



**HAL**  
open science

# QUELLE CONFIANCE ACCORDER AUX DONNÉES CMIP-6 POUR L'ESTIMATION DES PRÉCIPITATIONS ?

Jean Claude Bergès

► **To cite this version:**

Jean Claude Bergès. QUELLE CONFIANCE ACCORDER AUX DONNÉES CMIP-6 POUR L'ESTIMATION DES PRÉCIPITATIONS?. pp.1-6, 2022, Actes du 35e colloque annuel de l'Association Internationale de Climatologie ; Toulouse 6-9 juillet 2022. hal-03831651

**HAL Id: hal-03831651**

**<https://hal.science/hal-03831651>**

Submitted on 27 Oct 2022

**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.

# QUELLE CONFIANCE ACCORDER AUX DONNÉES CMIP-6 POUR L'ESTIMATION DES PRÉCIPITATIONS ?

**JEAN CLAUDE BERGÉS** <sup>1</sup>

<sup>1</sup>UMR PRODIG, Université Paris I Panthéon-Sorbonne, Bat. ODG, 75013 Paris,  
jean-claude.berges@univ-paris1.fr

**Résumé :** Cette étude porte sur la capacité des modèles globaux de circulation à restituer les apports pluviaux à une échelle régionale. À travers l'estimation de deux paramètres de distribution des précipitations, elle met en avant quelques difficultés rencontrées en utilisant cette source d'informations et propose une méthode permettant une amélioration sensible des résultats.

**Mots-Clés :** Précipitations, CMIP6, Erreurs d'estimation.

**Abstract :** Which accuracy of the CMIP6 precipitation forecast ? This study focuses on the quality of the regional precipitation forecast based on global circulation models. Analysing the estimation of two rainfall distribution parameters, the difficulties in merging models are underlined and an empirical approach is proposed to deal with models dispersion effects.

**Keywords :** Précipitation, CMIP6, Estimation errors.

## Introduction

Le CMIP (*Coupled Model Intercomparison Project*) est un programme initié par l'OMM pour soutenir les activités de l'IPCC (*Intergovernmental Panel on Climate Change*) en coordonnant les activités d'élaboration de modèles climatiques globaux et en organisant leur diffusion. Ces activités sont couplées avec les productions IPCC et actuellement le CMIP6 (Eyring *et al.*, 2016), sixième version de la base de données, contribue à fonder les conclusions de l'AR6 (*Assessment Report 6*), le plus récent rapport de cette institution.

La modification du régime de précipitations est un élément clé déterminant les impacts socio-économiques du changement climatique et le CMIP6 constitue la source de données la plus immédiatement accessible pour obtenir des informations quantitatives. Aussi la qualité de ces informations est une question d'autant plus importante que ces données constitueront des paramètres d'entrée de modèles agronomiques ou hydrologiques sensibles à divers paramètres de la distribution des précipitations.

## 1. Les enjeux de la prévision des précipitations

L'évaluation de la qualité des données de précipitations extraites du CMIP6 soulève deux types de difficultés. Les premières sont liées à l'objectif de ces modèles généraux de circulation (GCM) qui ne visent pas une information déterministe mais une prévision de l'évolution de la distribution des précipitations. Compte tenu de la longueur limitée des séries de données disponibles, éliminer l'effet de la variabilité naturelle du climat sur les paramètres de la distribution est un enjeu. Les autres difficultés proviennent des observations utilisées en référence. La pluie est un phénomène intermittent et, à toutes échelles, les intensités instantanées de précipitations se distribuent selon une loi de puissance. L'estimation quantitative des précipitations est donc une question complexe et les différents produits sont susceptibles de présenter des biais spécifiques.

Dans ce travail IMERG (*Integrated Multi satellite Retrievals for GPM*), le produit de synthèse du GPM (*Global Precipitation Mission*), sera considéré comme source de données de référence en contournant la question de ses biais d'estimation. Les paramètres de distribution des précipitations seront estimés sur la période 2015-2020. Cette période est définie pour des motifs contingents de disponibilité simultanée des données de référence et de modèles provenant des expériences de prévision. En effet, la base de données CMIP6 ne fournit pas de résultats de modèles en mode de prévision, excluant l'assimilation de données actuelles,

qu'à partir de 2015. Des motifs similaires de disponibilité de données ont conduit à définir la période 2045-2050 comme l'échéance de prévision à moyen terme.

Les mécanismes convectifs relèvent d'une échelle spatiale beaucoup plus fine que la grille de résolution des modèles de prévision climatique. Aussi ces derniers doivent introduire une représentation statistique de ces phénomènes à travers un processus de paramétrisation (Hourdin *et al.*, 2017) dont la pertinence jouera un rôle clé dans la qualité de restitution des précipitations.

## 2. Constitution du jeu de données et traitements préliminaires

Le portail de distribution du CMIP6 permet un accès facile aux données de modèles et une sélection rapide parmi les expériences. Les modèles sélectionnés correspondent à des expériences de prévision à moyen terme fournissant le flux de précipitations quotidien, les résolutions spatiales les plus fines étant privilégiées. Seuls les scénarios à plus haute émission, RCP8.5 et SSP5, ont été considérés. L'étude actuelle porte sur six modèles dont la dimension de grille varie de 25 km à 140 km : CMCC-CM2, CNRM-CM6, GFDL-ESM4, HiRAM-SIT, MRI-ESM2 et NCAR-CESM2. Certains modèles sont diffusés en plusieurs réalisations ou avec des conditions d'initialisation légèrement différentes ce qui permet, dans une certaine mesure, de prendre en compte la variabilité naturelle du climat. En pareil cas, une moyenne des diverses estimations dérivées du modèle a été effectuée.

Pendant la période considérée, les données de précipitations du produit GPM/IMERG (Huffman *et al.*, 2019) ne sont pas perturbées par des discontinuités majeures du dispositif d'observation. Dans sa version qualifiée de finale, ce produit, élaboré par la NASA, réalise une synthèse de diverses sources : micro-ondes passives provenant d'une constellation de satellites défilants, infrarouge thermique des géostationnaires et réseau de pluviomètres concentré par le GPCC (*Global Precipitation Climatology Center*). Ses algorithmes sont régulièrement remis à jour et ses données diffusées à l'échelle fine de la demi-heure et du dixième de degré. Cette source d'information présente toutefois une hétérogénéité spatio-temporelle ; le dispositif d'observation satellitaire évoluant dans le temps, les algorithmes d'estimation micro-ondes dépendant de l'état de surface et la couverture du réseau sol étant très inégale.

Les données ont été projetées sur une grille commune en quart de degré couvrant la totalité de la zone 60N/60S. Cette limite est imposée par le produit GPM qui ne fournit pas d'informations homogènes aux hautes latitudes.

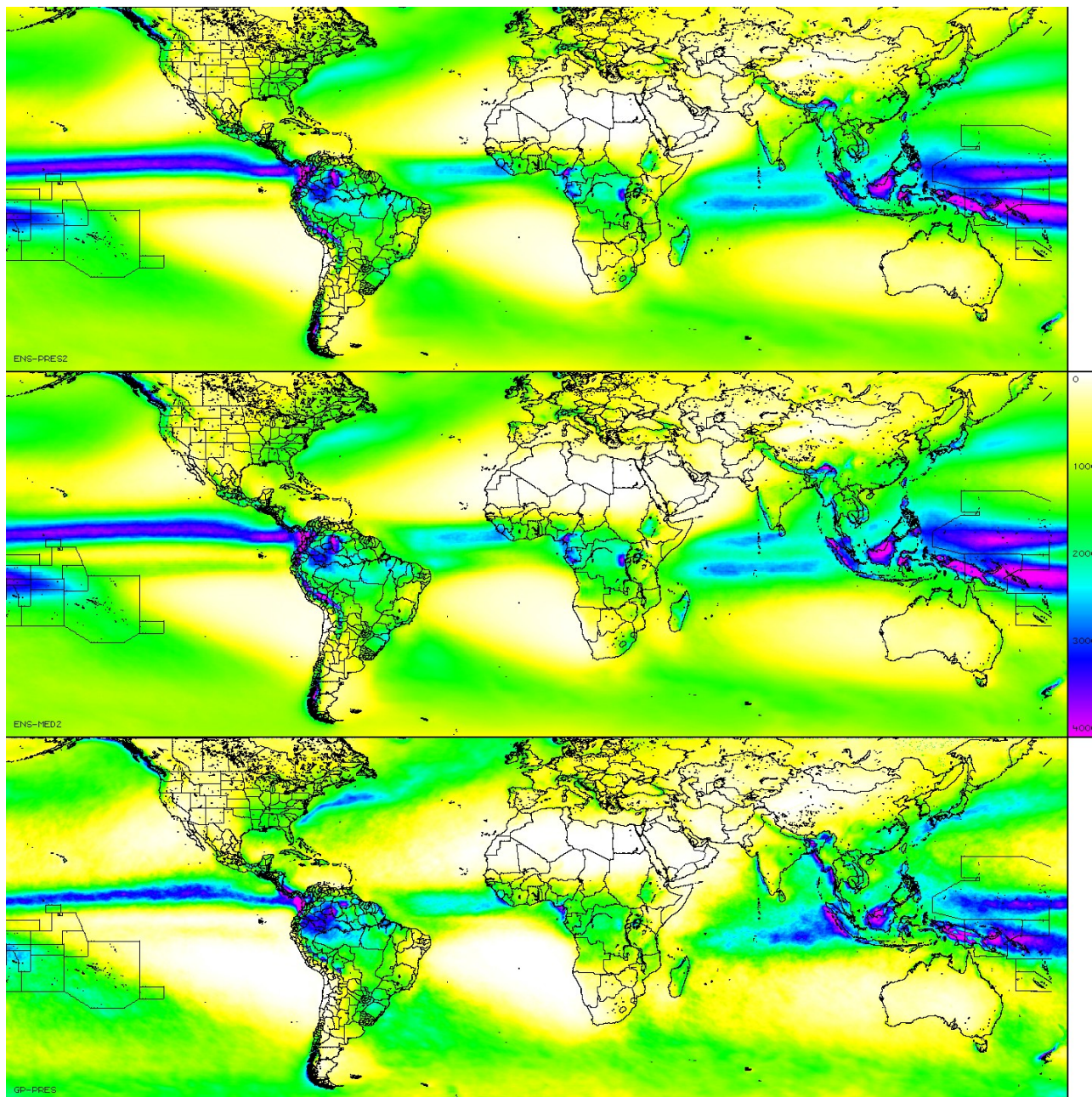
Deux paramètres de la distribution des précipitations ont été calculés : le cumul annuel de précipitations et le nombre de jours nécessaires pour cumuler 50% des pluies. Alors que le premier paramètre renseigne sur l'intensité du phénomène, le deuxième informe sur son intermittence. Il est obtenu en établissant pour chaque point de grille la série annuelle des précipitations quotidiennes, en triant cette série et en cumulant les valeurs jusqu'à dépasser la moitié du cumul annuel. La valeur de ce paramètre sera d'autant plus élevée que les pluies seront plus uniformément réparties. Bien que ce deuxième indice soit formellement indépendant du premier, sa valeur devient erratique dans les zones de faibles précipitations. Dans un premier temps, la synthèse des modèles a été effectuée par un simple moyennage en associant la même pondération à chaque modèle.

## 3. Analyse des résultats et retour sur la procédure de synthèse

La figure 1 présente le cumul annuel moyen de précipitations tel qu'estimé par cette procédure d'agrégation sur les périodes 2015-2020 (graphique du haut) et 2045-2050 (graphique du centre). Le graphique du bas est obtenu à partir du produit IMERG sur la période 2015-2020. Sur la période 2015-2020, cette procédure restitue la répartition des précipitations estimée par IMERG avec des gradients spatiaux plus atténués, ce qui est cohérent avec les échelles d'élaboration des produits. Toutefois, par rapport à IMERG, elle fait apparaître un biais positif de 9%. La répartition de ces erreurs n'est pas uniforme ; elles se concentrent aux basses latitudes avec un biais maximal de 1.3 mm/jour sur le quatrième parallèle sud (soit une erreur relative de 30%).

L'évolution à moyen terme des précipitations, telle qu'elle peut être déduite de cette figure, est en accord avec les conclusions IPCC. Un net renforcement des pluies sur la zone intertropicale de convergence apparaît clairement ainsi que, sur la plupart des zones de mousson, une augmentation sensible des

précipitations. Cette augmentation est particulièrement notable pour la mousson Ouest-Africaine à l'exception son extrémité occidentale. En revanche l'ouest du bassin méditerranéen évolue vers une situation de déficit pluviométrique. Toutefois le point important est que ces variations sont du même ordre de grandeur que l'erreur précédemment estimée.

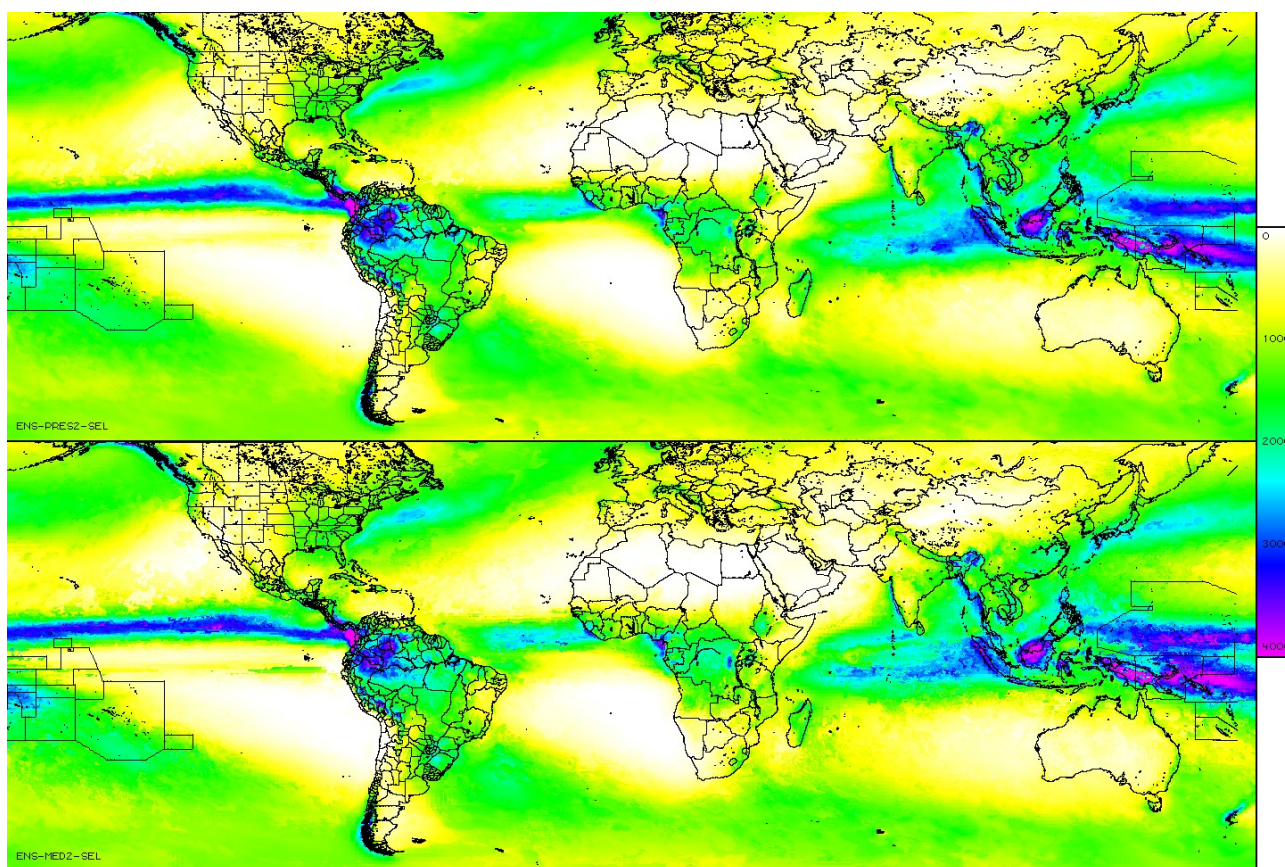


**Figure 1.** Moyenne du cumul annuel de précipitations (mm) sur la zone 60N/60S. Les images sont établies à partir de l'agrégation des GCM sur la période 2015-2020 (haut), de l'agrégation des GCM sur la période 2045-2050 (centre) et des estimations de précipitations GPM-IMERG sur la période 2015-2020 (bas).

Une procédure alternative permettant d'améliorer ces résultats est d'abandonner la pondération uniforme de la procédure d'agrégation au profit d'une sélection des modèles sur la base de leur adéquation locale aux observations. Yang *et al.* (2021) ont mis en œuvre une procédure de ce type pour évaluer les projections de précipitations sur la Chine. Une méthode plus élémentaire a été appliquée ici qui consiste à sélectionner en chaque point de grille le modèle le plus proche de la donnée IMERG. Le biais associé au cumul de précipitations estimé sur la période 2015-2020 est alors considérablement réduit et cette amélioration porte principalement sur la zone intertropicale où l'erreur relative maximale n'est plus que de 10%. La courbe des erreurs en fonction de la latitude ne présente plus le très net maxima à l'équateur. Cette observation pourrait être mis en relation avec la validité des schémas de convection des modèles qui mérite

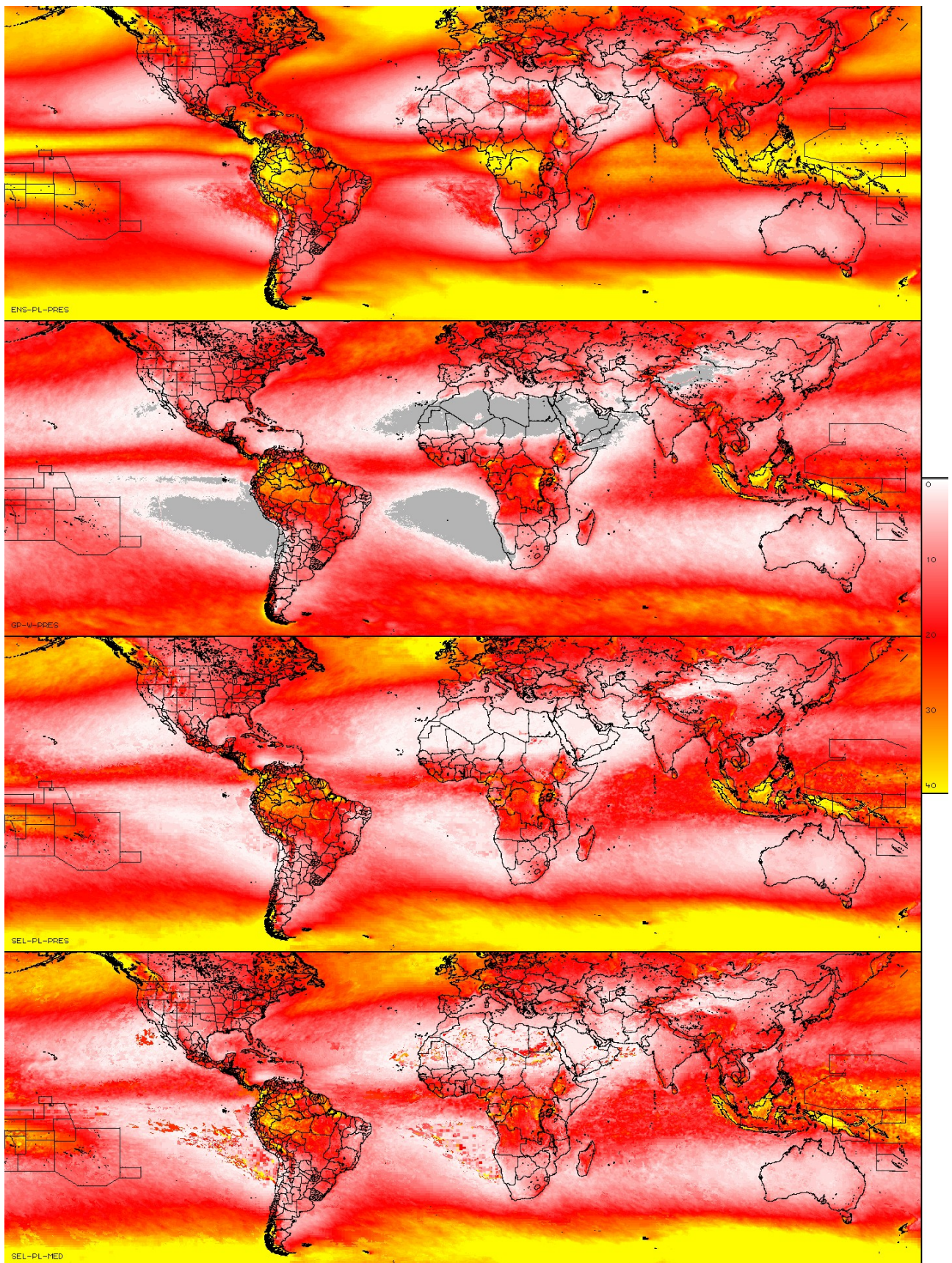
d'être évaluée. Bien qu'il soit, à ce stade, difficile d'associer un modèle à un régime de précipitations, quelques tendances émergent. Le modèle MRI-ESM2 semble le plus approprié pour prendre en compte le renforcement orographique qu'il restitue mieux que les modèles à très haute résolution CMCC-CM2 ou HiRAM-SIT. En zones de faibles précipitations (continentales ou maritimes), les modèles HiRAM-SIT et GFDL-CM4 présentent les biais les plus faibles. Enfin, les modèles MRI-ESM2, NCAR-CESM2 et CMCC-CM2 rendent le mieux compte des phénomènes pluviogènes en zones de mousson.

En figure 2 sont présentées les images obtenues par la méthode de synthèse précédemment décrite appliquée, d'une part, sur les projections pour la période 2015-2020 et, d'autre part, sur celles correspondant à la période 2045-2050. Bien que cette méthode sélectionne indépendamment pour chaque point de grille le modèle le plus approprié, la continuité spatiale des cartes générées est remarquable. Aucun gradient artificiel lié à une discontinuité entre modèles n'apparaît. En termes de changement de régime de précipitations, la différence la plus notable entre cette méthode de synthèse et la procédure d'agrégation est que les modifications des cumuls sont beaucoup plus faibles. Les zones en déficit ou en excès changent peu à l'exception des moyennes latitudes ; aucune tendance ne peut être mise en évidence sur l'Europe alors que la procédure précédente faisait apparaître un dipôle Est-Ouest ; l'Amérique du Nord présente un léger déficit au lieu d'un excédent. Les déficits en Méditerranée occidentale se maintient quoique d'une importance moindre.



**Figure 2.** Moyenne du cumul annuel de précipitations sur la zone 60N/60S. Les images sont établies à partir de l'agrégation des modèles suivant la procédure de sélection sur la période 2015-2020 (haut) et la période 2045-2050 (bas).

La figure 3 est similaire à la première figure mais porte sur le paramètre d'intermittence, le nombre de jours nécessaires pour cumuler 50% des précipitations. Cette statistique n'est calculée que lorsque le cumul pluviométrique dépasse 100 mm. En haut est représenté le résultat de la procédure d'agrégation sur 2015-2020 ; en centre-haut le même paramètre estimé par IMERG ; en centre-bas le paramètre extrait des modèles combinés par la méthode de sélection sur la période 2015-2020 et enfin en bas ce même



**Figure 3.** Moyenne du nombre de jours de l'année cumulant 50% des précipitations sur la zone 60N/60S. Les images sont établies à partir de la moyenne des GCM sur la période 2015-2020 (haut), des estimations de précipitations GPM-IMERG sur la période 2015-2020 (centre-haut), de la procédure par sélection de GCM sur la période 2015-2020 (centre-bas) et de la procédure par sélection de GCM sur la période 2045-2050 (bas). Les zones où la statistique n'est pas calculée figurent en gris.

indicateur sur la période 2045-2050. L'artefact évident sur la synthèse par moyenne des modèles est une massive surestimation par rapport à IMERG qui est de l'ordre de 75%. La synthèse par agrégation tend à répartir les pluies de manière beaucoup plus uniforme que les estimations IMERG. Ce biais est d'autant plus significatif, qu'alors que les estimations quantitatives de pluie sont imprécises pour les plus fortes intensités, la détection d'amas pluvio-gènes par les capteurs micro-ondes est non ambiguë. Aussi l'indicateur d'intermittence dérivé d'IMERG devrait être plus précis que le cumul de précipitations. Il est également notable que pour ce paramètre la dispersion entre les modèles est plus importante que pour le cumul de pluies.

La procédure de synthèse sélectionnant pour chaque nœud de grille le modèle le plus proche de la référence a été appliqué sur le paramètre d'intermittence. Mais alors que la contribution des divers modèles au cumul de précipitations se présentait comme une combinaison complexe, cette synthèse met en évidence la supériorité nette d'un modèle. Sur plus de 50% des points le modèle HiRAM-SIT fournit l'estimation la plus proche d'IMERG. Il mérite d'être signalé que la sélection du modèle local le plus approprié dépend étroitement du paramètre considéré. La comparaison des deux synthèses par ce dernier procédé met en évidence une tendance globale à l'augmentation des événements extrêmes, le nombre de jours nécessaires au cumul de 50% des pluies diminuant. La seule zone d'extension significative où le phénomène inverse se manifeste est le Nord-Est asiatique.

## Conclusion

Malgré les innovations récentes apportées par l'IPCC dans la diffusion des informations, le portail du CMIP reste un complément indispensable pour obtenir des données numériques. Mais, outre l'important volume de données, l'utilisateur est confronté à une dispersion des modèles particulièrement sensible pour les précipitations. Comme le suggère cette étude, cette même dispersion rend la procédure de moyennage aveugle, largement utilisée jusqu'ici, partiellement inefficace et une procédure alternative a été proposée qui s'attache à sélectionner un modèle localement pertinent. Des calculs élémentaires permettent d'obtenir des estimations apparemment plus réalistes de deux paramètres de la distribution des intensités de pluie. Toutefois cette proposition garde une nature plus heuristique qu'algorithmique. Définir une procédure opérationnelle supposerait tout d'abord qualifier les erreurs de la source de données de référence (le GPM-IMERG) et éventuellement lui substituer des données localement plus précises. La procédure elle-même pourrait gagner en sophistication en substituant à la simple sélection d'un modèle une somme pondérée. Enfin, le jeu de données de modèles utilisé a été constitué sur une base d'opportunité. Un mode de constitution plus systématique intégrant les données CORDEX serait également souhaitable. Ceci introduit bien sûr une complexité de traitement supplémentaire par rapport à la procédure élémentaire de synthèse fréquemment utilisée. Mais celle-ci ne permet plus de prendre en compte la richesse de l'information extraite du CMIP6.

## Bibliographie

Eyring V., Bony S., Meehl G.A., Senior C., Stevens B., Stouffer R. & Taylor K., 2016 : Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geoscience Model Development*, **9**, 1937-1958

Hourdin F., Mauritsen T., Gettelman A., Golaz J.C., Balaji V., Duan Q., Folini D., Ji D., Klocke D., Qian Y., Rauser F., Rio C., Tomassini L., Watanabe M. & Williamson D., 2017 : The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, **98(3)**, 589-602

Huffman G., Bolvin D., Braithwaite D., Hsu K., Joyce R., Kidd C., Nelkin E., Sorooshian S., Tan J., & Xie P., 2019 : Algorithm theoretical document ; NASA GPM ; Integrated Multi-satellite retrievals for GPM. NASA technical document ATBD V06, 38 p, <https://gpm.nasa.gov>

Yang X., Zhou B., Xu Y. & Han Z., 2021 : CMIP6 evaluation and projection of temperature and precipitation over china. *Advances in Atmospheric Research*, **38**, 817-830