à représenter. Une deuxième étape consiste à décider comment représenter ces phénomènes de manière simplifiée pour produire un schéma de base de données. Ce processus est défini en trois étapes : réalisation d'un modèle de données conceptuel, réalisation d'un modèle de données logique et d'un modèle de données physique. Les décisions les plus difficiles se situent au niveau conceptuel, car c'est ici que la sémantique du domaine est identifiée et définie. Un modèle conceptuel consiste en la création d'une structure représentant une image abstraite de la réalité à l'aide d'un formalisme. Dans le même sens, un modèle conceptuel est un modèle formé après la phase de conceptualisation et se référant à une composition des concepts utilisés pour permettre la compréhension et la simulation des éléments de la thématique étudiée. Au niveau logique, il s'agit d'associer les concepts développés dans le modèle conceptuel avec des structures de données manipulables (tables, graphes, tuples, etc.) en respectant des règles et des contraintes d'intégrité spécifiques. Concernant le niveau physique, il traite du stockage et de la récupération d'informations à partir d'un support numérique, il est géré par le système de gestion de base de données (SGBD).

La modélisation des données est un processus de conceptualisation à partir duquel une application, ou des phénomènes du monde réel, sont représentés numériquement. Ces approches de modélisation sont au cœur

de la théorie et de la pratique des bases de données. Dans un SIG, ce processus de modélisation est encapsulé dans un modèle de données géographiques (MDG), dont certaines des entités logiques du modèle conceptuel auront une expression spatiale. De plus, les entités seront associées par des relations spatiales qualitatives et quantitatives en général bien définies. Prenons l'exemple d'un type d'entité présent dans les applications de transport : une route. Un exemple de relation entre deux routes peut être une intersection si ces deux routes se croisent. Dans ce cas, intersection donne une classe de relation spatiale. La modélisation des données géographiques est décrite en utilisant les mêmes trois étapes que la modélisation des données conventionnelles. Cependant des éléments de conception supplémentaires sont introduits, concernant les types d'entités qui nécessitent une représentation spatiale, ainsi que la façon dont chacune sera représentée. Généralement les deux modèles spatiaux raster et vecteur sont les deux principales approches utilisées pour les modèles de données géographiques ([Worboys and Duckham, 2004](#)).

2.8.2 Modèle spatial de base de données

Selon ([Date, 2011](#)), un modèle de base de données (MBD) constitue un référentiel abstrait, logique et indépendant comprenant des structures de données et un ensemble d'opérateurs d'interrogation et de manipulation des données. Un modèle de base de données est en principe destiné à modéliser n'importe quelle application sans redondance ni perte d'information. Il ne devrait comprendre aucune référence préalable à l'implémentation physique décrivant la manière dont ces données doivent être stockées et organisées. On distingue plusieurs modèles de base de données tel que le modèle hiérarchique, le modèle relationnel, le modèle orienté objet ou à base de graphes. En faisant une analogie, si l'on considère qu'un modèle de base de données est équivalent à un langage de programmation, alors un modèle logique de données sera équivalent du programme écrit dans ce langage.

La manipulation de très larges volumes de données géo-référencées dans de nombreuses applications, a largement favorisé l'augmentation des capacités de gestion, d'interrogation, de traitement, et d'indexation de données spatiales dans les bases de données spatiales ([Noyon et al., 2007](#)).

([Güting, 1994](#)) définit une base de données spatiales comme un système de base de données qui manipule des types de données spatiales dans la structure de son modèle de données de base et les intègre dans un propre langage de requête. Elle doit fournir à minima des méthodes d'indexations spatiales et de jointures spatiales. Un tel système de gestion de base de données spatiales doit être capable d'implémenter les différentes représentations et abstractions que les SIG utilisent pour modéliser les phénomènes géographiques ([Codd, 1970](#);

[Astrahan et al., 1976](#)). L'arrivée des bases de données spatiales en tant que domaine de recherche et de développement spécialisé a commencé avec les travaux initiaux de ([Berman and Stonebraker, 1977](#); [Chang and Fu, 1980](#)), qui ont défini les grands principes d'une architecture de base de données spatiales, suggérant par exemple l'approche d'architecture en couches. L'évaluation et l'adaptation des langages de requête pour l'intégration de géométries et types spatiaux ([Frank, 1982](#)) et plusieurs propositions d'indexation des structures de données spatiales (ex. ([Stonebraker, 1983](#); [Guttman, 1984](#))) ont également été des étapes importantes pour les premiers développements des bases de données spatiales.

Différents travaux ont été menés pour le développement des modèles de base de données extensibles, avec l'apparition de SGBD tels que Postgres ([Stonebraker and Rowe, 1986](#)), Probe ([Dayal and Smith, 1986](#)), EXODUS ([Richardson and Carey, 1987](#)). Ces travaux ont conduit au développement du principe d'architecture intégrée, un tournant important dans le développement des technologies de bases de données spatiales. Cette architecture permet d'unifier les standards, et les principes de définition de l'ensemble des éléments et des opérations de traitement liées à un SGBD. L'intérêt est de consolider le développement d'extensions (dite "cartouche") en particulier pour la modélisation spatiale. Cette approche permet de définir les types des données spatiales et les fonctions associées ; par exemple, la fonction permettant l'union de deux objets géographiques est basée sur le même principe que la concaténation de deux chaînes de caractères de type string. Même chose pour les méthodes d'indexation pour un attribut non spatial (ex. B-tree) et un attribut spatial (ex. R-tree), les méthodes de jointures sort/merge avec la jointure bounding-box. Les mécanismes d'optimisation des requêtes ne doivent également pas distinguer les opérations spatiales des autres opérations. La grande force de cette architecture est de développer un modèle spatial dans un SGBD extensible, lui-même devenant extensible. Cette approche permet également de capitaliser à partir des modèles conventionnels sans avoir à multiplier les développements et ainsi d'ajouter des fonctionnalités manquantes et spécifiques (ex. nouveaux opérateurs spatiaux basés sur une algèbre spatiale, nouveau type de données, etc.). Une introduction complète et un examen de l'état de l'art de l'évolution des bases de données spatiales sont disponibles dans le travail de ([Güting, 1994](#)).

Bien que la technologie des bases de données spatiales constitue une composante essentielle pour des applications des SIG pour l'étude des phénomènes géographiques, il apparaît, à ce stade, que la plupart des bases de données spatiales conventionnelles sont insuffisantes pour analyser les phénomènes dynamiques. En effet, elles sont essentiellement basées sur une représentation "figée" des phénomènes géographiques, et non dynamique.

Pour représenter des données spatio-temporelles, plusieurs recherches proposent des modèles de données dont le but est de gérer la dynamique spatio-temporelle du monde réel et celles des phénomènes géographiques.

2.9 Vers des modèles spatio-temporels

Le besoin d'intégration de la notion du temps dans un modèle spatial a fait émerger un ensemble des modèles spatio-temporels issus de différents types de classification et de taxonomie. Une approche standard consiste à classifier les modèles par rapport à leurs fondements théoriques et les primitives de représentation utilisées.

Les motivations qui ont menées aux propositions actuelles sont diverses, mais peuvent être classées en 4 principales qui sont : 1) le type d'application, 2) la nature des représentations et des requêtes souhaitées, 3) la prise en compte des possibilités de raisonnement spatio-temporel et 4) le besoin de performance des traitements.

Cette section examine un ensemble de modèles spatio-temporels proposés dans la littérature. Le but n'est pas d'être exhaustif, néanmoins nous représentons six grandes catégories de modèles spatio-temporels : le modèle de trois domaines, modèles basés sur l'approche continue, modèles orientés événement et processus, modèles basés sur un langage conceptuel, modèles basés sur l'approche de graphe, et finalement modèles basés sur le concept d'objet mobile. Des études plus exhaustives des modèles de bases de données spatio-temporelles et des approches SIG temporelles sont disponibles dans les travaux de ([Peuquet \(2001\)](#); [Pelekis et al. \(2004\)](#); [Siabato et al. \(2018\)](#); [Carré \(2020\)](#)).

2.9.1 Three Domain Model

Ce modèle consiste à représenter séparément l'information thématique, spatiale et temporelle ([Yuan, 1994, 1999](#)). Une relation entre ces trois types d'information est identifiée à travers des liens logiques, et permet de décrire les processus géographiques liés à un phénomène étudié. Dans ce modèle, l'information thématique se réfère à un concept du monde réel indépendamment de sa localisation dans l'espace et le temps. Par exemple, le propriétaire d'une parcelle pourrait être considéré comme un attribut de parcelle, alors que dans ce modèle le propriétaire est considéré comme une entité sémantique en relation spatiale avec la parcelle (i.e., objet spatial) et un changement de cette parcelle est associé, pour une date considérée, à un objet temporel. L'avantage de ce modèle est qu'il supporte à la fois la représentation des objets

et leur évolution, et les processus d'interrogation thématique et/ou spatial et/ou temporel (Figure 2.4). Toutefois, le modèle se limite à la vision objet de l'espace sans présenter de mécanismes permettant l'intégration de la vision continue de l'espace.

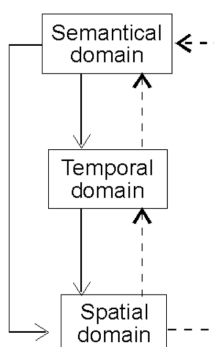


FIGURE 2.4 – Shéma conceptuel du three-domain-model (Yuan, 1994)

Dans un premier temps, ce modèle a été conçu comme un cadre conceptuel permettant de capturer et gérer l'information, et ses capacités de modélisation ont été confrontées à des cas réels complexes tels que les feux de forêt et la dynamique des cyclones (Yuan, 1994, 1996, 1999). Ensuite, il a été implémenté en tant que modèle de données spatio-temporelles afin d'améliorer la capacité des SIG à supporter des requêtes spatio-temporelles complexes. Cependant, le modèle n'est pas soutenu par une représentation formelle rigoureuse permettant de vérifier sa cohérence et de le rendre plus générique et extensible.

2.9.2 Modèles basés sur l'approche continue

Le premier modèle constituant de cette catégorie est le **modèle Snapshot** (Armstrong, 1988) qui consiste à capturer des états spatiaux sous forme de couches à un instant ou période donnée t_i . Chaque couche décrit la configuration spatiale d'une seule thématique (ex : occupation du sol, bâtiments) dont le temps associé (une date en générale) est considéré comme un attribut. Les états sont capturés avec des intervalles réguliers ou irréguliers, dont peuvent être conditionnés par des changements dans l'espace. Ce modèle propose une approche conceptuellement simple à implémenter pour intégrer le temps dans un SIG. Il permet de connaître l'état de chaque unité spatiale ou entité géographique en un instant donné. Cependant, ce modèle présente un ensemble de limitations dont les principales sont la redondance et le manque d'informations sur les changements (Peuquet and Duan, 1995; Claramunt and

Theriault, 1996). La redondance est liée au fait que toute l'information est reprise à chaque état, même en cas d'absence de changement. Les changements n'étant pas représentés explicitement dans un état, il est nécessaire pour les identifier de comparer deux états différents ce qui alourdit les traitements.

Une proposition permettant en partie de contrecarrer ces limitations a été proposé par (Peuquet and Duan, 1995). Dans le modèle **ESTDM**, à défaut de prendre un état pour chaque instant t_i , le modèle propose de prendre un état initial de l'espace en t_0 et d'enregistrer le changement sous forme d'une liste d'événement : c'est-à-dire que pour chaque instant t_i une liste d'événements désigne les changements produits par rapport à l'état t_{i-1} . Un événement est lié à une localisation (pixel en mode raster) et à un temps donné. Ce modèle est conçu pour manipuler les données raster, en revanche son adaptation au mode vecteur reste complexe. Avec cette approche, les changements sont explicitement repérés dans le temps, ainsi les requêtes à caractère temporel (ex. trouver tous les événements ayant eu lieu pendant une période choisie) seront faciles à gérer par le modèle (figure 2.5). Les capacités d'implémentation de ce type de modèles restent cependant relativement lourdes à mettre en oeuvre et à manipuler, notamment pour les gros volumes de données.

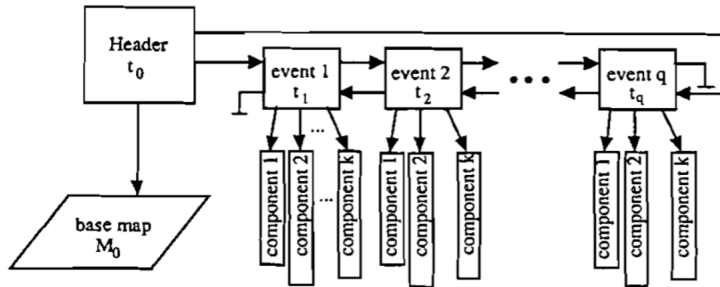


FIGURE 2.5 – La structure du modèle ESTDM est constituée d'un en-tête (header), d'une carte initiale (base map) et d'une liste d'événements (event list), l'événement étant lui-même formé de composants (components)

Un autre modèle proposé pour gérer la vision continue de l'espace est le système **TGrass** (Gebbert and Pebesma, 2014) qui étend l'approche snapshot avec l'utilisation d'un nouveau concept appelé *spatio-temporal data-set (STD)*. Il permet de gérer le mode raster, voxel (3D pixel) et vecteur map (TIN, diagramme de Voronoi), et de gérer les relations temporelles topologiques. Des snapshots en liaison temporelle peuvent être identifiés et les différences entre eux peuvent être calculées. Contrairement à ESTDM, TGRASS n'utilise pas de structure spécifique de données pour stocker les

différences entre les champs spatiaux. Le modèle n'enregistre pas explicitement des changements, cependant, il les supporte. Par contre, l'effort de calcul pour détecter les différences est cependant beaucoup plus élevé que dans l'approche ESTDM et le stockage des données n'est pas aussi efficace. Cette approche a abouti au développement d'un logiciel TGRASS comme extension du système GRASS ([Gebbert and Pebesma, 2014, 2017](#)).

Le modèle **Space-time Raster GIS data model** proposé par ([Zhao et al., 2015](#)) a pour but d'organiser des collections de données multi-temporelles en distinguant différentes sources de données raster. Ce modèle permet le suivi et l'analyse de séquences de changement pour chaque pixel à travers le temps, et est particulièrement efficace pour manipuler des requêtes spatio-temporelles qui permettent de trouver l'état d'une zone étude en un temps donné. Le modèle utilise des concepts orientée-objet (classes, méthodes, objets, etc.) afin de gérer les données raster multi-temporelles et de différentes sources. Le modèle est représenté par dix classes d'objets dont trois entre elles constituent le corps de ce modèle. Il s'agit de :

Space-time Object : cet objet est instancié pour capter toutes les informations historiques de différentes sources de données de chaque position (pixel) à travers le temps.

Space-time Dataset : cet objet représente toutes les informations historiques à travers le temps liées à une position (pixel) de l'image, mais pour une seule source de données.

Space-time Element : cet objet est instancié pour chaque position (pixel) de chaque image, il enregistre la valeur et le temps correspondants à cette position. Le diagramme de la (Figure 2.6) illustre la relation qui existe entre ces trois classes d'objets.

Chaque classe d'objets est caractérisée par des propriétés et des comportements qui servent à appliquer des requêtes spatio-temporelles et faire des analyses soit en accédant aux propriétés des objets ou en appelant leurs méthodes (i.e., fonctions). Il est ainsi possible de stocker les valeurs résultantes des requêtes exécutées.

Ce modèle présente l'avantage de pouvoir identifier les événements périodiques et prévoir les événements futurs, et détecter efficacement les changements au niveau de chaque pixel. Cependant, il comporte les mêmes inconvénients du modèle ESTDM à savoir l'incapacité de préserver les liens entre les événements pour définir différents processus spatio-temporels.

Dans l'ensemble, les modèles présentés ci-dessus, se limitent à la gestion

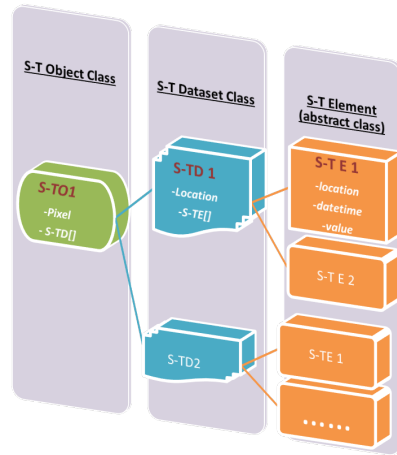


FIGURE 2.6 – Diagramme conceptuel montrant la relation de hiérarchie entre les trois classes d’objets : ST object class, ST Dataset Class et ST-Element Class

de la vision continue de l’espace dans le temps. Plus particulièrement ils adressent l’extension du mode raster dans le temps. De plus, et dans la plupart des cas, l’approche de modélisation est limitée à l’implémentation des schémas de données sans être soutenue par des modèles de représentation formels.

2.9.3 Modèles orientés événement et processus

Ces approches considèrent un changement comme étant une entité distincte du concept d’objet. Un événement ou un processus possède ses propres attributs et relations. Un changement est identifié comme un événement ou un processus (la différence entre les deux concepts était discutée dans la section 2.6). Les modèles décrits dans la section précédente, manipulent pour certain d’entre eux explicitement les événements (i.e., des changements), mais sans pouvoir y associer des processus. L’objectif, ici, consiste à pouvoir classer et qualifier les différents changements ainsi que de proposer des modèles capables de préserver explicitement les liens connus entre les événements et leurs conséquences. Dans ce sens, (Claramunt and Thériault, 1995) ont proposé une typologie des processus qui s’opèrent sur les entités géographiques à partir de 3 principales catégories de classes : déplacement, diffusion et restructuration. En parallèle un modèle de données plus conceptuel et basé sur le concept de "three domain framework" a été proposé, il intègre cette typologie et permet de représenter explicitement les types de processus et de séparer les différents domaines à savoir : spatial, thématique et temporel en

mettant des liens entre eux. (Thériault et al., 1999) prolongent ces travaux en proposant une taxonomie des changements qui décrivent les processus pour des ensembles d'entités géographiques. Assumant que pour modéliser un système il faut prendre en considération les objets et les événements, ainsi que les relations et les interactions entre les deux, (Worboys and Hornsby, 2004) propose le modèle *GEM* supportant cette dualité en définissant trois types d'entités : geospatial object, geospatial event, et geospatial setting (geo-settings). Cette dernière est considérée comme une caractéristique temporelle (instant ou événement), spatiale et/ou spatio-temporelle. Cette entité peut être attachée soit à un événement, soit à un objet. Ce modèle définit quatre relations événement-événement (ex. initiation, blocage, etc.) et cinq relations objet-événement (ex. création, division, etc.). Ces relations représentent la dépendance mutuelle des objets et des événements : les objets ont besoin d'une occurrence d'événement pour exister et les objets donnent un sens aux événements (Figure 2.7).

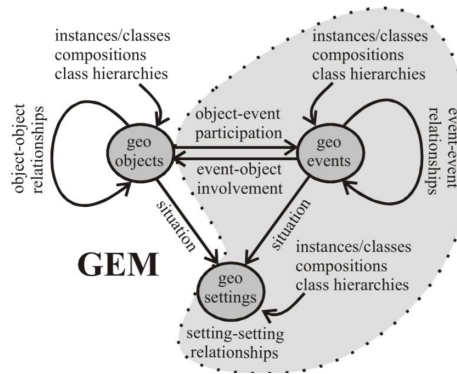


FIGURE 2.7 – Le modèle GEM : objets, événements et leurs interactions (Worboys and Hornsby, 2004)

Cette approche permet de modéliser la dynamique du domaine géographique en considérant les changements à l'intérieur des objets, les événements qui ont produit ces changements et la définition des relations de dépendance possibles. Si le potentiel de ce modèle a été montré sur un cas théorique, aucune implémentation sur des données réelles n'a été faite à ce stade afin de tester sa faisabilité (Siabato et al., 2018).

Une autre approche visant le même objectif, mais plus générique, est celle de (Ferreira et al., 2014). Cette approche propose une algèbre prenant en compte la notion d'événement et décrit la dynamique du domaine spatio-temporel. Avec cette algèbre, trois types de données sont définis comme des abstractions génériques basées sur des observations du monde réel : *séries temporelles*, *trajectoire* et *couverture*. Des types d'objets et d'événements

sont définis sur la base de ces abstractions. Un objet est une entité identifiable dont les propriétés spatiales et non spatiales peuvent changer, alors qu'un événement correspond à un épisode individuel avec un début et une fin. La seule relation établie entre un événement et un objet est qu'un événement peut impliquer un ou plusieurs objets. Les types et les fonctions de données proposés peuvent modéliser et capturer les changements dans une large gamme d'applications. Le modèle a été testé à l'aide de la bibliothèque de logiciels TerraLib (Câmara et al., 2000, 2008). Le modèle ne propose pas des types de données qui représentent les différents types de relations entre événement-événement et objet-événement comme celui de **GEM**, mais le modèle reste abstrait et extensible avec des spécifications génériques supportant le développement de ces types de données.

2.9.4 Modèles spatio-temporels basés sur un langage conceptuel

Dans ces approches, la majorité des travaux proposés sont des extensions des modèles déjà existants à savoir 1) le modèle entity-relationship (ER) (Chen, 1976), 2) la modélisation objet (Rumbaugh et al., 1991). Ils concernent également l'extension du langage unifié de modélisation (UML) (Booch, 1999). Le but est d'étendre ces approches avec des concepts et paradigmes permettant de capturer les propriétés spatio-temporelles de l'information.

Le Modèle Entité-Relation spatio-temporel (STER) proposé par (Tryfona and Jensen, 1999; Tryfona et al., 1999; Tryfona and Jensen, 2000), est une extension du modèle entité-relation (ER), qui est un langage conceptuel permettant de modéliser des concepts du monde réel. En particulier, l'ensemble des concepts contenus dans le modèle ER permet la capture et l'organisation de l'information en facilitant le passage à des structures de bases de données relationnelles.

L'approche STER étend les capacités de modélisation conceptuelle du modèle ER, lui conférant la possibilité de capturer la nature dynamique des bases de données spatio-temporelles. Il adresse en particulier trois objectifs : 1) capturer la position d'un objet dans le temps et dans l'espace, 2) décrire la sémantique temporelle des attributs et intégrer les cardinalités du temps, et 3) capturer les interrelations de la sémantique spatiale et temporelle. Il reste à mentionner que, même si l'intention de ce modèle était de construire un passage ou une transition depuis la réalité vers la base de données relationnelles, ce modèle reste aussi valable pour modéliser et capturer des connaissances ou des concepts généraux du monde réel et

pourrait être utilisé comme un principe d'architecture pour d'autres modèles.

Le **Modèle conceptuel pour des applications spatio-temporelles (MADS)**, développé par ([Parent et al., 1997, 1998](#); [Claramunt and Parent, 2003](#); [Parent et al., 2006](#)) a été proposé comme un langage de modélisation de base de données spatio-temporelle. Dans cette approche, la description des processus inclut les types de relations entre les objets, ce qui permet une meilleure description d'un phénomène spatio-temporel. Le modèle repose sur la définition de trois composants principaux. Premièrement des primitives spatiales, allant des plus basiques (point, ligne, surface) vers des niveaux abstractions plus importants et hiérarchiques avec Geo, simple Geo et complex Geo. Ensuite, des primitives représentant les relations typologiques (égalité, chevauchement, croisement, etc.). Et enfin, des attributs temporels valables en différentes granularités (débit, max, min, moyenne). Le modèle établit une base théorique pour reproduire des opérations de manipulation. Il prend en charge la définition des contraintes d'intégrité basées sur les relations spatiales entre les objets spatiaux ou d'autres conditions sur la géométrie des objets. L'avantage clé de ce modèle réside dans le fait que son pouvoir d'expressivité est élevé, et cela en raison de l'emploi du principe d'orthogonalité pour amender les structures de données par des constructions spatio-temporelles. Ce principe permet de rendre les différentes dimensions spatiales, temporelles et thématiques indépendantes, ce qui facilite la combinaison et/ou le détachement des différents concepts en fonction des spécifications de l'application. Il offre de plus, la possibilité de décrire et de représenter des attributs spatiaux continus (notion de champ). Dans ce modèle la description des processus est représentée sous format d'un type de relation entre les objets ce qui permet une description meilleure d'un phénomène spatio-temporel.

L'**Extended Spatio-temporal UML (XSTUML)**, proposé par ([Price et al., 2002](#)), est une extension du langage UML dont le but est de fournir une notation claire, simple et cohérente, pour capturer une sémantique alternative concernant le temps, l'espace et les processus. Ce modèle hérite des différentes spécifications et notations d'UML en augmentant sa flexibilité pour la modélisation des scénarios spatio-temporels. Cela a été réalisé par l'introduction de constructions appropriées qui peuvent être combinées et appliquées à différents niveaux du diagramme de classes UML. Plus précisément, il propose cinq éléments : (i) emprise spatiale ; (ii) existence d'objet et temps de transaction ; et (iii) trois types différents de données spatio-temporelles (changements temporels dans les emprises spatiales, changements thématiques dans l'espace ou le temps, et données spatiales compo-

sites changeant avec le temps et/ou localisation). Ces éléments peuvent être déployés au niveau des attributs, des objets et des associations.

Dans l'ensemble, bien que les modèles STER, MADS et XSTUML proposent une représentation formelle puissante, dotée de concepts et de constructions génériques, la principale critique qui peut être formulée est que ces langages sont essentiellement conceptuels et ne permettent pas une correspondance facile avec la représentation logique, ainsi qu'une intégration fluide avec les outils SIG disponibles.

2.9.5 Modèles basés sur une approche de graphe :

Le modèle **Model History graphe** proposé par (Renolen, 1997) recouvre une large gamme des comportements qui concernent des objets selon trois types : 1) des objets qui changent en continuité, 2) des objets statiques, mais dénotant des changements qui durent et 3) objets statiques, mais dénotant des changements soudains. L'auteur propose dans ce modèle une notation représentant : a) une version d'objet (un état statique) sous forme d'une boîte rectangulaire, b) une transition (un état qui change) sous forme d'une boîte ronde et c) un changement soudain sous forme d'une boîte circulaire (Figure 2.8).

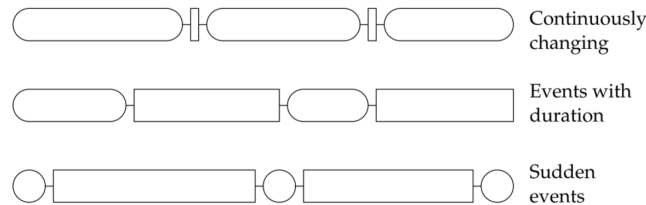


FIGURE 2.8 – Les trois types de comportements pour un objet temporel selon le modèle de Renolen (Renolen, 1997)

Les liens entre deux états représentent les relations successeur-prédécesseur entre l'état d'origine et l'état cible. Sept types de changements sont envisagés : création, réincarnation, destruction, altération, fusion/an-nexion, scission/déduction, et ré-allocation. Le modèle a l'avantage de permettre de dériver directement l'ensemble des relations temporelles et permet la représentation du temps absolu et relatif ainsi que le temps discret et continu. Le modèle permet de capturer les concepts de mouvement et de changement et prend en charge la plupart des types de requêtes spatio-temporelles. Le modèle exprime des événements tels que le "fractionnement" et la "fusion" et peut être étendu pour mieux caractériser la durée des

événements. Cependant ce graphe ne permet de gérer efficacement l'interaction entre les objets. Par exemple, les processus de scission/déduction impliquent généralement la création d'un objet supplémentaire et la modification d'un autre. De plus, il n'existe pas de méthode définie pour déterminer tous les processus et/ou mutations. En plus de ces lacunes, le modèle reste au niveau conceptuel avec un manque évident de formalisme rigoureux, et aucune proposition de mise en œuvre permettant de tester le modèle sur un cas réel n'est suggérée.

Le Modèle de graphe spatio-temporel pour représenter l'évolution des entités proposé par (Del Mondo et al., 2010; Del Mondo, 2011), permet la représentation du domaine spatio-temporel comme étant un réseau en définissant un graphe. Conceptuellement, il se base sur un modèle précédent qui permettrait de relier des évolutions d'entités cadastrales par un graphe (Spéry et al., 2001).

Le modèle conceptualise l'ensemble des entités qui composent l'espace géographique en étant les nœuds du graphe et les relations qui les relient dans l'espace et le temps en étant les arêtes. Ensuite, le modèle définit trois éléments fondamentaux pour la formalisation du graphe : un domaine temporel, un ensemble de relations spatiales (par exemple des relations topologiques), un ensemble de relations de filiation (appelées relations de continuation et relations de dérivation). Les entités ayant une emprise spatiale sont reliées entre elles par des relations spatiales, spatio-temporelles et de filiation et cela à travers la définition des relations binaires entre chaque pair à un même instant ou à deux instants différents. Après la construction du graphe, les auteurs proposent une représentation spatio-temporelle avec plusieurs opérateurs, d'abord pour la manipulation des graphes (intersection des graphes, union des graphes et jointure des graphes) et ensuite pour la détection des processus (division, fusion, naissance, destruction, etc.). Le modèle a été ensuite implémenté par (Thibaud et al., 2013) dans une base de données relationnelle dont l'application a concerné la modélisation de la dynamique des dunes sous-marines. Ainsi, le modèle fournit des éléments pour vérifier la cohérence du graphe et, par extension, la cohérence des données. Le modèle présente la limite de se focaliser complètement sur la vision objet de l'espace sans présenter d'autre mécanisme pour supporter l'aspect continu et distribué des phénomènes géographiques. Ainsi, le modèle n'offre pas de mécanisme ou implémentation générique avec des types de données spatio-temporelles pour gérer les changements spatiaux dans le temps (déplacement, contraction et/ou l'expansion de l'emprise spatiale, etc.) au niveau d'un objet.

2.9.6 Approche "Objet mobile"

Cette approche est favorisée par le développement des technologies des capteurs mobiles avec la difficulté de gérer et implémenter ces types de données avec les méthodes SIG traditionnelles. Le potentiel de cette approche en a également fait un sujet de recherche actif visant à modéliser la dynamique de certains phénomènes géographiques et notamment la représentation d'objets mobiles. Dans ce sens, une série de travaux de recherche a été proposée concernant 1) la modélisation des données (Xu and Güting, 2013), 2) les méthodes d'interrogation et de requêtes (Zheni et al., 2009; Ilarri et al., 2010; Elmongui et al., 2013; Jin et al., 2017) et 3) les méthodes d'interpolation spatio-temporelles et l'inférence de position (Wentz et al., 2010; Prager and Barber, 2012).

Des approches spécifiques de modélisation spatio-temporelle ont été proposées pour les objets mobiles, et par similarité, pour les phénomènes représentés par des géométries ponctuelles (Hornsby and Cole, 2007). Par contre, seules quelques études ont été menées sur les objets mobiles ayant une emprise spatiale bidimensionnels (i.e., des régions) (Güting et al., 2000; Yang and Claramunt, 2003; Noyon et al., 2007; Heinz and Güting, 2016, 2018, 2020). Le principe de ces méthodes est bien décrit dans les travaux de (Güting et al., 2000) et (Forlizzi et al., 2000). Ils ont proposé un modèle de données pour gérer les objets en mouvement (moving object) en introduisant un constructeur appelé "moving type" qui transforme les types statiques en types mobiles. Ce constructeur a été appliqué sur les primitives géométriques pour avoir des types comme : point mobile, ligne mobile et région mobile. Avec ces types de données abstraits associés à une collection d'opérations, le modèle devient capable de définir le mouvement continu des objets. Pour détailler davantage, et afin de mieux gérer les observations de mouvements avec ces types de données, les auteurs adoptent un modèle discret où ils décomposent l'ensemble des observations en un ensemble de séquences. Chaque séquence est composée d'un nombre fini d'unités, et une unité correspond à une paire de périodes $[t_i, t_j]$ avec la fonction qui relie chaque instant t_k dans cette période avec sa valeur. Le grand intérêt de ce modèle de données est qu'il permet à l'utilisateur de modéliser, stocker, récupérer et interroger les changements d'objets au fil du temps en prenant en compte la continuité de mouvement. Cependant, cette approche reste trop limitée quand il s'agit de gérer et intégrer les représentations spatiales continues. Ainsi, si la spécification logique et les aspects de mise en œuvre sont les points forts de ces types de modèles, il leur manque néanmoins une représentation formelle complète au niveau conceptuel qui permettrait de capturer facilement les informations liées à la dynamique des phénomènes

géographiques.

2.9.7 Analyse et critique globale des modèles

Il existe au final différentes approches permettant l'analyse des modèles spatio-temporels. (Pelekis et al., 2004) ainsi que (Siabato et al., 2018) fournissent un ensemble complet de critères sémantiques et fonctionnels pour évaluer chaque modèle. Par exemple, ils indiquent le type du modèle spatial supporté (i.e., continu ou discret), si le modèle propose une implémentation ou non, le type de structure de données spatio-temporelle supportée (point mobile, région mobile, etc.), ou encore si le modèle permet de répondre aux différents types de requêtes (spatiales, temporelles et spatio-temporelles). Ce dernier critère reste un moyen fiable et pratique à partir duquel nous pouvons tester et évaluer un SIG temporel.

(Yuan and McIntosh, 2002) ont proposé une typologie de 11 requêtes d'informations spatio-temporelles pour évaluer un modèle spatio-temporel : requêtes d'attribut, requêtes spatiales (trois types), requêtes temporelles (trois types) et requêtes spatio-temporelles (quatre types). Nous allons nous concentrer sur des requêtes spatio-temporelles qui caractérisent l'aspect multiaspect des phénomènes complexes (Yuan, 2001; Galton, 2004). Nous prenons deux exemples illustratifs de ces types de requêtes :

- S'interroger sur le possible impact des conditions atmosphériques (i.e., la dynamique du champ) sur la résilience d'une réserve biologique (i.e., l'évolution des objets) au cours d'un phénomène d'ouragan (i.e., un événement). Dans ce cas la résilience de la réserve biologique se mesure en caractérisant le changement de son emprise spatiale.
- S'interroger sur les possibles conséquences de l'évolution d'un champ de dunes marines (i.e., la dynamique des objets) sur la variation temporelle de la profondeur locale du fond marin (i.e., La dynamique du champ).

Ces exemples de requêtes illustrent une sorte de combinaison ou d'hybridation entre la vision du champ et celle d'un objet dans le temps. Par exemple, la première requête permet d'explorer l'impact d'une variation du champ sur l'évolution d'un objet, tandis que la deuxième requête illustre l'effet d'une évolution d'objet sur une dynamique du champ. Ces requêtes spatio-temporelles illustrent également la manière dont les concepts de continuant (c.-à-d., objets dont l'existence persiste dans le temps) et celui d'occurrent (objets dont l'existence est limitée dans le temps) devraient être pris en

compte de manière implicite surtout lors de la manipulation de l'approche objet dans le temps. Notre attention se porte donc sur la capacité des modèles à intégrer les deux visions objet et champ, et à prendre en compte les notions de continuant et occurrent.

Concernant les différents modèles et approches spatio-temporelles décrits précédemment, nous remarquons que tous ces modèles présentent une certaine dichotomie soit d'une part entre la vision champ et celle d'objet, soit d'autre part entre les concepts continuant et occurrent (i.e. événement). En général, les modèles ont pu réconcilier un aspect, mais ignorer l'autre. Ainsi, nous pouvons classer les modèles comme suit :

- **Modèles orientés vers une approche continuant** : il s'agit principalement du modèle de "three domain" (Yuan, 1994, 1999) ainsi que les modèles adoptant l'approche "objet mobile". Ces derniers définissent des structures et des opérateurs permettant de mieux décrire l'évolution des objets dans le temps. Par conséquent, ces modèles ne permettent pas d'intégrer suffisamment l'approche champ de l'espace dans le temps avec le concept d'objet mobile.
- **Modèles qui ont intégré l'approche occurrent/continuant** : il s'agit principalement des modèles de GEM (Worboys and Hornsby, 2004), de (Ferreira et al., 2014) ainsi que le modèle de graphe spatio-temporel de (Del Mondo, 2011). Ce dernier a l'avantage de gérer les relations spatiales et spatio-temporelles entre objets et d'en dériver des changements appelés processus. Cependant le modèle supporte pour l'essentiel une catégorie de processus surtout celle liée à l'identité (e.g. fusion, division, naissance, etc.) et ne permet pas de préserver explicitement des liens existents entre ces processus. Concernant les modèles de (Ferreira et al., 2014) et GEM, les deux préservent les liens entre événement/objet et événement/événement même si ces liens restent beaucoup plus basiques chez (Ferreira et al., 2014) en comparaison avec le modèle de GEM. Cependant, la majeure limitation de ces modèles est qu'ils ne permettent pas de gérer la vision champ dans le temps et son association avec l'approche objet/événement.
- **Modèles focalisés sur l'approche continue/raster** : Ces modèles ont proposé des techniques intéressantes pour une gestion plus efficace des données spatio-temporelles de type raster qui rendent possible la représentation du champ à travers le temps. Cependant, ils ignorent complètement la gestion de l'objet dans l'approche temporelle.

Reste l'approche qui consiste à proposer des langages de modélisation conceptuelle à savoir STER, UML et MADS. Ces langages sont plus orientés entité et objet et présentent l'avantage d'un niveau d'abstraction élevé ce qui les rend relativement capables de capturer les propriétés d'une vision continue ainsi que l'approche événement. Néanmoins, ces modèles ne spécifient pas des moyens ou des constructions permettant d'intégrer la notion hybride (champ et objet) dans un modèle de données spatio-temporel.

En général, la majeure limitation de ces modèles décrits dans cette section réside dans le fait qu'ils ne proposent pas une représentation explicite combinant la vision champ et objet dans le temps ce qui ne leur permet pas de s'adapter à certains aspects des phénomènes complexes, ces aspects qui ont été illustrés par les requêtes ci-dessus. De ce fait, dans la prochaine section, nous nous concentrerons davantage sur les modèles de données spatio-temporelles qui visent l'intégration des deux visions dans le temps que nous qualifierons comme des modèles hybrides.

2.10 Vers des modèles de données hybrides

Dans cette section, nous décrivons les différents travaux connexes sur les modèles de données hybrides. Nous les catégorisons selon la façon dont ils intègrent conceptuellement les visions champ et objet dans un modèle de données spatio-temporelles. Dans ce sens nous distinguons trois approches : l'approche basée sur l'intégration par primitive, l'approche basée sur l'intégration de concepts abstraits, et l'approche basée sur l'association des visions champ et objets.

2.10.1 Approche basée sur l'intégration par primitive

Cette approche est basée essentiellement sur la définition d'une unité ou primitive de base à travers laquelle on dérive de nouveaux concepts hybrides en utilisant le mécanisme d'agrégation.

Modèle pour représenter les phénomènes complexes distribués :

Ce modèle a été proposé par (Yuan, 2001). Pour intégrer les deux visions l'auteur propose un cadre conceptuel hiérarchique composé de trois couches dont celle appelée *state* est la plus basique. Avec cette couche les différentes données raster sont gérées. Les emprises spatiales des objets sont identifiées

dans chaque instant t et sont explicitement stockées. Ensuite des couches appelées processus composites sont construites dont chacune comporte tous les processus inclus dans un événement. Dans cette conception, un processus est formé par une séquence temporelle de *state*, et un événement est défini comme l'agrégation spatiale et temporelle de ses processus (Figure 2.9).

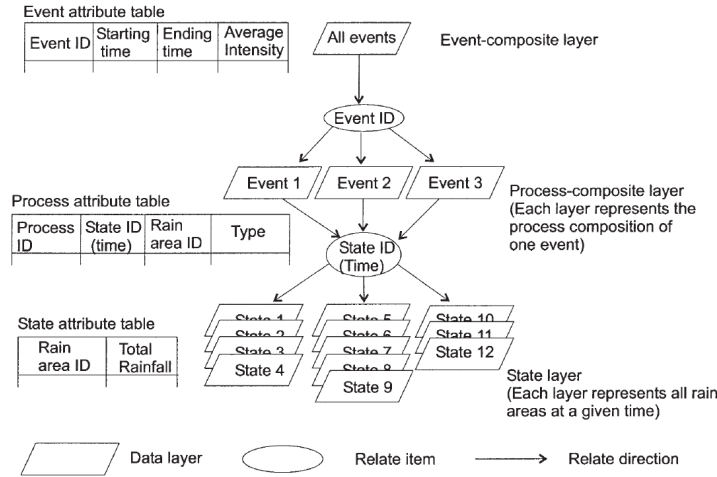


FIGURE 2.9 – Modèle logique montrant comment les données se sont organisées autour de state, processus et événement [Yuan \(2001\)](#)

En général, la vision objet est gérée et représentée à travers la notion de processus et d'événements où leurs attributs, leur durée, leur mouvement et leur fréquence sont explicitement stockés. La vision champ, qui représente les distributions spatio-temporelles d'un thème géographique (dans ce cas comme les précipitations), est gérée à travers les couches *state*. Le modèle offre un support direct pour interroger ces informations spatio-temporelles, car les événements et les processus sont explicitement gérés et liés entre eux. Néanmoins ce cadre conceptuel ne permet pas de gérer la notion du continuant sachant que les entités dans l'espace discret ne sont pas toujours considérées comme des processus (exemples des rivières, des dunes, une métropole, etc.) mais plutôt comme des endurants. Par conséquent, nous ne pouvons pas considérer qu'un événement est une agrégation spatio-temporelle de ces types d'objets. Le modèle ne gère pas explicitement les relations entre processus, et présente un schéma relationnel simple, qui n'est pas suffisamment générique pour permettre de l'étendre et l'enrichir avec d'autres concepts.

Un cadre pour améliorer la flexibilité sémantique pour l'analyse

des phénomènes distribués :

Ce modèle est proposé par (McIntosh and Yuan, 2005) et peut considérer comme une version améliorée du modèle de (Yuan, 2001) qui adopte la même approche duale pour représenter les phénomènes géographiques distribués. Ils proposent un schéma relationnel basé sur les événements, où la primitive de base est une entité appelée zone. Celle-ci est identifiée à chaque instant t_i dont elle possède une forme géométrique (polygone) et constitue une agrégation d'un sous-ensemble de cellules raster (i.e., pixels). Les pixels et les zones, ainsi que leurs relations, sont stockés dans des tables relationnelles. Le modèle construit trois types d'entités (séquences, processus et événements) en suivant la même hiérarchie que le modèle précédent, et en gardant le même processus d'agrégation, mais cette fois-ci, selon des règles basées sur des relations spatio-temporelles. Ces entités sont aussi stockées explicitement dans des tables relationnelles (Figure 2.10).

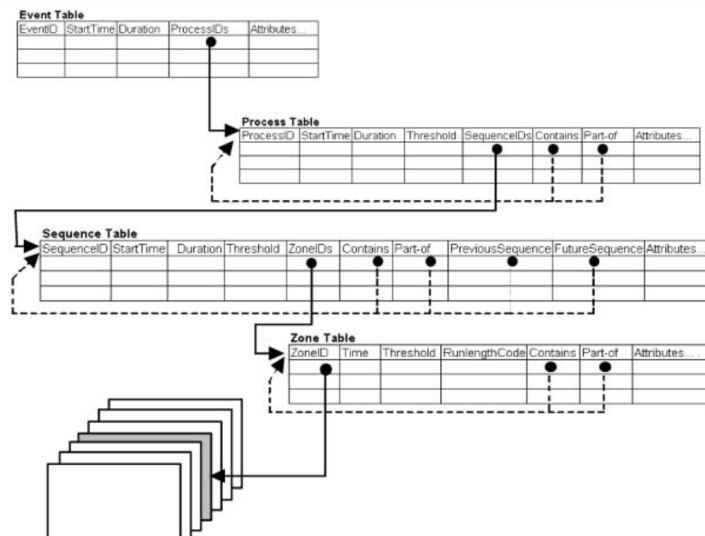


FIGURE 2.10 – Modèle de données logique pour représenter les concepts de zone, processus et événement (McIntosh and Yuan, 2005)

L'avantage de ce modèle est qu'il permet de représenter les relations spatiales et spatio-temporelles entre zones, événements et processus. Néanmoins il garde quelques limitations à savoir, le fait de ne pas pouvoir gérer des endurants, et que le modèle d'implémentation n'est pas générique, car il représente un schéma de données logiques composé seulement d'une collection de tables.

Modèle EDGIS :

(Pultar et al., 2010) développe le modèle EDGIS (‘Extended Dynamic GIS’) basé sur une forme atomique d’informations géographiques appelées *geoatoms*, celle-ci est définie par (Goodchild and Yuan, 2007) comme un tuple $\langle x, Z, z \rangle$ dont x constitue un point dans l’espace-temps associé à une propriété Z et une valeur z . (Goodchild and Yuan, 2007) proposent de construire ou de dériver la vision *géo-champ* de cette notion comme une agrégation des *geoatoms* selon une propriété Z . Un *géo-objet* peut être dérivé à travers une agrégation des *geoatoms* avec une ou des conditions sur la valeur de différentes propriétés Z . Ensuite, le concept de *géo-dipôle* est défini comme un tuple $\langle x_1, x_2, Z, z \rangle$ associant une paire de points spatio-temporels x_1 et x_2 à une propriété Z associée à sa valeur observée z . Une agrégation d’un ensemble de *géo-dipoles* sous une condition sur les valeurs de la propriété Z , conduit au concept d’objet-champ (voir définition section 2.3). Dans le modèle EDGIS le concept de *geo-atom* est matérialisé par celui de ”point spatio-temporel” (STP). L’approche EDGIS définit essentiellement trois structures de données : *STP*, *Features* et *themes*. *Features* identifie l’objet et *themes* le type d’objet (voir Figure 2.11). Ces structures permettent d’interroger des données spatio-temporelles par points, temps et attributs en intégrant l’approche objet matérialisée par le concept *features*. Chaque *features* est défini par son nom, un identifiant et pourrait appartenir à plusieurs *themes*. Le point le plus important est qu’elle stocke explicitement l’ensemble des identifiants des *STP* qu’elle contient, ce qui permettra de gérer l’hétérogénéité au sein de l’objet. Dans l’autre sens, chaque *STP* est caractérisé par ses coordonnées spatiales x et y et sa composante temporelle t . Un *STP* est associé à une séquence des valeurs observées dont chacune correspond à une propriété de l’espace (température, pression, élévation, etc.). Chaque *STP* est associé à l’ensemble des *features* (voir Figure 2.11).

Le problème avec cette approche réside dans le fait que la vision de l’objet (i.e., *feature*) n’est pas totalement dissociée de la vision champ (STP). En effet, les objets ne peuvent pas avoir leurs propres attributs indépendamment de ceux associés aux STPs. Le modèle peut apparaître basique, mais il est basé sur une compréhension solide du domaine géographique et permet de gérer un ensemble important de cas ou scénarios complexes. Plus particulièrement, il permet de gérer les requêtes d’interaction entre objets (OIQ). Les auteurs présentent une évaluation de la performance de leurs modèles en faisant un test avec une masse de données importantes de points spatio-temporels. Enfin, le modèle propose une interface utilisateur permettant d’interroger les données.

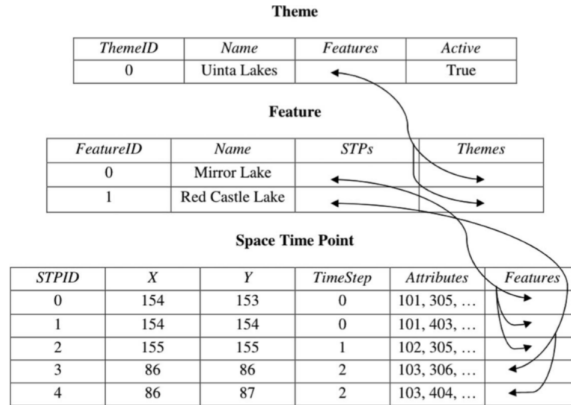


FIGURE 2.11 – Structure de données du modèle EDGIS (Pultar et al., 2010)

Modèle pour représenter des variations internes des objets mobiles :

Ce modèle est proposé par (Ibrahim and Kraak., 2014), dont l'objectif est de gérer des entités géographiques mobiles ayant des propriétés continues qui varient au sein de leur emprise spatio-temporelle. Il s'agit plus particulièrement d'enrichir des bases de données spatio-temporelles avec des types de données et des opérations capables d'étendre les capacités de leurs langages de requête.

L'unité basique sur laquelle ils construisent leur modèle est la notion d'observation. Celle-ci est définie comme une paire valeur (v) - instant (t). Le type de la valeur v est générique, il peut être primitif (ex. entier, décimal), géométrique (ex. point, polygone) ou composé en pixels (position + valeur numérique). Ensuite, sur la base de cette construction, est définie la notion de type mobile comme étant un ensemble ordonné, n , d'observations avec la fonction d'évolution (f). Cette dernière, décrit le changement entre deux observations ou plusieurs observations, dont l'évolution peut être constante, linéaire, sinusoïdal, etc. L'objectif de cette fonction est d'assurer la continuité et d'aider à interroger les données pour différentes granularités temporelles.

Sur cette base, les auteurs représentent un champ spatio-temporel qui est ensuite relié à un objet mobile. Pour cela leur modèle de données est basé essentiellement sur la définition de deux types de données : *mraster* (i.e., raster mobile), et *CType* (i.e., type composé). Concernant le type *mraster*, il représente l'ensemble des observations comme une collection de paires (v, t) où v est un pixel d'un raster. En appliquant le type mobile sur cette notion,

ils obtiennent une séquence ou une collection ordonnée d'observations raster $(t, pixel)$ qui pourrait être de 2 à n . Chaque collection est associée, à une fonction d'évolution qui décrit la façon dont les valeurs changent en pixels. Le deuxième type de données *CType* représente l'hétérogénéité d'un objet mobile à travers la composition de deux attributs. Le premier est discret, plus particulièrement géométrique, et correspond à son emprise mobile, et l'autre est continue et stocke les références pointant sur un ensemble d'enregistrements du type *mraster*. Ce dernier décrit les caractéristiques internes telles que la température à l'intérieur de cette emprise spatio-temporelle. A ce titre, la fonction *ST_snap* est proposée, qui prend comme argument un objet mobile hétérogène A de type *CType* et un instant t et renvoie la couverture raster qui se situe à l'intérieur de l'emprise mobile de l'objet A à l'instant t . Malgré le grand intérêt du modèle, le nombre d'opérateurs est limité et le modèle ne prend pas en compte le concept d'occurent/événement dans sa modélisation conceptuelle.

2.10.2 Approche basée sur une intégration à travers un concept abstrait

Il s'agit de définir un concept générique avec un niveau d'abstraction élevé et ensuite de se servir de mécanismes de spécialisation pour dériver des concepts permettant de représenter l'espace géographique.

Modèle G-champ :

(Liu et al., 2008) proposent un cadre conceptuel descriptif pour un modèle, appelé champ générique (G-champ). Ce concept est défini comme une association f entre un domaine spatio-temporel X constitué d'un ensemble de points ayant quatre dimensions (x, y, z et temps) avec l'ensemble des valeurs V dont son type est générique. Les auteurs proposent que l'image de G-champ, $img(f)$, qu'on peut définir comme le sous-ensemble $V_i \subseteq V$ de toutes valeurs associées à son domaine X , pourrait être catégorisée en plusieurs classes de valeurs, de telle sorte que pour chaque classe ou type (ex. scalaires, vecteurs, types composés, etc.), nous pourrions obtenir un nouveau type de champ (ex. champ de précipitation étant un champ scalaire, vitesse de vent étant un champ vectoriel). Le cas où la classe de valeurs appartient aux types d'objets, le concept hybride d'objet-champ (voir la définition dans la section 2.3) est dérivé. Le concept est caractérisé par un niveau d'abstraction assez puissant puisque les notions du champ et de l'objet-champ ne sont que des cas spéciaux du concept de G-champ.

Cependant et contrairement au travail de (Goodchild and Yuan, 2007), d'une part les auteurs ne proposent pas de modèle de données pour concrétiser et implémenter ce concept, et d'autre part, le travail ne traite pas l'évolution des objets dans le temps. De plus aucun moyen pour dériver la variation interne d'un objet à partir de G-champ n'a été proposé.

Un modèle de données basé sur la perspective Objet-champ pour les phénomènes géographiques mobiles :

Un autre travail qui rentre dans cette catégorie est celui de (Kim and Kiyoki, 2010), mais avec un processus inverse, qui considère le concept d'objet mobile comme étant un concept de haut niveau d'abstraction, et dérive la notion de champ mobile comme spécialisation de celui-ci où ce dernier hérite de toutes les caractéristiques et opérations d'un objet. Tout comme (Ibrahim 2016), ce modèle étend les modèles objets mobiles en proposant un modèle de données représentant la structure interne d'un objet mobile par le biais du champ continu. Le modèle définit un champ comme une fonction $f: S \rightarrow V_f(t)$ qui relie un domaine spatial S avec son codomaine $V(f)$, représentant l'ensemble des valeurs du champ f variant dans le temps. Pour prendre en considération tous les modèles spatiaux qui discrétisent un champ continu, le modèle définit le domaine S comme une classe générique qui pourrait être composée non seulement par des points, mais aussi par d'autres géométries (ex. ligne, polygone, polyèdres, etc.).

Au niveau du stockage dans la base de données spatiales, l'ensemble des observations liées à l'objet mobile est décomposé en un ensemble de séquences, et le champ mobile est divisé en un sous-ensemble d'unités où, pour chaque unité, une fonction d'interpolation temporelle est affectée. Chaque unité est composée d'une séquence de tuples $[(p, v1, t1), (p, v2, t2), (p, v3, t3), f_i]$ où chaque tuple est composé d'une géométrie p , d'une valeur v , ainsi qu'un instant t . Cette séquence de tuples est associée à une fonction f_i qui définit une méthode d'interpolation temporelle. Le champ mobile est ensuite associé à un ensemble d'opérations. Cet ensemble d'opérations est classé en plusieurs catégories : fonction de couverture (interpolation spatiale), fonction de déplacement (interpolation temporelle) et autres : superposition, classification, etc. Malgré le grand intérêt du modèle, l'ensemble des opérateurs hybrides permettant de relier les deux approches au niveau des requêtes complexes n'est pas complet. De même, le modèle ne tient pas compte du concept d'occurent/événement dans sa modélisation conceptuelle.

Modèle de données basé sur le concept de phénomène :

Ce modèle est proposé par (De Bakker, 2017) pour gérer les données et les phénomènes environnementaux qui combinent les agents (terme utilisé en écologie qui désigne les individus et les objets en science de l'information géographique) avec la notion du champ. Le modèle de données conceptuel donne une représentation uniforme de plusieurs types de données à un niveau d'abstraction élevé, où le concept de phénomène, appelé *phenomenon*, survole cette hiérarchie. Ce concept est modélisé comme un ensemble d'éléments qui partagent une collection des propriétés, mais aussi un domaine spatial et temporel de même type. La collection des propriétés est appelée *propertySet* dont chaque propriété est associée à la valeur appelée *value*. Le domaine correspond au type de l'emprise spatio-temporelle (région, ligne, point) qu'un ensemble de propriétés partagent (Figure 2.12).

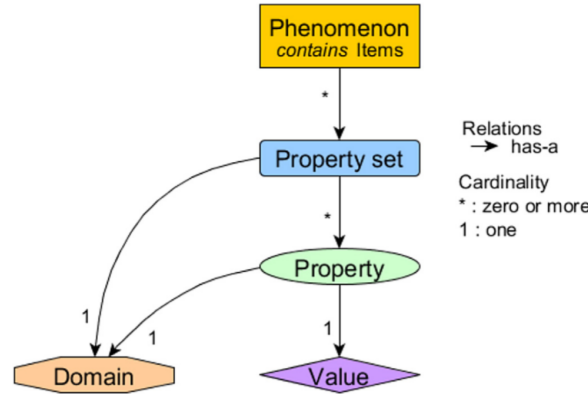


FIGURE 2.12 – Modèle de données conceptuel pour l'intégration de l'approche hybride (De Bakker, 2017)

Le modèle propose une séparation entre les attributs sémantiques et spatio-temporels (appelés domaine), qui tous les deux caractérisent un élément d'un phénomène. La notion de *propertyset* fait référence à un et un seul *domaine*, et un *domaine* appartient à un et un seul *propertyset*. Prenons l'exemple du phénomène d'évolution des réserves écologiques. Une réserve écologique est une zone ayant une emprise spatiale et temporelle bien définie considérée comme une instance du *domaine*. Cette instance fait référence à un et un seul *propertyset* qui est dans ce cas l'ensemble de propriétés qui caractérisent cette réserve biologique (i.e., le nombre d'espèces vivantes, la surface totale, etc.) En utilisant trois concepts *domaine*, *propertyset* et *value*, le modèle peut instancier le concept de champ (par exemple, l'élévation, la température), ainsi que le concept d'objet mobile (considéré dans l'article comme un agent). Plus précisément, le modèle définit la notion du champ

comme un type de données. Celui-ci peut avoir une ou plusieurs propriétés. Chaque propriété est représentée par un tableau bidimensionnel de paire pixel-temps. En ce qui concerne la vision objet, le modèle définit un type de données qui s'appelle agent mobile qui stocke, pour chaque agent, son empreinte spatiale ainsi que l'ensemble de propriétés avec leurs valeurs. Afin de permettre la liaison entre les deux visions objet et champ, le modèle permet, via le tableau bidimensionnel de paire pixel-temps, de représenter un champ à l'intérieur d'une empreinte spatio-temporelle d'un agent mobile. Ce modèle de données est utilisé comme entrée d'un programme en python implémentant les opérations de map algebra. L'utilisation de ce programme réduit les capacités de mises à jour et l'usage de fonctions et opérateurs d'interrogation spatiale (intersection, couvertures, etc.). De plus, le modèle entraîne une forte redondance de données, dans le cas de stockage de plusieurs propriétés différentes pour un même objet.

2.10.3 Approche basée sur l'association des visions champ et objets

Cette approche peut être illustrée par les travaux introduits par ([Galton, 2001a](#); [Cova, 2002](#)) où les notions de champ et celle d'objets sont conceptualisées indépendamment, puis associées par un mapping formel qui a abouti à un concept hybride d'objet-champ (voir la section 3). Malgré le fait que les visions de champ et d'objet sont bien appréhendées par ces modèles, l'idée principale n'est pas une conceptualisation hybride de l'espace, mais une correspondance entre les deux visions. De plus, ces modèles ne sont pas associés à la dimension temporelle. L'extension temporelle de ces concepts a été développée par ([Hamdani et al., 2019](#)). Dans notre cas, nous explorerons cette approche plus en détail dans le prochain chapitre.

2.11 Évaluation et analyse des modèles hybrides

Dans cette section, nous comparons les modèles spatio-temporels hybrides sur la base de caractéristiques qui génèrent un cadre d'évaluation commun. Pour cela, nous avons établi un ensemble de critères de comparaison qu'on classe en deux catégories : critères sémantiques et critères non fonctionnels.

Concernant les critères sémantiques, nous distinguons les éléments suivants :

- *Gestion de la vision continue dans le temps* : le modèle supporte la représentation de l'espace en champ en intégrant la composante temporelle. Autrement dit le modèle gère plusieurs propriétés, localisées et mesurées dans l'espace et dans le temps.
- *Gestion de la vision discrète dans le temps* : le modèle supporte la représentation de l'espace en entités ou objets intégrant la composante temporelle.
- *Intégration vision champ/vision discrète* : le modèle établit une liaison explicite entre les deux visions de l'espace dans le temps, et considère l'ensemble des interconnexions entre les deux visions.
- *Intégration continuant/occurent* : le modèle gère les deux concepts dans un seul modèle avec la présence d'une liaison explicite entre les deux visions de l'espace discret.
- *Granularité spatio-temporelle* : le modèle introduit des mécanismes permettant de supporter des requêtes dans différents niveaux de granularités spatiales et temporelles.
- *Gestion des relations spatiales dans le temps* : le modèle prend en compte des structures spécifiques ainsi que des opérations concernant les relations spatiales (ex. topologiques), temporelles et spatio-temporelles, et définit un langage de requête ou des opérateurs permettant des requêtes spatio-temporelles.

En ce qui concerne les critères non fonctionnels, dans notre cas, ils correspondent aux éléments suivants :

- *Développement formel* : le modèle présente un fondement formel lui permettant d'être générique avec un niveau d'abstraction élevé et vérifiable en termes de cohérence.
- *Respect de processus de modélisation de données* : le modèle est présenté sous forme conceptuelle, logique avec une implémentation.
- *Test d'application* : le modèle est testé sur un ou plusieurs cas d'études et s'il est confronté à un ensemble de scénarios.
- *Extensibilité* : le modèle présente un ensemble de types de données et

d'opérateurs génériques.

- *Présence des opérateurs et fonctions hybrides* : le modèle développe un ensemble de fonctions qui permettent d'explorer l'interaction dans le temps entre les deux modèles champ et objet.
- *Présence des opérateurs d'analyse spatio-temporelle* : le modèle implémente un ensemble de fonctions, mécanismes et méthodes qui permettent de manipuler, interroger et analyser les données spatio-temporelles.
- *Teste de performance et du passage à l'échelle du modèle* : Le modèle présente ses performances concernant le temps de réponse à différentes requêtes hybrides, en fonction de leur complexité et de la quantité des données prises en compte.

Les deux tableaux suivants comparent les différents modèles selon les deux catégories des critères (Tableaux 2.1 et 2.2) en adoptant les trois notes suivantes : (-) qui signifie *non Géré*, (+) signifie *Géré* et (++) qui signifie *Développé*.

Modèles de données Hybrides	Critères Sémantiques		Field + Temps	Objet + Temps	Intégration Field/Objet	Intégration Continuant/ Occurrent	Relation Spatio-temporelle	Granularité Spatio-temporelle
	Type de données							
EDGIS (Pultar et al. 2010)	Spatio-temporal point (STP)		++	++	+	-	-	-
Internal Varying of Moving Object (Ibrahim et al. 2014)	Observation, Mraster, Ctype		++	++	+	-	-	+
G-Field (Liu et al. 2008)	Generic Field, Object-field		++	+	+	-	-	-
Object-Field Perspective (Kim et al. 2010)	Moving Object, Moving Field		++	++	+	-	-	+
Agent and Field based Integration (De Bakker et al. 2017)	Phenomenon, Agent, Field		++	+	+	-	-	-
A framework for distributed phenomena (Mcintosh et al. 2005)	Point, Zone, Event, Process		++	++	+	-	++	-
Modèle Field-Object (Yuan 2001)	Field, Event, Process		+	+	+	-	-	-

TABLE 2.1 – Tableau comparatif évaluant les modèles hybrides selon les critères sémantiques

Tous les modèles hybrides intègrent les deux visions champ et objet dans un seul modèle mais aucun d'entre eux ne couvre toutes les caractéristiques escomptées.

Modèle Hybrides	Critères non fonctionnelles	Formelle	Processus de modélisation des données			Test d'application	Fonction Hybrides	Opérateur d'analyse spatio-temporelle	Extensibilité	Tests de Performance et scalability
			Conceptuel	Logique	Implémentation					
	<i>EDGIS (Pultar et al. 2010)</i>	+	++	++	++	++	-	-	-	++
	<i>Internal Varying of Moving Object (Ibrahim et al. 2014)</i>	+	-	++	++	+	+	+	+	-
	<i>G-Field (Liu et al. 2008)</i>	++	+	-	-	-	+	-	-	-
	<i>Object-Field Perspective (Kim et al. 2010)</i>	+	++	-	+	+	+	+	+	-
	<i>Agent and Field based Integration (De Bakker et al. 2017)</i>	-	++	++	++	+	-	++	+	-
	<i>A framework for distributed phenomena (Mcintosh et al. 2005)</i>	+	+	++	++	+	-	+	-	-
	<i>Modèle Field-Object (Yuan 2001)</i>	-	+	++	++	+	-	-	-	-

TABLE 2.2 – Tableau comparatif évaluant les modèles hybrides selon les critères non fonctionnels

Aucun ne couvre l'ensemble de l'attendu. En effet il n'est pas suffisant de combiner les deux conceptions, ni d'intégrer explicitement leurs liaisons dans un modèle logique et physique des données. Il est donc nécessaire de disposer d'un arsenal de mécanismes et d'opérateurs qui permettent d'explorer l'interaction entre la vision objet et la vision champ (continue). Il s'agit de permettre l'expression d'opérateurs de manipulation et d'analyse spatio-temporelle en les appliquant sur un phénomène géographique complexe. Enfin, il faut permettre de gérer plusieurs scénarios et répondre à des requêtes hybrides elles aussi très complexes. Aucun des modèles ne permet une intégration fine des notions de continuant et de occurrent. Par ailleurs, la majorité des modèles hybrides ne gère pas l'évolution temporelle des relations spatiales entre objets. Il en va de même pour la gestion de la granularité spatiale et temporelle, de même que pour la gestion des objets en mouvement.

Cette analyse comparative nous amène à définir clairement l'objectif de cette thèse qui consiste à développer une représentation hybride de l'espace géographique accompagné d'un modèle de données spatio-temporel qui va permettre de gérer l'ensemble des requêtes concernant les phénomènes dits complexes ou multi-aspects. Plus spécifiquement, nous pouvons articuler notre objectif autour des points suivants :

- Dépasser la dichotomie entre les représentations continues et discrètes de l'espace géographique au niveau formel, conceptuel et logique des données.

- Dépasser la dichotomie qui existe dans la représentation discrète de l'espace géographique entre le concept d'endurant (i.e., continuant) et celle de perdurant (i.e., occurrent) au niveau conceptuel et logique des données.
- Explorer l'émergence des concepts résultant de l'interconnexion entre la vision discrète et la vision continue dans le temps.
- Adopter une approche formelle générique, suivie d'un modèle de données développé selon une approche conceptuelle, logique avec une implémentation.
- Proposer un modèle extensible afin de pouvoir l'enrichir avec d'autres concepts.
- Offrir des mécanismes tels que des opérateurs qui permettent d'explorer l'interaction entre l'évolution d'objets et de champ. Coupler ces mécanismes à des analyses spatio-temporelles et les appliquer à un cas réel.
- Proposer l'évaluation des performances et du passage à l'échelle en essayant de tester les limites avec des quantités importantes de données spatio-temporelles.

Chapitre 3

Vers une représentation hybride de l'espace géographique

"I believe that mathematical reality lies outside us, that our function is to discover or observe it, and that the theorems which we prove, and which we describe grandiloquently as our 'creations', are simply our notes of our observations."

G.H. Hardy

Depuis l'émergence des sciences de l'information géographique, et comme cela a été rapporté dans le chapitre précédent, plusieurs auteurs ont successivement contribué aux modèles objet et champ. Néanmoins, si les deux approches sont structurellement différentes, elles offrent des perspectives complémentaires. De nombreux travaux ont introduit une conceptualisation hybride dans laquelle deux notions principales ont émergé : la notion d'objet-champ (Galton, 2001a; Cova, 2002; Liu et al., 2008; Cova, 2016) et la notion de champ-objet (Yuan, 2001; Goodchild and Yuan, 2007; Cova, 2016). Cependant, ces concepts hybrides méritent encore des développements formels supplémentaires qui donneront un support solide à leur représentation, et une meilleure intégration de la dimension temporelle. Sur la base de ces développements méthodologiques, il convient de montrer comment une représentation hybride pourrait contribuer à améliorer notre compréhension

des phénomènes géographiques. L'approche développée dans ce chapitre, formalise ces concepts hybrides et prend en compte une intégration étroite de la dimension temporelle. Le second objectif est d'explorer dans quelle mesure une représentation hybride peut fournir des mécanismes nouveaux pour étudier des phénomènes géographiques complexes. Cela est d'abord réalisé au niveau formel où deux perspectives de l'approche hybride sont définies, puis conceptuellement en utilisant un modèle de données et une spécification d'une série de catégories de requêtes qui manipulent les représentations basées sur l'objet et le champ. L'approche de modélisation est illustrée par des exemples concernant essentiellement les phénomènes d'évolution des dunes dans un environnement marin. En effet, cet exemple représente un cas pertinent de phénomènes environnementaux pour lesquels une représentation hybride a un intérêt.

3.1 Concepts spatio-temporels principaux

Considérons l'espace comme étant dual et étant formé d'un ensemble de points L ainsi que d'un ensemble d'entités géographiques G , ces dernières peuvent être soit des continuants C soit des objets perdurants appelés des occurrents O_c . Plus précisément, l'ensemble L désigne un domaine spatial qui est défini comme un ensemble de points spatiaux tel que $L = \{l_1, l_2, l_3, \dots, l_n\}$ et $\mathcal{P}(L)$ est un ensemble de tous les sous-ensembles de L .

Nous représentons le temps par le domaine $(T', <)$ dont T' est un ensemble des instants tels que $T' = \{t'_1, t'_2, t'_3, \dots, t'_n\}$ est ordonné par la relation $<$. Nous définissons ensuite T comme le sous-ensemble $T \subset T' \times T'$ constitué des paires d'instant (t'_i, t'_j) pour tout $t'_i \leq t'_j$. Un élément $t \in T$ pourrait dénoter un intervalle si et seulement si $t'_i < t'_j$ dont (t'_i) et (t'_j) constituent respectivement le début et la fin de l'intervalle t . Dans le cas où $t'_i = t'_j$, nous considérerons t comme un instant. Nous désignons $L \times T$ comme un ensemble de points spatio-temporels $(l, t) \in L \times T$.

Nous définissons la valeur v comme une paire constituée d'une propriété z et de sa mesure m (ex. $v = (\text{temperature}, 23.2)$). Par conséquent $V = (Z, M)$ sera défini comme l'ensemble des paires (z, m) . Nous exprimons le concept suivant $\mathcal{P}(V)$ comme l'ensemble de tous les sous-ensembles des valeurs V .

Nous définissons une observation o_b comme la paire $(t, f_o(t))$ telle que t correspond à un instant ou intervalle temporel et f_o correspond à une fonction qui associe pour chaque t une valeur $v \in V$ dont $f_o: T \rightarrow \mathcal{P}(V)$.

Avec cette définition, les éléments de $\mathcal{P}(V)$, et par conséquent les valeurs de f_o , peuvent correspondre à la valeur nulle, ou à une valeur singulière v , ou à une agrégation de valeurs V_i . Une valeur nulle est représentée par l'ensemble vide, alors que pour une valeur unique, $\mathcal{P}(V)$ est limité à un singleton (i.e., un ensemble contenant un seul élément), et il n'existe aucune restriction sur l'agrégation des valeurs V_i .

3.2 Définition d'un champ spatio-temporel

Nous proposons dans cette section une nouvelle redéfinition de la notion de champ qui étend celle qui a été définie. Le terme *field* pourra également utilisé pour se rapprocher des notions décrites dans la littérature. Le but est de formuler une définition de champs plus générique qui permet d'englober l'ensemble des propriétés spatialement distribuées (e.g. radiation, température, etc.). Formellement, un *field* f était défini par (Galton, 2001a) comme une fonction $f: L \rightarrow V$, où L correspond à son domaine et dénote un ensemble de localisations spatiales, et V son codomaine constitué d'une collection de valeurs qui dénotent des propriétés continues. Nous étendons cette définition pour la rendre plus générique avec $f: L \rightarrow \mathcal{P}(V)$. Le domaine L et son codomaine $\mathcal{P}(V)$ sont définis comme suit :

- Domaine L : ce domaine est défini comme un ensemble de localisations spatiales, considérées comme des primitives fondamentales pour le concept de *field*. Les éléments de L désignent soit des points distribués régulièrement dans l'espace (Figure 2.2.A) ou irrégulièrement (Figure 2.2.B), soit des partitions régulières de l'espace (des pixels par exemple) (Figure 2.2.C) ou irrégulières (TIN par exemple) (Figure 2.2.E).
- Codomaine $\mathcal{P}(V)$: comme indiqué ci-dessus, V désigne un ensemble de paires (z, m) tandis que $\mathcal{P}(V)$ désigne un ensemble de tous les sous-ensembles de V . Pour chaque localisation l , $\mathcal{P}(V)$ peut être exprimé comme $\{z_1, m_1, z_2, m_2, \dots, z_k, m_k\}$ où z_1, z_2, z_k représentent des propriétés continues de l'espace (ex. pression atmosphérique, élévation, courant de marée, etc.). Cela permet de préciser plusieurs propriétés pour un seul emplacement spatial l et donc pour un champ.

Suivant la manière dont les domaines L et T sont organisés, plusieurs fonctions pourraient être définies pour étendre la notion de champ (i.e., la fonction f), vers le domaine temporel. Nous définissons dans un premier

temps une fonction $state: L \times T \rightarrow \mathcal{P}(V)$ qui attribue un sous-ensemble de $\mathcal{P}(V)$ pour chaque tuple (l, t) issu du produit cartésien entre le domaine spatial et le domaine temporel $L \times T$.

Ensuite, et afin de dériver un champ temporel d'un champ spatial, nous nous appliquons le processus de *currying* qui transforme une fonction avec plus d'un paramètre en une séquence de fonctions, chacune avec un seul paramètre. On obtient la fonction $hist: L \rightarrow (T \rightarrow \mathcal{P}(V))$ qui, à partir d'un emplacement spatial l , renvoie un ensemble des paires temps-valeurs. Cette fonction peut être exprimée différemment avec $hist: L \rightarrow f_o(t)$ qui relie un domaine spatial L avec son codomaine $f_o(t)$ représentant l'ensemble des valeurs du *field* variant dans le temps.

Il est à noter que la définition de la deuxième fonction (i.e., *hist*) suppose qu'à tout moment nous avons le même échantillon de localisations spatiales l .

3.3 Définition d'un objet spatio-temporel

3.3.1 Continuant versus occurrent

Pour l'ensemble des entités géographiques composant l'espace (i.e., G) une différence est faite entre le concept de continuant et celui d'occurrent.

Tout d'abord, un continuant c est défini comme une entité qui persiste dans le temps, subit des changements mais préserve son identité. Ces changements reflètent une évolution de sa géométrie, de ses propriétés et de ses relations avec les objets voisins (ex. relations topologiques, relations cardinales, etc.). Un continuant possède un historique et existe dans son entier à chaque état, c'est à dire à chaque instant de sa vie. Les fonctions qui délivrent l'état et l'histoire d'un continuant c sont définies comme suit :

- $CState: C \times T \rightarrow \mathcal{P}(V)$: la fonction $CState$, associe à un continuant c pour chaque instant valide t , une valeur du domaine $\mathcal{P}(V)$. On appelle le tuple $(c, t) \in C \times T$ l'état du continuant c . En nous servant de la fonction $CState$ nous pouvons affecter plusieurs propriétés pour un seul état d'un continuant c .
- $CHist: C \rightarrow (T \times \mathcal{P}(V))$: la fonction $CHist$ renvoie pour chaque continuant c un ensemble de paires temps-valeur appartenant au domaine $(T \times \mathcal{P}(V))$. Cette fonction peut être exprimée différemment

avec $CHist: C \rightarrow f_o(t)$ qui relie un ensemble des continnants C avec son codomaine $f_o(t)$ représentant l'ensemble des valeurs variant dans le temps.

Un *occurent* désigne une entité qui s'étend sur toute une période de temps allant de son début à sa fin mais, contrairement à un continuant, ne peut pas subir de changement et surtout, n'existe pas dans son intégralité à un moment donné (Grenon and Smith, 2004; Galton, 2004, 2006). Les propriétés sémantiques d'un *occurent* o_c pourraient ainsi être représentées par la fonction suivante : $OState: O_c \rightarrow \mathcal{P}(V \times T)$. Cette fonction indique que pour chaque *occurent* o_c , nous associons un sous-ensemble de toutes les paires valeur-temps qui correspondent aux attributs de l'élément o_c .

Suite à cette dichotomie, puis en termes d'observation des phénomènes dynamiques dans l'espace géographique, nous considérons que l'ensemble de toutes les entités observées G est le résultat d'une union disjointe de tous les *occurents* O_c et de tous les continnants C . Nous pourrions représenter cela par la fonction $f_{ge}: C \cup O_c \rightarrow G$. Cette forme est considérée comme une généralisation des deux concepts de continuant et d'*occurent* à celui d'entité géographique. Comme cette dernière possède une emprise spatiale et temporelle, les deux concepts continuant et *occurent* héritent de ces caractéristiques spatio-temporelles où chacun a la possibilité d'avoir ses propres collections d'attributs.

3.3.2 Relations spatiales et spatio-temporelles

Outre les attributs qui caractérisent les objets géographiques, ces derniers sont caractérisés par l'ensemble des relations spatiales et spatio-temporelles qui les lient entre eux. Dans cette section nous formalisons un ensemble de concepts qui caractérisent cette propriété.

Nous définissons tout d'abord le domaine $H = C \times T$ comme l'ensemble des paires (c, t) qui représente l'état d'un continuant c à un temps t . Ensuite, nous définissons deux relations binaires R_s et R_{st} , que nous appelons relation spatiale et relation spatio-temporelle, sur le domaine H . Nous désignons $sp(h)$ comme l'empreinte spatiale de l'entité $h \in H$ et $\propto$ comme l'opérateur qui vérifie l'existence d'une relation spatiale entre deux empreintes spatiales.

La relation spatiale est définie comme suit :

$$R_s \subset H \times H \text{ avec}$$

$$R_s = \{(h_1, h_2) \mid sp(h_1) \propto sp(h_2) \wedge time(h_1) = time(h_2) \text{ tel que } h_1 \neq h_2\}$$

Cette relation contient seulement les paires des états (h_1, h_2) présents dans le même temps et dont leurs emprises spatiales sont en relation spatiale.

En ce qui concerne la relation spatio-temporelle, nous la définissons comme suit :

$$R_{st} \subset H \times H \text{ avec}$$

$$R_{st} = \{(h_1, h_2) \mid sp(h_1) \propto sp(h_2) \wedge time(h_1) < time(h_2)\}$$

Cette relation contient seulement les paires des états (h_1, h_2) présents dans deux temps consécutifs et dont leurs emprises spatiales sont en relation spatiale au travers du temps.

Nous dénotons R_a comme l'ensemble des paires (h_i, h_j) ayant une relation spatiale ou une relation spatio-temporelle : $R_a = R_s \cup R_{st}$.

Nous définissons une fonction de qualification $f_q: R_a \rightarrow V_q$ qui nous permet d'attribuer une valeur v_q à chaque relation $(h_i, h_j) \in R_a$. Ces valeurs peuvent être de deux types :

- des valeurs numériques pour représenter les types de relations métriques.
- des valeurs textuelles descriptives pour représenter les types de relations qualitatives comme les relations topologiques ou relation d'orientation.

L'image de la fonction f_q , dénotée $img(f_q)$, représente l'ensemble de toutes les valeurs v_q associées à l'ensemble des éléments de R_a .

La représentation de ces relations par un graphe spatio-temporel permet d'extraire plusieurs opérateurs et fonctions (Del Mondo et al., 2010; Del Mondo, 2011). Cependant, dans notre cas, nous nous limitons à deux fonctions qui nous semblent utiles dans notre démarche et qui nous aideront à formuler des requêtes spatio-temporelles. Le but de ce processus formel est d'aboutir aux deux fonctions spatio-temporelles. La première permet d'explorer les voisinages d'un continuant dans un instant ou intervalle donné

et la seconde permet d'explorer l'ensemble de routes d'un continuant pour un intervalle donné.

La première fonction est la fonction $f_{vg}(c_i, t)$:

$$f_{vg}(c_i, t) = \{(c_j, t', v_q) \mid \forall t' \in t, c_i \neq c_j \text{ tel que } (h_i, h_j) \in R_s \wedge v_q = f_q(h_i, h_j)\}$$

Nous précisons que $h_i = (c_i, t')$ et $h_j = (c_j, t')$ désignent deux états de deux continuant c_i et c_j qui sont présents au même instant temporel t' . La fonction f_{vg} renvoie tous les états du continuant c_j qui sont voisins au continuant c_i par la relation spatiale R_s dans chaque temps $t' \in t$. Chaque état renvoyé est associé avec une valeur v_q qui le qualifie. Par exemple si la relation spatiale est de type métrique, la valeur v_q va dénoter la distance entre les états de c_i et de c_j dans chaque instant t' .

La deuxième fonction renvoie l'ensemble des routes possibles pour un continuant c_i . Cette notion de route a été définie par (Del Mondo, 2011) au niveau d'un graphe spatio-temporel comme étant le chemin parcouru dans ce graphe renvoyant la séquence temporelle des noeuds (i.e., entités spatiales) situés entre le noeud initial et le noeud terminal.

Nous la définissons comme suit :

$$f_{rt}(c_i, t) = \{(c_j, t_j, v_q) \mid \forall t_i < t_j \in t \text{ tel que } (h_i, h_j) \in R_{st} \wedge v_q = f_q(h_i, h_j)\}$$

Nous précisons que $h_i = (c_i, t_i)$ et $h_j = (c_j, t_j)$ désignent deux états de deux continuants c_i et c_j qui sont présents dans deux instants consécutifs t_i et t_j .

La fonction f_{rt} renvoie tous les états du continuant c_j qui succèdent aux états du continuant c_i par la relation spatio-temporelle R_{st} dans chaque instant $t' \in t$. Chaque état renvoyé est associé avec une valeur v_q qui le qualifie.

3.4 Vision hybride du temps

Avant de traiter la vision hybride spatio-temporelle entre les deux visions de l'espace, nous proposons dans un premier temps d'introduire la vision hybride du temps (voir la sous-section 2.4.2). La notion de temps introduit une dichotomie entre *field* temporel et objet temporel (Galton, 2003, 2004). Par conséquent, nous présentons, l'équivalent temporel des deux concepts *objet-field* et *field-objet*.

3.4.1 Objet-champ temporel

Considérons l'ensemble d'objets temporels E dont chaque élément $e \in E$ possède une emprise temporelle, qui constitue l'équivalent de l'emprise spatiale d'un objet spatial. Nous représentons cette emprise temporelle par la fonction $E_t: E \rightarrow T \times T$. Cette fonction renvoie pour chaque objet temporel e un intervalle temporel $t = [t_i, t_j]$ dans lequel l'objet e persiste. Au cas où $t_i = t_j$ alors on considère que l'objet e a une emprise temporelle ponctuelle.

Nous définissons un *objet-field* temporel comme une paire ordonnée (t, of_e) dont t est un instant ou intervalle temporel et of_e est un *field* temporel dont ses valeurs correspondent au sous-ensemble d'objets temporels $E_i \subseteq E$, tel que :

$$of_e: T \rightarrow \mathcal{P}(E) \text{ avec}$$

$$of_e(t) = \{E_i \in \mathcal{P}(E) \mid \forall e \in E_i \text{ alors } t \cap E_t(e) \neq \emptyset\}$$

Une contrainte temporelle est appliquée sur les valeurs de la fonction of_e qui stipule que l'ensemble des temps t liés à un objet temporel $e \in E_i$ doivent forcément avoir une relation temporelle d'intersection avec son emprise temporelle renvoyée par la fonction $E_t(e)$. Cette définition préserve la relation fonctionnelle entre le domaine temporel et les événements e associés. Cela est particulièrement pertinent lorsqu'une période temporelle est attribuée à plusieurs objets temporels (i.e., dans le cas d'objets qui se chevauchent temporellement).

3.4.2 Champ-objet temporel

Nous définissons le concept appelé *field-objet* temporel comme étant une sorte de réciproque du concept d'*objet-field* temporel défini ci-dessus. Ce concept est défini comme une paire ordonnée (e, fo_e) où e est un objet temporel et fo_e est une fonction qui renvoie pour chaque élément e un ensemble d'observations o_b . Celle-ci (i.e., o_b) a été définie comme une paire (t, f_o) dont f_o associe pour chaque instant t une valeur $v \in V$. Formellement fo_e est défini comme suit :

$$fo_e: E \rightarrow O_b(t) \text{ avec}$$

$$fo_e(e) = \{(t, v) \in O_b(t) \mid t \cap E_t(e) \neq \emptyset\}$$

La même contrainte temporelle est appliquée sur les valeurs de la fonction fo_e qui stipule que l'ensemble des temps t faisant partie d'une observation

$o_b \in O_b$, et liés à un objet temporel e , doivent forcément avoir une relation temporelle d'intersection avec son emprise temporelle renvoyée par la fonction $E_t(e)$.

3.5 Vers une représentation spatio-temporelle Hybride

Notre état de l'art a montré que le développement de fonctions permettant la manipulation de l'approche spatio-temporelle Hybride était encore insuffisant pour permettre l'analyse de phénomènes géographiques complexes. Nous proposons ici de développer les concepts nécessaires pour explorer chaque aspect de l'interaction des deux visions de l'espace (i.e., *field* et objet). Dans cette section, nous présenterons cette interaction sous forme de diverses fonctions hybrides, puis nous les illustrerons à l'aide d'exemples avec une série d'interrogations associées à l'étude de la géodynamique marine.

3.5.1 Association des deux visions de l'espace *field* et objet à travers le temps

Dans un premier temps, il s'agit de définir une fonction de base permettant le "passage" entre les deux représentations de l'espace. Sur la base de la fonction $Cstate$, en remplaçant le domaine V par le domaine L (i.e., l'ensemble des localisations spatiales), nous obtenons la fonction suivante : $f_l: C \times T \rightarrow \mathcal{P}(L)$. Cette fonction renvoie un sous-ensemble de localisations spatiales $L_i \subseteq L$ associés à un continuant c pendant un domaine temporel valide t . Cette fonction relie explicitement l'ensemble des continnants avec le domaine L d'un *field*.

Pour matérialiser cette association, nous définissons la fonction $f_g: C \times T \rightarrow geom$ qui représente la composante spatiale d'un continuant c où *geom* correspond à la géométrie du c qui peut par exemple être un point, une ligne ou un polygone, etc.

Par la suite, nous définissons une contrainte spatio-temporelle qui restreint le résultat de la fonction f_l comme suit :

$$f_l(c, t) = \{L_i \in \mathcal{P}(L) \mid \forall l \in L_i \text{ alors } l \cap f_g(c, t) \neq \emptyset\}$$

La contrainte ci-dessus précise que pour un instant donné t , la composante spatiale du continuant c (i.e., $f_g(c, t)$) doit être spatialement intersectée avec les éléments appartenant à l'image de f_l étant donné les emplacements spatiaux l de L_i , tel que $L_i \subseteq L$ (Figure 3.1).

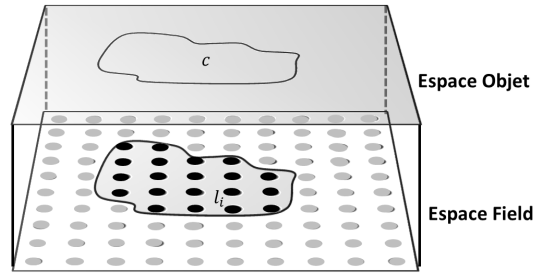


FIGURE 3.1 – Figure montrant l'association d'un continuant c avec une collection de localisations spatiales L_i

3.5.2 Vision Hybride : *Field* vers objet

Cette perspective hybride matérialise l'association entre les localisations spatiales qui constituent le domaine d'un *field* et une ou plusieurs entités géographiques $g \in G$, continuants ou occurrents. Cette dernière distinction conduit à dériver deux fonctions hybrides Ft_c et Ft_o . On distingue alors deux concepts :

Premièrement, le concept qui matérialise le passage d'un champ à un continuant est défini comme un triplet (l, t, Ft_c) où l est une localisation spatiale, t correspond à un instant ou intervalle temporel et Ft_c est la fonction renvoyant pour chaque paire (l, t) un sous-ensemble de continuants :

$$Ft_c: L \times T \rightarrow \mathcal{P}(C) \text{ avec} \quad (3.1)$$

$$Ft_c(l, t) = \{C_i \in \mathcal{P}(C) \mid \forall c \in C_i \text{ alors } l \in f_l(c, t)\}$$

Deuxièmement, le concept qui matérialise le un occurrent est défini sous la forme d'un triplet (l, t, Ft_o) où l est une localisation spatiale, en un instant ou intervalle temporel et Ft_o est la fonction hybride renvoyant pour chaque paire (l, t) un sous-ensemble d'occurrents :

$$Ft_o: L \times T \rightarrow \mathcal{P}(O_c) \text{ avec} \quad (3.2)$$

$$Ft_o(l, t) = \{O_i \in \mathcal{P}(O_c) \mid \forall o_c \in O_i \text{ alors } l \cap f_s(o_c) \neq \emptyset \wedge t \subseteq E_t(o_c)\}$$

Pour chaque collection de localisations spatiales $\{l_1, l_2, l_3, \dots, l_n\}$ et pour un temps donné t , les deux fonctions renvoient un sous-ensemble d'instances de GeoEntity, c'est-à-dire soit continuant, soit occurrent. Dans le cas d'un continuant (avec la fonction Ft_c) chaque localisation l doit appartenir au sous-ensemble de localisations renvoyé par la fonction $f_l(g, t)$. Dans le cas

d'un occurrent (avec la fonction Ft_o), les arguments de localisation et de temps doivent intersecter l'emprise spatiale (i.e., $fs(o_c)$) et temporelle (i.e., $E_t(o_c)$) de cet occurrent.

Les définitions de ces deux fonctions Ft_c et Ft_o préservent la relation fonctionnelle entre le *field* et les entités géographiques. En effet, une localisation spatiale l pourrait avoir une relation avec au moins un continuant et/ou un occurrent pour un temps donné t . Cela est particulièrement pertinent lorsqu'un élément spatio-temporel (l, t) est attribué à plusieurs continnants ou occurents (dans le cas où des entités géographiques se chevauchent spatialement et temporellement).

Ces deux fonctions permettent de décrire ou d'explorer l'impact des variations temporelles du *field* sur les propriétés de l'objet. Par exemple, cela pourrait permettre d'explorer comment le courant de marée (le *field*) impacte la vitesse de migration de dunes sous-marines (l'objet) dans une zone donnée.

Pour illustrer davantage le rôle de ces deux fonctions dans la manipulation des requêtes spatio-temporelles hybrides, prenons trois exemples :

Req_1) ***Recherchez l'ensemble des dunes ayant une superficie supérieure à 7 km² et dont la hauteur a augmenté de plus de 15 m entre deux instants t_1 et t_3 .***

Cette requête est décomposée en deux étapes comme suit :

$$Q := \{Ft_c(l, t) \mid hist(l)(t_3)[elev] - hist(l)(t_1)[elev] > 15\}$$

$$R := \{c \in Q \mid area(f_g(c, t_i)) > 7000 \text{ et } t_i \in [t_1, t_3]\}$$

Dans un premier temps, la requête utilise la fonction hybride Ft_c sur l'ensemble du champ. Pour chaque localisation spatiale l dont la différence de hauteur entre t_1 et t_3 est supérieure à 15 mètres, la fonction $Ft_c(l, t)$ identifie les entités c associées. Le résultat de cette étape est à affecter à la variable Q . L'étape suivante applique l'instance de R pour identifier le sous-ensemble des entités c qui ont une surface supérieure à 7km². Notons que Q et R sont deux variables arbitraires utilisées pour stocker le résultat de chaque étape de la requête.

Req_2) ***Recherchez la dune qui s'est rapprochée le plus de la dune c_A et dont la hauteur a diminué de plus de 2 m entre deux instants t_1 et t_2***

Nous appelons la fonction f_{vg} qui va renvoyer pour chaque dune l'ensemble de ses voisins en chaque instant et la distance qui les sépare.

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned} Q &:= \{ Ft_c(l, t) \mid hist(l)(t_2)[elev] - hist(l)(t_1)[elev] < -2 \} \\ R &:= \{ (c_i, rd) \mid c_i \in f_{vg}(c_A).c \wedge rd = f_{vg}(c_A)(Q.c, t_2) - f_{vg}(c_A)(Q.c, t_1) \} \\ W &:= \{ c_B \in R \mid c_B \in R.c_i \wedge R.rd = Min(R.rd) \} \end{aligned}$$

Dans un premier temps, la requête utilise la fonction hybride Ft_c sur l'ensemble du champ. Pour chaque localisation spatiale l dont la différence de hauteur entre t_2 et t_1 est supérieure à 2 mètres, la fonction $Ft_c(l, t)$ identifie les entités c associées. L'étape suivante consiste à calculer la grandeur rd qui évalue la variation de la distance entre la dune c_A et les dunes voisines entre les deux instants t_1 et t_2 à l'aide de la fonction f_{vg} . Ensuite, il s'agit d'extraire la dune c_B qui s'est le plus rapprochée de la dune c_A en utilisant la fonction $Min(R.rd)$.

Req_3) *Sur la période comprise entre t_1 et t_3 , recherchez les entités de type "orage" au cours desquelles la variation de la pente du fond marin affecté par l'orage a été supérieure à 30%.*

$$\begin{aligned} Q &:= \{ Ft_o(l, t) \mid |hist(l)(t_3)[pente] - hist(l)(t_1)[pente]| > 30 \} \\ R &:= \{ o_c \in Q \mid type(o_c) = "orage" \wedge t \cap E_t(o_c) \neq \emptyset \} \end{aligned}$$

Dans un premier temps, la requête applique la fonction hybride Ft_o . Pour chaque localisation spatiale l sur le fond marin, dont la valeur absolue de la différence de pente est supérieur à 30% entre t_2 et t_1 , la fonction $Ft_o(l, t)$ identifie les entités o_c associées. L'étape suivante applique l'instance de R pour identifier le sous-ensemble des occurrents o_c de type "orage" dont le temps de présence t doit appartenir à l'emprise temporelle d'au moins un occurrent.

3.5.3 Vision Hybride : Objet vers *field*

Cette seconde perspective, qui s'avère être proche de l'inverse de la première, concerne le lien entre l'objet et ses variations internes. Ces variations matérialisent la dynamique d'un *field* où à chaque localisation spatiale est affectée une valeur d'une propriété. Ainsi, on peut établir une correspondance étroite entre les entités géographiques et les espaces continus qui leur sont associés.

Dans le cas d'une dune sous-marine par exemple, on considère d'une part la dune comme un objet mobile avec son identité, sa forme, ses attributs (ex. longueur d'onde, hauteur moyenne, etc.) et ses relations avec d'autres objets (ex. une dune se trouve à l'intérieur d'un champ de dunes), et d'autre part comme un *field* avec des variations de surface telle que la profondeur ou la pente.

Des entités géographiques pouvant être modélisées sous forme d'un continuant et leurs propriétés définies comme un état ou un historique (voir section 3.3), deux fonctions supplémentaires appelées Ct_{fh} et Ct_{fs} sont définies comme suit :

La fonction Ct_{fh} attribue à chaque objet c et à un moment t (instant ou intervalle), un sous-ensemble de localisations spatiales avec leurs valeurs historiques associées. Ces valeurs dénotent des propriétés continues dans l'espace (ex. altitude, vitesse de vent, pression, etc.) dont le modèle permet de suivre l'évolution temporelle (cf Figure 3.2-a).

$$\begin{aligned} Ct_{fh}: C \times T &\rightarrow (L \rightarrow O_b) \quad \text{avec} \\ Ct_{fh}(c, t) &= \{(l, O_{bi}) \in (L, O_b) \mid \forall O_{bi} \cdot t_i \subseteq t\}, \\ &i) \ l \in f_l(c, t) \quad ii) \ O_{bi} \subseteq Hist(l) \end{aligned} \quad (3.3)$$

La valeur $Ct_{fh}(c, t)$ est contrainte par deux conditions : 1) la localisation spatiale l appartient au sous-ensemble des localisations spatiales renvoyées par la fonction f_l , 2) le sous-ensemble d'observations O_{bi} appartient à l'ensemble des paires temps-valeur (t, v) renvoyées par la fonction $Hist(l)$.

La fonction hybride Ct_{fs} , attribue à chaque objet c le sous-ensemble de localisations spatiales internes à cet objet pour un temps donné. Nous choisissons d'abord de définir le triplet $b = (l, t, state(l, t))$ dont l est un emplacement spatial, t représente un instant ou un intervalle temporel et $state(l, t)$ la valeur du *field* associée à l et t . Nous notons B comme l'ensemble des triplets notés b . Alors nous définissons la fonction Ct_{fs} comme suit :

$$\begin{aligned} Ct_{fs}: C \times T &\rightarrow \mathcal{P}(B) \quad \text{avec} \\ Ct_{fs}(c, t) &= \{B_i \in \mathcal{P}(B) \mid \forall b \in B_i \text{ alors } b \cdot t_i \subseteq t \wedge b \cdot l_i \in f_l(c, t)\} \end{aligned} \quad (3.4)$$

Cette fonction permet de décrire le champ interne de chacun des états d'un continuant c pour un intervalle temporel t . La fonction renvoie un sous-ensemble de triplets $b \in B$ dont chacun contient un emplacement

spatio-temporel (l, t) avec ses valeurs (données par la fonction $state(l, t)$) avec les conditions stipulant que 1) chaque instant $b.t_i$ doit être inclus dans l'intervalle t , et 2) chaque emplacement spatial $b.l_i$ a une relation spatiale avec l'emprise spatiale continue renvoyée par la fonction f_l (cf Figure 3.2-b).

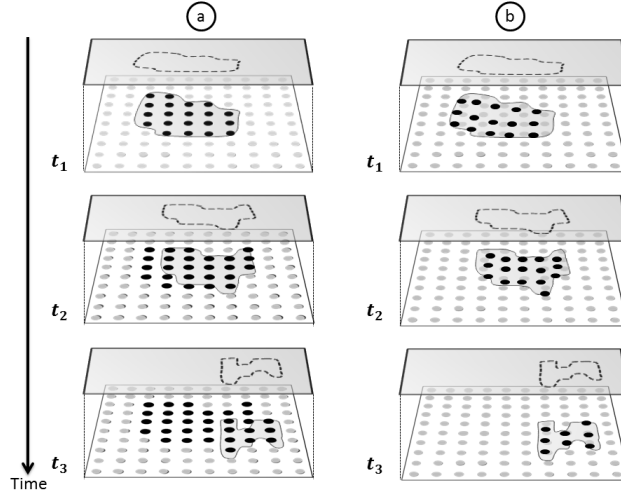


FIGURE 3.2 – Différence entre la fonction Ct_{fh} (a) et Ct_{fs} (b)

Nous définissons finalement la fonction Ot_{fs} , semblable à Ct_{fs} , mais pour des occurrents O_c . Cette fonction matérialise l'hétérogénéité spatio-temporelle d'un occurrent en représentant son *field* interne pendant un temps donné. Formellement, nous définissons cette fonction comme suit :

$$Ot_{fs}: O_c \times T \rightarrow \mathcal{P}(B) \quad \text{avec}$$

$$Ot_{fs}(c, t) = \{B_i \in \mathcal{P}(B) \mid \forall b \in B_i \text{ alors } b \cdot t_i \subseteq E_t(o_c) \wedge b \cdot l_i \cap f_s(o_c) \neq \emptyset\} \quad (3.5)$$

Cette fonction renvoie un sous-ensemble d'emplacements spatio-temporels associés à leurs valeurs (retournées par la fonction $State(l_i, t_i)$) pour chaque instant t_i dans l'intervalle t . l_i a une relation spatiale et temporelle avec l'emprise spatiale et temporelle d'un occurrent o_c renvoyée respectivement par les fonctions E_t et f_s .

Pour illustrer le rôle des fonctions Ct_{fh} et Ct_{fs} pour la manipulation des requêtes spatio-temporelles hybrides, prenons trois exemples :

Req_4) *Sur un champ de dunes et entre deux instants t_1 et t_2 , où sont localisées les augmentations d'altitude du fond marin et*

leur amplitude, liées au passage de la dune qui a eu le déplacement le plus important sur cette période ?

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned} Q &:= \{(c, dp) \mid dp = distance(ctr(c, t_1), ctr(c, t_2))\} \\ R &:= \{Ct_{fh}(Q.c, [t_1, t_2]) \mid Q.dp = Max(Q.dp)\} \\ W &:= \{(R.l, v) \mid v = (hist(R.l)(t_2)[elev] - hist(R.l)(t_1)[elev]) \wedge v < 0\} \end{aligned}$$

Cette requête illustre comment la dynamique d'une dune c (dans ce cas il s'agit de déplacements) peut influencer les propriétés du fond marin. Tout d'abord, l'ensemble des dunes avec leurs taux de déplacement respectifs est sélectionné et le résultat est attribué à la variable Q . Cela se fait en calculant la distance entre les centroïdes (i.e., ctr) de chaque dune c à deux instants différents t_1 et t_2 .

Deuxièmement, l'entrée ou l'argument passé à la fonction Ct_{fh} est la dune c ayant le taux de déplacement le plus important en appliquant la fonction Max sur l'attribut dp . Par conséquent, R reçoit la sortie de $Ct_{fh}(c, [t_1, t_2])$ qui représente un ensemble de paires ordonné $\langle l, hist(l) \rangle$. Enfin, la variable W reçoit le résultat de la variable R , avec lequel la sous-requête calcule de différence d'élévation et renvoie un ensemble de paires ordonnées (l, v) qui localisent et dimensionnent l'augmentation d'altitude du fond marin.

Req_5) *Sur un champ de dunes et entre deux instants t_1 et t_3 , quelle est l'amplitude des variations de pente sur le fond marin impacté par le rapprochement entre une dune c_A et la dune qui s'est le plus rapproché de cette dernière*

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned} Q &:= \{(c, dr) \mid c \in f_r(c_A).c \wedge dr = f_{vg}(c_A)(c, t_3) - f_{vg}(c_A)(c, t_1)\} \\ R &:= \{Ct_{fh}(c', [t_1, t_3]) \mid c' = c_A \vee c' \in Q.c \wedge Q.dr = Min(Q.dr)\} \\ W &:= \{(R.l, v) \mid v = |hist(R.l)(t_3)[pente] - hist(R.l)(t_1)[pente]|\} \end{aligned}$$

La première étape consiste à calculer la grandeur dr qui évalue la valeur de variation de la distance entre la dune c_A et les dunes voisines entre deux instants t_1 et t_3 à l'aide de la fonction f_{vg} . L'étape suivante consiste à sélectionner à la fois (avec l'opérateur logique $\vee$) la dune c_a et la dune qui

s'est le plus rapproché de la dune c_A en utilisant la fonction $Min(Q.dr)$. Le résultat est affecté à la fonction Ct_{fh} . Comme pour la requête 4, le résultat est un ensemble de paires ordonnées (l, v) qui déterminent les variations spatio-temporelles de la pente du fond marin.

Req_6) *Sur un champ de dunes et entre deux instants t_1 et t_2 , rechercher la localisation spatiale et temporelle de la pente la plus faible sur les flancs de la dune qui a subi le changement de superficie le plus important sur cette période ?*

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned} Q &:= \{(c, sp) \mid sp = superficie(c, t_2) - superficie(c, t_1)\} \\ R &:= \{Ct_{fs}(Q.c, [t_1, t_2]) \mid Q.sp = Max(Q.sp)\} \\ W &:= \{(R.l, R.t, v) \mid v = Min(state(R.l, R.t)[pente])\} \end{aligned}$$

Cette requête montre l'utilisation de la fonction hybride Ct_{fs} qui permet d'explorer l'hétérogénéité spatio-temporelle du *field* correspondant à la surface interne d'une entité. Tout d'abord, l'ensemble des dunes avec leur changement de superficie respectif sont sélectionnées et le résultat est attribué à la variable Q . Cela se fait en calculant la différence de superficie pour chaque dune dans deux instants différents t_1 et t_2 .

Deuxièmement, l'argument passé à la fonction Ct_{fs} correspond à la dune c ayant le changement de superficie le plus élevé en appliquant la fonction Max sur l'attribut sp . Par conséquent, R reçoit le résultat de $Ct_{fs}(c, [t_1, t_2])$ qui représente un ensemble de tuples $\langle l, t, v \rangle$. Enfin, la variable W reçoit le résultat de la variable R , avec lequel la sous-requête sélectionne l'emplacement spatio-temporel en se basant sur la valeur minimale de la pente.

Req_7) *Déterminer l'évolution dans le temps et l'espace de la valeur de pression barométrique minimale à l'intérieur d'un orage qui a lieu sur une ville donnée ?*

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned} Q &:= \{(o_c, t_i) \mid type(o_c) = "orage" \wedge geom(o_c) \sqcap geom(ville)\} \\ R &:= \{Ot_{fs}(Q.o_c, t) \mid t = [Min(Q.t_i), Max(Q.t_i)]\} \\ W &:= \{(R.l, R.t', v) \mid v = Min(R.v[pression]), \forall t' \in t\} \end{aligned}$$

Cette requête illustre l'utilisation de la fonction hybride Ot_{fs} . Dans un premier temps, l'événement orageux (l'entité o_c) est sélectionné avec l'ensemble des états qui correspondent à la période où l'emprise spatiale

de l'événement (i.e., $geom(o_c)$) croise l'emprise spatiale de la ville (i.e., $geom(ville)$). Le résultat est attribué à la variable Q .

Deuxièmement, la sous-requête construit l'intervalle temporel t correspondant à la durée totale de passage de l'événement sur la ville. L'événement o_c et l'intervalle t sont passés comme arguments à la fonction hybride Ot_{fs} dont le résultat est affecté à la variable R qui représente un ensemble de tuples $\langle l, t, v \rangle$. Enfin, la variable W reçoit le résultat de la variable R , avec lequel la sous-requête extrait dans le temps et l'espace la valeur de pression barométrique.

Les fonctions décrites dans cette perspective (i.e., objet vers *field*) permettent d'explorer l'impact possible de l'évolution d'un ou de plusieurs objets sur la dynamique d'un champ donné. Cependant, elles ne permettent pas de suivre l'amplitude de la variation du champ interne à l'objet évoluant dans le temps. C'est pourquoi une nouvelle catégorie de fonctions hybrides et de mécanismes supplémentaires doit être développée à cette fin.

3.5.4 Vision Hybride : Évolution du *field* interne d'un objet

De nombreux phénomènes géographiques sont représentés par un ensemble d'entités géométriques polygonales, ce qui a amené la communauté scientifique à proposer de nombreux travaux permettant de décrire et explorer la dynamique spatio-temporelle au travers de la variation de la dimension, de l'orientation, de la forme ainsi que le déplacement de ces étendues spatiales ([Sadahiro and Umemura, 2001](#); [Stefanidis et al., 2003](#); [Robertson et al., 2007](#)). Cependant ces travaux ne fournissent aucun mécanisme pour explorer la variation de la structure interne (i.e., champ) de ces objets, et ne permettent pas d'identifier la localisation de ces changements. Il est cependant évident que ces informations peuvent être utiles pour comprendre le processus géographique étudié.

Cette approche est intéressante, car elle met en évidence l'interrelation qui existe entre la dynamique de l'objet et le changement des conditions physiques qui s'y rapportent. Selon [Goodchild and Yuan \(2007\)](#) " *Lorsque le champ-objet se déplace et change de forme, il entraîne avec lui les champs qui lui sont associés, ce qui permet de mieux comprendre sa dynamique dans un système de coordonnées mobiles qui opère au sein de l'emprise spatiale d'un objet*".

Nous illustrons cela par deux requêtes spatio-temporelles hybrides :

- Comment évolue la pente à l'intérieur de l'emprise spatiale de la dune dont le taux de déplacement est le plus important sur un champ de dunes ?
- Comment évolue la valeur de la pente à l'intérieur de l'emprise spatiale de la dune qui se rapproche le plus d'une dune référente ?

Pour pouvoir répondre à de telles requêtes, le modèle hybride doit gérer non seulement l'hétérogénéité au sein de l'objet, mais aussi doit permettre d'explorer et caractériser la variation spatio-temporelle du *field* à l'intérieur d'un objet, quand ce dernier se déplace ou se déforme.

Le but de cette partie est de proposer un formalisme sous la forme d'une fonction hybride permettant de caractériser la variation du champ intérieur d'un objet mobile.

3.5.4.1 Définition d'un cadre de référence interne

Nous désignons P comme l'ensemble des emprises spatiales de type polygone associées à différents objets dans différents temps. Nous définissons ensuite un cadre de référence interne lié à un polygone p comme un ensemble d'éléments géométriques à partir desquels des positions et des distances sont mesurées. Formellement, nous le définissons comme une paire ordonnée (P, f_{pe}) tel que :

$$f_{pe} : p \longmapsto (e_1, e_2)$$

La fonction f_{pe} associe pour chaque polygone $p \in P$ deux éléments de références (e_1, e_2) à partir desquels la position d'un point s à l'intérieur de p pourraient être identifiée dans le plan. Ces éléments appartiennent à l'emprise spatiale du polygone p (ex. centroïde, frontière du polygone, ligne d'allongement principal passant par le centroïde, etc.). Nous mettons $I_F = (P, f_{pe})$. Pour permettre d'appliquer des métriques à l'intérieur d'un polygone, nous définissons un espace métrique relatif E_m comme un tuple ordonné :

$$E_m = ((S)_{p \in P}, M_1, M_2)$$

tel que $(S)_{p \in P}$ correspond à la collection des sous-ensembles de points de l'espace dont chaque sous-ensemble S_p est spatialement couvert par le polygone p . Nous désignons $(S)_{p \in P} \subseteq \mathcal{P}(S)$, cela signifie que $(S)_{p \in P}$ correspond à une collection de sous-ensembles de l'espace S tel que $(S)_{p \in P} = \{S_{p1}, S_{p2}, \dots, S_{pn}\}$. M_1 et M_2 sont deux métriques qui définissent le concept de

distance entre deux points appartenant à un sous-ensemble S_p . Pour exprimer la relation entre notre espace métrique E_m et le cadre de référence I_F , la position de chaque point $s \in S_p$ au sein du polygone p peut être exprimée par rapport au cadre défini I_F de telle sorte que : $M_1(e_1, s) \in \mathbb{R}$ et $M_2(e_2, s) \in \mathbb{R}$. Ces deux derniers éléments correspondent respectivement à la distance entre la référence e_1 et le point s calculée par la fonction M_1 , et à la distance entre la référence e_2 et le point s calculée par la fonction M_2 .

3.5.4.2 Fonction hybride liée à l'évolution de la structure interne d'un continuant

Dans cette sous-section nous nous proposons de définir la fonction Ct_{fr} qui permet de décrire l'évolution de structure interne d'un continuant. La structure interne, est considérée ici comme un espace continu de type "champ", et seulement une partition de domaine $L_i \subseteq L$ de ce champ est utilisée. Cette partition est identifiée par des paires de coordonnées relatives liées à l'emprise spatiale du continuant c à un instant donné t . Le suivi de l'évolution d'un continuant est ainsi réalisé via le suivi de l'évolution de chacun des éléments de la partition repérée pour chaque état de ce continuant. Ce principe est illustré par la figure 3.3.

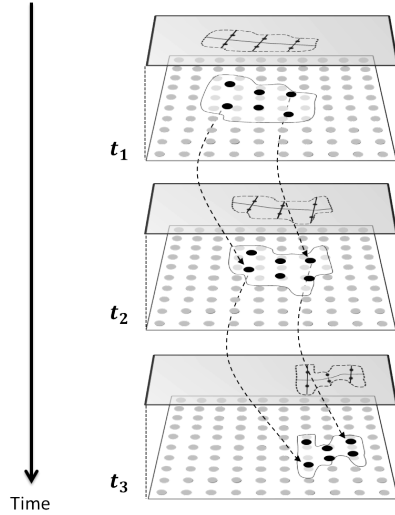


FIGURE 3.3 – Figure illustrant le suivi de la variation interne d'un objet en l'associant avec une partition de l'espace champ qui décrit la surface interne de l'objet à chaque temps.

Avant de passer à la définition de la fonction hybride Ct_{fr} , nous devons être capables de générer un sous-ensemble de points dans chacun des états

d'un continuant c afin de pouvoir suivre l'évolution dans le temps de la surface interne de cet objet. Ces points doivent être identifiables dans chacun des polygones p représentant la succession des états de c .

Pour cela, nous définissons la fonction f_{sp} qui génère un sous-ensemble de points S_p avec leurs coordonnées relatives à l'intérieur d'un polygone p :

$$f_{sp}: P \rightarrow (S_u \rightarrow \mathbb{N}^2)$$

$$f_{sp}(p)(s) = \{(a, b) \mid \forall s \in S_p \text{ et } S_p \subseteq S_u\}$$

L'ensemble S_u constitue l'union des éléments de tous les sous-ensembles de $(S)_{p \in P}$ tel quel $S_u = S_{p_1} \cup S_{p_2} \cup \dots \cup S_{p_n}$, où $p_1, p_2, \dots, p_n$ correspondent aux polygones appartenant à l'ensemble P .

Ensuite, nous considérons la fonction f_p qui associe à un objet c , à chaque instant valide t , une composante spatiale de type polygone. Cette fonction est définie formellement comme suit :

$$f_p: C \times T \rightarrow P$$

dans laquelle P correspond à l'ensemble des polygones présents dans l'espace. Pour deux polygones P_i et P_j , représentant deux état d'un même objet c , avec la fonction f_{sp} , nous pouvons générer deux sous-ensembles de points de taille égale, $f_{sp}(p_i)$ et $f_{sp}(p_j)$, tel que $p_i \in f_p(c, t_i)$ et $p_j \in f_p(c, t_j)$.

Nous disons alors que deux points (s_i, s_j) sont homologues de telle sorte que $s_i \cap p_i$ and $s_j \cap p_j$ si et seulement si $f_{sp}(p_i)(s_i) = f_{sp}(p_j)(s_j)$. Cela signifie que les deux points s_i et s_j présents à deux instants différents pour le même objet mobile c doivent avoir les mêmes paires de coordonnées relatives (a, b) . Pour associer les points générés dans un polygone p avec le domaine d'un champ L , nous définissons la fonction $\varphi: S_u \rightarrow L$, elle associe chaque point spatial s à un emplacement spatial $l \in L$.

Enfin, nous définissons notre fonction hybride Ct_{fr} comme suit :

$$Ct_{fr}: C \times T \rightarrow (L \times \mathbb{N}^2 \times T \rightarrow V_f) \quad \text{avec}$$

$$Ct_{fr}(c, t) = \{(l_i, k, t_i, v_f) \mid \forall t_i \subseteq t, \exists! s_i \in f_{sp}(p_i).s \text{ et } p_i \in f_p(c, t_i) \text{ tel que}$$

$$l_i = \varphi(s_i) \wedge k = f_{sp}(p_i)(s_i) \wedge v_f = state(l_i, t_i)\}$$
(3.6)

La fonction Ct_{fr} renvoie un sous-ensemble des emplacements spatio-temporels avec à la fois leurs paires de coordonnées relatives au sein de

l'emprise spatiale du continuant c , et la valeur contenue dans le *field* (données par la fonction $state(l, t)$). Le résultat de cette fonction est contraint par la condition qui stipule que, pour tout instant $t_i \subseteq t$, il existe pour chaque localisation spatiale l_i au sein de l'emprise spatiale $f_p(c, t_i)$, un point associé unique s_i (i.e., $l_i = \varphi(s_i)$) tel que les coordonnées relatives représentées par la valeur $k \in \mathbb{N}^2$ concernant la localisation spatiale l_i est l'image directe du point s_i par la fonction $f_{sp}(p_i)(s_i)$.

Pour illustrer l'emploi de cette fonction hybride, nous nous inspirons des requêtes 4 et 5, vues précédemment :

Req_8) *Comment évolue la pente sur les flancs de la dune qui a eu le déplacement le plus important dans un champ de dunes entre deux temps t_1 et t_2 ?*

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned} Q &:= \{(c, dp) \mid dp = distance(ctr(c, t_1), ctr(c, t_2))\} \\ U &:= \{Ct_{fr}(Q.c, [t_1, t_2]) \mid Q.dp = Max(Q.dp)\} \\ W &:= \{(U.l_i, U.l_j, dv) \mid U.l_i \cap f_p(c, t_1) \wedge U.l_j \cap f_p(c, t_2) \wedge U.k_i = U.k_j \\ &\quad \text{where } dv = |U.v_j[pente] - U.v_i[pente]|\} \end{aligned}$$

En premier lieu, l'ensemble des dunes avec leurs taux de migration respectifs sont sélectionnées et le résultat est attribué à la variable Q . Cela se fait en calculant la distance entre les centroïdes (i.e., ctr) de chaque dune c à deux instants différents t_1 et t_2 . Deuxièmement, l'entrée ou l'argument passé à la fonction Ct_{fr} est la dune c ayant le taux de déplacement le plus important en appliquant la fonction Max sur l'attribut dp . Par conséquent, U reçoit le résultat de la fonction $Ct_{fr}(c, [t_1, t_2])$ qui représente un ensemble de tuples $\langle l, k, t, v \rangle$. Ensuite la variable W désigne une relation ternaire $W \subseteq U \times U \times \mathbb{R}$ afin de regrouper les paires des localisations spatiales (l_i, l_j) avec leurs valeurs réelles, telles que leurs coordonnées k_i et k_j soient égales. Ceci indique que les localisations spatiales l_i et l_j appartiennent respectivement aux emprises spatiales $f_p(c, t_1)$ et $f_p(c, t_2)$ correspondant à deux états spatio-temporels différents de la dune. Enfin, la sous-requête calcule la valeur absolue de la différence de pente $dv \in \mathbb{R}$ et renvoie le tuple (l_i, l_j, dv) qui définit la localisation et la variation de pente entre t_1 et t_2 .

Req_9) *Comment évalue la pente sur les flancs d'une dune A et de la dune qui c'est rapprochée le plus de cette dune A , entre t_1 et t_2 ?*

Cette requête est différente de la requête 5 même si à première vue elles peuvent se ressembler. En effet, la requête 5 explore comment le rapprochement de deux dunes sélectionnées pourrait impacter la variation spatio-temporelle de la pente sur le fond marin, alors que dans cette requête, nous cherchons à déterminer comment évolue la pente à l'intérieur de chaque dune pendant le rapprochement.

Cette requête est décomposée en trois étapes comme suit :

$$\begin{aligned}
 Q &:= \{(c, dr) \mid c \in f_r(c_A).c \wedge dr = f_{vg}(c_A)(c, t_2) - f_{vg}(c_A)(c, t_1)\} \\
 U &:= \{Ct_{fr}(c', [t_1, t_2]) \mid c' = c_A \vee c' \in Q.c \wedge Q.dr = Min(Q.dr)\} \\
 W &:= \{(U.l_i, U.l_j, dv) \mid U.l_i \cap f_p(c', t_1) \wedge U.l_j \cap f_p(c', t_2) \wedge U.k_i = U.k_j \\
 &\quad \text{where } dv = |U.v_j[pente] - U.v_i[pente]| \}
 \end{aligned}$$

La première étape consiste à calculer la grandeur dr qui détermine la distance de rapprochement entre la dune A et les dunes voisines à l'aide de la fonction f_{vg} . L'étape suivante consiste à sélectionner à la fois (avec l'opérateur logique $\vee$) la dune c_A et la dune qui s'est le plus rapprochée de la dune A en utilisant la fonction $Min(Q.dr)$. Le résultat est passé en argument à la fonction Ct_{fr} . R reçoit le résultat de la fonction $Ct_{fr}(c', [t_1, t_2])$ qui correspond à un ensemble de tuples $\langle l, k, t, v \rangle$ issu des deux dunes sélectionnées et stockées dans la variable c' . Ensuite la variable W désigne une relation ternaire $W \subseteq U \times U \times \mathbb{R}$ afin de regrouper les paires de localisations spatiales (l_i, l_j) pour chacune des deux dunes avec leurs valeurs réelles tel que leurs coordonnées k_i et k_j soient égales. Enfin, la sous-requête calcule la valeur absolue de la différence de pente $dv \in \mathbb{R}$ et renvoie le tuple (l_i, l_j, dv) qui définit la localisation et la variation de pente entre t_1 et t_2 , sur les deux dunes sélectionnées

3.6 Schéma conceptuel pour des requêtes spatio-temporelles hybrides.

Les concepts hybrides développés ci-dessus, amènent à des fonctions dont chacune a un rôle spécifique et des mécanismes importants. Ces mécanismes permettent le passage d'une représentation basée sur un espace champ à une représentation basée sur un espace objet d'une manière plus fluide, et permettent d'explorer le réseau d'interconnexion entre les deux espaces. Ce réseau a pour objectif d'explorer l'hétérogénéité interne des objets, et l'interaction entre les deux espaces. Cette interaction amène deux perspectives :

1) explorer l'impact de la dynamique des champs sur la dynamique des objets ou 2) explorer l'impact de la dynamique des objets sur la dynamique du champ.

Nous avons illustré l'utilisation de ces fonctions avec neuf exemples de requêtes. Pour préciser la typologie de ces requêtes, nous nous appuyons sur les travaux de (Sinton, 1978). Selon lui, il existe une structure inhérente à l'information géographique. Il propose un cadre conceptuel générique permettant de catégoriser et de caractériser les requêtes spatio-temporelles. Nous proposons de faire le lien avec les fonctions hybrides que nous venons de définir, pour aider à sélectionner la plus adaptée. Pour (Sinton, 1978), une observation est extraite à partir de trois attributs : l'espace, le temps et le thème (le terme thème fait référence au phénomène ou à l'élément du monde réel observé). Il soutient que nous pouvons créer des généralisations d'informations géographiques en fonction de la façon dont ces attributs (espace, temps et thème) sont impliqués. D'une manière générale, les observations sont obtenues en : (1) fixant la valeur d'un attribut ; (2) faisant varier le deuxième attribut de manière contrôlée ; et (3) mesurant le troisième attribut, compte tenu des contraintes du deuxième attribut. Le lien entre la typologie proposée par (Sinton, 1978) et les fonctions hybrides que nous avons proposées est décrit dans le tableau 3.1.

Fixer	Contrôler	Mesurer	Fonctions
dynamique du <i>field</i>	location / temps	dynamique de continuant	$Ft_c()$
dynamique du <i>field</i>	location / temps	attributs d'occurrent	$Ft_o()$
évolution des attributs d'objet	continuant / temps	dynamique du <i>field</i>	$Ct_{fh}()$
évolution des attributs d'objet	continuant / temps	<i>field</i> interne de continuant	$Ct_{fs}()$
valeurs d'attributs d'occurrent	occurrent / temps	<i>field</i> interne d'occurrent	$Ot_{fs}()$
évolution des attributs d'objet	continuant / temps	variation du <i>field</i> interne	$Ct_{fr}()$

TABLE 3.1 – Schéma conceptuel montrant comment chaque fonction hybride pourrait être utilisée pour exprimer des requêtes hybrides.

Les requête Req_1 et Req_2 explorent le lien entre la dynamique du *field* (ex. la variation de profondeur du fond marin) et les changements observés sur les propriétés des objets (ex. la superficie ou la position de la dune) ou sur leurs relations (ex. le rapprochement des dunes par exemple). La requête Req_3 porte également sur la même interrogation, mais concerne une entité occurrente (i.e., un orage). Ces requêtes appliquent les fonctions Ft_c et Ft_o , qui permettent de fixer la variation du *field* (par exemple, une différence de profondeur supérieure à 2m), de contrôler l'espace et le temps et de mesurer

soit le changement de propriété d'un objet continuant, soit d'identifier un objet occurrent.

Les requêtes Req_4 et Req_5 fixent un changement de propriétés (i.e., attribut) (ex. déplacement) ou de relations spatiales entre objets (ex. rapprochement des dunes) et contrôlent les objets et le temps. Le résultat de la fonction Ct_{fh} permet de mesurer une variation sur le *field* (ex. variation d'altitudes ou de pentes sur le fond marin). Ces deux requêtes permettent d'explorer comment la dynamique d'objets pourrait impacter la dynamique du *field*.

Les requête Req_6 traitent de la structure interne d'une entité de type continuant (ex. mesure de la pente), dont on fixe la variation spatio-temporelle d'une propriété (ici, la superficie) et on contrôle les objets et le temps. Cette requête repose sur l'utilisation de la fonction Ct_{fs} . La requête Req_7 est très similaire à la précédente, à la différence que la manipulation concerne un objet occurrent (ici, un orage). Dans ce cas, on fixe la valeur d'un attribut (ici, le terme "orage"), on contrôle l'objet et le temps, et on mesure une propriété du *field* (ici, la pression atmosphérique).

Les deux dernières requêtes Req_8 et Req_9, fixent un changement d'attribut (ex. déplacement) ou de relations spatiales entre objets (ex. rapprochement des dunes), contrôlent les objets et le temps à l'aide de la fonction Ct_{fr} et mesure la variation interne du *field* des objets sélectionnés.

3.7 Conclusion et synthèse

Dans ce chapitre, comme première étape de notre travail, nous avons proposé un formalisme plus générique de la notion de champ spatio-temporel afin de prendre en compte un ensemble de propriétés spatialement distribuées. Nous modélisons ainsi conceptuellement la vision de l'objet de manière à ce qu'elle surmonte la dichotomie entre les notions de continuant et d'occurrent. Cette dichotomie est en fait surmontée par une généralisation des deux concepts vers la notion d'entité géographique dénommée G . Enfin, nous apportons une vision hybride de la notion de temps.

Notre vision hybride de l'espace-temps qui s'articule autour des points suivants :

- Dans un premier temps, une association est définie entre les deux représentations spatio-temporelles (i.e., champ et objet). Pour un continuant, cette association est explicite et est définie comme une fonction qui relie chacune des états spatio-temporels d'un continuant

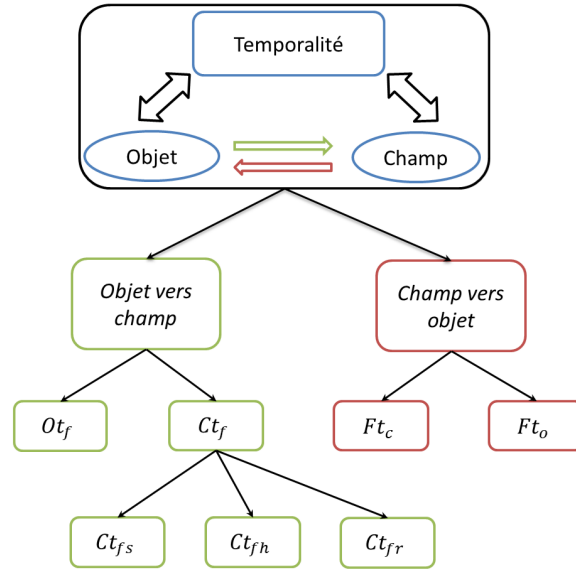


FIGURE 3.4 – Dérivation des fonctions hybrides issue de l’association bidirectionnelles entre des objets et des champs à travers le temps

(i.e., $C \times T$) avec le domaine spatial d’un champ (i.e., L). Quant à un *occurent*, il est en relation avec le domaine et l’image d’un champ spatio-temporel. Cette association sera directement gérée par une fonction hybride.

- Ensuite, à partir de cette association, nous avons formalisé une série de fonctions hybrides qui représentent des associations bidirectionnelles entre des objets et des champs. Il s’agit essentiellement d’une dérivation d’un ensemble de concepts résultant de l’interconnexion entre des visions discrètes et continues de l’espace à travers le temps. La figure 3.4 résume et classifie les fonctions hybrides selon le sens d’association entre les deux visions de l’espace. D’une part, quand le sens de l’association part de l’objet vers le champ, nous distinguons quatre fonctions hybrides : Ot_f , Ct_{fs} , Ct_{fh} et Ct_{fr} . D’autre part et quand le sens d’association entre les deux visions de l’espace part de l’objet vers le champ nous distinguons les deux fonctions hybrides suivantes : Ft_c et Ft_o .
- En dernier lieu, nous avons abordé le problème lié à la fonction Ct_{fr} . Celle-ci représente le concept par lequel on peut explorer et

caractériser la variation spatio-temporelle d'un champ à l'intérieur d'un objet lorsque ce dernier se déplace ou se déforme. En effet, cette notion nécessite la mise en place d'un formalisme pour un cadre de référence mobile. Ce cadre est constitué de trois composantes principales : 1) un cadre de référence interne (*CRI*) et 2) une fonction qui génère un ensemble de points à l'intérieur de chaque polygone représentant un état spatio-temporel d'un objet en mouvement (i.e., continuant). Cela permet ensuite d'identifier des points homologues qui consistent à la paire de points (s_i, s_j) appartenant aux polygones p_i et p_j dans deux instants t_i et t_j qui possèdent les mêmes coordonnées relatives en rapport avec leur *CRI* pré-établi.

Enfin, nous avons utilisé le cadre conceptuel générique proposé par (Sinton, 1978) qui catégorise les requêtes en fonction de l'emplacement, du temps et de la sémantique. L'objectif de l'introduction de ce cadre est de proposer une typologie des différentes requêtes hybrides spatio-temporelles. De plus, il indique les correspondances entre ces types de requêtes et les fonctions hybrides afin d'aider le lecteur à sélectionner la bonne fonction pour le bon scénario.

Chapitre 4

Vers un modèle spatio-temporel hybride

"The computer offers another kind of creativity. You cannot ignore the creativity that computer technology can bring. But you need to be able to move between those two different worlds."

Tadao Ando

Nous avons montré dans le chapitre précédent, par une approche formelle, appuyée par des exemples de requêtes spatio-temporelles, comment la combinaison de deux représentations orientées "champ" et "objet" permet une sorte de ramification de plusieurs concepts hybrides. Ces concepts, lorsqu'ils sont étendus par la composante temporelle, peuvent aider à améliorer notre compréhension de l'évolution de certains phénomènes géographiques complexes. Cependant, et malgré l'intérêt de cette approche formelle, il est encore nécessaire d'intégrer plus étroitement cette approche hybride de l'espace au niveau d'un modèle de données spatio-temporelles, et de proposer la mise en oeuvre des capacités de manipulation de données spatio-temporelles. Dans le chapitre précédent, les requêtes hybrides étaient exprimées de façon formelle. Dans ce chapitre, nous proposerons un ensemble de spécifications logiques avec de nouveaux types de données et d'opérations qui permettent d'étendre le langage de requête, afin qu'il puisse prendre en charge les requêtes spatio-temporelles hybrides.

4.1 Modèle de données hybride : vue globale

Notre objectif consiste à introduire un modèle de données hybrides dans les SIG dynamiques qui représente et intègre étroitement les deux représentations de l'espace, l'objet et le champ. Les exigences de base pour un tel modèle de données hybrides peuvent être rappelées comme suit :

- 1) Intégrer les dimensions temporelles dans une représentation des objets et des champs.
- 2) Dépassez la dichotomie entre la vision d'un continuant et celle d'un occurrent dans le modèle de données aux niveaux conceptuel et logique.
- 3) Proposer des mécanismes pour pouvoir associer explicitement ces deux conceptualisations de l'espace.
- 4) Développer et mettre en œuvre des opérateurs et des fonctions hybrides spatio-temporelles.

Cette section présente les schémas conceptuels et logiques du modèle de données hybrides qui est une forme de représentation numérique du modèle formel présenté dans le chapitre précédent. Nous retenons le processus général de modélisation des données illustré par la figure 4.1. L'objectif de ce processus est de montrer comment les concepts hybrides seront intégrés dans un modèle de données hybride.

Dans ce processus, nous distinguons deux niveaux de modélisation, dont le premier concerne les considérations conceptuelles (en haut à droite sur la figure 4.1) tandis que le deuxième niveau concerne plutôt les considérations liées à la représentation numérique (en bas à droite sur la figure 4.1).

Le modèle conceptuel formel est le premier niveau de ce processus et le plus haut niveau d'abstraction vu qu'il se concentre uniquement sur la sélection des concepts clés liés au domaine géographique, tout en cachant les détails de la structure et du traitement des données. Le modèle conceptuel hybride fournit plus de spécifications pour les données spatio-temporelles, mais en conservant toujours un lien étroit avec les concepts du domaine géographique.

Le deuxième niveau de ce processus est constitué en premier lieu par la couche logique, qui fournit une couche de liaison reliant le modèle concep-

tuel hybride au modèle d'implémentation. Cette couche permet une description abstraite de la structure et de l'ensemble des opérations de traitement des données. Elle ne contiendra que les spécifications formelles relatives aux différents types de données abstraites (abstract data type - ADT) avec leurs opérations associées.

En deuxième lieu, ce même niveau est composé d'un modèle d'implémentation ainsi que d'un modèle d'instanciation. Le modèle d'implémentation permet la mise en œuvre et le développement de différents types de structures de données avec leurs fonctions provenant de la couche logique. Nous encapsulons ce modèle par un schéma de base de données générique établi selon une architecture logicielle bien définie. Cette architecture permet d'étendre le langage de requête avec des opérations de manipulation et d'analyse de données spatio-temporelles permettant de gérer des requêtes hybrides. Le modèle d'instanciation consiste en un schéma de base de données alimenté par une collection de données spatio-temporelles qui permettront de fournir des résultats en réponse aux requêtes hybrides formulées à l'aide du langage d'interrogation de la base de données.

Notons que nous utiliserons les termes "objet" et "entité" de manière interchangeable dans la suite de cette section.

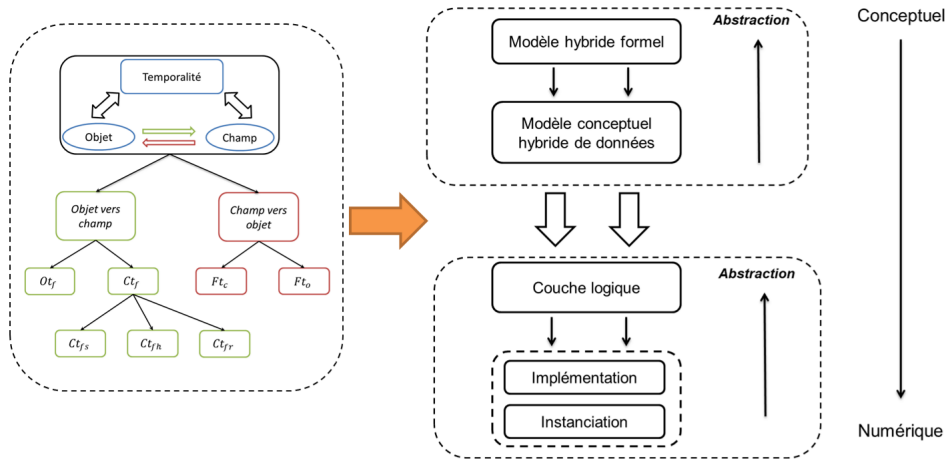


FIGURE 4.1 – Intégration de la représentation hybride selon un processus général de modélisation des données

4.2 Modèle Conceptuel

Nous proposons un modèle conceptuel de données vu comme un médiateur entre le modèle formel développé dans le chapitre précédent et le

modèle de base de données. Au niveau conceptuel, quatre notions essentielles sont utilisées, à savoir les classes, les attributs, les relations et les contraintes. Ces concepts nous permettent de saisir les principales caractéristiques liées au domaine géographique. La conception est présentée ci-après à l'aide de diagrammes de classes UML qui fournissent des symboles visuels et des notations pour représenter les types d'objets. Le modèle spatio-temporel proposé est conçu avec deux blocs principaux. Le premier représente la vue basée sur le champ tandis que le second représente la vue basée sur l'objet. La vue hybride consiste en une forme d'association de ces deux visions.

Dans cette approche de modélisation conceptuelle des données, nous nous appuyons sur quatre mécanismes clés caractéristiques du modèle conceptuel qui nous permettront d'intégrer notre approche hybride à un modèle de données spatio-temporel. Ce dernier sera ensuite qualifié de modèle de données hybride.

- *Mécanisme de généralisation* : il permet de supprimer les différences entre les types d'entités, en identifiant leurs caractéristiques communes et en les généralisant en une unique superclasse dont les types d'entités originaux sont des sous-classes spécifiques (Elmasri, 2008). Il est représenté en UML par une flèche vide pointant vers la superclasse (figure 4.3). Avec ce mécanisme nous dépassons la dichotomie continuant/occurent en matérialisant l'association entre une entité géographique (i.e., *GeoEntity*), un continuant et un occurrent.
- *Mécanisme d'agrégation* : il représente la relation entre un objet composé et ses parties. Il est schématisé par un diamant vide dans le diagramme d'UML (figure 4.2). Ce mécanisme, selon (Arlow and Neustadt, 2002), possède différentes sémantiques : il permet à l'objet et à ses parties d'exister, ou non, indépendamment l'un de l'autre, c'est-à-dire que la suppression de l'un n'entraîne pas forcément la suppression de l'autre. L'objet est en quelque sorte incomplet si certaines parties sont manquantes. Ce mécanisme joue un rôle clé dans notre approche de modélisation spatio-temporelle hybride puisqu'il nous permet de matérialiser formellement l'association entre la vision champ et la vision objet.
- *Mécanisme de composition* : La composition représente une relation entre le tout et ses parties telle que la classe A est composée des objets de la classe B. Elle est utilisée lorsque les objets de la classe A ne peuvent pas logiquement exister sans avoir au minimum un

objet de la classe B. En UML, cette relation est décrite par un diamant plein. Selon (Arlow and Neustadt, 2002), ce mécanisme adopte la sémantique suivante : i) les parties appartiennent à un seul composite à la fois ; ii) le composite peut libérer des parties, à condition que ces dernières soient reprises par un autre phénomène composite ; iii) si le composite est détruit, il doit soit détruire toutes ses parties, soit les transmettre à un autre composite. Avec ce mécanisme nous établissons la relation entre un continuant et ses états spatio-temporels représentés par la classe *CState* (figure 4.3) afin de permettre de gérer la variation de ses attributs dans le temps.

- *Mécanisme de spécialisation* : il correspond à un processus d'abstraction inverse de celui de la généralisation dans lequel nous faisons apparaître les différences entre plusieurs types d'entités, ce qui amène à définir un ensemble de sous-classes d'un type d'entité. Cet ensemble de sous-classes, qui forme une spécialisation, est défini sur la base de certaines caractéristiques distinctives des entités de la super classe (Elmasri, 2008). Avec ce mécanisme, nous pouvons spécialiser n'importe quel type d'entité présent dans le modèle hybride. Par exemple, la classe *FieldState* (figure 4.2), qui représente le concept d'un champ géographique, pourrait être spécialisée en deux sous-classes l'une dénote les conditions atmosphériques (ex. température, vitesse du vent) et l'autre représente les conditions hydrodynamiques (ex. vitesse du courant, etc.).

4.2.1 Modèle spatio-temporel basé sur la vision de champ

Nous présentons trois types d'entités de base qui représentent l'essence de l'approche de modélisation du modèle champ (Figure 4.2). Il s'agit essentiellement de représenter les concepts de domaine L et du codomaine $\mathcal{P}(V)$ avec la composante temporelle (voir section 3.2).

- **FieldLocation** : cette classe représente le domaine L d'un champ défini dans la section 3.2. Elle désigne l'ensemble des localisations spatiales l . Pour chaque localisation spatiale, un tuple est constitué d'un identifiant *locID* et d'un attribut *pos_loc*, de type *point*, qui représente les coordonnées de cette localisation spatiale. Notons qu'une localisation spatiale n'appartient pas nécessairement à l'ensemble des instances de la classe *FieldState*. Pour chaque localisation,

un pixel peut être généré par la fonction *LocToPixel()*. Celle-ci prend comme arguments un point géométrique s ainsi qu'un nombre entier a et retourne un rectangle (i.e., pixel) de surface $a \times a$ dont le point s constitue son centre.

- **FieldState** : cette classe affecte, pour une localisation spatiale l représentant une instance de la classe *FieldLocation* à un temps donné $time_fs$, des valeurs correspondantes à différentes propriétés (ex. température, pression, etc.). Afin de pouvoir gérer plusieurs propriétés, nous définissons un attribut composite multi-valeurs appelé *field_value*. Cet attribut composite permet d'assigner pour chaque tuple $(locID, time_fs)$ un ensemble de paires (nom de propriété : valeur) exprimé comme $\{z_1 : val_1, z_2 : val_2, \dots, z_k : val_k\}$ où z_1, z_2 jusqu'à z_k dénotent des propriétés différentes. Cette approche permet de représenter l'historique de plusieurs propriétés pour la même localisation spatiale. Notons également que toute instance de la classe *FieldState* est associée à une et une seule instance de la classe *FieldLocation*.
- **Coverage** : cette classe représente les caractéristiques d'une zone d'étude (i.e., une étendue spatiale) : le nom, la superficie, etc. Les caractéristiques sont enregistrées à l'aide de l'attribut *metadata*. Chaque étendue comprend un sous-ensemble de localisations spatiales l représenté par une relation d'agrégation. Notons qu'une localisation spatiale peut appartenir à plusieurs zones d'études.

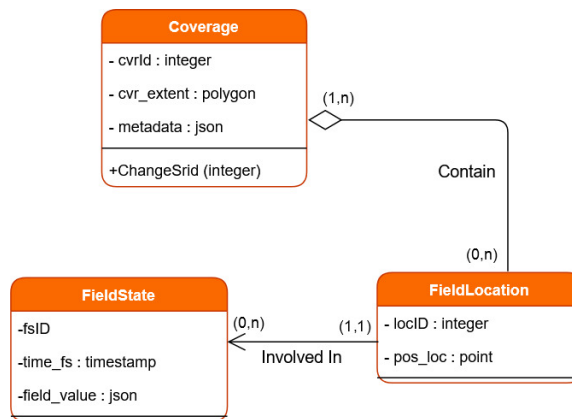


FIGURE 4.2 – Modèle conceptuel représentant la vision field.

4.2.2 Modèle spatio-temporel basé sur la vision objet

Nous présentons quatre types d'entités de base qui représentent l'essence de l'approche de la modélisation du modèle objet (Figure 4.3). Il s'agit essentiellement d'intégrer les concepts du continuant et d'occurrent (voir section 3.3) dans un modèle conceptuel de données.

- **GeoEntity** : il s'agit d'une classe abstraite qui représente un ensemble d'entités géographiques dont chacune est caractérisée par un attribut *gtype* dénotant son type (ex. dune, parcelle, métropole, etc.), une emprise spatiale (i.e., *spatialExtent*) et une emprise temporelle (i.e., *temporalExtent*). Chaque entité géographique, considérée comme instance de cette classe, doit être classée soit comme un continuant, soit comme un occurrent selon le contexte de l'étude. Par conséquent, une entité géographique ne peut pas être à la fois de type continuant et de type occurrent. De plus, chaque instance peut avoir une ou plusieurs relations avec d'autres entités. Ces relations peuvent être spatiales, spatio-temporelles.
- **Occurrent** : cette classe représente le concept d'occurrent comme défini dans la section 3.3.1. Elle constitue la première spécialisation de la classe *GeoEntity*, dont elle hérite tous ses attributs. Par conséquent, chaque instance de la classe *Occurrent* est caractérisée par son type, une emprise spatiale et une emprise temporelle. En plus des attributs hérités, la classe comporte l'attribut *ocrAttrset* qui représente l'ensemble des propriétés sémantiques de chaque instance de la classe *Occurrent*.
- **Continuant** : cette classe représente le concept de continuant comme défini dans la section 3.3.1. Elle constitue la deuxième spécialisation de la classe *GeoEntity*, dont elle hérite tous ses attributs. Par conséquent, chaque instance de la classe *Continuant* est caractérisée par son type, une emprise spatiale et une emprise temporelle. En plus des attributs hérités, la classe comporte l'attribut *ctn_identity* qui représente l'identité d'un continuant *c* qui lui permet de persister dans le temps. En outre, chaque instance de cette classe consiste en une composition de plusieurs instances de la classe *CState*.
- **CState** : cette classe représente l'ensemble des états spatio-temporels d'un continuant. Cette instance est représentée par un identifiant (i.e., *cstateID*), une géométrie (i.e., *cgeom*), un temps (i.e., *ctime*) et

un ensemble de propriétés sémantiques gérées par l'attribut *cattrSet* caractérisant chaque état spatio-temporel d'un continuant.

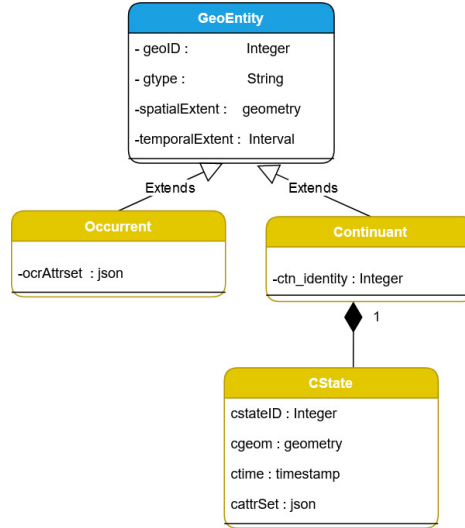


FIGURE 4.3 – Modèle conceptuel représentant la vision objet.

La principale différence entre les concepts Occurrent et Continuant, selon le modèle conceptuel de données de la figure 4.3, est que chaque instance de la classe *Occurrent* ne peut être enregistrée qu'une seule fois, alors que chaque instance de la classe *Continuant* peut être enregistrée plusieurs fois. En effet, plusieurs états d'un même Continuant peuvent être définis. Ceci est indiqué par la relation de composition entre la classe *Continuant* et la classe *CState*. Cela signifie que la classe *CState* sert de conteneur pour représenter l'ensemble des états de chaque Continuant. De plus, la relation de composition signifie que la destruction d'une instance de la classe *Continuant* implique directement la destruction de tous ses états qui sont présents en tant qu'instances de la classe *CState*.

4.2.3 Modèle conceptuel hybride

Le point clé de l'approche que nous proposons réside dans la matérialisation d'une relation entre les deux conceptualisations de l'espace champ et objet. Principalement, l'état d'un continuant (i.e., une instance de la classe *CState*) est associé avec le domaine d'un champ f représenté par les instances de la classe *FieldLocation*. Cette association est de type agrégation, illustrée sur la figure 4.4 par une liaison rouge avec un diamant vide à l'extrémité. Par ailleurs, une instance de la classe *Occurrent* peut

être associée à un champ spatio-temporel continu, correspondant à sa structure interne, par une relation spatio-temporelle avec la classe *FieldState* à l'aide de la fonction hybride Ot_{fs} (voir section 3.5.2). En d'autres termes, les classes *CState* et *Occurrent* matérialisent des instances qui peuvent être considérées comme des agrégats des instances des classes *FieldLocation* et *FieldState*. Notons que l'association *CState* et *FieldLocation* n'est pas de type composition car les champs et les objets peuvent exister indépendamment au niveau conceptuel. Par contre, si nous utilisons l'association de composition entre les deux classes, les instances de la classe *FieldLocation* seront strictement dépendantes de l'existence d'une instance de la classe *CState* qui représente un état d'une entité géographique de type continuant. Cette association est spécifiée par des contraintes spatio-temporelles qui ont été formellement définies dans la section 3.5.1 à l'aide de la fonction f_l .

L'une des particularités de l'approche de continuant est qu'elle offre la possibilité d'étudier les changements à différents temps, et de relier ces changements à l'évolution interne de ces continuants au niveau du champ. Un *occurent* peut être directement associé à sa structure interne en utilisant sa relation spatio-temporelle avec un champ spatio-temporel représenté par les instances de la classe *FieldState*.

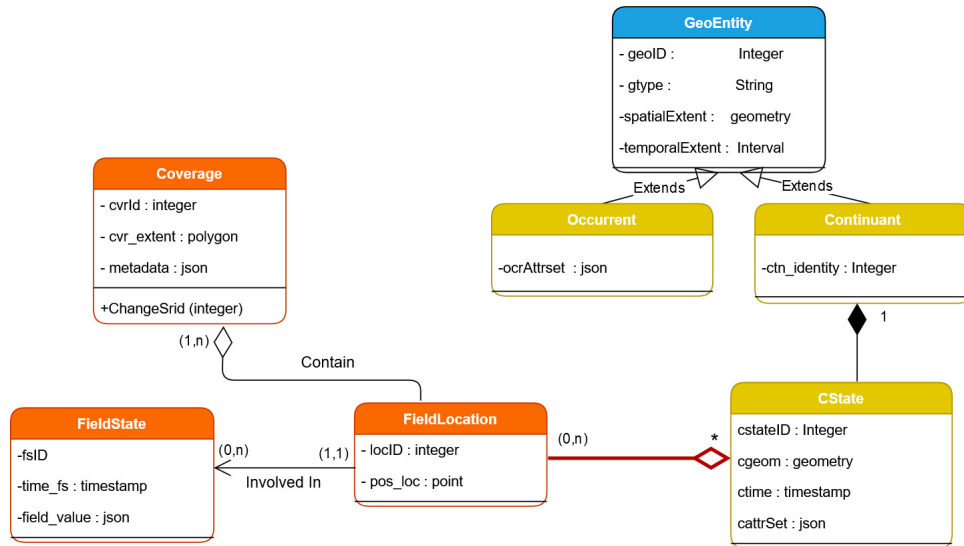


FIGURE 4.4 – Modèle conceptuel hybride représentant l'association entre les deux visions field et objet.

4.3 Modèle logique pour représenter l'approche hybride

Cette partie décrit la couche logique qui est le troisième niveau de modélisation (voir figure 4.1) permettant d'établir un lien entre le modèle conceptuel hybride et celui de l'implémentation. Dans cette couche, nous décrivons de manière abstraite les spécifications des structures des données avec leurs opérations associées à travers la notion d'ADT (*Abstract Data Type*). C'est à ce niveau que seront intégrées les fonctions hybrides, les opérations de manipulation des données, l'analyse spatiale et spatio-temporelle. Plus précisément, c'est à ce niveau que nous matérialisons les manières dont nous établissons la relation entre les fonctions hybrides et le modèle de données.

4.3.1 Types de données pour le modèle hybride

Un type de données abstrait, dénoté ADT (Abstract Data Type), permet de spécifier les abstractions de modélisation (Guttag and Horning, 1978). Il est défini par un ensemble de paramètres et une collection d'opérations. La spécification algébrique d'un ADT consiste en : (1) une description syntaxique qui définit le nom, le domaine et l'image des opérations; et (2) une spécification sémantique qui contient un ensemble d'axiomes sous forme d'équations qui relient les opérations entre elles (Guttag and Horning, 1978). Les opérations sont génériques et peuvent être appliquées pour différentes structures de données et différentes mises en œuvre. Notez que l'ADT peut également être défini comme un sous-type d'un autre ADT défini comme un super-type. Un sous-type hérite de la structure et du comportement de ses super types. Les instances ADT peuvent être stockées de manière persistante dans la base de données. Les points, les lignes et les polygones sont des exemples courants dénotant des ADT spatiaux.

Dans cette section nous introduisons trois nouveaux types de données qui seront dérivés du modèle conceptuel hybride (Figure 4.4). Ces types de données sont : *ctn*, *ocr* et *locset*.

Nous considérons une entité géographique g représentant une instance de la classe *GeoEntity* et définie comme le tuple $g = (geoID, gtype, spatialExtent, temporalExtent)$.

- **Ctn** : type de données qui représente les instances de la classe *Continuant* et est défini comme un type composé de l'identifiant d'une

instance de la classe *GeoEntity* avec une identité. Nous le définissons formellement comme suit :

$$Ctn := \{ (gid, ctn_identity) \mid gid \in geoID \wedge ctn_identity \in int \}$$

Ce type de données comporte plusieurs opérations comme montre la figure 4.5. C'est à ce niveau que les fonctions hybrides définies dans les sections précédentes seront mises en place.

- **Ocr** : type de données qui représente les instances de la classe *Occurrent* et est défini comme un type composé de l'identifiant d'une instance de la classe *GeoEntity* avec un attribut gérant l'ensemble de ses propriétés sémantiques. Nous le définissons formellement comme suit :

$$Ocr := \{ (gid, ocrAttrset) \mid gid \in geoID \wedge ocrAttrset \in json \}$$

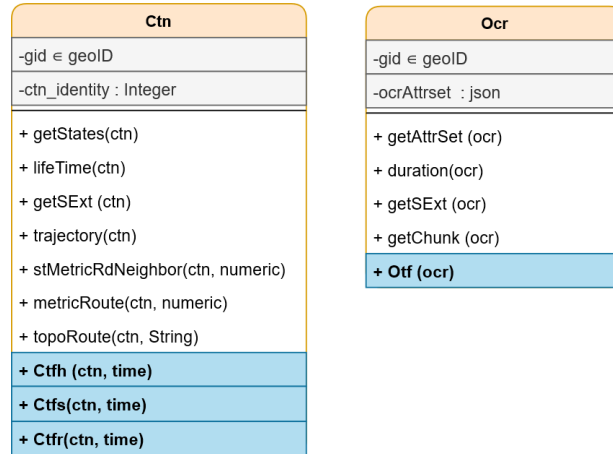


FIGURE 4.5 – L'ensemble des fonctions spatio-temporelles ordinaires et hybrides associées au deux types de données *Ctn* et *Ocr*.

Au niveau de ces deux types de données *Ctn* et *Ocr*, nous définissons un ensemble d'opérations. Plus particulièrement, c'est à ce niveau que nous définissons quatre fonctions hybrides qui sont illustrées par la couleur bleu dans la figure 4.5.

Le troisième type de données correspond à la matérialisation de la relation d'agrégation entre la classe *CState* et la classe *FieldLocation* (Figure 4.4). Nous définissons ce type de données comme suit :

- **LocSet** : il consiste en une structure de données qui gère un sous-ensemble des localisations spatiales (i.e., instances de la classe *FieldLocation*) définies pour chaque état d'un continuant c (i.e., instance de la classe *CState*). Plus précisément, ce type de données représente l'image de la fonction f_l (i.e., $img(f_l)$) qui a été définie dans la section 3.5.1. Rappelons que la fonction f_l renvoie un sous-ensemble de localisations spatiales L_i associées à un continuant c pendant un instant temporel valide t . Nous définissons formellement ce type de données comme suit :

$$LocSet := \{ L_i \mid L_i = img(f_l(c, t)) \}$$

Pour définir ce type de données au niveau logique, nous avons besoin de définir deux types de données intermédiaires, *locEntity* et *Rloc*, que nous définissons comme suit :

$$locEntity := \{ lce \in integer \mid lce \in locID \}$$

$$Rloc := \{ (lc, k_index, r_index) \mid lc \in locID \wedge k_index \in int \wedge r_index \in int \}$$

Le type *locEntity* se réfère à l'identifiant *locID* qui caractérise d'une façon unique chaque instance de la classe *FieldLocation*. Nous étendons cette définition à l'aide du type *Rloc* qui désigne une localisation spatiale relative. Ce type est caractérisé par l'identifiant *locID*, et est défini par deux paramètres *k_index* et *r_index* qui désignent le positionnement relatif de l'emplacement spatial référencé par *locID* dans l'emprise spatiale d'un continuant à un temps donné.

Nous définissons notre type de données *LocSet* comme suit :

$$LocSet := \{ [l] \mid l \in locEntity \vee l \in Rloc \}$$

Ce type de données est défini comme une agrégation d'éléments l qui sont soit de type *locEntity* ou de type *Rloc* (Figure 4.6).

A partir de *LocSet*, nous établissons un lien entre les deux visions champ et objet. Au niveau du type de données, deux fonctions hybrides sont spécifiées, Ft_c et Ft_o , qui sont illustrées dans la figure 4.6 via la couleur bleue.

4.3.2 Fonctions hybrides

Les six fonctions hybrides définies dans le chapitre précédent (i.e., Ct_{fh} , Ct_{fs} , Ct_{fr} , Ot_{fs} , Ft_c et Ft_o), qui représentent chacune un aspect de la

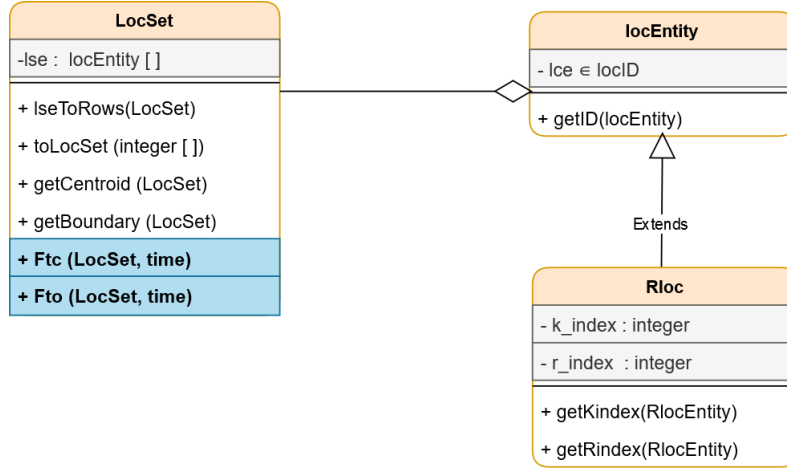


FIGURE 4.6 – Structure des type de données *LocSet*

dynamique des phénomènes géographiques, impliquent les trois types de données *Ctn*, *Ocr* et *LocSet*, comme le montrent les figures 4.5 et 4.6.

En ce qui concerne le type de données *Ctn*, trois fonctions hybrides sont spécifiées :

- Tout d'abord la fonction Ct_{fh} qui prend deux paramètres. Le premier est un objet c de type *ctn* et le second est le temps t (i.e., instant ou intervalle) et renvoie un sous-ensemble de localisations spatiales l avec leurs valeurs historiques associées (i.e., $\{v_1, v_2, \dots, v_n\}$). La signature logique de cette fonction est exprimée comme suit :

$$Ct_{fh}: (c, t) \longmapsto \text{set}(l, \text{set}(t', \{v_1, v_2, \dots, v_n\})) \quad \forall t' \in t$$

- Ensuite, la fonction Ct_{fs} prend aussi comme paramètres un objet c de type *ctn* et un instant ou un intervalle temporel t , et renvoie un sous-ensemble des localisations spatio-temporelles (l, t') avec leurs valeurs $\{v_1, v_2, \dots, v_n\}$ à l'instant ou l'intervalle concerné. La signature logique de cette fonction est exprimée comme suit :

$$Ct_{fs}: (c, t) \longmapsto \text{set}(\text{tuple}(l, t', \{v_1, v_2, \dots, v_n\})) \quad \forall t' \in t$$

- Enfin la fonction Ct_{fr} prend les mêmes paramètres que les deux fonctions précédentes et renvoie un sous-ensemble des localisations spatio-temporelles (l_r, t') avec leurs valeurs $\{v_1, v_2, \dots, v_n\}$. Plus particulièrement, l'élément l_r consiste en une localisation spatiale relative de type $Rloc$. La signature logique de cette fonction est exprimée comme suit :

$$Ct_{fr}: (c, t) \mapsto \text{set}(\text{tuple}(l_r, t', \{v_1, v_2, \dots, v_n\})) \mid \\ l_r \in Rloc \quad \forall t' \in t$$

Les définitions formelles de ces trois fonctions ont été établies dans les sous-sections 3.5.3 (pour les fonctions Ct_{fh} et Ct_{fs}) et 3.5.4.2 (pour la fonction Ct_{fr}).

En ce qui concerne le type de données Ocr , une seule fonction hybride est spécifiée :

- Il s'agit de la fonction Ot_{fs} qui prend comme paramètres un objet o_c de type Ocr et un instant ou un intervalle temporel t et renvoie un sous-ensemble de localisations spatio-temporelles (l, t') avec leurs valeurs $\{v_1, v_2, \dots, v_n\}$ décrivant l'hétérogénéité spatio-temporelle des entités géographiques de type *occurrent*. La signature logique de cette fonction est exprimée comme suit :

$$Ot_{fs}: (o_c, t) \mapsto \text{set}(\text{tuple}(l, t', \{v_1, v_2, \dots, v_n\})) \quad \forall t' \in t$$

Cette fonction matérialise l'hétérogénéité spatio-temporelle d'un *occurrent* en représentant son champ interne pour un temps donné. Ainsi, Cette fonction représente une forme d'extension spatiale de la notion de champ-objet temporel définie dans la sous-section 3.4.2. La définition formelle de la fonction Ot_{fs} a été établie dans la sous-section 3.5.3.

Les fonctions hybrides définies dans ces deux types de données Ctn et Ocr représentent la perspective allant des objets vers la représentation du champ. L'autre perspective hybride partant du champ vers les objets est représentée au niveau du type de données $LocSet$. Notamment, deux fonctions hybrides Ft_c et Ft_O sont spécifiées :

- La fonction Ft_c qui prend comme paramètres un objet de type $LocSet$ avec un instant ou un intervalle de temps et renvoie un sous-ensemble d'entités géographiques qui sont de type continuant (i.e., Ctn). La signature logique de cette fonction est exprimée comme suit :

$$Ft_c: (ls, t) \mapsto set(c)$$

- La fonction Ot_c qui prend comme paramètres un objet de type *LocSet* avec un instant ou un intervalle de temps et renvoie un sous-ensemble d'entités géographiques qui sont de type *occurrent* (i.e., Ocr). La signature logique de cette fonction est exprimée comme suit :

$$Ft_o: (ls, t) \mapsto set(o_c)$$

Nous pouvons remarquer que la fonction Ft_c constitue une forme d'extension temporelle de la notion d'objet-champ spatial définie dans la section 2.3. Plus précisément, le domaine de la fonction objet-champ spatial (i.e., L) a été étendu vers un domaine spatio-temporel (i.e., $L \times T$). Ainsi, le codomaine de cette fonction est passé d'un ensemble d'objets statiques (i.e., O) à un ensemble de continnants (i.e., C). D'autre part, nous pouvons également remarquer que la fonction Ft_o constitue une forme d'extension spatiale de la notion d'objet-champ temporel définie dans la sous-section 3.4.1. Plus précisément, le domaine de la fonction objet-champ temporel (i.e., T) a été étendu vers un domaine spatio-temporel (i.e., $L \times T$). Ainsi, le codomaine de cette fonction est passé d'un ensemble d'événements (i.e., E) dont chaque élément est caractérisé uniquement par son empreinte temporelle à un ensemble d'occurrents (i.e., O_c) dont chaque élément est doté de son empreinte spatio-temporelle.

Notons que les définitions formelles des deux fonctions Ft_c et Ft_o ont été établies dans le chapitre précédent, plus particulièrement dans la sous-section 3.5.2.

Au-delà de la définition formelle et de la spécification logique, ces fonctions hybrides sont caractérisées par les propriétés suivantes :

- Ce sont des fonctions spatio-temporelles : c'est-à-dire qu'elles prennent comme paramètres ou renvoient des valeurs spatio-temporelles.
- Ce sont des fonctions qui renvoient un sous-ensemble de tuples ou d'éléments en retour :

$$F: (p1, p2, \dots, p3) \mapsto set(tuple(l, t, v))$$

- Ce sont des fonctions qui peuvent prendre en paramètres un ou plu-

sieurs sous-ensembles d'éléments ou de tuples :

$$F: (\text{set}(\text{tuple}(l, t)), p2, p3, \dots, pn) \mapsto c$$

L'exemple de la fonction Ct_{fs} (voir sa signature dans cette section et sa définition formelle dans la section 3.5.3) qui peut prendre en paramètre un ensemble de tuples de type $(ctn, temps)$ dénotant un état d'un continuant c .

- Ce sont des fonctions d'ordre supérieur : définie comme une fonction qui prend une ou plusieurs fonctions comme paramètres, ou bien renvoie une fonction comme résultat (Stark, 1994) :

$$F1: a \mapsto b$$

$$F2: (F1(a), p2, p3, \dots, pn) \mapsto c$$

l'exemple de la fonction hybride Ft_c (voir sa signature dans cette section et sa définition formelle dans la section 3.5.2) qui accepte comme argument la fonction f_l qui renvoie les localisations spatiales liées à un état d'un continuant.

- Ce sont des fonctions polymorphes : la signature que nous avons établie pour les fonctions hybrides peut être surchargée par d'autres paramètres ou définir d'autres fonctions du même nom mais avec des paramètres différents.

4.3.3 Opération composées temporelles, spatiales et spatio-temporelles

En complément des fonctions hybrides, les types de données Ctn , Ocr , $LocSet$ définissent des fonctions composées temporelles, spatiales et spatio-temporelles, respectivement, et qui renforcent les capacités analytiques du modèle de données hybride.

L'exemple d'opérations composées spatiales concerne les deux fonctions $getCentroid()$ et $getBoundary()$ qui prennent comme paramètre un élément de type $LocSet$ et renvoient respectivement le centroïde et la frontière relatifs à toutes les localisations spatiales comprises dans le type $locSet$ et passées en argument à ces deux fonctions.

L'exemple d'opérations composées temporelles correspond aux fonctions $lifetime()$ et $duration()$, dont la première prend comme paramètre des objets

de type *ctn* et retourne la "durée de vie" d'un continuant *c*, tandis que la seconde prend comme paramètre une entité géographique de type *ocr* et retourne son étendue temporelle.

L'exemple d'opérations composées spatio-temporelles correspond aux fonctions *trajectory()*, *getChunk()*, *stMetricRdNeighbor()*, *metricRoute()*, *topoRoute()* dont :

- *trajectory()* : prend comme paramètre un objet *c* de type *Ctn* et renvoie un ensemble d'emprises spatio-temporelles (g, t) associées à l'évolution d'un continuant *c*. La signature logique de cette fonction est exprimée comme suit :

$$trajectory(): (c) \longmapsto set(tuple(g, t')) \quad \forall t' \in lifetime(c)$$

- *getChunk()* : prend comme paramètre une entité géographique *o_c* de type *Ocr* et renvoie ses emprises spatio-temporelles (g, t) qui correspondent à la durée temporelle de l'entité *o_c*. La signature logique de cette fonction est exprimée comme suit :

$$getChunk(): (o_c) \longmapsto tuple(g, t) \mid t = duration(o_c)$$

- *stMetricRdNeighbor()* : prend comme paramètres une entité géographique *c* de type *Ctn*, un instant ou un intervalle *t* et une valeur numérique *r*. Elle renvoie tous les états des continuants voisins (i.e., (c', t')) avec leurs distances *d* par rapport à l'entité *c* dans un rayon métrique défini par la valeur numérique *r*. La signature logique de cette fonction est exprimée comme suit :

$$stMetricRdNeighbor(): (c, t, r) \longmapsto set(tuple(c', t', d)) \quad \forall t' \in t$$

Cette opération est définie formellement par la fonction f_{vg} dans la sous-section [3.3.2](#).

- *topoRoute()* : prend comme paramètres une entité géographique *c* de type *ctn* avec un instant ou un intervalle *t*, et renvoie tous les états des continuants (i.e., (c', t')) ayant une relation spatio-temporelle topologique avec l'entité d'entrée. Chaque état renvoyé est associé avec

une valeur v_q qui spécifie le type de la relation spatio-temporelle (ex. touche, intersection, recouvrement, etc.). La signature logique de cette fonction est exprimée comme suit :

$$topoRoute(): (c, t) \longmapsto set(tuple(c', t', v_q)) \quad \forall t' \in t$$

Cette opération est définie formellement par la fonction f_{rt} dans la sous-section 3.3.2.

- *metricRoute()* : prend comme paramètres une entité géographique c de type *Ctn* avec un instant ou un intervalle t et une valeur numérique d , et renvoie tous les états successifs des continuants (i.e., (c', t')) ayant une relation spatio-temporelle métrique (i.e., selon une distance donnée d) avec l'entité d'entrée c . Chaque état renvoyé est associé à une valeur v_q qui, dans ce cas, précise la distance spatio-temporelle entre chacun des deux états successifs qui doit être inférieure ou égale à d . La signature logique de cette fonction est exprimée comme suit :

$$metricRoute(): (c, t, d) \longmapsto set(tuple(c', t', v_q)) \quad \forall t' \in t$$

D'autres fonctions supplémentaires de manipulation pourront être définies, à savoir, la fonction *ToLocSet()* qui transforme un tableau d'identifiants des instances de la classe *FieldLocation* (i.e., localisations spatiales L) vers un objet de type *LocSet*. En ce qui concerne la fonction *lscToRows()*, elle transforme un objet de type *locSet* vers un ensemble de tuples, dont chaque tuple correspond à une instance de la classe *FieldLocation* (voir les opérations liées au type *LocSet* dans la figure 4.6).

Il convient enfin de noter que le modèle reste indépendant des considérations d'implémentation telles que les langages de programmation, les plateformes et architectures logicielles. Néanmoins, dans la section suivante, nous proposons un modèle de mise en œuvre qui prendra la forme d'un schéma de base de données objet-relationnel générique en définissant l'architecture globale du système.

4.4 Modèle d'implémentation

Dans cette section nous introduisons notre architecture globale selon laquelle le modèle est implémenté. Nous définissons et instancions les différents types de structures de données pour représenter avec un niveau de détail plus élevé les spécifications provenant de la couche logique. Ces structures

de données, avec leurs fonctions, sont gérées à partir du schéma de base de données générique qui permet d'étendre le modèle objet-relationnel, plus particulièrement son langage de requête SQL avec des opérations de manipulation et d'analyse de données spatio-temporelles permettant de gérer des requêtes hybrides.

4.4.1 Architecture et plateforme

Nous présentons dans cette section l'architecture logicielle qui permet la mise en œuvre du modèle conceptuel de données hybride présenté dans la figure 4.4. La figure 4.7 montre les composants de cette architecture logicielle qui est essentiellement basée sur le système de base de données objet-relationnelle PostgreSQL qui offre divers types de structures de données telles que des tableaux (i.e., arrays) et des documents de type JSON et XML. PostgreSQL propose différents types de techniques d'indexation permettant d'améliorer l'efficacité des exécutions de requêtes SQL (ex., Btree, GIN, GIST). Il agit comme une plateforme extensible où l'utilisateur peut définir ses propres types de données et permet la programmation des routines grâce à son support de différents langages de programmation, à savoir PL/pgSQL, PL/Python, PL/Perl, C/C++ et PL/R.

En plus du système de gestion des bases données PostgreSQL, défini comme le cœur de notre architecture, celle-ci (i.e., notre architecture) est composée de différents composants (voir figure 4.7) à savoir :

- Composante spatiale : cette composante est gérée par l'extension spatiale de PostgreSQL, appelée PostGIS, qui prend en charge les modèles de données spatiales vectorielles et matricielles. C'est une exigence forte de notre approche surtout en terme d'intégration des données spatiales.
- Composante temporelle : elle correspond aux objets temporels, notamment les types de données temporelles et les opérateurs de manipulation des relations temporelles entre les intervalles de temps. Par exemple, afin de représenter les champs dans le temps, le cadre *TimeDBscale* gère un partitionnement dynamique des tables basé sur un attribut temporel. Afin de gérer l'évolutivité en termes de stockage de la base de données, notamment pour la table *FieldState* qui stocke des emplacements spatio-temporels avec leurs valeurs, nous utilisons l'extension *TimeDBscale*. Celle-ci permet le fractionnement dynamique de la table basé sur l'attribut temporel. Cela signifie que le sous-ensemble des tuples enregistrés pour un même temps seront

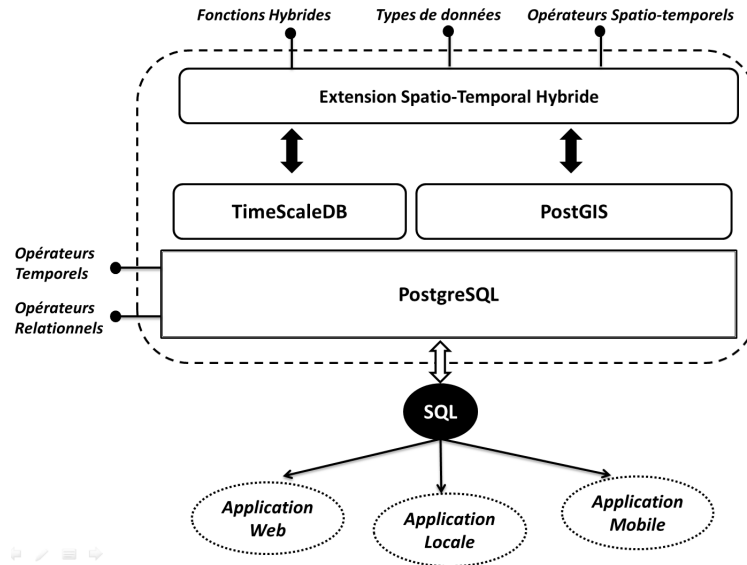


FIGURE 4.7 – Implémentation du modèle de données hybride avec une architecture basée sur PostgreSQL

stockés dans une nouvelle table. Notons que ces tables créées dynamiquement seront gérées par la seule table mère qui est *FieldState*.

- Composante hybride : outre les deux composantes spatiales et temporelles, nous avons développé une extension qui, non seulement exploite toutes les fonctionnalités offertes par les systèmes PostgreSQL et PostGIS, mais met également en œuvre une série de types de données, d'opérateurs et de fonctions hybrides tels que définis dans notre approche de modélisation. L'objectif principal est essentiellement d'étendre et d'enrichir le langage SQL par les opérateurs, les types de données et les fonctions hybrides présents dans notre modèle de données.

Avec cette extension, le modèle de données sera capable de supporter des requêtes hybrides spatio-temporelles en se basant sur la combinaison des deux modèles d'objet et de champ au sein d'une représentation intégrée des données.

- Composante de visualisation : l'avantage de PostgreSQL réside dans sa connectivité avec plusieurs autres plateformes logicielles à l'aide d'interfaces de programmation d'applications (API). C'est notamment le cas avec les logiciels de SIG tels que *QGIS* et *GRASS* qui permettent de représenter et manipuler des données géographiques.

Dans notre cas, et afin de matérialiser les différentes requêtes spatio-temporelles expérimentales, tous les résultats des requêtes hybrides seront présentés visuellement à l'aide du logiciel *QGIS*.

4.4.2 Schémas de base de données objet-relationnel hybride

Afin de poursuivre la mise en œuvre du modèle conceptuel hybride combinant les deux visions de l'espace champ et objet, un schéma objet-relationnel extensible a été spécifié et mis en œuvre. Nous allons montrer comment tous ces concepts de classes et de types de données peuvent être rassemblés dans un tel schéma. Nous créons trois tables pour la conception basée sur le champ et trois tables pour la conception basée sur l'objet. Une série d'attributs est ensuite spécifiée en fonction des types de données définis précédemment.

4.4.2.1 Schéma basé sur le modèle champ

La représentation du modèle conceptuel du champ dans un schéma logique est établie à travers les trois tables suivantes :

1. *FieldLocation* : (*locID* : *BigInt*, *pos_loc* : *Geometry(Point)*)
2. *FieldState* : (*fsID* : *BigInt*, *locID* : *Int*, *time_fs* : *Timestamp*, *field_value* : *Json*)
3. *Coverage* : (*cvrID* : *Int*, *cvr_extent* : *Int*, *coll_loc* : *int []*, *metadata* : *Json*)

La table *FieldLocation* gère l'ensemble des champs spatio-temporels en stockant les localisations spatiales avec leurs identifiants. Ces localisations spatiales sont gérées par l'attribut *locID*. Cet attribut va être lié à la table *FieldState* afin que ses valeurs soient associées à la fois à la composante temporelle (i.e., attribut *time_fstp*) pour représenter la notion de localisation spatio-temporelle, et à l'ensemble des valeurs correspondant aux différentes propriétés prises en compte.

Par conséquent, la structure de données la plus importante de ce bloc est la table *FieldState* qui gère toutes les localisations spatio-temporelles avec leurs valeurs. Il s'agit d'une nouvelle façon de représenter la notion de champ à l'aide de données semi-structurées. Le fait que nous définissions l'attribut *Field_value* comme document de type *json* pour chaque emplacement spatio-temporel (voir figure 4.8), rend le modèle plus générique dans le sens où

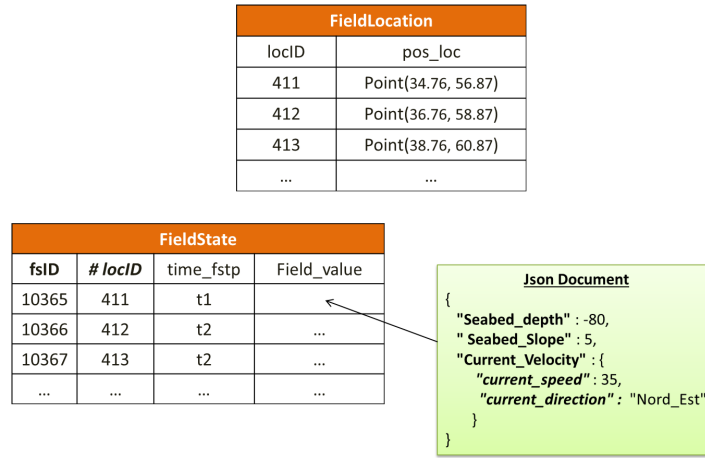


FIGURE 4.8 – Instantiation du modèle champ

il permet de gérer de multiples propriétés avec leurs valeurs, y compris les valeurs de forme scalaires et les valeurs de forme vectorielle.

4.4.2.2 Schéma basé sur le modèle Objet

La représentation du modèle conceptuel de l'objet est établie à travers les trois tables suivantes :

1. *GeoEntity* : (*geoID* : *Int*, *gtype* : *String*, *spatialExtent* : *Geometry*, *temporalExtent* : *tsrange*, *internLoc* : *LocSet*, *cattrSet* : *Json*)
2. *CState* : (*cstateID* : *Int*, *obj_ctn* : *ctn*, *cgeom* : *Geometry*, *ctime* : *Timestamp*, *internLoc* : *LocSet*, *cattrSet* : *Json*)
3. *OccurrentTable* : (*evID* : *Int*, *obj_ocr* : *ocr*)

La classe *GeoEntity* est implémentée dans le schéma comme étant une table relationnelle qui stocke explicitement le type, l'emprise spatiale et temporelle, à la fois des continnants et des occurrents. La relation de spécialisation est traduite par la migration de l'identifiant de la table *GeoEntity* vers les types de données *ctn* et *ocr*. Le type de données *ctn*, avec son identité, est défini au sein de la table *CState*, ce qui permet de lui associer différentes valeurs pour chaque instant (Figure 4.9). Par conséquent, un continuant peut posséder plusieurs états.

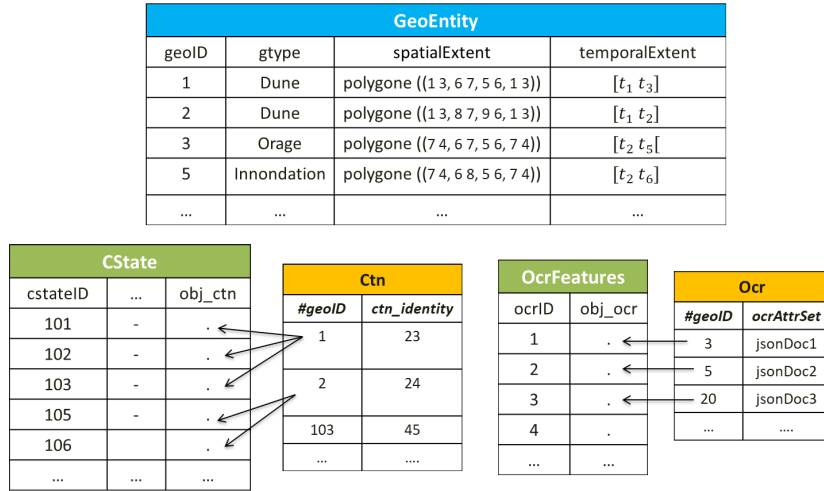
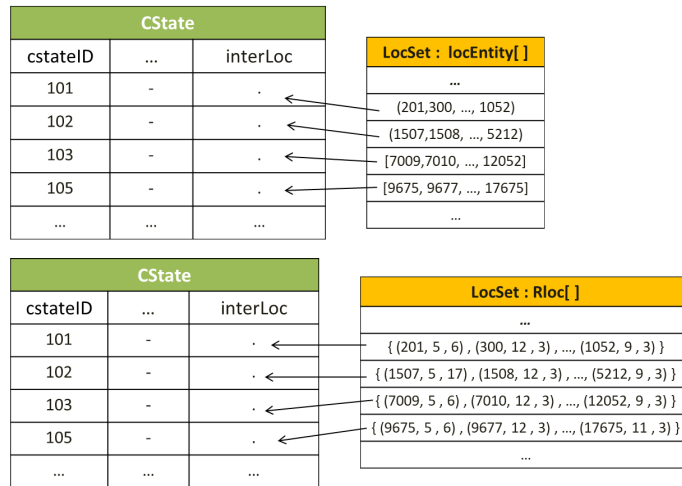


FIGURE 4.9 – Instanciation du modèle objet


 FIGURE 4.10 – Instantiation de la structure de donnée de type *LocSet* au sein de la table *CState*

L'association entre le schéma représentant la vision champ et le schéma représentant la vision objet est matérialisée à l'aide de l'attribut *internLoc* de type *Locset* dans la table *CState* (Figure 4.10).

Dans cette structure de données, l'utilisation de l'attribut *internLoc* permet une séparation entre les localisations spatiales (i.e., instances de la classe *FieldLocation*) et les valeurs des propriétés prises en compte (ex. pente du terrain, précipitation, etc.). Seules les références à un sous-ensemble de localisations spatiales sont stockées comme valeur de l'attribut *internLoc*. L'as-

sociation avec les valeurs du champ est assurée par les fonctions hybrides. Le fait de découpler les localisations spatiales de leurs valeurs, permet de gérer plusieurs champ pour un même continuant à partir de la création de l'attribut *internLoc*. Notons bien que cet attribut pourrait seulement stocker les références aux localisations spatiales *locEntity*, ou représenter ces localisations avec leurs coordonnées ou avec un index relatif *Rloc* comme le montre la deuxième structure de la figure 4.10.

4.4.2.3 Schéma hybride générique

Nous utilisons le concept de spécialisation qui a été défini dans le début de la section 4.2. de ce chapitre. Ce concept joue un rôle important en permettant à l'utilisateur de créer des schémas de base de données spécifiques pour un ou plusieurs phénomènes géographiques, en tirant partie de la structure générique de notre schéma hybride avec toutes les fonctionnalités qu'il contient.

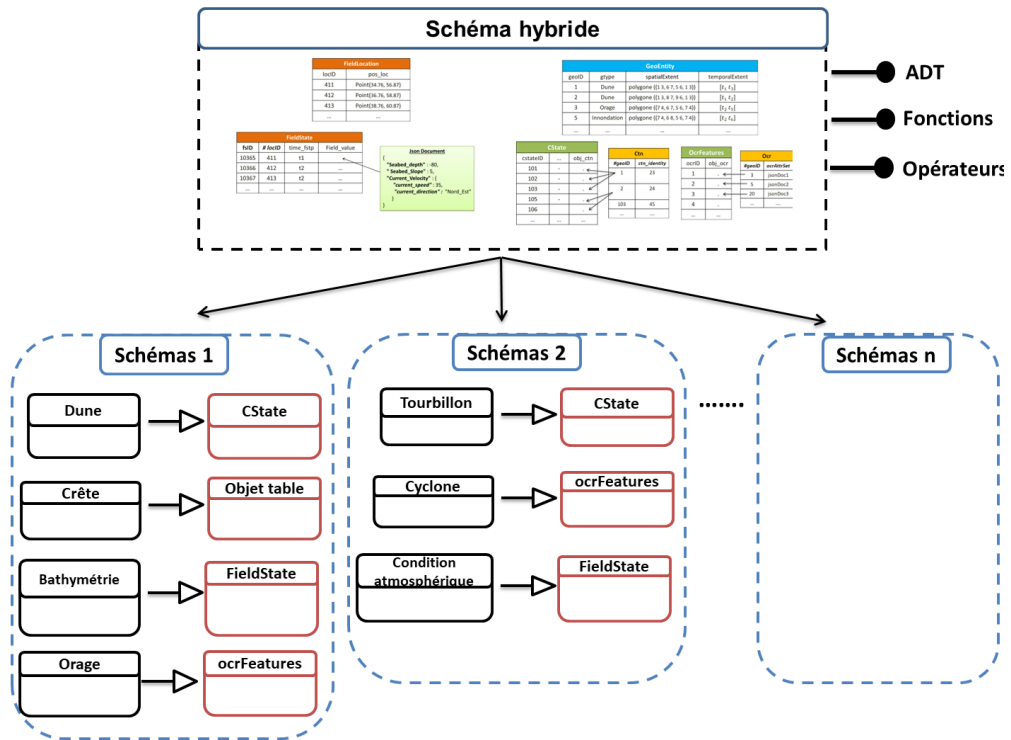


FIGURE 4.11 – Schéma générique de la base de données permettant d'implémenter le modèle hybride

La figure 4.11 montre l'exemple de deux schémas de base de données hy-

brides. Le schéma 1 permet de gérer des données décrivant l'interaction entre les dunes (i.e., continuant), un orage (i.e., occurrent) et la bathymétrie du fond marin (i.e., champ). Le schéma 2 gère les données décrivant l'interaction entre les tourbillons (i.e., continuant), les cyclones (i.e., occurrent) et les conditions atmosphériques (i.e., champ).

L'utilisateur peut définir un nombre n de schémas hybrides. Pour cela, il lui suffit de définir la correspondance de spécialisation (flèche avec un triangle vide) entre la table qui gère ses entités (i.e., dunes) et la table générique équivalente (i.e., *CState*). Au final, l'utilisateur hérite de tous les attributs et fonctions hybrides liés à cette structure. Cette opération de spécialisation est gérée par les trois opérateurs liés au schéma générique qui sont :

- *CreatefieldObject()* : cet opérateur prend comme arguments le type de l'entité (ex. dune, orage, etc) ainsi que la nature de l'entité géographique (i.e., continuant ou occurrent) et crée ensuite une table contenant des entités hybrides (i.e., continuant ou occurrent).

Exemple d'appel :

```
CreatefieldObject('dune', 'continuant');  
CreatefieldObject('orage', 'occurrent');
```

Si le deuxième argument passé à la fonction est "continuant", la nouvelle table créée héritera de la table de *CState*, alors que si le deuxième argument passé à la fonction est "occurrent", la nouvelle table créée héritera de la table de *ocrFeatures*.

- *Createfield()* : cet opérateur prend un seul argument dénotant le type de champ que nous voulons créer.

Exemple d'appel :

```
Createfield ('Bathymétrie');
```

A compter de son appel, une nouvelle table est créée dont le nom est celui passé en argument de la fonction. Notez qu'une seule table de type champ gèrera plusieurs variables, en l'occurrence la table *Bathymétrie* gèrera tous les paramètres liés à la surface du fond marin, c'est-à-dire la pente, la rugosité, la courbure, etc. En effet, la table *Bathymetry* héritera l'attribut *field_value* de la table *FieldState*.

Étant donné que, comme expliqué dans la section 4.2.1 du présent chapitre, l'attribut *field_value* est un document de type *json* (voir figure 4.8), cela permet de gérer plusieurs variables avec leurs valeurs pour chaque emplacement spatio-temporel.

A partir de cette implémentation, les types de données *ctn*, *ocr* et *LocSet* peuvent être spécialisés en définissant des sous-types qui héritent de toutes les propriétés des types originaux et en rajoutant d'autres nouvelles caractéristiques. L'utilisateur pourrait même définir d'autres fonctions pour chaque type de données de base, de manière à pouvoir redéfinir les mêmes fonctions hybrides existantes en les remplaçant par d'autres paramètres ou en définissant de nouveaux paramètres d'entrée pour les mêmes fonctions.

Ces caractéristiques rendent le modèle d'implémentation plus générique et extensible, lui permettant de gérer plusieurs études de cas et de l'enrichir avec d'autres fonctionnalités en fonction des besoins spécifiques des utilisateurs.

4.5 Conclusion et synthèse

Les travaux de ce chapitre ont conduit à la construction d'un modèle de données pour un SIG temporel qui prend en compte une représentation hybride de l'espace basée sur les notions d'objet et de champ, et la dichotomie entre les notions de continuant et d'occurrent. Ce processus de construction est à la fois conceptuel et logique.

D'un point de vue conceptuel, nous avons utilisé quatre mécanismes de modélisation afin de pouvoir relier notre représentation hybride à un modèle de données spatio-temporelles. Il s'agit des mécanismes de généralisation, d'agrégation, de composition et de spécialisation.

Au niveau logique, nous définissons trois nouveaux types de données qui sont dérivés du modèle conceptuel hybride. Ces types de données sont : 1) *Ctn* qui représente la notion du continuant, 2) *Ocr* qui représente la notion d'occurrent et 3) *LocSet* qui matérialise la relation d'agrégation entre des états d'un continuant (i.e., instances de la classe *CState*) avec le domaine d'un champ (représenté par la classe *FieldLocation*). La spécification de ces types de données est le point clé pour l'intégration complète de notre approche hybride dans un modèle logique. Cela est dû au fait que cette spécification présente les caractères suivants :

- Elle permet la représentation explicite et générique des deux notions

de continuant et d'occurrent au sein d'un modèle de données logique ;

- Elle permet de matérialiser la relation entre les modèles champ et objet ;
- Elle permet d'intégrer l'ensemble des fonctions hybrides au sein d'un modèle de données logique.

A ce stade, nous identifions les propriétés spécifiques qui caractérisent les fonctions hybrides. Celles-ci peuvent être appelées au niveau de la manipulation des données, car elles matérialisent toutes les relations établies entre les champs et les objets.

Par ailleurs, nous proposons une architecture logicielle ainsi qu'un schéma de base de données relationnelles-objets matérialisant une représentation logique qui supporte un modèle de données hybride. L'architecture définie est essentiellement basée sur le SGBD libre PostgreSQL et son extension spatiale PostGIS.

Chapitre 5

Experimentation et Application

” When scientists study the physical properties of atoms, such as mass and weight, they look and act like physical matter. However, when the same atoms are described in terms of voltage potentials and wavelengths, they exhibit the qualities and properties of energy (waves). The Universe is one indivisible, dynamic whole in which energy and matter are so deeply entangled it is impossible to consider them as independent elements.”

Bruce H. Lipton

Dans ce chapitre, nous allons mettre en œuvre l’implémentation de l’approche de modélisation développée dans les deux chapitres précédents. Cette approche est illustrée par un cas d’étude liée à un phénomène géographique complexe concernant la dynamique du fond marin. Cette complexité est due à la forme duale selon laquelle ce phénomène peut être appréhendé, d’une part à travers la perspective du champ en se focalisant sur la variabilité spatio-temporelle de la valeur de profondeur du fond marin et d’autre part en considérant un ensemble des dunes qui se déplacent et changent à la fois leurs formes géométriques et leurs relations spatiales et spatio-temporelles. Nous considérons cette thématique comme un exemple significatif et pertinent de phénomènes environnementaux pour lesquels une représentation

hybride présente un grand intérêt. L'objectif de ce chapitre n'est pas de faire une étude de ce phénomène géographique, mais plutôt une démonstration du rôle que peut jouer un modèle de données intégrant les notions de champ et d'objet, en permettant l'expression d'une série de fonctions hybrides à partir de requêtes SQL intégrées qui montrent le potentiel de notre approche de modélisation au niveau de la manipulation des données. L'accent est mis sur les requêtes spatio-temporelles décrivant l'interaction entre la dynamique de l'espace appréhendée avec une vision champ et objet, pour globalement étudier un phénomène dynamique de géomorphologie marine. Une série de tests de performance et d'évolution du système sont finalement présentés et montrent comment le modèle de données dans son ensemble permet l'exécution de différentes requêtes spatio-temporelles hybrides.

5.1 Méthodologie générale de l'expérimentation

La méthodologie de développement de cette expérimentation est basée sur deux axes comme le montre la figure 5.1. Ces deux axes comportent plusieurs étapes et utilisent le schéma de base de données hybride développé dans la section précédente pour mener à bien l'expérimentation.

Le premier axe, nommé **a** dans la figure, vise à explorer la capacité d'expressivité du modèle de données en termes de formulation de requêtes spatio-temporelles hybrides. Nous utilisons le langage de requêtes SQL, étant le langage de requêtes utilisé par la plateforme PostgreSQL, pour identifier les changements spatio-temporels au niveau des objets et au niveau du champ afin d'exprimer ces requêtes hybrides. Notons que Le langage SQL utilisé est étendu par l'ensemble des opérateurs et des fonctions de notre modèle hybride.

Le deuxième axe, nommé **b** dans la figure, vise à examiner les performances du modèle en termes de capacité et de temps de calcul pour chaque type de requêtes hybrides. Il s'agit donc d'examiner comment cette capacité est affectée lorsque la quantité de données interrogées augmente. Pour ce faire, nous générons une quantité massive de données spatio-temporelles en adoptant une stratégie pour effectuer une telle évaluation. Les mêmes requêtes exprimées dans l'axe **a** seront exécutées, mais avec le volume de données générées. Enfin, nous présentons les résultats sous forme d'un ensemble de graphes et une discussion portant sur les résultats obtenus.

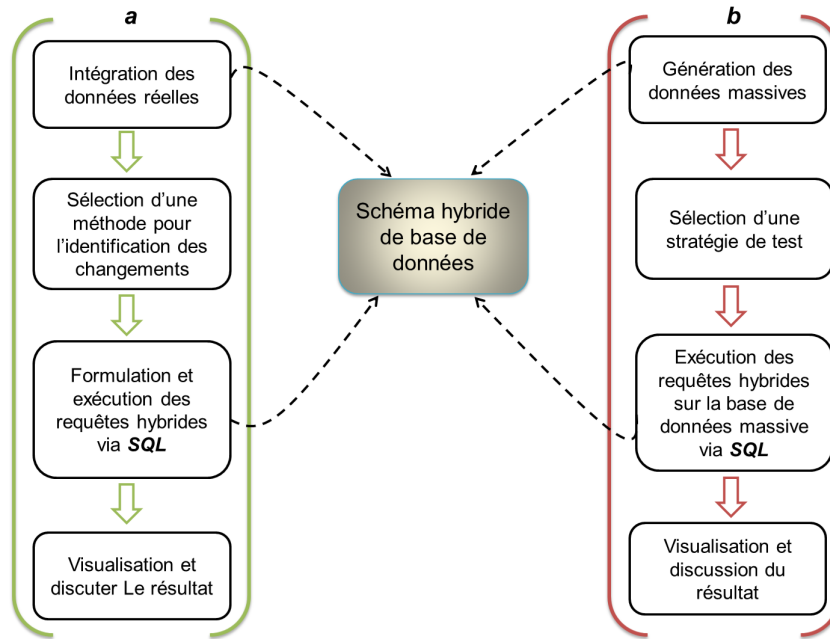


FIGURE 5.1 – Méthodologie d’expérimentation du modèle hybride

5.2 Contexte et zone d’étude

Le contexte pris en compte est lié à la volonté de comprendre la dynamique des fond marins en eau peu profonde. Focalisé sur l’observation des dunes sous-marines, l’objectif est de proposer un modèle capable de prendre en compte le processus de migration et d’interaction de ces formes de relief.

La zone d’étude est située à environ 10 km au large de la Bretagne Occidentale dans la région de jonction entre la Manche et l’océan Atlantique Nord, bien connue pour ses forts courants de marée et ses grandes houles (Figure 5.2). La zone d’étude est caractérisée par la présence d’un ensemble de dunes de sable marines (i.e., un champ de dunes).

La forme et le déplacement des dunes sont fortement corrélés à leur nature (grain, sédiment, taille) et aux facteurs hydrodynamiques (courants, marée, houle, vagues) (Le Bot, 2001; Wever, 2004). Ainsi, au fil du temps, les observations et la modélisation montrent que les déplacements varient en vitesse et en direction. Ceci est influencé par des phénomènes naturels dont l’intensité et l’impact sont plus ou moins importants. Il est très difficile d’estimer la dynamique sur une longue période, car les événements locaux spécifiques (une tempête par exemple) jouent un rôle aussi important que les phénomènes périodiques et plus réguliers causés par les courants de marée Le Bot (2001). Comme cette variabilité des facteurs hydrodynamiques est

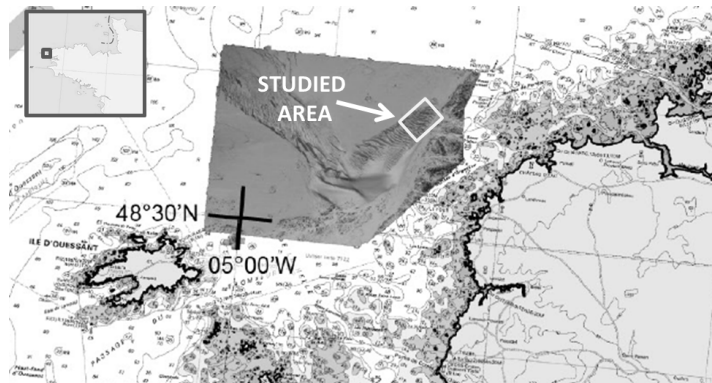


FIGURE 5.2 – zone d'étude

importante, il n'existe pas de lois générales sur la vitesse de déplacement et, par conséquent, sur la stratégie d'étude à mettre en oeuvre. Ainsi, nous pouvons considérer que l'ensemble des dunes liées à chaque forme de fond est unique et affecté par plusieurs processus superposés de durée et de force différentes.

Les études menées dans ce domaine suggèrent la méthodologie suivante : la compréhension des processus nécessite de prendre en considération plusieurs dunes et d'appréhender la dynamique des dunes tout en tenant compte de leur environnement géomorphologique et hydrodynamique ([Livingstone et al., 2007](#)).

Par conséquent, notre objectif est de montrer comment notre modèle hybride pourrait aider les scientifiques à améliorer leurs compréhensions de ce phénomène à travers des requêtes hybrides portant sur la complexité de l'interaction entre la variation spatio-temporelle des fonds marins représentée par les données de bathymétrie (dynamique des champs) et les évolutions des dunes marines dans le temps (dynamique des objets).

5.3 Aperçu de l'ensemble des données

Le jeu de données utilisé dans notre démonstration est un ensemble de levés bathymétriques enregistrés sur trois années, de 2009 à 2011, sous la forme de modèles numériques de terrain (MNT) avec un maillage de 2 m. Le premier levé a été effectué en février 2009 par le bureau hydrographique français, le second a été enregistré en août-septembre 2010 et le troisième en juillet 2011, tous deux par l'Institut Universitaire Européen de la Mer (respectivement dénommés t1, t2 et t3 dans les sections suivantes). Les observations permettent d'identifier 35 dunes. L'identification de ces dunes repose sur la

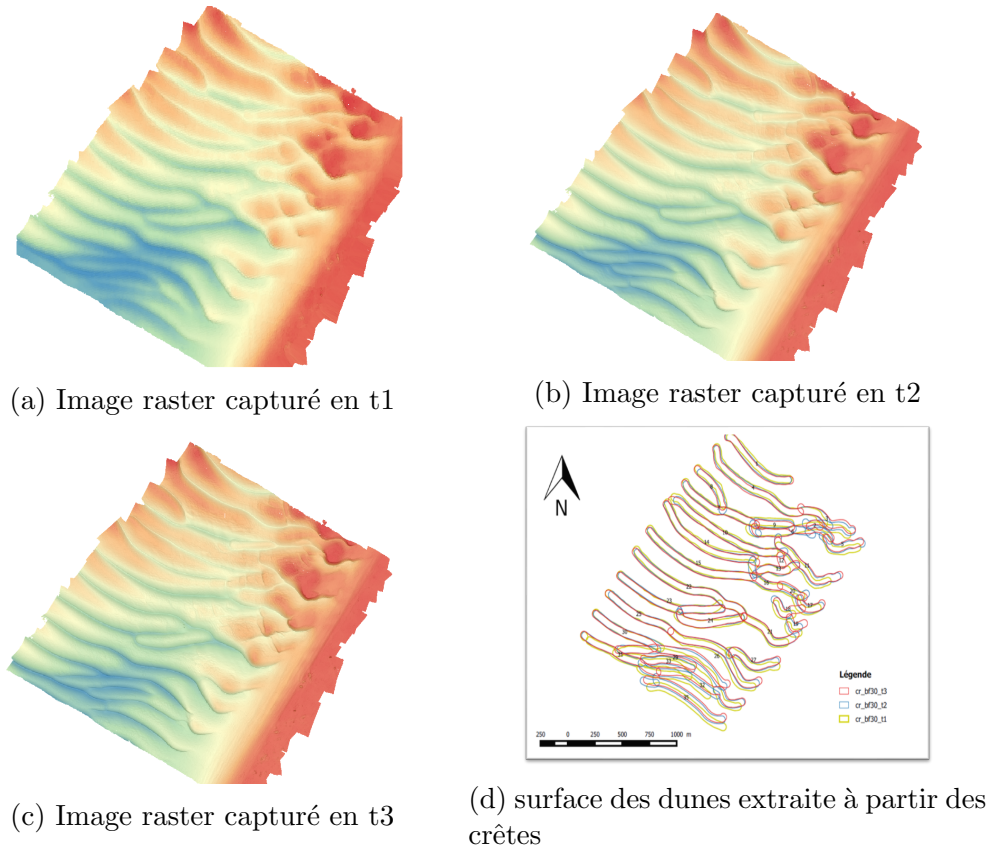


FIGURE 5.3 – Données de test

localisation des crêtes des dunes et des ruptures de pente à la base de leurs flancs.

5.4 Initialisation du modèle de données

Les méthodes courantes de description des changements de relief sont actuellement basées sur la comparaison d'un ensemble de MNE (modèles numériques d'élévation) enregistrés à des moments successifs ([van Dijk and Egberts, 2008](#); [Franzetti et al., 2013](#)). D'autres approches sont également basées sur des comparaisons de primitives géométriques comme la localisation remarquable (i.e., les points les plus élevés), la morphologie particulière des lignes (i.e., la ligne de crête, pied de pente) et les caractéristiques des profils bathymétriques (i.e., les profils en coupe transversale perpendiculaires à la ligne de crête) ([Garlan et al., 2013](#); [Thibaud et al., 2013](#); [Homrani et al., 2019](#)). En fait, et comme précédemment souligné, un modèle hybride est

nécessaire pour tenir compte de la dynamique du relief : La représentation de l'espace basée sur un champ correspond à la bathymétrie qui décrit une surface continue liée au fond marin, tandis que la représentation basée sur un objet est liée à l'identification de la dune à partir de la ligne de crête et/ou de ses pieds de pentes.

Dans notre cas, la mise en œuvre est basée sur le système de base de données objet-relationnel PostgreSQL, et son extension spatiale PostGIS. Cette extension supporte les modèles de données spatiales vecteur et raster, ce qui est une exigence forte de notre approche. Toutes les tables logiques représentant le cœur de nos schémas hybrides qui ont été introduites dans le chapitre précédent ont été implémentées dans PostgreSQL. Pour les aspects basés sur le champ, nous nous sommes concentrés sur la propriété de la profondeurs du fond marin. Les fonctions de type raster de Postgis (`ST_PixelAsPoints` et `ST_Value`) sont utilisées afin d'extraire une géométrie de point de chaque pixel avec les valeurs d'élévation correspondantes. Elles sont utilisées pour remplir les tables *FieldLocation* et *FieldState*. Pour la vue basée sur l'objet, nous enregistrons la dune sous la forme d'un polygone dans la table *CState*. Une dune est considérée comme une entité polygonale définie par une crête et par deux flancs (un flanc doux et un côté raide) limités par une rupture de pente à la base des flancs (i.e., le pied de pente). L'identification de la limite des flancs est parfois ambiguë et peut nécessiter l'utilisation de données de type sismique pour extraire la structure des emplacements sédimentaire à l'intérieur la dune. Dans notre cas, nous nous limitons à considérer l'objet dune à partir de sa ligne de crête. Pour en extraire un polygone, en tenant compte des dimensions des dunes présentes dans la zone d'étude, nous considérons arbitrairement une zone tampon de 30 m de part et d'autre de la crête pour identifier les entités dunaires (figure 5.3). Ainsi, nous identifions 35 dunes différentes. Certaines de ces dunes étant incomplètes dans notre jeu de données, nous conservons 21 enregistrements. Chacune de ces dunes est présente dans les trois jeux de données, pour un total de 63 instances dans la table *CState*.

5.5 Expression et visualisation des requêtes hybrides

Pour examiner l'expressivité du modèle de données, nous utilisons six requêtes spatio-temporelles qui décrivent six scénarios du monde réel. Nous qualifions ces requêtes de requêtes d'interaction hybride (HIQ). Les deux

premières requêtes explorent l’impact de la variation du champ dans le temps sur l’évolution de l’objet. Les deux requêtes suivantes suivront le processus inverse qui consiste à explorer l’impact possible de l’évolution de l’objet sur la variation du champ dans le temps. Enfin, les deux dernières requêtes explorent la variation du champ interne liée à l’évolution d’un objet au fil du temps. Avec ce modèle, nous exprimerons dans un premier temps les changements continus basés sur le champ et les changements basés sur l’objet en utilisant des constructions SQL, en particulier les fonctions de fenêtres et d’agrégation. Ensuite, nous exprimerons l’interaction entre les deux types de changements en utilisant l’une des trois fonctions hybrides FT_c , CT_{fh} , CT_{fr} dans chacune des requêtes.

Dans un premier temps il faut déterminer une méthode pour pouvoir identifier ou quantifier un changement tant du point de vue de la vision champ que celle objet.

5.5.1 Identification des changements champ et objet

L’identification des changements selon le point de vue champ s’effectuera en quantifiant les différences de valeurs enregistrées à des localisations spatio-temporelles successives (l_a, t_i) et (l_b, t_j) , portant sur une même composante spatiale (i.e., $l_a = l_b$) et des temps différents (i.e., $t_i < t_j$).

L’identification des changements en relation avec les objets se fera entre les états successifs (c, t_i) et (c, t_j) d’un même continuant c . Afin d’effectuer cette opération, nous utilisons un type de fonctions SQL appelées fonctions *fenêtres*. Plus précisément nous utilisons une variété de ce type de fonctions *lag over()*. Cette fonction est utilisée pour accéder à plusieurs lignes d’une table relationnelle sans utiliser de jointure interne.

Pratiquement, pour ce qui concerne l’observation de la morphologie du relief continue de notre zone d’étude, nous utilisons la méthode introduite par (Mitasova et al., 2011). Cette méthode consiste à estimer l’évolution de la profondeur du fond marin en calculant la régression linéaire des valeurs enregistrées à des temps successifs pour chaque localisation spatiale (i.e., pixel). Par conséquent, pour un intervalle de temps donné, les processus d’érosion où de déposition liés à la surface du fond marin peuvent être quantifiés par la pente de la fonction de régression. Cette opération est réalisée en utilisant la fonction d’agrégation SQL *regr_slope()*.

En ce qui concerne l’évolution des objets, nous nous concentrons sur deux types de changements : le premier est lié aux attributs géométriques

et le second concerne le changement des relations spatiales entre les objets. Pour les attributs géométriques, nous nous intéressons au processus de mouvement ainsi qu'aux variations de surface qui seront dérivées en utilisant une différence entre deux états de chaque continuant. Dans l'exemple du mouvement, il est calculé en se basant sur la distance euclidienne spatio-temporelle entre les centroïdes des deux états successifs d'un continuant donné. Pour cela nous ferons usage de la fonction de PostGIS *ST_distance*. Cette distance est calculée en prenant en compte les centroïdes des deux états de l'emprise spatiale de objets. Pour les relations spatiales entre objets, nous faisons appel à la fonction *stMetricRdNeighbor()* pour identifier le changement de relation spatiale (dans ce cas la relation métrique) entre plusieurs continuants pour un intervalle de temps donné (voir la section 4.3.3).

Pour assurer le processus d'agrégation spatiale dans nos requêtes, nous utilisons la fonction de PostGIS, *ST_collect*, qui nous permettra de rassembler plusieurs formes géométriques en une collection de géométries. Ainsi, nous pourrions associer pour chaque continuant, une collection de ces états géométriques en une seule instance et afin de faciliter ensuite la visualisation.

5.5.2 Analyse des capacités d'interrogation des modèles de données hybrides

Comme nous l'avons déjà mentionné, en particulier dans le deuxième chapitre, les modèles de données hybrides développés dans la littérature se concentrent sur l'étude des mouvements continus de la surface interne d'un objet. Ces requêtes visent simplement à interroger le sous-ensemble des points spatio-temporels avec leurs valeurs numériques à l'intérieur de l'étendue spatio-temporelle des objets en mouvement. Elles tentent le plus souvent de suivre l'évolution dans l'espace et le temps des valeurs extrêmes à l'intérieur de l'emprise spatiale de ces objets. Une requête comme " *Quelle est la vitesse maximale du courant superposé au déplacement d'une dune côtière pendant une tempête donnée ?*", effectuée en chaque instant une agrégation spatiale de la valeur du courant pour calculer la valeur maximale à chaque état et renvoie ensuite la trajectoire de cette valeur maximale.

L'inconvénient ici est que nous perdons la variabilité spatiale de tels phénomènes et que, par conséquent, l'interaction entre les deux espaces dynamiques (objet en mouvement et champ dynamique) ne peut être traitée.

Cette interaction pourrait être décrite selon trois perspectives, dont chacune est représentée par l'une des trois fonctions hybrides suivantes : $Ft_c()$, $Ct_{fh}()$ et $Ct_{fr}()$ (voir la Table 3.1 dans le chapitre 3). La première fonction $Ft_c()$ nous permet, en contrôlant les localisations spatiales (i.e., points de la table *FieldLocation*) et le temps, d'explorer l'impact de la dynamique du champ (dans notre cas : la variation de la surface du fond marin) sur les changements basés sur les attributs ou les relations entre objets (i.e., le déplacement ou rapprochement des dunes). Inversement, la deuxième fonction $Ct_{fh}()$, nous permet d'explorer l'impact de l'évolution des dunes sur la variation de la surface du fond marin, en contrôlant les dunes et le temps. Enfin la fonction $Ct_{fr}()$ permet de mesurer la variation du champ interne pour des objets en contrôlant les dunes et le temps.

La valeur ajoutée de l'approche hybride est qu'elle permet d'explorer la variabilité spatiale d'un phénomène donné à travers le temps. Une telle approche permet, pour une propriété donnée, d'effectuer une agrégation temporelle des valeurs de cette propriété à chaque localisation spatiale (e.g., calcul de la pente de la régression linéaire) et d'observer ensuite sa distribution spatiale.

Nous proposons maintenant de détailler la spécification des requêtes illustrant la manipulation des fonctions que nous avons définies, et l'utilisation de la base de données.

5.5.2.1 Explorer l'impact des changements de champ sur l'évolution des objets

Requête 1 : Sur les fonds marins qui se sont approfondies de plus de 0.7 mètre entre t_1 et t_3 , identifiez et ordonnez les quatre dunes qui ont eu le déplacement le plus important.

```
with cte as (select (C).* from
(select FtC (lse, '[2009-02-26, 2011-08-26]')) as C from
(select array_agg(ToLocSet(locID)) as lse from (select
locID, regr_slope(fval, foo.time_step) as erosion from
(select locID, time_fs, (field_value->>'elevation')::
numeric as fval, row_number() over (partition by locID
order by time_fs ) as time_step from FieldState) foo group
by locID having regr_slope(fval, foo.time_step) <-0.7 )
sq2)sq3)sq4)

select ctnFeature, moved_obj, sum_mig from
```

```
(select ctnFeature, st_collect(cgeom) as moved_obj, sum(
  ST_distance(ST_centroid(cgeom), ST_centroid(prv_geom))) as
  sum_mig from (select cte.ctnFeature, cte.cstateID, cte.
  time_, cs.cgeom, lag(cgeom) over (partition by (cte.
  ctnFeature).ctn_idt order by ctime) prv_geom from cte,
  CState cs where cte.cstateID = cs.cstateID )foo group by
  ctnFeature)foo2 order by sum_mig desc limit 4
```

La requête utilise la fonction $FTc()$ qui prend comme premier argument un sous-ensemble de localisations spatiales ayant des différences de profondeur inférieures à -0,7m calculées par la fonction *regr_slope* qui génère la pente de changement pour chaque localisation spatiale pendant la période sélectionnée (voir section 5.5.1). Ce sous ensemble est construit à l'aide de la fonction de PostgreSQL, *array_agg*, qui permet d'agréger plusieurs localisations spatiales en une seule collection. Le deuxième argument est un intervalle de temps pendant lequel les dunes évoluent. La fonction FTc renvoie la liste des dunes et de leurs géométries (i.e., *moved_obj*). La requête identifie ensuite les dunes ayant subi le plus grand taux de déplacement pendant l'intervalle de temps défini. Le résultat de la requête est illustré par la figure 5.4.

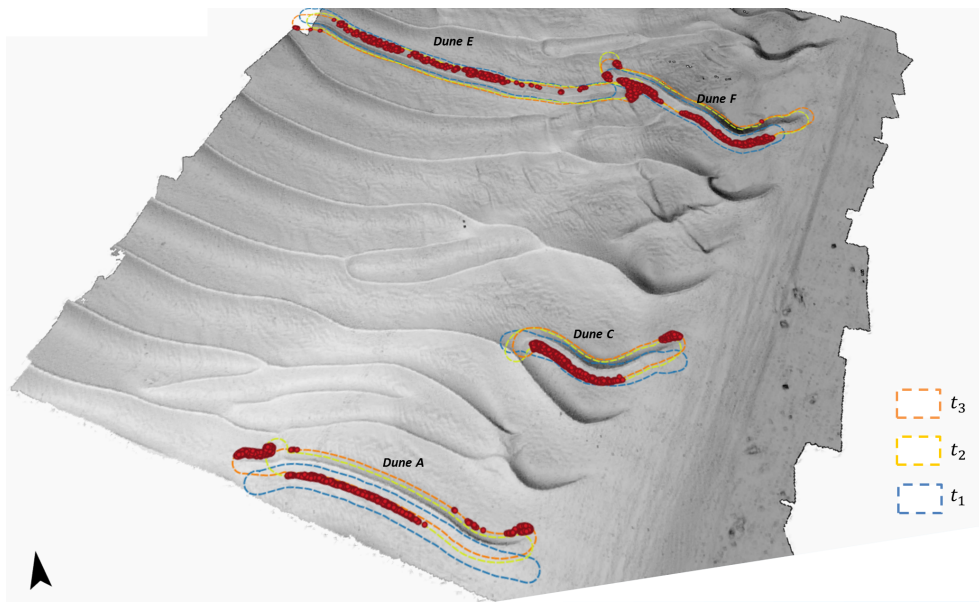


FIGURE 5.4 – Les quatre dunes les plus déplacées où les fonds marins ont subi une érosion supérieure à 0,7 m (localisées avec des points rouges)

Requête 2 : Sur les fonds marins qui se sont approfondis de plus de 1 mètre entre t_1 et t_3 , identifiez et ordonnez les deux paires de dunes qui se

sont le plus rapprochées.

```
with cte as ( select (C).* from(
select FTc (lse, '[2009-02-26, 2011-08-26]') as C from
  (select array_agg(ToLocSet(LocId)) as lse from
    (select LocId, regr_slope(fval, foo.time_step) as erosion
     from
      (select LocId, time_fstp, (field_value->>'elevation')::
       numeric as fval,row_number() over (partition by LocId
       order by time_fs ) as time_step from FieldState ) foo
     group by LocId having regr_slope(fval, foo.time_step) <-1)
    foo2)foo3)foo4)

select ctnFeature, ctn_vg, sum(dist_rate) as sum_rate from
(select *, (dist-lag(dist) over(partition by foo3.ctnFeature,
  foo3.ctn_2 order by foo3.ctime)) as dist_rate from

  (select distinct (stMetricRdNeighbor(ctnFeature,900)).*
   from
    (select ctnFeature from cte e1)foo)foo2 )foo3
  group by ctnFeature, ctn_vg order by sum(dist_rate) asc
  limit 2
```

Comme pour la requête précédente, l'idée consiste dans un premier temps à calculer la variation de profondeurs sur les reliefs continus du fond marin (i.e., variation supérieure à 1m) et de donner le résultat comme argument à la fonction *FTc*. Cette dernière renvoie comme résultat un ensemble de continuants qui sont affectés à la variable *cte*. La fonction *stMetricRdNeighbor()* est ensuite appelée qui, dans un premier temps, pour chaque dune en un instant *t*, sélectionne ses *n* voisins distants de moins de 900m, et ensuite calcule la distance avec chacun de ses voisins (i.e., *dist*). Ceci est représenté dans la figure 5.5 par différentes zones colorées. Les paires de dunes avec leurs taux de rapprochement ou d'éloignement sont ensuite sélectionnées (ex. les dunes L et E ont un taux de rapprochement de -27,45m alors que E et A ont -4,05m). Cela donne un sous-ensemble de paires avec leur taux de convergence ou de divergences (i.e., *dist_rate*), dont on conserve uniquement les deux plus importantes. Le résultat de cette requête est illustré par la figure 5.5.

5.5.2.2 Explorer l'impact des évolutions des objets sur le champ

Requête 3 : Indiquer les variations de profondeurs du fond marin entre t_1 et t_3 , associées à la présence de la dune dont la superficie a le plus évolué.

```
with ObjectAreaChange as (select ctnFeature, sum(ST_area(
  cgeom) - ST_area(prv_geom)) as SumAreaChange from (
  select ctnFeature, cstateID, ctime, cgeom,
```

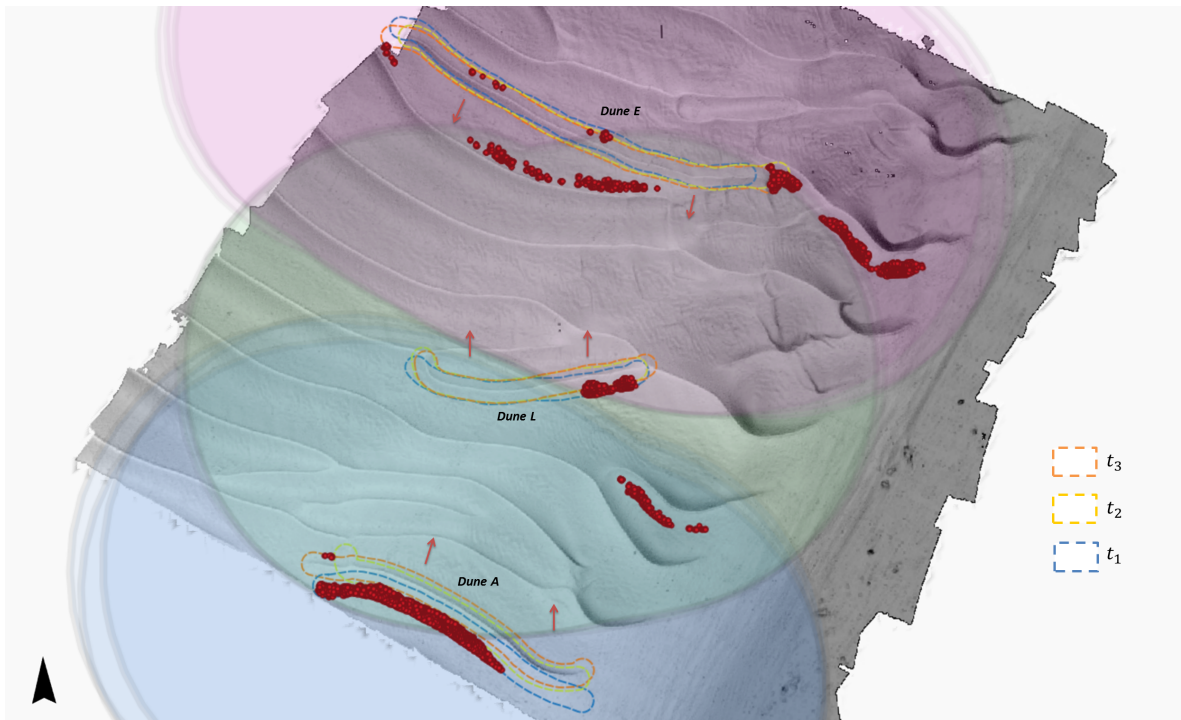


FIGURE 5.5 – Pour les localisations où l'érosion est supérieure à 1 m (situés avec des points rouges), identification les deux paires de dunes qui se sont le plus rapprochées (i.e., la dune E avec la dune L et la dune A avec la dune L). Les flèches indiquent les directions de déplacement.

```
lag(cgeom) over (partition by (ctnFeature).c_identity order
  by ctime) prv_geom from CState)foo group by ctnFeature)

select * from (select ctnFeature, cstateID, pos_loc,
  regr_slope(npval, time_step) as FieldChange from (select
  foo2.ctnFeature, id_cstp, locID, foo2.pos_loc, time_, (
  field_value->>'elevation')::numeric as npval, row_number()
  over (partition by locID order by time_ ) as time_step
  from
  (select ctnFeature, (CTfh(foo.gagg, '[2009-02-26,
  2011-08-2611]')).* as env from
  (select oc.ctnFeature, sum_area_change, array_agg(cgeom)
  as gagg from ObjectAreaChange oc, CState eh where
  sum_area_change = (select max(sum_area_change) from
  ObjectAreaChange oc) and oc.ctnFeature = oc.ctnFeature
  group by oc.ctnFeature, sum_area_change)foo)foo2)foo3 group
  by ctnFeature, locID, pos_loc )foo4 where elev_change >0
```

Cette requête présente un processus inverse des précédents où la requête

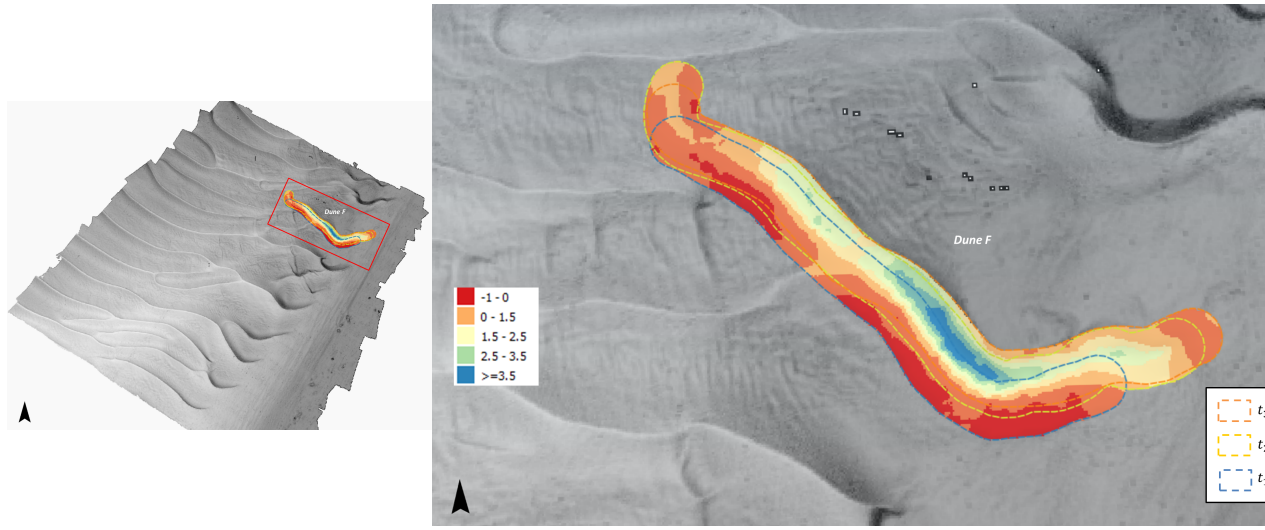


FIGURE 5.6 – Localisation et amplitude des variations de profondeur du fond marin entre t_1 et t_3 , associées à la présence de la dune dont la surface a le plus évoluée.

sélectionne d'abord la dune caractérisée par la plus grande variation de surface et la passe ensuite comme argument à la fonction $CTfh$. Celle-ci retourne le sous-ensemble des localisations spatiales couvert par le déplacement de la dune sur la période de temps spécifié. Enfin, avec une sous-requête (i.e., `foo3`), la variation de profondeur à chaque pixel est exprimée par la pente d'une régression linéaire sur les valeurs de profondeur. Le résultat de cette interrogation est illustré par la figure 5.6

Requête 4 : Indiquer les variations de profondeurs du fond marin entre t_1 et t_3 , associées à la présence de la paire de dunes ayant le taux d'éloignement le plus élevé.

```
with t1 as ( select ctnFeature, ctn_vg, sum(dist_rate) as
              sum_rate from
              (select *, (dist-lag(dist) over(partition by foo2.ctnFeature
              ,foo2.ctn_vg order by foo2.time_eph)) as dist_rate from

              (select distinct (stMetricRdNeighbor(ctn_ph,900)).* from
              (select ctnFeature from CState e1)foo)foo2 )foo3
              group by ctn_inp, ctn_vg order by sum(dist_rate) asc limit
              1)

t2 as (select ctnFeature from t1 union select ctn_vg as
```



```

ctnFeature from t1)

select * from (select ctnFeature, locID, pos_loc, regr_slope(
  npval, time_step) as FieldChange from
  (select foo2.ctnFeature, id_cstp, locID, foo2.pos_loc, time_
    , (field_value->>'elevation')::numeric as npval,
    row_number() over (partition by locID order by time_ ) as
    time_step from
  (select ctnFeature, (CTfh (foo.gagg, '[2009-02-26,
    2011-08-26]')).* as env from
  (select t2.ctnFeature, sum_rate, array_agg(cgeom) as gagg
    from t2, CState eh where t2.ctnFeature = eh.ctnFeature
    group by t2.ctnFeature, sum_rate)foo)foo2)foo3 group by
    ctnFeature, locID, pos_loc )foo4

```

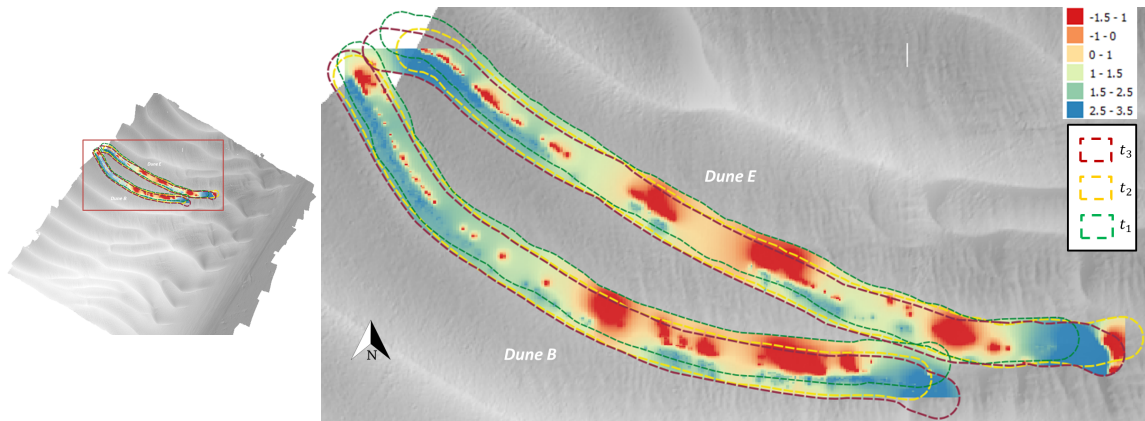


FIGURE 5.7 – Localisation et amplitude des variations de profondeur du fond marin entre t_1 et t_3 , associées à la présence de la paire de dunes ayant le taux d'éloignement le plus élevé.

Comme pour la requête 2, la fonction *stMetricRdNeighbor()* retourne les paires de dunes avec leur taux d'éloignement. La paire de dunes s'éloignant le plus rapidement est ensuite sélectionnée (i.e., les dunes E et F). Le résultat est donné ensuite en entrée à la fonction *CTfh* pour obtenir un sous-ensemble de localisations spatio-temporelles avec leurs valeurs de profondeur. Enfin, le changement continu est calculé par une régression linéaire. Le résultat de cette interrogation est illustré par la figure 5.7.

5.5.2.3 Explorer la variation spatio-temporelle du champ interne des objets

Requête 5 : indiquer les variations de profondeur et les déplacements sur la surface de la dune ayant le taux de déplacement le plus important entre t_1 et t_2 ainsi qu’une somme de surface inférieure de 15 hectares.

```
with step1 as (select ctnFeature, array_agg(row(cgeom,ctime)
:: st_geom):: st_geom [] as stg,    sum(prv_geom) as
mig_sum from (
select ctn_ph, cgeom , ctime, (ST_centroid(cgeom) <->
ST_centroid(lag(cgeom) over (partition by ctnFeature order
by ctime))) prv_geom from CState)foo group by ctnFeature
having sum(st_area(cgeom))/10000 <15 order by sum(
prv_geom) desc limit 1 ),

step2 as ( select ctnFeature, time_, pos_loc, k_ind, r_ind,
( p_val->>'elevation')::numeric as p_val from
(select ctnFeature, (CTfr(stg,8, 35, radians(0)::
numeric)).* from (select ctnFeature, stg from step1)doo)
loo)

select cidt, k_ind, r_ind, pos_loc, st_makeline(prv_pt,
pos_loc) as time_arrow, (p_val- pv_val) as diff, time_
from (select ctnFeature, k_ind, r_ind, time_, pos_loc, lag
(pos_loc) over w as prv_pt, p_val, lag(p_val) over w as
pv_val from step2
window w as (partition by ctnFeature, k_ind, r_ind
order by time_))doo2 where pv_val is not null
```

La requête sélectionne la dune caractérisée par le taux de déplacement le plus élevé et la transmet ensuite comme argument à la fonction **CTfr**. Celle-ci prend comme premier argument un ensemble de géométries estampillées (i.e., *stg*) représentant les états de la dune sélectionnée par la sous-requête *step1*. Les deux autres paramètres concernent les entrées de l’algorithme qui permet la génération des points dans l’emprise spatiale de la dune sélectionnée. Dans ce cas, nous choisissons à titre indicatif les valeurs 8 et 35 comme arguments à cette algorithme pour générer ensuite 280 (i.e. 8×35) points au total au sein dans l’emprise spatiale de la dune sélectionnée.

la fonction **CTfr** renvoie un ensemble de tuples points spatio-temporel avec leurs coordonnées relatifs *k_ind* et *r_ind* au sein de la dune sélectionnée. Ensuite, on associe des paires de points homologues (i.e., mêmes valeurs pour les coordonnées *k_ind* et *r_ind*). Enfin, la sous-requête calcule la différence de profondeur *diff* et renvoie un tuple (*pos_loc*, *time_arrow*, *diff*) qui indique la position du point, et la direction et l’amplitude de déplacement. Afin

de marquer la direction, nous construisons une ligne entre deux points de déplacement dans la dune à deux instants différents en utilisant la fonction *st_makeline*. Le résultat de cette interrogation est illustré par la Figure 5.8

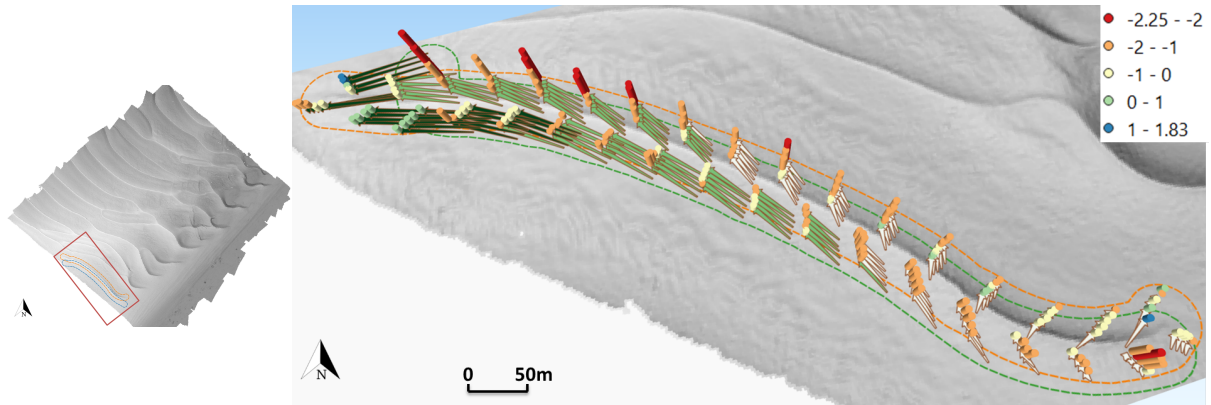


FIGURE 5.8 – variation de profondeur et vecteur de déplacement à la surface de la dune dont le taux de déplacement est le plus important entre t_1 et t_2 .

Requête 6 : indiquer les variations de profondeur et les déplacements sur la surface de la paire des dunes ayant eu le taux de rapprochement le plus élevé entre t_1 et t_3 .

```
with step1 as ( select ctnFeature, ctn_vg, sum(dist_rate) as
    sum_rate from
(select *, dist-lag(metric_dist) over(partition by foo4.
    ctnFeature,foo4.ctn_vg order by foo4.ctime) as dist_rate
    from

(select distinct (stMetricRdNeighbor(ctnFeature,600)).* from
(select ctnFeature from CState e1)foo)foo4 )foo5
    group by ctn_inp, ctn_vg order by sum(dist_rate) asc limit
    1),

step2 as (select ctnFeature as ctn_ph from t1 union select
    ctn_vg as ctn_ph from step1 )

step3 as ( select ctnFeature, time_, pos_loc, k_ind, r_ind, (
    p_val->>'elevation')::numeric as p_val from
    (select ctnFeature, (CTfr(stg,8, 35, radians(0)::
    numeric)).* from (select ctnFeature, stg from step1)do)
    loo)
```

```
select cidt, k_ind, r_ind, pos_loc, st_makeline(prv_pt,
pos_loc) as time_arrow, (p_val- pv_val) as diff, time_
from (select ctnFeature, k_ind, r_ind, time_, pos_loc, lag
(pos_loc) over w as prv_pt, p_val, lag(p_val) over w as
pv_val from step2
      window w as (partition by ctnFeature, k_ind, r_ind
order by time_))doo2 where st_makeline(prv_pt,ps) is not
null
```

Comme pour les requêtes 2 et 4, la fonction *stMetricRdNeighbor()* retourne les paires de dunes avec leur taux d'éloignement (i.e., dist). La recherche des paires de dunes est faite ici dans un rayon de 600 mètres, contrairement aux requêtes 2 et 4 pour les quelles le rayon était de 900 mètres. L'idée est de suivre le processus du rapprochement mais dans un rayon plus réduit afin de pouvoir identifier si, au cours de ce processus, l'une des dunes de la paire en question a pu affecter l'autre. La paire de dunes qui s'éloignent le plus rapidement est sélectionnée. Le résultat est ensuite donné en entrée à la fonction *CTfr*. Celle-ci renvoie un ensemble de tuples de points spatio-temporels, avec leurs coordonnées relatives *k_ind* et *r_ind* au sein de la paire de dunes sélectionnée. On associe ensuite les paires de points homologues (i.e., mêmes valeurs pour les coordonnées *k_ind* et *r_ind*) et la sous-requête calcule de différence de profondeur *diff*, et renvoie un tuple (*pos_loc*, *time_arrow*) qui indique la position du point, et l'amplitude et la direction de déplacement. Le résultat de cette interrogation est illustré par la Figure 5.9

5.6 Évaluation et Analyse des performances du modèle de données hybrides

Dans cette section et dans un premier temps, nous étudions la capacité de stockage liée aux deux structures de données les plus importantes de notre schéma hybride, à savoir les tables *FieldState* et *CState*. La première gère les données des champs spatio-temporels tandis que la seconde gère l'ensemble des états spatio-temporels des différentes continuants (ici les dunes). Nous examinons notamment l'évolution de la capacité de stockage à mesure que de nouvelles données sont introduites dans la base de données.

Dans un deuxième temps, nous concevons un test de performance pour un modèle de données hybride. Il consiste à évaluer les réponses du système aux requêtes hybrides et non en une comparaison avec la performance des autres modèles spatio-temporels. L'objectif est d'examiner comment le modèle de données réagit à l'augmentation progressive de la quantité de données insérées

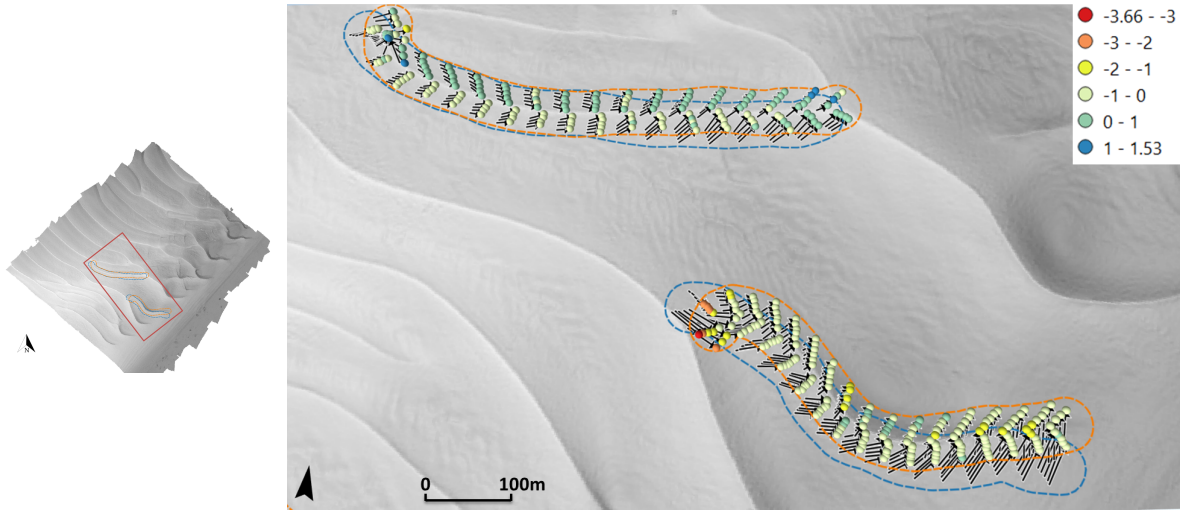


FIGURE 5.9 – Carte illustrant la variation de la profondeur du fond marin au sein des emprises spatiales de la paire des dunes ayant eu le taux de rapprochement le plus élevé entre deux instants t_1 et t_3 .

dans le système. Ces tests nous permettent d’identifier les parties du modèle de données qui doivent être mises à l’échelle et optimisées en identifiant les traitements les plus coûteux. Pour ce faire, nous utilisons les requêtes spatio-temporelles que nous venons de décrire dans la section 4 de ce chapitre. Pour éviter les redondances, nous considérons uniquement les quatre premières requêtes puisque les requêtes 5 et 6 sont similaires aux requêtes 3 et 4 en termes d’étapes de calcul et de complexité du calcul.

5.6.0.1 Analyse du stockage de schémas de base de données

Dans cette section, nous mettons l’accent sur la table *FieldState* qui gère toutes les localisations spatio-temporelles avec l’ensemble de leurs valeurs. La taille de la table est dépendante du nombre et de la dimension des propriétés (attributs) prises en compte et des types de données utilisés. Tous les types de données utilisés ont une taille constante, sauf le type de données *Json*. En prenant en compte la présence de valeurs numériques dans ce type de données, la taille des enregistrements dépendra 1) du nombre de propriétés de type champ, sachant que chacune de ces propriétés est gérée sous la forme d’un ensemble de tuples *nom-valeur* (voir la figure 4.8) et 2) du nombre de “digits” représentant la valeur, ce qui peut être exprimé par la formule : $(33 + \frac{n}{2}) \times m$ où 33 correspond au nombre d’octets occupé par une propriété (i.e., attribut) de type *Json* vide, n est le nombre de chiffres pour chaque

valeur et m le nombre de propriétés à l'intérieur du document *Json*.

En utilisant la formule simple et en supposant utiliser deux propriétés de type champ (ex. profondeur et pente), avec des valeurs numériques d'au maximum 12 chiffres, nous obtenons une taille de l'ordre de 78 octets. Par conséquent la taille totale d'une ligne de cette table est de $8(bigint) + 8(bigint) + 8(timestamp) + 78(json) = 102 \text{ octets}$ (voir la figure 4.8).

Le type de données *Json* pourrait être considéré comme coûteux en termes de taille de stockage, mais l'avantage est qu'il permet d'ajouter dynamiquement d'autres variables de type champ sans modifier la structure de la table.

Afin d'évaluer l'impact de la colonne *field_value* qui est de type *Json* sur le stockage de la table *FieldState*, nous traçons un graphique qui illustre comment la taille de stockage de cette table évolue au fur et à mesure que le nombre de propriétés (i.e., attributs) de type champ augmente à l'intérieur de la colonne *field_value* pour chaque localisation spatio-temporelle. Comme la taille de la table dépend également du nombre de lignes, nous faisons varier ce paramètre (i.e., nombre de lignes) (Figure 5.10).

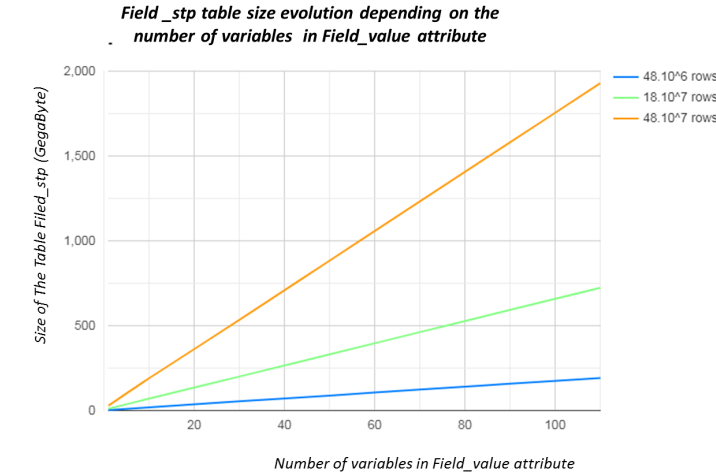


FIGURE 5.10 – Évolution de la taille du stockage de la table *FieldState* en fonction de l'attribut *field_value*

Comme le montre le graphe de la figure 5.10, quelque soit le nombre d'enregistrements (i.e., nombre de lignes), le volume de stockage augmente linéairement avec le nombre de propriétés de type *Json* dans la colonne *field_value*. Ce volume augmente d'autant plus fortement avec le nombre d'enregistrement et atteint rapidement 500 gigabits pour 48×10^7

enregistrements et cela sans compter la taille des index définis sur cette table.

Afin d’adapter la table à l’augmentation considérable de sa taille de stockage, nous utilisons un partitionnement temporel qui permet de diviser la table *FieldState* en plusieurs sous-tables sous format de blocs en fonction de l’attribut *time_fs*. Cette approche permet de gérer les enregistrements dans ces différentes sous-tables. Bien entendu, lors du processus d’interrogation de la base de données par l’utilisateur, toutes les requêtes sont adressées à la table parente *FieldState*. Pour obtenir de meilleures performances en terme de réponses aux requêtes spatio-temporelles, nous définissons trois index dont deux sont de type Btree, un sur l’attribut *locID* et l’autre sur l’attribut *time_fs* et un index Gist sur l’attribut *pos_loc* à l’intérieur de la table *FieldLocation*.

5.6.1 Préparation de l’expérimentation

Le modèle de données générique a été implémenté dans la plate-forme Postgresql 11 et a été spécifié en utilisant PL/pgSQL, un langage de programmation procédural qui permet d’étendre les fonctionnalités de Postgresql. Le modèle de données a été testé pour les mesures de performance tant pour le temps de calcul que pour la taille des données. Les tests ont été effectués sur un *HP Pavillon 14* avec un processeur *4,6 GHz Max*, 8ème génération, *Intel Core i7* avec 16 Go de RAM. Le système d’exploitation utilisé était *Ubuntu Linux 19.10*.

Pour effectuer le test, nous générons environ 48×10^7 localisations spatio-temporelles avec leurs valeurs de données. Si chaque localisation est considérée comme le centroïde d’une cellule de grille (i.e., pixel) de 2m de résolution, ce nombre pourrait couvrir une zone de 180 km^2 en dix instantanés de type raster. Cela équivaut à la superficie de l’île d’Aruba. Pour donner une idée, pour 5m de résolution, ce nombre de localisations couvrirait la superficie de l’île de Jersey (118 km^2) pour 100 instantanés de type raster. Notons que dans notre modèle, chaque localisation spatio-temporelle correspond à une ligne dans la table *FieldState*. Le nombre des propriétés de type champ (e.g., profondeur, pente, etc.) pour chaque localisation spatio-temporelle (i.e., ligne) est théoriquement illimité. Mais dans la pratique, à chaque fois que le nombre de propriétés augmente, cela a un coût sur la taille de chaque ligne, comme le montre le graphe dans la figure 5.10. Pour ce qui concerne le nombre d’états contenu dans la table *CState*, nous générons 15015 enregistrements qui couvrent la surface des 180 km^2 en dix instants temporels différents. La taille totale de la base de données en Postgresql est de 75GB.

5.6.1.1 Évaluation des mesures de performance dans le temps pour les requêtes hybrides

Avec cette évaluation expérimentale, nous observons les performances en temps d'exécution pour chacune des quatre requêtes introduites précédemment en relation avec le nombre de localisations spatiales ou d'objets impliqués.

Il est à noter que le temps de réponse d'une requête peut varier en fonction de plusieurs paramètres liés au planificateur et exécuteur des requêtes au sein du PostgreSQL. Afin de ne pas biaiser nos résultats, nous exécutons chaque requêtes 5 fois et nous conservons la moyenne concernant le temps de réponse des quatre requêtes (voir figures 5.11 et 5.12).

Les graphes de la figure 5.11 illustrent l'évolution du temps de calcul des requêtes qui utilisent la fonction Ftc , tandis que les graphes de la figure 5.12 illustrent les performances pour des requêtes hybrides utilisant la fonction Ct_{fh} .

La différence importante entre ces requêtes provient du fait que les requêtes 1 et 3 intègrent une sous-requête qui traite des changements temporels des propriétés des objets (ici la position ou la superficie) tandis que les requêtes 2 et 4 intègrent une sous-requête qui traite des changements de relations spatiales entre les objets (ici le rapprochement ou l'éloignement).

Pour les requêtes 1, 2 et 3, nous fixons le nombre d'objets (i.e., instances de la table $CState$) à 15015, alors que nous proposons trois configurations pour la requête 4 (1000, 8000 et 15015). En effet, outre l'interrogation des relations entre les objets (i.e., instances de la table $CState$), la requête 4 doit également parcourir toutes les instances de la table $CState$. Par conséquent, les performances du système ne dépendent pas seulement du nombre d'emplacements spatio-temporels à parcourir, mais aussi du nombre d'instances dans la table $CState$ (i.e., entités). Par conséquent, faire varier le nombre de ces instances a un effet significatif sur la tendance d'évolution de la courbe du graphique, ce qui nous amène à proposer trois courbes différentes spécialement pour cette requête (cf. figure 5.12 : (b)).

Sur la figure 5.11, pour les deux requêtes, nous observons une augmentation du temps de calcul avec le nombre de localisations spatiales. Cette augmentation est presque linéaire pour la requête 1, alors que pour la requête 2, nous voyons un plateau entre 100×10^6 et 250×10^6 localisations spatiales. De même, nous constatons que la requête 2, qui implique les relations spatiales entre objets (ici le rapprochement entre deux dunes), a

5.6. ÉVALUATION ET ANALYSE DES PERFORMANCES DU MODÈLE DE DONNÉES HYBRIDES

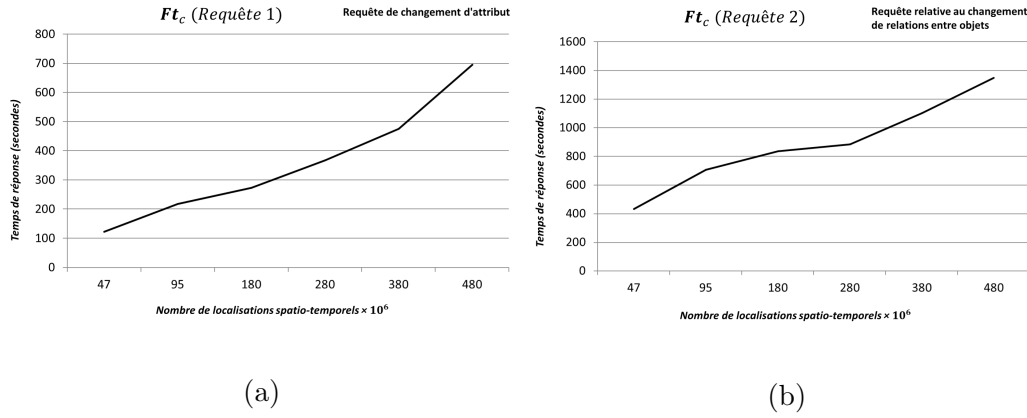


FIGURE 5.11 – Graphes mettant en évidence l'évolution du temps de calcul en fonction du nombre de localisations spatio-temporelles des requêtes qui utilisent la fonction Ft_c

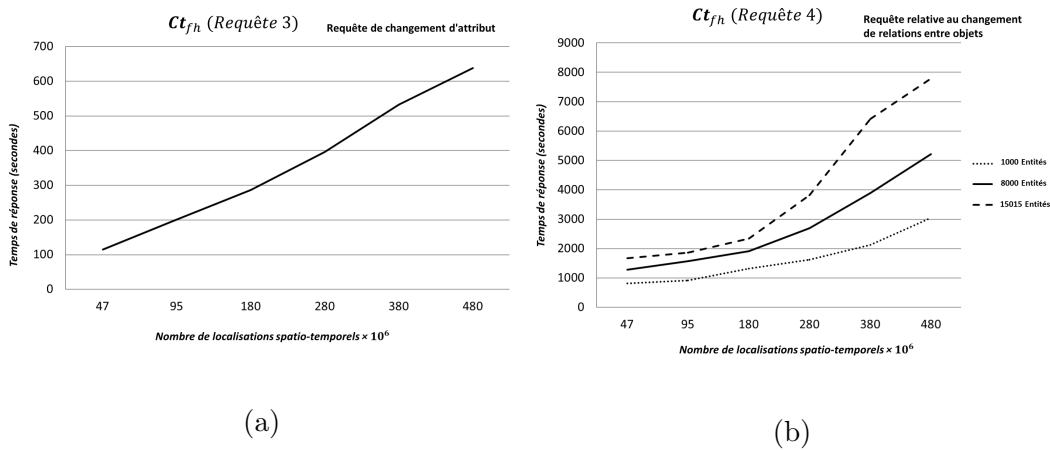


FIGURE 5.12 – Graphiques indiquant l'évolution du temps de calcul en fonction du nombre de localisations spatio-temporelles des requêtes hybrides utilisant la fonction Ct_{fh} .

un coût plus élevé, avec une moyenne de 14 min 47 s contre 4 min 4 s pour la requête 1. Dans le détail, le temps de calcul pour la requête 2 est souvent le double de celui observé pour la requête 1.

Concernant les requêtes qui impliquent la fonction CT_{fh} (i.e., requêtes 3 et 4), nous remarquons également une augmentation du temps de calcul avec le nombre de localisations spatiales (cf figure 5.12). Cette augmentation est linéaire pour la requête 3 alors qu'elle augmente plus fortement entre 150 x 10⁶

et 350×10^6 localisations spatiales pour la requête 4. Pour un nombre d'objets égal à 15150 (courbes orange) la requête 4 réclame 10 fois plus de temps que la requête 3, et environ 4 fois plus que la requête 2 qui implique pourtant, elle aussi, les relations spatiales. En comparant les temps de calcul des requêtes 1 et 2 avec ceux des requêtes 3 et 4, on peut dire que la fonction Ct_{fh} est plus coûteuse que la fonction Ft_c . En faisant varier le nombre d'entités (i.e., les enregistrements de la table $CState$) dans la requête 4, nous pouvons clairement noter que plus le nombre d'objets interrogés augmente, moins l'évolution du graphe est linéaire.

D'après ce test d'évaluation concernant l'évolution de performance de calcul pour des requêtes hybrides, nous pouvons en déduire deux principales conclusions :

- L'utilisation des fonctions SQL pour déterminer les changements des valeurs d'attributs des objets, ou pour le calcul de la variation des valeurs dans une représentation de type champ, est bien gérée par le système. Cependant, l'appel à des fonctions traitant des relations spatiales et spatio-temporelles entre les objets a un impact important et augmente les temps de calcul. Afin de réduire cet effet, nous suggérons de développer un sous-modèle qui pourrait stocker explicitement toutes les relations spatiales et temporelles et utiliser des index de base de données pour rendre la réponse à ces requêtes beaucoup plus rapide. Bien entendu, la taille du stockage en base de données augmentera significativement, en revanche, le temps de réponse aux requêtes sera considérablement amélioré.
- La façon dont le schéma représentant le modèle champ est conçu dans la base de données et l'utilisation du partitionnement temporel pour la table *FieldState*, permettent aux sous-requêtes qui calculent la variation temporelle du champ avec des fonctions hybrides d'être relativement efficaces. Cependant, les performances sont limitées en particulier pour les bases de données spatio-temporelles massives de plus de 1TB relatives à de grandes échelles géographiques avec des données raster de haute résolution. La solution proposée dans les bases de données géographiques est un partitionnement spatial et spatio-temporel. Malheureusement, PostgreSQL n'offre pas de stratégie efficace pour y parvenir, alors que d'autres systèmes de données comme GeoMesa ([Hughes et al., 2015](#)), GeoSpark ([Yu et al., 2015](#)), SpatialSpark ([You et al., 2015](#)), qui soutiennent cette stratégie proposent des solutions alternatives de mise en œuvre pour gérer une telle quantité de données. Il sera donc intéressant de mettre en œuvre ce modèle hybride en lien avec l'une de ses solutions et de voir comment l'ensemble

du système réagit à chacune des requêtes hybrides.

5.7 Conclusion et synthèse

Dans ce chapitre, nous avons introduit une méthodologie expérimentale qui nous donne la possibilité d'évaluer notre modèle sur deux aspects : 1) l'expressivité en termes de requêtes spatio-temporelles et 2) les performances de calcul.

En ce qui concerne le premier aspect, six requêtes hybrides ont été élaborées pour répondre à des questionnements en relation avec la dynamique des morphologies sous-marines. Cette thématique est considérée comme un contexte d'application valide pour une expérimentation des capacités de représentation et de traitement de notre approche duale de modélisation. Les interrogations permettent d'explorer l'interaction entre la dynamique du fond marin en tant que surfaces continues et la dynamique des dunes sous-marines en tant d'objets géographiques composant l'espace. Une base de données spatio-temporelles a été construite et mise en oeuvre en utilisant des données réelles. Les requêtes hybrides ont été formulées avec le langage SQL étendu par les opérateurs et fonctions que nous avons définis. Les résultats montrent la capacité de mise en oeuvre de notre modèle, son interfaçage avec les bases de données géographiques et la validité des réponses obtenues sous la forme de représentations de type cartographique.

Quant au deuxième aspect, nous avons conçu un test expérimental pour un modèle de données hybride. L'objectif est d'examiner comment le modèle de données réagit en termes de réponse à différentes requêtes hybrides et à l'augmentation de la quantité de données insérées dans le système. Ce test montre que les types de requêtes hybrides n'ont pas les mêmes réponses au fil du temps et n'évoluent pas de la même manière à mesure que le volume de données augmente. De même, les tests montrent que les requêtes qui traitent de relations entre les entités sont plus coûteuses. Enfin, ces tests permettent d'identifier les parties du modèle de données qui doivent être mise à l'échelle et optimisée en identifiant des requêtes très coûteuses.

Conclusion générale

Synthèse de la contribution

L'objectif de cette recherche a consisté à développer une représentation hybride de l'espace géographique qui prenne en compte une intégration fine de la dimension temporelle soutenue par un modèle de données spatio-temporel permettant de spécifier un large éventail de requêtes qualifiées d'hybrides. Il s'agissait également d'explorer et de valider dans quelle mesure une représentation hybride pourrait d'une part fournir une modélisation applicable et sémantiquement riche, et d'autre part permettre la définition de mécanismes de manipulation au niveau d'un langage d'interrogation pour l'étude des phénomènes géographiques dits complexes ou à multiples composantes.

Dans cette optique, les travaux développés ont donné lieu à un ensemble de contributions qui apportent des réponses aux questionnements scientifiques identifiés dans notre introduction :

Un cadre formel

Tout d'abord, en ce qui concerne la forme d'une représentation hybride, nous avons choisi un modèle mathématique formel abstrait, que nous avons illustré par des exemples représentatifs afin de demeurer aussi générique que possible tout en gardant une indépendance totale de toute considération d'implémentation. Les différents exemples proposés, concernant la dynamique des fonds marins littoraux et des dunes associées, démontrent le potentiel de notre modélisation et participent également à la validation de nos concepts.

Le développement d'un tel modèle nous a permis de produire un ensemble de contributions que nous résumons dans les apports suivants :

-
- Nous proposons une formalisation générique duale intégrant les composantes d’une modélisation de type ”champ” (vision continue de l’espace) et d’une modélisation de type ”objet” (vision discrète de l’espace). Le concept de champ spatio-temporel permet de formaliser un ensemble de propriétés spatialement distribuées. Ce même formalisme intègre et modélise conceptuellement la vision objet et la dichotomie entre les notions de continuant et d’occurent.
Enfin, avec ce formalisme, nous établissons une association entre les deux modèles champ et objet. Pour les continuants, cette association se traduit par une forme de relation spatiale reliant les états de ces continuants à la composante spatiale d’un champ (i.e., L). En revanche, pour les occurrents, cette association se traduit par une forme de relation spatio-temporelle reliant chaque occurrent au domaine spatio-temporel d’un champ (i.e., $L \times T$).
 - Nous proposons une dérivation d’un ensemble de concepts hybrides résultant de l’interconnexion entre les visions discrète et continue de l’espace, et ce à travers le temps. Il s’agit principalement 1) d’explorer le potentiel impact de la dynamique d’un champ sur l’évolution des objets (i.e., champ vers objet), 2) d’explorer le potentiel impact de l’évolution d’un ou plusieurs objets sur la dynamique d’un champ (i.e., objet vers champ) et 3) d’explorer la variation interne d’un champ au sein d’un objet au cours de l’évolution de ce dernier.
 - Nous formalisons une série de fonctions avancées (Ft_c , Ft_o , Ct_{fs} , Ct_{fh} et Ct_{fr}) qui représentent des associations bidirectionnelles entre des objets et des champ. Ces fonctions, appelées hybrides, vont au-delà d’une simple exploration de la distribution interne d’un champ dans un objet donné, et manipulent également l’ensemble des interconnexions entre les champs et les objets à travers le temps. Cela nous permet d’améliorer les capacités de conceptualisation et de manipulation d’une représentation spatio-temporelle hybride, en particulier pour examiner les relations étroites résultantes de l’évolution des objets et des champs.
 - Nous mettons en place un formalisme pour enrichir la représentation spatio-temporelle hybride avec un cadre de référence interne. Il s’agit d’explorer et de caractériser la variation spatio-temporelle d’un champ au sein d’un objet lorsque ce dernier se déplace ou se déforme. Nous définissons ce cadre de référence interne permettant de localiser de manière unique un point au sein d’une emprise spatiale d’un objet
-

en mouvement. Par ailleurs, nous montrons comment ce cadre interne peut être exploité formellement par une fonction hybride afin de pouvoir suivre la variation des champs successifs d'un continuant et de ses emprises spatiales successives.

- Nous classifions les fonctions hybrides en référence au travail de [Sinton \(1978\)](#) qui catégorise les requêtes en fonction de l'emplacement, du temps et de la sémantique. Ce schéma génère cinq catégories de requêtes hybrides permettant de typer les manipulations de champs et d'objets, vus comme des possibilités de manipulation complémentaires pour des requêtes spatio-temporelles.

Un modèle

Nous proposons d'associer le cadre formel à un modèle de données spatio-temporel hybride implémentable. Ce choix est motivé par le besoin d'étendre un outil de SIG temporel par un ensemble de mécanismes, une structure de données et un ensemble d'opérations capables de répondre à des requêtes qualifiées d'hybrides. A ce niveau les contributions peuvent être mentionnées par les apports suivants :

- Nous avons employé quatre mécanismes de modélisation conceptuelle de données. Le premier est un mécanisme de généralisation avec lequel nous avons surmonté le problème de dichotomie continuant/occurent en matérialisant l'association entre une entité géographique, un continuant et un occurrent. Le second est un mécanisme d'agrégation qui joue un rôle clé dans notre approche de modélisation spatio-temporelle hybride dans la mesure où il permet de matérialiser formellement l'association entre la vision d'un champ et celle d'un objet. Quant au troisième mécanisme, celui de composition, il nous permet d'établir une relation entre un continuant et ses états spatio-temporels afin de gérer la variation de ses attributs dans le temps. Enfin, nous déployons un quatrième mécanisme de spécialisation qui sert à spécialiser tout type d'entités présent dans le modèle hybride. Par exemple, la classe *FieldState* (figure 4.2), qui représente le concept de champ, peut être spécialisée en deux sous-classes, l'une dénoterait les conditions atmosphériques (ex. température, vitesse du vent) et l'autre représenterait les conditions hydrodynamiques (ex. vitesse du courant, pression, etc.).
- Au niveau logique, nous définissons trois nouveaux types de données

qui sont dérivés du modèle conceptuel hybride. Ces types de données sont les suivants : *Ctn*, *Ocr* et *LocSet*. À travers ces trois types de données, nous décrivons six fonctions hybrides. Ces dernières ont des caractéristiques spécifiques et peuvent être appelées au niveau de la manipulation des données car elles matérialisent l'ensemble des relations établies entre les champs et les objets.

- Nous proposons une architecture logicielle ainsi qu'un schéma de base de données objet-relationnel matérialisant une représentation logique qui supporte un modèle de données hybride.

L'implémentation et l'évaluation des performances

Les éléments du modèle et la base de données associée sont entièrement implémentés au sein de PostgreSQL et de son extension spatiale PostGIS. Les fonctions et les opérations sont intégrées au sein du langage SQL pour prendre en charge les requêtes hybrides. Les contributions sont les suivantes :

- Nous validons la mise en oeuvre de la base de données géographiques avec des données réelles issues de levés bathymétriques. Le cas d'étude concerne l'évolution de dunes sous-marines avec d'une part des objets spatiaux qui se déplacent, changent de forme et surtout modifient leurs relations spatiales, et d'autre part le relief sous-marin dont les profondeurs et les pentes évoluent dans le temps.
- Nous implémentons les 6 types de fonctions hybrides et nous produisons plusieurs requêtes complexes interrogeant à la fois l'information spatiale continue et discrète. Les résultats sont exprimés sous la forme de représentations de type cartographique qui décrivent et permettent de quantifier la dynamique du fond marin et des dunes sous-marines associées.
- Nous concevons un test expérimental pour un modèle de données hybride. Il consiste à évaluer la réponse en temps de calcul de différents types de requêtes hybrides. L'objectif est 1) de montrer que les types de requêtes hybrides n'ont pas les mêmes réponses en temps de calcul et n'évoluent pas de la même manière lorsque la taille des données augmente, en particulier lorsque nous traitons le mouvement des objets ou les changements de relation entre les objets ; 2) d'identifier la partie du modèle de données qui doit être mise à l'échelle et optimisée en identifiant les requêtes à coût élevé ; et 3)

de permettre aux futurs modèles de données hybrides de comparer leurs performances de calcul avec les résultats de notre propre modèle.

Discussion et analyse critique des réponses apportées aux questionnements scientifiques

Bien que la question principale, qui consiste à mener des recherches sur le raisonnement spatial et sur une forme de représentation de l'espace et du temps avec une vision hybride, pourrait apparaître en première impression comme abstraite et éloignée des besoins pratiques concernant l'utilisation de l'information géographique, elle conduit, en fin de compte, à lever des verrous conceptuels et opérationnels nécessaires que nous pouvons résumer comme suit :

- Du point de vue conceptuel, le fait d'aborder le questionnement de la modélisation hybride ouvre la porte à de nouvelles réflexions concernant l'appréhension de phénomènes géographiques complexes, et la mise en évidence de nouveaux types d'interrogations spatio-temporelles. Celles-ci concernent en particulier l'interrelation entre phénomènes continus et phénomènes discrets.
Plus précisément, la recherche menée dans cette thèse recèle et met en valeur plusieurs aspects. Le premier doit permettre d'explorer le potentiel effet de la dynamique d'un champ sur la dynamique d'un objet. Inversement, le deuxième aspect vise à explorer le potentiel impact d'une dynamique des objets sur la dynamique d'un champ. Enfin, le troisième aspect concerne l'évolution dans le temps du champ interne inclu dans l'emprise spatio-temporelle d'un objet en mouvement. Ces aspects sont utiles car ils améliorent la capacité à décrire des phénomènes géographiques complexes localement et globalement, et aident à mieux comprendre les relations entre différents types de processus spatio-temporels.
- Du point de vue plus technique et opérationnel, les travaux réalisés ont abouti à un modèle de données, et de nouvelles structures de données, offrant un plus grand potentiel pour la manipulation de données et de requêtes spatio-temporelles. Ce modèle a été conçu et développé de telle sorte qu'il puisse être extensible et utilisé principalement pour la mise en oeuvre d'une nouvelle génération

de SIG temporel capable de supporter les données relatives aux évolutions des objets en conjonction avec d'autres sources de données qui capturent la dynamique des propriétés continues en prenant en compte l'interaction entre les deux formes de dynamiques. Cette dernière perspective de développement pourrait tirer partie des avancées considérables des technologies modernes de l'information et pourrait permettre de mieux gérer, interroger et analyser des données spatio-temporelles pour des phénomènes complexes.

Le modèle présenté dans notre recherche a apporté des réponses ainsi que des solutions conceptuelles et logiques pour le développement d'un modèle hybride de l'espace et du temps. Cependant, la complexité d'un tel sujet reste néanmoins importante et laisse subsister de nombreux verrous scientifiques :

- Le modèle ne gère pas de façon satisfaisante le changement d'un niveau de granularité spatio-temporelle à un autre, cela restant une problématique complexe tant conceptuellement, formellement et logiquement. A titre d'exemple, nous pouvons mettre en parallèle la temporalité des phénomènes tempétueux et celle de l'évolution du mouvement des dunes sous-marines. En effet les tempêtes ont une dynamique rapide (heures ou jours), alors que le déplacement des dunes est un phénomène qui doit être perçu que dans une temporalité plus étendue (années) car elle dépend aussi de phénomènes plus lents telle que la dynamique des courants marins par exemple. De même, le modèle est pensé pour que les enregistrements des localisations spatiales et de l'état des entités soient synchrones.
- Actuellement, le modèle ne fait pas de différence explicite entre les notions d'événement et de processus et leurs relations avec la notion de champ, de continuant et d'occurrent ainsi qu'avec la vision hybride en général. Cette préoccupation de recherche n'est pas nouvelle. Une spécification plus précise de ces deux notions à un premier niveau conceptuel, puis de ses implications formelles et logiques, pourrait également impacter les perspectives de développement et de mise en oeuvre de notre approche de modélisation.
- La question de l'identité n'a pas été abordée ; elle est simplement impliquée dans le modèle du continuant et a été modélisée comme un attribut défini par l'utilisateur. Par conséquent, le cadre de référence interne proposé ne traite pas encore le cas où un objet se fragmente

ou lorsqu'un objet se fusionne avec un autre, et le modèle, en général, n'offre pas de moyens qui permettent de soutenir une intégration avancée de cette notion.

- La définition de la fonction Ctf_h n'est pas assez générique, car elle opère sur des champs spatio-temporels dont les localisations spatiales sont fixes à travers le temps. L'ambition est de pouvoir généraliser la définition de cette fonction afin qu'elle puisse englober également les champs dont la répartition de leur échantillon des localisations spatiales (i.e., leur domaine spatial) n'est pas la même à chaque instant.
- Les tests de performances ont montré que le modèle hybride aboutit à des résultats relativement satisfaisants. Néanmoins, en ce qui concerne les requêtes hybrides qui explorent les changements dans les relations spatiales entre les objets, le modèle de données et le système éprouvent de grandes difficultés à répondre rapidement à ces requêtes. Cela est lié au fait que cette opération est déléguée à la fonction *metricRdNeighbor* qui évalue les relations pendant la phase d'exécution de la requête.

Au final, il convient de rappeler qu'un des avantages principaux du modèle est qu'il est pleinement extensible, de sorte que de nouvelles structures de données peuvent être intégrées et que des fonctions hybrides peuvent être redéfinies pour incorporer de nouveaux paramètres ainsi que des routines.

Extensions de recherche et perspectives applicatives

Plusieurs extensions de recherche peuvent être proposées afin de compléter ce travail de thèse avec les orientations suivantes :

- Généraliser notre approche afin qu'elle puisse gérer des requêtes spatio-temporelles hybrides à différents niveaux de granularité spatio-temporelle. Cela imposera, au niveau conceptuel, d'identifier comment les propriétés spatiales et les relations entre les objets et les champs peuvent être tracées. Il est clair que si le modèle de données hybrides vise à pouvoir prendre en charge des requêtes hybrides spatio-temporelles à plusieurs échelles, il doit opter pour

le déploiement de stratégies d'interpolation spatio-temporelle. Ces stratégies permettront non seulement de traiter des requêtes hybrides portant sur un phénomène géographique présentant différents niveaux de granularité, mais aussi de combiner plusieurs phénomènes géographiques complexes qui chacun présente différents niveaux de granularité.

- Élargir la définition de la fonction hybride Ct_{fh} de manière à ce qu'elle ne concerne pas le seul cas où les localisations spatiales qui décrivent l'évolution d'un champ continu soient fixes au travers de la succession des enregistrements de ce champ. Cela ouvre la porte à d'autres types de manipulations et d'analyses algébriques locales, zonales et globales du champ spatio-temporel. Cela pourrait être utile dans le cas où différents échantillons de données décrivant la dynamique de phénomènes géographiques pour une même étendue spatiale, seraient disponibles pour différents instants.
- Élargir les définitions et les concepts avec la gestion de l'incertitude concernant les propriétés spatio-temporelles (et les relations) des champs et des objets. Étudier l'implication de cet élargissement sur les fonctions hybrides, et par conséquent sur les capacités d'interrogation spatio-temporelle du modèle de données.
- Proposer une architecture de gestion des données massives non centralisée pour implémenter notre modèle hybride. On peut par exemple proposer une distribution du stockage et de la gestion des données au sein des différents types de SGBD, où chaque type de structure de données sera géré par le SGBD le plus approprié. Dans ce cas, les requêtes seraient gérées par un "framework" offrant un accès commun à ces diverses sources de données, ainsi que la possibilité de joindre des données entre ces sources. Le but d'une telle architecture consisterait à améliorer les performances de calcul des requêtes hybrides en optant pour le calcul distribué et parallèle. Ces techniques sont déjà implémentées et employées par les plateformes *Big Data* en termes de gestion et traitement des données spatiales. Il s'agit de faire évoluer le système lorsque la quantité de données devient de plus en plus importante et clairssemée. De plus, une telle architecture permet de déléguer certains traitements, dont ceux liés aux changements des relations spatiales et spatio-temporelles. Dans ce cas, l'idée est d'utiliser un environnement ou des outils plus adaptés, en l'occurrence un environnement plus optimal en termes de

traitement et d'analyse des représentations de graphes.

- Élargir le champ d'applications de notre approche hybride à d'autres études de cas et avec d'autres outils et modèles d'analyse spatio-temporelle. Nous avons en effet traité uniquement des objets polygonaux et le cas de variabilités géomorphologiques. L'extension au suivi d'objets de type points ou polygones et/ou en 3D est à entrevoir, de même que l'analyse de dynamiques urbaines, météorologiques ou d'objets mobiles. On peut évoquer en particulier le suivi de contaminations épidémiologiques et de pollutions atmosphériques en lien avec le déplacement d'individus ou de véhicules.

Bibliographie

- Allen, J. F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11) :832–843.
- Allen, J. F. and Ferguson, G. (1994). Actions and events in interval temporal logic. *Journal of logic and computation*, 4(5) :531–579.
- Arlow, J. and Neustadt, I. (2002). Uml and the unified process : practical object-oriented analysis and design. 2002. *Sanchez, " Resolution of composite fuzzy relation equations" Inform and Control*, 30 :38–48.
- Armstrong, M. P. (1988). Temporality in spatial databases.
- Astrahan, M. M., Blasgen, M. W., Chamberlin, D. D., Eswaran, K. P., Gray, J. N., Griffiths, P. P., King, W. F., Lorie, R. A., McJones, P. R., Mehl, J. W., et al. (1976). System r : relational approach to database management. *ACM Transactions on Database Systems (TODS)*, 1(2) :97–137.
- Baumann, R., Loebe, F., and Herre, H. (2012). Ontology of time in gfo proceedings of the 7th international conference on formal ontology in information systems. FOIS.
- Berman, R. R. and Stonebraker, M. (1977). Geo-ouel : a system for the manipulation and display of geographic data. In *Proceedings of the 4th annual conference on Computer graphics and interactive techniques*, pages 186–191.
- Booch, G. (1999). Uml in action. *Communications of the ACM*, 42(10) :26–28.
- Burrough, P. A. (1986). Principles of geographical. *Information Systems for Land Resource Assessment*. Clarendon Press, Oxford.
- Burrough, P. A. and Frank, A. (1996). *Geographic objects with indeterminate boundaries*, volume 2. CRC Press.

- Câmara, G., Souza, R., Pedrosa, B., Vinhas, L., Monteiro, A. M. V., Paiva, J., Carvalho, M. d., and Gattass, M. (2000). Terralib : Technology in support of gis innovation. In *II Workshop Brasileiro de Geoinformática, GeoInfo2000*, volume 2, pages 1–8. São Paulo.
- Câmara, G., Vinhas, L., Ferreira, K. R., De Queiroz, G. R., De Souza, R. C. M., Monteiro, A. M. V., De Carvalho, M. T., Casanova, M. A., and De Freitas, U. M. (2008). Terralib : An open source gis library for large-scale environmental and socio-economic applications. In *Open source approaches in spatial data handling*, pages 247–270. Springer.
- Carré, C. (2020). *Conception d’une ontologie spatio-temporelle dédiée au suivi de l’évolution des objets géographiques*. PhD thesis.
- Chang, S. K. and Fu, K. S. (1980). *Pictorial information systems*. Springer.
- Chen, P. P.-S. (1976). The entity-relationship model—toward a unified view of data. *ACM transactions on database systems (TODS)*, 1(1) :9–36.
- Chrisman, N. (1978). Concepts of space as a guide to cartographic data structures. *Harvard papers on geographic information systems*, 8 :1–17.
- Claramunt, C., de Sède, M., Prelaz-Droux, R., and Vidale, L. (1993). Integration of heterogeneous gis for environmental planning and management, gis technology and applications. In *World Scientific Pub., Singapour 1993, Hongjun Lu and Beng Chin Ooi (eds.)*, pages 404–415.
- Claramunt, C. and Mainguenaud, M. (1995). Dynamic and flexible vision of a spatial database. In *Proceedings of the DEXA’95 Int. Workshop on Database and Expert System Applications*, pages 483–492. Omnipress, London, United-Kingdom.
- Claramunt, C. and Parent, C. (2003). Modelling concepts for the representation of evolution constraints. *Computers, environment and urban systems*, 27(3) :225–241.
- Claramunt, C., Parent, C., and Thériault, M. (1997). Design patterns for spatio-temporal processes. searching for semantics : Data mining, reverse engineering, s. spaccapieta and f. marganski, eds.
- Claramunt, C. and Stewart, K. (2015). Special issue on spatio-temporal theories and models for environmental, urban and social sciences : where do we stand ?

- Claramunt, C. and Thériault, M. (1995). Managing time in gis an event-oriented approach. In *Recent advances in temporal databases*, pages 23–42. Springer.
- Claramunt, C. and Theriault, M. (1996). Toward semantics for modelling spatio-temporal processes within gis. *Advances in GIs Research I*, pages 27–43.
- Codd, E. (1970). A relational model of data for large relational databases. *Communications of the ACM*, 13(6) :77–87.
- Cole, J. P. and King, C. A. (1968). Quantitative geography : Techniques and theories in geography. Technical report.
- Couclelis, H. (1992). People manipulate objects (but cultivate fields) : beyond the raster-vector debate in gis. In *Theories and methods of spatio-temporal reasoning in geographic space*, pages 65–77. Springer.
- Cova, Thomas J., G. M. F. (2002). Extending geographical representation to include fields of spatial objects. *International Journal of geographical information science*, 16(6) :509–532.
- Cova, T. J. (2016). Data Model, F-Objects and O-Fields. *International Encyclopedia of Geography : People, the Earth, Environment and Technology*, pages 1–5.
- Date, C. J. (2011). *SQL and relational theory : how to write accurate SQL code*. ” O’Reilly Media, Inc.”.
- Dayal, U. and Smith, J. M. (1986). Probe : A knowledge-oriented database management system. In *On knowledge base management systems*, pages 227–257. Springer.
- De Bakker, Merijn P., D. J. K. S. O. e. a. (2017). Design and demonstration of a data model to integrate agent-based and field-based modelling. 89 :172–189.
- Del Mondo, G. (2011). *Un modèle de graphe spatio-temporel pour représenter l’évolution d’entités géographiques*. PhD thesis.
- Del Mondo, G., Stell, J. G., Claramunt, C., and Thibaud, R. (2010). A graph model for spatio-temporal evolution. *J. UCS*, 16(11) :1452–1477.

- Egenhofer, M. J., Frank, A. U., and Jackson, J. P. (1989). A topological data model for spatial databases. In *Symposium on Large Spatial Databases*, pages 271–286. Springer.
- Egenhofer, M. J. and Herring, J. (1990). Categorizing binary topological relations between regions, lines, and points in geographic databases. *The*, 9(94-1) :76.
- Elmasri, R. (2008). *Fundamentals of database systems*. Pearson Education India.
- Elmongui, H. G., Mokbel, M. F., and Aref, W. G. (2013). Continuous aggregate nearest neighbor queries. *GeoInformatica*, 17(1) :63–95.
- Ferreira, K. R., Camara, G., and Monteiro, A. M. V. (2014). An algebra for spatiotemporal data : From observations to events. *Transactions in GIS*, 18(2) :253–269.
- Forlizzi, L., Güting, R. H., Nardelli, E., and Schneider, M. (2000). A data model and data structures for moving objects databases. In Chen, W., Naughton, J. F., and Bernstein, P. A., editors, *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, May 16-18, 2000, Dallas, Texas, USA*, pages 319–330. ACM.
- Frank, A. (1982). Mapquery : Data base query language for retrieval of geometric data and their graphical representation. *ACM SIGGRAPH Computer Graphics*, 16(3) :199–207.
- Frank, A. U. (1996). Qualitative spatial reasoning : Cardinal directions as an example. *International Journal of Geographical Information Science*, 10(3) :269–290.
- Franklin, W. R. (1973). Triangulated irregular network program.
- Franzetti, M., Le Roy, P., Delacourt, C., Garlan, T., Cancouët, R., Sukhovich, A., and Deschamps, A. (2013). Giant dune morphologies and dynamics in a deep continental shelf environment : example of the banc du four (western brittany, france). *Marine Geology*, 346 :17–30.
- Galton, A. (1997). Space, time, and movement. In *Spatial and Temporal Reasoning*, pages 321–352. Springer.
- Galton, A. (2001a). A formal theory of objects and fields. In *International Conference on Spatial Information Theory : Lecture Notes in Computer Science*, volume 2205, pages 458–473. Springer, Berlin, Heidelberg.

- Galton, A. (2001b). Space, time, and the representation of geographical reality. *Topoi*, 20(2) :173–187.
- Galton, A. (2003). Desiderata for a spatio-temporal geo-ontology. In *International Conference on Spatial Information Theory*, pages 1–12. Springer.
- Galton, A. (2004). Fields and objects in space, time, and space-time. *Spatial cognition and computation*, 4(1) :39–68.
- Galton, A. (2006). Processes as continuants. In *Thirteenth International Symposium on Temporal Representation and Reasoning (TIME'06)*, pages 187–187. IEEE.
- Galton, A. (2016). The ontology of time and process. *Third Interdisciplinary School on Applied Ontology, Bozen-Bolzano, Italy*.
- Galton, A. and Worboys, M. (2005). Processes and events in dynamic geo-networks. In *International Conference on GeoSpatial Semantics*, pages 45–59. Springer.
- Garlan, T., Cartier, A., Franzetti, M., Le Roy, P., Duarte, J., Pombo, J., Peix, M., Guyomard, P., Le Faou, Y., Gabelotaud, I., et al. (2013). Complex morphology and organisation of dunes in a giant dunes field. In *Proc. MARID 2013*, pages 113–118.
- Gebbert, S. and Pebesma, E. (2014). A temporal gis for field based environmental modeling. *Environmental Modelling & Software*, 53 :1–12.
- Gebbert, S. and Pebesma, E. (2017). The grass gis temporal framework. *International Journal of Geographical Information Science*, 31(7) :1273–1292.
- Gold, C. M. and Condal, A. R. (1995). A spatial data structure integrating gis and simulation in a marine environment. *Marine Geodesy*, 18(3) :213–228.
- Goodchild, M. F. (1989). Modeling error in objects and fields. *Accuracy of spatial databases*, 113.
- Goodchild, M. F. (2001). A geographer looks at spatial information theory. In *International Conference on Spatial Information Theory*, pages 1–13. Springer.

- Goodchild, M. F. (2003). Geographic information science and systems for environmental management. *Annual Review of Environment and Resources*, 28.
- Goodchild, M. F. (2004). Giscience, geography, form, and process. *Annals of the Association of American Geographers*, 94(4) :709–714.
- Goodchild, M. F. and Yuan, May, C. T. J. (2007). Towards a general theory of geographic representation in gis. *International journal of geographical information science*, 21(3) :239–260.
- Grenon, P. and Smith, B. (2004). Snap and span : Towards dynamic spatial ontology. *Spatial cognition and computation*, 4(1) :69–104.
- Güting, R. H. (1994). An introduction to spatial database systems. *the VLDB Journal*, 3(4) :357–399.
- Güting, R. H., Böhlen, M. H., Erwig, M., Jensen, C. S., Lorentzos, N. A., Schneider, M., and Vazirgiannis, M. (2000). A foundation for representing and querying moving objects. *ACM Transactions on Database Systems (TODS)*, 25(1) :1–42.
- Guttag, J. V. and Horning, J. J. (1978). The algebraic specification of abstract data types. *Acta informatica*, 10(1) :27–52.
- Guttman, A. (1984). R-trees : A dynamic index structure for spatial searching. In *Proceedings of the 1984 ACM SIGMOD international conference on Management of data*, pages 47–57.
- Hamdani, Y., Thibaud, R., and Claramunt, C. (2019). A hybrid temporal gis representation for coastal dynamics. In *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 584–587.
- Heinz, F. and Güting, R. H. (2016). Robust high-quality interpolation of regions to moving regions. *GeoInformatica*, 20(3) :385–413.
- Heinz, F. and Güting, R. H. (2018). A data model for moving regions of fixed shape in databases. *International Journal of Geographical Information Science*, 32(9) :1737–1769.
- Heinz, F. and Güting, R. H. (2020). A polyhedra-based model for moving regions in databases. *International Journal of Geographical Information Science*, 34(1) :41–73.

- Homrani, S., Le Dantec, N., Floc'h, F., Franzetti, M., Delacourt, C., Sedrati, M., and Winter, C. (2019). Multi time-scale morphological evolution of a shell sand, dune bank in a shallow mesotidal environment. *MARID VI*, page 121.
- Hornsby, K. and Egenhofer, M. J. (1997). Qualitative representation of change. In *International conference on spatial information theory*, pages 15–33. Springer.
- Hornsby, K. and Egenhofer, M. J. (2000). Identity-based change : a foundation for spatio-temporal knowledge representation. *International journal of geographical information science*, 14(3) :207–224.
- Hornsby, K. S. and Cole, S. (2007). Modeling moving geospatial objects from an event-based perspective. *Transactions in GIS*, 11(4) :555–573.
- Hornsby, K. S. and Yuan, M. (2008). *Understanding dynamics of geographic domains*. CRC Press.
- Hughes, J. N., Annex, A., Eichelberger, C. N., Fox, A., Hulbert, A., and Ronquest, M. (2015). Geomesa : a distributed architecture for spatio-temporal fusion. In *Geospatial Informatics, Fusion, and Motion Video Analytics V*, volume 9473, page 94730F. International Society for Optics and Photonics.
- Ibrahim, Ahmed, U. T. and Kraak, M.-J. (2014). Representing internal varying characteristics of moving objects. In Springer, C., editor, *East European Conference on Advances in Databases and Information Systems*, pages 207–218. Springer, Cham.
- Ilarri, S., Mena, E., and Illarramendi, A. (2010). Location-dependent query processing : Where we are and where we are heading. *ACM computing surveys (CSUR)*, 42(3) :1–73.
- Jin, M., Claramunt, C., and Wang, T. (2017). A map-matching approach for travel behavior analysis. In *2017 4th International Conference on Systems and Informatics (ICSAI)*, pages 1405–1410. IEEE.
- Kim, K.-S. and Kiyoki, Y. (2010). An object-field perspective data model for moving geographic phenomena. In Springer, editor, *International Conference on Database Systems for Advanced Applications*, pages 410–421. Springer.

- Langran, G. and Chrisman, N. R. (1988). A framework for temporal geographic information. *Cartographica : The International Journal for Geographic Information and Geovisualization*, 25(3) :1–14.
- Lardon, S., Libourel, T., and Cheylan, J.-P. (1999). Concevoir la dynamique des entités spatio-temporelles. 9 :45–66.
- Le Bot, S. (2001). *Morphodynamique de dunes sous-marines sous influence des marées et des tempêtes : processus hydro-sédimentaires et enregistrement : exemple du Pas-de-Calais*. PhD thesis, Lille 1.
- Ledoux, H. and Gold, C. (2004). Modelling oceanographic data with the three-dimensional voronoi diagram. In *ISPRS 2004-XXth Congress, Istanbul, Turkey*, volume 2, pages 703–708.
- Liu, Y., Goodchild, M. F., Guo, Q., Tian, Y., and Wu, L. (2008). Towards a general field model and its order in GIS. *International Journal of Geographical Information Science*, 22(6) :623–643.
- Livingstone, I., Wiggs, G. F., and Weaver, C. M. (2007). Geomorphology of desert sand dunes : a review of recent progress. *Earth-science reviews*, 80(3-4) :239–257.
- Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W. (2005). *Geographic information systems and science*. John Wiley & Sons.
- Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W. (2015). *Geographic information science and systems*. John Wiley & Sons.
- Mark, D. M. et al. (2000). Geographic information science : Critical issues in an emerging cross-disciplinary research domain. *URISA-WASHINGTON DC-*, 12(1) :45–54.
- McIntosh, J. and Yuan, M. (2005). A framework to enhance semantic flexibility for analysis of distributed phenomena. *International Journal of Geographical Information Science*, 19(10) :999–1018.
- Mitas, L. and Mitasova, H. (1999). Spatial interpolation. *Geographical information systems : principles, techniques, management and applications*, 1(2).
- Mitasova, H., Hardin, E., STAREK, M. J., and al. (2011). Landscape dynamics from LiDAR data time series. *Geomorphometry*, pages 3–6.

- Noyon, V., Claramunt, C., and Devogele, T. (2007). A relative representation of trajectories in geographical spaces. *GeoInformatica*, 11(4) :479–496.
- Parent, C., Spaccapietra, S., and Zimányi, E. (2006). *Conceptual modeling for traditional and spatio-temporal applications : The MADS approach*. Springer Science & Business Media.
- Parent, C., Spaccapietra, S., Zimanyi, E., Donini, P., Plazanet, C., and Vangenot, C. (1998). Modeling spatial data in the mads conceptual model. In *Proc. of the 8th Int. Symp. on Spatial Data Handling, SDH'98*, number CONF.
- Parent, C., Spaccapietra, S., Zimányi, E., Donini, P., Plazanet, C., Vangenot, C., Rognon, N., Pouliot, J., and Crausaz, P.-A. (1997). Mads : un modèle conceptuel pour des applicationsspatio-temporelles. *Revue internationale de géomatique*, 7(ARTICLE) :317–352.
- Pelekis, N., Theodoulidis, B., Kopanakis, I., and Theodoridis, Y. (2004). Literature review of spatio-temporal database models. *The Knowledge Engineering Review*, 19(3) :235–274.
- Peuquet, D. J. (1984). A conceptual framework and comparison of spatial data models. *Cartographica : The International Journal for Geographic Information and Geovisualization*, 21(4) :66–113.
- Peuquet, D. J. (1994). It's about time : A conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of american Geographers*, 84(3) :441–461.
- Peuquet, D. J. (2001). Making space for time : Issues in space-time data representation. *GeoInformatica*, 5(1) :11–32.
- Peuquet, D. J. (2002). *Representations of space and time*. Guilford Press.
- Peuquet, D. J. and Duan, N. (1995). An event-based spatiotemporal data model (estdm) for temporal analysis of geographical data. *International journal of geographical information systems*, 9(1) :7–24.
- Plumejeaud, C., Gensel, J., Villanova-Oliver, M., Ben Rebah, M., and Vergnaud, G. (2009). Modélisation de hiérarchies territoriales multiples. In *Colloque International de Géomatique et d'Analyse Spatiale (SAGEO 2009)*, pages 25–27.

- Prager, S. D. and Barber, J. J. (2012). Modeling unobserved true position using multiple sources and information semantics. *International Journal of Geographical Information Science*, 26(1) :15–37.
- Price, R., Tryfona, N., and Jensen, C. S. (2002). Extending uml for space-and time-dependent applications. In *Advanced Topics in Database Research, Volume 1*, pages 342–366. Igi Global.
- Pullar, D. V. (1988). Towards the defection and use of topological relations among spatial objects. In *Proceeding of the third International Symposium on Spatial Data Handling, IGU, Colombus, 1988*.
- Pultar, E., Cova, T. J., Yuan, M., and Goodchild, M. F. (2010). EDGIS : a dynamic GIS based on space time points. *International Journal of Geographical Information Science*, 24(3) :329–346.
- Renolen, A. (1997). Modelling spatiotemporal information : The spatiotemporal object model. *Norwegian University of Science and Technology, Trondheim, Norway*, pages 1–22.
- Richardson, J. E. and Carey, M. J. (1987). Programming constructs for database system implementation in exodus. In *Proceedings of the 1987 ACM SIGMOD international conference on Management of data*, pages 208–219.
- Robertson, C., Nelson, T. A., Boots, B., and Wulder, M. A. (2007). Stamp : spatial-temporal analysis of moving polygons. *Journal of Geographical Systems*, 9(3) :207–227.
- Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorensen, W. E., et al. (1991). *Object-oriented modeling and design*, volume 199. Prentice-hall Englewood Cliffs, NJ.
- Sadahiro, Y. and Umemura, M. (2001). A computational approach for the analysis of changes in polygon distributions. *Journal of Geographical Systems*, 3(2) :137–154.
- Siabato, W., Claramunt, C., Ilarri, S., and Manso-Callejo, M. Á. (2018). A survey of modelling trends in temporal gis. *ACM Computing Surveys (CSUR)*, 51(2) :1–41.
- Sinton, D. (1978). The inherent structure of information as a constraint to analysis : Mapped thematic data as a case study. *Harvard Papers on Geographic Information Systems*.

- Smith, B. and Varzi, A. C. (1997). Fiat and bona fide boundaries : Towards an ontology of spatially extended objects. In *International Conference on Spatial Information Theory*, pages 103–119. Springer.
- Smith, B. and Varzi, A. C. (2000). Fiat and bona fide boundaries. *Philosophical and phenomenological research*, pages 401–420.
- Snodgrass, R. and Ahn, I. (1985). A taxonomy of time databases. *ACM Sigmod Record*, 14(4) :236–246.
- Spéry, L., Claramunt, C., and Libourel, T. (2001). A spatio-temporal model for the manipulation of lineage metadata. *Geoinformatica*, 5(1) :51–70.
- Sriti, M., Thibaud, R., and Claramunt, C. (2005). A fuzzy identity-based temporal gis for the analysis of geomorphometry changes. In *Journal on Data Semantics III*, pages 81–99. Springer.
- Stark, I. D. B. (1994). *Names and higher-order functions*. PhD thesis, Cite-seer.
- Stefanidis, A., Eickhorst, K., Agouris, P., and Partsinevelos, P. (2003). Modeling and comparing change using spatiotemporal helixes. In *Proceedings of the 11th ACM international symposium on Advances in geographic information systems*, pages 86–93.
- Stonebraker, M. (1983). Application of abstract data types and abstract indices to cad data bases. In *Pcoc. Annual Meeting Database Week*, pages 107–115.
- Stonebraker, M. and Rowe, L. A. (1986). The design of postgres. *ACM Sigmod Record*, 15(2) :340–355.
- Thériault, M. and Claramunt, C. (1999). La représentation du temps et des processus dans les sig : une nécessité pour la recherche interdisciplinaire. *Représentation de l'espace et du temps dans les SIG, Revue internationale de géomatique*, 9 :67–99.
- Thériault, M., Claramunt, C., and Villeneuve, P. Y. (1999). A spatio-temporal taxonomy for the representation of spatial set behaviours. In *International Workshop on Spatio-Temporal Database Management*, pages 1–18. Springer.
- Thibaud, R., Del Mondo, G., Garlan, T., Mascaret, A., and Carpentier, C. (2013). A spatio-temporal graph model for marine dune dynamics analysis and representation. *Transactions in GIS*, 17(5) :742–762.

- Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography*, 46(sup1) :234–240.
- Tryfona, N., Busborg, F., and Borch Christiansen, J. G. (1999). starer : a conceptual model for data warehouse design. In *Proceedings of the 2nd ACM international workshop on Data warehousing and OLAP*, pages 3–8.
- Tryfona, N. and Jensen, C. S. (1999). Conceptual data modeling for spatio-temporal applications. *Geoinformatica*, 3(3) :245–268.
- Tryfona, N. and Jensen, C. S. (2000). Using abstractions for spatio-temporal conceptual modeling. In *Proceedings of the 2000 ACM symposium on Applied computing-Volume 1*, pages 313–322.
- van Dijk, T. A. G. P. and Egberts, P. (2008). The variability of sand wave migration in the north sea. In *Marid 2008. Third international workshop on marine and river dune dynamics 1-3 April 2008, Leeds, UK*, pages 63–67. SHOM.
- Wachowicz, M. and Healey, R. G. (1994). Towards temporality in gis. *Innovations in GIS*, 1 :105–115.
- Wang, J., Duckham, M., and Worboys, M. (2016). A framework for models of movement in geographic space. *International Journal of Geographical Information Science*, 30(5) :970–992.
- Watson, D. E. (1992). Contouring : A guide to the analysis and display of spatial data, tarrytown, ny.
- Wentz, E. A., Peuquet, D. J., and Anderson, S. (2010). An ensemble approach to space–time interpolation. *International Journal of Geographical Information Science*, 24(9) :1309–1325.
- Wever, T. F. (2004). Bedforms and bedform migration : a data review. In *Proc. Marine and River Dune Dynamics*, pages 330–337.
- Whitrow, G. J. (1989). *Time in history : Views of time from prehistory to the present day*. Oxford University Press, USA.
- Worboys, M. and Hornsby, K. (2004). From objects to events : Gem, the geospatial event model. In *International Conference on Geographic Information Science*, pages 327–343. Springer.

- Worboys, M. F. (1994). Unifying the spatial and temporal components of geographical information. In *Information Symposium on Spatial data Handling*.
- Worboys, M. F. and Duckham, M. (2004). *GIS : a computing perspective*. CRC press.
- Xu, J. and Güting, R. H. (2013). A generic data model for moving objects. *GeoInformatica*, 17(1) :125–172.
- Yang, Y. and Claramunt, C. (2003). A process-oriented multi-representation of gradual changes. *Journal of Geographic Information and Decision Analysis*, 7(1) :1–13.
- You, S., Zhang, J., and Gruenwald, L. (2015). Large-scale spatial join query processing in cloud. In *2015 31st IEEE International Conference on Data Engineering Workshops*, pages 34–41. IEEE.
- Yu, J., Wu, J., and Sarwat, M. (2015). Geospark : A cluster computing framework for processing large-scale spatial data. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 1–4.
- Yuan, M. (1994). Wildfire conceptual modeling for building gis space-time models. In *proceedings of GIS/LIS*, volume 94, pages 860–869.
- Yuan, M. (1996). Temporal gis and spatio-temporal modeling. In *Proceedings of Third International Conference Workshop on Integrating GIS and Environment Modeling, Santa Fe, NM*, volume 33.
- Yuan, M. (1999). Use of a three-domain representation to enhance gis support for complex spatiotemporal queries. *Transactions in GIS*, 3(2) :137–159.
- Yuan, M. (2001). Representing complex geographic phenomena in gis. *Cartography and Geographic Information Science*, 28(2) :83–96.
- Yuan, M. and McIntosh, J. (2002). A typology of spatiotemporal information queries. In *Mining Spatio-Temporal Information Systems*, pages 63–81. Springer.
- Zeiler, M. (1999). *Modeling our world : the ESRI guide to geodatabase design*. ESRI, Inc.

- Zhao, Z., Shaw, S.-L., and Wang, D. (2015). As pace-time raster gis data model for spatiotemporal analysis of vegetation responses to a freeze event. *Transactions in GIS*, 19(1) :151–168.
- Zheni, D., Frihida, A., Ghezala, H. B., and Claramunt, C. (2009). A semantic approach for the modeling of trajectories in space and time. In *International Conference on Conceptual Modeling*, pages 347–356. Springer.
- Zimmermann, K. and Freksa, C. (1996). Qualitative spatial reasoning using orientation, distance, and path knowledge. *Applied intelligence*, 6(1) :49–58.

Titre : Un modèle spatio-temporel hybride de SIG temporel : application à la géomorphologie marine

Mots clefs : SIG temporel, modèle de données hybride, Bases de données spatio-temporelles, champ, objet, Dynamique, géomorphologie marine.

Grâce au développement permanent de la technologie des capteurs géo-localisés et des infrastructures d'acquisition de données, les chercheurs en sciences environnementales et urbaines disposent de nouvelles possibilités pour étudier les phénomènes géographiques à différentes échelles temporelles et spatiales. En effet, le large éventail de données géographiques désormais disponibles offre de multiples possibilités d'exploration des données spatio-temporelles pour une meilleure compréhension de la nature de la dynamique des processus environnementaux sous-jacents. Face à de telles opportunités, des défis majeurs se manifestent et soulèvent cependant de nombreuses questions méthodologiques liées à la conceptualisation et à la représentation numérique de l'espace-temps.

L'objectif de cette recherche consiste à développer une modélisation hybride de l'espace géographique qui réconcilie les approches champs et objet et qui prenne en compte une fine intégration de la dimension temporelle. Sur la base de ce modèle de données spatio-temporel hybride, notre approche spécifie un large éventail de requêtes spatio-temporelles

qualifiées d'hybrides et qui permet d'exploiter le potentiel du modèle développé. La finalité consiste à explorer et de valider dans quelle mesure une telle représentation hybride puisse fournir une modélisation faisable et sémantiquement riche, et d'autre part de permettre la définition de mécanismes de manipulation au niveau d'un langage d'interrogation pour l'étude de phénomènes géographiques complexes. Le contexte de validation proposé est celui d'un phénomène de dynamique de géomorphologie marine. Ce cas est considéré comme un exemple pertinent de phénomène environnemental pour lequel une représentation hybride a un sens. Une série d'expériences de performance et d'extensibilité sont finalement rapportées et montrent comment l'ensemble du modèle se comporte.

Title: Spatio-temporal hybrid model for temporal GIS: Application to marine geomorphology

Keywords: Temporal GIS, Hybrid data model, spatio-temporal databases, Field, Object, Dynamic, Marine geomorphology.

With the continuous development of geo-localized sensor technology and data acquisition infrastructures, environmental and urban science researchers have new opportunities to study geographic phenomena at different temporal and spatial scales. Indeed, the wide range of geographic data now available offers multiple possibilities for exploring spatio-temporal data for a better understanding of the nature of the dynamics of underlying environmental processes. In front of such opportunities, major challenges emerge and raise many methodological issues related to the conceptualization and digital representation of space-time.

The objective of this research is to develop a hybrid model of the geographical space that reconciles the field and object approaches and that takes into account a fine integration of the temporal dimension. On the basis of this hybrid spatio-temporal data model, our approach specifies a wide range of spatio-temporal queries qualified as hybrid and which allows exploiting the potential of the developed model.

The aim is to explore and validate to what extent such a hybrid representation can provide a feasible and semantically rich model, and on the other hand to allow the definition of manipulation mechanisms at the level of a query language for the study of complex geographic phenomena. The proposed validation context is that of a marine geomorphology dynamics phenomenon. This case is considered as a relevant example of an environmental phenomenon for which a hybrid representation makes sense. A series of performance and scalability experiments are finally reported and show how the whole model behaves.