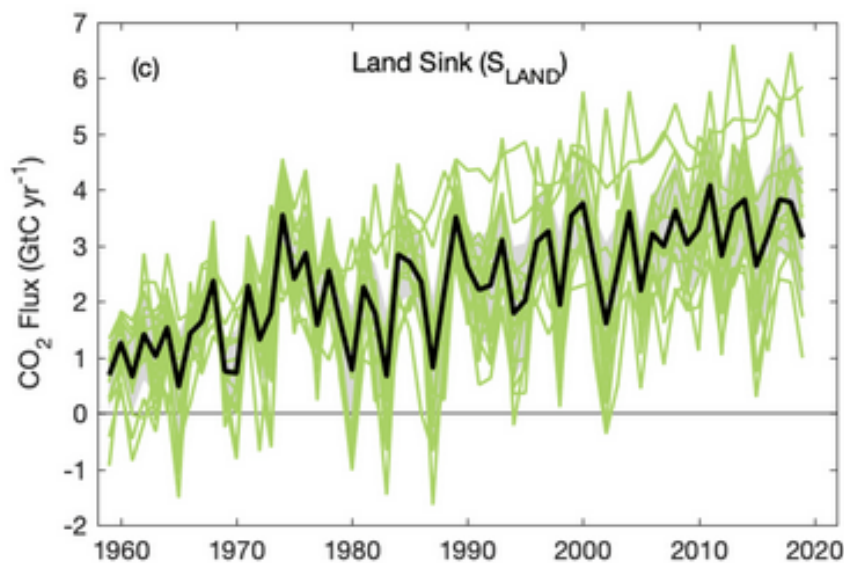


INSA TOULOUSE
DÉPARTEMENT GÉNIE PHYSIQUE

RAPPORT DE STAGE

Rôle des forçages atmosphériques dans
l'estimation des stocks et puits de carbone
par les modèles de surface.



Etudiant :

Lucas HARDOUIN

Tuteur INSA :

Julian CARREY

Tuteurs :

Christine DELIRE

Bertrand DECHARME

christine.delire@meteo.fr

bertrand.decharme@meteo.fr

Table des matières

1	Introduction	7
1.1	Présentation de la structure d'accueil	7
1.2	Contexte scientifique et problématique associé	8
1.2.1	Le carbone sur terre	8
1.2.2	Carbone et réchauffement climatique	9
1.2.3	Estimation du puits de carbone continental	10
1.3	Objectifs	13
2	Données et méthodologie	14
2.1	Données	14
2.1.1	Modèles	14
2.1.2	Forçages	16
2.1.3	Variables	18
2.2	Traitement de données	20
2.2.1	Pré-traitement avec Climate Data Operator (CDO)	20
2.2.2	Traitement avec Python	21
2.2.3	Analyse de la variance (ANOVA)	22
3	Résultats	26
3.1	Les flux de carbone globaux	26
3.1.1	Exemple de la production primaire brute	26
3.1.2	Respiration et production primaire nette	30
3.1.3	Le puits de carbone continental	32
3.2	Échelle Régionale	36
4	Conclusion	40
A	Variabilité autour des modèles des flux de carbone	43
B	Répartition de la variabilité des flux de carbone	44
C	Estimation des stocks de carbone par les modèles	45
D	Distribution spatiale de la variabilité	46
E	Évolution temporelle de flux secondaires	48
F	Comparaison des estimations de la production primaire brute entre le CNRM et le NCAR	49
G	Questionnaire d'étonnement	50

Table des figures

1	Principaux flux et réservoirs mis en jeu dans le cycle du carbone. Les flèches noires correspondent aux flux du cycle du carbone "naturel", tandis que les flèches rouges sont liées à l'activité humaine. Science Education Resource Center (SERC) at Carleton College [2].	10
2	Représentation des principaux échanges du cycle du carbone global et de ses perturbations dues à l'activité humaine. Les valeurs sont moyennées pour la décade 2010-2019. Global Carbon Budget 2020 [17].	11
3	Représentation du cycle du carbone dans le modèle ISBA-CTRIP. Delire et al. (2020) [12].	15
4	Variabilité relative des variables de précipitation (<i>pr</i>), d'humidité spécifique (<i>huss</i>), d'humidité relative (<i>hurs</i>), de rayonnement incident grande longueur d'onde (<i>rlds</i>) et petite longueur d'onde (<i>rsds</i>), de vitesse du vent à 10m (<i>sfcWind</i>) et de température de surface à 2m (<i>tas</i>), moyennées de 2003 à 2012, pour les forçages GSWP3, CRUNCEP et Princeton.	17
5	Humidité spécifique moyenne sur la décade 2003-2012 pour le forçage GSWP3. Différence de l'humidité spécifique entre GSWP3 et CRUNCEP (gauche) et GSWP3 et Princeton (droite) pour la décade 2003-2012.	18
6	Évolution de la production primaire brute de 1960 à 2012 estimée par les modèles du CNRM (bleu), MPI (rouge), BCC (cyan), MOHC (violet), IPSL (vert) et NCAR (orange) pour les forçages GSWP3 (traits pleins), CRUNCEP (tirets) et Princeton (pointillés).	21
7	Moyenne des estimations de l'évolution de la production primaire brute pour les modèles du CNRM (bleu), MPI (rouge), BCC (cyan), MOHC (violet), IPSL (vert) et NCAR (orange).	27
8	Moyenne des estimations de l'évolution de la production primaire brute pour les forçages atmosphériques GSWP3 (trait plein bleu foncé), CRUNCEP (tirets oranges) et Princeton (pointillés verts) à partir des données des six modèles de surface.	28
9	Dispersion de la moyenne temporelle de la production primaire brute selon les forçages (à gauche), selon les modèles (au centre) et pourcentage de variabilité expliqués par les forçages, les modèles et les interactions pour la production primaire brute (à droite).	29
10	Dispersion de la moyenne temporelle (1960-2012) selon les forçages de la respiration autotrophe (<i>Ra</i>) à gauche, la production primaire nette (<i>NPP</i>) au centre, et la respiration hétérotrophe (<i>Rh</i>) à droite.	30

11	Dispersion de la moyenne temporelle (1960-2012) selon les modèles de la respiration autotrophe (Ra) à gauche, la production primaire nette (NPP) au centre, et la respiration hétérotrophe (Rh) à droite.	31
12	Pourcentage de variabilité explicable par les forçages, modèles et interactions pour la respiration autotrophe (Ra) à gauche, la production primaire nette (NPP) au centre, et la respiration hétérotrophe (Rh) à droite.	31
13	Dispersion de la moyenne temporelle de GPP-Ra-Rh selon les forçages (à gauche), selon les modèles (au centre) et pourcentage de variabilité expliqués par les forçages, les modèles et les interactions pour le flux GPP-Ra-Rh (à droite).	33
14	Évolution temporelle des émissions liées aux feux de forêt pour les simulations des modèles du CNRM (bleu), du NCAR (orange) et du MPI (rouge) forcé par GSWP3 (trait plein), CRUNCEP (tireté) et Princeton (pointillé).	34
15	Dispersion de la moyenne temporelle du puits de carbone continental selon les forçages (à gauche), selon les modèles (au centre) et pourcentage de variabilité expliqués par les forçages, les modèles et les interactions pour le puits de carbone continental (à droite).	35
16	Distribution spatiale de la variabilité pour la production primaire brute. Les régions plus sombre sont celles qui contribuent le plus à la production globale.	36
17	Distribution spatiale de la variabilité pour la production primaire nette.	37
18	Distribution spatiale de la variabilité pour le puits de carbone continental.	38
19	Dispersion de la moyenne temporelle des différents flux pour les modèles du CNRM (bleu), du MPI (Rouge) et du NCAR (orange).	43
20	Variabilité expliqué par les forçages (violet), modèles (orange), et interactions (rose) pour la production primaire brute (a), la respiration autotrophe (b), la production primaire nette (c), la respiration hétérotrophe (d), GPP-Ra-Rh (e), le flux de chaleur latente HFLS (f), les émissions liées aux feux de forêt (g) et le puits de carbone continental (h).	44
21	Moyenne des estimations du carbone stocké pour les modèles du CNRM (bleu), MPI (rouge), BCC (cyan), MOHC (violet), IPSL (vert) et NCAR (orange).	45
22	Distribution spatiale de la variabilité pour la respiration autotrophe (a), la respiration hétérotrophe (b), GPP-Ra-Rh (c) et fFire (d).	47
23	Moyenne par modèle de l'estimation des émissions du aux changements d'usage des sols (a) et de l'estimation du flux de chaleur latente HFLS (b) pour les modèles du CNRM (bleu), MPI (rouge), MOHC (violet), IPSL (vert) et NCAR (orange).	48

24	à gauche : différences entre les estimations de la production primaire brute faite par le modèle du CNRM : CRUNCEP – GSWP3 (haut), Princeton – GSWP3 (centre) et Princeton – CRUNCEP (bas). à droite : différences entre les estimations des modèles du CNRM et du NCAR forcé par CRUNCEP (haut), Princeton (milieu) et GSWP3 (bas).	49
----	--	----

Acronymes

ANOVA ANalysis Of VAriance. 22

CDO Climate Data Operator. 1, 20

CMIP6 Sixième Projet d'Intercomparaison des Modèles Couplés. 13, 14

CNRM Centre National de Recherche Météorologique. 7, 8, 25, 27, 41

CNRS Centre National de Recherche Scientifique. 7

CTRIP CNRM version of Total Runoff Integrating Pathways. 8

ESGF Earth System Grid Federation. 14, 19

ESM Earth System Model. 8, 11, 14, 15

f_{Fire} émissions liées aux incendies. 18, 19, 40

f_{Luc} émissions liées aux changements d'usage des sols. 19

GCP Global Carbone Project. 10, 12, 26, 41

GES Gaz à Effet de Serre. 7, 9

GPP production primaire brute. 19, 27–31, 36–40

IPCC Intergovernmental Panel on Climate Change. 14, 22

ISBA Interaction Soil-Biosphere-Atmosphere. 8

LS3MIP Land Surface, Snow and Soil moisture Model Intercomparaison Project. 14, 16

MPI Max Planck Institute. 25

NBP production nette du biome. 19

NCAR National Center for Atmospheric Research. 25, 27, 41

NEP production nette de l'écosystème. 19, 41

NPP production primaire nette. 19, 31, 38, 40

Ra respiration autotrophe. 19, 31, 40

Rh respiration hétérotrophe. 19, 31, 40

SOAC-DC Sciences de l'Océan, de l'Atmosphère et du Climat, parcours Dynamique du Climat. 7

Remerciements

Je tenais avant tout à remercier chaleureusement mes encadrants, les Drs. Christine Delire et Bertrand Decharme, pour le soutien, la confiance et la bienveillance qu'ils m'ont accordé durant ce stage. Leurs conseils et leurs encouragements m'ont permis de progresser rapidement et de me sentir confiant dans mon travail. J'ai beaucoup appris sur les sciences du climat et sur le rôle des surfaces grâce à eux, de part les lectures qu'ils m'ont proposées et les explications qu'ils m'ont apportées. Je souhaitais plus particulièrement les remercier de m'avoir ouvert les portes de la recherche, notamment en me laissant le temps de préparer le concours de l'école doctorale, malheureusement sous une autre direction que la leur, de m'avoir permis de préparer un article à leur côté, et surtout de veiller à ce que tout se passe bien pour moi à l'avenir.

Ensuite, je souhaitais remercier le CNRM, le groupe GMGEC et toutes les personnes que j'ai pu y rencontrer. Merci aux membres du deuxième étage pour les bons moments et les pauses cafés. Je remercie plus particulièrement Raphael et Maya, les doctorants de mes maîtres de stage, pour avoir partagé avec moi leurs codes Python et m'avoir ainsi épargné de longues heures à programmer. Je les remercie aussi pour les heures, étrangement moins longues, passées à manger, à boire, à jouer ou à grimper, avec les autres "djeuns" du CNRM : Brice, Morgane, Simon, Antoine, Gabou, Arthur, Arnaud, Léa, Emilie, Nils et tous ceux avec qui j'ai pu échanger.

Enfin, il me tient à coeur de remercier les personnes, qui par leur présence et leur soutien, m'ont permis de mener à bien ce stage. D'abord, tous les copains et toutes les copines, en passant par les rencontres de cette année aux proches ami.e.s du collège et lycée, ou à mes supers camarades de l'INSA, qui m'ont permis de me changer les idées et ne se sont pas endormis (ou du moins ont essayé) quand je leur racontais mes projets. Ma famille bien sûr, qui a toujours été là pour moi et qui constitue ma première source d'inspiration. Cassandre, pour me supporter au quotidien, et rendre ce dernier meilleur. Et, pour terminer, je remercie du fond du coeur les personnes qui ont pris le temps de relire ce rapport ainsi que mes précédents écrits. Elles doivent ruminer contre mes enseignantes de primaire.

1 Introduction

Ce rapport a pour but de présenter une synthèse des travaux et activités menés au cours de mon stage de fin d'étude, pour l'obtention du double diplôme d'ingénieur Génie Physique et du master Sciences de l'Océan, de l'Atmosphère et du Climat, parcours Dynamique du Climat (SOAC-DC), où je me suis intéressé au rôle des forçages atmosphériques dans l'estimation des stocks et puits de carbone par les modèles de surface. Dans un premier temps, je décris le cadre et le contexte scientifique de ce stage, ainsi que les objectifs attendus. Par la suite, les données utilisées et les méthodes que j'ai développées pour répondre à ces objectifs sont explicitées. Une troisième partie présente ensuite les résultats obtenus, en incluant une analyse et une discussion. Enfin, je termine ce rapport par une double conclusion, présentant, d'une part, un regard critique et un compte rendu des travaux sur le plan scientifique, et, d'autre part les apports professionnels et personnels de ce stage.

1.1 Présentation de la structure d'accueil

Ce stage a été effectué au Centre National de Recherche Météorologique (CNRM), une unité mixte de recherche sous la co-tutelle du Centre National de Recherche Scientifique (CNRS) et de Météo-France, situé sur le campus de la météopole de Toulouse (42 avenue Gaspard Coriolis), Toulouse, France). Les actions de recherche du CNRM concernent essentiellement les prévisions des phénomènes météorologiques, notamment les phénomènes extrêmes et la météorologie urbaine, mais aussi l'étude du climat et des perturbations climatiques, les cycles biogéochimiques et plus particulièrement celui de l'eau, ainsi que les relations aux interfaces océans, atmosphère, continents et cryosphère.

Plus particulièrement, j'ai intégré le Groupe de Météorologie Grande Échelle et Climat (GM-GEC) qui s'intéresse aux interactions pouvant exister entre les différentes échelles spatiales, aux interactions entre les composantes du système climatique, aux réponses de ce système aux forçages anthropiques ainsi qu'aux sources d'incertitudes affectant la variabilité et la prévisibilité du climat futur. Ces travaux sont essentiellement menés grâce à la modélisation du climat et de l'atmosphère, à l'échelle globale et régionale, ainsi que par la collaboration avec d'autres groupes de modélisation dans le cadre d'exercices d'intercomparaison.

J'ai évolué au sein de l'équipe Étude du Système Terre, (EST), qui est principalement en charge de la modélisation du système terre dans son ensemble, c'est à dire en prenant en compte les transports de masses et d'énergie en plus de la circulation générale. Les membres de l'équipe s'intéressent à l'impact sur le climat des flux de Gaz à Effet de Serre (GES) et d'aérosols, et inversement aux rétroactions du changement climatique sur ces flux. Des travaux de recherche portent également sur l'étude de techniques d'atténuation du changement climatique via la géo-ingénierie, en particulier sur la gestion des surfaces continentales.

Pour ce faire, l'équipe EST contribue notamment au développement du modèle de surface continentale Interaction Soil-Biosphere-Atmosphere (ISBA) et de transport fluvial : CNRM version of Total Runoff Integrating Pathways (CTRIP) [11], inclus dans la plate-forme de modélisation SURFace EXternalisé (SURFEX) de Météo-France. En particulier, la version système Terre d'ISBA-CTRIP [12], mis en place par mes tuteurs, les Drs. Christine Delire, et Bertrand Decharme, permet une représentation complète de la partie continentale du cycle du carbone, prenant en compte les échanges biogéochimiques entre l'atmosphère, le sol, et la végétation, en plus des processus physiques liés au climat et à l'hydrologie.

Le système terre est modélisé grâce à l'Earth System Model (ESM) du CNRM, CNRM-ESM2-1 [48]. Ce modèle de système terre se base sur le modèle couplé de circulation générale, CNRM-CM6-1, qui grâce au coupleur OASIS-MCT représente la dynamique atmosphérique (ARPEGE-Climat v6.3) et océanique (NEMO) ainsi que les interactions avec les surfaces (ISBA-CTRIP "physique"), les lacs (FLake), et la glace de mer (GELATO). Ce coeur "physique" est ensuite complété par la représentation du cycle du carbone continentale (via ISBA-CTRIP) et océanique (PISCES), de la chimie atmosphérique (REPROBUS), et des aérosols (TACTIC), développés par d'autres équipes de recherche du CNRM.

1.2 Contexte scientifique et problématique associé

1.2.1 Le carbone sur terre

Le carbone est un élément chimique indispensable, et intrinsèque, à la vie telle que nous la connaissons. Son numéro atomique (6) lui permet de former jusqu'à quatre liaisons covalentes, parfois doubles ou triples, lui permettant de jouer le rôle de charpente moléculaire du vivant. Complété par l'hydrogène, l'oxygène, mais aussi parfois l'azote, le phosphore, ou le soufre, il est la base de macromolécules, qui peuvent, notamment, s'apparier entre elles pour former le code génétique : l'ADN.

On estime à environ 2000 Gigatonne (Gt) la quantité de carbone présente dans la *biosphère*, qu'on définit comme l'ensemble des écosystèmes existants sur terre, que ce soit à la surface de la *lithosphère*, dans l'*hydrosphère* ou dans l'*atmosphère*. Pourtant, la majorité du carbone est en réalité inorganique, et se rencontre le plus souvent dans la croûte terrestre sous forme de roche sédimentaire (50 à 80 millions de Gt), ce qui fait de la lithosphère le premier réservoir de carbone terrestre (99,6 % du carbone total). En seconde position vient l'hydrosphère, et plus particulièrement les océans profonds, avec 38 000 Gt de carbone inorganique dissous. Le plus petit de ces quatre grands réservoir étant l'atmosphère, avec environ 800 Gt de carbone, que l'on rencontre le plus souvent sous forme de dioxyde de carbone (CO₂).

1.2.2 Carbone et réchauffement climatique

Le CO_2 est un gaz inorganique ne représentant qu'une toute petite fraction (0.041%) de l'air que nous respirons. On exprime souvent la concentration de CO_2 atmosphérique en partie par million (ppm) plutôt qu'en pourcentage, soit 410 ppm ici. Comme d'autres gaz présents dans l'atmosphère, il contribue à ce que l'on nomme l'effet de serre, c'est à dire l'interception d'une partie du rayonnement thermique infrarouge sortant émis par la terre. Cela a pour effet d'augmenter la température en surface, et ainsi d'obtenir les conditions climatiques, — propices au vivant —, que nous connaissons.

Cependant, depuis le début de l'âge industriel, la concentration des GES dans l'atmosphère ne cesse d'augmenter. Par exemple, la concentration de CO_2 atmosphérique est passée de 277 ppm en 1750 à plus de 410 ppm aujourd'hui [17], soit une augmentation de 48%. Celle du méthane (CH_4), un hydrocarbure (organique) ayant un pouvoir réchauffant 9 fois supérieur à celui du CO_2 , est passé de 722 parties par milliard (ppb) en 1750 à 1870 ppb (1,870 ppm) aujourd'hui. L'augmentation de la concentration de ces GES induit une augmentation de l'effet de serre et donc un réchauffement du climat moyen. Ce réchauffement est par ailleurs amplifié de manière dramatique par différentes rétroactions du système climatique, une des plus importante étant l'augmentation de la concentration de vapeur d'eau (un autre GES) dans l'atmosphère.

Le CO_2 serait, à lui seul, responsable de 66% du forçage attribuable au GES, ce qui fait de l'étude de ce gaz un enjeux primordial pour comprendre le climat et prévoir son évolution. Pour comprendre l'évolution de la concentration de CO_2 atmosphérique, il faut s'intéresser au cycle biogéochimique du carbone, c'est à dire à l'ensemble des processus de transport et de transformation du carbone sur terre. L'étude des cycles biogéochimiques, notamment celui du carbone, a toujours été d'une grande utilité pour comprendre les disponibilités et la transformation des éléments dans la nature, notamment pour l'agriculture puis, par la suite, pour le développement de l'industrie et la recherche d'hydrocarbures [22]. Aujourd'hui, ce sont finalement les problématiques environnementale et climatique, provoquées par l'industrialisation, qui renouvellent l'intérêt que l'on porte à ces cycles.

A notre échelle de temps, les échanges de carbone ont essentiellement lieu aux interfaces océan-atmosphère et atmosphère-surface continentale. Sans perturbation, en moyenne spatiale et temporelle, le système est à l'équilibre : le carbone est échangé en permanence entre les réservoirs, mais les flux entrants (sources) et sortants (puits) sont égaux. L'activité humaine vient cependant perturber cette équilibre en rajoutant des sources de carbone pour l'atmosphère. La figure 1 présente une version schématique de ce cycle du carbone "perturbé", les flèches en rouge représentant les émissions d'origine anthropique.

Cependant, l'augmentation de la concentration de CO_2 atmosphérique s'accompagne également d'un accroissement de l'efficacité des puits de carbone océaniques et continentaux. Ce phénomène est expliqué, d'une part, par des mécanismes physiques, liés à la dissolution du

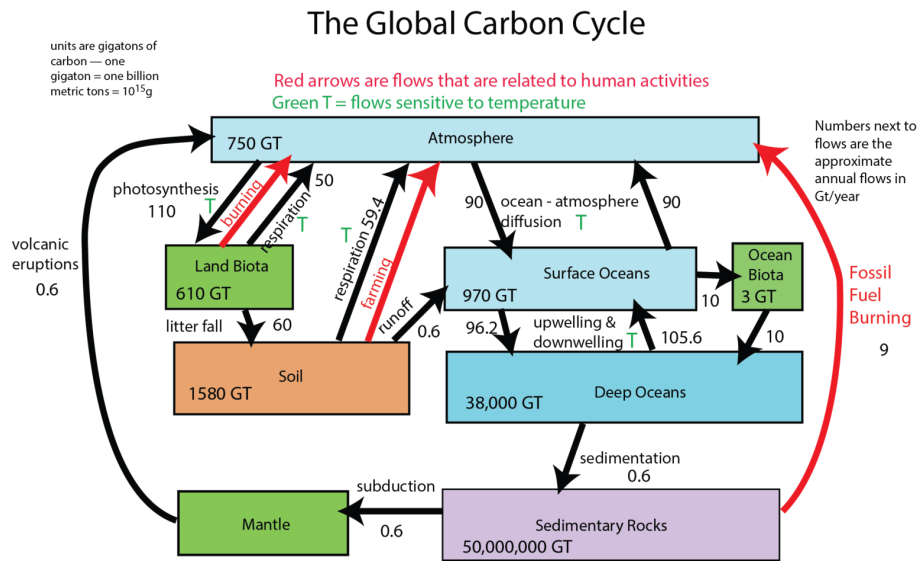


FIGURE 1 – Principaux flux et réservoirs mis en jeu dans le cycle du carbone. Les flèches noires correspondent aux flux du cycle du carbone "naturel", tandis que les flèches rouges sont liées à l'activité humaine. Science Education Resource Center (SERC) at Carleton College [2].

carbone dans l'océan de surface, et, d'autre part, par des processus biologiques, concernant la transformation du CO_2 inorganique en matière organique par les organismes autotrophes, comme les plantes ou le phytoplancton. On estime que depuis le début de l'âge industriel, ce sont 32% des émissions de CO_2 qui ont été absorbés par les surfaces continentales, et 24 % par les océans [16]. Des explications supplémentaires sur le cycle du carbone, sur ses perturbations et sur le rôle des surfaces continentales, peuvent être trouvés dans le rapport bibliographique à destination de l'INSA Toulouse [20], réalisé en début de stage .

1.2.3 Estimation du puits de carbone continental

Des groupements de chercheurs, comme le Global Carbon Project (GCP), formé en 2001, ont pour vocation de comprendre le cycle du carbone, et s'intéressent à l'évolution de la concentration de CO_2 , de CH_4 mais aussi de protoxyde d'azote (N_2O), dans l'atmosphère. Depuis 2013 [33], le GCP publie des rapports annuels détaillant l'évolution des puits, des sources et de la concentration de CO_2 atmosphérique, de 1959 à nos jours [31], [17]. La figure 2 présente les principaux échanges de carbone globaux pour la décennie 2010-2019.

En ce qui concerne le puits de carbone continental, S_{Land} , c'est à dire la quantité de CO_2 absorbé annuellement par les surfaces, il était d'abord estimé par le GCP de manière diagnostique, en fonction des émissions de carbone lié à la combustion ou l'oxydation d'énergie fossile, E_{Fos} , et des émissions et absorption de carbone liées au changement d'usage des sols, E_{Luc} , auxquelles on retranche la part des émissions qui vont contribuer à l'augmentation du taux de

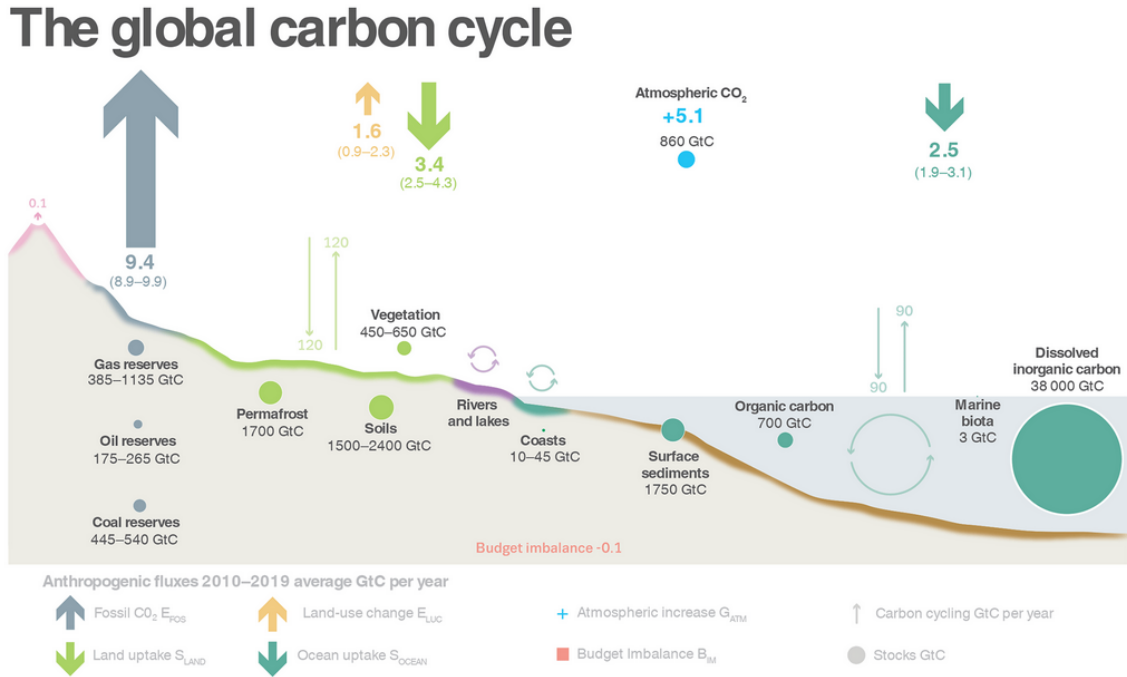


FIGURE 2 – Représentation des principaux échanges du cycle du carbone global et de ses perturbations dues à l'activité humaine. Les valeurs sont moyennées pour la décennie 2010-2019. Global Carbon Budget 2020 [17].

CO₂ atmosphérique, G_{Atm} , et celle absorbée par le puits océanique, S_{Ocean} , soit :

$$S_{Land-ancien} = E_{Fos} + E_{Luc} - G_{atm} - S_{Ocean} \quad (1.1)$$

Cependant, depuis 2017 [32], la méthodologie a été révisé, et S_{Land} est désormais calculé en utilisant les estimations de différents modèles de surfaces continentales. Nous détaillerons par la suite la méthode de calcul utilisé.

Les premiers modèles de surface sont apparus avec le début de la modélisation climatique [36], pour fournir aux modèles de circulation atmosphérique une représentation réaliste des échanges d'énergie, d'eau et de quantité de mouvement ayant lieu aux interfaces [10]. Pour correctement représenter l'hydrométéorologie des surfaces continentales, ces modèles ont rapidement intégré une représentation du couvert végétal [9], car l'évaporation des hydrométéores interceptés par la canopée, ainsi que la transpiration des végétaux, jouent un rôle essentiel dans le cycle hydrologique. Plus récemment, les risques induits par le changement climatique ont conduit à l'élaboration de modèles nouvelle génération dits système terre (ESM), et donc au besoin de représenter les cycles biogéochimiques, en particulier celui du carbone, dans les modèles de surface continentale [15].

En effet, si la biosphère continentale absorbe une partie du CO₂ atmosphérique, l'efficacité de ce puits va directement dépendre des conditions environnementales et climatiques. Les variations de température, d'humidité et de rayonnement vont grandement influencer l'efficacité de la photosynthèse, mais aussi les émissions de carbone liées à la respiration des végétaux et

du sol [3]. Par exemple, des températures suffisamment élevées sont nécessaires pour réaliser la transformation du CO_2 en matière organique, et la production primaire, c'est à dire la vitesse à laquelle la biomasse est synthétisée, augmente dans un premier temps avec la température. Cependant, cet effet bénéfique arrive rapidement à saturation, tandis que la température augmente également de manière exponentielle la respiration des plantes et du sol. Par conséquent, l'efficacité du puits de carbone tend à diminuer au delà d'un seuil de température critique. On comprend alors bien, d'une part, les rétroactions que peut avoir le changement climatique sur le cycle du carbone, et, d'autre part, l'importance que les conditions météorologique peuvent avoir sur ce puits.

Pourtant, si, dans le dernier rapport du GCP [17], 17 modèles de surface sont utilisés pour calculer S_{Land} — ce qui permet de représenter l'incertitude lié à la paramétrisation des processus dans les modèles — un seul et même forçage atmosphérique est utilisé pour l'ensemble des simulations. Un forçage atmosphérique est un ensemble de données, le plus souvent multi-horaire, contenant l'évolution des précipitations, de l'humidité, de la température de surface, de la vitesse du vent ou encore du rayonnement incident aux courtes (visible) et grandes (infrarouge) longueurs d'onde, et ce pour chaque point de grille. C'est donc à partir de ces données climatiques, construites à partir d'observations et de techniques de réanalyse, que les modèles estiment ces flux de carbone.

Cependant, le forçage atmosphérique constitue en soit une source d'incertitude. D'abord, parce que les observations comprennent nécessairement des erreurs par rapport à la réalité. On retrouve, dans ces erreurs d'observation, des erreurs liées à la mesure (erreurs instrumentales, erreurs dans le protocole de mesure), mais aussi des erreurs liées à la représentativité de l'observation, qui peut être très proche de la réalité à un point de mesure donné, sans pour autant correctement représenter l'environnement autour (par exemple, la température est le plus souvent mesurée en plaine, et décrit donc mal la réalité des reliefs ou du couvert végétal qui l'entoure). Ensuite, à cause de la mauvaise distribution spatio-temporelle des observations, autrement dit du manque de donnée. Ce manque de données est en partie pallié par des techniques de réanalyse, qui à partir d'une ébauche (une simulation de climat), corrigent les observations pour produire un ensemble de données complet pour chaque point de grille et pas de temps. On doit alors considérer des erreurs liées à la résolution du modèle, un point de grille représentant seulement une valeur moyenne sur une surface de plusieurs centaines voir milliers de km^2 , et un pas de temps pouvant couvrir plusieurs heures voir plusieurs jours. Enfin, les simulations climatiques utilisées pour combler ces observations sont elles mêmes soumises à l'erreur liées aux imperfections des modèles atmosphériques (on parle d'erreur d'ébauche), tandis que des biais peuvent aussi provenir de la technique d'analyse utilisée pour créer le forçage atmosphérique.

L'incertitude liée au forçage atmosphérique suscite alors plusieurs interrogations, qui ont

motivé la réalisation de ce stage. Comment cette incertitude se répercute-t-elle sur les estimations des modèles de surface ? Est-elle la même pour toutes les variables du cycle du carbone ? Est-elle significative par rapport aux différences causées par la représentation des processus dans les modèles ? Cette incertitude est-elle la même d'un modèle à l'autre ? En somme : quel est le rôle des forçages atmosphériques dans l'estimation des stocks et flux de carbone par les modèles de surface ? Notons que, si les incertitudes entre les modèles ont été estimées à de nombreuses reprises, en comparant les modèles aux observations [41] ou entre eux [51], seulement quelques études ont estimé l'incertitude liée aux forçages atmosphériques dans le cycle du carbone [23, 42], mais jamais en la comparant à l'incertitude liée aux modèles de surface.

1.3 Objectifs

Les objectifs de mon stage ont évolué au cours de son avancement et des découvertes que nous avons faites. L'objectif initial était d'évaluer l'incertitude liée aux forçages atmosphériques sur les flux et stocks de carbone à l'échelle globale, en comparant les simulations réalisées par différents modèles utilisant différents forçages, pour un ensemble de variables relatives au cycle du carbone. Cette objectif ayant été accompli, les objectifs et activités suivantes ont été réalisés durant ce stage :

- Mener des travaux bibliographiques sur le cycle du carbone et sur sa paramétrisation dans les modèles de surfaces.
- Réaliser un traitement de données sur des produits de simulations, réalisées dans le cadre du Sixième Projet d'Intercomparaison des Modèles Couplés (CMIP6) pour observer et mettre en valeur le rôle des forçages atmosphériques dans l'estimation des stocks et flux de carbone.
- Définir des variables d'intérêt, relatives au puits de carbone continentale.
- Quantifier l'incertitude totale sur ces variables à l'échelle globale. Utiliser des outils statistiques pour définir la variabilité attribuable aux modèles, aux forçages, et aux variations dans la réponse des modèles aux forçages atmosphériques.
- Observer le rôle des forçages atmosphériques à l'échelle régionale pour identifier les principales sources de variabilité et observer sa distribution spatiale. Transposer les outils statistiques utilisés à l'échelle globale sur des moyennes régionales.
- Préparer une publication des résultats.

2 Données et méthodologie

2.1 Données

Les données utilisées au cours de ce stage ont été réalisées par différents organismes internationaux dans le cadre du Sixième Projet d'Intercomparaison des Modèles Couplés, CMIP6 [14]. Ces simulations sont notamment exploitées pour préparer le sixième cycle des Assessment Report (AR6) de l'Intergovernmental Panel on Climate Change (IPCC). Plus particulièrement, les données que j'ai utilisées font partie de l'exercice Land Surface, Snow and Soil moisture Model Intercomparaison Project (LS3MIP) [25], qui concerne les simulations relatives aux surfaces continentales. Ces données sont publiques et disponibles sur le site de l'Earth System Grid Federation (ESGF) au format NetCDF (Network Common Data Form) [45]. Ce format auto-descriptif permet de stocker des champs multidimensionnels et des métadonnées.

Pour les besoins de mon étude, j'ai fait le choix de n'utiliser que les données produites par des modèles utilisant, au moins, deux forçages différents pour simuler le cycle du carbone. Par ailleurs, j'ai uniquement considéré les forçages ayant été utilisés par, au moins, deux modèles différents. Ces critères me permettent d'utiliser les simulations de **six modèles** de surface continentale, utilisant jusqu'à **trois forçages atmosphériques distincts**, détaillés ci-après.

2.1.1 Modèles

Les modèles de surface continentale simulant le cycle du carbone sont le plus souvent destinés à être utilisés comme composants de modèle de système terre. Même s'ils n'ont pas été couplés au modèle atmosphérique dans le cadre de ces simulations, ils n'ont parfois pas de nom, et sont donc englobés dans l'ESM associé. Les modèles de surfaces continentales utilisés sont ceux développés par :

CNRM : *Le Centre National de Recherches Météorologiques* (CNRM), avec l'ESM CNRM-ESM2-1 [48] et son modèle de surface, ISBA-CTRIP [11] [12], présenté en introduction. La figure 3 représente la modélisation du cycle du carbone dans la dernière version d'ISBA-CTRIP. Comme les autres modèles de surface, ISBA-CTRIP représente bien sûr les processus d'absorption du carbone par la photosynthèse, les émissions liées à la respiration des organismes autotrophes (les végétaux) et la respiration hétérotrophe du sol, ainsi que le stockage du carbone dans la végétation, dans la litière, et dans le sol. Par rapport à la précédente version, utilisée pour CMIP5, une paramétrisation relativement simple des feux de forêt a été ajoutée. Ces derniers ont en effet un rôle important pour le climat, avec les émissions de carbone et d'aérosols qui leur sont associées. Enfin, les modèles de surface ne représentent pas les végétaux par espèces mais par "types de fonctionnement" (Plant Functional Types (PFT)). Des valeurs caractéristiques, comme par exemple celle du taux de photosynthèse maximale, sont ensuite attribués à chaque type de plante, afin

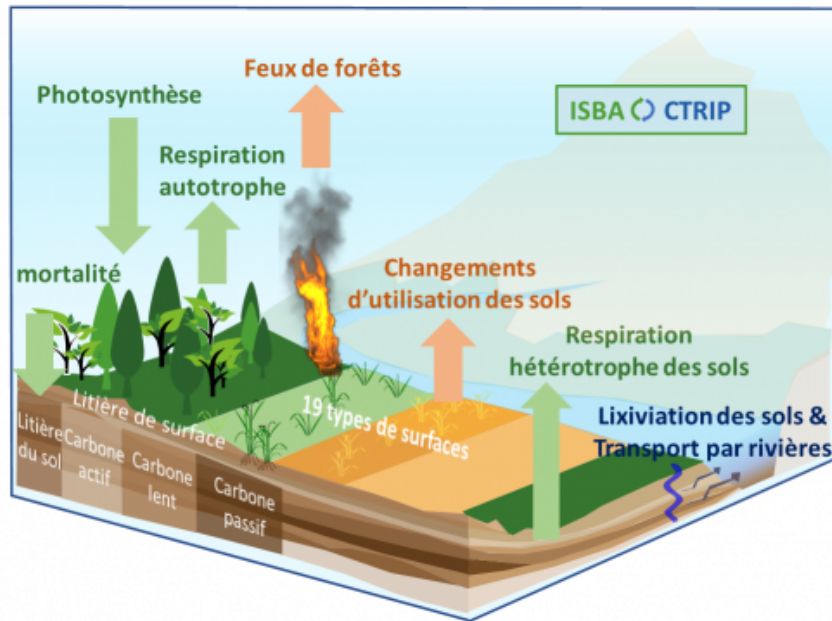


FIGURE 3 – Représentation du cycle du carbone dans le modèle ISBA-CTRIP. Delire et al. (2020) [12].

de représenter les différences physiologiques. Pour ISBA-CTRIP, il y a 16 PFTs possible, 9 arbres, 1 type d'arbuste, 3 types d'herbacées et 3 types de cultures. La grille de surface a une résolution de $1.4^\circ \times 1.4^\circ$.

MPI : Le *Max Planck Institute* for Meteorology (MPI-M), situé en Allemagne, avec MPI-ESM1-2-LR [19]. Le modèle de surface associé est JSBACH [37]. Une différence notable avec ISBA-CTRIP est l'implémentation d'un module de feu de forêt, SPITFIRE [53] [28], qui permet une représentation mécanistique des processus liés aux incendies, ainsi qu'une représentation du cycle de l'azote [18]. Les simulations ont été réalisées en version basse résolution (1.9°).

NCAR : Le *National Center for Atmospheric Research* (NCAR), centre de recherche américain qui développe notamment l'ESM CESM2 en open-source [8]. Le modèle communautaire de surface continentale, CLM5 [30], possède également un module de feu performant [35], prenant en compte un grand nombre de processus et de répercussions [34]. Les données sont disponibles sur une grille latitude-longitude de 0.9° par 1.25° .

BCC : Le *Beijing Climate Center* (BCC) en Chine, avec le modèle BCC-CSM2-MR [56]. Les surfaces sont représentées par BCC-AVIM (Atmosphere-Vegetation-Interaction-Model), qui simule le cycle du carbone, mais ne prend pas en compte les émissions liées aux feux de forêt.

MOHC : Le *Met Office Hadley Center* (MOHC), britannique, avec l'ESM UKESM1-0-LL [49], qui représente de manière précise le cycle du carbone, mais sans simuler les émissions liées aux feux de forêts.

IPSL : Enfin, l'*Institut Pierre Simon Laplace*, à Paris, qui développe IPSL-CM6A-LR [4], et son modèle de surface, ORCHIDEE. Un modèle de feu (SPITFIRE[53]-ORCHIDEE [57, 58]) a été incorporé au modèle de surface, mais les données relatives aux émissions de CO₂ ne sont pas disponibles pour les simulations CMIP6.

2.1.2 Forçages

Les forçages atmosphériques utilisés par les modèles de surface pour leurs simulations proviennent de :

GSWP3 : La troisième version du Global Soil Wetness Project (GSWP3), qui fournit ici des données atmosphériques de 1850 à 2014, avec une résolution de 3h sur une grille de $0.5^\circ \times 0.5^\circ$ [27]. Ce forçage a été construit à partir de plusieurs ensembles de données atmosphériques et de produits de réanalyse [7]. Il est le forçage de base pour les simulations LS3MIP, et a par conséquent été utilisé par tous les modèles.

CRUNCEP : La base de données utilisée est CRU-JRA [21], construite à partir de la combinaison de deux jeux de données préexistants. Les données d'observation mensuelles du Climate Research Unit (CRU), sur une grille de $0.5^\circ \times 0.5^\circ$ [39] couvrant une période de 1901 à 2002, permettent de fournir une moyenne mensuelle pour ajuster les données de la Japanese Reanalysis (JRA) produites par la Japanese Meteorological Agency (JMA). À l'origine, le produit de réanalyse du National Centers for Environmental Prediction (NCEP) (grille de $2.5^\circ \times 2.5^\circ$), couvrant la période 1948–2010 [26] était utilisé, ce qui lui vaut son nom. La version utilisée couvre la période de 1850–2014. Ce forçage a été utilisé par 5 modèles (ceux du CNRM, MPI, NCAR, MOHC et IPSL) sur 6 pour réaliser les simulations.

Princeton : Enfin, le forçage atmosphérique global Princeton [50], développé en combinant le produit de réanalyse du NCEP avec des données d'observations de précipitation, de température et d'éclairement. Les données sont disponibles avec une résolution de 3h sur une grille de $0.5^\circ \times 0.5^\circ$. Les données utilisées pour CMIP6 couvrent une période de 1850 à 2012. Seulement 4 modèles (CNRM, MPI, NCAR, et BCC) ont utilisé ce forçage.

Bien que les méthodes de construction des forçages atmosphériques soient relativement similaires, il existe des différences notables entre ces forçages. La figure 4 présente la variabilité relative autour de certaines variables atmosphériques. On observe, notamment, une variabilité très importante pour l'humidité spécifique et relative.

Nous rappelons que l'humidité relative correspond au rapport de la pression partielle de la vapeur d'eau contenue dans l'air, sur la pression de vapeur saturante, c'est à dire la pression à laquelle les phases gazeuse et liquide sont à l'équilibre, (il y a autant d'évaporation que de

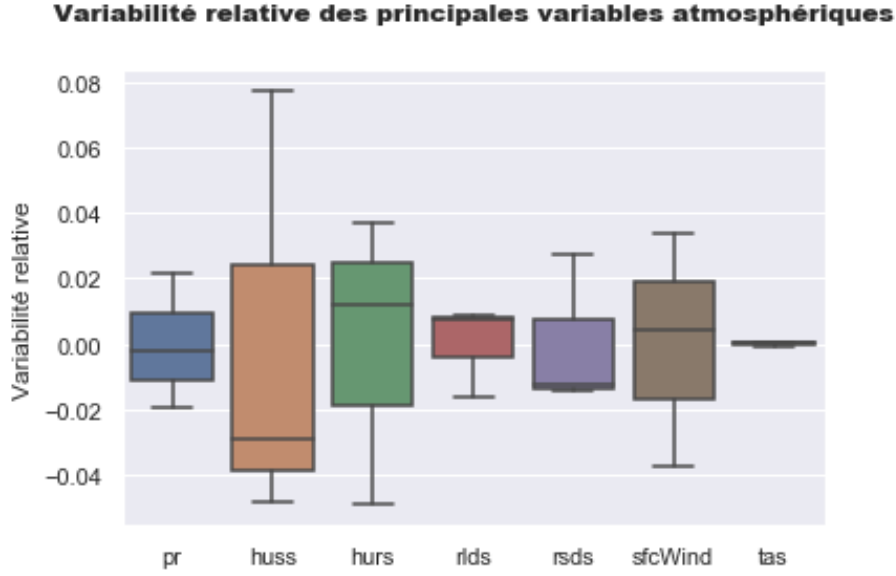


FIGURE 4 – Variabilité relative des variables de précipitation (*pr*), d’humidité spécifique (*huss*), d’humidité relative (*hurs*), de rayonnement incident grande longueur d’onde (*rlds*) et petite longueur d’onde (*rsds*), de vitesse du vent à 10m (*sfcWind*) et de température de surface à 2m (*tas*), moyennées de 2003 à 2012, pour les forçages GSWP3, CRUNCEP et Princeton.

condensation) soit :

$$hurs = \frac{P_{vap}}{P_{sat}(T)} \times 100 \quad (2.1)$$

Pour une même quantité de vapeur d’eau, l’humidité relative augmente lorsque la température diminue.

L’humidité spécifique est définie comme le rapport de la masse d’eau dans l’air sur la masse d’air humide, exprimé en g d’eau par kg d’air. Ainsi, contrairement à l’humidité relative, l’humidité spécifique se conserve avec la température et l’altitude sous réserve que l’eau ne change pas d’état (pas de condensation ou d’évaporation). Elle est néanmoins difficile à mesurer en réalité, et peut être approximée comme :

$$huss \approx \frac{0.622P_{vap}}{P - 0.378P_{vap}} \times 1000 \quad (2.2)$$

La figure 5 présente les valeurs moyennes d’humidité spécifique en point de grille pour la décade 2003-2012 pour le forçage GSWP3, ainsi que la différence d’humidité spécifique entre GSWP3 et CRUNCEP, et GSWP3 et Princeton. On observe que pour le forçage Princeton, l’humidité spécifique est globalement plus élevée que pour GSWP3 et CRUNCEP (GSWP3 semble plus humide au Sahara, mais la végétation et donc l’absorption de CO₂ est quasi nulle dans cette zone). Par ailleurs, l’humidité spécifique est en général plus élevée avec GSWP3 qu’avec CRUNCEP au niveau des forêts équatoriales. Ces tendances se confirment lorsqu’on

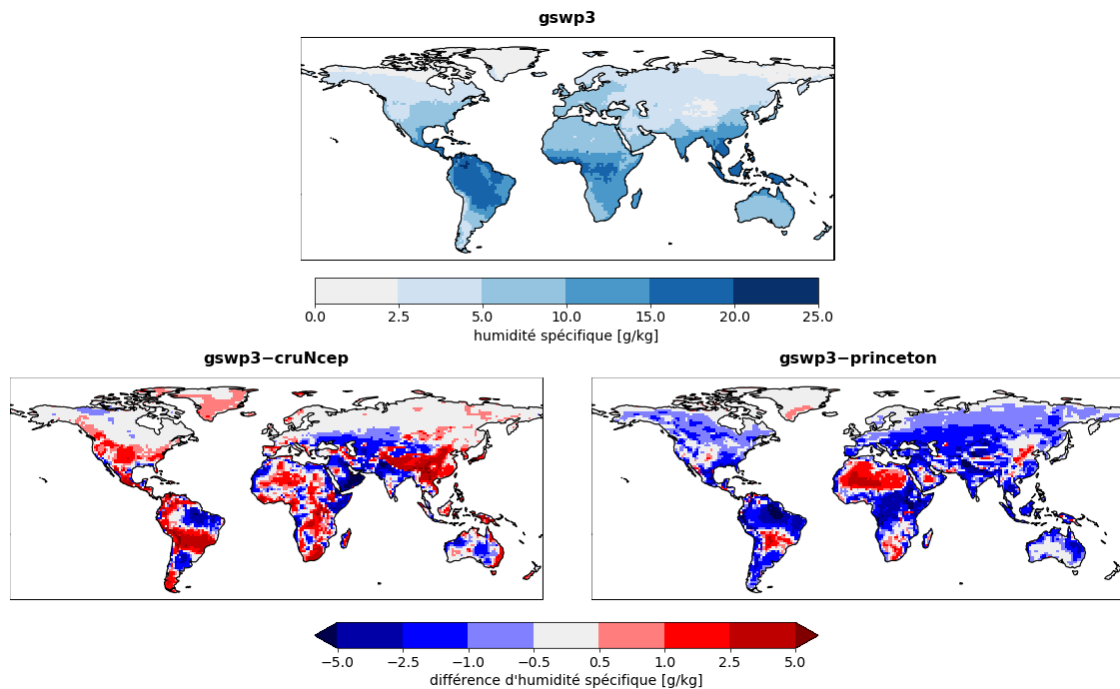


FIGURE 5 – Humidité spécifique moyenne sur la décade 2003-2012 pour le forçage GSWP3. Différence de l'humidité spécifique entre GSWP3 et CRUNCEP (gauche) et GSWP3 et Princeton (droite) pour la décade 2003-2012.

regarde l'humidité relative plutôt que l'humidité spécifique : pour les régions où la végétation est présente, l'humidité est plus élevée avec le forçage Princeton qu'avec GSWP3, qui est généralement plus humide que CRUNCEP. Les différences de précipitation entre les forçages sont assez faibles régionalement, la variabilité relative que l'on peut observer sur la figure 4 pour pr est essentiellement due à des différences de précipitation à l'intérieur du cercle arctique, c'est à dire pour des zones pauvres en végétation. Une des principales différences entre les forçages est donc l'humidité de l'air, qui est nettement plus élevée avec Princeton, et un peu plus élevée avec GSWP3, par rapport à CRUNCEP.

D'autres différences entre les forçages atmosphériques subsistent. Si les températures de surface varient assez peu, il y a des différences notables sur le rayonnement incident aux courtes et grandes longueurs d'onde. On a un rayonnement incident plus élevé aux courtes longueurs d'onde et plus faible aux grandes longueurs d'onde avec CRUNCEP qu'avec GSWP3 et Princeton, ces derniers ayant des valeurs sensiblement équivalentes. Enfin, en ce qui concerne la vitesse du vent, on observe de grandes différences régionales entre les forçages, sans pouvoir discuter d'une tendance globale.

2.1.3 Variables

Le rôle des forçages atmosphériques a été évalué sur une quinzaine de variables relatives au cycle du carbone. Certaines variables comme les émissions liées aux incendies (f_{Fire}) ou

la transpiration ne sont pas disponible pour tous les modèles. Un certain nombre de données était manquantes sur le site de l'ESGF en début de stage, mais nous avons pu contacter les institutions concernées pour les obtenir. On peut distinguer trois grands groupes de variables qui ont été étudiées :

1. Les flux de carbone :

GPP : La *production primaire brute (GPP)*, c'est à dire la quantité de carbone transformée en biomasse par les organismes autotrophes, par unité d'aire et de temps. Pour les plantes, c'est le résultat brut de la photosynthèse.

Ra : La *respiration autotrophe (Ra)* caractérise les émissions de CO₂ liées à la respiration des producteurs primaires.

NPP : La *production primaire nette (NPP)*, c'est la différence entre la production primaire brute et la respiration autotrophe ($NPP = GPP - Ra$). C'est le flux de CO₂ qui est réellement capté chaque année par la végétation.

Rh : La *respiration hétérotrophe (Rh)* correspond à la respiration des micro-organismes du sol qui décomposent la biomasse morte.

f_{Fire} : Les *émissions liées aux incendies (f_{Fire})*.

NEP : La *production nette de l'écosystème (NEP)*, c'est le flux de CO₂ qui est capté par la biosphère continentale ($NEP = NPP - Rh - f_{Fire}$).

f_{Luc} : Les *émissions liées aux changements d'usage des sols (f_{Luc})*. Le carbone stocké dans la biomasse peut être libéré vers l'atmosphère (complètement ou en partie) lors de changements d'usage des sols. À l'inverse certains processus (reboisement de cultures) peuvent stocker du carbone.

NBP : La *production nette du biome (NBP)*, correspond à la *NEP* à laquelle on soustrait l'ensemble des flux anthropiques, comme les flux liés aux changements d'usage des sols (*f_{Luc}*) mais aussi parfois à la récolte des cultures (*f_{Harvest}*). *NBP* représente donc le flux final, réellement absorbé dans la biosphère, lorsqu'on retranche l'ensemble des émissions.

2. Les stocks de carbone :

cVeg : Le *carbone stocké dans la végétation (cVeg)*.

cLitter : Le *carbone stocké dans la litière (cLitter)*, c'est la biomasse morte se trouvant à la surface et dans le sol.

cSoil : Le *carbone stocké dans le sol (cSoil)*.

3. Les échanges d'eau et d'énergie :

tran : La *transpiration (tran)* correspond à l'évaporation de l'eau puisé dans le sol dans les végétaux. Elle se fait essentiellement au niveau des stomates, les pores situés au

niveau de feuilles, où les échanges gazeux avec l’atmosphère sont réalisés. En état de stress hydrique, les végétaux limitent leur transpiration en fermant les stomates, ce qui réduit l’efficacité de la photosynthèse. Les cycles de l’eau et du carbone sont donc intimement liés chez les végétaux.

evpsbl : L’*évapotranspiration (evpsbl)* correspond au flux d’eau transféré vers l’atmosphère par l’évaporation du sol, l’évaporation des précipitations interceptés (neige ou pluie) et la transpiration des végétaux.

hfls : Le *flux de chaleur latente montant à la surface (hfls)*, c’est l’échange d’énergie entre la surface et l’atmosphère permettant le changement d’état de l’eau lors de l’évapotranspiration, que ce soit par sublimation ou vaporisation. Comme, sur les surfaces, la majorité des échanges se fait par vaporisation, le flux de chaleur latente est approximativement proportionnel à l’évapotranspiration.

2.2 Traitement de données

Les stocks et flux de carbone calculés par les modèles se présentent pour la plupart sous forme de données mensuelles, parfois journalières, déterminées pour chaque maille du modèle. La période d’étude peut aller au maximum de janvier 1850 à décembre 2014, parfois séparée en plusieurs fichiers, mais peut être réduite selon les modèles et les forçages. Les fichiers netCDFs récupérés sur le site de l’ESGF contiennent des valeurs de la variable estimée sur trois dimensions (longitude, latitude et temps) et de nombreuses méta-données comme l’institution, le nom de la variable ou encore son unité. Les stocks et flux sont exprimés en unité d’aire, soit respectivement en kg.m^{-2} et $\text{kg.m}^{-2}.\text{s}^{-1}$. Les flux d’énergie (chaleur latente) sont exprimés en W.m^{-2} .

2.2.1 Pré-traitement avec CDO

Un pré-traitement des données a été effectué avec la bibliothèque CDO[46], qui contient de nombreux opérateurs spécifiquement développés pour traiter des données climatiques, notamment au format netCDF. Par exemple, lorsque les données sont contenues dans plusieurs fichiers, la commande `cdo cat` me permet de les concaténer pour obtenir un fichier netCDF qui contient les données mensuelles sur toute la période d’étude. J’ai ensuite réalisé des scripts `shell` me permettant de convertir des données mensuelles (ou journalières le cas échéant) en moyenne annuelle avec `cdo yearmonmean (cdo yeardaymean)`, de manière à masquer la variabilité saisonnière des flux.

Pour l’étude des flux globaux, j’ai d’abord créé, pour chaque modèle, un fichier NetCDF contenant l’aire de surface continentale par maille. Pour cela j’ai utilisé les fichiers `areacella`, qui contiennent l’aire de chaque maille, et `sftlf`, qui contient la fraction de surface continentale occupée par maille. J’ai ensuite multiplié ces masques par les fichiers contenant les moyennes annuelles

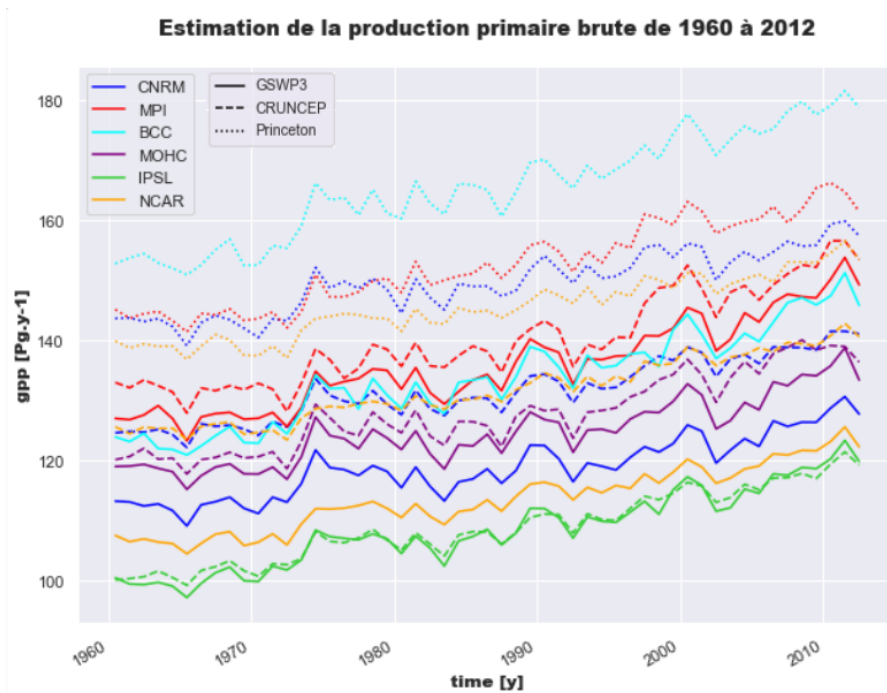


FIGURE 6 – Évolution de la production primaire brute de 1960 à 2012 estimée par les modèles du CNRM (bleu), MPI (rouge), BCC (cyan), MOHC (violet), IPSL (vert) et NCAR (orange) pour les forçages GSWP3 (traits pleins), CRUNCEP (tirets) et Princeton (pointillés).

des variables pour créer de nouvelles données contenant la masse (kg) ou le flux ($\text{kg}\cdot\text{s}^{-1}$) par maille.

Afin de comparer les estimations régionales faites par les modèles, j'ai également utilisé la commande `cdo remapcon` de manière à avoir tous les flux sur la même grille. Le choix a été fait d'utiliser la grille atmosphérique du CNRM ($\sim 1.4^\circ \times 1.4^\circ$), car j'avais facilement accès aux données des forçages atmosphériques et à l'ensemble des produits de simulations réalisés par CNRM-ESM2-1 sous ce format. J'ai donc créé des fichiers NetCDF "remap" pour l'ensemble des estimations des variables réalisées par les autres modèles, à partir des fichiers de moyenne annuelle par unité d'aire. Au besoin, j'ai ensuite pu utiliser le fichier contenant l'aire de surface continentale par maille de la grille du CNRM pour créer de nouveaux fichiers "remap" sans unité d'aire. Ce changement de grille peut néanmoins entraîner une légère perte d'information, par conséquent j'ai préféré utiliser les fichiers avec la grille d'origine plutôt que les fichiers "remap" pour réaliser l'analyse des résultats globaux.

2.2.2 Traitement avec Python

Les données pré-traitées sont ensuite analysées avec Python. Pour exploiter les fichiers netCDF, j'ai utilisé la bibliothèque `xarray`[24], qui permet de conserver les dimensions, le nom des variables et les autres méta-données au lieu de simplement créer un tableau multidimensionnel de valeur. Pour réaliser la somme globale des stocks et flux, j'ai par exemple utilisé

l'opération somme (*sum*) sur les dimensions de longitude et latitude, ce qui me donne un nouveau tableau de données n'ayant plus qu'une seule dimension, le temps. J'ai ensuite converti les unités pour obtenir des stocks et flux globaux respectivement en Pg (10^{15} g) et Pg.y^{-1} (y pour année). J'ai ensuite pu représenter l'évolution temporelle de ces variables avec la bibliothèque *matplotlib*, comme représenté figure 6 et réaliser des opérations simples sur les tableaux de valeur avec *numpy* et *panda*. J'ai par la suite utilisé les bibliothèques *seaborn* et *statsmodels*[47] pour réaliser des analyses et test statistiques sur mes données.

Pour visualiser les variables sous forme de cartes, j'ai choisi d'utiliser la bibliothèque *cartopy* [38] qui présente une grande facilité d'utilisation. Afin de mener une analyse des flux par régions plutôt que par points de grille, j'ai également réalisé un masque sur la grille du CNRM, me permettant d'isoler les 30 régions sub-continentales définies dans le rapport SREX de l'IPCC [52]. Ces régions sont celles partageant un même type de climat, ayant des caractéristiques de surface similaire, et abritant les mêmes types de végétations.

2.2.3 Analyse de la variance (ANOVA)

Après avoir simplement observé le rôle des forçages atmosphériques dans l'estimation des stocks et flux de carbone, j'ai cherché une manière de quantifier l'incertitude attribuable aux forçages, tant à l'échelle globale que régionale. J'ai proposé d'utiliser la méthode de l'ANALYSIS Of VARIANCE (ANOVA) pour y parvenir.

Le but de l'analyse de variance est de comparer des moyennes de groupes, rassemblant différentes valeurs d'une variable *quantitative* (par exemple : la taille à l'âge adulte), et de déterminer si les différences observés entre les moyennes de ces groupes sont dues au hasard ou non. Ces groupes sont en effet définis à partir de facteurs *qualitatifs* (par exemple : le sexe biologique) ayant différents niveaux (masculin, féminin). Le rôle de l'ANOVA est donc de déterminer si des facteurs *qualitatifs* ont une influence sur les variations d'une variable *quantitative* (Le sexe biologique a-t-il une influence sur la taille des individus à l'âge adulte ?) Il est également possible de s'intéresser à l'influence d'autres facteurs (exemple : la nutrition), ayant plusieurs niveaux (carence en vitamines, régime équilibré, surplus de calories) et même d'observer le rôle des interactions qu'il peut y avoir entre ces facteurs.

Dans le cadre de mon stage, deux facteurs vont déterminer la valeur des estimations obtenues. Le **modèle** qui a produit la simulation, puisque les différents choix de paramétrisation des processus vont largement influencer les valeurs obtenues, et bien sûr le **forçage** atmosphérique utilisé, qui est l'objet de notre étude. Pour l'ANOVA à deux facteurs, possédant respectivement I et J niveaux, on utilise le modèle statistique suivant pour exprimer toutes valeurs y_{ij} que l'on obtient lorsque on applique les niveaux i et j des facteurs :

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \quad (2.3)$$

où μ représente une valeur de référence, α_i une valeur réelle représentant l'influence du niveau i du premier facteur, par exemple le **modèle**, β_j une valeur réelle représentant l'influence du niveau j du second facteur, le **forçage**, et ϵ_{ijk} l'erreur résiduelle, soit l'écart par rapport à μ qui n'est pas expliqué par α_i et β_j pour la valeur y_{ij} . On cherche à minimiser la somme des carrés de ces erreurs résiduelles, autrement dit à représenter le mieux possible l'influence des facteurs à travers les termes α_i et β_j .

Cependant, ce modèle statistique ne prend pas en compte les interactions qui peuvent exister entre les facteurs, par exemple le fait que pour un modèle i on obtient de meilleurs résultats avec le forçage j qu'avec j' , alors qu'avec le modèle i' , on obtient un meilleur résultat avec j' qu'avec j . Afin de différencier ces interactions d'une erreur résiduelle liée au hasard, il est nécessaire d'avoir au moins deux valeurs, y_{ijk} et $y_{ijk'}$, pour chaque combinaison modèle-forçage (i, j) . En introduisant cet indice de répétabilité k , on peut maintenant écrire :

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \quad (2.4)$$

avec $(\alpha\beta)_{ij}$ la valeur réelle représentant l'interaction du couple (i, j) . En pratique, on ne représente ces interactions que si elles sont significatives, sinon, on utilise le modèle statistique décrit par l'équation 2.11. La significativité du rôle des interactions, tout comme celle des facteurs, peut être déterminée avec un test de Fisher [6], que nous ne détaillerons pas dans ce rapport.

Afin de déterminer le pourcentage de variabilité expliqué par le modèle statistique, on peut utiliser la somme des carrés des écarts à la moyenne, (SCE), c'est à dire la variance multipliée par le nombre de degrés de liberté (le nombre d'interaction entre les groupes, ce qui correspond au nombre de niveaux de facteurs moins 1). Dans le cas où l'on a deux facteurs à I et J niveaux, avec K valeurs par combinaison, la somme des carrés des écarts à la moyenne totale (SCE_T), autrement dit la dispersion totale autour de la moyenne, vaut :

$$SCE_T = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (y_{ijk} - \bar{y}_{...})^2 = \sum_{ijk} (\bar{y}_{ij.} - \bar{y})^2 + \sum_{ijk} (y_{ijk} - \bar{y}_{ij.})^2 \quad (2.5)$$

avec \bar{y} la moyennes de l'ensemble des valeurs y_{ijk} , et $\bar{y}_{ij.}$ la valeur moyenne obtenue pour le modèle i avec le forçage j . La part de la dispersion qui n'est pas expliquée par le modèle, la somme des carrés des écarts à la moyenne résiduelle (SCE_R), est :

$$SCE_R = \sum_{ijk} (y_{ijk} - \bar{y}_{ij.})^2. \quad (2.6)$$

Par conséquent, pour la part de dispersion expliquée par le modèle statistique, SCE_M , on

a :

$$SCE_M = SCE_T - SCE_R = \sum_{ijk} (\bar{y}_{ij.} - \bar{y})^2. \quad (2.7)$$

On définit alors R^2 , le pourcentage de variabilité expliqué par le modèle, comme :

$$R^2 = \frac{SCE_T - SCE_R}{SCE_T} = \frac{SCE_M}{SCE_T}. \quad (2.8)$$

Dans le cas d'un plan *complet* et *équirépeté*, c'est à dire que pour chaque combinaison (i, j) , il existe un même nombre k de répétitions, on peut décomposer SCE_M comme :

$$\sum_{ijk} (\bar{y}_{ij.} - \bar{y})^2 = \sum_{ijk} (\bar{y}_{i..} - \bar{y})^2 + \sum_{ijk} (\bar{y}_{.j.} - \bar{y})^2 + \sum_{ijk} (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y})^2 \quad (2.9)$$

avec $\bar{y}_{i..}$ la valeur moyenne obtenue avec le modèle i , et $\bar{y}_{.j.}$ la valeur moyenne pour le forçage j . Le premier terme correspond donc à la somme au carré des écarts entre les moyennes obtenues pour les modèles i et la moyenne globale, soit SCE_{mod} , le second à la somme au carré des écarts entre les moyennes obtenues pour les forçages j et la moyenne globale, SCE_{for} , et le dernier à la somme au carré des écarts à la moyenne des interactions, SCE_{int} . On peut alors réécrire 2.5 comme :

$$SCE_T = SCE_M + SCE_R = SCE_{mod} + SCE_{for} + SCE_{int} + SCE_R \quad (2.10)$$

Par conséquent, on peut maintenant calculer la variabilité expliquée par les modèles, les forçages, les interactions, ou les résidus, soit :

$$R_X^2 = \frac{SCE_X}{SCE_T} \quad (2.11)$$

Pour mettre en pratique la théorie de l'ANOVA, j'ai du définir un nouveau plan d'expérience, de manière à ce qu'il soit *complet*, *équirépeté*, avec un taux de répétitions supérieur à 1. En effet, mon plan n'est pas complet, puisque tous les modèles n'ont pas utilisé trois forçages différents pour réaliser leurs simulations, et certaines variables ne sont pas estimées par tous les modèles, comme indiqué sur le tableau 1. Par ailleurs, je ne possède qu'une seule simulation par modèle, $K = 1$.

Une solution pour obtenir plusieurs valeurs par simulation, et donc par combinaison (i, j) , aurait pu être de réaliser l'analyse de variance sur les valeurs de la série temporelle, et donc d'avoir un indice de k années. L'inconvénient majeur de cet approche est que si je ne considère qu'un petit nombre K d'années, l'erreur résiduelle augmente du fait de la variabilité inter-annuelle. Si au contraire je considère un grand nombre K , l'influence de la variabilité

	CNRM	MPI	NCAR	BCC	MOHC	IPSL
GSWP3	Complet	Suffisant	Complet	Incomplet	Incomplet	Incomplet
CRUNCEP	Complet	Suffisant	Suffisant	Absent	Incomplet	Incomplet
Princeton	Complet	Suffisant	Suffisant	Incomplet	Absent	Absent

TABLE 1 – Données disponibles pour les différentes combinaisons forçages-modèles. "Complet" indique que des données sont disponibles pour l'ensemble des variables, "suffisant" que les variables majeures ont toutes été estimées, "incomplet" que certaines variables majeures n'ont pas été estimées, et "absent" qu'aucune variable n'est disponible.

inter-annuelle diminue, mais l'erreur résiduelle devient encore plus importante du fait de l'augmentation de la concentration de CO₂ atmosphérique, qui tend à augmenter la valeur des flux estimés entre l'année 1 et l'année K . L'approche que j'ai choisie est donc de réaliser mon analyse de variance sur les valeurs moyennes des simulations, sur une période donnée, et de faire l'hypothèse que la moyenne d'une autre simulation, utilisant le même modèle et le même forçage, donnerait des résultats similaires. Cette hypothèse me semble acceptable puisque les modèles sont exécutés en mode forcé, il n'y a pas de variabilité interne. J'ai par ailleurs pu confirmer cette hypothèse sur le modèle du MPI qui a effectué deux simulations en utilisant un même forçage atmosphérique, et qui présentait une différence infime entre les produits des deux simulations.

Deux possibilités s'offraient à moi pour la réalisation d'un plan complet. La première, que j'ai mise en place pour tester l'analyse de variance, a été de compléter les simulations manquantes par les valeurs moyennes que nous aurions probablement obtenues, au vu des simulations réalisées par les autres modèles et de la tendance observée. En effet, en calculant les écarts moyens entre deux estimations du même modèle, (par exemple celles forcées par Princeton et GSWP3), on peut compléter les valeurs manquantes pour les modèles n'ayant pas réalisé de simulation avec un des forçages (pour Princeton ici). Cette approche permet d'avoir une bonne idée de la variabilité attribuable aux facteurs. Néanmoins, elle représente mal les interactions qu'il peut y avoir, puisque même si on observe une tendance pour les simulations utilisant le même forçage, on ne peut pas prévoir comment un modèle particulier va réagir.

La seconde option a été permise plus tard au cours de mon stage, lorsque les produits de simulations réalisés par le NCAR ont été publiés. J'ai alors choisi de réduire mon plan d'expérience, en n'utilisant que les modèles ayant utilisés trois forçages atmosphériques différents, à savoir les modèles du CNRM, du MPI et du NCAR. Par ailleurs, ces modèles ont été les seuls à publier des données pour les émissions de carbone liées aux feux de forêt, qui rentrent pourtant dans le calcul des flux nets absorbés par la biosphère. Cette approche m'a permis de mieux représenter les interactions, et de réaliser l'ANOVA sur l'ensemble des variables d'intérêt. De plus, cela me permet d'avoir le même nombre de niveaux (3) et donc de degrés de liberté (2) par facteurs. En contrepartie, cette approche ne me permet pas d'exploiter toutes les données à ma disposition et peut rendre mes résultats moins significatifs.

3 Résultats

La section suivante présente les principaux résultats obtenus durant ce stage, en se concentrant sur la question des flux de carbone et plus particulièrement de la production primaire et du puits de carbone continental.

Les résultats obtenus pour les stocks de carbone sont présentés en annexe. En effet, bien qu'étant une résultante des flux, le stockage du carbone dépend beaucoup des choix de paramétrisation faits par les modèles, en particulier pour la représentation de la litière et du sol. Par conséquent, la variabilité due aux modèles est très importante pour les stocks, et la variabilité due aux forçages est directement explicable par la variabilité observée sur les flux.

3.1 Les flux de carbone globaux

Pour chacun des flux de carbone étudiés, j'ai dans un premier temps choisi de représenter la dispersion autour des modèles et des forçages à l'échelle globale, avant de réaliser une analyse de variance. Dans un premier temps, je présente les résultats obtenus pour la production primaire brute, en tenant compte des résultats intermédiaires pour expliquer la démarche effectuée sur l'ensemble des flux. Par la suite, je présente uniquement les résultats finaux des autres flux de carbone avant de discuter du puits de carbone continental.

3.1.1 Exemple de la production primaire brute

Pour l'ensemble des variables, j'ai choisi de travailler sur la période allant de janvier 1960 à décembre 2012. J'ai choisi de commencer en 1960 car les données du modèle du BCC semblaient présenter des anomalies avant 1950, et parce que les résultats du GCP commençaient en 1960, ce qui facilitait la comparaison des résultats. Notons également que ces dates coïncident avec les premières mesures de CO₂ atmosphérique qui ont commencé en 1957 à Mauna Loa, sur l'archipel d'Hawaï, et sont depuis complétées par d'autres stations et accessibles sur le site de la *National Oceanic and Atmospheric Administration* (NOAA) [13].

J'ai choisi de m'arrêter en 2012 plutôt qu'en 2014, car le forçage atmosphérique Princeton s'arrête à cette date. Cela me permet d'avoir des simulations de durée équivalente, ce qui facilite les opérations de comparaison que je peux effectuer entre les données. Par exemple, j'ai dans un premier temps moyenné entre elles les simulations provenant d'un même modèle, et calculé pour chaque pas de temps l'écart type associé. La figure 7 présente les estimations moyennes des différents modèles pour la production primaire brute et l'intervalle de confiance à $\pm 1\sigma$ (68%). On peut remarquer, tout d'abord, que ces "simulations moyennes" ne se superposent pas, il y a donc bien un rôle des modèles, lié aux choix de paramétrisation des processus, dans l'estimation de la production primaire brute. Ensuite, on remarque que pour la plupart des modèles, l'écart type autour de ces moyennes semble significatif, puisque il est de l'ordre de grandeur de l'écart

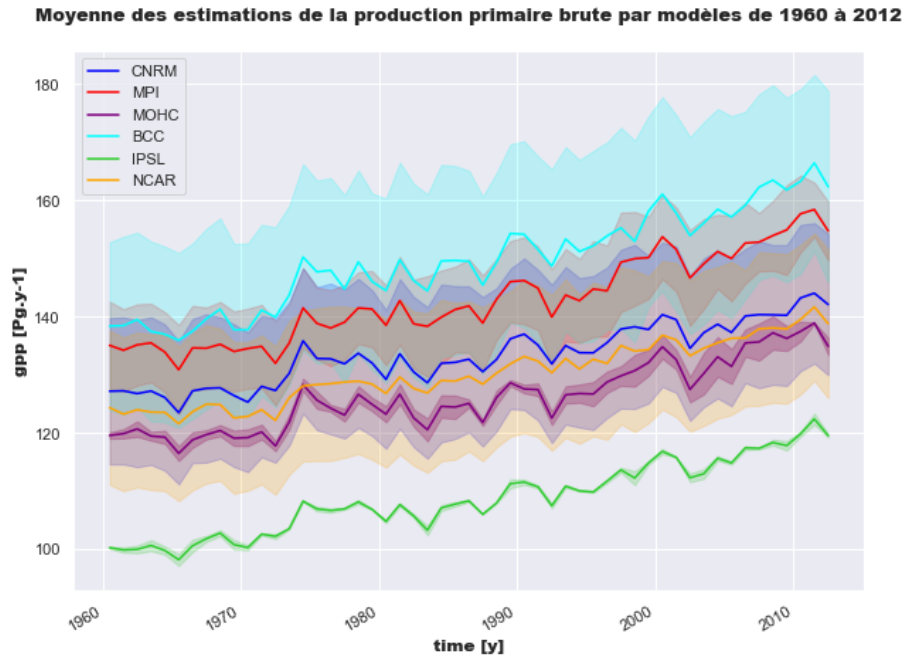


FIGURE 7 – Moyenne des estimations de l'évolution de la production primaire brute pour les modèles du CNRM (bleu), MPI (rouge), BCC (cyan), MOHC (violet), IPSL (vert) et NCAR (orange).

entre les moyennes. Il y a donc, au moins, une autre source de variabilité entre les simulations, à priori les forçages atmosphériques. Enfin, on voit que les écarts types dépendent des modèles considérés. Par exemple, les modèles du BCC, du CNRM ou du NCAR semblent présenter un grand écart type et leurs intervalles de confiances se chevauchent, tandis que les modèles de l'IPSL et dans une moindre mesure du MOHC présentent très peu de variabilité autour de leur valeur moyenne, les simulations produites par le modèle de l'IPSL étant particulièrement invariantes et isolées des autres.

De manière similaire, j'ai moyenné entre elles les simulations utilisant le même forçage atmosphérique et calculé l'écart type associé, comme on peut le voir figure 8, toujours pour la GPP. Là encore, les estimations moyennes ne se superposent pas, ce qui met en valeur le rôle des forçages dans la variabilité. Les simulations réalisés avec Princeton se distinguent particulièrement des autres, avec une valeur moyenne nettement plus élevée, un écart type plus faible, et un intervalle de confiance qui ne chevauche pas celui des simulations réalisées avec CRUNCEP et GSWP3. Ces deux dernières semblent par ailleurs avoir des valeurs plus proches, bien que la valeur moyenne obtenue avec CRUNCEP est un peu plus importante que celle obtenue avec GSWP3.

Pour expliquer les variations dans l'estimation de la GPP selon les forçages atmosphériques, on peut se rappeler les différences entre les variables météorologique des forçages, présentées figure 4. En particulier, nous avons constaté que le forçage atmosphérique Princeton était, globalement, plus humide. Cela a tendance à favoriser la photosynthèse, car la différence de potentiel hydrique entre l'air et le sol est moindre, et les plantes ont moins besoin de fermer leurs

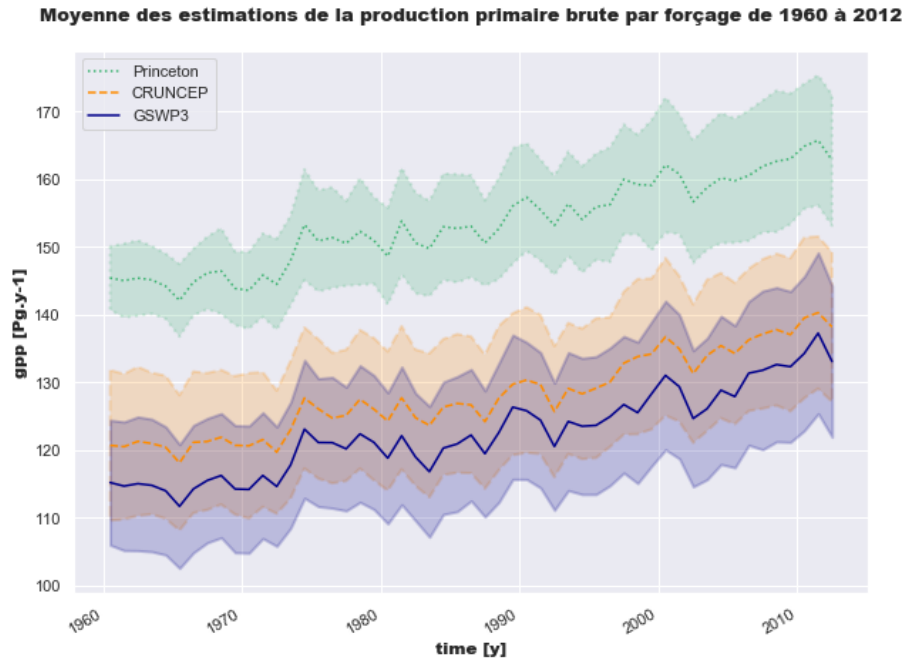


FIGURE 8 – Moyenne des estimations de l'évolution de la production primaire brute pour les forçages atmosphériques GSWP3 (trait plein bleu foncé), CRUNCEP (tirets oranges) et Princeton (pointillés verts) à partir des données des six modèles de surface.

stomates pour retenir la même quantité d'eau. Cela nous fournit une hypothèse convaincante pour justifier les valeurs plus élevées des simulations réalisées en utilisant le forçage atmosphérique Princeton. Pour CRUNCEP et GSWP3, bien que le deuxième soit un peu plus humide que le premier, on peut noter que le rayonnement aux courtes longueurs d'onde, est plus important pour CRUNCEP. Ce sont ces photons, plus énergétiques, qui sont utilisés par les végétaux pour réaliser la photosynthèse. Cette différence pourrait donc justifier l'écart observé pour les estimations de la GPP. Néanmoins, rappelons que de nombreux facteurs atmosphériques peuvent influencer le cycle du carbone, et qu'il existe des relations complexes entre ces variables. Ces justifications restent donc hypothétiques et nécessiteraient un travail plus approfondi, qui dépasse le cadre de ce rapport.

Il est difficile d'explicitier les différences dans les estimations moyennes de la GPP globale entre les modèles, qui sont liées à des choix de paramétrisation auxquels nous n'avons pas toujours accès. On peut néanmoins s'interroger sur les différences entre les écarts types obtenus pour les moyennes des modèles, autrement dit : la sensibilité des modèles aux forçages atmosphériques. Un modèle ayant un faible écart type, comme par exemple celui du MOHC ou de l'IPSL, semble à première vue peu sensible aux conditions climatiques pour estimer la production primaire brute, au contraire des modèles du BCC, du CNRM, du NCAR, et dans une moindre mesure, du MPI.

Néanmoins, ces résultats sont à relativiser car, comme nous avons vu, tous les modèles n'ont pas utilisé les trois forçages atmosphériques, et en particulier Princeton, le forçage qui

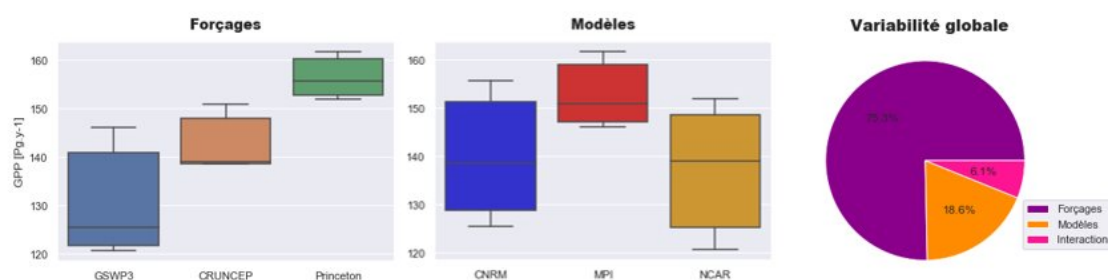


FIGURE 9 – Dispersion de la moyenne temporelle de la production primaire brute selon les forçages (à gauche), selon les modèles (au centre) et pourcentage de variabilité expliqués par les forçages, les modèles et les interactions pour la production primaire brute (à droite).

se distingue le plus des deux autres, n'a pas été utilisé par l'IPSL et le MOHC. Par ailleurs, ces modèles peuvent être plus sensible à l'humidité (on s'attendrait dans ce cas à obtenir des valeurs très importante avec Princeton), ou moins sensible au rayonnement incident que les autres modèles, ce qui aurait tendance à rapprocher les estimations obtenues avec GSWP3 et CRUNCEP et justifier les résultats obtenus.

Ce mauvais équilibre entre les forçages atmosphérique utilisés par les modèles peut aussi, par exemple, expliquer l'écart type plus petit autour de la moyenne des simulations utilisant Princeton, puisque on ne prend en compte que quatre simulations, et qu'en particulier on ne considère pas l'IPSL, qui a une valeur moyenne assez éloigné des autres modèles. (Mais une simulation forcée par Princeton pour le modèle de l'IPSL aurait justement, à priori, rehaussé cette moyenne). C'est également à cause de ces problèmes de représentativité des données que j'ai choisi de définir un nouveau plan d'expérience, complet et équilibré, pour réaliser mes comparaisons et en particulier l'ANOVA.

La figure 9 représente la dispersion de la moyenne temporelle (1960-2012) de la production primaire brute, autour des forçages et des modèles, pour les simulations du nouveau plan d'expérience, sous la forme de boîte à moustache. On observe ainsi encore plus clairement les différences entre les forçages, en particulier que l'estimation de la GPP est plus élevée avec Princeton qu'avec CRUNCEP, qui donne dans la majorité des cas des valeurs plus importantes qu'avec GSWP3. Pour les modèles, on voit que les médianes sont relativement proches, mais que la dispersion est importante, en particulier pour les modèles du CNRM et NCAR par rapport à celui du MPI. Cela montre que le rôle des forçages atmosphériques est peut être plus important que celui des modèles dans l'estimation de ce flux, et que les modèles ne présentent pas tous la même sensibilité aux forçages, il existe donc bien des termes d'"interactions" qui expliquent la variabilité. Le diagramme à droite de la figure 9 présente le pourcentage de variabilité expliqué par les forçages atmosphériques, les modèles ayant réalisés les simulations, et les termes d'interaction, obtenus à partir des rapports SCE_X/SCE_T en se basant sur les équations 2.10, 2.11 de l'ANOVA. On constate que dans le cas de la GPP moyenne, sur la période

1960-2021, c'est 75% de la variabilité qui est expliqué par les forçages atmosphériques, 19% par les modèles et 6% par les termes d'interactions. On obtient des ordres de grandeur relativement similaire lorsqu'on s'intéresse à des périodes plus courtes comme par exemple la GPP moyenne sur la décade de 2003 à 2012. On peut donc en conclure que les forçages atmosphériques ont un rôle significatif, et même prépondérant, dans l'estimation de la production primaire brute globale.

3.1.2 Respiration et production primaire nette

Nous avons montré que les forçages atmosphériques jouent un rôle important dans l'estimation de la production primaire brute, c'est à dire sur la quantité de carbone assimilé, par an, par la photosynthèse. On peut également s'interroger sur le rôle des forçages atmosphériques dans l'estimation des émissions de carbone liée à la respiration des plantes et du sol. Nous présenterons ici directement les résultats sous forme de boîtes à moustaches, pour les valeurs moyennes des flux entre 1960 et 2012.

La figure 10 présente la dispersion de la respiration autotrophe R_a , de la production primaire nette NPP, et de la respiration hétérotrophe R_h selon les forçages. On observe une différence nette entre les forçages pour R_a , avec toujours les valeurs les plus élevés obtenues avec Princeton et les plus faibles avec GSWP3. La tendance observée est relativement similaire pour NPP et R_h , mais l'écart entre les médianes est moins important et la dispersion inter quartile plus grande. Les tendances observées pour NPP et R_h sont très similaires, ce qui est peu surprenant puisque à l'équilibre, ces flux sont égaux. Il est assez normal qu'on retrouve cette tendance selon les forçages atmosphériques, compte tenu des différences observées entre les variables atmosphériques et des résultats obtenus pour la GPP. D'une part, parce que si la production primaire brute est plus importante, la quantité de biomasse augmente, donc le flux associé à la respiration de cette biomasse, R_a , augmente également. Par ailleurs, il y a aussi une plus grande quantité de biomasse qui meurt et se décompose, entraînant une augmentation de R_h . D'autre part, parce que des conditions plus humides vont avoir tendance à augmenter la respiration, en particulier celle du sol, en accélérant la vitesse de décomposition. Cela peut également expliquer

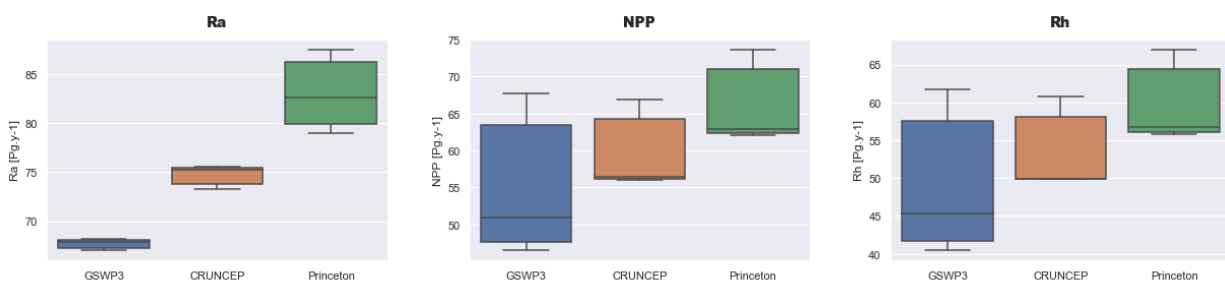


FIGURE 10 – Dispersion de la moyenne temporelle (1960-2012) selon les forçages de la respiration autotrophe (R_a) à gauche, la production primaire nette (NPP) au centre, et la respiration hétérotrophe (R_h) à droite.

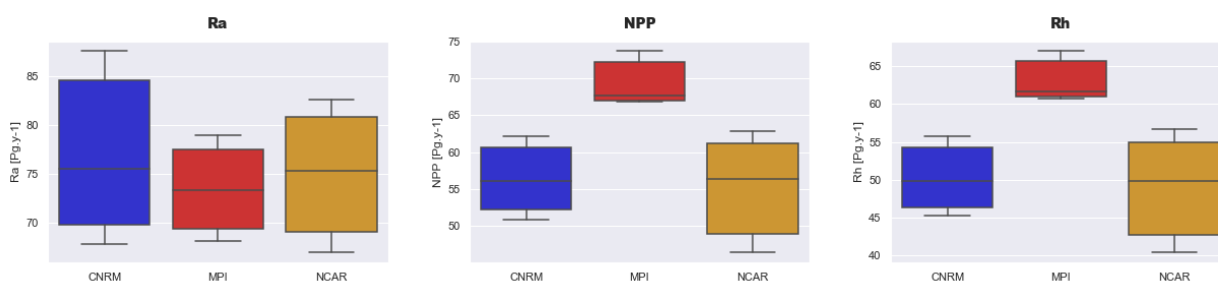


FIGURE 11 – Dispersion de la moyenne temporelle (1960-2012) selon les modèles de la respiration autotrophe (Ra) à gauche, la production primaire nette (NPP) au centre, et la respiration hétérotrophe (Rh) à droite.

que le rôle des forçages soit moins important pour NPP (et Rh) que pour Ra et GPP, puisque des conditions atmosphériques favorables au développement de la végétation augmentent aussi bien la GPP que Ra, ces effets vont en partie se compenser lorsque on calcule $NPP = GPP - Ra$.

La figure 11 présente quant à elle la dispersion des flux pour les modèles. On remarque que les médianes sont encore très proches pour Ra et que le modèle du CNRM semble plus sensible aux forçages atmosphérique que celui du NCAR, qui l'est encore une fois plus que celui du MPI. Par contre, on observe une différence plus nette entre les médianes pour NPP et Rh, ce qui indique une variabilité liée au modèle plus importante. On remarque aussi une moins grande sensibilité aux forçages pour NPP et Rh, surtout pour le modèle du CNRM, qui comme nous l'avons vu peut être expliquée par les effets antagonistes des forçages qui ont tendance à se compenser dans le calcul des flux nets. Les résultats de l'ANOVA présentés figure 12 montrent clairement que les forçages atmosphériques jouent un rôle prépondérant (presque 90%) dans la variabilité de la respiration autotrophe, mais ils ne sont responsables que d'environ 30% de cette variabilité pour NPP et Rh contre plus de 60% expliquée par les modèles. Le rôle des interactions reste aux alentours de 6% pour l'ensemble des flux. Même si les forçages atmosphériques ne représentent plus toujours la première source de variabilité, leur rôle reste indéniablement significatif dans l'estimation des principaux flux de carbone par les modèles de surface.

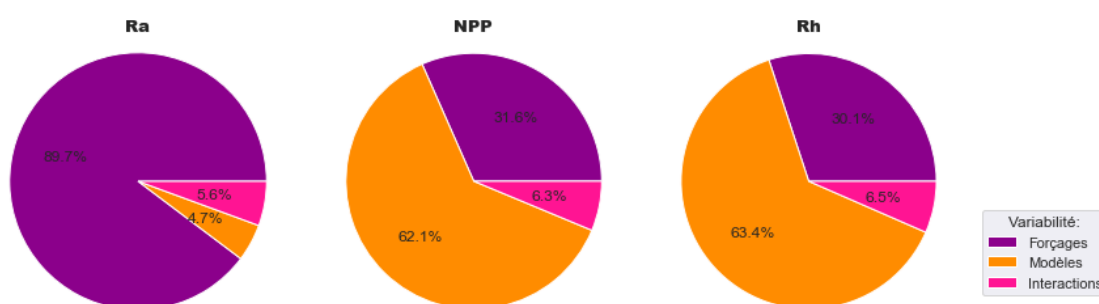


FIGURE 12 – Pourcentage de variabilité explicable par les forçages, modèles et interactions pour la respiration autotrophe (Ra) à gauche, la production primaire nette (NPP) au centre, et la respiration hétérotrophe (Rh) à droite.

3.1.3 Le puits de carbone continental

Nous allons à présent nous intéresser au rôle des forçages atmosphériques dans l'estimation du puits de carbone continental. Nous pouvons dans un premier temps nous interroger sur la définition de ce puits. En effet, le GCP le définit comme le flux net de carbone entrant dans la végétation, en excluant tous les processus d'origine anthropique liés à l'usage et au changement d'usage des sols. Il est estimé en utilisant un masque invariant pour le couvert végétal correspondant à l'état pendant l'âge pré-industriel, le flux net estimé par les modèles étant ainsi affranchi des changements anthropiques, et est appelé S_{Land} . Cependant, les simulations CMIP6 que j'ai utilisées n'utilisent pas ce protocole, et représentent les changements d'usage des sols ainsi que les émissions associées f_{Luc} . Il est néanmoins possible de se rapprocher de l'estimation de S_{Land} réalisée par le GCP, soit à partir de la productivité nette du biôme NBP auxquels on ajoute f_{Luc} , soit à partir de la productivité nette de l'écosystème NEP qui correspond plutôt bien à la définition de S_{Land} , soit en calculant directement le flux $GPP - Ra - Rh - f_{Fire}$. Selon les définitions données par CMIP6, ces trois quantités sont supposés être équivalentes. Cependant, en pratique, j'ai pu constater des différences, essentiellement dues à des problèmes d'interprétation des définitions. J'ai donc décidé, pour chaque modèle, de recalculer ce puits de carbone avec la troisième méthode, en partant de la production primaire et en retranchant les émissions de carbone naturelles.

J'ai commencé par m'intéresser aux flux $GPP-Ra-Rh$, car tous les modèles n'ont pas estimé les émissions liées aux feux de forêt. Ce flux prend en compte la plus grande partie des échanges de carbone terrestre, et peut donner une première idée sur la variabilité dans l'estimation de S_{Land} . La figure 13 représente la dispersion de la moyenne temporelle (1960-2012) de ce flux, pour les forçages et les modèles, ainsi que le pourcentage de variabilité expliqué par les différents facteurs. On observe toujours que les valeurs estimées sont plus élevées avec Princeton qu'avec GSWP3, mais la différence avec CRUNCEP est moins évidente. Cela peut être expliqué par le fait que CRUNCEP est moins humide que Princeton, et donc que la respiration hétérotrophe est moins importante pour une même quantité de biomasse morte. La médiane des modèles est très proche, et la dispersion inter quartile importante. Le diagramme représentant la répartition de la variabilité montre que le rôle des modèles est peut important pour ce flux (7%) contrairement aux forçages (52%). On voit cependant que les interactions représentent 41% de la variabilité, alors que les modèles semblent tous trois aussi sensibles aux forçages si l'on se réfère uniquement aux boîtes à moustaches. Cette valeur est pourtant parfaitement explicable si on observe les valeurs moyennes obtenues pour chaque simulation pour la période 1960-2012 donné dans le tableau 2. Pour le modèle du CNRM, on trouve une valeur légèrement plus forte avec Princeton qu'avec CRUNCEP, mais ces deux valeurs sont bien supérieures à GSWP3. Pour le modèle du MPI, la valeur obtenue avec Princeton est largement supérieure aux deux autres, alors que pour le NCAR, c'est avec CRUNCEP qu'on obtient la valeur la plus importante. Ainsi, pour certains

modèles (NCAR, CNRM), l'humidité va beaucoup augmenter la respiration et la décomposition de la matière organique, au point que même si la production primaire est plus importante avec Princeton, on obtient des résultats similaires voir supérieurs avec CRUNCEP pour ce flux. Ces résultats soulignent des désaccords entre les modèles, sur la réponse du puits de carbone continental, lorsqu'on fait varier les conditions atmosphériques. Afin de prévoir l'évolution de la concentration de CO_2 atmosphérique, il est pourtant nécessaire de pouvoir estimer l'évolution de ce puits dans le contexte du changement climatique.

	GSWP3	CRUNCEP	Princeton
CNRM	5,67	6,23	6,33
MPI	6,01	6,06	6,68
NCAR	6,03	6,53	6,14

TABLE 2 – Valeurs moyennes de GPP-Ra-Rh sur la période 1960-2012 pour chaque simulation.

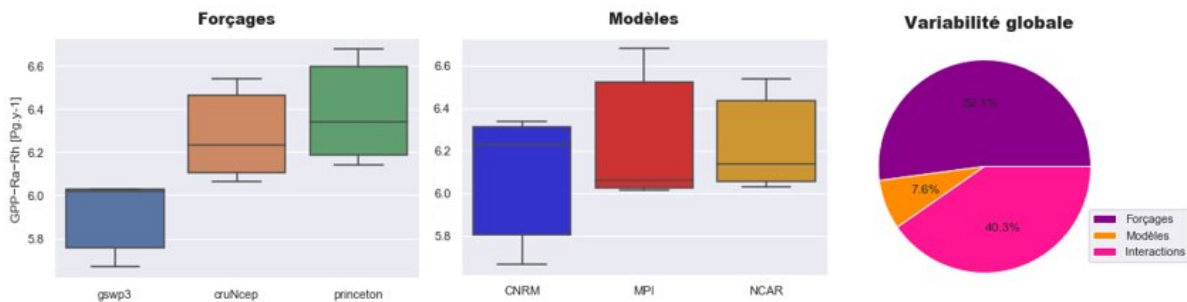


FIGURE 13 – Dispersion de la moyenne temporelle de GPP-Ra-Rh selon les forçages (à gauche), selon les modèles (au centre) et pourcentage de variabilité expliqués par les forçages, les modèles et les interactions pour le flux GPP-Ra-Rh (à droite).

Nous allons à présent nous intéresser aux émissions liées aux feux de biomasses, qui rentrent dans le calcul de S_{Land} . Ces incendies, qui ne concernent pas les feux liés à l'activité humaine, émettent entre 2 et 3 Pg de carbone par an [54], mais présentent une forte variabilité inter-annuelle. En plus du CO_2 , ces feux de forêts émettent d'autres types de gaz et d'aérosols[1], ce qui leur confère un rôle important dans le système climatique. Plusieurs techniques ont été développées pour estimer les émissions liées aux feux de biomasse, notamment des bases de données issues d'observations satellites comme la Global Fire Emissions Database (GFED) [44], complétées par différents indicateurs comme les concentrations de certains gaz dans la troposphère, dans la glace, ou encore en se basant sur le charbon sédimentaire pour obtenir des jeux de données étendus jusqu'au début de l'âge industriel. Une telle base de données (BB4CMIP), commençant en 1750, a été développée pour CMIP6 [55], en utilisant également les estimations moyennes de six modèles de feu. Ces simulations ont été réalisées dans le cadre du projet de comparaison des modèles de feu FireMIP [43], auxquels ont notamment participé les modèles du NCAR [35], du MPI (JSBACH-SPITFIRE [29]) et de l'IPSL (ORCHIDEE-SPITFIRE [57, 58]) en utilisant le forçage atmosphérique CRUNCEP.

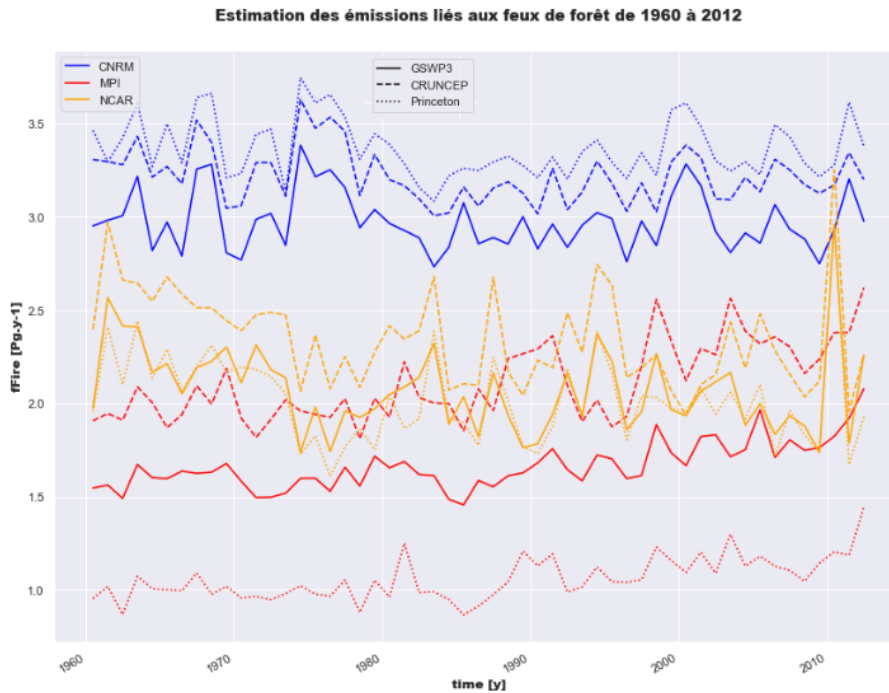


FIGURE 14 – Évolution temporelle des émissions liées aux feux de forêt pour les simulations des modèles du CNRM (bleu), du NCAR (orange) et du MPI (rouge) forcé par GSWP3 (trait plein), CRUNCEP (tireté) et Princeton (pointillé).

La figure 14 présente les estimations des émissions de carbone liées aux feux de forêts réalisées par les modèles du CNRM, du MPI et du NCAR pour différents forçages. On peut observer une grande différence entre les modèles, et, en particulier, des valeurs qui semblent anormalement élevées pour le modèle du CNRM, et un peu trop faible pour le MPI. Pour expliquer ces différences, on peut rappeler que le modèle de feu d'ISBA-CTrip est relativement simple, et qu'il ne calcule pas des occurrences de feu, mais des fractions de biomasse brûlée par réservoir et par grille, qui dépendent de la résistance au feu de la biomasse considérée [12]. Pour JSBACH-SPITFIRE, une explication avancée par le MPI met en cause des interactions inattendues entre le cycle de l'azote, la dynamique de la végétation et le module de feu SPITFIRE. En effet, SPITFIRE permet de recalculer la concentration d'azote organique dans l'atmosphère et dans le sol, qui diminue après un incendie. Comme le modèle JSBACH prend en compte la limitation par l'azote, il en résulte que la productivité primaire baisse dans les zones où les incendies sont récurrents, ce qui conduit à une fraction de biomasse moins importante, donc, par rétroaction, moins d'incendies, et finalement des émissions de CO_2 et des surfaces brûlées sous-estimées par rapport aux observations, comme le rapporte l'article de Mauritsen et al. [37]. Ces grandes différences entre les valeurs moyennes des modèles leur confèrent un rôle très important (85%) dans la variabilité de ces émissions. Pourtant, on peut remarquer visuellement un rôle des forçages et des interactions. Le modèle du CNRM semble suivre la tendance déjà observée pour la production primaire et la respiration : comme il y a plus de biomasse produite avec Princeton qu'avec CRUNCEP et GSWP3, il y a également plus de biomasse brûlée et donc

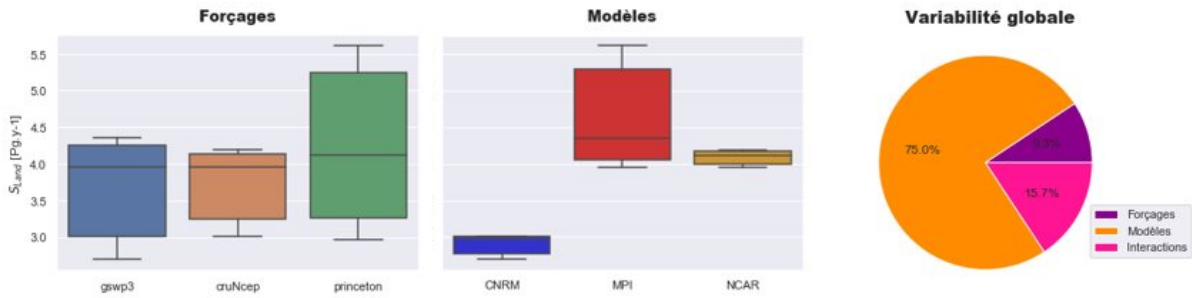


FIGURE 15 – Dispersion de la moyenne temporelle du puits de carbone continental selon les forçages (à gauche), selon les modèles (au centre) et pourcentage de variabilité expliqués par les forçages, les modèles et les interactions pour le puits de carbone continental (à droite).

d'émissions de carbone associées. Cependant, pour les simulations réalisées par les modèles du MPI et du NCAR, qui calculent la probabilité d'incendie, c'est avec Princeton qu'on observe les émissions les moins importantes. En effet, on peut avancer l'idée que lorsque l'atmosphère est plus humide, la probabilité d'incendie diminue fortement, et les émissions de carbone liées aux feux de forêt sont plus faibles, même si la quantité de biomasse disponible pour la combustion est plus importante avec Princeton. Cela expliquerait également pourquoi on obtient les valeurs les plus élevées avec CRUNCEP, qui arrive second en terme de biomasse mais est un peu moins humide que GSWP3. Les valeurs très faibles trouvées avec Princeton pour le MPI semblent néanmoins excessives, et pourraient résulter d'une part des interactions avec le cycle de l'azote expliqué plus haut, et d'autre part d'un problème de sensibilité de la probabilité d'incendie à l'humidité.

Lorsqu'on retranche, pour chaque simulation, f_{Fire} au flux $GPP - Ra - Rh$, on obtient les résultats présentés figure 15, pour la dispersion et les sources de variabilité dans l'estimation de S_{Land} . Du fait que l'incertitude liée aux modèles soit très forte pour f_{Fire} , et que l'amplitude de cette variabilité ($\approx 1 \text{ Pg.y}^{-1}$) soit grande devant l'amplitude de la variabilité pour $GPP - Ra - Rh$ ($\approx 0,1 \text{ Pg.y}^{-1}$), on observe que ce sont essentiellement les modèles (75%) qui expliquent la variabilité de S_{Land} . Les forçages et les termes d'interactions restent significatifs, avec respectivement 9% et 16% de la variabilité expliquée. On remarquera surtout que l'incertitude provient essentiellement des émissions liées aux feux de forêts, puisque c'est seulement à ce stade que le rôle des forçages devient secondaire. Nous aurions pu choisir d'utiliser les données d'observations, ou les données du NCAR qui semblent plus en accord avec la réalité, pour estimer f_{Fire} et S_{Land} pour l'ensemble des simulations, ce qui conduirait à un rôle plus important des forçages atmosphériques. Néanmoins, les émissions liées aux feux de forêts dépendent notamment des réservoirs de biomasse, des types de végétation et bien sûr des équations de feux utilisées. Des différences importantes subsistent entre les modèles pour ces paramètres (voir annexe C pour les réservoirs), et il est donc logique que l'incertitude liée au modèle soit importante pour les émissions liées aux feux de forêt, et par conséquent sur l'estimation du

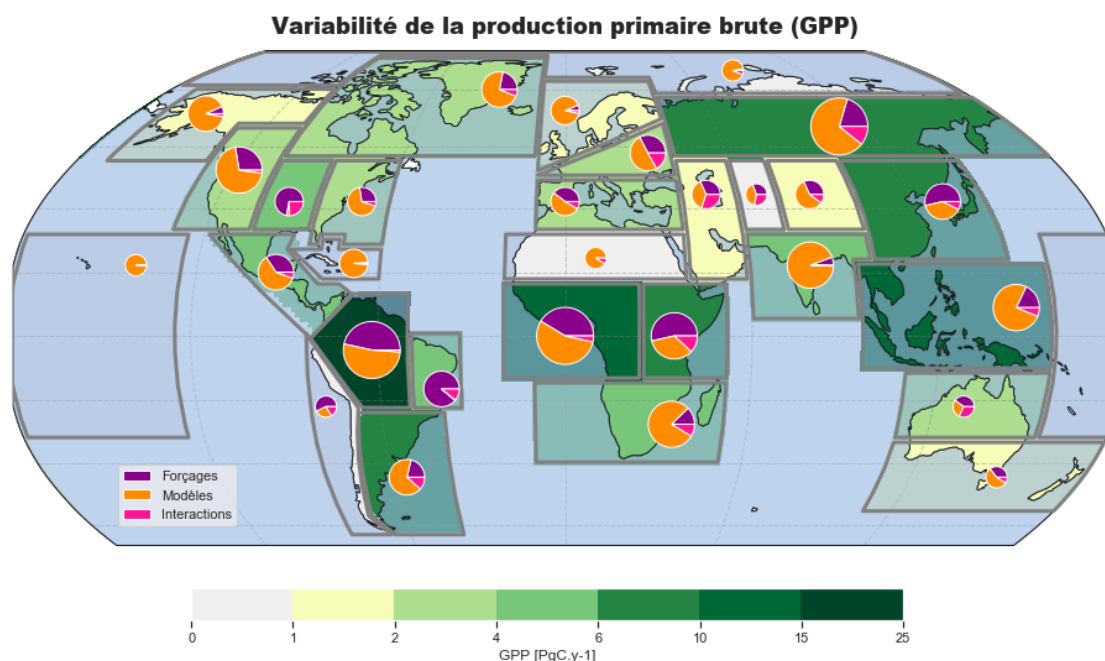


FIGURE 16 – Distribution spatiale de la variabilité pour la production primaire brute. Les régions plus sombre sont celles qui contribuent le plus à la production globale.

puits de carbone continental. Ces modules doivent donc être améliorés et intégrés à l'ensemble des modèles de surface de manière à mieux représenter le cycle du carbone mais aussi les autres cycles biogéochimiques. En particulier, il existe de nombreuses interactions entre les activités humaines, les feux de forêts et le changement climatique [5] [40], il est donc nécessaire d'améliorer leur représentation dans les modèles, de manière à prévoir leur évolution et leur rétroactions sur le climat.

3.2 Échelle Régionale

Je me suis par la suite intéressé à la distribution régionale de la variabilité, avec pour objectif initial de déterminer les zones où la responsabilité des forçages dans l'incertitude était la plus importante. La figure 16 présente les résultats obtenus pour la production primaire brute, pour les différentes régions terrestres définies dans le rapport SREX de l'IPCC [52]. Les plus gros diagrammes, que l'on peut observer dans les forêts équatoriales et aux hautes latitudes de l'Asie, correspondent aux zones où l'amplitude de la variabilité est la plus importante. On peut remarquer que, pour la plupart des régions, les forçages expliquent une part de variabilité moins importante que les modèles, contrairement à ce que l'on avait pu observer pour la GPP à l'échelle globale figure 9. Lorsqu'on compare les cartes de la production primaire brute entre les modèles, données en Annexe F, on s'aperçoit qu'il existe des grands désaccords entre les simulations sur l'impact des différentes régions, et en particulier entre les forêts équatoriales et les hautes latitudes. Les valeurs moyennes de GPP trouvées par le CNRM aux hautes latitudes sont par exemple très inférieures aux valeurs moyennes trouvées par le NCAR, et on observe

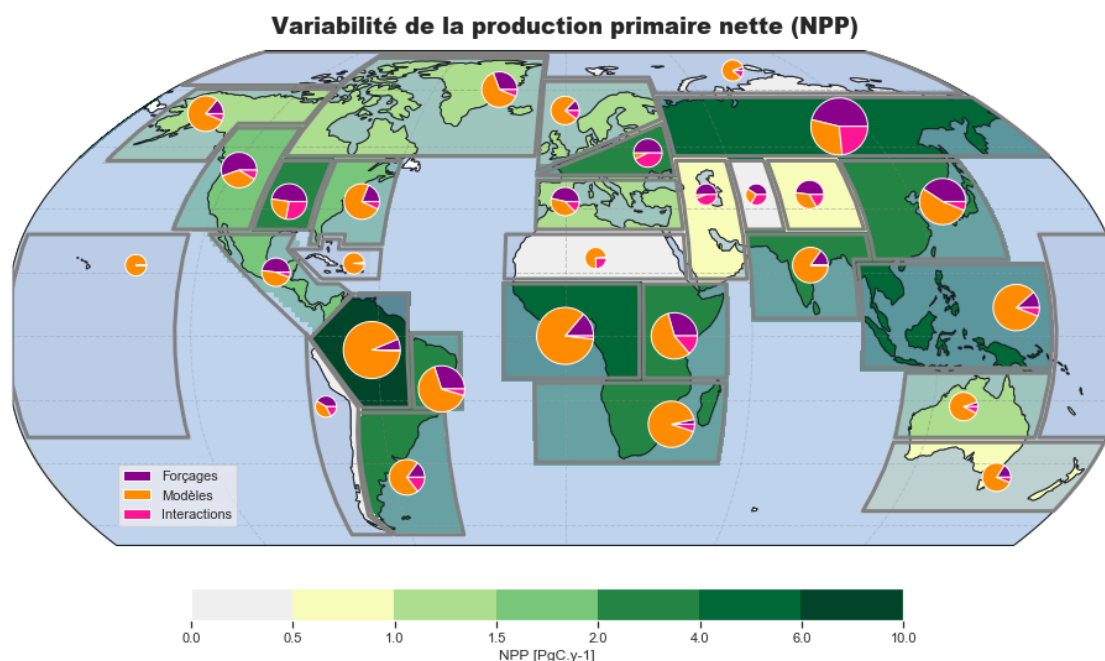


FIGURE 17 – Distribution spatiale de la variabilité pour la production primaire nette.

le phénomène inverse au niveau de la forêt équatoriale. Ces différences vont en grande partie se compenser lorsqu'on somme l'ensemble des régions pour obtenir le flux global, ce qui réduit la variabilité due aux modèles et donc augmente la variabilité des forçages. Le rôle des forçages est beaucoup plus marqué dans les forêts équatoriales qu'aux hautes latitudes pour la GPP, une explication pouvant être la difficulté d'obtenir des données atmosphériques représentatives dans ces zones, ainsi que la période de croissance qui est plus longue (toute l'année contre quelques mois), ce qui laisse plus de temps aux estimations pour diverger. Le rôle des forçages semble cependant moins important lorsqu'on regarde l'Inde et l'Asie du Sud Est. On peut avancer le rôle de la mousson, qui est plutôt bien représenté par les forçages car les précipitations mesurées sont très importantes, et l'erreur relative associée plus faible. Pour l'Asie du Sud Est, une erreur peut également provenir de la maille des modèles, en effet le modèle du NCAR a une résolution presque deux fois supérieure à celui du MPI, le modèle du CNRM étant à l'entre deux. Par conséquent, la fraction de végétation par maille peut beaucoup différer dans les régions côtières, au point que certaines îles ne sont parfois pas représentées. En dernier point, on peut souligner que la variabilité liée aux modèles est presque inexistante dans les plaines au centre de l'Amérique du Nord. Ces zones sont essentiellement dédiées à l'agriculture, et présentent le même type de végétation sur de grandes zones, ce qui est par ailleurs bien documenté en terme de superficie et d'espèce (on a le plus souvent des monocultures). Ces zones bien connues, mais présentant peu de variation de relief et de végétation, sont plus faciles à modéliser, ce qui peut expliquer les résultats obtenus.

La figure 17 présente les résultats pour la production primaire nette (NPP). On obtient par ailleurs des résultats assez similaires pour la respiration hétérotrophe (Rh). Le rôle des forçages

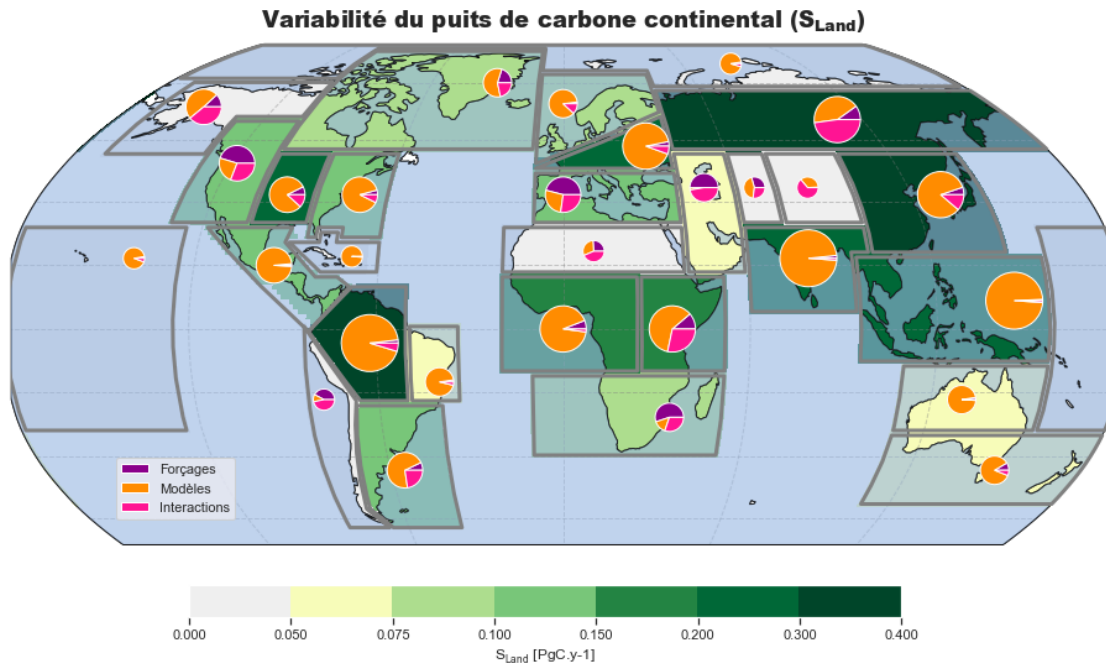


FIGURE 18 – Distribution spatiale de la variabilité pour le puits de carbone continental.

devient moins important dans les forêt équatoriales, mais beaucoup plus important aux hautes latitudes. Dans les deux cas, on peut avancer une hypothèse en expliquant les différences entre les modèles obtenus précédemment pour la GPP. Dans la forêt équatoriale, la quantité de biomasse et donc de carbone dans la végétation est sensiblement la même. En effet, la croissance et la mortalité de la végétation a lieu tout au long de l'année, la forêt est presque à l'équilibre. Les différences observées dans les estimations de la GPP proviennent essentiellement des choix faits par les modèles pour limiter la photosynthèse en fonction de l'humidité et de l'éclairement, plutôt que des différences dans la quantité de végétation initiale. Par conséquent, pour la respiration autotrophe, qui dépend essentiellement de la quantité de biomasse, on peut s'attendre à avoir des estimations assez similaires entre les modèles dans ces zones. En retranchant ces valeurs de respiration proches, on accentue les différences entre les modèles déjà observées. C'est le phénomène inverse qui est observable aux hautes latitudes : les variations entre les modèles dans l'estimation de la GPP sont en partie liées à des différences dans l'estimation de la biomasse produite chaque année, qui engendrent, de manière proportionnelle, des écarts dans l'estimation de R_a . Lorsqu'on calcule NPP, ces écarts se compensent en partie, et le biais final entre les modèles est réduit, ce qui laisse finalement une part plus importante aux forçages dans l'explication de la variabilité.

Enfin, la figure 18 présente les résultats pour le puits de carbone continental. On retrouve majoritairement une incertitude liée aux modèles, comme nous pouvions déjà le voir à l'échelle globale (figure 15). Pour le justifier, on peut rappeler les incertitudes sur le rôle des différentes régions, déjà visibles dans GPP et NPP, qui sont amplifiées lorsqu'on prend en compte la

respiration du sol. On remarquera l'influence de la mousson et de la taille de la maille pour l'Inde et l'Asie du sud-est, ainsi que le rôle des feux de forêts, qui est particulièrement important en Afrique, en Océanie et aux hautes latitudes asiatiques. Le rôle des forçages ne cesse de diminuer car, comme nous l'avons vu, les écarts observés pour la GPP engendrent le même type d'écart pour R_a et R_h . Leurs effets vont donc avoir tendance à se compenser lorsqu'on se rapproche du flux net. Puisque l'effet de l'humidité augmente particulièrement la respiration du sol, mais que ce phénomène n'est pas pris en compte de la même manière par tous les modèles, la tendance observée entre les forçages peut même être modifiée, ce qui génère des termes d'interactions importants qu'on peut observer aux hautes latitudes, dans le bassin méditerranéen et en Afrique du sud et de l'est. Rappelons que ces termes d'interactions sont également visibles à l'échelle globale pour les feux de forêts, ces derniers venant amplifier ce phénomène dans les régions concernées.

4 Conclusion

Les travaux menés au cours de ce stage ont permis de montrer le rôle des forçages atmosphériques dans l'estimation des stocks et flux de carbone par les modèles de surface. Dans un premier temps, une description qualitative des estimations réalisées par les modèles a mis en valeur l'impact des forçages. En particulier, j'ai pu observer que les forçages ont un plus grand impact sur les flux que sur les stocks de carbone, qui dépendent beaucoup des choix de représentation des modèles (Annexe C), et que tous les modèles ne présentaient pas la même sensibilité aux forçages atmosphériques.

Par la suite, j'ai utilisé l'analyse de variance pour quantifier la part de variabilité attribuable aux forçages dans les estimations globales des flux, en utilisant les simulations de trois modèles de surface réalisées avec trois forçages atmosphériques différents. Ces résultats ont montré que les forçages avaient un rôle largement significatif et parfois même dominant dans la décomposition de l'incertitude des flux de carbone, avec par exemple $\sim 75\%$ de la variabilité expliquée pour la GPP, $\sim 89\%$ pour Ra et $\sim 30\%$ pour Rh. Le rôle des forçages semble moins important lorsqu'on considère le puits de carbone continental, avec $\sim 9\%$ de la variabilité expliquée contre $\sim 75\%$ pour les modèles, mais une analyse plus poussée révèle que ce sont essentiellement les émissions liées aux feux de forêt qui expliquent ces résultats. En effet, les valeurs de f_{Fire} trouvées par les modèles du CNRM et du MPI sont assez éloignées des observations. Améliorer ces modèles de feux devrait être un objectif important pour la partie continentale du septième projet d'intercomparaison des modèles couplés (CMIP7), de manière à réduire l'incertitude sur les estimations du puits de carbone actuel mais aussi de son évolution dans le contexte du changement climatique, à cause notamment des épisodes de sécheresse intense conduisant à des incendies violents que l'on peut déjà observer aujourd'hui.

La réponse des modèles aux différents forçages atmosphériques a également été quantifiée via les termes d'interaction. Une interaction relativement faible, autour de 6%, a été trouvée pour GPP, Ra, Rh et NPP, contre 41% pour GPP-Ra-Rh et 16% pour S_{LAND} . En effet, des processus comme la photosynthèse ou la respiration vont être plus ou moins modifiés selon le modèle considéré, mais toujours de la même manière, lorsqu'on fait varier le forçage atmosphérique (valeurs d'interaction faible). Lorsqu'on fait la différence de la photosynthèse et de la respiration, on n'obtient pas la même réponse entre les modèles, selon que ce soit la photosynthèse ou la respiration qui a été le plus affecté par le changement des paramètres atmosphériques, on a alors un terme d'interaction élevé (un même forçage ne donne pas les valeurs les plus élevées pour tous les modèles). Notons que les émissions liées aux feux de forêt présentent également une grande sensibilité aux forçages atmosphériques, puisque des variables climatiques comme l'humidité de l'air peuvent avoir des effets antagonistes en augmentant la quantité de biomasse disponible d'un côté, et en réduisant la probabilité d'incendie de l'autre.

Enfin, l'analyse de variance a été appliquée à l'échelle régionale, dans l'optique de déterminer

les régions du globe avec le plus d'incertitude, et où les forçages atmosphériques contribuaient le plus à la variabilité. Ces résultats ont montré que le rôle des forçages était fortement réduit par rapport à ce que l'on pouvait observer en global, pour la plupart des régions et quel que soit le flux considéré. En comparant les estimations régionales selon les modèles et selon les forçages (Annexe F), j'ai pu montrer que les variations des flux de carbone causées par les différences entre les forçages étaient généralement du même signe pour toutes les régions du globe, tandis que les modèles étaient en opposition avec par exemple des valeurs plus élevées dans la forêt tropicale pour le CNRM, et des valeurs plus élevées dans l'hémisphère Nord pour le National Center for Atmospheric Research (NCAR). Lorsqu'on fait la somme de ces différences pour obtenir le flux global, les différences entre les forçages s'amplifient, tandis que celles entre les modèles se compensent. On peut conclure que si le rôle des forçages atmosphériques est très important à l'échelle globale et devrait être pris en compte dans les estimations du cycle du carbone, c'est encore la structure des modèles qui domine l'incertitude dans la plupart des régions. Néanmoins, la part de variabilité attribuable aux forçages reste significative pour les estimations de la production primaire brute dans les forêts inter-tropicales, et les forçages semblent jouer un rôle important pour la production nette aux hautes latitudes. Une prise en compte des forçages atmosphériques à l'échelle régionale peut donc également être préconisée, d'autant plus que l'utilisation de modèles régionaux plutôt que globaux devrait largement réduire l'incertitude attribuable à la résolution et à la représentation de la végétation.

La principale limite de cette étude est, selon moi, le nombre de simulations utilisées pour parvenir à ces conclusions. J'ai été plusieurs fois confronté au manque de données durant ce stage, puisque les données du NCAR et de l'IPSL ont été déposées plusieurs mois après le début des travaux, et que certaines estimations de variables étaient absentes ou erronées pour les autres modèles. Le choix de n'utiliser que les modèles utilisant les trois forçages atmosphériques était, selon moi, nécessaire pour quantifier correctement le rôle des forçages et des interactions. Cependant, le hasard a fait que les modèles de l'IPSL et du BCC ont été exclus de l'étude sur ces critères, tandis que ces deux modèles fournissaient respectivement les plus faibles et plus fortes valeurs moyennes et écart inter-quartile pour la plupart des flux. Les valeurs obtenues durant ce stage ne sont donc pas complètement représentatives de l'ensemble des modèles de surface. On peut, selon moi, leur attribuer un indice de confiance moyen malgré ces lacunes. D'une part, parce que des résultats relativement similaires à ceux présentés dans ce rapport ont été retrouvés en attribuant des valeurs approchées pour les simulations manquantes, et d'autre part, parce que les modèles de surface du BCC et de l'IPSL ont régulièrement des comportements d'outsider pour la modélisation du cycle du carbone. Une seconde limite pourrait être le calcul du puits de carbone continental sous la forme du NEP, qui n'est pas strictement équivalent à l'estimation de S_{LAND} faite par le GCP, même si nos résultats concernent l'ensemble des flux de carbone terrestre. Par ailleurs, rappelons que le rôle des forçages atmosphériques n'avait jamais été analysé de la sorte pour le cycle du carbone. Les résultats obtenus sont donc suffisamment

forts et inattendus pour être partagés avec la communauté scientifique, en encourageant d'autres travaux palliant ces limites.

Cette expérience m'a beaucoup apporté sur le plan personnel et professionnel. D'abord, je pense qu'elle a parfaitement répondu à mes attentes en termes d'initiation à la recherche. J'ai eu l'opportunité d'acquérir des connaissances nouvelles sur le cycle du carbone continental, sur sa représentation dans les modèles de surface, et de manière plus générale sur le fonctionnement des modèles de climat et la représentation des processus physiques et biologiques. J'ai pu développer des compétences propres au métier de chercheur, notamment en recherche bibliographique, en rédaction et en programmation pour l'analyse de données, ainsi que découvrir de nouveaux outils informatiques et statistiques, et entraîner mon esprit d'analyse, mon regard critique et ma curiosité. Ce stage m'a surtout permis d'intégrer un laboratoire de recherche et de côtoyer différents chercheurs, ingénieurs et doctorants, qui ont partagé avec moi leur passion pour ce travail, leurs conseils et leur enthousiasme, ainsi que de très bons moments. En particulier, mes tuteurs m'ont beaucoup appris, orienté et encouragé, tout en me laissant une grande autonomie dans mon travail en m'accordant leur confiance. Ce stage est également l'aboutissement de cinq années d'études en tant qu'ingénieur physicien, bien qu'il fasse essentiellement écho à mon dernier semestre de spécialisation au sein du master sciences de l'océan, de l'atmosphère et du climat. Ce semestre m'a en effet formé aux enjeux liés au cycle du carbone et ses interactions avec le climat, et m'a donné envie d'effectuer ce stage afin de travailler sur cette problématique, tout en cherchant à en apprendre davantage. Néanmoins, ce sont aussi les compétences humaines, techniques et les innombrables opportunités offertes par la formation INSA qui ont contribué au bon déroulement de ce stage, et m'ont donné envie de poursuivre mon parcours dans la recherche.

A Variabilité autour des modèles des flux de carbone

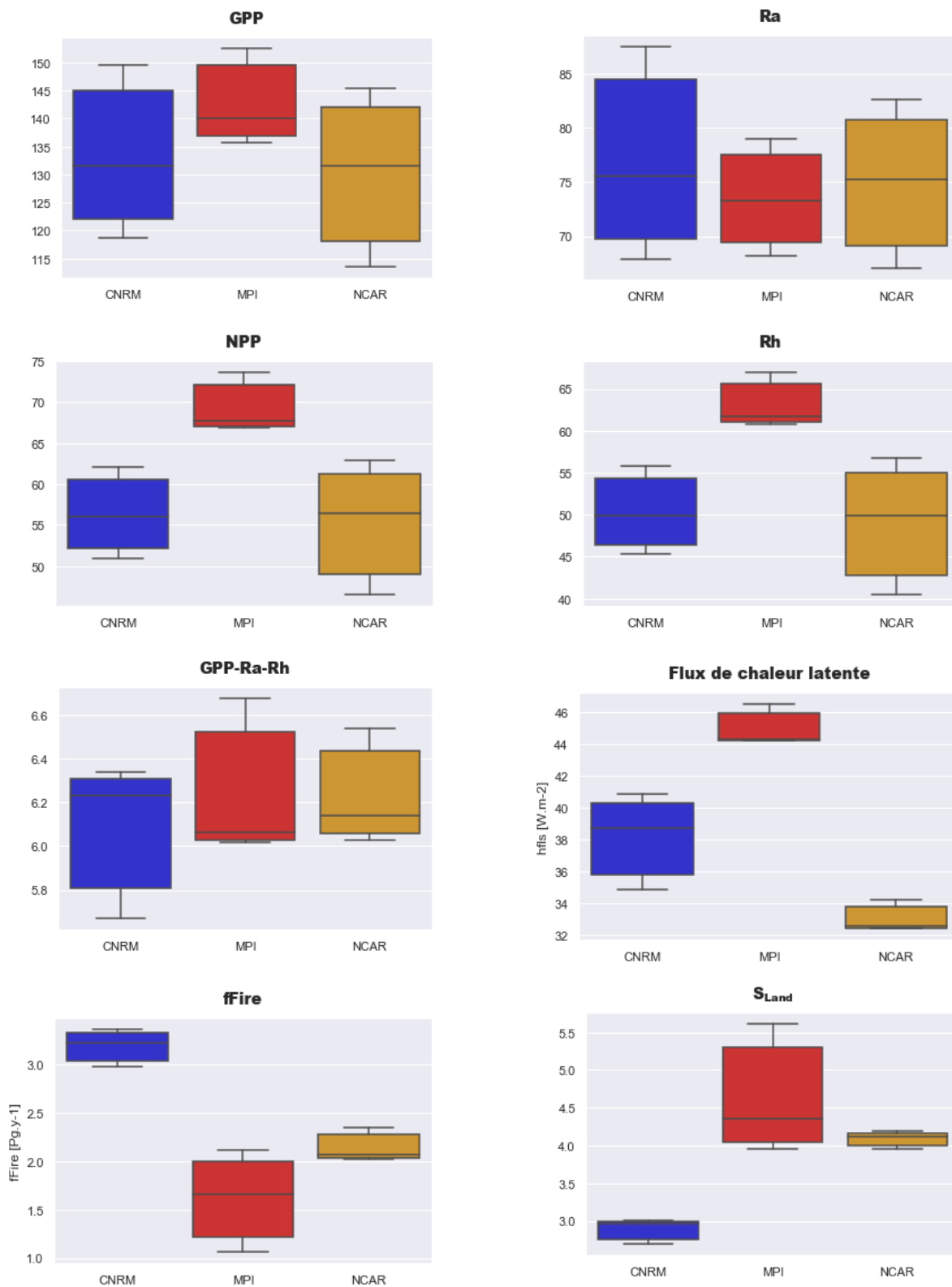


FIGURE 19 – Dispersion de la moyenne temporelle des différents flux pour les modèles du CNRM (bleu), du MPI (Rouge) et du NCAR (orange).

B Répartition de la variabilité des flux de carbone

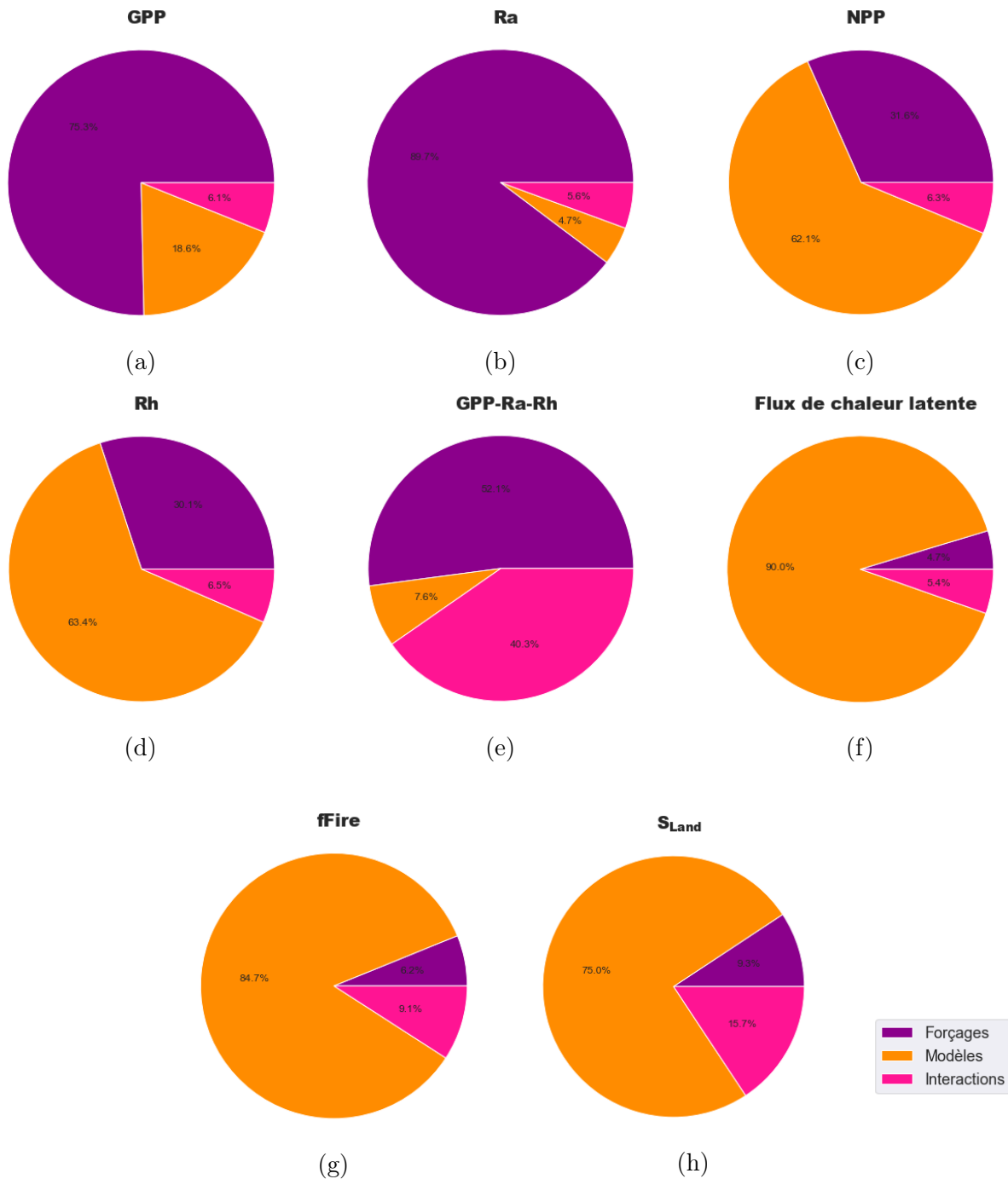
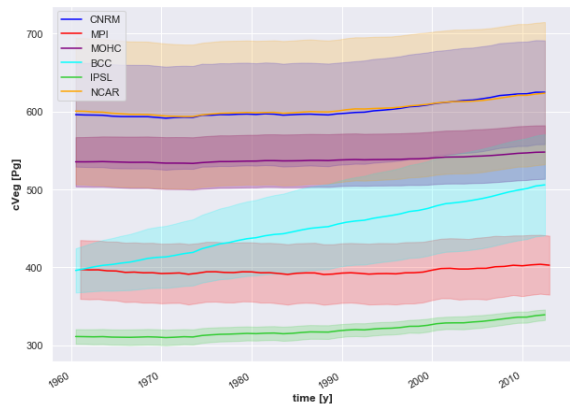
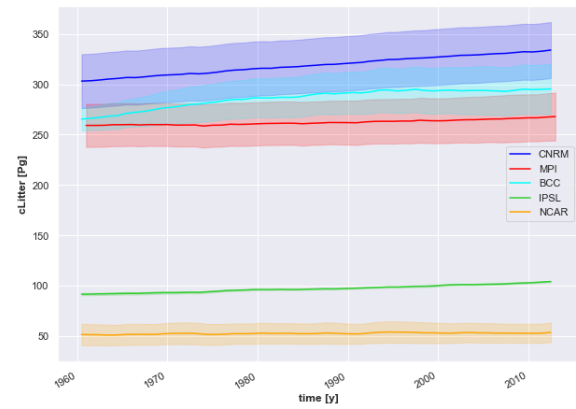


FIGURE 20 – Variabilité expliquée par les forçages (violet), modèles (orange), et interactions (rose) pour la production primaire brute (a), la respiration autotrophe (b), la production primaire nette (c), la respiration hétérotrophe (d), GPP-Ra-Rh (e), le flux de chaleur latente HFSL (f), les émissions liées aux feux de forêt (g) et le puits de carbone continental (h).

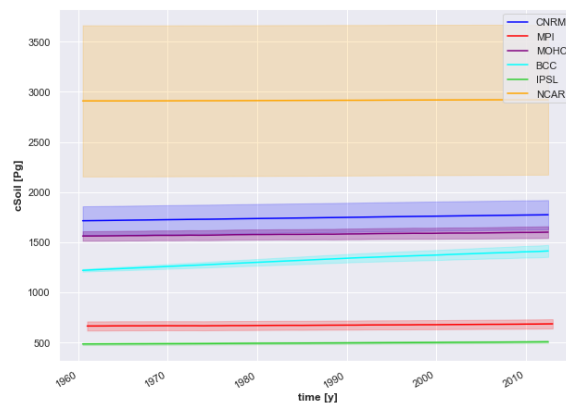
C Estimation des stocks de carbone par les modèles



(a) cVeg



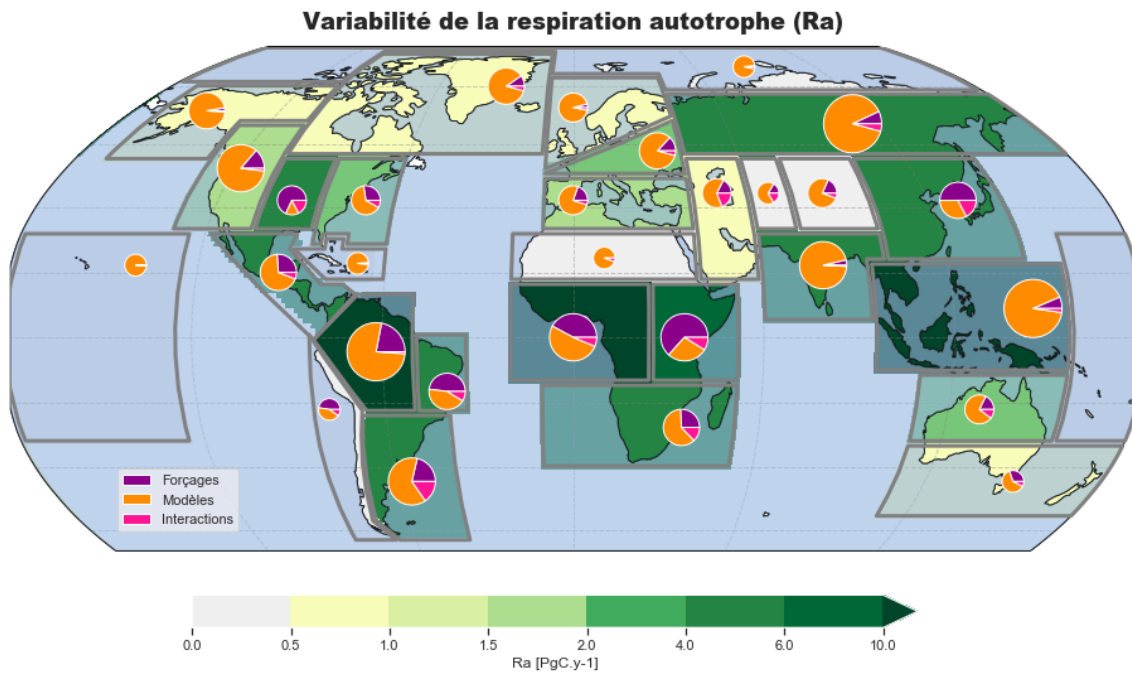
(b) cLitter



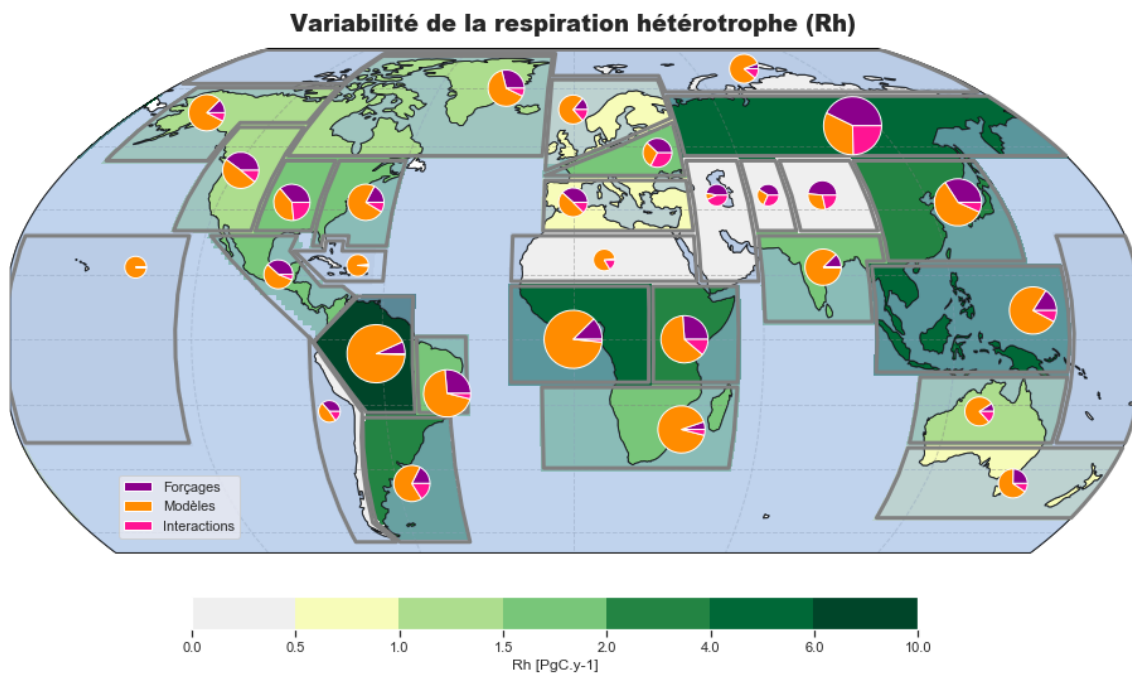
(c) cSoil

FIGURE 21 – Moyenne des estimations du carbone stocké pour les modèles du CNRM (bleu), MPI (rouge), BCC (cyan), MOHC (violet), IPSL (vert) et NCAR (orange).

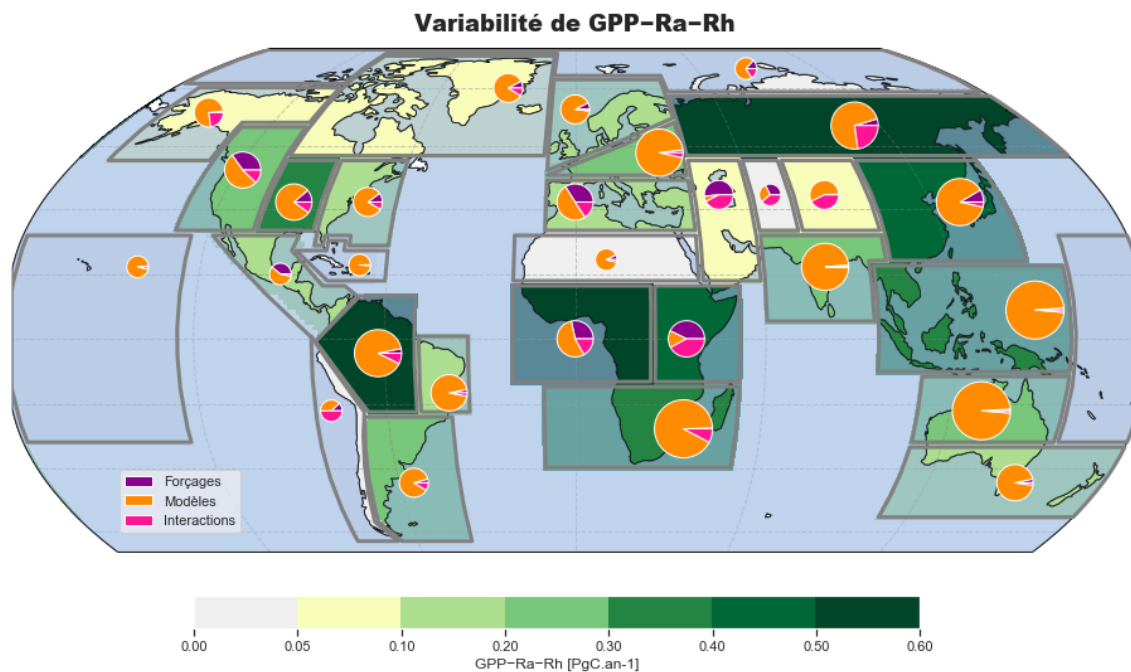
D Distribution spatiale de la variabilité



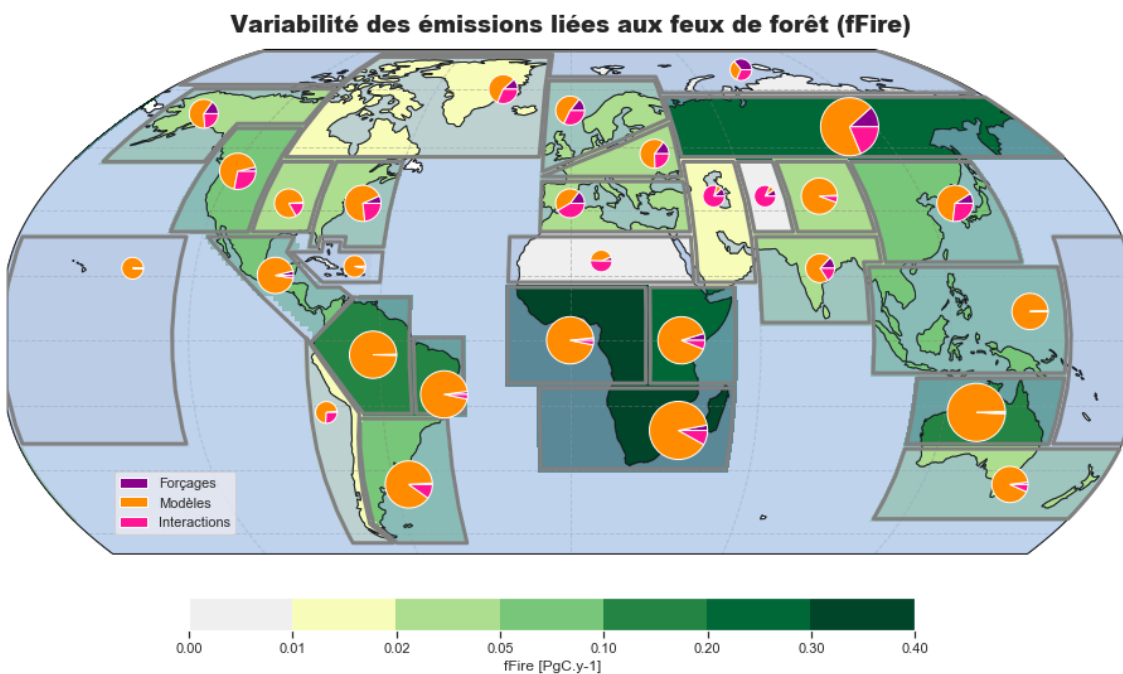
(a)



(b)



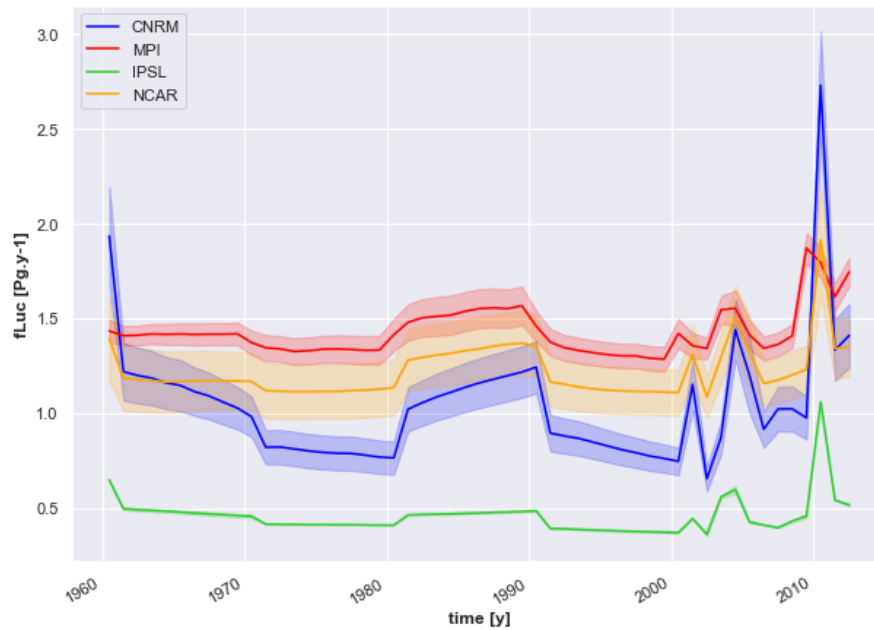
(c)



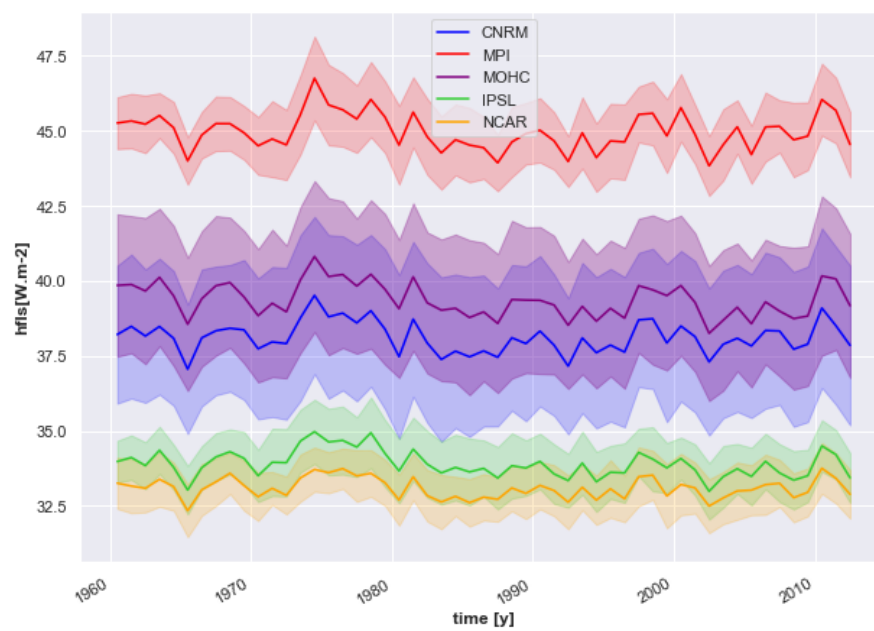
(d)

FIGURE 22 – Distribution spatiale de la variabilité pour la respiration autotrophe (a), la respiration hétérotrophe (b), GPP-Ra-Rh (c) et fFire (d).

E Évolution temporelle de flux secondaires



(a) fLuc



(b) HFLS

FIGURE 23 – Moyenne par modèle de l'estimation des émissions du aux changements d'usage des sols (a) et de l'estimation du flux de chaleur latente HFLS (b) pour les modèles du CNRM (bleu), MPI (rouge), MOHC (violet), IPSL (vert) et NCAR (orange).

F Comparaison des estimations de la production primaire brute entre le CNRM et le NCAR

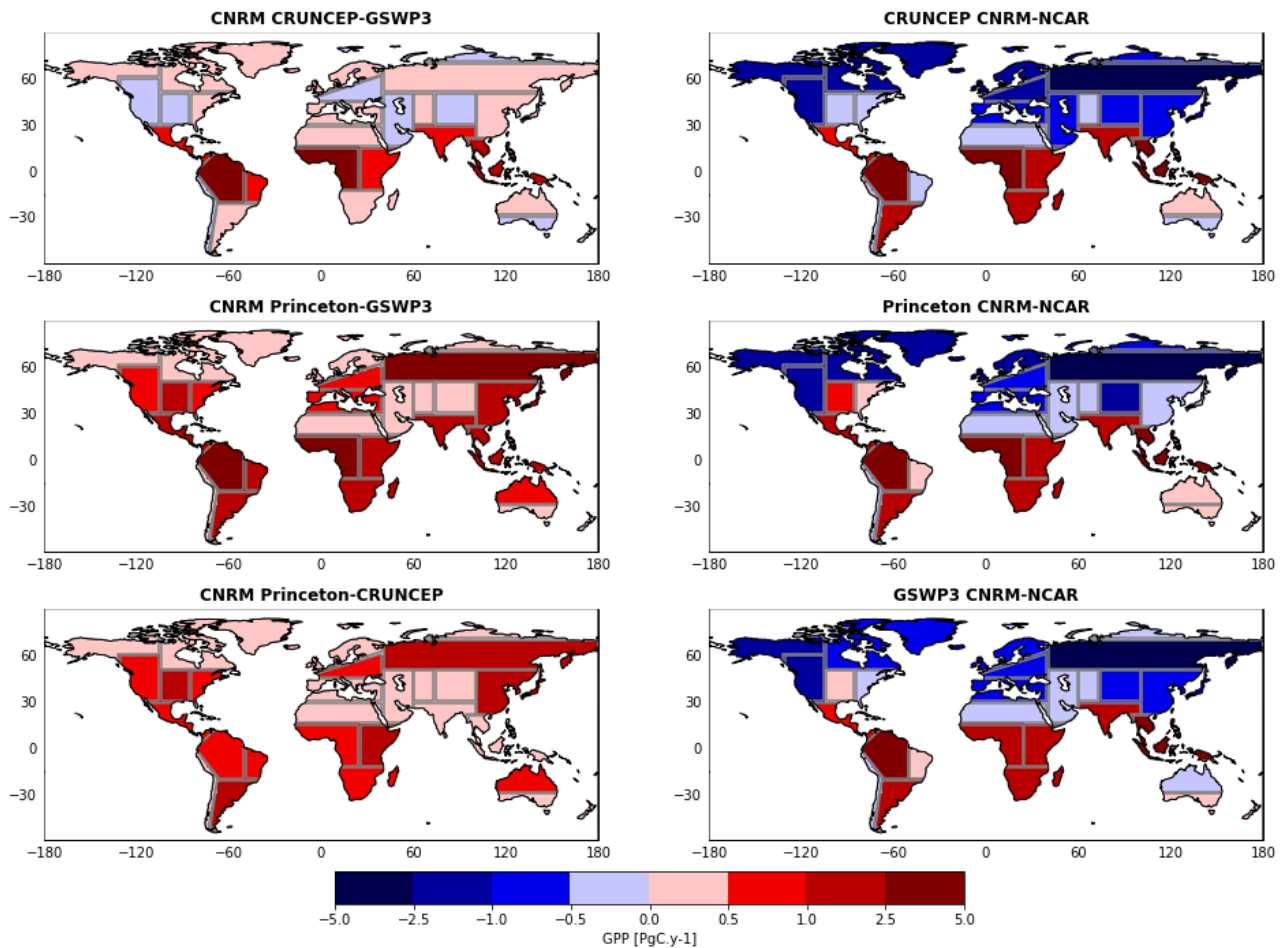


FIGURE 24 – à gauche : différences entre les estimations de la production primaire brute faite par le modèle du CNRM : CRUNCEP – GSWP3 (haut), Princeton – GSWP3 (centre) et Princeton – CRUNCEP (bas). à droite : différences entre les estimations des modèles du CNRM et du NCAR forcé par CRUNCEP (haut), Princeton (milieu) et GSWP3 (bas).

G Questionnaire d'étonnement

1. *Quelle(s) forme(s) d'innovation récente(s) avez-vous pu identifier dans l'entité qui vous héberge en stage ?*

Le Centre National de Recherche Météorologique est, comme son nom l'indique, un centre de recherche. Ses membres contribuent à améliorer sans cesse nos connaissances, notre compréhension et nos manière de prévoir la météo et le climat. Cela constitue en un sens une innovation incrémentale, puisque les chercheurs contribue à améliorer un état existant du savoir. Il arrive aussi que l'innovation soit radicale, en introduisant de nouveaux outils comme récemment avec l'utilisation de l'intelligence artificiel pour la prévision météorologique. Néanmoins, l'innovation n'a ici pas de but économique, les chercheurs, doctorants, ingénieurs, techniciens et personnel administratif contribuent en premier lieu à produire et développer les connaissances scientifiques.

2. *Comment l'innovation est-elle mise en œuvre dans cette entité ? Décrivez et expliquez*

L'"innovation" est mise en oeuvre en améliorant les connaissances, les modèles, l'assimilation de données et les systèmes de prévisions et en introduisant de nouveaux outil.

3. *Par qui l'innovation est-elle mise en œuvre dans l'entité (un département, un service, un métier, une fonction...) ? Expliquez.*

L'innovation résulte indirectement du travail des chercheurs et doctorant, est mise en place par les chercheurs, technicien et ingénieurs, mais c'est toute l'entité qui est dédié à la création de connaissance.

4. *Essayez de déterminer à quelle échelle (petite, grande) l'innovation est mise en œuvre : concerne-t-elle tous les domaines et services, ou est-elle très ponctuelle ?*

Tout les domaines et services introduise de l'innovation selon la définition d'amélioration de l'état existant, elle est donc à grande échelle. Certains services sont dédiés à l'introduction de nouvelles méthode et outil, comme la branche travaillant sur l'amélioration des prévisions par l'intelligence artificiel.

5. *Qu'apporte cette innovation et à qui est-elle bénéfique ?*

La connaissance produite est bénéfique et partagée à l'ensemble de l'humanité. En particulier, l'amélioration des prévisions climatiques est supposé alerter le grand public et les décideurs politiques sur les perturbations (réchauffement, fonte, sécheresse, etc) engendré par la civilisation thermo-industriel. Les innovations en prévisions météorologique sont également bénéfique pour tous, elles ont toujours été d'un grand intérêt pour des secteurs comme le transport (rou-tier, bateau, aviation) et l'agriculture, mais aussi pour prévoir des catastrophes naturel comme celles que nous avons pu observer cet été avec des crues et des incendies.

6. *Comment cette innovation se situe-t-elle par rapport à la notion de développement durable et aux enjeux sociétaux actuels ?*

Cette innovation permet, notamment, de mieux comprendre les enjeux liés au changement climatique. Elle atteste de ces changements et essaye de prévoir leur évolution et leurs conséquences sur d'autres composantes du système terre (océans, surfaces, vivant, sociétés humaine). Elle ne présente pas directement de solutions à ce changement, majoritairement provoqué par l'augmentation de la concentration de gaz à effet de serre, mais tire la sonnette d'alarme et invite à se mettre en mouvement de toute urgence pour réduire nos émissions. Certains services sont également concernés par d'autres enjeux sociétaux, comme la pollution et la présence de contaminants dans l'atmosphère. (D'autres laboratoires de Toulouse travaillent plus spécifiquement sur la pollution des océans et des sols).

7. A l'inverse, avez-vous perçu des limites ou des réticences face à cette innovation ? A-t-elle eu des impacts négatifs ?

On peut voir quelques aspects négatifs à l'innovation de la connaissance et de la recherche. L'utilisation des bâtiments, du matériel informatique, et des différents outils scientifiques consomme de l'énergie et émet des GES. L'envoi de nouveaux satellites pollue l'environnement et l'espace et émet également une grande quantité de GES. Enfin, l'amélioration des prévisions et des modèles peut parfois demander plus de ressources et d'énergie consommées par les super-calculateurs. Néanmoins, ces opérations sont le plus souvent justifiées et nécessaires : elles nous permettent justement d'observer les changements de notre environnement et de notre climat, nous permettent de prévoir leur évolution, et nous alertent sur les dangers de nos modes de production et de consommation. Même si certaines méthodes pourraient être optimisées, notamment dans l'utilisation des bâtiments (trop chauffés en hiver, climatisés en été), voir l'envoi de certains satellites, les innovations introduites au CNRM contribuent positivement à répondre aux enjeux climatiques.

8. Si vous étiez embauché demain dans cette entité, quelle serait la première innovation que vous introduiriez ?

Si j'avais été embauché au CNRM en tant que doctorant, j'aurais contribué à améliorer la représentation du cycle du carbone dans le modèle de surface du CNRM, en particulier en introduisant et développant un nouveau modèle de feu.

Références

- [1] M. O. Andreae and P. Merlet. Emission of trace gases and aerosols from biomass burning. *Global biogeochemical cycles*, 15(4) :955–966, 2001.
- [2] D. Bice. Unit 9 reading : The global carbon cycle.
- [3] G. B. Bonan. *Ecological climatology : concepts and applications*. Cambridge University Press, 2008.
- [4] O. Boucher, J. Servonnat, A. L. Albright, O. Aumont, Y. Balkanski, V. Bastrikov, S. Bekki, R. Bonnet, S. Bony, L. Bopp, et al. Presentation and evaluation of the ipsl-cm6a-lr climate model. *Journal of Advances in Modeling Earth Systems*, 12(7) :e2019MS002010, 2020.
- [5] D. M. Bowman, J. Balch, P. Artaxo, W. J. Bond, M. A. Cochrane, C. M. D’antonio, R. DeFries, F. H. Johnston, J. E. Keeley, M. A. Krawchuk, et al. The human dimension of fire regimes on earth. *Journal of biogeography*, 38(12) :2223–2236, 2011.
- [6] G. E. Box. Non-normality and tests on variances. *Biometrika*, 40(3/4) :318–335, 1953.
- [7] G. P. Compo, J. S. Whitaker, P. D. Sardeshmukh, N. Matsui, R. J. Allan, X. Yin, B. E. Gleason, R. S. Vose, G. Rutledge, P. Bessemoulin, et al. The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137(654) :1–28, 2011.
- [8] G. Danabasoglu, J.-F. Lamarque, J. Bacmeister, D. Bailey, A. DuVivier, J. Edwards, L. Emmons, J. Fasullo, R. Garcia, A. Gettelman, et al. The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, 12(2), 2020.
- [9] J. W. Deardorff. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *Journal of Geophysical Research : Oceans*, 83(C4) :1889–1903, 1978.
- [10] B. Decharme, C. Delire, and A. Boone. La représentation des surfaces continentales dans la modélisation du climat à météo-france. *La Météorologie*, 108 :67–76, 2020.
- [11] B. Decharme, C. Delire, M. Minvielle, J. Colin, J.-P. Vergnes, A. Alias, D. Saint-Martin, R. Sférian, S. Sénési, and A. Voldoire. Recent changes in the isba-ctrip land surface system for use in the cnrm-cm6 climate model and in global off-line hydrological applications. *Journal of Advances in Modeling Earth Systems*, 11(5) :1207–1252, 2019.
- [12] C. Delire, R. Sférian, B. Decharme, R. Alkama, J.-C. Calvet, D. Carrer, et al. The global land carbon cycle simulated with isba-ctrip : Improvements over the last decade. *Journal of Advances in Modeling Earth Systems*, 12, 2020.
- [13] E. Dlugokencky and P. Tans. Trends in atmospheric carbon dioxide, national oceanic and atmospheric administration.
- [14] V. Eyring, S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor. Overview of the coupled model intercomparison project phase 6 (cmip6) experimental design and organization. *Geoscientific Model Development*, 9(5) :1937–1958, 2016.

- [15] P. Friedlingstein, L. Bopp, and P. Cadule. Changement climatique et cycle du carbone. *La Météorologie*, 2007.
- [16] P. Friedlingstein, M. Jones, M. O’sullivan, R. Andrew, J. Hauck, G. Peters, W. Peters, J. Pongratz, S. Sitch, C. Le Quéré, et al. Global carbon budget 2019. *Earth System Science Data*, 11(4) :1783–1838, 2019.
- [17] P. Friedlingstein, M. O’Sullivan, M. W. Jones, R. M. Andrew, J. Hauck, et al. Global carbon budget 2020. *Earth System Science Data*, 12(4) :3269–3340, 2020.
- [18] D. S. Goll, A. J. Winkler, T. Raddatz, N. Dong, I. C. Prentice, P. Ciais, and V. Brovkin. Carbon–nitrogen interactions in idealized simulations with jsbach (version 3.10). *Geoscientific Model Development*, 10(5) :2009–2030, 2017.
- [19] O. Gutjahr, D. Putrasahan, K. Lohmann, J. H. Jungclaus, J.-S. v. Storch, N. Brüggemann, H. Haak, and A. Stössel. Max planck institute earth system model (mpi-esm1. 2) for the high-resolution model intercomparison project (highresmip). *Geoscientific Model Development*, 12(7) :3241–3281, 2019.
- [20] L. Hardouin. Rôle des surfaces continentales dans le cycle du carbone global.
- [21] I. Harris. Cru jra v1. 1 : A forcings dataset of gridded land surface blend of climatic research unit (cru) and japanese reanalysis (jra) data, january 1901–december 2017, university of east anglia climatic research unit, centre for environmental data analysis, 2019.
- [22] J. I. Hedges. Global biogeochemical cycles : progress and problems. *Marine Chemistry*, 39 :67–93, 1992.
- [23] J. A. Hicke. Ncep and giss solar radiation data sets available for ecosystem modeling : Description, differences, and impacts on net primary production. *Global Biogeochemical Cycles*, 19(2), 2005.
- [24] S. Hoyer and J. Hamman. xarray : N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, 5(1), 2017.
- [25] B. v. d. Hurk, H. Kim, G. Krinner, S. I. Seneviratne, C. Derksen, T. Oki, H. Douville, J. Colin, A. Ducharne, F. Cheruy, et al. Ls3mip (v1. 0) contribution to cmip6 : the land surface, snow and soil moisture model intercomparison project–aims, setup and expected outcome. *Geoscientific Model Development*, 9(8) :2809–2832, 2016.
- [26] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. The ncep/ncar 40-year reanalysis project. *Bulletin of the American meteorological Society*, 77(3) :437–472, 1996.
- [27] H. Kim, S. Watanabe, E. Chang, K. Yoshimura, J. Hirabayashi, J. Famiglietti, and T. Oki. Global soil wetness project phase 3 atmospheric boundary conditions (experiment 1)[data set], data integration and analysis system (dias), 2017.

- [28] G. Lasslop, V. Brovkin, C. H. Reick, S. Bathiany, and S. Kloster. Multiple stable states of tree cover in a global land surface model due to a fire-vegetation feedback. *Geophysical Research Letters*, 43(12) :6324–6331, 2016.
- [29] G. Lasslop, K. Thonicke, and S. Kloster. Spitfire within the mpi earth system model : Model development and evaluation. *Journal of Advances in Modeling Earth Systems*, 6(3) :740–755, 2014.
- [30] D. M. Lawrence, R. A. Fisher, C. D. Koven, K. W. Oleson, S. C. Swenson, G. Bonan, N. Collier, B. Ghimire, L. van Kampenhout, D. Kennedy, et al. The community land model version 5 : Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, 11(12) :4245–4287, 2019.
- [31] C. Le Quéré, R. Andres, T. Boden, T. Conway, R. Houghton, et al. Global carbon budget 1959-2011. *Earth System Science Data*, 5 :165–185, 2013.
- [32] C. Le Quéré, R. M. Andrews, P. Friedlingstein, S. Slitch, J. Pongratz, et al. Global carbon budget 2017. *Earth System Science Data*, 10 :405–448, 2018.
- [33] C. Le Quéré, G. Peters, R. Andres, R. Andrews, T. Boden, P. Clais, P. Friedlingstein, et al. Global carbon budget 2013. *Earth System Science Data*, 6 :235–263, 2014.
- [34] F. Li and D. M. Lawrence. Role of fire in the global land water budget during the twentieth century due to changing ecosystems. *Journal of Climate*, 30(6) :1893–1908, 2017.
- [35] F. Li, X. Zeng, and S. Levis. A process-based fire parameterization of intermediate complexity in a dynamic global vegetation model. *Biogeosciences*, 9(7) :2761–2780, 2012.
- [36] S. Manabe. Climate and the ocean circulation : I. the atmospheric circulation and the hydrology of the earth’s surface. *Monthly Weather Review*, 97(11) :739–774, 1969.
- [37] T. Mauritsen, J. Bader, T. Becker, J. Behrens, M. Bittner, R. Brokopf, V. Brovkin, M. Claussen, T. Crueger, M. Esch, et al. Developments in the mpi-m earth system model version 1.2 (mpi-esm1. 2) and its response to increasing co2. *Journal of Advances in Modeling Earth Systems*, 11(4) :998–1038, 2019.
- [38] Met Office. *Cartopy : a cartographic python library with a matplotlib interface*. Exeter, Devon, 2010 - 2015.
- [39] T. D. Mitchell and P. D. Jones. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology : A Journal of the Royal Meteorological Society*, 25(6) :693–712, 2005.
- [40] M. A. Moritz, M.-A. Parisien, E. Batllori, M. A. Krawchuk, J. Van Dorn, D. J. Ganz, and K. Hayhoe. Climate change and disruptions to global fire activity. *Ecosphere*, 3(6) :1–22, 2012.
- [41] S. Piao, S. Sitch, P. Ciais, P. Friedlingstein, P. Peylin, X. Wang, A. Ahlström, A. Anav, J. G. Canadell, N. Cong, et al. Evaluation of terrestrial carbon cycle models for their

- response to climate variability and to co2 trends. *Global change biology*, 19(7) :2117–2132, 2013.
- [42] B. Poulter, D. Frank, E. Hodson, and N. Zimmermann. Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the co 2 airborne fraction. *Biogeosciences*, 8(8) :2027–2036, 2011.
- [43] S. S. Rabin, J. R. Melton, G. Lasslop, D. Bachelet, M. Forrest, S. Hantson, J. O. Kaplan, F. Li, S. Mangeon, D. S. Ward, et al. The fire modeling intercomparison project (fire-mip), phase 1 : experimental and analytical protocols with detailed model descriptions. *Geoscientific Model Development*, 10(3) :1175–1197, 2017.
- [44] J. Randerson, G. Van Der Werf, L. Giglio, G. Collatz, and P. Kasibhatla. Global fire emissions database, version 4.1 (gfedv4). *ORNL DAAC*, 2015.
- [45] R. Rew, E. Hartnett, J. Caron, et al. Netcdf-4 : Software implementing an enhanced data model for the geosciences. In *22nd International Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology*, volume 6, 2006.
- [46] U. Schulzweida and L. Kornblueh. Cdo climate data operators, 2012.
- [47] S. Seabold and J. Perktold. statsmodels : Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.
- [48] R. Séférian, P. Nabat, M. Michou, D. Saint-Martin, A. Voltaire, J. Colin, B. Decharme, C. Delire, S. Berthet, M. Chevallier, et al. Evaluation of cnrm earth system model, cnrm-esm2-1 : Role of earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, 11(12) :4182–4227, 2019.
- [49] A. A. Sellar, J. Walton, C. G. Jones, R. Wood, N. L. Abraham, M. Andrejczuk, M. B. Andrews, T. Andrews, A. T. Archibald, L. de Mora, et al. Implementation of uk earth system models for cmip6. *Journal of Advances in Modeling Earth Systems*, 12(4) :e2019MS001946, 2020.
- [50] J. Sheffield, G. Goteti, and E. F. Wood. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of climate*, 19(13) :3088–3111, 2006.
- [51] S. Sitch, P. Friedlingstein, N. Gruber, S. D. Jones, G. Murray-Tortarolo, A. Ahlström, S. C. Doney, H. Graven, C. Heinze, C. Huntingford, et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, 12(3) :653–679, 2015.
- [52] I. SREX. Managing the risks of extreme events and disasters to advance climate change adaptation. *A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*, edited by : Field, CB, Barros, V., Stocker, TF, Qin, D., Dokken, DJ, Ebi, KL, Mastrandrea, MD, Mach, KJ, Plattner, G.-K., Allen, SK, Tignor, M., and Midgley, PM, Cambridge University Press, Cambridge, UK, and New York, NY, USA, 2012.

- [53] K. Thonicke, A. Spessa, I. Prentice, S. P. Harrison, L. Dong, and C. Carmona-Moreno. The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions : results from a process-based model. *Biogeosciences*, 7(6) :1991–2011, 2010.
- [54] G. R. Van der Werf, J. T. Randerson, L. Giglio, G. Collatz, M. Mu, P. S. Kasibhatla, D. C. Morton, R. DeFries, Y. v. Jin, and T. T. van Leeuwen. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric chemistry and physics*, 10(23) :11707–11735, 2010.
- [55] M. J. Van Marle, S. Kloster, B. I. Magi, J. R. Marlon, A.-L. Daniau, R. D. Field, A. Arneth, M. Forrest, S. Hantson, N. M. Kehrwald, et al. Historic global biomass burning emissions for cmip6 (bb4cmip) based on merging satellite observations with proxies and fire models (1750–2015). *Geoscientific Model Development*, 10(9) :3329–3357, 2017.
- [56] T. Wu, Y. Lu, Y. Fang, X. Xin, L. Li, W. Li, W. Jie, J. Zhang, Y. Liu, L. Zhang, et al. The beijing climate center climate system model (bcc-csm) : The main progress from cmip5 to cmip6. *Geoscientific Model Development*, 12(4) :1573–1600, 2019.
- [57] C. Yue, P. Ciais, P. Cadule, K. Thonicke, S. Archibald, B. Poulter, W. Hao, S. Hantson, F. Mouillot, P. Friedlingstein, et al. Modelling the role of fires in the terrestrial carbon balance by incorporating spitfire into the global vegetation model orchidee–part 1 : simulating historical global burned area and fire regimes. *Geoscientific Model Development*, 7(6) :2747–2767, 2014.
- [58] C. Yue, P. Ciais, P. Cadule, K. Thonicke, and T. Van Leeuwen. Modelling the role of fires in the terrestrial carbon balance by incorporating spitfire into the global vegetation model orchidee–part 2 : Carbon emissions and the role of fires in the global carbon balance. *Geoscientific Model Development*, 8(5) :1321–1338, 2015.