



SORBONNE UNIVERSITÉ
ÉCOLE DOCTORALE 129 - SCIENCES DE L'ENVIRONNEMENT D'ÎLE-DE-FRANCE
LABORATOIRE DE MÉTÉOROLOGIE DYNAMIQUE

Thèse de doctorat pour obtenir le grade de Docteur de Sorbonne Université

Spécialité : Modélisation, méthodes numériques pour l'étude du système climatique

L'art et la science de l'ajustement des modèles de climat : de la théorie à la pratique

Présentée par

Maëlle COULON--DECORZENS

Encadrée par

Frédéric HOURDIN, Najda VILLEFRANQUE

Devant un jury composé de :

Hélène CHEPFER

Professeure des universités, Sorbonne Université, LMD

Jacques GAUTRAIS

Chargé de Recherche, CNRS, CRCA et Laplace

Gerhard KRINNER

Directeur de recherche, CNRS, IGE

Quentin LIBOIS

Chargé de Recherche, Météo-France, CNRM

Pierre NABAT

Chargé de Recherche, Météo-France, CNRM

Aurore VOLDOIRE

Ingénieure Divisionnaire des Travaux de Météo-France, CNRM

Frédéric HOURDIN

Directeur de Recherche, CNRS, LMD

Najda VILLEFRANQUE

Chargée de recherche, CNRS, CNRM

Présidente

Rapporteur

Rapporteur

Examineur

Examineur

Examinatrice

Directeur de thèse

Co-encadrante de thèse

Table des matières

I	Introduction	1
I.1	Contexte	1
I.1.1	Qu'est-ce qu'un modèle de climat et à quoi sert-il ?	1
I.1.2	Concrètement, qu'est-ce qu'il y a dans un modèle de climat ?	2
I.1.3	Comment comprend-on le climat avec un modèle ?	3
I.1.4	Comment savoir qu'un modèle de climat simule un climat réaliste ?	3
I.1.5	La question de l'ajustement des modèles de climat	5
I.1.5.1	Pourquoi a-t-on besoin d'ajuster les modèles de climat ?	5
I.1.5.2	Vers une science de l'ajustement des modèles	6
I.1.5.3	Les méthodes d'ajustement objectives	7
I.1.5.4	Stratégie d'ajustement du modèle climatique de l'IPSL	8
I.2	Positionnement de la thèse	9
II	Le modèle de circulation général LMDZ	13
II.1	Principes de modélisation des GCMs atmosphériques	13
II.1.1	Les équations primitives de la météorologie	15
II.1.2	Séparation d'échelles de l'écoulement atmosphérique	16
II.1.2.1	Décomposition de Reynolds	16
II.1.2.2	Séparation entre cœur dynamique et paramétrisations physiques	18
II.2	Quelques éléments de configuration de LMDZ6A	20
II.3	Les paramétrisations physiques de LMDZ	22
II.3.1	Turbulence de petite échelle	23
II.3.2	Paramétrisation de la convection peu profonde dans LMDZ	25
II.3.3	Schéma de condensation grande échelle	28
II.3.3.1	Paramétrisation du recouvrement vertical des précipitations	30
II.3.4	Paramétrisation de la convection profonde	31
II.3.4.1	L'efficacité de l'auto-conversion d'eau liquide nuageuse en pluie	32
II.3.4.2	Déclenchement de la convection profonde	33
II.3.4.3	La fermeture convective	34

II.3.4.4	Energie et puissance de soulèvement fournies par les thermiques	34
II.3.4.5	Le modèle de poches froides	35
II.3.4.6	Le déclenchement stochastique	36
II.3.5	Transfert radiatif	37
III	History Matching with Iterative Refocusing	39
III.1	Introduction	40
III.2	La méthode d'History Matching with Iterative Refocusing	42
III.2.1	Préambule	42
III.2.1.1	Quelques mots sur le cadre mathématique	42
III.2.1.2	Philosophie de la méthode	42
III.2.2	Définition du problème	43
III.2.3	Les étapes de la méthode HMIR	46
III.3	L'outil <code>htexplo</code>	49
III.3.1	Présentation générale de l'outil <code>htexplo</code>	49
III.3.2	Exemple d'application avec un modèle simple	52
III.3.2.1	Construction du modèle simple	52
III.3.2.2	Première expérience de tuning : une métrique et un paramètre	53
III.3.3	Convergence de l'expérience	59
III.3.3.1	Deuxième expérience de tuning du modèle conceptuel de climat	61
III.3.3.2	Exploitation des ensembles de simulations	62
III.3.3.3	Visualisation des espaces de paramètres libres	64
III.4	Tuning de LMDZ avec <code>htexplo</code>	66
III.4.1	Tuning dans le cadre 1D/LES	66
III.4.2	Tuning de LMDZ en global	69
III.4.2.1	Présentation des métriques	70
III.4.2.2	Résultats	72
IV	Tuning visant les observations du SIRTa : configura- tion de référence	75
IV.1	Introduction	75
IV.1.1	Histoire de l'expérience	76
IV.1.2	Plan	79
IV.2	Motivations	79
IV.2.1	Importance du couplage entre l'atmosphère et la surface	79
IV.2.2	Réduire les biais classiques dans la simulation du climat de surface par les GCMs	80
IV.2.3	Exploiter les données sur site	81

IV.3 Observations du SIRTA et configuration de LMDZ	84
IV.3.1 Description du site du SIRTA	84
IV.3.2 Traitement des données ReOBS et des sorties de simulation	84
IV.3.3 Une première configuration de LMDZ	86
IV.3.3.1 Grille zoomée	87
IV.3.3.2 Guidage	88
IV.3.3.3 Durée des simulations, spin-up et choix de l'année	89
IV.3.3.4 La question de la représentativité des observations du site du SIRTA	91
IV.3.3.5 Diagnostique de l'albédo au SIRTA et forçage de l'albédo dans LMDZ	92
IV.3.3.6 Sensibilité des simulations aux aérosols	95
IV.4 Variabilité interne	95
IV.4.1 Ensemble aux états initiaux perturbés	97
IV.4.2 Sensibilité au guidage	99
IV.5 Variabilité interne des précipitations	101
IV.5.1 Variabilité interne des précipitations convectives	102
IV.5.2 Étude régionale des précipitations autour du SIRTA	103
IV.6 Couper la rétro-action de l'humidité du sol sur l'atmosphère en imposant le facteur d'aridité	107
IV.6.1 Diagnostic du facteur d'aridité dans les observations du SIRTA	107
IV.6.1.1 Diagnostic de l'évaporation potentielle à partir des observations du SIRTA	108
IV.6.1.2 Diagnostic du facteur d'aridité	110
IV.6.1.3 Comparaison du facteur d'aridité diagnostiqué et simulé au SIRTA	114
IV.6.2 Simulations avec un facteur d'aridité imposé	116
IV.6.2.1 Derniers ajustements sur la configuration	116
IV.6.2.2 Impact du découplage de l'hydrologie au SIRTA	120
IV.7 Configuration finalement retenue	126
IV.8 Conclusion	126

V Tuning visant les observations du SIRTA : première expérience de tuning	129
V.1 Protocole expérimental	129
V.1.1 Choix des paramètres libres de l'expérience	130
V.1.2 Préconditionnement avec les métriques uni-colonne	130
V.1.3 Choix des métriques au SIRTA et de leurs tolérances à l'erreur	132

V.2 Résultats	136
V.2.1 Analyse des meilleurs simulations	143
V.2.2 Evaluation sur d'autres années	144
V.3 Place dans la chaîne de tuning actuelle	149
V.3.1 Effet du préconditionnement 1D sur les simulations au SIRTA	149
V.3.2 3D	154
V.4 Conclusion	154
VI Rayonnement	155
VI.1 Transfert radiatif avec ecRad	155
VI.1.1 Prise en compte des nuages pour le transfert radiatif dans ecRad	156
VI.1.1.1 Hypothèses de prise en compte des nuages disponibles dans ecRad	157
VI.1.1.2 Traitement des hypothèses par les solveurs	161
VI.1.2 ecRad et le cadre SCM/LES	167
VI.1.2.1 Calculs radiatifs de référence par Monte Carlo sur des LES	169
VI.1.2.2 Utiliser les calculs de référence Monte Carlo pour évaluer et tuner les modèles de transfert radiatif de ecRad	172
VI.2 Experience compensation d'erreur rayonnement - nuage	175
VI.3 Rayonnement 3D	202
VII Synthèse sur la pratique du tuning avec htexplo pour la modélisation du climat	203
VII.1 Discussion du 25 septembre	203
VII.2 Reflexions 19 juin 2025	203
VII.3 Reflexions 12 Novembre 2024	204
VIII Conclusion	207
A Résumé des paramètres libres	223
B Emulation par processus gaussiens	229
B.1 Cadre théorique des Processus Gaussiens	229
B.1.1 Définition d'un processus stochastique	229
B.1.2 Définition d'un processus gaussien	230
B.1.3 Apprentissage	231

B.1.4	Prédiction et incertitude de prédiction	232
B.2	Mise en œuvre dans htexplo	234
B.2.1	Apprentissage	234
B.2.2	Prédiction	235
B.2.2.1	Si le point auquel on réalise la prédiction est strictement un des points sondés	236
B.2.2.2	Si le point auquel on réalise la prédiction est infiniment proche de l'un des points sondés	237
B.2.2.3	Si le point auquel on réalise la prédiction est très loin de tous les points sondés	237
B.2.3	Vérification	237

Table des figures

II.1	Illustration d'un état atmosphérique simulé par LMDZ avec une grille grossière (image réalisée par Laurent Fairhead et Frédéric Hourdin en 1995). On voit le découpage de l'atmosphère en mailles, avec une discrétisation en intervalles réguliers de longitude et latitude. L'épaisseur de l'atmosphère est très exagérée sur cette représentation. Les couleurs représentent la température dans chaque maille et les flèches les vents horizontaux	14
II.2	Grille horizontale du modèle en projection Eckert-Greifendorff qui conserve l'aire des mailles. Une maille sur est deux colorée, soit en bleu foncé pour les mailles océaniques, soit en une couleur correspondant à l'altitude moyenne de la maille. On voit l'aire des mailles diminuer à mesure que l'on s'éloigne de l'équateur. . . .	21
II.3	Epaisseur des niveaux verticaux d Z (km) en fonction de l'altitude du milieu du niveau Z (km) pour les grilles à 79 niveaux en noir et 95 niveaux en rouge, issue de Hourdin et al. (2021). A gauche les profils sont tracés pour toute la colonne atmosphérique et à droite dans les trois premiers kilomètres	21
II.4	(a) Image photo-réaliste réalisé avec htrdr(Villefranque et al., 2019) d'une scène de cumulus du cas ARMCU simulée par la LES Meso-NH. Les cellules de convection peu profonde sont schématisées en rose, et la turbulence de petite échelle en violet. Les nuages de cumulus se situent dans la couche d'inversion au dessus des ascendances d'airs. (b) Schéma de la représentation de la couche limite convective (turbulence petite échelle et thermique effectif) dans une colonne de LMDZ. (c) Profils idéalisés d'une couche limite atmosphérique convective (adapté de Hourdin (2005))	26
II.5	(a) Distribution du déficit de saturation $Q(s)$ à 1200 d'altitude, pour la 8 ^e heure d'un cas de cycle diurne de cumulus continentaux au site ARM dans les grandes plaines états-uniennes (le cas ARMCU). L'histogramme est calculé sur les données d'une simulation LES de ce cas. La courbe verte est la gaussienne de l'environnement ; la courbe rouge celle du thermique et la bleue, distribution retenue pour la paramétrisation statistique des nuages dans LMDZ, la somme des deux (Jam et al., 2013) (b) Profil de la fonction $\xi(P)$ utilisée pour le calcul de l'écart type de la distribution de $Q(s)$ quand les thermiques ne sont pas actifs (adapté de Madeleine et al. (2020))	29

II.6	Schéma décrivant la paramétrisation de condensation grande échelle prenant en compte le recouvrement vertical des précipitations, sur quelques couches nuageuses d'une colonne du modèle. Il n'est pas montré ici (1) le cas où le nuage est recouvert par un nuage plus haut (disjoint), où la couche du sommet du nuage pourra recevoir un flux de précipitation ciel-clair venant de ce dernier et sur une fraction calculée par le modèle de recouvrement vertical des nuages (2) le cas où le flux de précipitation ciel-clair est trop faible par rapport à la fraction de maille qu'il occupe (d'un facteur RI) et commence à décroître. Sur le schéma, les couches de nuage sont disposées de manière à se recouvrir au maximum, c'est-à-dire à minimiser leur couverture nuageuse.	30
II.7	Photos et schémas illustrant les concepts véhiculés dans la construction du modèle de convection et de son couplage avec les processus sous-nuageux. En haut à gauche : photo de cumulonimbus au-dessus des plaines du Texas Panhandle, avec localisation de l'enclume, de la base de l'ascendance adiabatique et des proches froides (juin 2023 — Photo : Jack Sillin, utilisée permission) ; en haut à droite : cumulonimbus vu d'en haut avec le sommet de l'enclume qui s'étale sous la tropopause, les nuages bourgeonnant qui overshoot au milieu et des cumulus de couche limite autour (photo prise par l'astronaute Tim Peake depuis la Station spatiale internationale durant la mission Principia (2016), crédit : ©ESA/NASA, utilisée conformément aux conditions d'utilisation du site de l'ESA) ; en bas à gauche : photo d'un haboob au Sahel, tempête de poussières créée par le front de rafale d'une poche froide (crédit : © Françoise GUICHARD / Laurent KERGOAT / CNRS Images) ; en bas à droite : schéma conceptuel du schéma de convection et de son couplage avec les processus sous-nuageux.	31
III.1	Schéma des espaces de paramètres libres Λ , \mathcal{A} , \mathcal{A}_{f_1} et \mathcal{A}_{f_2}	45
III.2	Schéma de l'évolution des espaces de paramètres libres au cours de deux itérations avec la méthode d'HMIR	47
III.3	Logigramme de l'outil htexplo	51
III.4	Relation entre le rayonnement solaire et infrarouge global au sommet de l'atmosphère et le paramètre libre FALLV dans quatre modèles de climat	53
III.5	Evolution des espaces des paramètres libres et des métriques au cours des trois premières vagues de tuning de l'exemple simple (une métrique f deux paramètres)	55
III.6	Illustration de la sensibilité de la prédiction par les émulateurs au nombre de points sondes (exemple métrique OLR)	57
III.7	Figure automatique de l'outil htexplo : Plots_Metrics.pdf	58
III.8	Figure automatique de l'outil htexplo : Plots_L00.pdf	59
III.9	Evolution au cours des vagues de la fraction du NROY par rapport à l'hypercube initial dans l'exemple simple à une métrique	60
III.10	Evolution au cours des vagues de la fraction du NROY par rapport à l'hypercube initial dans l'exemple à deux puis trois métriques	62
III.11	Evaluation de la dispersion des scores des PPEs de l'exemple à deux métriques et quatre paramètres	63
III.12	Zoom sur la matrice d'in vraisemblance	64
III.13	Matrices d'in vraisemblance des cinquième et dixième itérations de l'expérience à deux métriques et quatre paramètres	65
III.14	Profils de température potentielle, humidité spécifique et fraction nuageuse de trois cas de convection peu profonde (ARMCU, RICO et SANDU) pour quatre PPEs issu d'une expérience de tuning contraintes par des métriques sur ces mêmes cas, et avec une grille verticale à 95 niveaux (Hourdin et al., 2021)	68

III.15	Métriques radiatives utilisées pour le tuning de LMDZ	72
III.16	Métriques radiatives utilisées pour le tuning de LMDZ	72
III.17	Métriques radiatives utilisées pour le tuning de LMDZ	73
III.18	Métriques radiatives utilisées pour le tuning de LMDZ	73
III.19	Métriques radiatives utilisées pour le tuning de LMDZ	73
IV.1	Biais de température atmosphérique à deux mètres dans les ensembles multi-modèle de CMIP5 et CMIP6	82
IV.2	Biais de température atmosphérique à deux mètres dans différentes versions du modèle de l'IPSL, issue de Boucher et al. (2020)	83
IV.3	Photos du site du SIRTa	85
IV.4	Grille utilisée pour les simulations zoomées	87
IV.5	Humidité du sol sur la période 2015-2020 dans les observations et cinq simulations avec la physique LUDO-L95	90
IV.6	Evolution de l'albédo du 10 au 25 avril 2018 diagnostiqué sur les données horaires du SIRTa	93
IV.7	Evolution de l'albédo en moyenne mensuelle sur l'année 2018, diagnostiqué à partir des observations du SIRTa	94
IV.8	Evolution des albédos dans le visible et le proche infra-rouge sur l'année 2018 en fréquence horaire, dans les simulations avec ou sans albédo imposé sur la maille du SIRTa	95
IV.9	Evolution en moyenne mensuelle sur l'année 2018 (a) du flux solaire montant à la surface (b) du flux solaire descendant à la surface (c) du flux solaire net à la surface (d) du flux radiatif total (solaire + IR) net à la surface, dans les simulations avec et sans albédo imposé sur la maille du SIRTa	96
IV.10	Evolution en moyenne mensuelle de janvier à septembre 2018 (a) du flux solaire ciel-clair descendant à la surface (b) du flux radiatif totale (solaire + infrarouge) net à la surface pour une simulation de contrôle sans aérosol, et la même simulation avec des aérosols climatiques	96
IV.11	Evolution en moyenne mensuelle (a) du flux solaire descendant à la surface (b) des précipitations (c) du flux de chaleur latente (d) de l'humidité du sol à 10 cm de profondeur de deux ensemble perturbé aux états initiaux (PISE) (un en changeant les états initiaux de LMDZ et l'autre d'ORCHIDEE)	98
IV.12	Evolution en moyenne mensuelle (a) du flux solaire descendant à la surface (b) des précipitations (c) du flux de chaleur latente (d) de l'humidité du sol à 10 cm de profondeur (e) de l'humidité spécifique à 2 mètres et (f) de la température atmosphérique à 2 mètres, de janvier à août 2018. Les trois PISE sont tracées (guidage en vent, vent + température et vent + température + humidité relative)	100
IV.13	Evolution en moyenne mensuelle sur l'année 2018 (a) des précipitations (b) des précipitations issue de la paramétrisation de la convection profonde (c) des précipitations issu de la paramétrisation grande échelle, sur la maille du SIRTa ; pour le deux PISE où les états initiaux de LMDZ et ORCHIDEE sont perturbés.	102
IV.14	Cartes des précipitations moyennes du mois de juin 2018, pour les réanalyses Coméphore regrillées sur la grille du modèle, la simulation de contrôle et deux membre du PISE atmosphérique. Les cartes des précipitations convectives et grande échelles de la simulation de contrôle sont également tracées.	104

IV.15	Evolution en moyenne mensuelle sur l'année 2018 des précipitations moyennes (a) sur 9 mailles (b) sur 25 mailles (c) sur 49 mailles autour de la maille du SIRTa ; (d) des précipitations issues de la paramétrisation de la convection profonde en moyenne sur 25 mailles (e) des précipitations issue de la paramétrisation des précipitations grandes échelles en moyenne sur 25 mailles ; dans la simulation de contrôle et les deux PISE	105
IV.16	Evolution en moyenne mensuelle sur l'année 2018 des écarts-type spatiaux des précipitations sur 25 mailles (a) précipitations totales (b) précipitations issues de la paramétrisation de la convection profonde (c) précipitations issues de la paramétrisation de condensation grande échelle. Dans la simulation de contrôle et les deux PISE	106
IV.17	Evolution à fréquence horaire sur le mois de janvier 2018 (a) du coefficient d'échange κ diagnostiqué avec l'Equation IV.20 avec masque (rouge) et sans masque (noir) (b) de l'évaporation potentielle dignostiquée avec l'Equation IV.15 dans les observations du sirta (noir), simulée par le modèle (jaune) et diagnostiquée avec l'Equation IV.15 dans le modèle (bleu) (c) du facteur d'aridité diagnostiqué avec l'Equation IV.10 (noir) et simulé par le modèle (jaune) (d) du flux de chaleur latente observé au SIRTa (noir) et reconstruit avec l'Equation IV.10 (rouge). La simulation est réalisée avec la physique CMIP6-L79, l'albédo imposé sur la maille du SIRTa, un guidage en vent, température et humidité.	111
IV.18	Mêmes graphiques que la Figure IV.17 pour le mois de Juin 2018	112
IV.19	Evolution journalière sur l'année 2018 (a) du facteur d'aridité simulé et diagnostiqué au SIRTa, avec test des moyennes glissantes, et (b) de l'humidité du sol à 10 cm, 30 cm et 100 cm de profondeur, dans les observations du SIRTa, la simulation de contrôle et son PISE associé	114
IV.20	Comparaison des simulations avec et sans facteur d'aridité imposé sur la maille du SIRTa, ainsi qu'avec un facteur d'aridité imposé dans un rayon de 100 km autour du SIRTa	117
IV.21	Carte du facteur d'aridité, en moyenne mensuelle pour le mois de juillet 2018, dans la simulation où le facteur d'aridité est imposé dans les mailles situées dans un rayon de 100 km autour de la maille du SIRTa. La simulation est effectuée avec la physique CMIP6-L79, un guidage en vent, température et humidité et l'albédo imposé sur la maille du SIRTa.	118
IV.22	Evolution en moyenne journalière de l'humidité du sol sur l'année 2018 à 10 cm, 30 cm et 100 cm de profondeur, sur la maille du SIRTa : (a) pour la configuration avec β imposé sur les mailles dans un rayon de 100 km autour de la maille du SIRTa ; (b) pour la configuration avec en plus l'humidité du sol imposée à saturation en dessous de 1,3 mètres de profondeur et une discrétisation du sol à 22 couches	118
IV.23	Comparaison des simulations avec le facteur d'aridité imposé sur 100 km et celles ayant en plus l'humidité du sol imposée à saturation en dessous de 1,3 mètres de profondeur, un sol discrétisé à 22 couches ainsi qu'un indice de surface foliaire journalier imposé	119
IV.24	Comparaison de deux PISE avec une configuration où le facteur d'aridité imposé sur 100 km, l'humidité du sol imposé à saturation sous 1,3 mètres et l'indice de surface foliaire imposé. Les deux PISE diffèrent dans leur constante de guidage.	120
IV.25	Comparaison de deux PISE, celui de la configuration finale (avec notamment β imposé sur une région autour de la maille du SIRTa), et celui de la configuration sans β imposé ; sur les variables liées aux flux turbulents.	121

IV.26	Même figure que Figure IV.25 pour : (a) le flux solaire net à la surface (W.m^{-2}); (b) le flux infra-rouge net à la surface (W.m^{-2}); (c) l'effet radiatif solaire des nuages à la surface (W.m^{-2}); (d) l'effet radiatif infra-rouge des nuage à la surface (W.m^{-2}); (e) le flux infra-rouge descendant à la surface (W.m^{-2}); (f) le flux infra-rouge descendant à la surface ciel clair (W.m^{-2}); (g) le flux infra-rouge montant à la surface.	123
IV.27	Même figure que Figure IV.25 pour : (a) la quantité d'eau liquide intégrée sur la verticale (<i>Liquid Water Path</i> en g.m^{-2}); (b) la couverture nuageuse des nuages bas; (c) la température atmosphérique à deux mètres ($^{\circ}\text{C}$); (d) l'humidité spécifique à deux mètres (g.kg^{-1}); (e) l'humidité relative à deux mètres (%)	124
IV.28	Effet du β imposé sur la variabilité interne des précipitations	125
V.1	Evolution de la fraction du NROY restante en fonction du numéro de la vague sur les 40 vagues du préconditionnement 1D. En bleu pour l'expérience avant restriction de l'espace de paramètres initial, et en rouge après restriction.	133
V.2	Evolution en moyenne mensuelle sur l'année 2018 de (a) l'effet radiatifs SW des nuages à la surface (b) effet radiatif LW des nuages à la surface (c) des précipitations moyennes sur 25 mailles autour de la maille du SIRTa. En gris claire l'ensemble aux paramètres perturbées de la vague 41 de l'expérience, en gris foncé l'ensemble aux conditions initiales perturbé, et bleu la simulation de contrôle de l'expérience (configuration SIRTa), en noir les observations du SIRTa et en noir pointillé les observations radar pour les précipitations	134
V.3	Evolution en moyenne journalière du 22 au 30 juin 2018 de (a) l'effet radiatif SW des nuages à la surface (b) des précipitations moyennes sur 25 mailles autour de la maille du SIRTa. En gris claire l'ensemble aux paramètres perturbées de la vague 41 de l'expérience, en gris foncé l'ensemble aux conditions initiales perturbé, et bleu la simulation de contrôle de l'expérience (configuration SIRTa), en noir les observations du SIRTa et en noir pointillé les observations radar pour les précipitations	135
V.4	Score (Equation V.1) et score par métrique pour l'ensemble des simulations du PPE41 en bleu et du PPE42 en marron. Les croix noires marquent les valeurs minimales et maximales des scores des deux PPE. Les valeurs des scores des 10 meilleures simulations sont indiqués avec des symboles.	137
V.5	Evolution en moyenne journalière du 22 mai au 30 juin 2018 (a) de l'effet radiatif SW des nuages à la surface, (b) de la couverture nuageuse intégrée. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa.	138
V.6	Evolution en moyenne mensuelle sur l'année 2018 (a) de l'effet radiatif SW des nuages à la surface, (b) de l'effet radiatif LW des nuages à la surface, (c) des précipitations moyennes sur 25 mailles, (d) de l'écart-type spatial des précipitations sur 25 mailles, (e) des précipitations issues de la paramétrisation de convection profonde (moyennes sur 25 mailles), (f) des précipitations issues de la paramétrisation de condensation grande échelle. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa.	139
V.7	Evolution en moyenne mensuelle sur l'année 2018 (a) de la quantité d'eau liquide (b) de la quantité d'eau glacée (c) de la couverture nuageuse intégrée totale (d) de la couverture nuageuses des nuages bas (e) de la couverture nuageuse des nuages moyens (f) de la couverture nuageuse des nuages haut. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle.	140

V.8	Evolution en moyenne journalière du 22 mai au 30 juin 2018 (a) des précipitations en moyennes sur 25 mailles, (b) des précipitations issu de la paramétrisation de convection profonde, en moyenne sur 25 mailles, (c) des précipitations issues de la paramétrisation de condensation grande échelle, en moyenne sur 25 mailles. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir pointillées les observations radar.	141
V.9	Evolution en moyenne mensuelle sur l'année 2018 (a) du flux radiatif net à la surface (b) de la somme des flux turbulents de surface (c) du flux de chaleur sensible (d) de la température à 2 mètre. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa	142
V.10	Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 41, en bleu claire les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa	145
V.11	Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 41, en bleu claire les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa	146
V.12	Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 41, en bleu claire les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa	147
V.13	Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 41, en bleu claire les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa	148
V.14	Evolution en moyenne mensuelle sur l'année 2018 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa	150
V.15	Evolution en moyenne mensuelle sur l'année 2018 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa	151
V.16	Evolution en moyenne journaliere du 22 mai au 31 juin 2018 de (a) CRESWSFC [<i>c'est plus visible sur le CRELWSFC donc je changerais</i>] (b) /bin/bash : q : command not found En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa	152
V.17	Evolution en moyenne mensuelle sur l'année 2018 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa	153

VI.1	Schéma issu de Grailet et al. (2025), illustrant l'architecture de ecRad. Les 5 grandes boîtes correspondent aux 5 modules de ecRad. Les boîtes pointillées correspondent aux entrées de ecRad. La boîte "Thermodynamics" correspond aux profils de température et de pression aux inter-couches. Les flèches en pointillé modélisent les structures de données internes d'ecRad. La boîte en pointillés en arrière-plan représente les étapes d'interpolation utilisées pour passer de la grille du modèle d'atmosphère à celle de ecRad et vice versa.	156
VI.2	Ce tableau est issu de Hogan and Bozzo (2018). Il répertorie les principales options possibles pour les différents modules de ecRad. En gras sont indiqués les options choisies pour la version de ecRad intégrée dans l'IFS.	157
VI.3	Schéma de trois hypothèses possibles de recouvrement des nuages dans les colonnes d'un GCM, issu de Hogan and Shonk (2008). A gauche le recouvrement aléatoire, au milieu le recouvrement exponentiel-aléatoire avec une facteur w égale à 0.6 pour toutes les couches de nuage adjacent et à droite recouvrement maximum-aléatoire. La ligne verticale en pointillé indique la couverture nuageuse totale de la colonne.	158
VI.4	Rendu physiquement réaliste d'un stratocumulus complètement couvrant, à gauche vu du dessous et à droite vu du dessus. On voit des zones plus claires et d'autres plus foncées, traduisant des différences d'épaisseur optique du nuage. Vues du dessous, les zones plus claires correspondent à des zones d'épaisseur optique plus fines. Vu du dessus, c'est l'inverse. Le stratocumulus est issu du cas FIRE simulé avec le modèle LES Meso-NH Lafore et al. (1998); Lac et al. (2018) et la visualisation est rendue possible par le code de transfert radiatif htrdr développé dans la thèse de Villefranque (2019).	159
VI.5	Schéma issu de Schäfer et al. (2016), représentants les effets radiatifs 3D associés aux nuages dans le solaire. (a) <i>side leakage or escape</i> (b) <i>side illumination</i> ou effet d'ombre (c) <i>entrappment</i> ou piégeage	160
VI.6	Schéma issu de Villefranque (2019) (a) nuage homogène avec recouvrement maximal, ce qui correspond aux hypothèses du schéma de transfert radiatif de LMDZ dans sa version CMIP6, (b) nuage homogène avec recouvrement exponentiel (c) nuage hétérogène illustrant le solver TripleClouds avec un recouvrement exponentiel (d) nuage hétérogène illustrant la méthode McICA avec un recouvrement exponentiel	161
VI.7	Schéma représentant l'approximation des colonnes indépendantes (ICA). Cette approximation stipule que le flux radiatif moyen d'une colonne atmosphérique est égal à la moyenne des flux radiatifs calculés indépendamment dans les N sous-colonnes découpant la colonne initiale (où chaque maille de chaque sous-colonne est optiquement homogène); lorsque le nombre de sous-colonnes tend vers l'infini. L'image des nuages est un rendu physiquement réaliste effectué avec htrdr (Villefranque et al., 2019) de la 8ème heure de la simulation du cas ARMCU effectuée par le modèle LES Meso-NH.	162
VI.8	Schéma représentant la méthode de McICA pour le calcul du transfert radiatif. Etape 1 : Les profils d'entrées, fournis par le GCM, associés aux hypothèses de recouvrement vertical et d'hétérogénéité horizontale des nuages, sont utilisés pour générer N sous-colonnes atmosphériques. Etape 2 : Un tirage aléatoire est réalisé pour déterminer la longueur d'onde associée à chaque sous-colonne, puis le transfert radiatif monochromatique est résolu dans chaque sous-colonne avec un modèle à deux flux. Etape 3. Les profils de flux radiatifs dans les domaines solaires (ou IR) finaux sont calculés comme la moyenne des profils des sous-colonnes.	162

VI.9 Schéma adapté de Shonk and Hogan (2008) (a) fonction de densité de probabilité p_{Q_c} du contenu en eau condensée q_c , de moyenne $\overline{q_c}$ (b) fonction de répartition associée. q_c^{fin} (respectivement $q_c^{\text{épais}}$) est la valeur du contenu en eau condensée du nuage optiquement fin (respectivement épais), situé à la distance $(1 - f^{\text{fin}})\overline{q_c}$ (respectivement $(f^{\text{épais}} - 1)\overline{q_c}$) de la moyenne. La valeur asymptotique de la fonction de répartition de l'eau condensée est la fraction nuageuse α	164
VI.10 Schémas représentant les flux radiatifs paramétrisés par le solveur SPARTACUS. (a) Lien entre processus et flux radiatifs paramétrisés dans SPARTACUS. Les trois flèches oranges représentent les flux radiatifs déjà paramétrisés par le solveur TripleClouds. Les flèches jaunes et rouges représentent les flux radiatifs liés aux effets 3D, les flèches rouges pour les flux radiatifs spécifiques à l'IR. (b) et (c) Schéma adapté de Hogan et al. (2016) représentant les échanges radiatifs modélisés dans SPARTACUS entre deux régions a et b, qui peuvent être du ciel-clair et/ou des nuages optiquement fin et/ou épais. Les flèches pointillées schématisent les échanges déjà modélisés dans TripleClouds, celles en traits pleins les échanges ajoutés dans Spartacus (b) Le flux radiatif arrivant en haut d'une région par diffusion peut être soit réfléchi (R) soit transmis (T) dans la même région (R_{aa} , R_{bb} , T_{aa} , T_{bb}) ou dans une autre région (R_{ab} , R_{ba} , T_{ab} , T_{ba}) (c) l'émission radiative dans une région dans l'IR (E) contribue au flux radiatif ascendant (E^{\uparrow}) et descendant (E^{\downarrow}) de la même région (E_{aa} , E_{bb}) ou d'une autre région (E_{ab} , E_{ba}).	165
VI.11 Schéma issu de Fielding et al. (2020) montrant la disposition des nuages (en noir) dans des pixels de $1 \times 1 \text{ km}^2$ pour différentes fraction nuageuses (colonne) (a-e) en utilisant la formulation VI.8 avec une taille effective des nuages $C_s = 1 \text{ km}$ et (f-j) en utilisant la formulation VI.9 avec un espacement effectif des nuages $C_x = 4.6 \text{ km}$. On voit le nombre de nuages augmenter avec la fraction nuageuse (a-c) puis le nombre de trous diminuer avec la fraction nuageuse (d-e) lorsque l'on utilise la formulation avec une taille effective des nuages fixe. Lorsque l'on utilise la formulation avec un espace effectif fixe, on voit une augmentation de la taille des nuages avec la fraction nuageuse (f-g) puis une diminution de la taille des trou avec l'augmentation de la fraction nuageuse (i-j)	166
VI.12 Schéma issu de Hogan et al. (2019) représentant les différents traitements possibles pour le piégeage dans SPARTACUS. Deux nuages de fraction 0,5 et épais d'une couche sont situés à la couche 3 et à la couche 6 de la colonne atmosphérique. Leur recouvrement est aléatoire. L'atténuation du rayonnement solaire descendant par les nuages est montrée avec les différents niveaux de gris. (a) cas sans piégeage (b) avec traitement explicite du piégeage et (c) avec traitement maximum du piégeage.	167
VI.13 Schéma illustrant le cadre de travail SCM/LES (ou 1D/LES), pour l'évaluation, le développement et le tuning des GCMs. Les sorties de simulation d'une colonne de GCM peuvent être comparées aux sorties de simulations LES tridimensionnelles qui servent de références pour les variables de physique atmosphériques et nuageuses, et pour le tuning des paramètres libres du GCM (Couvreur et al., 2021; Hourdin et al., 2021). Les champs de LES convertis en propriétés optiques sont utilisés pour calculer des références radiatives par Monte Carlo (MC3D), qui peuvent ensuite être comparées à un calcul de transfert radiatif à deux flux utilisant comme profil d'entrée le profil moyen de la LES 3D (Villefranche et al., 2021) ou le profil du SCM (ici).	168

VI.14	Schéma illustrant l'algorithme analogue de Monte Carlo. (a) Quelques chemins parcourus par les photons dans la colonne atmosphériques sont schématisés ici. Un capteur transparent chargé de compter les photons arrivant sur sa surface (depuis le haut ou le bas) est placé à l'altitude z_0 . L'image utilisée est un rendu photo réaliste produit par htrdr sur une scène de la simulation LES du cas ARMCU par Méso-NH. (b) Visualisation des chemins des photons (en blanc et vue de face) effectivement parcourus lors d'un calcul de transfert radiatif par Monte Carlo avec htrdr dans une portion de champ de cumulus. Les nuages sont repérables aux endroits où les photons sont fortement diffusés (diffusion de Mie) et changent ainsi souvent de direction sur de courtes distances.	170
-------	---	-----

Chapitre I

Introduction

Remarque lecture : Ce chapitre est très abouti - reste du travail sur les exemples, la biblio et l'orthographe

I.1 Contexte

I.1.1 Qu'est-ce qu'un modèle de climat et à quoi sert-il ?

Un modèle de climat est un modèle numérique dont l'objectif est la compréhension du climat et sa prévision. Il est souvent composé de plusieurs modèles qui communiquent entre eux : modèle d'atmosphère, d'océan et de surface des continents. Les modèles d'atmosphère et d'océans sont des GCMs (*General Circulation Model*), c'est-à-dire des modèles qui simulent la circulation du fluide (eau ou air) sur l'ensemble du globe. Les GCMs atmosphériques sont nés dans les années 1950, en même temps que les premiers ordinateurs, d'abord pour prévoir le temps, puis pour simuler et étudier le climat. Que ce soit pour une application en météorologie ou en climat, le GCM va toujours simuler l'état de l'atmosphère (température, humidité, nuages, etc.) toutes les quelques secondes à minutes, c'est-à-dire qu'il va toujours simuler la météorologie. La différence consiste en la manière dont est faite l'expérience. Pour une application en prévision du temps, l'état initial de la simulation va être très important, et les simulations vont être très courtes, typiquement d'une quinzaine de jours. Pour une application climatique, ce sont les états des régimes du système sur des échelles à minima décennales qui vont importer. Le climat d'une planète est alors défini comme les statistiques des grandeurs atmosphériques (température, humidité, nuages, etc.) sur ces échelles de temps. Autrement dit, le climat est une propriété émergente de la météorologie simulée par le modèle, mais il n'est pas modélisé directement : ce sont les processus météorologiques qui le sont.

Dans le contexte du changement climatique actuel, les modèles de climat ont deux objectifs principaux. (1) Fournir les projections climatiques les plus fiables et utiles possibles pour informer la société du changement climatique (2) Nous aider à comprendre le fonctionnement du climat et du changement climatique, et à comprendre l'impact des processus atmosphériques sur le climat et leur évolution avec le changement climatique. Le développement des modèles de climat est toujours sous-tendu par ces deux objectifs. Depuis une trentaine d'années, le calendrier de la recherche est rythmé par les exercices d'intercomparaison des modèles de climat (CMIP pour *Coupled Model Intercomparison Project*) en phase avec les rapports du GIEC (Groupe d'experts Intergouvernemental sur l'Évolution du Climat). La place des modèles de climat dans la recherche est d'autant plus importante que c'est le seul moyen d'expérimentation disponible pour comprendre le climat dans sa globalité, puisqu'il n'existe pas de méthode permettant à ce jour de reproduire l'ensemble des processus responsables du climat dans un laboratoire.

I.1.2 Concrètement, qu'est-ce qu'il y a dans un modèle de climat ?

Les modèles de climat sont basés sur une représentation discrétisée de l'atmosphère, des océans et des surfaces des continents. Ces trois milieux sont découpés spatialement en maille, qui font typiquement $100 \times 100 \text{ km}^2$ sur l'horizontal, pour quelques mètres à kilomètres de haut. Le modèle calcule alors dans chaque maille et à chaque pas de temps (toutes les secondes à minutes) les valeurs prises par les variables d'états du modèle, qui servent à décrire le milieu. Ces calculs suivent un ensemble d'équations décrivant la physique et parfois la chimie et la biologie du système modélisé.

Dans les GCMs atmosphériques, les mouvements de l'air sont décrits par les équations primitives, version modifiée des équations de Navier-Stokes avec l'hypothèse que l'air est un fluide parfait s'écoulant dans une mince couche sphérique en rotation. Ces équations sont discrétisées en temps et en espace et sont résolues explicitement par le modèle : c'est ce que l'on appelle le cœur dynamique du modèle. De nombreux processus atmosphériques ne sont pas décrits par ces équations, soit parce qu'ils ont lieu à des échelles spatiales inférieures à la taille des mailles du modèle (c'est le cas de la formation des nuages par exemple), soit parce qu'ils sont d'une nature différente (comme pour le rayonnement et la thermodynamique). Ces deux catégories de processus sont alors prises en compte par un ensemble de sous-modèles appelés paramétrisations : c'est ce que l'on appelle la physique du modèle. Ce découpage en deux parties distinctes, cœur dynamique et paramétrisations physiques, rejoint un découpage spatial des équations. La dynamique assure les transferts horizontaux ainsi que la partie résolue ou grande échelle des échanges verticaux, tandis que les paramétrisations physiques assurent les transferts verticaux non résolus par le cœur dynamique. Le même découpage existe dans les GCMs océaniques.

Les paramétrisations modélisent l'effet des processus atmosphériques non résolus par le cœur dynamique sur les variables d'état du modèle (température, humidité, vent etc.) via des équations heuristiques alliant théorie, phénoménologie et statistiques. Elles sont appelées paramétrisations par la communauté scientifique sans doute parce qu'elles contiennent un certain nombre de paramètres. Les processus sous-mailles paramétrisés dans les GCMs atmosphériques sont généralement : la turbulence de petite échelle, la convection peu profonde à l'origine des nuages de cumulus et stratocumulus, la convection profonde à l'origine des cumulonimbus et des orages, la condensation et la précipitation de l'eau liquide et glacée, et les ondes d'échelle inférieure à la taille des mailles. S'ajoute à ces paramétrisations, les paramétrisations permettant de lier les différents modèles entre eux (les échanges d'eau et de chaleur avec les surfaces océaniques et continentales, le freinage de l'air par la surface) ainsi que les paramétrisations du transfert radiatif. Ces dernières modélisent le transport du rayonnement solaire à travers l'atmosphère et le transport et l'émission du rayonnement infra-rouge par la surface terrestre et l'atmosphère. Les théories physiques fondamentales régissant la dynamique, la thermodynamique et le transfert radiatif dans l'atmosphère sont bien établies. En revanche, les paramétrisations des processus sous-maille et du transfert radiatif en présence de nuages font l'objet de recherches actives. En effet, un certain nombre de ces processus ne peuvent être facilement dérivés des théories physiques, conduisant à une diversité d'approches pour modéliser le même processus. Les panaches convectifs conduisant à la formation des cumulus et stratocumulus dans la couche limite atmosphérique peuvent par exemple être modélisés comme un panache convectif moyen dans une maille (par exemple Jam et al. (2013) ou Tiedtke (1989)) ou comme un ensemble de panaches décrits statistiquement (par exemple Arakawa and Schubert (1974)). La précipitation de l'eau liquide des nuages est paramétrisée de manière plus ou moins complexe : certains modèles la font dépendre simplement de la quantité d'eau condensée dans la maille (Madeleine et al., 2020) quand d'autres modèles paramétrisent les processus de coalescence ou de collection [*A priori il y aurait CESM2 mais je m'y connais pas assez pour savoir bien vérifier ça*]. Cette diversité d'approches possibles impacte significativement la simulation du climat et de la sensibilité cli-

matique (hausse des températures globales en réponse à une énergie ajoutée dans le système Terre) et constitue l'une des motivations majeures de la création des projets d'intercomparaison des modèles de climat. Ces projets déterminent un protocole de simulation strict, suivi par un ensemble de modèles de climat, afin d'isoler l'impact des choix de paramétrisations et des valeurs de paramètres associés sur la simulation du climat et du changement climatique. C'est une manière de quantifier l'incertitude des modèles et de leurs projections climatiques.

I.1.3 Comment comprend-on le climat avec un modèle ?

Les modèles de climat se sont complexifiés au fur et à mesure de leur développement. Leur complexité est le reflet de la complexité du climat terrestre et de cette caractéristique : des processus de très petites échelles peuvent avoir une influence sur le climat d'une région voire du globe. Cette imbrication des processus de différentes échelles spatiales et temporelles rend à la fois les modèles de climat globaux nécessaires - ce sont les seuls modèles pouvant imbriquer ces différentes échelles - mais les rendent aussi très complexes. Le climat ainsi simulé résulte de l'interaction entre tous les processus, conditionnée par de grands équilibres ou quasi-équilibres entre eux. Établir des liens de causalité dans ce cadre peut s'avérer délicat : c'est ce que certains.e.s épistémologues appellent le holisme des modèles (Lenhard and Winsberg, 2010). Une grande partie de la recherche en modélisation consiste alors à trouver des stratégies pour contourner cette difficulté. Par exemple, pour savoir si les précipitations du Brésil viennent de l'évaporation de la forêt amazonienne ou non, on pourra raser cette forêt dans le modèle et regarder l'impact sur les pluies brésiliennes (voire par exemple Walker et al. (1995)). Pour savoir si un processus a une influence sur le climat, on peut le paramétrer et comparer des simulations sur des échelles climatiques avec ou sans la paramétrisation de ce processus. S'il y a des changements significatifs entre les deux simulations, ils pourront être attribués au processus paramétrisé *[mettre un exemple ? Sur les stratocumulus ?]*. Le poids de certains processus sur le climat peut également être étudié en faisant des simulations avec différentes valeurs de poids sur les paramétrisations, contrôlées par leur paramètres. On pourra par exemple comparer les climats de simulations ayant une dépendance de la taille des cristaux de glace des nuages à la température plus ou moins forte. *[trouver un exemple avec citation !]* C'est ce que l'on appelle les études de sensibilité du climat aux paramètres libres du modèle.

Désintriquer l'effet de certains processus sur le climat peut également passer par la désactivation de certaines boucles de rétro-action. Le modèle de climat peut ainsi être utilisé sans le modèle d'océan, ou en imposant la dynamique atmosphérique. Des modèles idéalisés peuvent aussi être utilisés pour comprendre certains phénomènes. Le modèle peut par exemple être utilisé en imposant une température de surface ne dépendant que de la latitude, ou en mode aqua-planète, c'est-à-dire sans continent (voir par exemple Oueslati and Bellon (2013)). Il est aussi possible de réaliser des simulations en ne considérant qu'une seule colonne de modèle, dont les conditions aux bords sont forcées afin de reproduire des situations typiques ou idéalisées, telles que le développement de cumulus de beau temps en journée sur les continents par exemple (Lenderink et al., 2004; Rio and Hourdin, 2008). Les chercheuses s'appuient alors sur ce que l'on appelle une hiérarchie de modèle, allant de la colonne atmosphérique (ou océanique ou continentale) isolée au système climatique complet.

I.1.4 Comment savoir qu'un modèle de climat simule un climat réaliste ?

Une partie de la confiance accordée aux projections climatiques faites par les GCMs est basée sur leur capacité à bien simuler le climat terrestre présent et passé. Comment fait-on, concrètement, pour juger si la simulation du climat par le modèle est bonne ou mauvaise ? Les différents centres de modélisation du climat adoptent, en pratique, une diversité d'approches,

qui se retrouve par exemple dans les publications des versions de référence de leurs modèles pour les exercices CMIP (Hourdin et al. (2006, 2013, 2020) pour les versions de CMIP4, CMIP5 et CMIP6 du GCM atmosphérique LMDZ et Boucher et al. (2020); Dufresne et al. (2013) pour celles du modèle couplé de l'IPSL (Institut Pierre-Simon Laplace)). La question de l'évaluation des modèles est un domaine de recherche actif.

Les GCMs sont tout d'abord jugés sur leur capacité à simuler correctement les climats du passé, sensiblement différents du climat actuel, et qui sont connus grâce à des observations indirectes appelés proxys climatiques. Les proxys climatiques les plus connus sont les carottes de glace, qui permettent de connaître la composition de l'atmosphère, c'est-à-dire les gaz qu'elle contient et leurs isotopes, sur plusieurs centaines de milliers d'années. L'étude des climats du passé sur des échelles de temps géologiques, à partir des proxys climatiques et des simulations par des modèles de climats, est un domaine de recherche à part entière désigné sous le terme de paléoclimatologie.

Les GCMs sont également jugés sur leur capacité à simuler correctement le climat des dernières décennies, pour lequel on dispose d'un ensemble assez vaste d'observations. Tout d'abord, un ensemble de satellites observe la Terre en permanence. Ces satellites embarquent des instruments qui captent le rayonnement émis par l'ensemble du système Terre. Cela permet d'avoir une observation de ce rayonnement, de première importance pour le climat, mais aussi de remonter à d'autres variables atmosphériques comme l'humidité ou la couverture nuageuse. Ces dernières années, des instruments d'observation actifs envoyant des pulses de lasers (LIDAR) ont été embarqués par des satellites. Ces instruments permettent d'observer plus profondément dans les couches atmosphériques opaques au rayonnement et d'avoir accès aux variations verticales d'eau dans les nuages par exemple. Des simulateurs d'observables ont été développés pour transformer les sorties des GCMs et les rendre comparables aux signaux observés par les satellites (par exemple Bodas-Salcedo et al. (2011)).

L'atmosphère est aussi observée in situ. Des stations météorologiques sont installées dans tous les pays du monde et mesurent en continu un certain nombre de grandeurs (température, pression, humidité, vent etc.) nécessaires à la prévision du temps. Deux fois par jour, des capteurs sont embarqués dans des ballons qui montent jusqu'à une trentaine de kilomètres dans l'atmosphère afin de mesurer les variations de ces grandeurs sur la verticale. Des sites de mesures atmosphériques sont également développés à des fins de recherche pour l'étude du climat. Ces sites, moins nombreux, possèdent plus d'instruments, permettant d'observer par exemple la composition des nuages, le rayonnement arrivant à la surface, la chimie de l'atmosphère, etc. Ces sites s'organisent en réseau pour partager des protocoles d'observations communs et fournir ainsi des observations comparables entre les sites. Ils sont également utilisés pour la calibration des instruments embarqués dans les satellites (Driemel et al., 2018). L'atmosphère est aussi observée depuis la surface des océans par des bouées météorologiques fixes ou dérivantes. Enfin, des campagnes de mesures sont régulièrement organisées à des fins de recherche, avec des objectifs d'observation et de compréhension des processus physiques, chimiques ou biogéochimiques précis. La campagne StratoClim-2 a par exemple envoyé en 2021 une vingtaine de ballons pressurisés à dériver pendant plusieurs mois entre 18 et 20 km d'altitude, afin de comprendre les processus d'échange de vapeur d'eau entre la troposphère et la stratosphère (AERIS et al., 2019). La campagne de mesure AWACA (*Atmospheric Water Cycle over Antarctica*) déploie des stations de mesures atmosphériques en Antarctique, continent très peu observé, afin de mieux comprendre les processus atmosphériques spécifiques des régions polaires (IPSL et al., 2023).

Les différents types d'observations de la Terre sont utilisés pour produire ce que l'on appelle des climatologies du système Terre, c'est-à-dire des données statistiques en tout point du globe décrivant certains aspects du climat (température, précipitations, humidité de l'atmosphère et du sol, etc.). Ces climatologies regroupent la description du climat actuel et sont largement

utilisées pour l'évaluation des modèles de climat.

Des simulations de référence sont également utilisées pour évaluer les modèles de climat. Des LES (*Large Eddies Simulations* ou simulation des grands tourbillons) sont utilisées pour simuler à haute résolution des portions de l'atmosphère ou de l'océan. Une partie des processus sous-maille paramétrisés dans les GCMs est alors explicitement simulé par la LES, qui sert ainsi de référence pour la simulation de ces processus par les paramétrisations du modèle. Des calculs de transferts radiatifs sur les scènes nuageuses simulées par les LES ont récemment été développés en utilisant les méthodes de Monte Carlo (Villefranque et al., 2019). Ces calculs de références permettent de connaître avec fiabilité les flux radiatifs associés aux nuages simulés par la LES et ouvrent la voie à de nouvelles études sur la modélisation des interactions entre nuages et rayonnement. Des simulations avec des CRM (*Cloud Resolving Model* ou modèle resolvant les nuages) régionaux ou globaux, à résolution kilométrique, peuvent également être utilisées comme référence pour la modélisation des systèmes convectifs et leur organisation spatiale.

I.1.5 La question de l'ajustement des modèles de climat

I.1.5.1 Pourquoi a-t-on besoin d'ajuster les modèles de climat ?

Les paramétrisations des modèles font intervenir un grand nombre de paramètres (typiquement plusieurs centaines pour un modèle de climat complet), appelés paramètres libres. Ils sont généralement peu contraints par la théorie ou les observations. Ces paramètres peuvent être des facteurs de proportionnalité, des temps de vie, des paramètres géométriques (taille des gouttes par exemple) ou caractéristiques d'un élément (inertie thermique, albédo de la neige) etc. Ils dépendent de la formulation des paramétrisations physiques et ne sont donc généralement pas partagés entre les modèles de climat. La valeur de ces paramètres est fixe pendant toute la durée des simulations. Idéalement ces valeurs sont fixes sur tout le globe mais il arrive qu'elles varient régionalement, en fonction du type de surface par exemple. La question de l'ajustement, ou tuning, du modèle est la question de l'attribution d'une valeur à cet ensemble de paramètres libres. Cette question n'est pas spécifique aux modèles de climat. Elle est commune à la modélisation numérique en général, et est souvent désignée par le terme de calibration. Elle est vue comme une étape à part entière dans le développement d'un modèle numérique, et à fortiori, d'un modèle de climat.

L'ajustement des modèles de climat est fait de manière à ce que le modèle simule un climat réaliste. Autrement dit, les valeurs des paramètres libres sont choisies de manière à ce que le modèle simule un climat en accord avec les observations ou autres références disponibles, sur un certain nombre de caractéristiques mesurées à l'aide de métriques (les précipitations moyenne sur une région par exemple). Savoir définir ce qu'est un climat réaliste est une question de recherche à part entière, mais une certitude est partagée par tous les centres de modélisation : le modèle climatique doit simuler une température moyenne d'équilibre proche des observations lorsque l'énergie reçue du soleil est proche de sa valeur réelle (Hourdin et al., 2017). Le système Terre à l'équilibre doit également renvoyer autant d'énergie vers l'espace (sous forme de réflexion du rayonnement solaire et d'émission de rayonnement thermique) que l'énergie qu'il reçoit du soleil, à condition qu'il conserve numériquement l'énergie, ce qui n'est pas toujours le cas. Le système climatique n'est pas exactement équilibré du fait du changement climatique. Une partie de la chaleur est absorbée par les océans avec une constante de temps qui varie de la vingtaine à des milliers d'années. Cette absorption de chaleur par les océans provoque un déséquilibre énergétique du système autour de $0,5$ à 1 W.m^{-2} , bien en dessous des incertitudes d'observations (de quelques W.m^{-2} , Loeb et al. (2009)). La précision attendue sur le bilan énergétique des modèles de climat est ainsi de l'ordre du dixième de W.m^{-2} , ce qui correspond à une précision de l'ordre du dixième de degrés sur la simulation de la température globale (un changement d'un

W.m^{-2} sur le bilan énergétique provoque un changement de 0,5 à 1,5 °C sur la température moyenne globale). Une partie des processus influençant le bilan d'énergie du système Terre ne pourra jamais être simulée, ni même observée, avec suffisamment de précision pour que le modèle de climat soit “naturellement” à l'équilibre énergétique, c'est-à-dire sans intention délibérée qu'il le soit. C'est le cas des nuages, dont l'effet radiatif global net (-20 W.m^{-2}) n'est connu qu'à quelques W.m^{-2} près. Même les simulations avec les plus hautes résolutions (quelques mètres pour les LES) doivent paramétrer un certain nombre de processus de très fine échelle qui influencent la réflectivité et l'émissivité des nuages. Les tailles de gouttes de nuages, allant de quelques nanomètres à quelques millimètres, ne pourront par exemple jamais être simulées explicitement par les modèles de climat globaux, et leur influence sur la réflectivité des nuages est de plusieurs W.m^{-2} , bien au dessus de la précision recherchée [on a une ref pour ça ?]. Il n'est donc pas possible d'espérer un jour pouvoir contraindre les valeurs des paramètres libres du modèle uniquement par la théorie ou par l'observation. Les modèles de climat auront toujours besoin d'être ajustés afin qu'ils simulent, à minima, une température moyenne globale en accord avec les observations à quelques dixièmes de degrés près, et un bilan énergétique à quelques dixièmes de W.m^{-2} près.

I.1.5.2 Vers une science de l'ajustement des modèles

Cette question de l'ajustement des modèles a été identifiée dès les premiers travaux de modélisation du climat (Manabe and Wetherald, 1975), mais est longtemps restée au second plan. Les méthodologies utilisées pour ajuster les modèles de climat ont en effet peu été publiées et n'étaient pas demandées dans la documentation des modèles participants aux exercices d'intercomparaisons internationaux CMIP. Ce manque d'intérêt peut s'expliquer par le fait que la question de l'ajustement des modèles était vue par une partie des chercheuses comme une question purement technique, un “bidouillage” inévitable pour avoir un modèle “opérationnel”, voire même un moyen de masquer les défauts des modèles. Il peut peut-être aussi s'expliquer par la peur que cet ajustement ne renforce les arguments de ceux qui remettent en cause la validité des projections relatives au changement climatique. La question de l'ajustement des modèles a connu un net regain d'intérêt il y a une dizaine d'années. Un colloque international a été organisé en 2014 afin de partager les pratiques de tuning des différents centres de modélisation et d'explicitier les besoins et les questions scientifiques liés au tuning des modèles (WCRP, 2017). Une volonté forte de ce colloque était de faire de cette question de l'ajustement des modèles une science à part entière de la recherche en modélisation climatique, qui doit donc être étudiée, publiée et débattue dans la communauté scientifique (Hourdin et al., 2017).

La question des compensations d'erreurs est une question centrale dans le tuning des modèles. Lors du réglage du modèle, on souhaite au maximum que le modèle simule le bon climat pour les bonnes raisons. Notre confiance en les projections climatiques repose en effet sur ces deux aspects : la capacité du modèle à simuler correctement le climat présent, et le fait que le modèle soit basé sur la bonne combinaison d'équations physiques. Une mauvaise combinaison d'équations physiques peut conduire à une bonne simulation du climat actuel. Pire, cette mauvaise combinaison d'équations physiques pourrait être provoquée, intentionnellement ou non, pour que le modèle simule correctement le climat actuel. Par exemple, supposons que la formulation des équations conduisant à la simulation des nuages de cumulus contiennent des erreurs fondamentales (dites erreurs structurelles), et supposons que ces erreurs impliquent une sous-estimation moyenne de la couverture nuageuse des cumulus, responsable du renvoi d'une fraction non négligeable du rayonnement solaire vers l'espace. Le bilan énergétique du modèle pourra quand même être correctement simulé, en choisissant par exemple des valeurs de paramètres libres conduisant à une réflectivité de ces nuages irréaliste grande (grâce à une toute petite taille de goutte par exemple). La question de savoir à quel point ces compensations d'erreurs

sont désirables ou non, et où est-ce qu'elles sont acceptables ou pas, est aujourd'hui une question ouverte. La seule chose qui est sûre est que le modèle doit simuler un bilan énergétique global au dixième de W.m^{-2} près, au prix de compensation d'erreur si besoin.

On peut alors définir comme un enjeu de recherche le fait de s'assurer que les processus soient correctement simulés par le modèle tout en garantissant une bonne simulation du climat. Cela peut se faire en ajustant le modèle à différentes étapes du développement : pendant le développement de la paramétrisation (au plus proche des processus), puis au niveau de chaque composante du modèle indépendamment les unes des autres, et enfin au niveau du modèle climatique complet. La hiérarchie de modèles utilisée pour son développement pourrait ainsi également être utilisée pour l'ajustement du modèle. La question de l'arbitrage entre "simuler le bon climat" et "simuler les bons processus" pourra alors se traduire concrètement dans l'importance accordée aux métriques du modèle complet par rapport à celle accordée aux métriques orientées processus (sur des portions découplées du modèle) lors de l'ajustement du modèle.

La question des compensations d'erreur rejoint la question du surajustement (*overfitting* ou *overtuning*) des modèles numériques en général, qui arrive dans des contextes où l'ajustement est fait à un degré plus fin que les incertitudes de comparaison entre le modèle et les références utilisées, ou alors lorsqu'une partie trop partielle du modèle est ajustée. Ce surajustement est identifié comme une des causes principales de la lenteur du développement de nouvelles paramétrisations, et donc de l'amélioration des modèles de climat. Introduire une nouvelle paramétrisation dans le modèle conduisait en effet souvent à une moins bonne simulation du climat, alors que celui-ci a moins d'erreurs structurelles, car cela casserait les compensations d'erreurs introduites lors de l'ajustement du modèle. Il faut alors réajuster le modèle à chaque fois qu'une nouvelle paramétrisation est introduite. Cet ajustement a longtemps été mené sur la base d'essai-erreur, en faisant varier un à deux paramètres à la fois, ce qui le rendait fastidieux, chronophage et très peu efficace. Cela explique aussi pourquoi le tuning a longtemps été vu comme un frein au développement des modèles.

I.1.5.3 Les méthodes d'ajustement objectives

L'ajustement des modèles est un problème à grande dimension, tant au niveau du nombre de paramètres libres présents dans les modèles (de l'ordre de plusieurs centaines, souvent réduit à quelques dizaines dans le processus d'ajustement), qu'au niveau des caractéristiques climatiques que l'on souhaite que le modèle simule correctement (les métriques). Au vu de cette difficulté, les centres de modélisation ont adoptés, à partir des années 2010, des méthodes objectives d'ajustement de leur modèle, c'est-à-dire des méthodes basées sur un cadre mathématique ou statistique bien établi, développées par les communautés des statistiques, de l'ingénierie et de l'informatique. On distingue généralement deux types d'approches pour les méthodes d'ajustements. La première approche utilise un cadre d'optimisation, où l'ajustement consiste à minimiser une fonction des paramètres libres du modèle appelée fonction coût, qui mesure la distance entre les observations et la simulation. La deuxième approche, connue sous le nom de quantification des incertitudes, utilise des modèles statistiques (ou émulateurs) pour relier le modèle climatique aux observations, en quantifiant explicitement les principales sources d'incertitudes présentes dans la comparaison entre modèle et observations (incertitudes des observations, incertitudes dues aux conditions initiales et aux limites incertaines du modèle, incertitude structurelle du modèle due à sa formulation incomplète ou incorrecte). Ces approches sont basées sur des méthodes d'inférences bayésiennes, qui fournissent une incertitude autour des valeurs des paramètres libres. Les deux méthodes ont comme avantage de rendre la question du tuning objective, contrairement aux méthode d'essai-erreur, et ouvrent ainsi la voie à l'exploration des questions scientifiques liées à l'ajustement des modèles. Elles font toutes les deux face à la problématique très pragmatique des ressources en calcul, motivant ainsi à chercher des méthodes et des stratégies de

tuning accessibles en terme de coût de calcul.

La diversité des modèles de climat, inhérente à la complexité du système climatique et aux choix faits dans les paramétrisations des modèles, ainsi que la diversité des priorités de chaque centre de modélisation, implique une diversité dans l’ajustement des modèles. Cette diversité se retrouve dans les choix des paramètres libres qui vont être utilisés pour ajuster le modèle, ainsi que dans les métriques utilisées pour l’ajustement. Elle permet d’explorer au mieux les incertitudes des projections climatiques des modèles, évaluées lors des projets d’intercomparaison des modèles de climat. L’utilisation de méthodes d’ajustement objectives ne remplace en aucun cas la subjectivité de l’exercice d’ajustement, et sa diversité entre les différents centre de modélisation. Elle implique toutefois que cette expertise soit explicitée. L’arrivée de ces méthodes ouvre la voie à une meilleure connaissance de l’incertitude liée aux choix des valeurs des paramètres libres sur la simulation du climat et du changement climatique. Le tuning pourrait alors même devenir une aide et un guide au développement des modèles, en tant que méthode pour identifier les erreurs structurelles des modèles, indépendantes des valeurs des paramètres libres choisis.

I.1.5.4 Stratégie d’ajustement du modèle climatique de l’IPSL

Suite à la conférence internationale de 2014, la question de l’ajustement des modèles de climat a été (ré)ouverte dans les deux centres de modélisation climatique français : IPSL en région parisienne et Météo-France à Toulouse. Cette question s’est structurée dans le projet ANR HighTune “Amélioration et calibration des paramétrisations des nuages de couche limite à partir de simulations haute résolution”, réunissant, de 2016 à 2020, des statisticien·e·s, des expert·e·s du rayonnement, des modélisateur·ices du climat et de l’atmosphère et des spécialistes des processus atmosphériques. Ce projet a permis l’arrivée dans la communauté de modélisation du climat, d’une méthode objective d’ajustement semi-automatique des modèles issue de la communauté des quantifications d’incertitudes. Cette méthode, nommée “History Matching with Iterative Refocusing” (HMIR), vise à déterminer le sous-espace des paramètres libres en accord avec un ensemble de métriques à une tolérance à l’erreur prêt. Elle procède par échantillonnages itératifs de l’espace des paramètres libres, en utilisant des émulateurs gaussiens des métriques pour accélérer cette exploration. Elle est implémentée dans l’outil de tuning `htexplo` et est utilisée tout au long de cette thèse.

Une proposition forte de ce projet est de combiner des métriques orientées processus avec des métriques climatiques. Les métriques orientées processus portent sur des simulations atmosphériques uni-colonne forcées pour reproduire des situations météorologiques typiques, et qui ont comme référence des simulations tri-dimensionnelles haute-résolution (LES) ayant les mêmes forçages au bord de leur domaine. Les métriques climatiques portent sur le rayonnement et le bilan d’énergie au sommet de l’atmosphère du modèle atmosphérique non couplé au modèle océanique. L’objectif de cette proposition est double. Tout d’abord, il s’agit de garantir que la méthode d’ajustement des paramètres conduit à la bonne simulation du climat pour les bons processus. Ensuite, cette stratégie profite de l’aspect itératif de la méthode d’ajustement pour pré-contraindre l’ensemble des paramètres libres avec des simulations uni-colonne très peu coûteuses, et avoir ainsi un espace de paramètres libres moins grand à explorer pour contraindre l’espace des paramètres avec les métriques portant sur les simulations climatiques, coûteuses en temps de calcul. La preuve de concept de cette stratégie d’ajustement du modèle est faite sur l’ajustement des paramètres libres liés aux paramétrisations des couches limites faiblement convective du modèle atmosphérique de l’IPSL LMDZ. La comparaison entre les simulations uni-colonne et les sorties des simulations LES est implémentée de manière automatique dans l’outil de tuning `htexplo`, permettant un pré-conditionnement rapide de l’espace des paramètres libres.

Cette preuve de concept ouvre la voie à de nombreuses perspectives de recherche et d'enrichissement de cette stratégie de tuning en deux étapes, que cette thèse propose d'aborder.

I.2 Positionnement de la thèse

Le fait de considérer l'ajustement des modèles de climat comme une science à part entière dans les groupes de modélisation est très récent. Cette question a été clairement explicitée à l'échelle internationale lors du Workshop de 2014 sur le tuning, qui en pave les premières directions de recherche (Hourdin et al., 2017). Dans les deux centres de modélisation français, la méthode objective d'HMIR a été choisie pour l'ajustement de leurs modèles de climat globaux respectifs IPSL-CM et ARPEGE-Climat. L'implémentation de cette méthode dans l'outil `htexplo`, et les premières études qui en ont découlé, ont eu lieu juste avant le début de cette thèse (Couvreur et al., 2021; Hourdin et al., 2021; Villefranque et al., 2021). De nombreuses perspectives de recherche restent à explorer concernant l'ajustement des modèles de climat avec des méthodes objectives, et en particulier avec la méthode d'HMIR. Cette thèse a pour objectif d'explorer les potentialités de cette méthode pour l'ajustement des modèles de climat, et pour la modélisation du climat en général, en l'appliquant au GCM atmosphérique LMDZ. C'est sans doute une des premières thèses à avoir comme objet d'étude central l'ajustement d'un GCM, ce qui constitue une étape dans l'élaboration de cette nouvelle science.

La méthode d'HMIR a été choisie pour son potentiel à aborder de nouvelles questions scientifiques. Elle présente l'avantage de pouvoir combiner des métriques portant sur différentes configurations du même modèle, offrant une gamme de possibilités de métriques très large pour l'ajustement des paramètres du modèle.

Actuellement, l'ajustement de LMDZ se fait d'un côté au plus proche des processus avec des métriques portant sur des cas uni-unicolonne de convection peu profonde. Puis dans un second temps, des métriques sont ajoutées sur les flux radiatifs au sommet de l'atmosphère de simulations atmosphériques globales forcées, i. e. dans lesquelles les températures de surface de l'océan sont imposées. Le premier jeu de métriques a pour but de garantir la bonne simulation des processus, ici de couche limite convective, et le deuxième de garantir la bonne simulation du climat. Ces deux jeux de métriques pourraient être enrichis au sein de cette stratégie en deux parties. Côté simulation du climat global, on pourrait s'imaginer contraindre une plus grande diversité de caractéristiques du climat, telles que des caractéristiques du climat à la surface (température, précipitations, humidité, rayonnement etc.) qui sont très observées et dont l'enjeu sociétal est fort, ou de viser spécifiquement les caractéristiques de phénomènes grande échelle tels que les moussons ou ENSO (*El-Niño Southern Oscillation*). On pourrait également enrichir la stratégie avec des métriques visant à réduire les biais systématiques du modèle atmosphérique, comme par exemple la simulation d'une double ITCZ (ou ZCIT pour Zone de Convergence InterTropicale), la trop faible variabilité des précipitations, les précipitations trop intenses sur le continent maritime, etc. Côté processus, on pourrait ajouter des métriques sur une plus grande diversité de cas uni-colonne, et contraindre ainsi une plus grande diversité de processus atmosphériques ; tel que les cas de couches limites stables pour la turbulence de petite échelle ou les cas de convection profonde plus ou moins idéalisés.

Combiner des métriques proches des processus avec des métriques orientées climat global est une manière d'éviter les compensations d'erreurs, c'est-à-dire d'éviter que le climat soit bien simulé pour les mauvaises raisons (c'est-à-dire pour les mauvais processus). La méthode d'HMIR a un fort potentiel à la fois pour étudier les compensations d'erreurs - auxquelles une expérience de cette thèse est dédiée - mais aussi pour les éviter lors de l'ajustement du modèle. Pour ce dernier point, l'enjeu est à la fois de se doter de métriques au plus proche des processus, mais également d'assurer une continuité entre ces métriques et les métriques climatiques globales

dans la stratégie d’ajustement du modèle. Cela peut se faire en s’appuyant sur des configurations intermédiaires entre les configurations uni-colonnes et la configuration climatique globale du modèle, comme cela est déjà pratiqué pour le développement et l’évaluation de LMDZ.

Parmi la gamme des configurations intermédiaires possibles, nous proposons d’étudier le potentiel d’une configuration dite zoomée-guidée, où la dynamique atmosphérique est forcée par une technique de guidage et où la grille du modèle est affinée sur une région d’intérêt. La technique de guidage permet d’évaluer les paramétrisations physiques du modèle en imposant une succession de séquences météorologiques (synoptiques) proches des observations, et de séparer ainsi les erreurs causées par les paramétrisations physiques de celles causées par la dynamique atmosphérique du modèle. Cette séparation est moins forte que dans les utilisations uni-colonnes, un certain degré de liberté étant malgré tout laissé aux couplages physique-dynamique rapides, à l’intérieur de la région zoomée. Les simulations explorent en revanche un beaucoup plus grand nombre de situations météorologiques. L’utilisation de telles configurations pour l’ajustement du modèle pourrait ainsi permettre d’éviter, de contrôler et éventuellement de mieux comprendre les compensations d’erreurs qui peuvent émerger entre les parties dynamiques et physiques du modèle.

L’utilisation d’une grille zoomée avec guidage est fréquente pour simuler le climat d’une région d’intérêt, avec une grille plus fine sur cette région que la grille standard du modèle tout en limitant le coût de calcul des simulations. L’association de ces deux techniques permet d’imiter le fonctionnement d’un modèle régional à aire limitée avec le GCM atmosphérique LMDZ, en forçant la dynamique sur le bord du domaine du zoom grâce à un guidage fort. De telles configurations sont fréquemment utilisées pour la comparaison du modèle à des observations atmosphériques in-situ de campagnes de mesures ou de sites instrumentés. Les perspectives offertes par l’utilisation de ces données in-situ comme référence pour l’ajustement du modèle sont explorées dans cette thèse avec les observations du site instrumenté du SIRTa (Site Instrumental de Recherche par Télédétection Atmosphérique), situé en région parisienne, et une configuration de LMDZ guidée et zoomée sur ce site instrumenté. Nous abordons également dans ces chapitres la possibilité de cibler spécifiquement le couplage entre le modèle d’atmosphère LMDZ et celui de surfaces continentales ORCHIDEE, avec l’objectif d’ajuster conjointement les paramètres libres des deux modèles (et non plus séparément comme dans la stratégie de tuning actuel).

L’autre grand aspect que nous abordons dans cette thèse concerne l’ajustement des flux radiatifs. L’enjeu de continuité dans l’ajustement du modèle est particulièrement fort sur le transfert radiatif, qui est de première importance depuis l’échelle des processus jusqu’à l’échelle plus globale. Contraindre les flux radiatifs à l’échelle des processus sur des simulations uni-colonnes est aujourd’hui possible grâce aux calculs de références radiatives Monte Carlo développés par Villefranque et al. (2019), qui fournissent des références radiatives associées aux scènes nuageuses simulées par LES. De plus, l’amélioration de la simulation des nuages bas dans LMDZ grâce au développement d’une paramétrisation de la couche limite convective en flux de masse (Hourdin et al., 2002; Hourdin, 2005; Jam et al., 2013; Hourdin et al., 2019), et l’arrivée du nouveau modèle de transfert radiatif ecRad (Hogan and Bozzo, 2018) proposant notamment différentes hypothèses de prise en compte des nuages par le modèle de transfert radiatif, constituent aujourd’hui un contexte très favorable à l’étude des interactions entre nuages bas et rayonnement dans LMDZ. Nous proposons de les étudier dans la perspective de comprendre les compensations d’erreurs possibles entre les différentes paramétrisations des nuages bas et du transfert radiatif, qui peuvent émerger lors de l’ajustement des flux radiatifs du modèle global. Ces compensations d’erreurs sont notamment suspectées d’être à l’origine du biais de nuages bas trop peu couvrant et trop réfléchissant présent dans de nombreux modèles de climat globaux (Konsta et al., 2022; Nam et al., 2012). Ces expériences sont également l’occasion de mettre en pratique une première étude des compensations d’erreurs avec la méthode d’HMIR. Le travail

d'interfaçage du nouveau modèle de transfert radiatif ecRad avec LMDZ a eu lieu au cours de cette thèse, et cette étude est donc également motivée par le besoin de réajuster (tuner) le modèle après l'introduction de cette nouvelle paramétrisation.

Cette thèse a été l'occasion de pratiquer la méthode d'HMIR pour le tuning de LMDZ, et également d'observer l'appropriation et la pratique de cette méthode dans le centre de modélisation de l'IPSL. Cette thèse comporte donc une part de réflexivité sur les expériences mises en place spécifiquement à cette occasion mais également concernant des expériences de tuning auxquelles j'ai contribué, de façon plus ou moins significative, pendant cette période. Ces expériences couvrent un large spectre, allant du développement de nouvelles paramétrisations à la mise en place du modèle couplé pour le prochain exercice d'intercomparaison des modèles de climat, enrichissant une prise de recul sur notre pratique du tuning et de la modélisation du climat en général.

Ce manuscrit est organisé de la manière suivante. Le modèle de circulation générale LMDZ est présenté Chapitre II et la méthode d'HMIR Chapitre III. L'expérience de tuning visant les observations du site instrumenté du SIRTa est présentée sur deux chapitres distincts. Le premier, Chapitre IV, expose le travail de mise en place d'une configuration de référence de LMDZ, nécessaire à la réalisation de l'expérience de tuning visant les observations du SIRTa, présentée ensuite Chapitre V. Après une brève introduction du modèle de transfert radiatif ecRad, le Chapitre VI présente une étude des compensations d'erreurs entre les paramétrisations de la physique des nuages et le calcul du transfert radiatif, ainsi qu'une évaluation de ecRad en climat global. Une synthèse sur les pratiques de tuning pour la modélisation du climat est présentée Chapitre VII, suivie des conclusions et perspectives de ce travail de thèse Chapitre VIII.

Chapitre II

Le modèle de circulation général LMDZ

Remarque lecture : Relecture finale en cours - excepté partie transfert radiatif

LMDZ est un modèle dit de circulation générale (GCM pour l'acronyme anglais *General Circulation Model* utilisé dans ce document). C'est la composante atmosphérique du modèle climatique de l'IPSL. Il a été historiquement développé au LMD (Laboratoire de Météorologie Dynamique) pour la simulation des atmosphères terrestres et planétaires (Vénus, Mars, Titan etc) dans les années 1980 - 1990 (Forget et al., 2024; Hourdin, 2024). Il est toujours activement développé au LMD, et a gardé cette spécificité d'être conjointement développé pour simuler les atmosphères terrestre, planétaire et même exoplanétaires. Le centre de modélisation du climat de l'IPSL coordonne le développement du modèle climatique global pour la simulation du climat terrestre, dans lequel LMDZ est couplé avec le modèle de circulation général océanique NEMO (Madec, 2008; Rousset et al., 2015), avec le modèle de surface continental ORCHIDEE (Cheruy et al., 2020; Krinner et al., 2005) et avec les modèles de chimie atmosphérique INCA pour la troposphère (Hauglustaine et al., 2004) et REPROBUS pour la stratosphère (Jourdain et al., 2008). Chaque exercice d'intercomparaison des modèles de climat est l'occasion de fixer une version robuste, validée et publiée du modèle climatique global ainsi que de chacune de ses composantes. La version du modèle de référence utilisée dans cette thèse, et décrite dans ce chapitre, est celle du sixième exercice d'intercomparaison des modèles de climat (dit CMIP6 pour *Coupled Model Intercomparison Project*), appelé IPSL-CM6A pour le modèle climatique global (Boucher et al., 2020; Mignot et al., 2021) et LMDZ6A pour le modèle atmosphérique (Hourdin et al., 2020). En pratique, cette configuration a été fixée en 2017. Pour certaines expériences, des versions plus récentes de LMDZ sont utilisées, incluant des nouvelles paramétrisation et/ou une grille verticale plus fine.

Ce chapitre introduit les éléments de modélisation nécessaires à la compréhension des expériences exposées dans les parties suivantes de la thèse. Il expose Section II.1 les principe de modélisation du modèle de circulation générale LMDZ, puis quelques éléments de la configuration de LMDZ6A Section II.2 et détaille enfin Section II.3 ses principales paramétrisations : turbulence de petite échelle, modèle du thermique, convection profonde et transfert radiatif. Les paramétrisations sont présentées de manière à mettre en avant les paramètres libres utilisés par la suite dans les expériences d'ajustement du modèle.

II.1 Principes de modélisation des GCMs atmosphériques

Celles et ceux qui seraient intéressé.e.s par une introduction pédagogique et historiques peuvent se reporter à Guillemot and Hourdin (2021).

Les GCMs atmosphériques ont pour but de simuler l'évolution de l'état de l'atmosphère

sur l'ensemble du globe pendant une durée déterminée (une journée, une année, cent ans etc). Pour ce faire, ils intègrent numériquement, dans le temps et dans l'espace, un certain nombre d'équations régissant l'écoulement de l'air, ses échanges de chaleurs (sa thermodynamique) et de rayonnement (le transfert radiatif).

Les variables nécessaires à la description de l'évolution de l'état de l'atmosphère dans le modèle sont appelées variables d'état et sont généralement : les trois composantes du vent (U pour le vent zonal, V pour le vent méridien et W pour le vent vertical¹), la température de l'air notée T et l'humidité spécifique de l'air notée q_v (masse de vapeur d'eau par masse d'air sec), ainsi que les différentes phases de l'eau condensée. A ces variables peuvent s'ajouter un certain nombre de variables internes du modèle, destinées à des calculs intermédiaires ou à des diagnostics d'intérêt.

Intégrer numériquement les équations d'évolution de l'état de l'atmosphère signifie : (1) découper spatialement l'atmosphère en un ensemble de portions, appelées maille du modèle, (2) découper le temps en petite portions appelées pas de temps (une minute par exemple) (3) calculer à l'aide d'équations les variables d'états du modèle dans chaque maille de l'atmosphère et à chaque pas de temps, en commençant par le début, et ce pour toute la durée de la simulation. Pour une simulation d'un an avec un pas de temps de 15 min, le modèle calcule $4 \times 24 \times 365 = 35040$ états de l'atmosphère successivement en chaque point du maillage spatial.

La Figure II.1 illustre le découpage de l'atmosphère en maille, avec la température de chaque maille représentée en couleur et les vents horizontaux représentés par des flèches.

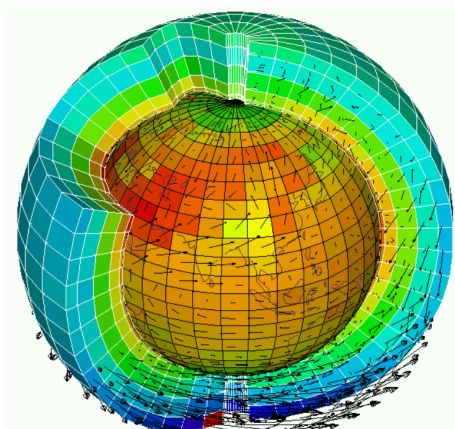


FIGURE II.1 – Illustration d'un état atmosphérique simulé par LMDZ avec une grille grossière (image réalisée par Laurent Fairhead et Frédéric Hourdin en 1995). On voit le découpage de l'atmosphère en mailles, avec une discrétisation en intervalles réguliers de longitude et latitude. L'épaisseur de l'atmosphère est très exagérée sur cette représentation. Les couleurs représentent la température dans chaque maille et les flèches les vents horizontaux

Modéliser l'atmosphère inclut toutes les étapes aboutissant à la simulation numérique de l'état de l'atmosphère. Alain Lahellec et Jean-Yves Granpeix, chercheurs au LMD, proposent de présenter de façon conceptuelle la modélisation comme une hiérarchie de 5 mondes ou niveaux (Guillemot and Hourdin, 2021) :

- le monde des phénomènes ou des processus. Il correspond à ce que l'on est capable d'observer des phénomènes atmosphériques tel que les tempêtes, la formation des nuages, les moussons etc.

1. Le vent vertical n'est pas à proprement parler une variable pronostique dans les modèles hydrostatiques, mais c'est alors la pression de surface qu'il faut ajouter comme variable d'état.

- le monde des théories, où ce que l'on a observé est transformé en explication cohérente (par exemple, la formation des cumulus est provoqué par l'ascendance de masse d'air chaudes et humides venant de la surface) et en principe physique (conservation de l'énergie, de la masse, etc).
- le monde des équations, où les théories sont traduites en équations mathématiques, comme par exemple les équations de Navier-Stokes pour la mécanique des fluides ou les équations du transfert radiatif pour le rayonnement
- le monde numérique (à nouveau dans la sphère des mathématiques), qui traite de la discrétisation des équation en temps et en espace
- le monde informatique, qui inclut l'écriture en code informatique des équations discrétisées dans un langage informatique (fortran pour LMDZ), l'adaptation des algorithmes aux types de machines (depuis les ordinateur portables jusque aux centres de calculs pour LMDZ) et puces électroniques utilisées pour faire les calculs (actuellement des CPU pour *Central Processing Unit* et prochainement des GPU pour *Graphics Processing Unit*), la gestion des flux d'entrées-sorties, les post-traitements, etc.

Toutes ces étapes mobilisées dans le développement du modèle font l'objet de recherche, l'objectif final étant d'expliquer ce qui se passe dans le monde des phénomènes.

II.1.1 Les équations primitives de la météorologie

Les équations primitives de la météorologie sont les équations décrivant les mouvements d'un fluide géophysique à la surface d'une planète en rotation, et sont la base des équations des GCMs océaniques et atmosphériques, tant pour la Terre que pour d'autres planètes. Ces équations sont dérivées des équations de Navier-Stokes, décrivant les mouvements des fluides par conservation du moment, auxquels sont associées les lois de conservations de la masse, de la chaleur et de n'importe quel traceur atmosphérique (espèces transportées par l'air comme la vapeur d'eau par exemple). Pour aboutir aux équations primitives à partir des équations de Navier-Stokes, un certain nombre d'hypothèses spécifiques aux caractéristiques de l'atmosphère sont appliquées et simplifient les équations de Navier-Stokes : 1) on applique tout d'abord l'approximation dite de couche mince, tenant parti du fait que l'épaisseur du fluide est très petite (99% de la masse de l'atmosphère est contenue dans les 100 premiers kilomètres) devant le rayon de la planète (6400 km pour la Terre) ; 2) l'air est de plus considéré comme un fluide parfait, dont la viscosité est négligée et dont l'écoulement est supposé adiabatique ; enfin, 3) l'approximation hydrostatique qui est utilisé dans LMDZ, ce qui n'est pas le cas de tous les GCMs. Cette approximation stipule que, sur la verticale, les forces de pressions sont à l'équilibre avec les forces de pesanteur. Cette approximation est vérifiée quand on moyenne les quantités concernées sur des dimensions horizontales supérieures à une dizaine voire vingtaine de kilomètres. L'aire des mailles du modèle ne doit donc pas être inférieure à environ $20 \times 20 \text{ km}^2$ afin de rester dans la gamme de validité de cette hypothèses.

Les équations primitives découlent fondamentalement d'équation de conservation pour :

$$\text{la quantité de mouvement : } \frac{D\vec{v}}{Dt} + \frac{1}{\rho}\vec{\text{grad}}(P) - \vec{g} + 2\vec{\Omega} \wedge \vec{v} = 0 \quad (\text{II.1})$$

$$\text{la masse de l'air : } \frac{D\rho}{Dt} + \rho\text{div}(\vec{v}) = 0 \quad (\text{II.2})$$

$$\text{la température potentielle : } \frac{D\theta}{Dt} = \frac{Q}{C_p} \left(\frac{P}{P_0} \right)^{\frac{C_p}{R_v}} \quad (\text{II.3})$$

$$\text{des concentrations d'espèces traces : } \frac{Dq}{Dt} = S_q \quad (\text{II.4})$$

Dans ces équations, $\frac{D}{Dt}$ désigne la dérivée temporelle lagrangienne, c'est-à-dire le long de la trajectoire de l'air. \vec{v} est le vecteur vent (de composante U, V, W), ρ la masse volumique de l'air, P la pression, g l'accélération de la pesanteur, $\vec{\Omega}$ le vecteur de rotation de la Terre (un tour par jour). La température potentielle θ est une variable conservée dans les mouvements adiabatiques en absence de changement de phase de l'eau. Elle se définit comme la température de l'air ramenée à une pression de référence $P_0 = 10^5 \text{ Pa}$, $\theta = T \left(\frac{P}{P_0} \right)^{\frac{C_p}{R_v}}$. Le terme Q rend compte des différentes sources de chaleur dues par exemple à la condensation de la vapeur d'eau, ou à l'absorption ou l'émission de rayonnement. $C_p = 1004 \text{ J.K}^{-1}.\text{kg}^{-1}$ est la chaleur spécifique de l'air et $R_v = 287 \text{ J.K}^{-1}.\text{kg}^{-1}$ la constante des gaz parfait pour l'air. Enfin q représente un traceur quelconque, S_q étant le terme source du traceur. Certaines de ces espèces traces peuvent être purement passives. D'autres, comme notamment les différentes phases de l'eau, rétroagissent sur la thermodynamique et le transfert radiatif.

L'approximation de couche mince intervient dans les termes géométriques de la dérivée lagrangienne ou encore dans le fait de négliger l'agrandissement des longueurs horizontales dans les calculs de gradient quand on s'éloigne de la surface. L'approximation hydrostatique revient à remplacer la composante verticale de la première équation (écrite sous forme vectoriel en toute généralité) par l'équation hydrostatique $\partial P / \partial z = -\rho g$.

II.1.2 Séparation d'échelles de l'écoulement atmosphérique

Les équations primitives sont discrétisées en temps et en espace afin de pouvoir être intégrées numériquement par le modèle. Les non linéarités des équations primitives atmosphériques impliquent que les phénomènes d'échelle inférieur à la discrétisation spatiale et temporelle influent sur l'écoulement de l'air à plus grande échelle et donc sur l'évolution des variables d'état à grande échelle. L'influence des petites échelles sur les plus grandes échelles est explicitement formalisé par la décomposition de Reynolds, ingénieur et physicien irlandais de la fin du XIX^e siècle.

II.1.2.1 Décomposition de Reynolds

Dans la décomposition de Reynolds, les grandeurs physiques décrivant l'écoulement d'un fluide sont décomposées en une partie moyenne et une fluctuation. Dans ce cadre, l'écoulement de l'air est considéré comme aléatoire. L'opérateur moyenne est alors défini comme la moyenne des réalisations de l'écoulement, appelée moyenne d'ensemble. Soit $\phi(\vec{x}, t)$ une grandeur physique fonction de la position \vec{x} et du temps t et $\phi_i(\vec{x}, t)$ la grandeur physique associée à la i^{e} réalisation de l'écoulement. La moyenne d'ensemble, dite aussi moyenne de Reynolds, notée $\bar{\phi}(\vec{x}, t)$ est telle que $\bar{\phi}(\vec{x}, t) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n \phi_i(\vec{x}, t)$. Pour un fluide compressible tel que l'air, cette moyenne doit être pondérée par la masse volumique, soit $\bar{\phi} = \overline{\rho\phi}/\bar{\rho}$. Les fluctuations de ϕ , notée ϕ' , sont

définies par rapport à cette moyenne, soit $\phi' = \phi - \bar{\phi}$. On note que la moyenne d'ensemble de la fluctuation ϕ' est nulle par définition.

La décomposition de Reynolds sert en mécanique des fluides à séparer différentes échelles spatio-temporelles caractéristiques de l'écoulement d'un fluide et est approprié aux situations où l'écoulement présente une claire séparation d'échelle. Cette échelle de séparation spatio-temporelle est liée à la moyenne de Reynolds via l'hypothèse d'ergodicité, qui stipule que la moyenne d'ensemble d'une grandeur ϕ est égale à sa moyenne spatiale sur une échelle caractéristique L et à sa moyenne temporelle sur une échelle caractéristique T correspondante. En se ramenant pour simplifier à un problème à une dimension spatiale, on a ainsi

$$\bar{\phi}(x, t) = \frac{1}{L} \int_{x-L/2}^{x+L/2} \phi(x, t) dx = \frac{1}{T} \int_{t-T/2}^{t+T/2} \phi(x, t) dt.$$

Prenons par exemple un ciel de cumulus, typique des après-midi de beau temps de nos latitudes. La moyenne de l'écoulement vu par un observateur placé en un point x de l'espace pendant une heure est considérée par cette hypothèse comme égale à la moyenne à un instant donné dans un rayon de 20 km autour de l'observateur. L'observateur verra en effet défiler un certain nombre de nuages entrecoupés de ciel clair au point x , correspondant en moyenne à une photo instantanée du ciel environnant le point x . La décomposition de Reynolds sert alors à séparer la partie grande échelle de l'écoulement, c'est-à-dire d'échelle supérieure à L et à T , de la partie petite échelle de l'écoulement, c'est-à-dire d'échelle inférieure à L et à T . La partie grande échelle de l'écoulement est alors décrite par les grandeurs physiques moyennées avec la moyenne de Reynolds, et les fluctuations décrivent la partie petite échelle de l'écoulement.

Avant d'appliquer la décomposition de Reynolds aux équations primitives atmosphériques, on écrit les équations primitives sous forme flux. La forme lagrangienne, utilisée dans les équations II.1, II.2, II.3, II.4, n'est pas adaptée pour les GCM qui possèdent un maillage spatial fixe. La dérivée lagrangienne est transformée en dérivée partielle temporelle à \vec{x} fixé avec :

$$\frac{D\phi}{Dt} = \frac{\partial\phi}{\partial t} + \vec{v} \cdot \overrightarrow{\text{grad}}(\phi) \quad (\text{II.5})$$

C'est la forme advective des équations, où $\vec{v} \cdot \overrightarrow{\text{grad}}(\phi)$ est l'advection de ϕ par \vec{v} . En multipliant les équations par la masse volumique de l'air et en utilisant l'équation de conservation de la masse sous sa forme advective on montre que :

$$\rho \frac{D\phi}{Dt} = \frac{\partial \rho \phi}{\partial t} + \text{div}(\rho \vec{v} \phi) \quad (\text{II.6})$$

avec $\rho \vec{v} \phi$ le flux de ϕ . C'est la forme flux des équations.

Prenons l'équation de conservation d'un traceur sous sa forme lagrangienne, Equation II.4. On peut écrire cette équation sous sa forme flux comme suit.

$$\frac{\partial \rho q}{\partial t} + \text{div}(\rho \vec{v} q) = \rho S_q \quad (\text{II.7})$$

On applique l'opérateur de moyenne statistique à l'Equation II.7. Cet opérateur commute avec les dérivées partielles de temps et d'espace. On peut alors écrire :

$$\frac{\partial \widetilde{\rho q}}{\partial t} + \text{div}(\widetilde{\rho \vec{v} q}) = \widetilde{\rho S_q} \quad (\text{II.8})$$

Par définition de l'espérance des processus stochastique q et S_q on a : $\widetilde{\rho q} = \rho \bar{q}$ et $\widetilde{\rho S_q} = \rho \bar{S_q}$, donc :

$$\frac{\partial \rho \bar{q}}{\partial t} + \text{div}(\rho \overrightarrow{\bar{v} q}) = \rho \bar{S_q} \quad (\text{II.9})$$

Et par décomposition de Reynolds des variables q et \vec{v} on a :

$\overline{\vec{v}q} = (\overline{\vec{v}} + \overline{\vec{v}'}) (\overline{q} + \overline{q}') = \overline{\vec{v}}\overline{q} + \overline{\vec{v}'q'}$ car $\overline{q'} = 0$ et $\overline{\vec{v}'} = 0$ par définition. L'Equation II.9 peut alors s'écrire :

$$\frac{\partial \overline{\rho q}}{\partial t} + \text{div}(\overline{\rho \vec{v}q}) = -\text{div}(\overline{\rho \vec{v}'q'}) + \overline{\rho S_q} \quad (\text{II.10})$$

$\overline{\rho \vec{v}'q'}$ est le flux turbulent de q . Il est non nul lorsque les fluctuations de \vec{v} sont corrélées aux fluctuations de q , ce qui est le cas pour les mouvements de turbulences de petites échelles ou de convection atmosphérique par exemple. On note que cette équation a la même forme que l'Equation II.7, avec un terme de divergence du flux turbulent en plus, que l'on peut associer à un terme source. Cette équation montre et explicite l'influence des fluctuations de petite échelle sur l'évolution des traceurs associés à l'écoulement de grande échelles \overline{q} .

La même méthode (opération de moyenne et décomposition de Reynolds) peut être appliquées aux autres équations de conservation (en passant par une transformation en forme flux). Les termes moyens ayant la même forme que l'équation initiale, les équations de conservation sont formellement inchangées à condition d'ajouter au second membre un terme source apparent de la forme

$$S_{\phi}^{\text{Reynolds}} = -\frac{\text{div}(\overline{\rho \vec{v}'\phi'})}{\rho} \quad (\text{II.11})$$

avec $\phi = u$ ou v pour l'Equation II.1, $\phi = \theta$ pour l'Equation II.3 et $\phi = q$ pour l'Equation II.4.²

II.1.2.2 Séparation entre cœur dynamique et paramétrisations physiques

Dans le modèle, la résolution explicite des équations a lieu jusqu'à l'échelle de leur discrétisation, notée T pour la discrétisation temporelle (de l'ordre du quart-d'heure), L pour la discrétisation horizontale (de l'ordre de la centaine de kilomètre) et Z pour la discrétisation verticale (de l'ordre de la dizaine de mètre au kilomètre).

Les échelles horizontales et verticales sont différenciées dans l'atmosphère à cause de l'influence de la gravité sur la partie verticale de l'écoulement qui provoque une stratification importante de l'écoulement sur la verticale. Les grandeurs physiques ont alors des variations beaucoup plus importantes sur la vertical que sur l'horizontal, ce qui nécessite une résolution beaucoup plus fine pour correctement représenter explicitement leurs variations verticales, typiquement de quelques dizaines de mètres dans les premier kilomètre à 1 km dans la moyenne troposphère.

Dans les GCMs, la décomposition de Reynolds ne s'applique ainsi que sur l'horizontale. Elle est utilisée pour séparer l'échelle explicitement résolue par le modèle, de l'échelle non résolue appelée aussi échelle sous-maille. Cette séparation d'échelle conduit à séparer le modèle en deux parties : le cœur dynamique et les paramétrisations physiques.

Le cœur dynamique a pour mission de résoudre la partie grande échelle de l'écoulement, c'est-à-dire celle d'échelle supérieure à l'échelle de la maille horizontale. Pour ce faire, il va intégrer numériquement les équations de conservation (Equation II.1 à II.4) sans termes sources (c'est à dire avec $Q \equiv 0$ et $S_{\text{trac}} \equiv 0$).

Les paramétrisations physiques ont quant à elle l'objectif d'estimer les termes sources des équations, à savoir le taux de chauffage Q , les termes sources des traceurs S_q et les termes de fluctuations sous-maille (divergence de flux turbulents).

2. Dans les équations finales, les notations de moyennes $\overline{\phi}$ ou $\tilde{\phi}$ sont traditionnellement systématiquement omises, sauf pour le terme croisé des fluctuations turbulentes $\overline{\vec{v}'\phi'}$.

Les paramétrisations sont en fait des modèles de ces termes sources, qui vont exprimer ces derniers fonction des variables grandes échelles du modèle via l'utilisation d'un certain nombre de variables internes (comme la fraction nuageuse ou les précipitations par exemple) et de paramètres libres (facteur de proportionnalité, propriété géométriques ou caractéristique etc).

Elles reposent sur l'hypothèse que les processus impliqués dans ces termes sources sont statistiquement homogènes sur l'horizontal, c'est-à-dire que les distributions statistiques de ces processus et leur moments sont invariants sur l'horizontal à l'intérieur des mailles du modèle. Cette hypothèse permet de décrire les grandeurs physiques à l'intérieur d'une maille comme des distributions statistiques ou à l'aide de leur moment. Elle implique que les dérivées partielles des flux turbulents sur l'horizontal sont nulles. Les termes de divergence de flux turbulent sont ainsi réduit à une dérivée partielle selon la verticale. La contribution des termes de Reynolds comme terme source au membre de droite des équation de conservation se réduit alors à :

$$S_{\phi}^{\text{Reynolds}} = -\frac{1}{\rho} \frac{\partial \overline{\rho w' \phi'}}{\partial z} \quad (\text{II.12})$$

Il faut ajouter à la représentation des mouvements sous mailles, le terme de chauffage radiatif Q_R ainsi que les termes de condensation (c) et sublimation (e), induisant un puits $c - e$ sur l'eau vapeur, et un terme de chauffage associé. Ce terme de chauffage vaut $L_v(c - e)$ si toute l'eau condensée est liquide. En présence de glace, ce terme devient plus compliqué. Cependant, si on suppose que la fraction f_g d'eau solide est inchangée pendant le processus d'évaporation ou de condensation, on peut l'écrire $(L_v + f_g L_f)(c - e)$, où L_f est la chaleur latente de fusion de l'eau. Ces termes sources sont souvent écrit dans la littérature de la convection en notant le terme de chauffage non radiatif Q_1 , en introduisant un terme équivalent Q_2 pour la source de vapeur d'eau, traduite en chauffage, et enfin parfois un terme équivalent Q_3 pour la quantité de mouvement :

$$\frac{D\vec{v}_h}{Dt} + \frac{1}{\rho} \text{grad}_h(P) - \vec{g} + 2f \vec{k} \wedge \vec{v}_h = -\frac{1}{\rho} \frac{\partial \overline{\rho w' \vec{v}_h'}}{\partial z} = Q_3 \quad (\text{II.13})$$

$$C_p \frac{D\theta}{Dt} = Q_R + (L_v + f_g L_f)(c - e) - C_p \frac{1}{\rho} \frac{\partial \overline{\rho w' \theta'}}{\partial z} = Q_R + Q_1 \quad (\text{II.14})$$

$$\frac{Dq_v}{Dt} = e - c - \frac{1}{\rho} \frac{\partial \overline{\rho w' q'_v}}{\partial z} = -Q_2/L_v \quad (\text{II.15})$$

$$\frac{Dq_c}{Dt} = c - e - \frac{1}{\rho} \frac{\partial \overline{\rho w' q'_c}}{\partial z} \quad (\text{II.16})$$

Remarque : dans cette écriture Q_2 est un puits d'humidité traduit conventionnellement en chaleur avec une constante L_v qui elle ne tient pas compte de la glace.

On suppose également que les processus impliqués dans les termes sources paramétrisés atteignent leur état stationnaire, en équilibre avec l'écoulement de grande échelle, en un temps très inférieure à celui du pas de temps du modèle, de manière à ce que leur régime transitoire puisse être négligé : c'est l'hypothèse de quasi-stationnarité. Pour le transfert radiatif par exemple, cette hypothèse est bien vérifiée dans la mesure où le temps de propagation du rayonnement dans l'atmosphère est négligeable devant le pas de temps du modèle, qui est de l'ordre du quart d'heure.

Finalement, la séparation entre cœur dynamique et paramétrisations physique repose sur une représentation de l'écoulement atmosphérique où :

- l'écoulement de grande échelle, résolu par le cœur dynamique, s'organise essentiellement de manière horizontale, dans des couches fines, et avec des mouvements verticaux lents (de l'ordre de 1 cm.s^{-1}) équilibrant les divergences horizontales de l'écoulement grande échelle

- l'écoulement turbulent (ayant des mouvements verticaux beaucoup plus rapides que l'écoulement vertical de grande échelle) et le chauffage, par rayonnement et par condensation ou évaporation de l'eau, interagissent localement avec cet écoulement grande échelle par des transferts verticaux, gérés par les paramétrisations physiques du modèle.

Le choix de la discrétisation numérique spatio-temporelle du GCM, certes guidée par des considérations de coût de calcul, vise plus fondamentalement à se placer dans une zone où les échelles caractéristiques de l'écoulement atmosphériques sont aussi distinctes que possible. Cela permet de séparer au mieux les processus qui doivent être paramétrisés de ceux qui sont explicitement résolus par le cœur dynamique du modèle. L'échelle de coupure observée dans l'écoulement atmosphérique entre quelques dizaines et quelques centaines de kilomètres permet ce compromis, et est utilisée par tous les GCMs atmosphériques. Les processus sous-mailles paramétrisés sont alors la turbulence de fine échelle (1-100 m) et les processus convectifs (0,1-10 km), et les processus explicitement résolus par le cœur dynamiques du modèle sont ceux des systèmes météorologiques d'échelle synoptiques (500-5000 km), tel que les anti-cyclones et dépressions.

Certains phénomènes, tel que les fronts, les cyclones ou les systèmes convectifs organisés, ont des échelles caractéristiques proches de l'échelle de coupure et sont ainsi difficilement simulés par les GCMs à moins de monter en résolution jusqu'à quelques dizaines de kilomètres.

L'hypothèse d'homogénéité statistique horizontale des processus sous-maille, nécessaire en théorie à leur paramétrisation, n'est de plus pas toujours vérifiée. Elle dépend de la taille caractéristique du processus paramétrisé (plus elle est proche de l'échelle de coupure, plus il sera difficile qu'il soit statistiquement homogène) mais aussi de la situation météorologique.

Cette séparation d'échelle des GCMs, et le cadre sous-jacent de développement des paramétrisations, est un cadre idéalisé d'une puissance souvent mésestimée, qui a sans doute fondé le succès de ces modèles depuis les années 1990. Tout d'abord, cette séparation permet d'avoir une discrétisation spatio-temporelle beaucoup plus lâche que les processus simulés par le modèle, assurant ainsi l'accessibilité en terme de coût de calcul des simulations. Cela a notamment permis aux GCMs d'assurer la simulation des projections climatiques du prochain siècle, nécessitant en pratique de l'ordre du millier d'années de simulation afin notamment d'équilibrer l'océan profond du modèle couplé océan-atmosphère. De plus, la représentation des processus sous-mailles sous forme de paramétrisations implique une compréhension de ces processus et de leur interaction avec l'écoulement grande échelle résolu, et permet de faire le lien entre processus sous-maille et simulation des propriétés émergentes du climat. Elle est à ce titre un formidable carde de compréhension des interactions d'échelle au sein de l'atmosphère [Citation du papier Rio et Hourdin sur les interactions d'échelle] et pour la compréhension des phénomènes météorologiques et climatiques.

II.2 Quelques éléments de configuration de LMDZ6A

La version de LMDZ utilisée pour CMIP6, LMDZ6A, est présentée dans Hourdin et al. (2020).

La grille horizontale du modèle est constituée de points également répartis en longitude et en latitude. La surface des mailles est ainsi plus importante dans les tropiques qu'aux pôles. La résolution standard utilisée est constituée de 144 points en longitude et 143 en latitude, soit des intervalles de 2,5 ° de longitude et 1,25 ° de latitude, et des mailles de 200x200 km² à l'équateur et 10x10 km² aux pôles.

La grille verticale est constituée de 79 points, avec un dernier niveau vers 80 km dans la haute mésosphère. Les mailles sont plus fines proche de la surface, avec une croissance presque linéaire de l'épaisseur des mailles dans les trois premiers kilomètres, d'épaisseur égale à $\Delta z \simeq 0,11z$. C'est en effet dans cette zone que les variations verticales sont les plus fortes, donc que le maillage

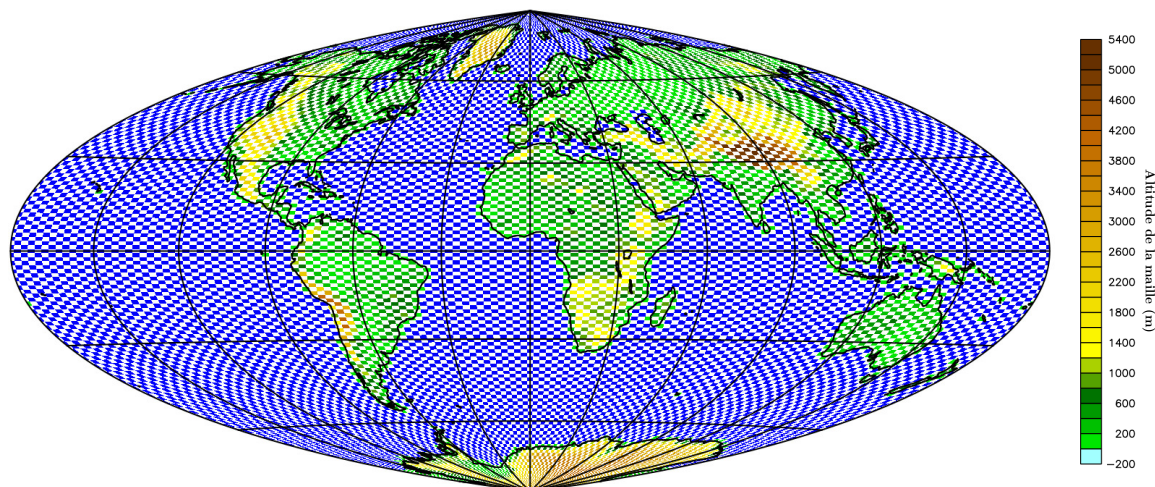


FIGURE II.2 – Grille horizontale du modèle en projection Eckert-Greifendorff qui conserve l'aire des mailles. Une maille sur est deux coloriée, soit en bleu foncé pour les mailles océaniques, soit en une couleur correspondant à l'altitude moyenne de la maille. On voit l'aire des mailles diminuer à mesure que l'on s'éloigne de l'équateur.

verticale nécessite d'être particulièrement fin. Une grille verticale constituée de 95 niveaux est également utilisée dans une partie des expériences de cette thèse. Cette grille est introduite par Hourdin et al. (2019) pour l'étude des stratocumulus, puis est légèrement modifiée par Hourdin et al. (2021). Elle est basée sur la grille à 79 niveaux avec un raffinement de la descriptions des cinq premiers kilomètres de l'atmosphère. Dans les trois premiers kilomètres, l'épaisseur des mailles a une croissance presque linéaire égale à $\Delta z \simeq 0,067z$. Les profils d'épaisseur des niveaux verticaux des grilles à 79 et 95 niveaux sont tracés Figure II.3.

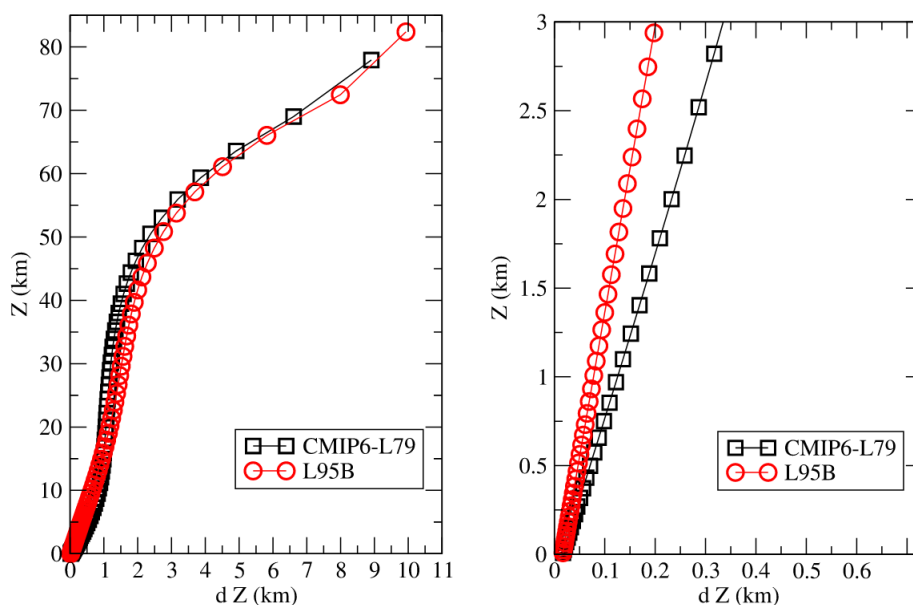


FIGURE II.3 – Epaisseur des niveaux verticaux dZ (km) en fonction de l'altitude du milieu du niveau Z (km) pour les grilles à 79 niveaux en noir et 95 niveaux en rouge, issue de Hourdin et al. (2021). A gauche les profils sont tracés pour toute la colonne atmosphérique et à droite dans les trois premiers kilomètres

Deux pas de temps sont utilisés pour la discrétisation temporelle des équations. Le cœur

dynamique du modèle utilise en schéma numérique de discrétisation temporelle de saute-mouton explicite (*leapfrog* en anglais) avec un pas de temps δt . Dans ce schéma, la variable d'état au temps $t + \delta t$ est calculée suivant l'équation : $\phi(t + \delta t) = \phi(t - \delta t) + 2\delta t M_{dyn}[\phi(t)]$. Le terme $M_{dyn}[\phi(t)]$ correspond à l'évaluation des équations du cœur dynamique (équations II.13, II.14, II.15 et II.16 sans terme source), mises sous la forme $\frac{\partial \phi}{\partial t} = M_{dyn}[\phi]$ puis discrétisées spatialement. Ce terme est aussi appelé terme de tendance de la dynamique.

Tous les 5 ou 7 pas de temps, un schéma numérique de Matsuno est appliqué afin de réconcilier les solutions correspondant aux pas de temps pairs et impairs. C'est à ce moment que les paramétrisations physiques sont appelées, avec un schéma explicite d'ordre 1. La variable d'état au temps $t + \delta t$ est alors calculée avec l'équation suivante :

$$\phi(t + \delta t) = \phi(t) + M_{dyn}[\phi(t)]\delta t + M_{phy}[\phi(t)]\Delta t \quad (\text{II.17})$$

avec Δt le pas de temps de la physique et $M_{phy}[\phi]$ l'évaluation des termes sources par les paramétrisations physiques. Ce terme est aussi appelé terme de tendance de la physique. Dans CMIP6, les paramétrisations de la physique sont évaluées toutes les 15 minutes et tous les 7 pas de temps de la dynamique, soit $\Delta t = 7\delta t = 15$ min.

En pratique, les paramétrisations physiques les plus coûteuses en temps de calcul ne sont pas appelées à tous les pas de temps de la physique afin de gagner en temps de calcul. La paramétrisation de la convection profonde est ainsi appelée tous les deux pas de temps de la physique (soit toutes les 30 minutes) et le transfert radiatif (très coûteux numériquement) tous les six pas de temps de la physique (soit toutes les 1h30). Avec la grille verticale à 79 niveaux, une simulation climatique d'une année prend environ 700 heures, soit un mois sur un processeur CPU. Des techniques de parallélisations des calculs, tirant notamment partie de l'indépendance des colonnes de la partie physique du GCM, permettent d'abaisser le temps de retour des résultats à une vingtaine d'années simulées par jour.

II.3 Les paramétrisations physiques de LMDZ

Les paramétrisations physiques de LMDZ sont appelées de manière séquentielle. Chaque paramétrisation calcule la tendance physique de son processus, et la somme des tendances de chaque paramétrisation constitue le terme de tendance physique $M_{phy}[\phi(t)]$ de l'Equation II.17.

Parce que les paramétrisations de la convection sont conçues pour travailler sur l'eau totale (c'est-à-dire sur la somme des rapports de mélange de l'eau sous forme de gaz, de liquide et de glace), la totalité de l'eau liquide et de la glace est évaporée au début de chaque appel à la physique (Madeleine et al., 2020). Un découpage est alors fait entre les paramétrisations qui s'occupent du transport vertical de l'eau totale (turbulence de petite échelle et convection peu profonde) et la paramétrisation qui s'occupe de la formation des nuages, c'est-à-dire du calcul des rapports de mélange de l'eau liquide et de la glace à partir du rapport de mélange de l'eau totale. Une exception est faite pour la paramétrisation de la convection profonde, qui s'occupe à la fois du transport vertical de l'eau et de sa condensation, car ces deux processus sont fortement couplés dans ce cas précis. Ce découpage est basé sur l'hypothèse que le temps de vie des condensats est négligeable devant le pas de temps de la physique (15 min), ce qui est généralement le cas pour les nuages d'eau liquide mais plus discutable pour les nuages de glace ou de phase mixte. Les nuages sont alors des variables semi-prognostiques dans LMDZ, car bien qu'ils soient réévaporés et reformés à chaque appel à la physique, les rapports de mélange d'eau liquide et de glace qui les constituent sont advectés par le cœur dynamique.

L'ordre d'appel des paramétrisations est le suivant :

- Evaporation de l'eau liquide et de la glace

- Turbulence de petite échelle, couplage et appel aux schémas de surface (ORCHIDEE pour les surfaces continentales et NEMO pour les surfaces océaniques)
- Convection profonde, dont condensation, nuages et précipitation
- Convection peu profonde
- Condensation et précipitation des nuages grande échelle et de convection peu profonde
- Transfert radiatif
- Processus orographiques avec des effets de freinage ou de déviation et la génération d'onde de gravité (orographiques et non orographique) déferlant dans la haute atmosphère
- Traceurs

Cette section ne suit pas l'ordre d'appel des paramétrisations dans LMDZ. Les paramétrisations de la turbulence petite échelle (Section II.3.1, de la convection peu profonde (Section II.3.2) et de la condensation grande échelle (Section II.3.3) seront présentées dans un premier temps. Puis la paramétrisation de la convection profonde sera introduite Section II.3.4, suivie de la paramétrisation du transfert radiatif Section II.3.5. Les paramétrisations des processus orographiques et des traceurs ne sont pas présentés. Une attention particulière a été portée à introduire les paramètres libres utilisés dans les expériences d'ajustement de cette thèse. Les principales équations dans lesquelles ils interviennent sont présentées dans cette section, et un tableau récapitulatif est disponible Annexe A. Ils sont notés en gras dans cette section.

II.3.1 Turbulence de petite échelle

Dans la couche limite, l'écoulement de l'air est très majoritairement turbulent, c'est-à-dire qu'il est constitué de pleins de tourbillons, allant de l'échelle caractéristique de la couche limite (de l'ordre du kilomètre ou de la demi-heure) à celle de la dissipation (de l'ordre du centimètre ou de la seconde).

La turbulence qui nous intéresse ici est une turbulence de petite échelle, désorganisée, dont les tourbillons vont de la centaine de mètre à l'échelle de la dissipation, avec une période comprise entre 10 secondes et 15 minutes. Cette turbulence est visible dans l'atmosphère lorsque l'on regarde la fumée d'une bougie par exemple, ou l'aspect bourgeonnant des panaches de fumées ou de certains nuages (typiquement les cumulus et les cumulonimbus). C'est aussi celle qu'on ressent au décollage d'un avion. Cette turbulence provoque un mélange local d'air donc un mélange des caractéristiques qu'il transporte (température, humidité, quantité de mouvement, etc).

Le transport moyen d'une variable ϕ par la turbulence de petite échelle est modélisé par une équation de diffusion, où le flux (de Reynolds) de la variable d'état ϕ transporté par la turbulence est proportionnel au contre-gradient local de cette variable d'état. Cette formulation est analogue à celle de la diffusion moléculaire due au mouvement brownien des molécules. Seule la composante verticale du transport par ce flux turbulent est paramétrisée selon

$$\overline{\rho w' \phi'} = -\rho K_z \frac{\partial \phi}{\partial z} \quad (\text{II.18})$$

où K_z est la diffusivité turbulente.

La fermeture de cette paramétrisation consiste à relier K_z aux variables d'état du modèle. Il existe une diversité d'approche en la matière et celle implémentée dans LMDZ6A est une fermeture d'ordre 1.5, basée sur le travail de Yamada (1983) et améliorée par Vignon et al. (2017). Cette fermeture consiste à exprimer le coefficient d'échange turbulent

$$K_z = l_m S(RI) \sqrt{TKE} \quad (\text{II.19})$$

comme le produit de trois termes :

- la longueur de mélange l_m définie comme la taille effective des tourbillons, supposés isotropes dans une couche limite neutre.
- Une fonction de stabilité $S(RI)$ qui modélise la déformation, ou l'anisotropie, des tourbillons, en fonction de la stabilité de la couche limite quantifiée par le nombre de Richardson RI . Les tourbillons sont plus étirés vers le haut dans des couches limites instables, et ont donc des tailles effectives plus grandes ($S(RI) > 1$ quand $RI < 0$) et inversement pour des couches limites stable, où leur taille effective est plus petite ($S(RI) < 1$ quand $RI > 0$).
- L'énergie cinétique turbulente, notée TKE pour *Turbulent Kinetic Energy* en anglais, définie par $TKE = \frac{1}{2}(\overline{u'^2} + \overline{v'^2} + \overline{w'^2})$ et calculée à partir de son équation prognostique

Les formulation de la longueur de mélange, des fonctions de stabilités et de l'équation prognostique de l'énergie cinétique turbulente utilisées dans LMDZ6A sont détaillées dans Vignon (2017). La formulation du coefficient d'échange pour le moment K_m (c'est-à-dire pour les composantes horizontales du vent u et v) est légèrement différente de celles pour la chaleur et l'humidité K_h .

Ce sont aussi ces petits tourbillons qui vont transporter l'eau et la chaleur de la surface vers l'atmosphère. La surface peut être un continent, un océan ou de la banquise, et elle est toujours considérée comme une surface solide. On fait l'hypothèse que la couche d'atmosphère la plus proche de la surface, épaisse de quelques millimètres, est en équilibre avec la surface, et a ainsi une température égale à celle de la surface T_s et est saturée en humidité ($q_{sat}(T_s)$). Les tourbillons vont alors mélanger cette petite couche d'atmosphère avec les premiers centimètres et mètres de l'atmosphère, c'est-à-dire avec la première couche verticale du modèle, d'une vingtaine de mètres d'épaisseur. Les flux de chaleurs et d'humidité induit par ce mélange tourbillonnaire sont respectivement appelés flux turbulents de chaleur sensible et flux d'évaporation (ou flux turbulent de chaleur latente) à la surface.

En intégrant l'Equation II.18 entre la couche millimétrique et le milieu de la première couche du modèle (situé à environ 10 m), on montre que les flux de chaleur sensible (respectivement de chaleur latente) sont proportionnels à la différence de température (respectivement d'humidité) entre la couche millimétrique et le premier niveau du modèle. Les coefficients de proportionnalité sont formellement établis pour la couche limite atmosphérique par la théorie des similitudes de Monin-Obukhov (Monin and Obukhov, 1954).

Cette théorie s'applique lorsque les flux turbulents sont constants, ce qui est le cas dans les premiers mètres de l'atmosphère (les flux turbulents varient typiquement de l'ordre de 10% dans le premier 10% de la couche limite, c'est-à-dire dans les 10 à 300 premiers mètres). La couche atmosphérique où les flux turbulents sont constants s'appelle la couche limite de surface. La théorie des similitudes de Monin-Obukhov donne alors une formulation des gradients verticaux du vent horizontal U (Equation II.20) et de température potentielle virtuelle θ_v (Equation II.21) dans la couche limite de surface, écrit ici dans le cas d'une couche limite stable par soucis de simplification :

$$\frac{\partial U}{\partial z} = \frac{u^*}{\kappa z} \Leftrightarrow U = \frac{u^*}{\kappa} \ln \left(\frac{z}{z_{0m}} \right) \quad (\text{II.20}) \quad \frac{\partial \theta_v}{\partial z} = \frac{\theta_v^*}{\kappa z} \Leftrightarrow \theta_v - \theta_{v,s} = \frac{\theta_v^*}{\kappa} \ln \left(\frac{z}{z_{0h}} \right) \quad (\text{II.21})$$

Avec $\kappa = 0,4$ la constante de Von Kármán, u^* la vitesse de friction, θ_v^* l'échelle turbulente de température, z_{0m} la hauteur à laquelle le vent horizontal s'annule, appelé rugosité de la surface et égale environ au dixième de la taille des obstacles présents à la surface et z_{0h} la hauteur où la température virtuelle est égale à la température virtuelle de la surface. z_{0m} et z_{0h} sont toujours situés en dessous du premier niveau du modèle.

Le flux turbulent de chaleur sensible à la surface s'écrit $\overline{\rho w' \theta_v'^s} = -\rho C_p u^* \theta_v^*$, avec C_p la capacité thermique massique de l'air, qui vaut $1004 \text{ J.kg}^{-1}.\text{K}^{-1}$. En utilisant l'Equation II.20 pour déterminer u^* et l'Equation II.21 pour déterminer θ_v^* , on montre que le flux de chaleur sensible *SENS* s'écrit selon l'Equation II.22.

$$SENS = \overline{\rho w' \theta_v'^s} = -\rho C_p |\vec{V}| \underbrace{\frac{\kappa^2}{\ln(z/z_{0m})\ln(z/z_{0h})} f_{sh}(RI)}_{C_h} (\theta_v - T_s) \quad (\text{II.22})$$

Avec $|\vec{V}|$ le module du vent horizontal dans la première couche du modèle, θ_v la température potentielle virtuelle dans la première couche du modèle, et C_h un coefficient d'échange qui dépend d'une fonction de stabilité $f_{sh}(RI)$ de l'atmosphère. De la même manière, on montre que le flux turbulent d'humidité sur une surface d'eau libre (ou évaporation potentielle) E_{pot} s'écrit selon l'Equation II.23.

$$E_{pot} = \overline{\rho w' q'^s} = -\rho |\vec{V}| \underbrace{\frac{\kappa^2}{\ln(z/z_{0m})\ln(z/z_{0h})} f_{sh}(RI)}_{C_h} (q_{air} - q_{sat}(T_s)) \quad (\text{II.23})$$

Avec q_{air} le rapport de mélange de vapeur d'eau au premier niveau du modèle. Sur les continents, l'évaporation $E = \beta E_{pot}$, avec β le facteur d'aridité qui prend en compte le fait que la surface n'a pas toujours assez d'eau pour apporter le flux d'humidité E_{pot} à l'atmosphère. En multipliant le flux d'évaporation par la chaleur latente de vaporisation $L_v = 2\,250 \text{ kJ.kg}^{-1}$, on obtient le flux de chaleur latent : $LAT = L_v E$.

Pour des soucis de stabilité numérique, et compte tenu de la longueur du pas de temps de la physique, un schéma Euler implicite est nécessaire pour la discrétisation temporelle des équations de la turbulence de petite échelle. Cela implique un couplage numérique avec les modèles de surface, et explique que l'appel aux modèles de surfaces continentales et d'océan se fasse dans la paramétrisation de la turbulence petite échelle.

A noter qu'une revisite de la paramétrisation de la turbulence de petite échelle a été menée au cours d'un ateliers de travail collectif auquel j'ai participé pendant ma thèse (Vignon et al., 2024). Le développement de cette nouvelle paramétrisation, actuellement en cours de test dans des configurations climatiques de LMDZ, est arrivée trop tardivement pour être intégrés aux simulations présentées dans cette thèse.

II.3.2 Paramétrisation de la convection peu profonde dans LMDZ

Dans la couche limite, des mouvements d'air organisés en rouleaux ou en ascendances peuvent apparaître pendant la journée. C'est ce que l'on appelle la convection peu profonde. Ces ascendances d'air organisées, appelées panaches thermiques ou plus simplement thermiques, se forment dans la couche limite de surface et se dissipent en haut de la couche limite dans la couche d'inversion (définie par un profil de température potentielle qui croît rapidement avec l'altitude). Elles sont schématisées Figure II.4 (a). Les thermiques font typiquement une à quelques centaines de mètres de diamètre pour 1 à 2 km de haut, et sont espacées de un à deux kilomètres. Ces ascendances organisées transportent la chaleur et l'humidité accumulées dans la couche de surface en journée en haut de la couche limite. Elles mélangent efficacement la couche située au dessus de la couche limite de surface, appelée couche limite mélangée, et sont responsables des profil neutre de température potentielle et d'humidité dans cette partie de l'atmosphère (Figure II.4 (c)). Si les conditions sont propices, l'apport d'humidité par les thermiques dans l'environnement froid de la couche d'inversion conduit à la formation de nuage de cumulus ou de stratocumulus.

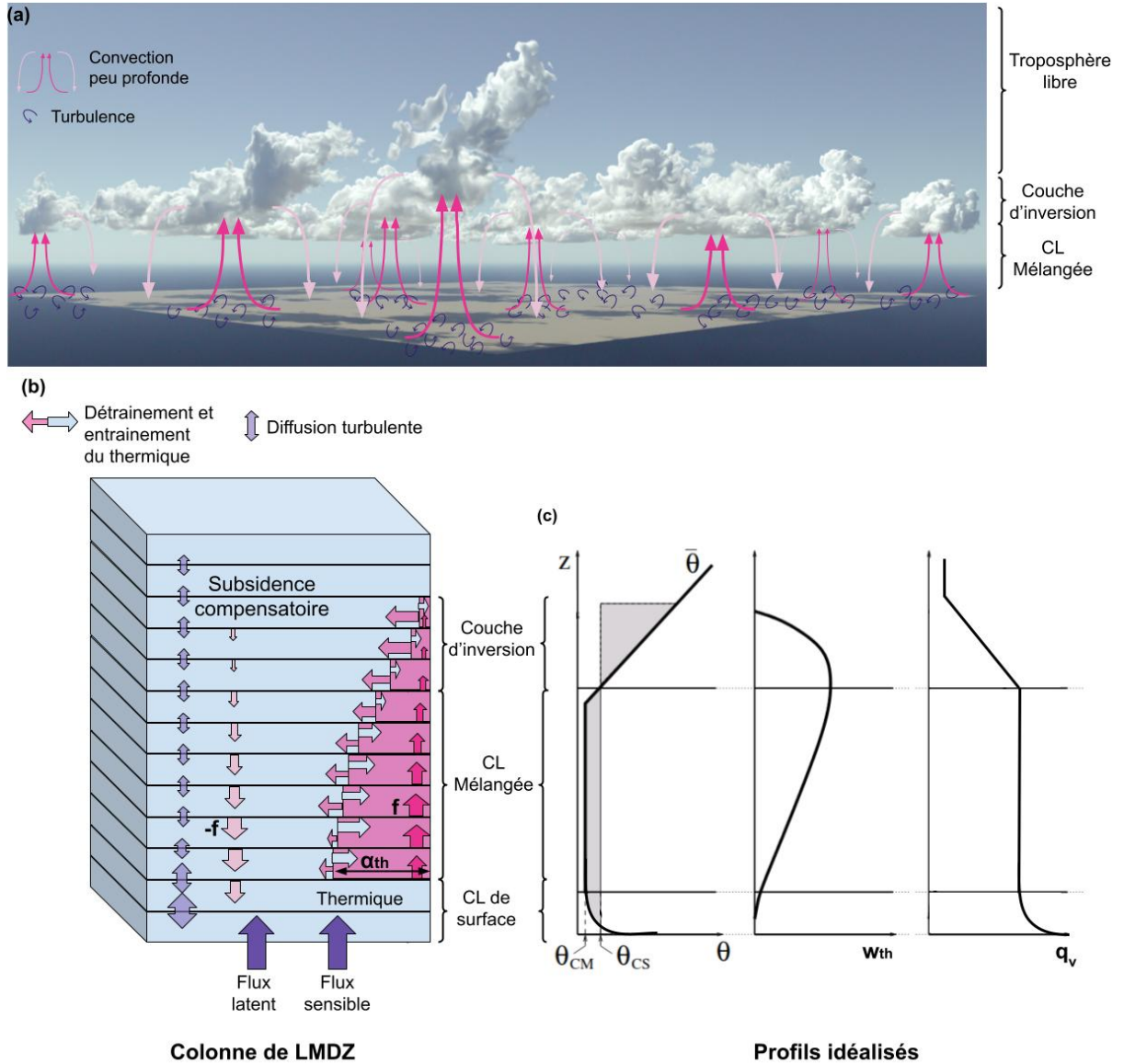


FIGURE II.4 – (a) Image photo-réaliste réalisé avec htrdr (Villefranque et al., 2019) d'une scène de cumulus du cas ARMCU simulée par la LES Meso-NH. Les cellules de convection peu profonde sont schématisées en rose, et la turbulence de petite échelle en violet. Les nuages de cumulus se situent dans la couche d'inversion au-dessus des ascendances d'airs. (b) Schéma de la représentation de la couche limite convective (turbulence petite échelle et thermique effectif) dans une colonne de LMDZ. (c) Profils idéalisés d'une couche limite atmosphérique convective (adapté de Hourdin (2005))

Les thermiques sont paramétrisés dans LMDZ par un modèle en flux de masse. Ils sont introduit dans LMDZ par Hourdin et al. (2002) et plus largement décrit dans Hourdin (2005). Leur paramétrisation est revisitée par Hourdin et al. (2019) pour améliorer la simulation des stratocumulus associés aux couches limites convectives, et un tuning spécifique de cette paramétrisation est proposé par Hourdin et al. (2021). L'ensemble des thermiques est modélisé par un thermique moyen ou effectif, stationnaire, transportant un flux de masse $f = \rho \alpha_{th} w_{th}$ vers le haut, avec α_{th} la fraction de la maille occupée par le thermique (typiquement 10%) et w_{th} sa vitesse verticale. Le thermique est entouré d'un atmosphère légèrement subsident pour compenser l'ascendance d'air, de flux de masse $-f$. Il échange des masse d'air latéralement avec son environnement, sous forme d'entrainement e (flux d'air de l'environnement qui entre dans le thermique), qui nourrit

le thermique, et de détrainement d (flux d'air du thermique qui sort dans l'environnement) qui le dissipe. Ce sont ces échanges qui vont conditionner sa formation (avec un entrainement maximale en bas du thermique) et sa dissipation (avec un détrainement maximale dans la couche d'inversion). La paramétrisation des thermiques dans LMDZ ne s'occupe que du mélange des variables d'état. La formation de nuage et éventuellement de précipitations liées à la présence du thermique sont gérées par la paramétrisation dite de condensation de grande échelle, décrite Section II.3.3. La variable d'état pour l'humidité dans la paramétrisation du thermique est donc le rapport de mélange de l'eau total q_t .

Le transport d'une variable d'état ϕ par le thermique est régi par l'équation

$$\frac{\partial f \phi_{th}}{\partial z} = e \phi - d \phi_{th} + \rho S_{\phi th} \quad (\text{II.24})$$

avec ϕ_{th} la variable d'état dans le thermique et $S_{\phi th}$ une source éventuelle de ϕ dans le thermique. La conservation de la masse d'air (l'équation précédente pour $\phi \equiv 1$ et $S_{th} \phi \equiv 0$) s'écrit

$$\frac{\partial f}{\partial z} = e - d \quad (\text{II.25})$$

Le transport vertical total de la variable d'état ϕ s'écrit dans une couche limite convective comme la somme du transport convectif et du mélange par diffusion turbulente :

$$\overline{\rho' w' \phi'} = f(\phi_{th} - \phi) - K_z \frac{\partial \phi}{\partial z} \quad (\text{II.26})$$

La couche limite convective est ainsi paramétrisée par un schéma mixte de diffusion turbulente et de flux de masse (dit *Eddy-Diffusivity-Mass-Flux scheme* en anglais), représenté Figure II.4 (b).

Le vitesse verticale du thermique w_{th} est déduite par l'Equation II.24 pour le terme source $S_{wth} = \mathbf{A1}B - \mathbf{A2}w_{th}^2$. Les paramètres libres $\mathbf{A1}$ et $\mathbf{A2}$ contrôlent respectivement la contribution de la flottabilité B dans l'accélération du panache et son freinage.

Les taux d'entrainement $\epsilon = e/f$ et de détrainement $\delta = d/f$ sont paramétrisés comme suit :

$$\epsilon = \max \left[0, \frac{\mathbf{B1}}{1 + \mathbf{B1}} \left(\frac{\mathbf{A1}.B}{w_{th}^2} - \mathbf{A2} \right) \right] \quad (\text{II.27})$$

$$\delta = \max \left[0, \frac{\mathbf{A1}.B1}{1 + \mathbf{B1}} \frac{B}{w_{th}^2} + \mathbf{CQ} \left(\frac{\Delta q_t / q_t}{w_{th}^2 / w_0^2} \right)^{1/4} \right] \quad (\text{II.28})$$

avec $w_0 = 1 \text{ m.s}^{-1}$.

Ils dépendent de la flottabilité du thermique B et de sa vitesse verticale w_{th} . Le détrainement dépend en plus du contraste d'humidité entre le thermique et son environnement, Δq_t , dont la contribution est modulée par le paramètre libre \mathbf{CQ} .

Enfin, la formulation de la flottabilité du thermique est modifiée dans la formule du détrainement afin d'améliorer la simulation des stratocumulus par le schéma, ceux-ci étant trop facilement dissipés à cause de l'overshoot³ (Hourdin et al., 2019).

$$B' = g \cdot \frac{\theta_{v,th}(z) - \theta_v(z + z.\mathbf{DZTH})}{\theta_v(z + z.\mathbf{DZTH})} \quad (\text{II.29})$$

Avec θ_v la température potentielle virtuelle de l'environnement et $\theta_{v,th}$ celle du thermique. Avec cette formulation, le détrainement est "averti" de l'inversion avant de l'atteindre, et commence à détrainier à la distance $z \times \mathbf{DZTH}$ sous l'inversion. \mathbf{DZTH} est le paramètre libre de la distance à laquelle le détrainement connaît la température au dessus de lui.

3. Vraiment pas de traduction disponible en français pour ce mot qui signifie qu'un panache convectif passe au dessus de l'inversion et pénètre la couche stable au dessus de la couche limite convective, dans une région très stable où il est rapidement freiné.

II.3.3 Schéma de condensation grande échelle

Le rôle du schéma de condensation grande échelle est de condenser et de précipiter l'eau en excès dans l'atmosphère. Cet excès est typiquement soit transporté depuis la surface par la paramétrisation des thermiques, soit advecté des mailles voisines ou des couches plus basses par la dynamique. Le schéma de condensation grande échelle de LMDZ6A est décrit dans Hourdin et al. (2021) et Madeleine et al. (2020).

Le schéma part du haut de la colonne atmosphérique et parcourt les niveaux verticaux jusqu'à la surface. Les étapes successives du calcul sont décrites pour un niveau quelconque d'altitude z_k .

Etape 1 : Réévaporation

Le flux de précipitation P venant du niveau z_{k+1} est évaporé selon l'équation Equation II.30, basé sur les travaux de Klemp and Wilhelmson (1978) et de Schlesinger et al. (1988).

$$\frac{\partial P}{\partial z} = -\mathbf{EVAP} \left(1 - \frac{q_t}{q_{sat}} \right) \sqrt{P} \quad (\text{II.30})$$

Avec q_{sat} l'humidité à saturation et \mathbf{EVAP} le paramètre libre contrôlant l'efficacité de cette évaporation. On note que le flux de précipitation issu d'une couche k peut traverser plusieurs couches du modèle.

Etape 2 : Formation des nuages

Pour éviter un schéma de nuage en tout ou rien, un schéma statistique de nuages est utilisé. Il est basé sur une paramétrisation de la distribution sous-maille $Q(s)$ de l'écart à la saturation, $s = q_t - q_{sat}$. On utilise comme variable s plutôt que de l'eau elle même, pour tenir compte en partie des corrélations entre les champs d'eau et de température. La fraction nuageuse est calculée comme la fraction de la maille où l'eau se condense (lorsque s est positif) :

$$\alpha = \int_0^\infty Q(s) ds \quad (\text{II.31})$$

et l'eau condensée q_c , liquide et solide, est donnée par

$$q_c = \int_0^\infty s Q(s) ds \quad (\text{II.32})$$

La fonction de densité de probabilité $Q(s)$ est paramétrisée par deux formes, suivant que la paramétrisation des thermiques a été activée dans la maille au cours du pas de temps ou pas.

Si les thermiques ont été activées, la densité de probabilité est modélisée par une bi-gaussienne, avec un mode représentant la variabilité de s dans le thermique et l'autre dans l'environnement du thermique (Jam et al., 2013). Les moyennes des deux gaussiennes sont les sursaturations dans le thermique $\overline{s_{th}}$ et dans l'environnement $\overline{s_{env}}$, calculées à partir des variables conservées, eau et températures potentielles, moyennes du thermique. Les écarts types des gaussiennes, sont paramétrisés par les équations

$$\sigma_{s,th} = \mathbf{BG2}(\alpha_{th} + 0.001)^{-\gamma_1} (\overline{s_{th}} - \overline{s_{env}}) + b\overline{q_{th}} \quad (\text{II.33})$$

$$\sigma_{s,env} = \mathbf{BG1} \frac{\alpha_{th}^{-\gamma_2}}{1 - \alpha_{th}} (\overline{s_{th}} - \overline{s_{env}}) + b\overline{q_{env}} \quad (\text{II.34})$$

où $\gamma_1 = 0.4$, $\gamma_2 = 0.6$, $b = 2.10^{-3}$, $\mathbf{BG1}$ et où $\mathbf{BG2}$ sont les paramètres libres, utilisés pour l'ajustement dans certaines expériences de cette thèse.

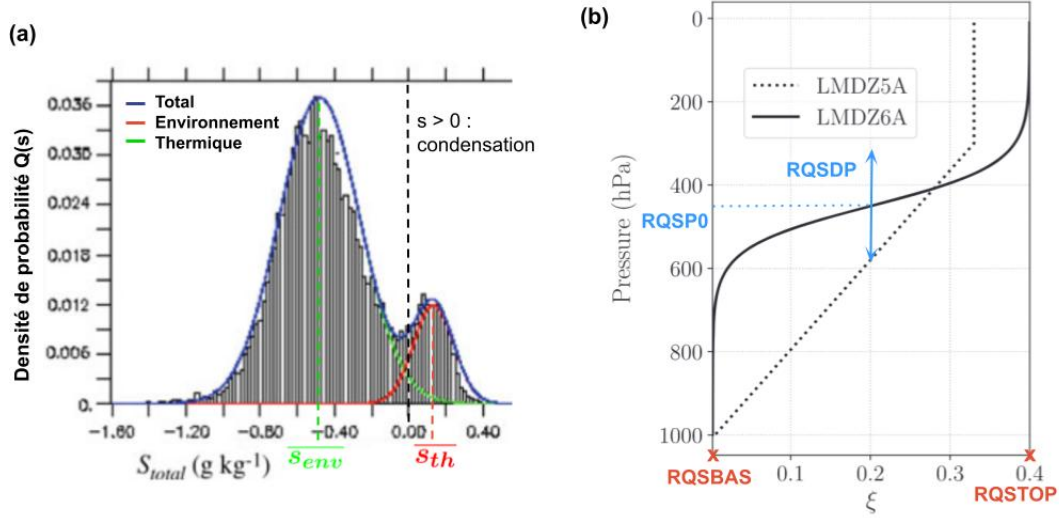


FIGURE II.5 – (a) Distribution du déficit de saturation $Q(s)$ à 1200 d'altitude, pour la 8^e heure d'un cas de cycle diurne de cumulus continentaux au site ARM dans les grandes plaines états-uniennes (le cas ARMCU). L'histogramme est calculé sur les données d'une simulation LES de ce cas. La courbe verte est la gaussienne de l'environnement ; la courbe rouge celle du thermique et la bleue, distribution retenue pour la paramétrisation statistique des nuages dans LMDZ, la somme des deux (Jam et al., 2013) (b) Profil de la fonction $\xi(P)$ utilisée pour le calcul de l'écart type de la distribution de $Q(s)$ quand les thermiques ne sont pas actifs (adapté de Madeleine et al. (2020))

Si les thermiques ne sont pas actifs dans la maille, la densité de probabilité $Q(s)$ est une lognormale d'écart type $\sigma = \xi q_t$. Avec ξ une fonction tangente hyperbolique de la pression (Figure II.5), réglée par plusieurs paramètres : **RQSBAS** la valeur de ξ à la surface, **RQSTOP** la valeur de ξ au sommet de l'atmosphère, **RQSP0** la hauteur de la transition de la tangente hyperbolique et **RQSDP** la largeur de la transition.

La phase de l'eau nuageuse est ensuite calculée, en fonction de la température. Le nuage est de phase mixte entre $T_{min} = -30$ °C et $T_{max} = 0$ °C, avec une fraction d'eau liquide $x_{liq} = \sqrt{(T - T_{min}) / (T_{max} - T_{min})}$.

Etape 3 : auto conversion et précipitation

Une partie de l'eau nuageuse est convertie en précipitation et est donc perdue par le nuage. La perte d'eau liée aux précipitations pour les nuages d'eau liquide suit l'équation

$$\frac{dq_l}{dt} = -\frac{q_l}{\mathbf{CLTAU}} \left[1 - e^{-(q_l^{in}/\mathbf{CLC})^2} \right] \quad (\text{II.35})$$

issue de Sundqvist (1978), où q_l est le rapport de mélange d'eau liquide dans la maille et q_l^{in} l'eau liquide ramenée à la fraction nuageuse c'est-à-dire que $q_l^{in} = q_l/\alpha$ en kg d'eau liquide par kg d'air nuageux. La conversion de l'eau nuageuse en précipitations commence à être efficace lorsqu'elle atteint le seuil **CLC**, et cette conversion se fait avec le temps caractéristique **CLTAU**.

La perte de glace liée aux précipitations suit l'équation

$$\frac{dq_i}{dt} = -\frac{1}{\rho} \frac{\partial \rho q_i w_i}{\partial z} \quad (\text{II.36})$$

où q_i est le rapport de mélange de glace dans la maille et w_i la vitesse de chute des cristaux de glace. Cette vitesse est proportionnelle à la vitesse de chute terminale des cristaux de glace $w_{i,\infty}$ suivant l'Equation II.37 (Heymsfield, 1977; Heymsfield and Donner, 1990), avec **FALLV** un paramètre libre.

$$w_i = \mathbf{FALLV} w_{i,\infty} \text{ avec } w_{i,\infty} = 3,29(\rho q_i)^{0,16} \quad (\text{II.37})$$

II.3.3.1 Paramétrisation du recouvrement vertical des précipitations

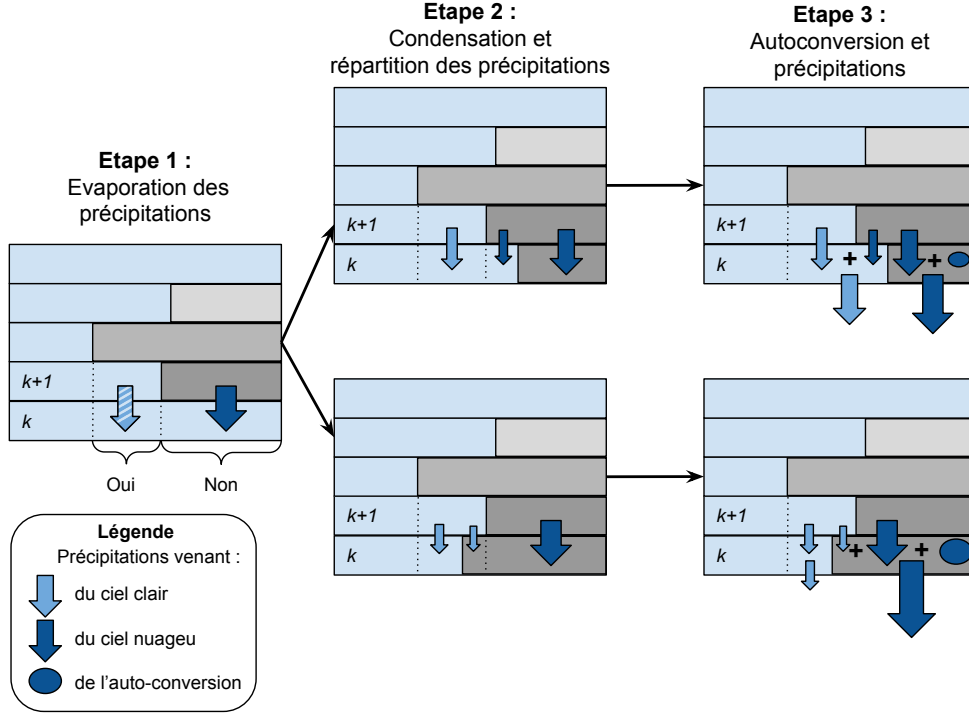


FIGURE II.6 – Schéma décrivant la paramétrisation de condensation grande échelle prenant en compte le recouvrement vertical des précipitations, sur quelques couches nuageuses d’une colonne du modèle. Il n’est pas montré ici (1) le cas où le nuage est recouvert par un nuage plus haut (disjoint), où la couche du sommet du nuage pourra recevoir un flux de précipitation ciel-clair venant de ce dernier et sur une fraction calculée par le modèle de recouvrement vertical des nuages (2) le cas où le flux de précipitation ciel-clair est trop faible par rapport à la fraction de maille qu’il occupe (d’un facteur **RI**) et commence à décroître. Sur le schéma, les couches de nuage sont disposées de manière à se recouvrir au maximum, c’est-à-dire à minimiser leur couverture nuageuse.

Une version plus récente du schéma de condensation grande échelle est utilisée pour certaines expériences de cette thèse. Cette paramétrisation modélise le recouvrement vertical des précipitations avec les nuages (Jakob and Klein, 2000; Touzé-Peiffer, 2021), et est schématisés Figure II.6. Elle prend en compte le fait que les précipitations du niveau $k + 1$ peuvent arriver dans la partie nuageuse ou dans la partie ciel-clair de la maille k . Les précipitations arrivant dans la partie nuageuse de la maille ne s’évaporent pas car elles arrivent dans un air supposé déjà saturé, contrairement à celles arrivant dans le ciel clair. Après condensation, les flux de précipitations venant du ciel clair et nuageux sont réparties dans la nouvelle partition ciel-clair/ciel-nuageux de la maille k , désormais connue. Puis l’étape d’autoconversion est appliquée au ciel nuageux, et ajoute une nouvelle source dans le flux de précipitation nuageux communiqué à la maille du dessous. Un seuil **RI** est introduit sur l’intensité des précipitations ciel-clair. A partir de ce seuil, la fraction de la maille dans laquelle ont lieu les précipitations ciel-clair décroît linéairement avec le flux de précipitation ciel-clair, afin d’éviter que cette fraction potentiellement grande passe à zéro d’une couche à l’autre. Ce seuil **RI** est un paramètre libre.

Notons que cette paramétrisation nécessite d'introduire un modèle de recouvrement vertical des nuages, ce qui n'était pas le cas jusqu'à présent, à part pour le transfert radiatif. Le modèle choisi est le modèle maximum-aléatoire, dans lequel les couches nuageuses adjacentes se recouvrent de manière à minimiser leur couverture nuageuse totale, alors que les couches de nuages séparées par du ciel-clair se recouvrent de manière aléatoire (décrit plus en détail Section VI.1.1.1).

II.3.4 Paramétrisation de la convection profonde

La convection profonde est le processus à l'origine de la formation des cumulonimbus, les nuages d'orages. Cette convection est caractérisée par des mouvements d'air ascendants qui, une fois enclenchés au sommet de la couche limite convective, sont renforcés par la chaleur libérée par la condensation de l'eau. Les courants ascendants au cœur des cumulonimbus atteignent souvent plusieurs dizaines de m.s^{-1} . La convection profonde atteint souvent la tropopause, soit une vingtaine de kilomètre à l'équateur. Lorsque le nuage atteint l'altitude de flottabilité neutre, il s'étale et forme une enclume. Cette convection est toujours accompagnée de nuage et de précipitations.

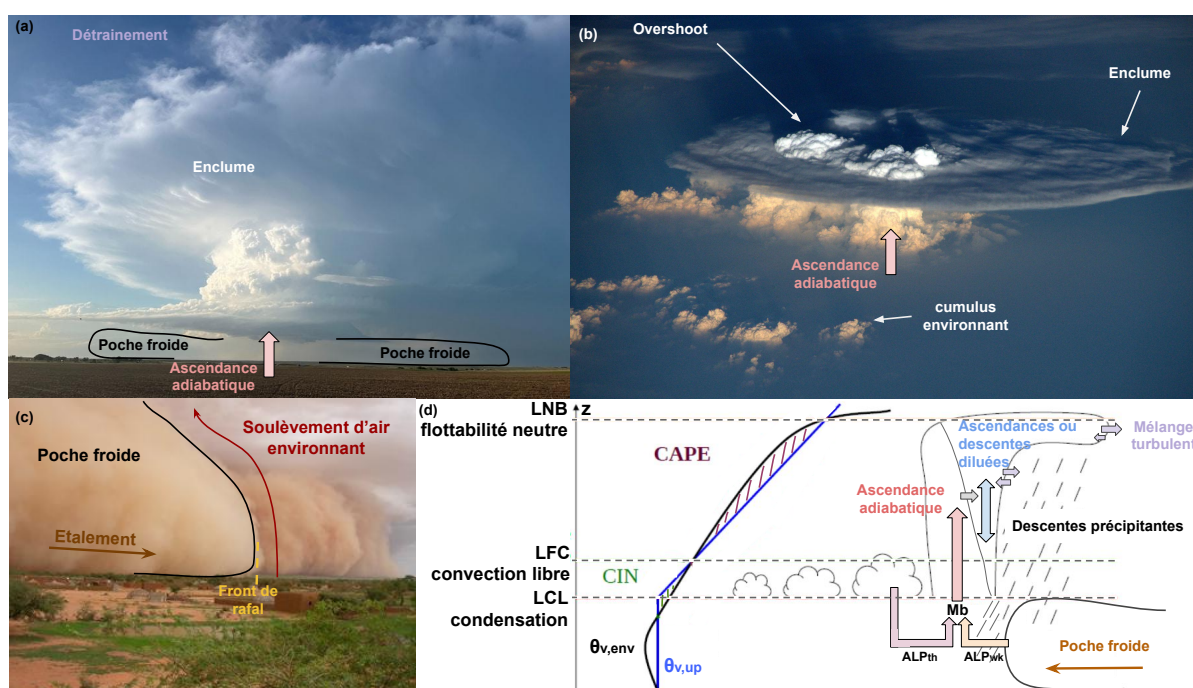


FIGURE II.7 – Photos et schémas illustrant les concepts véhiculés dans la construction du modèle de convection et de son couplage avec les processus sous-nuageux. En haut à gauche : photo de cumulonimbus au-dessus des plaines du Texas Panhandle, avec localisation de l'enclume, de la base de l'ascendance adiabatique et des poches froides (juin 2023 — Photo : Jack Sillin, utilisée permission); en haut à droite : cumulonimbus vu d'en haut avec le sommet de l'enclume qui s'étale sous la tropopause, les nuages bourgeonnants qui overshoot au milieu et des cumulus de couche limite autour (photo prise par l'astronaute Tim Peake depuis la Station spatiale internationale durant la mission Principia (2016), crédit : ©ESA/NASA, utilisée conformément aux conditions d'utilisation du site de l'ESA); en bas à gauche : photo d'un haboob au Sahel, tempête de poussières créée par le front de rafale d'une poche froide (crédit : © Françoise GUICHARD / Laurent KERGOAT / CNRS Images); en bas à droite : schéma conceptuel du schéma de convection et de son couplage avec les processus sous-nuageux.

La paramétrisation de la convection profonde dans LMDZ6A est basée sur une version modifiée du schéma en flux de masse de d'Emanuel (Emanuel, 1991, 1993), modifiée par (Grandpeix

et al., 2004) et accompagnée d’une paramétrisations des poches froides (Grandpeix and Lafore, 2010; Grandpeix et al., 2010).

Tout comme pour la convection peu profonde, la paramétrisation de la convection profonde représente une population de cumulonimbus qui se produirait dans une maille par un cumulonimbus effectif. Le transport convectif dans ce cumulonimbus effectif est modélisé par des flux de masse et des matrices d’échange d’air entre les trois zones considérée comme constituant un cumulonimbus. Elles sont schématisés Figure II.7 (d).

Une ascendance adiabatique qui n’entraîne pas d’air latéralement au dessus de la base du nuage, mais est “épluchée” petit à petit. Elle est supposée suffisamment rapide pour emporter avec elle l’eau liquide ou solide condensée en son sein.

Une population d’ascendances ou descentes diluées, lesquelles sont saturées, créée en mélangent une fraction d’air épluchée de l’ascendance adiabatique avec de l’air de l’environnement selon une PDF imposée. Elle est découpée en bins, puis chaque parcelle d’air correspondant à un bin est “envoyée” à son niveau (couche) de flottabilité neutre, créant ainsi une matrice dont chaque terme est un échange d’air entre deux couches du modèle. Avant de former ces mélanges, on fait précipiter l’ascendance adiabatique dans un troisième compartiment, puis à nouveau on fait précipiter l’eau en excès de la saturation avant que le flux ascendant ou descendant ne détraîne dans l’environnement. Les PDF utilisées pour fabriquer les ascendances mélangées ont été modifiées par Grandpeix et al. (2004) par rapport à la version originale du schéma d’Emanuel.

Les descentes précipitantes ou descentes insaturées qui reçoivent l’ensemble des précipitations formées lors de l’épluchage ou lors du déentraînement dans l’environnement. Une fraction de ces descentes s’effectue partie en dehors des nuages, dans de l’air non saturé en humidité, permettant leur évaporation. Cette évaporation de quantités très importantes de précipitations force des mouvements subsidants rapides. En dessous de la base des nuages, c’est l’intégralité des précipitation qui est hors des nuages. Leur réévaporation est à l’origine des courants de densité, ou poches froides, créées sous les cumulonimbus. Un paramètre libre important de ces descentes est la fraction de la surface de la maille qu’elles couvrent. Cette fraction **SIGDZ** est imposée et constante sur la verticale.

En pratique, les tendances convectives Q_1 et Q_2 sont séparées en deux parties : une partie saturée et une partie insaturée. Les tendances saturées Q_1^{sat} et Q_2^{sat} tiennent compte de l’ascendance adiabatique, des ascendances et descentes diluées et d’une subsidence compensatoire. Les tendances insaturées Q_1^{unsat} et Q_2^{unsat} tiennent compte des descentes insaturées ainsi qu’une ascendance compensatoire de ces descentes.

II.3.4.1 L’efficacité de l’auto-conversion d’eau liquide nuageuse en pluie

C’est est un des processus clé, qui module fortement l’intensité de la convection profonde ainsi que les nuages et l’humidité de la haute troposphère.

Cette efficacité de précipitation EP est comprise entre 0 et une valeur maximum, paramètres ajustable EP_{max} , selon

$$EP = \min \left[\max(1 - \frac{\epsilon}{q_c}, 0), EP_{\text{max}} \right] \quad (\text{II.38})$$

où q_c est le rapport de mélange de l’eau condensée et ϵ est donné par :

$$\epsilon = \text{ELCV} \quad \text{si } T_c > 0 \quad (\text{II.39})$$

$$= \text{ELCV}(1 - T_c/\text{TLCV}) \quad \text{sinon} \quad (\text{II.40})$$

où T_c est la température en degrés Celsius, **TLCV** (-55°C par défaut) est une température critique, toujours négative, renforçant l'efficacité de précipitation pour les températures très froides, et **ELCV** (par défaut 0.003 kg/kg) une concentration critique d'eau condensée. Si $q_c < \text{ELCV}$, alors l'efficacité de précipitation est systématiquement nulle.

Le paramètre **OMEPMAX** $= 1 - EP_{\max}$ (pour *One Minus EP_{max}*) est un des paramètres les plus sensibles pour le tuning du modèle global. On choisit de régler **OMEPMAX** plutôt que EP_{\max} car il contrôle plus directement la fraction minimum d'eau liquide non précipitée dans la partie haute du cumulonimbus, qui deviendra l'humidité de l'air détrainée de la convection. De plus, EP_{\max} prend des valeurs très légèrement inférieures à 1. L'efficacité maximale de l'auto-conversion est généralement atteinte en haut du nuage de convection profonde.

Dans la haute troposphère tropicale ce détrainement d'eau nuageuse est la principale source d'humidité (avec l'advection verticale à grande échelle). Si les conditions sont réunies, cette eau va se transformer en nuage haut (cirrus); et c'est le schéma de condensation grande échelle qui va s'en charger. C'est là qu'intervient le paramètre **FALLV**, comme un facteur sur la vitesse terminale des cristaux de glace. Plus **FALLV** est grand, plus les cristaux tombent vite, donc moins il y a de nuages haut et donc moins il va rester d'humidité quand le nuage va se dissiper. On peut donc dire que **OMEPMX** règle la source d'humidité dans la haute troposphère et **FALLV** son puits.

Dans les zones où on a à la fois source et puits d'humidité, les valeurs de ces deux paramètres peuvent se compenser lors d'un exercice de tuning pour à la fin avoir la bonne humidité dans la haute troposphère. Cette humidité va ensuite être transportée par la dynamique dans les moyennes latitudes. Si maintenant on se place à un endroit où il n'y a plus de convection profonde, il n'y a plus que **FALLV** qui joue; toujours comme un facteur sur le puits d'humidité de la haute troposphère. Ainsi, dans ces régions, les deux facteurs ne se compensent plus. **FALLV** va donc contrôler à quel point l'humidité de la haute troposphère (générée par la convection profonde dans les tropiques) va être transportée vers les plus hautes latitudes. Une valeur faible de **FALLV** résultant en une haute troposphère extra tropicale plus humide et plus nuageuse (Fester et al., en préparation).

II.3.4.2 Déclenchement de la convection profonde

La convection profonde est un processus à seuil. La colonne atmosphérique peut être instable au regard de l'instabilité convective mais il faut en général que les particules d'air ascendantes franchissent une inhibition convective (ou CIN) à la base des nuages. Pour que la convection profonde se déclenche, il faut que de l'air chaud de la couche limite soit soulevé suffisamment haut pour atteindre le niveau de convection libre (LFC pour *Level of Free Convection*).

Une spécificité des travaux développés au LMD depuis plus d'une vingtaine d'années autour de Jean-Yves Grandpeix est d'avoir mis la question du déclenchement et du contrôle de l'intensité (fermeture) convective au cœur de ses recherches. Dans ce que l'on appelle la "Nouvelle physique" de LMDZ, qui correspond aux versions de LMDZ à partir de la version LMDZ5B développée pour CMIP5 (Hourdin et al., 2013), le déclenchement de la convection profonde dépend de processus sous nuageux. Ce déclenchement dépend en particulier de l'intensité des mouvements convectifs de la couche limite, représentés au travers du modèle du thermique (Section II.3.2), et des poches froides créées sous les cumulonimbus par la réévaporation des pluies convectives. Ces poches froides, en s'étalant horizontalement, soulèvent de l'air à l'avant de leur "front de rafale" des parcelles d'air, pouvant déclencher de nouvelles cellules convectives.

Pour formaliser le contrôle du déclenchement par les poches froides, Grandpeix and Lafore (2010) introduisent le concept d'énergie disponible pour la convection, ou **ALE** pour *Available Lifting Energy*. Cette énergie peut venir soit des thermiques (ALE_{th}), soit des poches froides

(ALE_{wk}). Il faut donc, pour que la convection profonde se déclenche, satisfaire à l'inégalité :

$$ALE = \max(ALE_{th}, ALE_{wk}) > CIN \quad (\text{II.41})$$

II.3.4.3 La fermeture convective

Une fois la convection déclenchée, reste à déterminer son intensité, caractérisée dans LMDZ par le flux de masse à la base du nuage. Cette étape constitue la fermeture du schéma de convection profonde.

Ce flux de masse à la base du nuage, M_b , est défini à partir de la puissance de soulèvement disponible, notée ALP pour *Available Lifting Power*, fournie soit par les thermiques (ALP_{th}) soit par les poches (ALP_{wk}). Cette énergie disponible à la base de la convection doit être suffisante pour que l'énergie cinétique à la base de la convection profonde soit non nulle, ce qui donne la relation de fermeture suivante :

$$M_b = \frac{ALP_{bl} + ALP_{wk}}{2.w_b^2 + |CIN|} \quad (\text{II.42})$$

où w_b , la vitesse verticale minimale exigée à la base du nuage pour que la convection ait lieu, est une fonction de la pression (croissant en s'éloignant de la surface) faisant intervenir deux paramètres libres supplémentaires :

$$w_b = \mathbf{WBSRF} + \frac{\mathbf{WBTOP}}{1 + 500/(P_s - P_{LFC})} \quad (\text{II.43})$$

où P_s est la pression à la surface et P_{LFC} la pression au niveau de convection libre ou LFC. Cette formule avait été introduite dans la version 5B de LMDZ, première version dite "nouvelle physique" développée à l'occasion de l'exercice CMIP5, pour prendre en compte via la hauteur de la base du nuage convectif un contraste océan continent (la base des nuages est plus haute sur continent que sur océan (Hourdin et al., 2013)).

II.3.4.4 Énergie et puissance de soulèvement fournies par les thermiques

L'énergie cinétique disponible fournie par les thermiques est calculée à partir de la vitesse maximum des panaches thermiques sous le niveau de condensation $W_{th,max*}$,

$$ALE_{th} = \frac{1}{2} W_{th,max*}^2 \quad (\text{II.44})$$

La puissance disponible ALP_{th} peut être formulée comme un débit massique multiplié par une énergie cinétique (comme un flux d'énergie cinétique) :

$$ALP_{bl} = \mathbf{ALPBLK} \frac{1}{2} \overline{\rho w'^3} \quad (\text{II.45})$$

où \mathbf{ALPBLK} est l'efficacité de la transmission de la puissance des thermiques à la convection profonde. Plus ce facteur est grand, puis la puissance fournies par les thermiques est grande, donc plus le flux de masse de la convection profonde à la base (M_b) est grand (et dépendant des thermiques), donc plus la convection est intense.

Si on remarque que la vitesse verticale décrite par le modèle du thermique vaut uniformément w_{th} dans le thermique et $\alpha w_{th}/(1 - \alpha)$ à l'extérieur (pour que les flux de masse ascendant et subsidant se compensent), on montre facilement que

$$\overline{w'^3} = \left[\alpha - \frac{\alpha^3}{(1 - \alpha)^2} \right] w_{th}^3 = \frac{\alpha(1 - 2\alpha)}{(1 - \alpha)^2} w_{th}^3$$

II.3.4.5 Le modèle de poches froides

Le modèle de poches froides développé par Grandpeix and Lafore (2010); Grandpeix et al. (2010) est une des grandes originalités du modèle du LMD. Il suppose une population uniforme de poches circulaires de même rayon r , même hauteur, et s'étalant à une vitesse C^* . Les poches ont une densité surfacique imposée, différente sur continent (8 poches par $1000 \times 1000 \text{ km}^2$ dans LMDZ6A) et sur océan (paramètre réglage **WDENSO** égal par défaut à 10 poches par $100 \times 100 \text{ km}^2$). Les poches sont de plus caractérisées par des profils verticaux d'écart de température potentielle $\delta\theta(z)$, d'humidité $\delta q_v(z)$ et de vitesse verticale $\delta\omega(z)$ entre la poche et son environnement. Si on note X_{wk} la valeur de la grandeur X dans la poche (wake) et X_{ex} sa valeur en dehors de la poche, alors $\delta X = X_{wk} - X_{ex}$.

L'écart de température entre la poche et son environnement est créé par la réévaporation des précipitation dans les poches. Cet écart est en général maximum (le plus négatif) près de la surface et s'annule à une hauteur h_{wk} , typiquement de quelques centaines de mètres sur océan et jusqu'à 2 km sur continent. Dans le modèle de poche, le profil de différentiel de vitesse verticale $\delta\omega$ décroît linéairement en fonction de la pression depuis la hauteur h_{wk} jusqu'à la surface où il s'annule. Dans cette partie, on suppose que l'étalement horizontal compense exactement la subsidence. Au dessus de h_{wk} , $\delta\omega$ décroît linéairement jusqu'à un niveau arbitraire 600 hPa. Dans cette partie haute, on suppose que l'accroissement de la vitesse verticale quand on s'approche de h_{wk} est compensé par un entrainement latéral e_{wk} .

De l'écart de température on déduit une énergie d'effondrement de la poche appelée *WAPE* (pour *Wake Available Potential Energy*) et donnée par :

$$W A P E = g \int \frac{\delta\rho}{\bar{\rho}} = -g \int_0^\infty \frac{\delta\theta_v}{\theta_v} dz \quad (\text{II.46})$$

dont on déduit la vitesse d'étalement

$$C^* = \gamma \sqrt{2 W A P E} \quad (\text{II.47})$$

où $\gamma = 0.33$, ρ est la masse volumique de l'air et θ_v sa température potentielle virtuelle.

La vitesse d'étalement permet à son tour de calculer l'évolution de la fraction de surface couverte par les poches σ_{wk} :

$$\frac{\partial \sigma_{wk}}{\partial t} = 2\pi r C^* D_{wk} = 2C^* \sqrt{\pi D_{wk} \sigma_{wk}} \quad (\text{II.48})$$

On a deux équations pronostiques pour l'évolution de $\delta\theta$ et δq_v à savoir :

$$\begin{cases} \partial_t \delta\theta = \frac{\delta Q_1^{cv} + \delta Q_1^{wk}}{C_p} - \bar{\omega} \partial_p \delta\theta - \frac{K_{gw}}{\tau_{gw}} \delta\theta, \\ \partial_t \delta q = \frac{\delta Q_2^{cv} + \delta Q_2^{wk}}{L_v} - \bar{\omega} \partial_p \delta q \end{cases} \quad (\text{II.49})$$

C'est la réévaporation des précipitations dans les courants insaturés qui est le moteur premier du développement des poches. Ce processus est traduit dans le modèle en attribuant le terme de chauffage Q_1^{unsat} à l'intérieur des poches alors que Q_1^{sat} agit à l'extérieur des poches. De façon cohérente avec ce découpage, on suppose que la partie saturée du schéma convectif voit les profils extérieurs aux poches, et les descentes insaturées l'intérieur des poches. Dans LMDZ6A, on suppose de plus que les thermiques ne s'activent que dans la fraction de la maille située à l'extérieur des poches froides. Le modèle du thermique induit donc également une tendance sur les écarts poche-environnement opposée à la tendance moyenne sur la maille. Au bout du

compte, le contraste de tendances convectives (peu profonde et profonde) entre les poches et leur environnement s'écrit :

$$\begin{cases} \delta Q_1^{cv} = \frac{Q_1^{unsat}}{\sigma_{wk}} - \frac{Q_1^{sat}}{1-\sigma_{wk}} - Q_1^{th} \\ \delta Q_2^{cv} = \frac{Q_2^{unsat}}{\sigma_{wk}} - \frac{Q_2^{sat}}{1-\sigma_{wk}} - Q_2^{th} \end{cases} \quad (\text{II.50})$$

Le modèle de poches lui-même chauffe et humidifie de façon différenciée l'air à l'intérieur et à l'extérieur des poches, sous l'effet d'entraînement latéral d'air de l'environnement au dessus de h_{wk} , de subsidence dans les poches et d'ascendance compensatoire dans l'environnement :

$$\begin{cases} \frac{\delta Q_1^{wk}}{C_p} = -\frac{e_w}{\sigma_{wk}} \delta \theta - \delta \omega \partial_p \bar{\theta} - (1 - 2\sigma_{wk}) \delta \omega \partial_p \delta \theta \\ -\frac{\delta Q_2^{wk}}{L_v} = -\frac{e_w}{\sigma_{wk}} \delta q - \delta \omega \partial_p \bar{q} - (1 - 2\sigma_{wk}) \delta \omega \partial_p \delta q \end{cases} \quad (\text{II.51})$$

Les termes $-\bar{\omega} \partial_p \delta$ dans l'Equation II.49 viennent palier en partie le fait que les écarts δ ne sont pas transportés par la dynamique jusque là. On prend donc en compte, dans les paramétrisations, la partie verticale de l'advection grande échelle pour palier en partie ce manque.

Pour finir, le dernier terme, présent uniquement dans de la partie en θ de l'Equation II.49, correspond à la réduction des contrastes en température par les ondes de gravité avec un coefficient spécifié comme le rapport d'une efficacité K_{gw} sur une constante de temps τ_{gw} qui dépend de la fraction surfacique des poches et de la fréquence de Brunt-Väisälä.

C'est en s'étalant que les poches froides vont soulever de l'air et fournir une énergie et une puissance de soulèvement à la convection profonde. L'énergie disponible pour le soulèvement est prise proportionnelle à C^{*2} et se ramène finalement à

$$ALE_{wk} = W A P E \quad (\text{II.52})$$

La puissance de soulèvement est calculée comme un facteur d'efficacité **ALPWKK** fois la puissance horizontale fournie au niveau du front de rafale, produit pour chaque poche du débit de masse au niveau du front de rafale entourant la poche $2\pi\rho r h_{wk} C^*$ par l'énergie cinétique massique en $C^{*2}/2$, donnant finalement :

$$ALP_{wk} = \mathbf{ALPWKK} \rho D_{wk} C^{*3} h_{wk} \pi r \quad (\text{II.53})$$

II.3.4.6 Le déclenchement stochastique

La paramétrisation du déclenchement de la convection a évolué depuis la version initiale de Grandpeix and Lafore (2010) à l'occasion de la thèse de Nicolas Rochetin (Rochetin et al., 2014). Dans le cas où le déclenchement est réalisé par les panaches thermiques, en plus du fait que l'énergie disponible dépasse l'inhibition ($ALE_{th} > CIN$), on demande qu'au moins un cumulus soit suffisamment grand (taille évaluée comme la surface de la base du cumulus).

On peut plaquer plusieurs images derrière ce critère supplémentaire de déclenchement, liées à la façon dont on souhaite distinguer les paramétrisation de la convection peu profonde et profonde dans le modèle. La première est simplement de dire qu'on demande à "la convection profonde d'être profonde" c'est à dire d'atteindre typiquement au moins 5 km de haut. On peut aussi voir la convection profonde comme le processus qui conduit à la formation de nuages. On peut aussi se dire qu'on demande à la convection d'atteindre le point de congélation, à partir du quel le dégagement de chaleur latente dans les ascendance devient nettement plus grand. Enfin, on peut voir derrière ce critère supplémentaire comme lié au contrôle de la profondeur des cumulus par l'humidité troposphérique. En effet, une des fortes limitations de l'extension verticale des cumulus, est le fait qu'en entrainant de l'air sec de la troposphère libre, l'eau

nuageuse se réévapore, provoquant à la fois une dissipation du nuage et un refroidissement des parcelles d'air, limitant ou stoppant leur ascension. Ce processus de limitation devient de moins en moins opérant au fur et à mesure que les cumulus grossissent, isolant le cœur du cumulus de l'air environnant.

En pratique, La convection va se déclencher si il y a dans la maille au moins un cumulus plus grand qu'une taille seuil **STRIG**, paramètre ajustable. Pour savoir si c'est le cas, on utilise une approche stochastique d'où l'appellation de déclenchement stochastique. Rochetin et al. (2014) ont pu montré, dans des simulations LES, que la distribution de surface des thermiques à la base des nuages, est composée de deux modes, avec des distributions exponentielles. Seule la "deuxième population", celle des plus gros nuages, peut conduire à un déclenchement. Pour le déclenchement stochastique, on utilise le panache moyen effectif du thermique pour déduire les paramètres de cette distribution de tailles. On fait un tirage aléatoire d'un nombre R entre 0 et 1 que l'on compare à la probabilité de non déclenchement suivante :

$$P_{\Delta t} = \left[\left(1 - \exp \left(-\frac{\mathbf{STRIG}}{S_2} \right) \right)^{N_2} \right]^{\Delta t / \tau} \quad (\text{II.54})$$

Si $R > P_{\Delta t}$ alors la convection profonde est déclenchée. Cette comparaison entre R et $P_{\Delta t}$ revient à tirer une surface de thermique et de la comparer à **STRIG**. Dans le code, $\tau = \text{tau_trig_deep}$ ou tau_trig_shallow et vaut 1800 ou 600 s.

Ce tirage aléatoire est nécessaire quand la probabilité de déclenchement est inférieure à un événement par maille. Si on suppose que la probabilité de déclenchement est uniforme horizontalement, ce déclenchement stochastique va permettre d'avoir la bonne statistique de déclenchement réalisée sur un grand nombre de mailles.

[Maëlle : j'ai relu jusqu'à la]

II.3.5 Transfert radiatif

On décrit très sommairement ici le code de transfert radiatif utilisé encore aujourd'hui en standard dans LMDZ, sachant que le chapitre ?? fera une beaucoup plus large part à la présentation du nouveau code de rayonnement Ecrad actuellement en cours de test.

Le modèle radiatif actuel (comme le nouveau) est en fait le code développé au CEPPMT pour le modèle de prévision IFS.

Ce modèle se base sur l'hypothèse plan-parallèle et sur une équation à deux flux (two stream) dans laquelle on sépare les photons entre des photons qui montent et des photons qui descendent, ou autrement dit le rayonnement montant (ou descendant) est traité en une grandeur flux unique représentant l'intégrale du rayonnement sur le long de directions pointant vers le bas (respectivement vers le haut).

En ce qui concerne l'absorption où l'émission par les gaz (notamment de la vapeur d'eau, du dioxyde de carbone, et de l'ozone) le traitement spectral est très différent pour le rayonnement thermique (essentiellement infra-rouge) et le rayonnement solaire.

Pour le rayonnement thermique, on néglige la diffusion et on suppose qu'il se propage selon un angle moyen zénithal θ donné par $1/\cos\theta = 1.67$. C'est ce qu'on appelle l'approximation diffuse. Le code RRTM (Rapid Radiative Transfer Model) importé pour faire l'intégration spectrale est un modèle dit k-corrélé, dans lequel l'intégration sur les fréquences est remplacée par une intégrations sur des coefficients d'extinction, sur un certain nombre d'intervalles spectraux. En l'occurrence 16 dans la versions utilisée dans LMDZ6A. [Ajouter citation de Mlawer et al., 1997]

Concernant le rayonnement solaire, celui-ci est décomposé en un rayonnement directionnel provenant du soleil, et d'un rayonnement diffus utilisant là aussi un angle moyen (le même que

pour l'infra rouge[Vérifier]. L'intégration spectrale dans cette partie solaire ne repose pas dans la version actuelle sur une approche k-corrélée mais sur un modèle en bande large (6 au total) dans lesquelles les transmittivité sont calculées en fonction de la masse d'atmosphère traversée et d'une masse pondérée par la pression (pour tenir compte des effets non linéaires de saturation de l'absorption) Ce schéma est décrit dans [Ajouter citation de Fouquart and Bonnel (1980)].

Les paramétrisations des nuages et de la convection exposées plus haut fournissent au code radiatif des profils verticaux de fractions nuageuse et de contenu en eau des nuages en moyenne sur la maille horizontale. En pratique, dans la version actuelle de LMDZ, on calcule séparément des profils pour la convection profonde et pour le schéma de condensation à grande échelle (qui rend compte également des nuages de convection peu profonde). Quand les deux types de nuages co-existent, on somme les fraction nuageuses (en bornant la somme à 1) et on effectue une moyenne pondérée des contenus en eau.

Il faut alors faire des hypothèses quant aux propriétés micro (taille de gouttes, phase) et macro physiques (géométrie) des nuages.

La partie microphysique est très sommaire. La répartition entre phase solide et liquide de l'eau condensée est imposée en fonction de la température. Pour l'eau liquide, un rayon moyen est utilisé. Il est soit constant sur le globe, soit il est relié à une concentration en aérosols dans la maille. Plus la concentration en aérosols pouvant servir de noyaux de condensation est grande, plus l'eau liquide condense en un grand nombre de petites gouttes d'eaux, rendant le nuage plus réfléchissant pour le rayonnement solaire et plus opaque (et donc émetteur) pour le rayonnement infra-rouge. Pour les cristaux de glace, le rayon effectif r_{ei} (ice effectif radius) est imposé comme une fonction linéaire de la température en degrés celsius, T_c , entre une valeur minimum $r_{ei,min}$ ($=16 \mu m$ par défaut) pour $T_c < -81 \text{ }^\circ C$, et une valeur maximum $r_{ei,max}$ ($=61.29 \mu m$ par défaut), atteinte pour $T_c = 0 \text{ }^\circ C$:

$$r_{ei} = \mathbf{REI} \times \max[(r_{ei,max} - r_{ei,min}) * T_c / 81.4 + r_{ei,max}, r_{ei,min}] \quad (\text{II.55})$$

Le paramètre **REI** est un facteur introduit intentionnellement pour le tuning.

Pour la prise en compte de la géométrie des nuages, le code sépare une colonne ciel clair et une colonne nuageuse dans la maille. Pour connaître la surface couverte par la colonne ciel-clair, il est nécessaire de faire des hypothèses sur le recouvrement des nuages entre deux couches du modèle atmosphérique. Dans la version standard de LMDZ, on utilise la version standard dite maximum-aléatoire, dans laquelle on suppose que les nuages de deux couches nuageuses adjacentes se recouvrent au maximum (avec l'idée qu'il s'agit en fait des mêmes nuages sur les deux couches) alors que deux couches nuageuses séparées par une couche sans nuages se recouvriront aléatoirement.

Les questions relatives à cette prise en compte de la géométrie des nuages seront largement discutées dans le chapitre ??.

Chapitre III

La méthode d'History Matching with Iterative Refocusing

Remarque lecture : Chapitre presque aboutie : j'ai encore quelques remarques à prendre en compte, et à reprendre la dernière partie

L'équipe de développement de LMDZ s'est récemment dotée d'une méthode objective d'ajustement de ses paramètres libres lors de l'ANR HighTune entre 2016 et 2020. C'est la méthode d'History Matching with Iterative Refocusing (HMIR) qui est choisie et implémentée dans l'outil `htexplo`. Cette méthode est basée sur une approche bayésienne de quantification des incertitudes. Elle a été développée par Williamson et al. (2013, 2015, 2017) pour la calibration du modèle de climat du Royaume-Uni HadCM3 (Gordon et al., 2000; Pope et al., 2000), puis du modèle d'océan NEMO (Madec, 2008). Ce type de méthode a initialement été développé pour l'industrie pétrolière et la simulation des puits de pétrole dans les années 90 (Craig et al., 1996; Cumming and Goldstein, 2010), puis a été transposée à d'autres types de modèles, notamment par Vernon et al. (2010) pour un modèle de formation de galaxie et Edwards et al. (2011) pour un modèle de climat de complexité intermédiaire. La méthode d'ajustement en elle-même est aveugle au type de modèle que l'on souhaite calibrer. Le choix de la méthode et son utilisation dépendent en revanche des spécificités du modèle et du système modélisé. Pour les modèles de climat, la méthode d'ajustement choisie doit être adaptée à des modèles coûteux en temps de calcul et ayant un grand nombre de paramètres incertains, ce qui est le cas de la méthode HMIR.

La méthode porte différents noms suivant les articles. History matching, fait référence au fait qu'un modèle de nature prédictive est ajusté sur des données passées, dites historiques. On pourrait le traduire par "ajustement à l'historique" ou "calage sur l'historique". Iterative refocusing fait référence à l'aspect itératif de la méthode, qui réduit progressivement l'espace des paramètres acceptables au cours d'itérations aussi appelées vagues de tuning. Nous avons retenu ici l'acronyme HMIR History Matching with iterative refocusing qui souligne ces deux aspects de la méthode.

La méthode d'HMIR est actuellement utilisée par les deux centres de modélisation climatique français, l'IPSL et Météo-France, pour la mise en place de la configuration de référence de leur modèle de climat pour le prochain exercice d'intercomparaison des modèles de climat, CMIP7. C'est la première fois que cette méthode est utilisée pour la préparation d'un modèle à un exercice d'intercomparaison des modèles de climat. Elle a cependant été utilisée a posteriori de l'exercice CMIP6 pour vérifier que la méthode HMIR était bien capable de calibrer le modèle LMDZ aussi bien que la calibration traditionnelle par essai-erreur, sur les paramètres libres entrant en jeux dans les paramétrisations de la couche limite convective (Hourdin et al., 2021).

Ce chapitre présente la méthode d'HMIR telle qu'elle est comprise et pratiquée aujourd'hui dans la communauté de modélisation du climat. Ce chapitre bénéficie pleinement du recul ap-

porté par les expériences de cette thèse. A ce titre, il est en partie un résultat de la thèse, et pas seulement une présentation des outils et méthodes.

Après une brève introduction, ce chapitre présente la méthode d'History Matching with Iterative Refocusing (Section III.2) puis l'outil `htexplo` dans lequel elle est implémentée (Section III.3). Enfin, l'utilisation d'`htexplo` pour le tuning de LMDZ sera présentée section Section III.4.

III.1 Introduction

L'arrivée des méthodes objectives d'ajustement des modèles dans la communauté de la modélisation du climat date d'une quinzaine d'année. Deux types de méthodes, issues de la recherche en statistiques, informatique ou de l'ingénierie, ont été utilisées pour la calibration de modèles de climat.

La première famille de méthodes est celle de l'optimisation. La question de la calibration se pose alors comme la recherche du vecteur de paramètres libres conduisant à la meilleure simulation du climat. Dans ce cadre, l'approche consiste à définir une mesure de la distance entre simulation et référence, au travers d'une fonction communément appelée fonction coût. Une méthode classique d'optimisation issue du calcul variationnel, de type algorithme de descente de gradient, peut alors être appliquée afin de trouver le vecteur de paramètres minimisant cette fonction coût. Cette famille de méthode est au cœur des méthodes d'assimilation variationnelle (Talagrand and Courtier, 1987) qui ont connu un grand succès en prévision météorologique dans les années 90. Elles sont dans ce cadre utilisées pour déterminer l'état initial des prévisions, appelé analyse, et non pas les valeurs des paramètres libres du modèle de prévision du temps. La méthode 4D-Var, adoptée en 1997 par le Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT, également connu sous son acronyme anglais ECMWF), a en particulier conduit à une amélioration significative des prévisions météorologiques. Elle repose sur un algorithme de descente de gradient, dans lequel le gradient généralisé du modèle par rapport à ses conditions initiales est estimé en propageant à rebours dans le temps l'adjoint du modèle direct (Courtier et al., 1994). Elle est aussi utilisée pour construire les produits dits de réanalyses du CEPMMT (dont les produits ERA-interim et ERA5 utilisés dans cette thèse), constitué des analyses produites a posteriori avec une chaîne d'assimilation unique appliquée aux dernières décennies. Ces méthodes d'assimilation variationnelle ont par la suite été appliquées à la calibration des paramètres libres de plusieurs modèles de climat ces dernières années (Bellprat et al., 2012, 2016; Yang et al., 2013; Zou et al., 2014; Zhang et al., 2015).

L'utilisation de ces méthodes pour la calibration des modèles de climat est critiquée par une partie de la communauté scientifique des staticiens, qui considèrent que choisir un vecteur de paramètre sur un critère plus fin que l'incertitude de comparaison entre simulations et références ne peut pas se justifier. De fait, cette incertitude se répercute sur les valeurs des vecteurs de paramètres, et la méthode de calibration utilisée doit permettre d'informer sur cette incertitude paramétrique. L'enjeu n'est plus de déterminer le meilleur vecteur de paramètres, mais ceux qui permettent de simuler un climat compatible avec les références, en tenant compte des incertitudes de comparaison entre simulations et références. Les méthodes abordant la question de la calibration sous cet angle sont qualifiées de méthodes de quantification des incertitudes (UQ pour *Uncertainty Quantification*). L'ensemble de simulation correspondant à cet ensemble de vecteur de paramètre constitue ce que l'on appelle communément un ensemble aux paramètres physiques perturbés, nommé PPE pour *Perturbed Physical Ensemble* en anglais. Il peut être utilisé pour quantifier l'incertitude paramétrique du modèle, sous la contrainte que tous ces vecteurs de paramètres conduisent à ce que l'on aura défini comme étant une bonne simulation du climat, et compte tenu des incertitudes de comparaison simulations-références. Puisqu'il n'y

a pas de raison objective de privilégier un vecteur de paramètre de cet ensemble plutôt qu'un autre, il n'y a pas de raison de fixer une seule configuration de référence du modèle. Les centres de modélisations du climat devraient en toute rigueur proposer un ensemble de configurations à leurs utilisatrices, représentatif de l'incertitude paramétrique de leur modèle, et participer aux exercices CMIP avec cet ensemble de configurations afin d'instruire l'impact de l'incertitude paramétrique sur la simulation du climat futur. Ce dernier point s'avère très coûteux en temps de calcul et n'a pas encore été exploré à ce jour. Ces méthodes, plus récentes que l'optimisation classique, ont été utilisées par Bonnet et al. (2025); Edwards et al. (2011); Hourdin et al. (2021); Williamson et al. (2013) pour la calibration de modèles de climat, ou par Jackson et al. (2008); Rougier (2007) pour instruire l'incertitude paramétrique de leur modèle.

Notons que l'utilisation de PPEs pour l'analyse de modèles contenant des paramètres libres incertains est classique, et a été appliquée pour l'analyse de modèles climatiques avant d'être utilisée pour la calibration de modèles tel que la méthode d'HMIR le propose. Les PPEs ont en particulier été utilisés pour explorer l'incertitude paramétrique des projections climatiques futures (Bernstein and Neelin, 2016; Collins et al., 2006; Peatier et al., 2022; Sexton et al., 2021; Yamazaki et al., 2021), avec des études limitants les membres de leur PPE à ceux dont le climat simulé était considéré comme acceptable vis à vis des observations (Neelin et al., 2010; Hourdin et al., 2023).

Ces deux familles de méthodes (optimisation ou quantification des incertitudes) partagent un certain nombre de caractéristiques, dont celle de devoir définir une mesure de la distance entre simulations et références. A l'heure actuelle, il n'existe pas de standard pour la définition de cette distance. Un certain nombre de pratiques, dont celle d'utiliser le flux radiatif net au sommet de l'atmosphère et/ou les températures moyennes globales, comme critère de selection des paramètres, se retrouvent néanmoins dans plusieurs groupes de modélisation de climat (Hourdin et al., 2017). Le fait que chaque centre de modélisation explicite et communique cette mesure de distance est une avancée importante, qui pourra permettre de documenter cette diversité d'approche et de les comparer.

Les deux types de méthodes divergent en revanche dans leur manière de traiter les incertitudes. Dans le premier cas, l'incertitude de comparaison entre simulation et référence est prise en compte dans la fonction coût, de manière à accorder moins d'importance aux critères plus incertains. Dans le second, l'incertitude est traitée de manière plus fondamentale, de telle sorte que la méthode d'ajustement renseigne également sur l'incertitude paramétrique du modèle après calibration. Les méthodes basées sur la quantification des incertitudes permettent de plus d'identifier les cas où aucun vecteur de paramètres n'est compatible avec la gamme d'incertitude de comparaison entre simulation et référence. Cela permet ainsi d'informer sur la capacité ou non d'un modèle à simuler un certain nombre de caractéristiques en s'affranchissant de l'incertitude paramétrique, et ainsi d'informer sur ce que l'on appelle l'erreur structurelle du modèle. Ce n'est pas le cas des méthodes d'optimisation qui vont toujours trouver un vecteur de paramètres qui minimise la fonction coût, même si in fine ce vecteur de paramètres conduit à la simulation de certaines caractéristiques trop éloignées de leur valeur de référence.

C'est notamment pour ces raisons que les deux centres de modélisation français ont choisi de se doter d'une méthode objective d'ajustement faisant partie de la famille des quantifications d'incertitude. C'est la méthode d'HMIR qui a été retenue pour sa capacité à traiter la question des compensations d'erreurs.

III.2 La méthode d'History Matching with Iterative Refocusing

III.2.1 Préambule

Cette partie présente la méthode d'HMIR du point de vue d'une modélisatrice du climat, et telle que je l'explique aux collègues de ma communauté scientifique. Elle n'a pas l'objectif de poser rigoureusement le cadre mathématique sur lequel cette méthode est basée, bien que quelques éléments soient présentés dans la section suivante. En effet, l'appropriation de ce cadre par moi-même et ma communauté n'est pas complète, ce qui est inhérent à la pratique d'une recherche interdisciplinaire mais n'est en aucun cas un frein à une utilisation éclairée de la méthode pour l'ajustement de LMDZ. La force de la collaboration avec les spécialistes de cette méthode réside en effet sur leur capacité à nous transmettre les notions clés pour bien comprendre et utiliser cette méthode dans toute sa finesse sans forcément en devenir des spécialistes à notre tour. Cette méthode est un objet vivant de recherche scientifique dans la communauté des statistiques, qui est nourri, de façon symétrique, par son application à divers types de modèles dont ceux de climat (?) *[j'ai pas trouvé xu2023 : tu peux me donner le titre ?]* en profitant de l'expertise de ces modèles existant dans notre communauté. Ce chapitre n'est pas seulement une introduction à une méthode bien établie. Il est également une brique dans le processus en cours d'acculturation entre les deux communautés scientifiques.

III.2.1.1 Quelques mots sur le cadre mathématique

La méthode d'HMIR est une méthode d'inférence statistique. En ce sens, elle vise à contraindre les valeurs des paramètres libres, vus comme des variables aléatoires ; en connaissant la réalisation d'un échantillon de ces paramètres, c'est-à-dire en connaissant la valeur de la distance entre simulation et références pour un nombre fini de vecteurs de paramètres libres. L'approche d'inférence statistique la plus commune est l'approche bayésienne. Elle consiste à spécifier une distribution de probabilité d'une variable aléatoire a priori, puis à estimer une distribution de probabilité a posteriori, après avoir eu connaissance d'un ensemble de réalisation. Dans le cadre du tuning d'un modèle, ces distributions de probabilité correspondent à la probabilité qu'un vecteur de paramètre conduise à la simulation d'une distance simulation - référence donnée. L'inférence bayésienne classique place les distributions de probabilité au centre de son approche. Il peut être cependant difficile de spécifier ces distributions de probabilités, notamment dans le cadre de problème complexes à grandes dimensions telle que la calibration des modèles de climats. Il n'est de plus pas toujours utile de connaître précisément la distribution de probabilité recherchée : parfois, la connaissance des espérances et des variances et covariances suffisent. Dans ces situations, le cadre probabiliste développé par le statisticien de Finetti dans les années 1930 (de Finetti, 2017), désigné aussi sous le nom de statistiques bayésiennes linéaires, peut être utile. Il propose une approche qui permet d'ajuster les espérances, variances et covariances à partir d'un échantillon connu, sans avoir à spécifier la distribution de probabilité complète de la variable aléatoire associée. C'est pourquoi, aucune distribution de probabilité ne sera spécifiée dans la suite de ce chapitre. Seul les espérances, variances et covariances sont traitées par la méthode d'HMIR. Pour plus de détail sur les statistiques bayésiennes linéaires, on pourra se référer à Goldstein (1986); Goldstein and Wooff (2007) ou encore à Whittle (2000). Cette approche peut être vue comme un cas limite de l'inférence statistique classique.

III.2.1.2 Philosophie de la méthode

La méthode d'History Matching with Iterative Refocusing (HMIR) vise à éliminer les régions de l'espace des paramètres libres ne permettant pas de simuler correctement un certain nombre

de métriques scalaires définies par l'utilisateur. Le résultat final est alors un espace de paramètres libres conduisant à la simulation correcte ou acceptable de ces métriques. La méthode procède par itérations, en éliminant progressivement les parties de l'espace qui conduisent à des simulations considérées comme incorrectes. Elle utilise des émulateurs de la famille des Processus Gaussiens afin d'apprendre le lien entre métriques et paramètres ; et d'éliminer ainsi les parties de l'espace des paramètres non valides à un coût de calcul acceptable. C'est notamment leur capacité à fournir une quantification de l'incertitude associée à leur prédiction qui a motivé ce choix d'émulateur. La méthode d'HMIR peut être vue comme une manière d'identifier rapidement les régions de l'espace des paramètres qui conduisent à une trop mauvaise simulation du climat telle que définie par les métriques, afin de réserver les ressources de calcul disponibles à l'exploration de l'espace le plus plausible.

En toute rigueur, la simulation d'une métrique par le modèle est considérée comme correcte si sa différence avec la référence est inférieure à l'incertitude de comparaison simulation-référence multipliée par un seuil k , souvent pris égal à trois (règle des trois sigmas). Cette incertitude comprend des contributions de différentes natures : l'incertitude de la référence, l'incertitude liée à l'incertitude des conditions initiales de la simulation (appelée aussi variabilité interne) et de ses forçages, et enfin l'incertitude structurelle du modèle. Ces trois sources d'incertitudes sont indépendantes. L'incertitude paramétrique n'est pas incluse dans cette mesure de l'incertitude simulation-référence car elle est explorée et quantifiée par la méthode d'HMIR. En pratique, il est rare que l'ensemble des incertitudes puisse être quantifié. L'incertitude structurelle du modèle est en particulier très rarement connue et n'est même pas fermement définie. Elle peut être vue comme l'incertitude inhérente au fait de modéliser un système : erreur due au traitement numérique des équations, aux approximations faites pour établir ces équations, à la limite actuelle de la connaissance, au coût de calcul qui conduit à ne pas modéliser certains processus, ou encore aux erreurs dans la formulation du modèle. La quantification des autres sources d'incertitudes peut également s'avérer très difficile pour diverses raisons, et même l'incertitude observationnelle est souvent très mal connue. C'est le cas par exemple des produits satellites qui fournissent des observations du rayonnement terrestre, très souvent utilisés comme référence pour la calibration des GCMs. Toutes ces considérations poussent à introduire la notion de tolérance à l'erreur, comme l'erreur que l'on tolère sur chacune des métriques, en s'assurant qu'elle soit supérieure aux incertitudes que l'on aura su quantifier. La simulation d'une métrique sera alors considérée comme correcte si sa différence avec la référence est inférieure à cette tolérance à l'erreur multipliée par le seuil k .

Tous les choix effectués dans la méthode sont très conservatifs, c'est-à-dire que toutes les précautions sont prises pour ne pas éliminer des parties de l'espace des paramètres libres qui pourraient être bonnes. De plus, la méthode est construite de manière à ce que le surajustement (ou la compensation d'erreurs) est évité si l'incertitude de comparaison entre simulation et références est entièrement connue, ce qui est en pratique rarement le cas. Cependant, garantir que du surajustement ne peut pas être introduit par la même méthode en elle-même, mais uniquement par la manière dont la méthode est utilisée (c'est-à-dire par le choix des métriques et de leur tolérance à l'erreur) est une grande force de la méthode, qui est exploitée dans la suite de cette thèse pour étudier les éventuelles compensations d'erreurs dans LMDZ (voire Chapitre VI).

III.2.2 Définition du problème

Les vecteurs seront notés en gras, les espaces et les échantillons seront notés en capitale, et l'indice de l'itération sera noté en exposant. La Table III.1 récapitule les notations utilisées ici.

On a un modèle avec un nombre donné de paramètres libres incertains, notés λ^i . On souhaite trouver les vecteurs de paramètres libres $\boldsymbol{\lambda} = (\lambda^1, \dots, \lambda^N)$ permettant au modèle de

Variable	Type	Nom de la variable
N	entier	Nombre de paramètres libres que l'on souhaite ajuster
M	entier	Nombre de métriques
w_i	entier	Numéro de l'itération
λ	scalaire	Un paramètre libre quelconque du modèle
λ_i	scalaire	Le i^e paramètre libre que l'on souhaite ajuster
$\lambda_{min}^i, \lambda_{max}^i$	scalaire	Bornes inférieure et supérieure d'exploration du paramètre λ^i
$f_i(\boldsymbol{\lambda})$	scalaire	Valeur de la métrique pour une simulation effectuée avec le vecteur de paramètre $\boldsymbol{\lambda}$
r_i	scalaire	Valeur de la référence associée à la i^e métrique f_i
T_i	scalaire	Tolérance à l'erreur associée à la i^e métrique f_i
$\mu_i(\boldsymbol{\lambda}_j)$	scalaire	Prédiction de l'émulateur de la i^e métrique pour le vecteur de paramètre $\boldsymbol{\lambda}_j$
$\sigma_i(\boldsymbol{\lambda}_j)$	scalaire	Incertitude autour de la prédiction $\mu_i(\boldsymbol{\lambda}_j)$ de l'émulateur de la i^e métrique
$I(\boldsymbol{\lambda}_j)$	scalaire	Invraisemblance du vecteur $\boldsymbol{\lambda}_j$, déterminée par l'Equation III.5
$\boldsymbol{\lambda}$	vecteur	Un vecteur de paramètres libre du modèle, de composante $\lambda^1, \dots, \lambda^N$
Λ	ensemble	Hypercube initial des paramètres libres, de dimension N et égal à $[\lambda_{min}^1, \lambda_{max}^1] \times [\lambda_{min}^2, \lambda_{max}^2] \times \dots \times [\lambda_{min}^N, \lambda_{max}^N]$
\mathcal{A}	ensemble	Sous-ensemble de valeurs de paramètres libres conduisant à la simulation correcte des M métriques à leur tolérance près, tel que défini par l'Equation III.1
$\Lambda_{10N}^{w_i}$	échantillon	Echantillon de taille $10N$ de vecteurs de paramètres libres (N composantes), $\Lambda_{10N}^{w_i} = \{\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_{10N}\}$
F_i	échantillon	Ensemble des valeurs de la i^e métrique simulée pour les paramètres libres de l'échantillon $\Lambda_{10N}^{w_i}$. $F_i = \{f_i(\boldsymbol{\lambda}_1), \dots, f_i(\boldsymbol{\lambda}_{10N})\}$

TABLE III.1 – Variables de la partie 2 du chapitre htexplo [est-ce que je met ce tableau ici ou en annexe ?]

simuler correctement un certain nombre de caractéristiques traduites en métriques scalaires (la température moyenne à la surface du globe pour un GCM par exemple). On doit donc choisir a priori les N paramètres libres que l'on souhaite ajuster, ainsi que la gamme de valeurs plausibles $\lambda_{min}^i, \lambda_{max}^i$ pour chacun des paramètres. Cela définit un espace des paramètres libres sous forme d'un hypercube de dimension N le nombre de paramètres libres, noté $\Lambda = [\lambda_{min}^1, \lambda_{max}^1] \times [\lambda_{min}^2, \lambda_{max}^2] \times \dots \times [\lambda_{min}^N, \lambda_{max}^N]$

On définit ensuite les M métriques scalaires f_i que l'on souhaite que le modèle simule correctement. Ces métriques sont associées à des valeurs des références r_i qui peuvent être issues d'observations par exemple. A chaque métrique est associée une tolérance à l'erreur T_i . On considère que le modèle simule correctement une métrique si sa différence avec la référence est inférieure à un seuil k multiplié par la tolérance à l'erreur. Ce seuil est généralement pris égal à trois, conformément à la règle des trois sigmas (Pukelsheim, 1994), qui stipule que pour toute distribution de probabilité continue, unimodale et à variance finie, au moins 95% des valeurs de la distribution se situent dans un intervalle de plus ou moins trois écarts-types de

la moyenne de la distribution (aussi connue sous le nom d'inégalité de Vysochanskij–Petunin). Ainsi, si la tolérance à l'erreur correspond à l'incertitude de comparaison entre simulation et référence pour chacune des métriques, et en supposant que les distributions des écarts entre simulation et référence sont unimodale, au plus 5% des vecteurs de paramètres libres seront considérés comme faux alors qu'ils sont bons en réalité.

L'objectif de la méthode d'HMIR est de déterminer le sous-espace des paramètres libres \mathcal{A} tel que l'erreur sur chacune des métriques f_i soit inférieure à k fois la tolérance à l'erreur T_i . L'espace des paramètres libres \mathcal{A} est ainsi défini comme suit :

$$\begin{aligned} \mathcal{A} &= \{\boldsymbol{\lambda} \in \Lambda \text{ tel que } \forall i \in \llbracket 1; M \rrbracket, |f_i(\boldsymbol{\lambda}) - r_i| < kT_i\} \\ \Leftrightarrow \mathcal{A} &= \left\{ \boldsymbol{\lambda} \in \Lambda \text{ tel que } \max \left(\frac{|f_1(\boldsymbol{\lambda}) - r_1|}{T_1}; \dots; \frac{|f_M(\boldsymbol{\lambda}) - r_M|}{T_M} \right) < k \right\} \end{aligned} \quad (\text{III.1})$$

On note que cette définition n'autorise pas les compensations d'erreur entre les métriques. Il suffit en effet que le vecteur de paramètre $\boldsymbol{\lambda} \in \Lambda$ conduise à simuler une des métriques trop éloignée de sa référence pour l'exclure de l'espace des paramètres libres recherché \mathcal{A} , même si toutes les autres métriques ont des valeurs très proches de la référence. Ceci différencie assez fondamentalement cette méthode d'une méthode de calibration par optimisation où l'on chercherait, avec les mêmes notations, le minimum d'une fonction coût de type :

$$f(\boldsymbol{\lambda}) = \sum_{i=1}^N \frac{(f_i(\boldsymbol{\lambda}) - r_i)^2}{T_i^2} \quad (\text{III.2})$$

L'espace \mathcal{A} défini par l'Equation III.1 peut aussi être vu comme l'intersection des espaces de paramètres libres permettant de simuler correctement chaque métrique (Figure III.1), et s'écrit alors :

$$\mathcal{A} = \bigcap_{i \in \llbracket 1; M \rrbracket} \mathcal{A}^{f_i}; \text{ avec } \forall i \in \llbracket 1; M \rrbracket, \mathcal{A}^{f_i} = \left\{ \boldsymbol{\lambda} \in \Lambda \text{ tel que } \frac{|f_i(\boldsymbol{\lambda}) - r_i|}{T_i} < k \right\} \quad (\text{III.3})$$

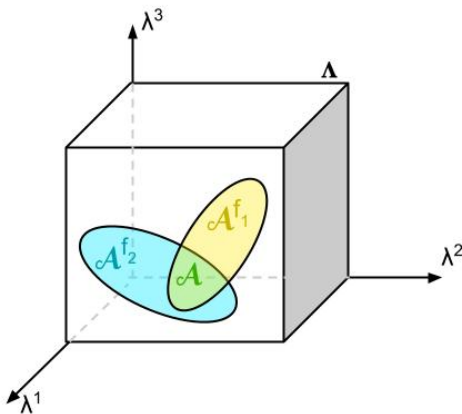


FIGURE III.1 – Schéma des espaces de paramètres libres à $N=3$ dimensions pour une expérience avec deux métriques. L'espace initial des paramètres libres est un cube $\Lambda = [\lambda_{min}^1, \lambda_{max}^1] \times [\lambda_{min}^2, \lambda_{max}^2] \times [\lambda_{min}^3, \lambda_{max}^3]$. L'espace des paramètres libres acceptables \mathcal{A} est tracé en vert. Il peut être défini comme l'intersection de l'espace des paramètres acceptables de la première métrique, \mathcal{A}^f en jaune, avec l'espace des paramètres libres acceptables de la deuxième métrique, \mathcal{A}^f en bleu. Les espaces \mathcal{A}^f et leur intersection \mathcal{A} sont de topologie quelconque.

Les espaces \mathcal{A} et \mathcal{A}^{f_i} sont de topologie quelconque, et peuvent être discontinus. Si aucun des vecteurs de paramètres libres de l'hypercube Λ ne peut remplir les contraintes à leurs tolérance à l'erreur près, l'espace final \mathcal{A} sera vide. Si au contraire ils les remplissent tous, l'espace \mathcal{A} sera égal à l'hypercube Λ . Le choix des paramètres libres que l'on souhaite tuner et leurs bornes d'exploration sont des choix importants dans la mise en place de l'expérience, qui conditionnent

son résultat. En effet, l'espace des paramètres compatible avec les références n'est cherché que dans l'hypercube de départ.

Finalement, concevoir une expérience de tuning avec la méthode HMIR consiste à définir les paramètres que l'on souhaite ajuster avec leur plage d'exploration ainsi que les métriques que l'on souhaite que le modèle simule correctement avec leurs valeurs de référence et leurs tolérances à l'erreur. La méthode d'HMIR consiste alors à déterminer l'espace \mathcal{A} correspondant à cette expérience.

III.2.3 Les étapes de la méthode HMIR

La méthode d'HMIR est divisée en cinq étapes. L'évolution des espaces de paramètres libres au cours des différentes étapes est illustrée Figure III.2.

1. Echantillonnage de l'hypercube

L'outil commence par échantillonner l'hypercube des paramètres Λ de manière uniforme. Cet échantillonnage est fait avec une méthode d'échantillonnage par hypercube latin, méthode classique d'échantillonnage des hypercubes. Cette méthode permet d'assurer que l'échantillonnage est bien réparti dans tout l'hypercube. La taille de l'échantillon est généralement de 10 fois le nombre de paramètres. Cet échantillon est noté $\Lambda_{10N}^1 = \{\lambda_1, \dots, \lambda_{10N}\}$ et est schématisé en trois dimensions Figure III.2 étape 1.1.

2. Simulations et calcul des métriques

Pour chacun des $10N$ vecteurs de paramètres de l'échantillon Λ_{10N}^1 , on fait tourner le modèle et on calcule les M métriques sur les simulations correspondantes. On aboutit ainsi à des ensembles $F_i^1 = \{f_i(\lambda) \text{ tel que } \lambda \in \Lambda_{10N}^1\}$.

3. Construction des émulateurs

Un émulateur est un modèle statistique d'apprentissage automatique, conçu pour imiter le comportement de modèles complexes avec un temps de calcul très court. Les émulateurs utilisés dans cette méthode sont des modèles statistiques des métriques, permettant de prédire la valeur des métriques sur tout l'espace des paramètres libres Λ . Un émulateur est ainsi construit pour chacune des M métriques, en utilisant les ensembles F_i^1 comme ensembles d'apprentissage. Ils jouent en quelques sortes le rôle d'interpolateurs entre les points de l'ensemble d'apprentissage F_i^1 . Il existe une grande diversité d'émulateurs dans la littérature, et ceux utilisés dans cette méthode sont des processus gaussien, qui ont la particularité de fournir une incertitude autour de leur prédiction. De plus, ils se satisfont d'un petit ensemble d'apprentissage, de l'ordre de 10 fois le nombre de dimensions de l'espace sur lequel ils portent, soit $10N$. C'est pourquoi ils sont bien adaptés pour l'ajustement de modèles coûteux en temps de calcul et de grande dimension comme les GCMs. Le cadre mathématique de l'émulation par processus gaussien et leur application à la méthode d'HMIR sont présentés dans l'Annexe B.

4. History matching

Un deuxième échantillonnage de l'espace des paramètres Λ est fait par hypercube latin, mais cette fois-ci avec une taille d'échantillon beaucoup plus grande (typiquement de 10^5 à 10^6 échantillons, Figure III.2 étape 1.2). Pour chaque vecteur de paramètres de cet échantillon, l'émulateur fournit une prédiction ou estimation de la métrique sous forme de l'espérance d'une variable aléatoire $\mu_i(\lambda)$ en même temps que son écart type qui quantifie l'incertitude associée, noté $\sigma_i(\lambda)$. On parlera parfois de "métrique émulée".

La mesure de distance utilisée pour savoir si la métrique émulée est proche ou non de la référence est une fonction d'invraisemblance notée $I(\lambda)$. Elle est calculée avec l'Equation III.5,

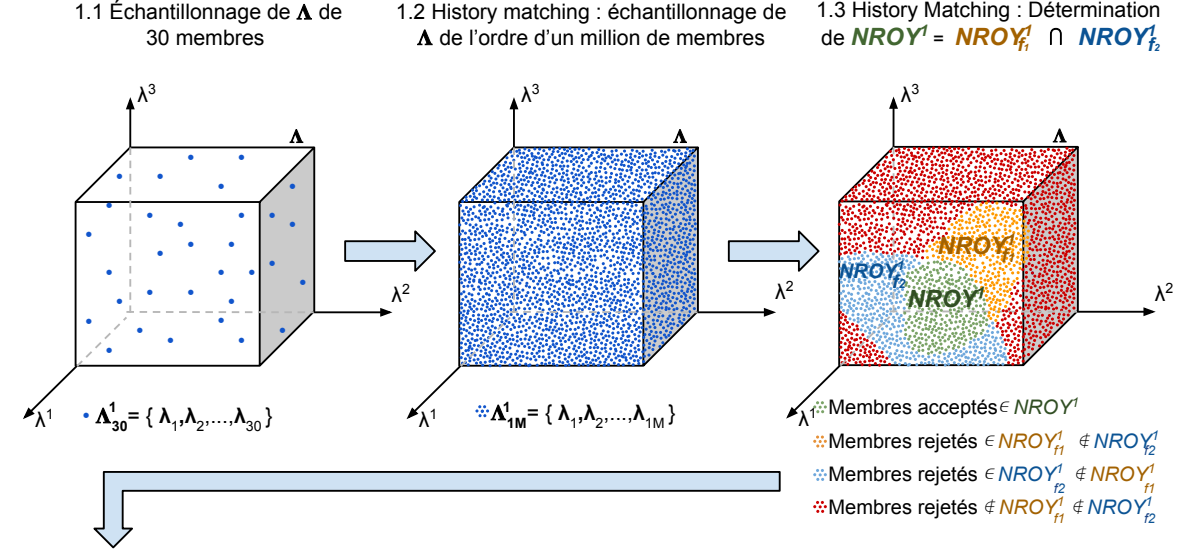
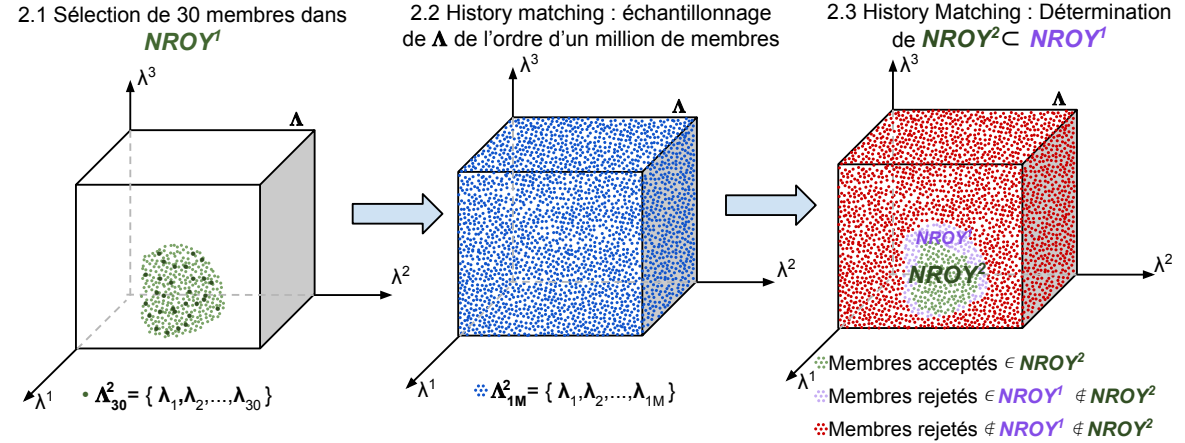
Première itération**Deuxième itération**

FIGURE III.2 – Schéma des espaces de paramètres libres pour $N=3$ au cours des deux premières itérations de la méthode d'HMIR. On suppose que l'espace des paramètres recherché \mathcal{A} est celui de la Figure III.1. (1.1) Un premier échantillonnage de l'hypercube initial des paramètres $\Lambda = [\lambda_{min}^1, \lambda_{max}^1] \times [\lambda_{min}^2, \lambda_{max}^2] \times [\lambda_{min}^3, \lambda_{max}^3]$ est effectué. Cet échantillon est noté $\Lambda_{30}^1 = \{\lambda_1, \dots, \lambda_{30}\}$, et sert à construire les bases d'apprentissage des deux émulateurs F_1^1 pour la première métrique et F_2^1 pour la deuxième métrique. (1.2) Un deuxième échantillonnage de Λ d'environ un million de point est ensuite effectué. (1.3) Les points de cet échantillon sont ensuite triés en fonction de leur invraisemblance. Si elle est inférieure à k ils sont conservés et constituent le $NROY^1$. Sinon, ils sont rejetés. On montre ici que le $NROY^1$ est l'intersection du $NROY$ correspondant à la première et à la deuxième métrique ($NROY^1 = NROY_{f_1}^1 \cap NROY_{f_2}^1$) (2.1) 30 vecteurs de paramètres appartenant au $NROY^1$ sont sélectionnés aléatoirement pour constituer l'échantillon Λ_{30}^2 servant à construire les bases d'apprentissage des émulateurs de la deuxième itération F_1^2 et F_2^2 . On remarque que l'échantillon Λ_{30}^2 est plus dense que l'échantillon Λ_{30}^1 sur l'espace des paramètres qui s'approche de l'espace recherché \mathcal{A} , ce qui conduit à des incertitudes de prédiction des émulateurs plus faibles à la deuxième itération sur cette partie de l'espace. (2.2) Un nouvel échantillonnage de Λ d'environ un million de point est effectué. (2.3) Les points de cet échantillon sont à nouveau triés en fonction de leur invraisemblance, calculée avec les émulateurs construits à la première puis à la deuxième itération. Si ces vecteurs de paramètres conduisent à une invraisemblance supérieure à k pour une des deux itérations ils sont rejetés. Sinon, ils constituent le $NROY$ de la deuxième itération, noté $NROY^2$. Une partie des points contenus dans le $NROY^1$ est rejetée à la deuxième itération du fait de la plus faible incertitude des émulateurs de cette itération.

comme le maximum de l'in vraisemblance associée à chaque métrique $I_i(\boldsymbol{\lambda})$ (Equation III.4).

$$I_i(\boldsymbol{\lambda}) = \frac{|\mu_i(\boldsymbol{\lambda}) - r_i|}{\sqrt{\sigma_i^2(\boldsymbol{\lambda}) + T_i^2}} \quad (\text{III.4})$$

$$I(\boldsymbol{\lambda}) = \max \{I_1(\boldsymbol{\lambda}); \dots; I_M(\boldsymbol{\lambda})\} = \max \left\{ \frac{|\mu_1(\boldsymbol{\lambda}) - r_1|}{\sqrt{\sigma_1^2(\boldsymbol{\lambda}) + T_1^2}}, \dots, \frac{|\mu_M(\boldsymbol{\lambda}) - r_M|}{\sqrt{\sigma_M^2(\boldsymbol{\lambda}) + T_M^2}} \right\} \quad (\text{III.5})$$

On note que ces formules correspondent bien à la définition de notre problème (Equation III.1), en remplaçant la métrique simulée par le modèle $f_i(\boldsymbol{\lambda})$ par sa valeur prédite par l'émulateur $\mu_i(\boldsymbol{\lambda})$, et en prenant en compte l'incertitude de la prédiction de l'émulateur $\sigma_i(\boldsymbol{\lambda})$ en plus de la tolérance à l'erreur T_i .

L'in vraisemblance peut être faible soit parce que la métrique émulée est proche de la valeur de référence, soit parce que l'incertitude associée à la prédiction de l'émulateur est élevée. En revanche, une in vraisemblance élevée traduit un écart entre métrique émulée et référence supérieur à la tolérance fixée pour au moins une métrique, même en tenant compte de l'incertitude de la prédiction. On procède alors par élimination des vecteurs de paramètres in vraisemblables, dès lors que leurs in vraisemblance est supérieure au seuil k . Les prédictions issues des processus gaussiens suivent une loi normale de moyenne $\mu_i(\boldsymbol{\lambda})$ et d'écart type $\sigma_i(\boldsymbol{\lambda})$ (voir Annexe B). Avec un seuil de rejet $k = 3$, seuls 0,3 % des vecteurs de paramètres corrects sont à tort rejetés lorsque l'incertitude de la prédiction domine la tolérance; ce choix est donc particulièrement conservatif. À l'inverse, lorsque la tolérance à l'erreur est supérieure à l'incertitude de prédiction, environ 5% des vecteurs de paramètres corrects sont éliminés, en supposant que la distribution des écarts entre valeurs émülées et référence soit continue et unimodale (Pukelsheim, 1994).

On définit ainsi un espace des paramètres non encore exclus appelé NROY^1 pour *Not Ruled Out Yet space* (et 1 pour la première itération), contenant tous les vecteurs de paramètres échantillonnés dont l'in vraisemblance est en dessous d'un certain seuil fixé k (Equation III.6).

$$\text{NROY}^1 = \{\boldsymbol{\lambda} \in \Lambda \text{ tel que } I(\boldsymbol{\lambda}) < k\} \quad (\text{III.6})$$

Le NROY ainsi déterminé peut aussi être vu comme l'intersection des NROY^1 déterminés pour chaque métrique (Figure III.2 étape 1.3). C'est-à-dire qu'on détermine un NROY pour chacune des métriques f_i , noté $\text{NROY}_{f_i}^1$, en utilisant la formule de l'in vraisemblance de la métrique (Equation III.4). On a donc $\text{NROY}_{f_i}^1 = \{\boldsymbol{\lambda} \in \Lambda \text{ tel que } I_{f_i}(\boldsymbol{\lambda}) < k\}$. Le NROY de l'ensemble des métriques peut alors s'écrire :

$$\text{NROY}^1 = \bigcap_{i \in [1;M]} \text{NROY}_{f_i}^1 \quad (\text{III.7})$$

5. Itérations

Pour que le NROY corresponde à l'espace des paramètres permettant que le modèle simule des métriques proches de leur référence à leur tolérance à l'erreur près, c'est-à-dire à l'espace \mathcal{A} que l'on cherche, il faut en théorie que les incertitudes des prédictions des émulateurs soient négligeables devant les tolérances à l'erreur. L'incertitude de la prédiction de l'émulateur dépend de sa distance par rapport aux vecteurs dont on connaît la valeur des métriques. Elle devient négligeable lorsque cette distance est faible devant la distance caractéristique de décorrélation (distance à partir de laquelle la valeur de la métrique en un point $\boldsymbol{\lambda}_i$ de l'espace n'est plus significativement corrélée à la valeur de la métrique en un autre point $\boldsymbol{\lambda}_j$).

Ainsi, plus l'espace des paramètres est densément échantillonné pour construire les bases d'apprentissage F_i des émulateurs, plus les incertitudes des prédictions des émulateurs diminuent. C'est-à-dire que pour diminuer cette incertitude, on peut soit augmenter la taille de la base d'apprentissage, soit diminuer le volume de l'espace des paramètres. C'est la deuxième option qui est choisie. Le NROY obtenu à la fin de l'étape de l'history matching, NROY^1 , est inclus dans l'hypercube de départ Λ et a donc un volume plus faible que celui-ci. En échantillonnant à nouveau ce NROY^1 avec $10N$ vecteurs de paramètres pour lesquels on fait tourner le modèle, on aboutit à des ensembles d'apprentissage plus denses pour construire les émulateurs et émuler les métriques (Figure III.2 étape 2.1). On peut à nouveau effectuer l'étape d'history matching avec ces nouveaux émulateurs, et aboutir à un nouvel NROY, appelé NROY^2 , qui est inclus dans le NROY précédent. Dans ce NROY^2 , on aura enlevé une partie des vecteurs qui n'avaient pas été considérés comme invraisemblables dans le NROY précédent car l'incertitude de la prédiction de l'émulateur était trop grande. Dans l'idéal, on itère ainsi jusqu'à ce que les incertitudes des émulateurs deviennent négligeables devant les tolérances à l'erreur, pour aboutir à l'espace recherché \mathcal{A} .

Les itérations sont aussi appelées vagues de tuning. Elles comprennent toutes les étapes 2 à 4. La première itération est la seule pour laquelle on fait un premier échantillonnage de l'hypercube initial des paramètres pour créer l'échantillon Λ_{10N}^1 (étape 1). Ces itérations donnent le nom "iterative refocusing" à cette méthode de tuning car à chaque itération, on se rapproche un peu plus de l'espace des paramètres recherché \mathcal{A} , qui est toujours inclus dans les NROY des itérations successives.

Il est important de noter que l'on ne connaît pas la topologie des différents NROY. La seule manière de connaître ces espaces est de faire l'étape d'history matching. Ainsi, peu importe l'itération dans laquelle on se trouve, on commence toujours l'étape 4 en échantillonnant l'hypercube initial des paramètres Λ , car c'est le seul espace que l'on sait échantillonner. Cependant, l'émulateur construit à la n^{e} itération a été construit uniquement sur une base d'apprentissage incluse dans le NROY^{n-1} . Ses prédictions ne sont donc pas valides en dehors de cet espace. Le calcul de l'invraisemblance et le rejet des vecteurs de paramètres libres se fait en pratique en itérant sur les émulateurs des vagues précédentes (en commençant par celui construit à la première vague). A la fin de l'étape d'history matching, on conserve $10N$ vecteurs de paramètre ayant une invraisemblance inférieure au seuil k afin de pouvoir lancer l'itération suivante. Ces vecteurs sont choisis aléatoirement. Si on obtient moins de $10N$ vecteurs de paramètres ayant une invraisemblance inférieure au seuil k , on recommence l'étape d'history matching, et ce jusqu'à trouver $10N$ vecteurs de paramètres. On voit ici une des limitations de la méthode : si l'espace des paramètres recherché \mathcal{A} est très petit devant l'espace initial Λ , alors l'étape d'history matching devient trop longue et trop coûteuse en temps de calcul. Avec les moyens dont nous disposons actuellement, nous ne pouvons pas explorer des espaces dont le volume par rapport à l'hypercube initial est inférieur à 10^{-7} . *[Dire recherche en cours sur ce sujet avec TinyDancer et la nouvelle version de Daniel bientôt ? ?]*

III.3 L'outil htexplo

La méthode d'HMIR a été implémentée dans l'outil **htexplo** (High Tune EXPLOrer) lors de l'ANR HighTune. Il est depuis développé en continu par ses utilisateures.

III.3.1 Présentation générale de l'outil htexplo

L'outil **htexplo** prend en entrée :

- un fichier contenant le nom des paramètres libres que l'on souhaite ajuster, avec leurs bornes d'exploration et leur valeur par défaut (à titre indicatif uniquement)
- un fichier contenant les valeurs des références et des carrés des tolérances à l'erreur associées à chaque métrique
- le seuil k , le nombre d'itérations, le nombre de simulations que l'on veut réaliser à chaque vague (on utilise en pratique souvent $10N$)
- à chaque itération w_i (on utilise ici la lettre w pour wave en anglais), la valeur des métriques produites par le modèle pour l'échantillon $\Lambda_{10N}^{w_i}$, qui formeront les ensembles d'apprentissage de chaque émulateur $F_i^{w_i}$.

A chaque vague w_i , l'outil sort automatiquement un certain nombre d'informations :

- Les valeurs des paramètres libres échantillonnés pour lancer les simulations, $\Lambda_{10N}^{w_i}$
- Une figure appelée matrice d'invraisemblance permettant de visualiser les points retenus ou rejetés lors de l'étape d'History Matching pour des espaces de dimensions quelconques
- La fraction volumique de NROY par rapport à l'hypercube initial
- Des graphiques pour contrôler la qualité des émulateurs (test LOO pour *Leave One Out*)
- Un certain nombre de scripts de visualisation des résultats sont également partagés dans l'outil, et peuvent être utilisés au besoin

Ces sorties seront expliquées via un exemple dans la section suivante. Le fonctionnement général de l'outil est schématisé sur la Figure III.3

Lorsque les métriques de l'expérience portent sur des simulations uni-colonne, l'installation du GCM, le lancement des ensembles de simulations de chaque vague ainsi que le calcul des métriques se fait automatiquement par l'outil. Les sorties des LES utilisées comme références sont alors rapatriées automatiquement, et les références et incertitudes des références sont calculées automatiquement sur les ensembles de LES rapatriées (voir Couvreur et al. (2021) et Section III.4.1). Sinon, il faut stopper l'outil pour aller lancer les simulations avec les vecteurs de paramètres de $\Lambda_{10N}^{w_i}$, calculer les métriques sur ces simulations par la méthode de son choix puis les fournir à l'outil. L'automatisation de l'outil `htexplo` à tout type de configuration de GCMs n'a pas été fait en priorité à la fois parce que l'outil n'a pas encore été porté sur les centres de calculs sur lesquels tournent les simulations 3D et parce que le coût des simulations numériques 3D fait qu'on préfère jusque là avoir une phase de contrôle de la bonne réalisation de ces simulations. Il s'agit cependant d'un axe de développement de l'outil pour le futur.

L'outil est développé au maximum de manière orthogonale, c'est-à-dire qu'il est découpé en différentes parties ayant un objectif bien défini et étant indépendantes de la structure interne des autres parties. Leur enchaînement est assuré par un script de lancement unique. En pratique dans `htexplo`, deux scripts de lancements sont disponibles : un spécifique à l'utilisation de `htexplo` pour les GCMs, et l'autre utilisable par n'importe quel modèle. Le principe d'orthogonalité est très répandu en développement informatique. La modularité qu'il apporte facilite le débogage et le développement. Par exemple, pour introduire une nouvelle métrique sur un cas 1D, il suffira de modifier le script de calcul associé ; sans avoir besoin de connaître les autres parties du code. Cette modularité est particulièrement utile pour deux raisons.

Tout d'abord, tout ce qui concerne les émulateurs est développé par la communauté des statisticiens sous le pilotage de Daniel Williamson, et leurs codes, écrits sous `R`, `python` et `c++` sont directement téléchargés et utilisés dans `htexplo`. Ce choix technique répond à la vision de l'interdisciplinarité portée ici et exposée plus haut, dans laquelle on souhaite profiter au maximum de la recherche effectuée dans l'autre communauté sans prétention à en devenir nous même des experts. Bien que le suivi des versions successives de leur code ne soit pas automatique, la modularité de l'outil facilite ainsi la bascule d'une version des émulateurs à une autre, afin

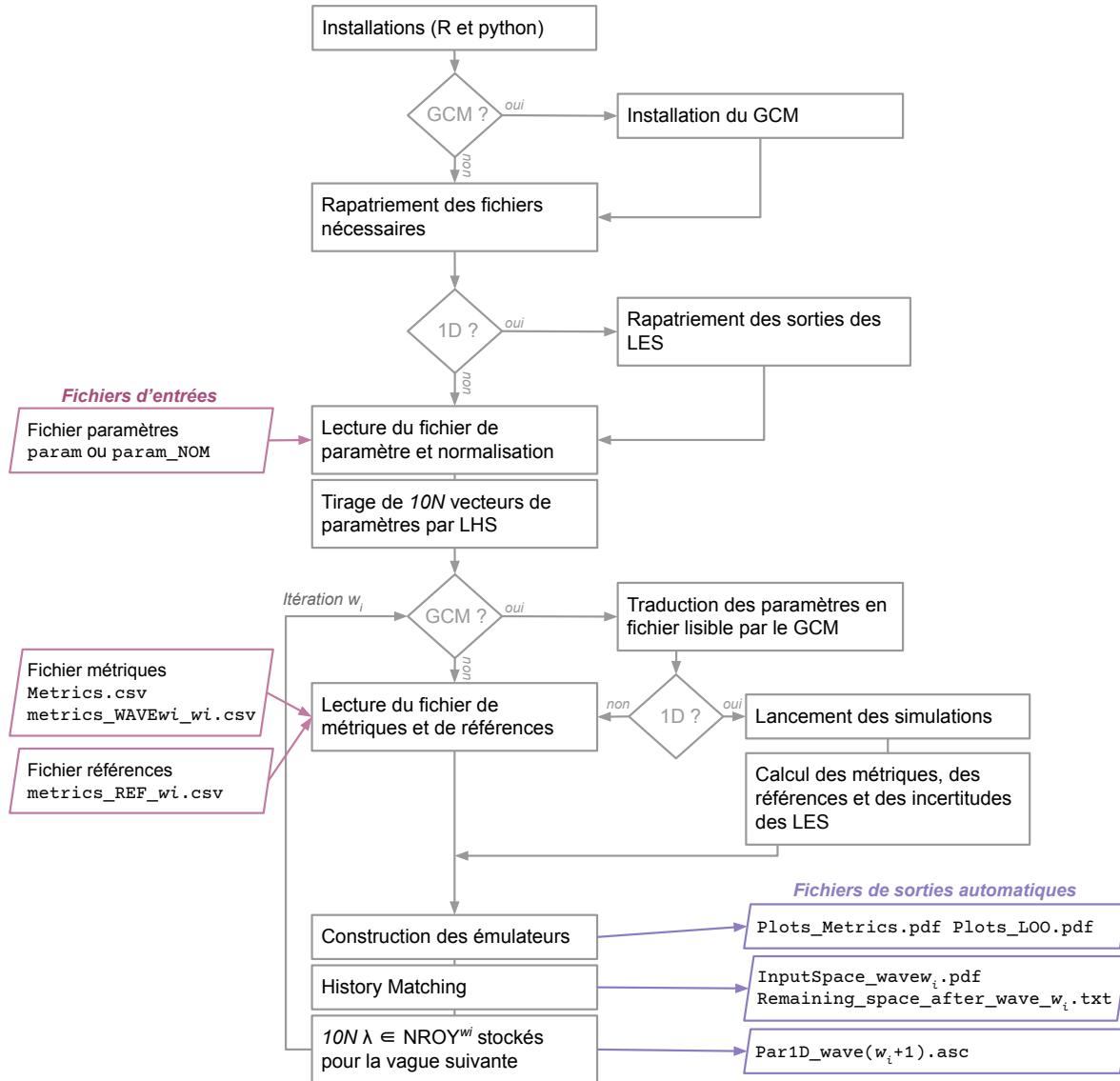


FIGURE III.3 – Logigramme de l’outil de tuning **htexplo**. Les actions effectuées par l’outil sont décrites dans les rectangles. Les actions menées en cas d’utilisation de métriques portant sur un GCM et/ou d’utilisation de métriques portant sur des simulations uni-colonnes (dites 1D dans ce schéma) sont spécifiées à partir des points de décisions représentés sous formes de losanges. Les entrées nécessaires à la réalisation d’une action sont spécifiées dans des parallélogrammes roses, et les sorties automatiques de l’outils sont spécifiées dans des parallélogrammes violet. Le traitement spécifique des métriques radiatives sur les cas uni-colonne est omis par raison de simplicité.

de profiter rapidement de leurs recherches sans être expert-es de leurs codes. La modularité est également très utile pour tout ce qui concerne l’interfaçage avec les GCMs. Chaque GCM pourra utiliser l’outil, moyennant la création des fichiers d’interface qui concernent l’installation automatique du GCM depuis `htexplo`, la traduction des vecteurs de paramètres en fichier lisible par le GCM, le lancement des simulations sur des cas 1D et le calcul des métriques. Ces derniers points sont rendus possibles par l’utilisation d’un format de fichier `netcdf` commun sur les sorties de modèles uni-colonne et de LES, standardisé par la communauté DEPHY. Ainsi, une fois que la traduction des sorties du GCM au format commun est codée, tous les scripts de calcul des métriques sur les cas 1D pourront être utilisés indifféremment du modèle qu’on souhaite calibrer.

`htexplo` est versionné sous `svn` (pour *SubVersioN*) dans une logique de développement à branche unique. Cette logique de développement, largement expérimentée et pensée pour le développement du modèle LMDZ, est adapté au développement en continu, et fonctionne bien avec des codes orthogonaux. Elle permet de bénéficier directement des développements et corrections apportés par d’autres personnes. Cette pratique, pour être efficace, doit s’accompagner de tests automatiques, dits de non régression, fait après chaque *commission* (commit en anglais), afin de vérifier qu’un bug n’a pas été introduit et que les principales fonctionnalités de l’outil n’ont pas été dégradées. L’automatisation de tels tests, opérationnels depuis une quinzaine d’année pour LMDZ, a été réalisée plus récemment pour `htexplo`, en même temps que la liste des utilisateurs s’accroissait.

Enfin, l’outil `htexplo` est distribué sous licence GPL v3 (*General Public License version 3*). Cette licence encadre le développement de logiciel open source, dont le code source est accessible gratuitement à tou-te-s, et qui peut être librement adapté et redistribué avec modification, à la seule condition que la redistribution reste sous licence GPL v3. L’outil est disponible avec la commande `svn checkout http://svn.lmd.jussieu.fr/HighTune/trunk HighTune`. Seul-es les utilisateur-es autorisé-es peuvent commettre (commit) leurs modifications sur la branche du projet hébergé à cette adresse.

III.3.2 Exemple d’application avec un modèle simple

Afin d’illustrer le fonctionnement de la méthode d’HMIR dans `htexplo`, un modèle simple du bilan radiatif de la Terre est construit ici. Ce modèle met en équation le rayonnement solaire et infrarouge global, en moyenne climatique, au sommet de l’atmosphère (TOA pour *Top Of Atmosphere*), et comprend un certain nombre de paramètres libres.

III.3.2.1 Construction du modèle simple

Le bilan radiatif au sommet de l’atmosphère dans l’infrarouge est égal au rayonnement infrarouge sortant au sommet de l’atmosphère, noté par la suite OLR pour *Outgoing Longwave Radiation*. Il peut être décomposé en une partie ciel clair, OLR_{clr} , et une partie nuageuse $CRELW_{TOA}$. On suppose dans ce modèle que la partie ciel clair est un paramètre libre que l’on notera **b** dans la suite de l’exemple. La partie nuageuse est modélisée par la somme d’une fonction notée f , et de la valeur observée $O_{crelw} = -30 \text{ W.m}^{-2}$ (Loeb et al., 2009). La formulation de cette fonction s’inspire d’une figure tirée de Hourdin et al. (2017) inter-comparant, dans plusieurs modèles de climat, la dépendance de l’OLR à la valeur d’un paramètre libre **FALLV**, en facteur de la vitesse de chute des cristaux de glace nuageux défini dans LMDZ par l’Equation II.37. La figure est reproduite ici Figure III.4. Cette dépendance croissante de l’OLR à **FALLV** peut se comprendre par le fait qu’une augmentation de **FALLV** conduit à une chute plus rapide des cristaux de glace, donc à des nuages hauts moins épais et moins persistants. Les cirrus, qui constituent une partie importante des nuages de glace sur Terre, piègent le rayonnement

infrarouge émis par les basses couches de l'atmosphère plus chaude donc plus énergétiques, et émettent à leur tour un rayonnement infrarouge vers l'espace moins énergétique car ils sont très froids. Leur effet radiatif dans l'infrarouge est donc réchauffant pour la planète. Ainsi, des cirrus moins épais ou moins persistants, dus à des grandes valeurs de **FALLV**, conduisent à un OLR plus important, via un effet radiatif global des nuages dans l'infrarouge moins réchauffant. Par souci de simplicité, **FALLV** sera noté **a** dans cet exemple. La fonction est choisie égale à $f(\mathbf{a}) = 10 \log(2\mathbf{a} - 1)$, de manière à ce que $f(1)=0$. On a alors l'équation suivante pour l'OLR :

$$OLR = OLR_{clr} + CRELW_{TOA} = \mathbf{b} + O_{crelw} + f(\mathbf{a}) \quad (\text{III.8})$$

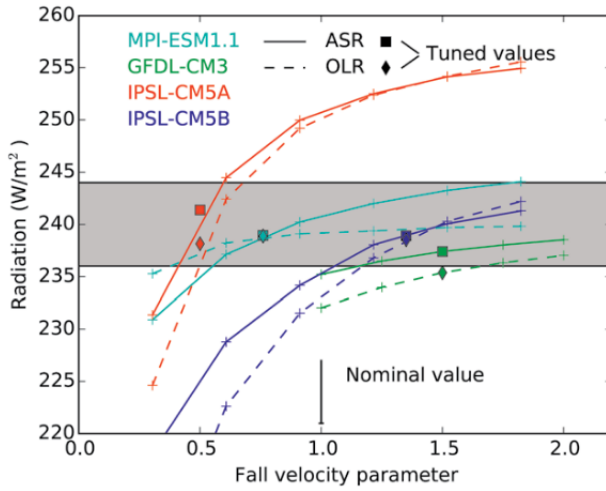


FIGURE III.4 – Flux radiatif au sommet de l'atmosphère (W.m^{-2}) en fonction du paramètre libre **FALLV** pour quatre modèles. Les courbes en trait plein représentent le rayonnement solaire absorbé par la Terre (ASR), et les courbes en pointillé le rayonnement infrarouge émis par la Terre (OLR), tous les deux comptés positivement. La partie grisée représente la valeur de références de ces grandeurs avec leur incertitudes ($240 \pm 4 \text{ W.m}^{-2}$). Les symboles carrés et losanges indiquent les valeurs choisies pour la configuration finale du GCM, après tuning. Cette figure est directement issue de Hourdin et al. (2017).

Le bilan radiatif global dans le solaire est égal au rayonnement solaire reçu au sommet de l'atmosphère moins celui qui est renvoyé vers l'espace. Cette grandeur correspond au rayonnement solaire absorbé par la Terre, noté ASR pour *Absorbed Solar Radiation*. Le rayonnement solaire reçu en moyenne au sommet de l'atmosphère est noté $S_0/4$, avec S_0 la constante solaire, égale à 1368 W.m^{-2} . Le rayonnement solaire renvoyé par la Terre vers l'espace peut être décomposé en une partie ciel-clair (ASR_{clr}) et une partie nuageuse ($CRESW_{TOA}$). La partie ciel-clair est modélisé par un albédo, noté **c**, de telle sorte que $ASR_{clr} = \mathbf{c} \times S_0/4$. Hourdin et al. (2017) montre également que l'ASR dépend de manière logarithmique à **FALLV** (Figure III.4). En effet, des valeurs plus grandes de **FALLV** conduisent à des cirrus moins couvrants et moins persistants donc à moins de réflexion du rayonnement solaire vers l'espace par ces nuages. L'effet radiatif global des nuages dans le solaire est modélisé comme étant une fraction **d** de leur effet radiatif dans l'infrarouge. Finalement, l'ASR est modélisé par :

$$\begin{aligned} ASR &= S_0/4 - SW_{TOA}^\uparrow = S_0/4 - (SW_{clr}^\uparrow + CRESW_{TOA}) \\ &= S_0/4 - (\mathbf{c} \times S_0/4 - \mathbf{d} \times CRELW_{TOA}) \\ &= (1 - \mathbf{c})S_0/4 + \mathbf{d} \times (O_{crelw} + f(\mathbf{a})) \end{aligned} \quad (\text{III.9})$$

III.3.2.2 Première expérience de tuning : une métrique et un paramètre

On commence par une expérience très simple où l'on souhaite ajuster les paramètres libres **a** et **c** du modèle conceptuel présenté à la section précédente.

On suppose, compte tenu des connaissances que l'on ait, que : $\mathbf{a} \in [0, 5; 20]$ et $\mathbf{c} \in [0, 1; 0, 3]$. L'hypercube de départ est donc le rectangle $\Lambda = [0, 5; 20] \times [0, 1; 0, 3]$. On souhaite ajuster ces

paramètres de sorte que l'OLR simulé par le modèle soit compatible avec la valeur de l'OLR observée $r_{OLR} = 240 \text{ W.m}^{-2}$. On suppose que ces observations sont connues à $\pm 2 \text{ W.m}^{-2}$ près. L'incertitude de l'observation de l'OLR est en réalité très difficile à estimer, et le produit d'observations utilisé en pratique pour le tuning de LMDZ fournit une incertitude un peu plus grande (voir Section III.4.2). On suppose de plus que la variabilité interne du modèle est connue et est de l'ordre de 1 W.m^{-2} . On prend alors une tolérance à l'erreur $T_{OLR} = \sqrt{2^2 + 1^2} = \sqrt{5} \simeq 2,24$. Dans ce premier exemple, on a une seule métrique, notée par la suite f_{OLR} , qui est modélisée comme la somme de l'OLR calculée avec l'Equation III.8 et d'un bruit blanc uniforme compris entre -1 et 1. Ce bruit est rajouté afin d'imiter la variabilité interne du système climatique. La métrique f_{OLR} ne dépend que du paramètre libre **a**. Ce choix est fait pour illustrer dans un premier temps de manière très simple la méthode d'HMIR, et notamment le fonctionnement des émulateurs, sans avoir à réfléchir dans des espaces à plusieurs dimensions.

Dans cet exemple très simple, la métrique est directement une sortie du modèle, lui-même fonction analytique du paramètre libre. Quand on effectue un tuning du modèle de climat, les moyennes climatiques du rayonnement au sommet de l'atmosphère sont des propriétés émergentes du climat terrestre, qui résultent d'un ensemble complexe de processus et de leur couplage. Les métriques portant sur ces variables et utilisées pour le tuning d'un GCM intègrent donc une grande partie des équations du GCM. Les dépendances de ces métriques à certains paramètres libres, comme **FALLV**, ne sont donc pas directes, contrairement au modèle conceptuel de cette partie.

Dans cet exemple simple, on peut déterminer l'espace recherché théoriquement, en résolvant l'équation suivante. On choisi un seuil de rejet $k = 3$.

$$\begin{aligned} OLR &= r_{OLR} \pm k T_{OLR} \\ \Leftrightarrow OLR_{clr} + O_{crelw} + 10 \log(2\mathbf{a} - 1) &= r_{OLR} \pm k T_{OLR} \\ \Leftrightarrow \mathbf{a} &= \frac{1}{2} \left[\exp \left(\frac{r_{OLR} - OLR_{clr} - O_{crelw} \pm k T_{OLR}}{10} \right) + 1 \right] \\ &\simeq 0,81 < \mathbf{a} < 1,69 \end{aligned}$$

On a donc :

$$\mathcal{A} = \left\{ (\mathbf{a}, \mathbf{c}) \in \Lambda \text{ tel que } \frac{|f_{OLR}((\mathbf{a}, \mathbf{c})) - r_{OLR}|}{T_{OLR}} < k \right\} = [0,81; 1,69] \times [0,1; 0,3] \quad (\text{III.10})$$

Trois vagues de tuning sont exécutées avec ce protocole expérimental et une base d'apprentissage contenant 10 points à chaque vague.

La Figure III.5 (a) montre les bases d'apprentissage des émulateurs pour chaque vague de l'expérience dans l'espace des paramètres Λ à deux dimensions. Ces ensembles contiennent 10 points chacun et sont notés $\Lambda_{10}^1 \in \Lambda$, $\Lambda_{10}^2 \in NROY^1$ et $\Lambda_{20}^3 \in NROY^2$. Au fur et à mesure des itérations, l'espace rempli par les vecteurs de paramètres de ces ensembles se rapproche de l'étape d'history matching (50 000), et leur statut rejeté ou accepté lors de cette étape ; pour les trois vagues de l'expérience. C'est une manière de visualiser les différents NROY, qui ne sont connus que par échantillonnage. On voit ainsi que le $NROY^3$ correspond à l'espace des paramètres libres recherché \mathcal{A} . On s'attend donc à ce que les points échantillonnés pour construire les bases d'apprentissage des émulateurs de la vague 4 soient tous dans \mathcal{A} , avec donc $f_{OLR} = r_{OLR} \pm 3T_{OLR}$. On note également sur cette figure que $NROY^1$ (qui correspond aux points non rejetés à la vague 1) est inclus dans $NROY^2$ (qui correspond aux points non rejetés à la vague 2). Et que $NROY^2$ est lui même inclus dans le $NROY^3$. Le $NROY^2$ est ici un espace disjoint. La méthode d'HMIR n'a en effet aucun mal à détecter ce genre d'espace, puisqu'elle ne fait aucune hypothèse sur la topologie des NROYS.

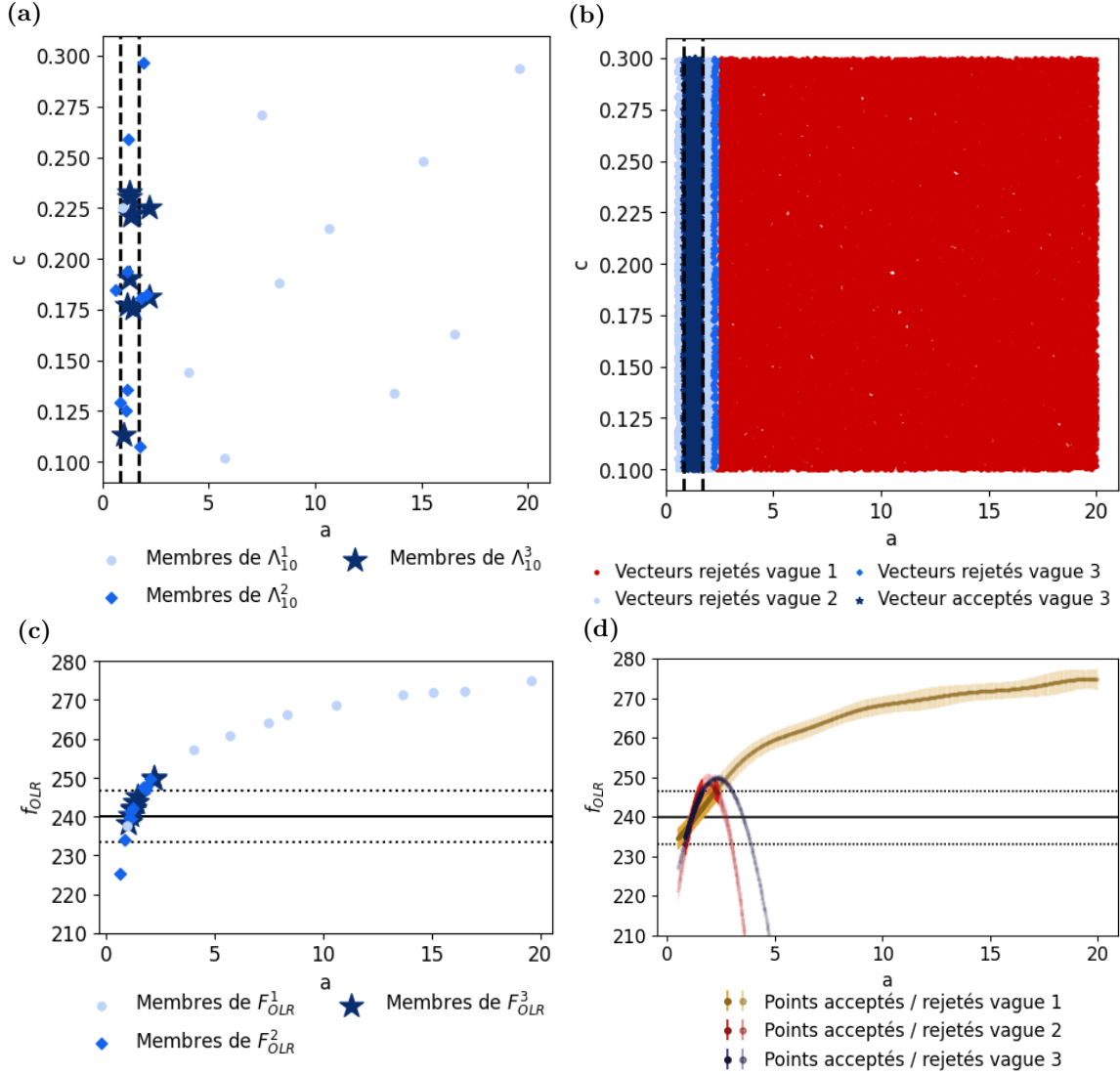


FIGURE III.5 – (a) et (b) Evolution des espaces des paramètres libres à deux dimensions au cours des trois vagues de l'expérience de tuning. L'espace recherché est délimité par les deux courbes pointillées. La figure (a) montre les vecteurs de paramètres des bases d'apprentissage des émulateurs ($\Lambda_{10}^1 \in \Lambda$, $\Lambda_{10}^2 \in NROY^1$ et $\Lambda_{10}^3 \in NROY^2$). (b) montre les points échantillonnés lors de l'étape d'history matching (50 000) et leur acceptation ou rejet dans les trois NROY. Au fur et à mesure des vagues, l'espace des points acceptés se rapproche de l'espace recherché. Les figures (c) et (d) montrent la métrique f en fonction du paramètre a au cours des trois vagues, avec en noir $r_{OLR} \pm 3T_{OLR}$. La figure (c) montre les points des bases d'apprentissage, c'est-à-dire la valeur de la métrique pour les vecteurs de paramètres montrés sur la figure (a). Au fur et à mesure des vagues, les points sont échantillonnés de plus en plus proche de $r_{OLR} \pm 3T_{OLR}$. Enfin, la figure (d) montre les prédictions des émulateurs de f_{OLR} des trois vagues de tuning, et l'incertitude de leurs prédictions, sur 1000 points tirés au hasard dans Λ . En clair, les points sont rejetés, et en foncés les points sont acceptés. Cette figure illustre la décroissance de l'incertitude des prédictions des émulateurs au cours des vagues. On note également l'aspect disjoint du $NROY^2$, également visible sur la figure (b). On remarque enfin que les prédictions des émulateurs des vagues 2 et 3 sont très loin des valeurs de f_{OLR} au dessus d'une certaine valeur de a , et sont rejetés bien que dans l'intervalle $r_{OLR} \pm 3T_{OLR}$. Cela s'explique par le fait que les points au dessus de $a \simeq 2,5$ sont rejetés par l'émulateur de la première vague. Cette partie de l'espace n'est donc pas échantillonnée pour construire les émulateurs des vagues 2 et 3. Les prédictions des émulateurs dans ces régions n'ont donc aucun sens malgré leur faible incertitude, et les points de ces régions sont rejetés par l'émulateur de la vague 1 pour toutes les vagues suivantes.

Les valeurs de la métrique f_{OLR} pour les points des bases d'apprentissage Λ_{10}^1 , Λ_{10}^2 et Λ_{20}^3 , c'est-à-dire les ensembles F_{OLR}^1 , F_{OLR}^2 et F_{OLR}^3 , sont tracés Figure III.5 (c). Ils sont de plus en plus proche de $r_{OLR} \pm 3T_{OLR}$ au fur et à mesure des vagues. Enfin, la Figure III.5 (d) montre les prédictions μ_{OLR} des émulateurs de f_{OLR} pour chaque vague ainsi que l'incertitude associée sous forme de barre d'erreur à $\mu_{OLR} \pm 3\sigma_{OLR}$. Les valeurs associées à 1000 vecteurs de paramètres, échantillonnés par méthode d'hypercube latin dans Λ , sont tracés sur cette figure. Les vecteurs rejetés à cette vague sont tracés en clair, et les vecteurs acceptés sont tracés en foncé. Cette figure illustre plusieurs choses. Tout d'abord, on voit clairement que les prédictions de l'émulateur de f_{OLR} de la première vague forment une courbe proche de la formule analytique de f_{OLR} , qui est donc bien apprise par l'émulateur. On note ensuite que l'incertitude des prédictions des émulateurs diminue au fur et à mesure des vagues. Cela s'explique par le fait que les bases d'apprentissage des émulateurs sont de plus en plus denses dans les NROY successifs. On retrouve sur cette figure le fait que le $NROY^2$, tracé Figure III.5 (b) soit disjoint. Le fait que $NROY^3$ soit indiscernable de \mathcal{A} sur la Figure III.5 (b) est dû au fait que les incertitudes des prédictions des émulateurs (Figure III.5 (d)) sont très faibles devant la tolérance à l'erreur de la métrique. L'émulateur de la troisième vague se comporte ainsi comme le modèle pour la métrique f_{OLR} et sur cette région de l'espace des paramètres.

Notons enfin que les prédictions des émulateurs des vagues 2 et 3 sont très éloignées des valeurs de f_{OLR} attendues pour \mathbf{a} supérieur à environ 2,5, et ces points sont rejetés même si leur prédiction est égale à $r_{OLR} \pm 3\sqrt{\sigma_{OLR}^2 + T_{OLR}^2}$. Les vecteurs de paramètres dont $\mathbf{a} > 2,5$ sont en fait rejetés par l'émulateur de la première vague. Ainsi, les émulateurs des vagues 2 et 3 n'ont aucun points de leur base d'apprentissage dans cette région, et ne sont donc pas pertinents malgré une faible incertitude de prédiction. C'est pour cela que, à chaque vague, les émulateurs de toutes les vagues précédentes sont utilisés pour déterminer si un vecteur de paramètres est acceptable ou non. Ainsi, pour déterminer si un vecteur de paramètre est acceptable ou non à la i ème itération, l'outil commence par interroger le premier émulateur. Si le point n'est pas rejeté par le premier émulateur, il interroge le deuxième émulateur, et ainsi de suite jusqu'à ce que le point soit rejeté par un des émulateurs. S'il n'est jamais rejeté, il appartient au NROY de la i ème vague.

En termes d'interprétation de ces résultats, imaginons que l'on aboutisse à la Figure III.5 sans avoir eu connaissance de la forme de f_{OLR} et de sa dépendance paramétrique ; ce qui est généralement le cas lorsque l'on tune un modèle de climat. On pourrait déduire de cette figure que f_{OLR} dépend uniquement du paramètre libre \mathbf{a} parmi les paramètres libres variés dans l'expérience (donc parmi \mathbf{a} et \mathbf{c}), et que cette dépendance à la forme d'une fonction logarithmique croissante de \mathbf{a} . On pourrait en conclure que, parmi les paramètres libres variés, les variations d'OLR sont dominées par les variations de **FALLV** donc par les variations de la simulations des nuages haut. Ces conclusions peuvent paraître anedoctiques pour cette expérience très simple, mais peuvent, dans des cas plus réel, fournir des informations robustes sur le comportement du modèle.

Sensibilité au nombre de points de la base d'apprentissage des émulateurs

Afin de mieux comprendre la dépendance de la construction des émulateurs au nombre de points de la base d'apprentissage, on refait la même expérience avec seulement trois points sondes (au lieu de dix dans l'expérience précédente). On change également légèrement la formulation des émulateurs utilisé, en désactivant la prise en compte de la pépite (ou *nugget* en anglais, voir Section B.2.1), afin d'avoir un comportement un peu plus simple et intuitif des émulateurs. On fait cette expérience deux fois pour voir l'influence du tirage aléatoire. La Figure III.6 montre les prédictions de l'émulateur de f_{OLR} pour la première vague de chacune des expériences, ainsi que les incertitudes des prédictions en bleu. On note la diminution de l'incertitude des prédictions

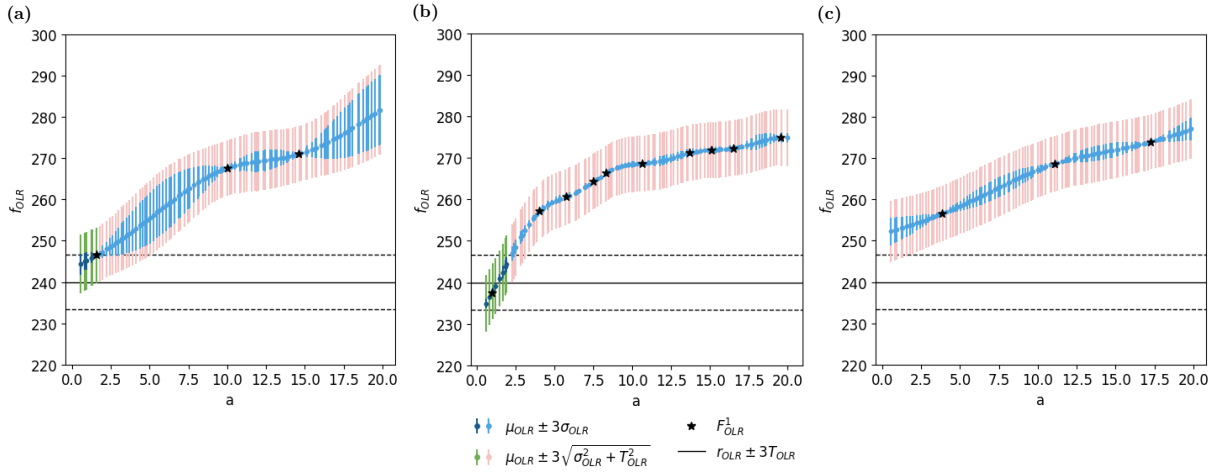


FIGURE III.6 – Prédiction de l’émulateur de f_{OLR} de la première vague, sur 100 vecteurs de paramètres échantillonnés par hypercube latin, et avec une pépité nulle dans les émulateurs : (a) et (c) exemples avec une base d’apprentissage de trois éléments (b) avec la base d’apprentissage de l’expérience précédente à dix éléments. Les incertitudes des prédictions des émulateurs multipliée par le seuil $k = 3$ sont tracées en bleu clair lorsque le point est rejeté et en bleu foncé s’il est accepté. L’incertitude totale du point multipliée par le seuil $k = 3$, c’est à dire $3\sqrt{\sigma_{OLR}^2 + T_{OLR}^2}$, est tracée en rouge clair lorsque le point est rejeté et en vert foncé s’il est accepté. Les éléments de F_{OLR}^1 sont indiqués avec des étoiles. L’incertitude de la prédiction des émulateurs est nettement plus faible lorsque l’émulateur a eu connaissance de la valeur de f_{OLR} sur dix points, par rapport à l’apprentissage fait sur trois points. Lorsque l’apprentissage est fait sur trois points, on note que l’incertitude de la prédiction est plus importante au fur et à mesure que l’on s’éloigne des points, ainsi qu’au bord du domaine, et qu’elle diminue proche des points d’apprentissage. La figure (c) montre l’aspect aléatoire du tirage des vecteurs de paramètres. Lorsque l’on tire trop peu de vecteurs de paramètres par rapport au nombre de dimensions, on peut tomber sur des jeux de paramètres qui ne permettent pas de déterminer correctement le NROY.

des émulateurs lorsque le nombre de points de la base d’apprentissage augmente. Dans les deux expériences où la base d’apprentissage ne contient que trois points, on voit nettement l’incertitude de la prédiction de l’émulateur augmenter lorsque le vecteur de paramètres est loin des points où l’on connaît la valeur de la métrique, ainsi qu’au bord du domaine. Cela reste le cas dans l’expérience à dix points d’apprentissage, mais est moins visible car les incertitudes de prédictions sont beaucoup plus petite du fait du nombre du points.

On note également que l’incertitude de la prédiction est très faible au voisinage des points d’apprentissage, et serait nulle si l’on tirait exactement le même point que le point d’apprentissage (ce qui est un événement de probabilité nulle). Lorsque l’émulateur utilise une pépité, ce qui est le cas par défaut dans `htexplo` l’incertitude de la prédiction n’est plus très faible au voisinage des points d’apprentissage. La pépité rajoute en effet une part d’incertitude irréductible aux émulateurs, qui permet d’éviter du sur-apprentissage dans leur construction. Cette pépité est un hyperparamètre, c’est-à-dire qu’elle est apprise pendant l’étape de construction de l’émulateur à partir de la base d’apprentissage. On voit sur la première vague de la Figure III.5 (d), qui diffère de la Figure III.6 (b) uniquement par la présence d’une pépité, que l’incertitude des prédictions est toujours plus grande, et en particulier au niveau des points sondes où l’incertitude de prédiction n’est jamais très faible.

La Figure III.6 (c) montre que si l’on utilise la méthode d’HMIR avec trop peu de points par rapport au nombre de paramètres que l’on souhaite ajuster, l’aspect aléatoire du tirage peut conduire à des bases d’apprentissage qui ne permettent pas de construire des émulateurs pertinents. Ainsi, dans ce cas, le NROY est considéré comme vide à la première itération alors que l’on sait que l’espace des paramètres recherché n’est pas vide.

Figures automatiques

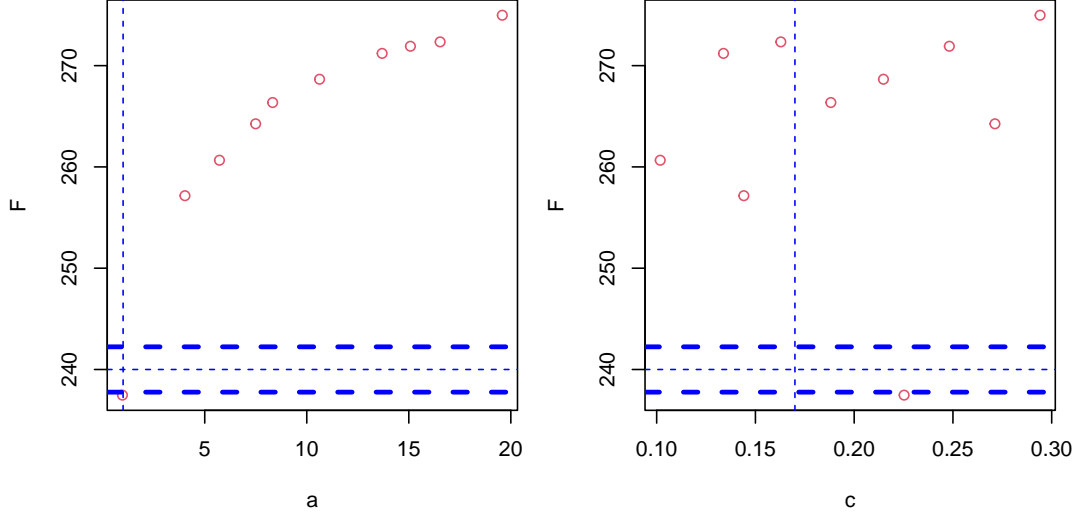


FIGURE III.7 – Figure automatique de l’outil `htexplo`, `Plots_Metrics.pdf`. Cette figure montre les ensembles d’apprentissage F^i , c’est-à-dire les valeurs de chaque métrique en fonction de chaque paramètre. Ici est tracée la métrique f_{OLR} en fonction du paramètre a à gauche et en fonction du paramètre c à droite, pour la première vague de l’expérience. Les valeurs de la référence $r_{OLR} \pm T_{OLR}$ sont tracées en pointillés sur l’horizontale. La valeur du paramètre par défaut est indiqué en pointillé sur la verticale.

L’outil `htexplo` sort un certain nombre de figures de manière automatique grâce à des scripts de post-traitement et visualisation au développement duquel j’ai contribué pendant ma thèse. Tout d’abord, à chaque vague, les valeurs des métriques calculées sur les simulations, c’est-à-dire les ensembles F^i , sont tracées en fonction des différents paramètres libres de l’expérience. Cette figure, appelée `Plots_Metrics.pdf` dans l’outil, est montrée sur la Figure III.7 pour 10 tirages à la première vague de l’exemple illustré plus haut.

L’outil sort également un test pour vérifier que les émulateurs ont bien été construits. Ce test classique pour ce type d’outil s’appelle LOO pour *Leave One Out*. C’est un test standard et rapide pour vérifier la bonne capacité de prédiction des émulateurs. Il consiste à enlever un vecteur de paramètre λ_j de la base d’apprentissage Λ_{10N}^{wi} et à réaliser une prédiction en ce point avec l’émulateur¹. Puisque les émulateurs utilisés sont des processus gaussiens, la distribution de leur prédiction suit une loi normale de moyenne $\mu(\lambda_j)$ et d’écart-type $\sigma(\lambda_j)$. L’intervalle de confiance à 95% est donc $[\mu(\lambda_j) - 2\sigma(\lambda_j); \mu(\lambda_j) + 2\sigma(\lambda_j)]$. Ainsi, si les émulateurs sont fiables, leur intervalle de confiance à 95% contient la valeur de la métrique calculée sur la simulation avec le vecteur λ_j avec une probabilité de 95%. Ce test est réalisé pour chaque élément de la base d’apprentissage Λ_{10N}^{wi} . La Figure III.8 montre le `Plots_LOO.pdf` de la première vague de l’expérience de tuning. Les valeurs des prédictions avec leurs incertitudes, $\mu_{OLR}(\lambda_j) \pm 2\sigma_{OLR}(\lambda_j)$, sont tracées pour chaque λ_j de Λ_{10}^1 . Les valeurs des métriques calculées en ces points sont également tracées sur cette figure, en vert lorsqu’elles sont dans l’intervalle de confiance à 95% et en rouge sinon.

[Fred pourquoi tu as supprimé : Les capacités prédictives des émulateurs sont considérés comme correcte si 95% des valeurs des métriques sont dans l’intervalle de leur prédiction associé, c’est-à-dire lorsque 95% des points de la figure BLABLA sont vert. ?]

Ici, il n’y a que 10 points, ce qui est un peu juste pour faire des statistiques. On peut tout de même considérer que s’il y avait moins de huit points vert (ou plus de deux points rouges),

1. A noter que les hyperparamètres des émulateurs ne sont pas réappris avec cette nouvelle base d’apprentissage contenant un point en moins, voir Section B.2.3

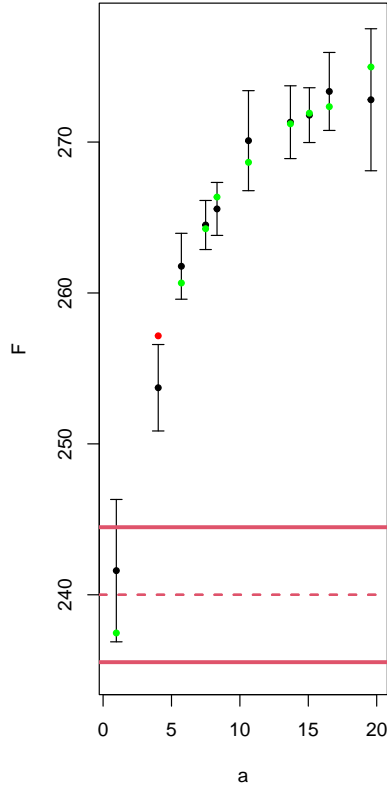


FIGURE III.8 – Résultat du test LOO (*Leave One Out*), issu de `Plots_L00.pdf` pour la première vague de l'expérience. Ici, l'outil a détecté automatiquement que la métrique $F = f_{OLR}$ ne dépend que du paramètre \mathbf{a} . C'est pourquoi il n'y a qu'un graphique. Chaque point de couleur correspond à la valeur de la métrique au point de la base d'apprentissage, tracée à gauche de la Figure III.7. Les points noirs correspondent aux prédictions faites par l'émulateur, en ayant retiré le point de la base d'apprentissage pour faire cette prédiction. Les barres d'erreur correspondent à deux fois l'incertitude autour de cette prédiction. La ligne rouge pointillée correspond à la cible, r_{OLR} , et les traits pleins rouges à deux fois la tolérance à l'erreur de part et d'autre de la cible ($2T_{OLR}$). Les points de F_{OLR}^1 sont mis en vert s'il sont dans l'intervalle de confiance à 95% de la prédiction, c'est-à-dire si $f_{OLR}(\lambda_j) \in [\mu_{OLR}(\lambda_j) - 2\sigma_{OLR}(\lambda_j); \mu_{OLR}(\lambda_j) + 2\sigma_{OLR}(\lambda_j)]$. Sinon, ils sont colorés en rouge.

la capacité prédictive de l'émulateur serait remise en question. Il est à noter que trop de points verts est également un mauvais signe. Par exemple, pour une base d'apprentissage contenant 100 points, il devrait y avoir 5 points en dehors de l'intervalle de confiance à 95%. S'il y en a trop ou trop peu, on considérera que l'émulateur a de mauvaises capacités prédictives. Pour en savoir plus sur l'émulation par processus gaussien, et le test LOO implémenté dans `htexplo` on pourra se référer à l'Annexe B.

La figure `Plots_L00.pdf` permet d'avoir une idée rapide de l'ordre de grandeur de l'incertitude des prédictions des émulateurs. La figure ne contient que les paramètres qui ont été détectés comme influençant la métrique. Un certain nombre de tests de dépendance entre métriques et paramètres sont faits pour construire les émulateurs, et si un paramètre est détecté comme n'influençant pas du tout la métrique, alors il est retiré. Cela permet de construire des émulateurs sur des espaces ne gardant que les dimensions utiles, et d'être ainsi plus efficace. C'est pour cela que le `Plots_L00.pdf` de l'expérience ne contient qu'un graphique montrant la dépendance au paramètre \mathbf{a} . Ici l'émulateur n'est construit que sur une dimension.

La troisième figure, qui est la matrice d'invraisemblance, est expliquée un peu plus loin, Section III.3.3.3.

III.3.3 Convergence de l'expérience

Le nombre d'itérations à effectuer n'est pas déterminé automatiquement par l'outil mais est fixé par l'utilisatrice. Comment savoir combien d'itérations sont nécessaires pour déterminer l'espace des paramètres libres compatibles avec les métriques? Cette question est celle de la convergence de l'expérience. Idéalement, on peut considérer que l'expérience a convergé lorsque le NROY correspond à l'espace des paramètres libres recherché **A faire** :

Compléter la phrase : j'ai pas compris

Cela se produit lorsque, pour chaque métrique, les incertitudes des prédictions des émulateurs

sont négligeables devant la tolérance à l'erreur. On appelle cette convergence la convergence stricte. Dans l'exemple précédent, la convergence est atteinte au bout de trois itérations.

Il arrive cependant fréquemment que l'incertitude des émulateurs ne devienne jamais négligeable devant la tolérance à l'erreur. On considère alors que l'expérience a convergé lorsque le NROY n'évolue plus significativement d'une vague à l'autre. Pour l'évaluer, on regarde la fraction du volume du NROY restant par rapport au volume de l'hypercube initial ; qui correspond au nombre de vecteurs de paramètres acceptés divisé par le nombre de paramètres échantillonnés lors de l'étape d'history matching. Avec `htexplo`, cette information est conservée à la fin de l'étape d'history matching dans les fichiers `Remaining_space_after_wave_wi.txt`, avec `wi` le numéro de la vague. Le script python `plot_NROY.py`, mis à disposition dans `htexplo`, pourra être utilisé pour tracer l'évolution de cette fraction au cours des vagues. On note que la fraction volumique du NROY n'évolue plus lorsque la convergence stricte est atteinte. La Figure III.9 montre l'évolution de cette fraction sur l'exemple de la section précédente pour cinq itérations. La fraction volumique du NROY n'évolue plus significativement après la troisième itération, ce qui est le signe de la convergence.

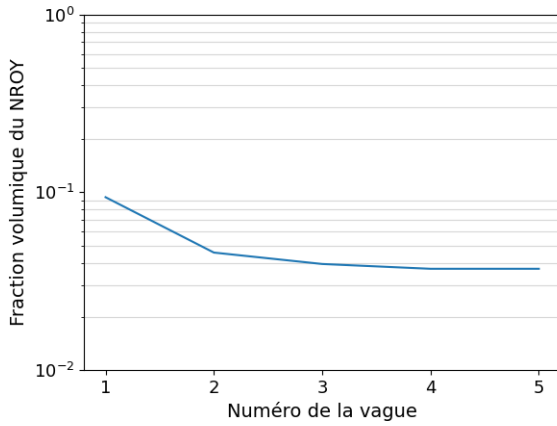


FIGURE III.9 – Evolution au cours des itérations de la fraction du volume du NROY par rapport au volume de l'hypercube initial, pour l'expérience à une métrique et deux paramètres. Après la troisième itération, cette fraction n'évolue plus significativement, on peut donc considérer que l'expérience a convergé à la troisième itération.

Cas limites

Il existe deux cas limites possibles au problème tel qu'il est posé Section III.2.2. Le cas où le NROY est vide et le cas où les NROYs successifs sont égaux à l'hypercube initial Λ .

La définition d'un NROY vide dépend en pratique du nombre de tirages que l'on peut faire à l'étape d'history matching, c'est-à-dire des capacités de calcul à notre disposition. Avec les moyens dont nous disposons et les codes actuels, nous ne pouvons pas explorer des espaces dont le volume est inférieur à une fraction 10^{-6} - 10^{-7} du volume de l'hypercube initial. En dessous de cette limite, le NROY sera considéré comme vide, bien qu'il soit possible qu'il soit juste très petit relativement à l'hypercube initial. Dans le cas où le NROY est théoriquement vide, on peut encore distinguer plusieurs cas de figure. Le NROY peut être théoriquement vide à cause de la définition de l'hypercube initial, c'est-à-dire à cause de la définition des bornes d'exploration des paramètres libres et/ou du choix des paramètres libres inclus dans l'expérience. Dans notre exemple simple, si l'on fixe une borne d'exploration minimale pour \mathbf{a} qui est trop grande (par exemple si l'on explore \mathbf{a} entre 10 et 50) ou bien si l'on n'inclue pas \mathbf{a} comme paramètre libre à ajuster et que sa valeur fixe est trop éloignée de 1, alors le NROY sera vide bien qu'il existe des vecteurs de paramètres, en dehors de l'hypercube initial, qui conduisent à une simulation correcte de f_{OLR} .

L'outil `htexplo` offre la possibilité d'explorer les paramètres libres de manière linéaire ou logarithmique, et ce afin d'explorer efficacement les paramètres dont les bornes d'exploration

balayent plusieurs ordres de grandeur. Cela est un bon moyen de déterminer l'ordre de grandeur des paramètres très incertains, afin d'éviter de passer à côté de l'espace des paramètres recherché.

Il se peut également que le modèle ne soit jamais capable de simuler l'ensemble de métriques à la tolérance à l'erreur près. L'expérience de tuning nous renseigne dans ce cas sur ce que l'on appelle parfois l'erreur structurelle du modèle, c'est-à-dire sur ses limites intrinsèques. On pourra alors, si besoin, recommencer l'expérience avec des tolérances à l'erreur plus grandes, ou avec moins de métriques.

En revanche, si le NROY est égal à l'hypercube initial Λ , c'est que tous les vecteurs de paramètres conduisent à une simulation acceptable des métriques étant données leurs tolérances à l'erreur. Si l'on souhaite affiner la sélection de vecteurs de paramètres libres, on pourra par exemple rajouter des métriques ou diminuer les tolérances à l'erreur (si cela a un sens).

Dans tous les cas, nous avons à notre disposition des ensembles de simulations aux paramètres perturbés, appelés communément PPE pour *Perturbed Physical Ensemble*, qui peuvent être exploités. C'est ce que nous allons illustrer dans la section suivante, avec une deuxième expérience de tuning en exemple.

III.3.3.1 Deuxième expérience de tuning du modèle conceptuel de climat

On reprend le modèle conceptuel de bilan radiatif de la Terre présenté Section III.3.2.1, en faisant varier cette fois les quatre paramètres libres $\mathbf{a} \in [0, 5; 3]$, $\mathbf{b} \in [250; 290]$, $\mathbf{c} \in [0, 1; 0, 3]$ et $\mathbf{d} \in [1; 3]$. L'espace initial des paramètres est désormais l'espace $\Lambda = [0, 5; 3] \times [250; 290] \times [0, 1; 0, 3] \times [1; 3]$, qui est un hypercube à quatre dimensions. On commence par faire cinq itérations avec deux métriques : la métrique f_{OLR} précédente avec une tolérance à l'erreur $T_{OLR} = 2,24 \text{ W.m}^{-2}$ et une valeur de référence $r_{OLR} = 240 \text{ W.m}^{-2}$; et une métrique sur l'ASR notée f_{ASR} . De la même manière que pour l'OLR, f_{ASR} est modélisée par l'ASR calculé avec l'Equation III.9, auquel est rajouté un bruit blanc uniforme compris entre -1 et 1 et décorrélié du bruit sur f_{OLR} . La référence associée à cette métrique est $r_{ASR} = 240,5 \text{ W.m}^{-2}$. Elle est plus élevée que r_{OLR} de $0,5 \text{ W.m}^{-2}$ car le déséquilibre énergétique actuel du système Terre est estimé autour de $0,5 \text{ W.m}^{-2}$. La tolérance à l'erreur autour de cette métrique, T_{ASR} est également prise égale à $2,24 \text{ W.m}^{-2}$, soit la racine de la somme quadratique de l'estimation de la variabilité interne de l'ASR (1 W.m^{-2}) et de l'estimation de l'incertitude autour de la référence (prise à 2 W.m^{-2} pour cet exemple, en réalité très mal connue). Après ces cinq premières itérations, on refait cinq itérations en rajoutant une métrique sur l'effet radiatif des nuages dans le solaire, appelée f_{CRE} , égale à $\mathbf{d} \times CRELW_{TOA} = \mathbf{d} \times (O_{crelw} + f(\mathbf{a}))$. Sa référence r_{CRE} vaut -50 W.m^{-2} avec une tolérance à l'erreur $T_{CRE} = 2 \text{ W.m}^{-2}$ (Loeb et al., 2009) On garde un seuil $k = 3$ tout au long de cette expérience. Le système d'équations résolu par la méthode d'HMIR s'écrit alors :

$$f_{OLR} = \mathbf{b} + O_{crelw} + f(\mathbf{a}) + \text{Bruit}_{OLR} = r_{OLR} \pm 3T_{OLR} \quad (\text{III.11})$$

$$f_{ASR} = (1 - \mathbf{c})^{S_0/4} + \mathbf{d} \times (O_{crelw} + f(\mathbf{a})) + \text{Bruit}_{ASR} = r_{ASR} \pm 3T_{ASR} \quad (\text{III.12})$$

$$f_{CRE} = \mathbf{d} \times (O_{crelw} + f(\mathbf{a})) = r_{CRE} \pm 3T_{CRE} \quad (\text{III.13})$$

On note que $f_{ASR} = (1 - \mathbf{c})^{S_0/4} + \text{Bruit}_{ASR} + f_{CRE}$.

Les valeurs de référence des paramètres libres peuvent être estimées à partir des données d'observations du bilan énergétique de la Terre. Tout d'abord, l'effet radiatif global dans l'infrarouge vaut -30 W.m^{-2} et l'OLR vaut 240 W.m^{-2} . Par construction du modèle de l'OLR, la valeur de référence pour la paramètre \mathbf{a} est 1. Donc $\mathbf{b} = 240 + 30 = 270 \text{ W.m}^{-2}$. Les paramètres libres \mathbf{c} et \mathbf{d} peuvent être dérivés du bilan radiatif observé dans le domaine solaire. En effet, sur les 342 W.m^{-2} reçu par l'atmosphère du soleil, environ 30% sont renvoyés vers l'espace et 20% sont

absorbé par l'atmosphère. Parmi ces 30%, environ 13% sont renvoyé par les nuages, et le reste, 17%, est renvoyé par la surface de la Terre, les gaz atmosphériques et les aérosols contenus dans l'atmosphère. On a donc par définition de \mathbf{c} : $\mathbf{c} \simeq 0,17$. Enfin, $\mathbf{d} = -CRESW_{TOA}/CRELW_{TOA} = 50/30 \simeq 1,7$.

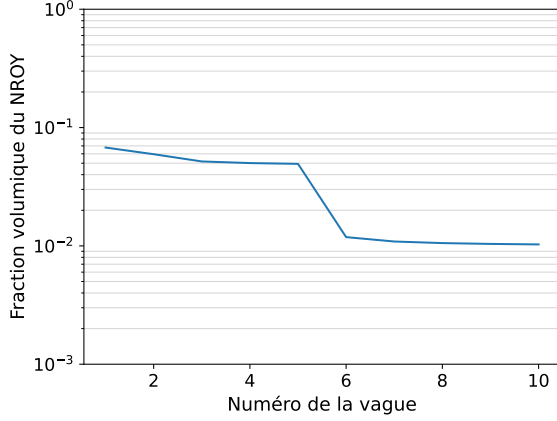


FIGURE III.10 – Evolution au cours des itérations de la fraction du volume du NROY par rapport au volume de l'hypercube initial, pour l'expérience à quatre paramètres et deux puis trois métriques. L'ajout d'une nouvelle métrique à la vague six fait chuter la fraction volumique du NROY. Une partie de l'espace conduisant à la bonne simulation des deux premières métriques ne permet pas de simuler correctement la troisième métrique.

L'évolution de la fraction volumique du NROY au cours des vagues est montrée en Figure III.10. L'ajout d'une nouvelle métrique à la sixième vague de l'expérience fait chuter la fraction volumique du NROY. Le NROY étant l'intersection des NROY associés à chacune des métriques, sa fraction ne peut que diminuer avec l'ajout de nouvelles métriques.

III.3.3.2 Exploitation des ensembles de simulations

Chaque ensemble de simulations réalisées pour constituer les bases d'apprentissage des émulateurs sont des PPEs. Lorsque l'expérience a convergé, toutes les simulations qui constituent le PPE de la dernière vague simulent les métriques à k fois leur tolérance à l'erreur près. On peut visualiser l'étendue des valeurs des métriques calculées sur les simulations des PPEs en traçant leur distance normalisées par métrique (Equation III.14) et pour l'ensemble des métriques (Equation III.15).

$$S_{f_i}(\boldsymbol{\lambda}) = \frac{|f_i(\boldsymbol{\lambda}) - r_i|}{T_i} \quad (\text{III.14})$$

$$S(\boldsymbol{\lambda}) = \max \{S_{f_1}(\boldsymbol{\lambda}), \dots, S_{f_M}(\boldsymbol{\lambda})\} \quad (\text{III.15})$$

On peut utiliser ces grandeurs comme des scores, qui permettent d'évaluer à quel point un membre de l'ensemble simule bien une ou plusieurs métriques. Des petits scores indiquent bien qu'une simulation est plus proche de la référence qu'une autre, contrairement aux invraisemblances qui peuvent être faibles à cause d'une forte incertitude de prédiction de l'émulateur. Contrairement à l'usage classique des scores, ici, plus le score est petit, meilleure est la simulation. On peut également attribuer un score unique à chaque simulation, comme étant le maximum des scores de toutes les métriques. Ce score est souvent utilisé pour classer les simulations les unes par rapport aux autres ; les meilleures simulations ayant ce "score maximal" le moins élevé. On appellera ce score maximal sur les métriques, "le score" de la simulation pour simplifier. Les scores des vagues 1, 5, 6 et 10 sont tracés sur la Figure III.11.

On voit sur cette figure que la dispersion des ensembles est fortement réduite entre la vague 1 et la vague 5. Pour les deux premières métriques, les scores sont presque tous compris entre -3 et 3 à la vague 5, signe que l'expérience a convergé avec ces deux métriques. La plus petite valeur obtenue pour le score maximum diminue également entre la vague 1 et la vague 5, signe que

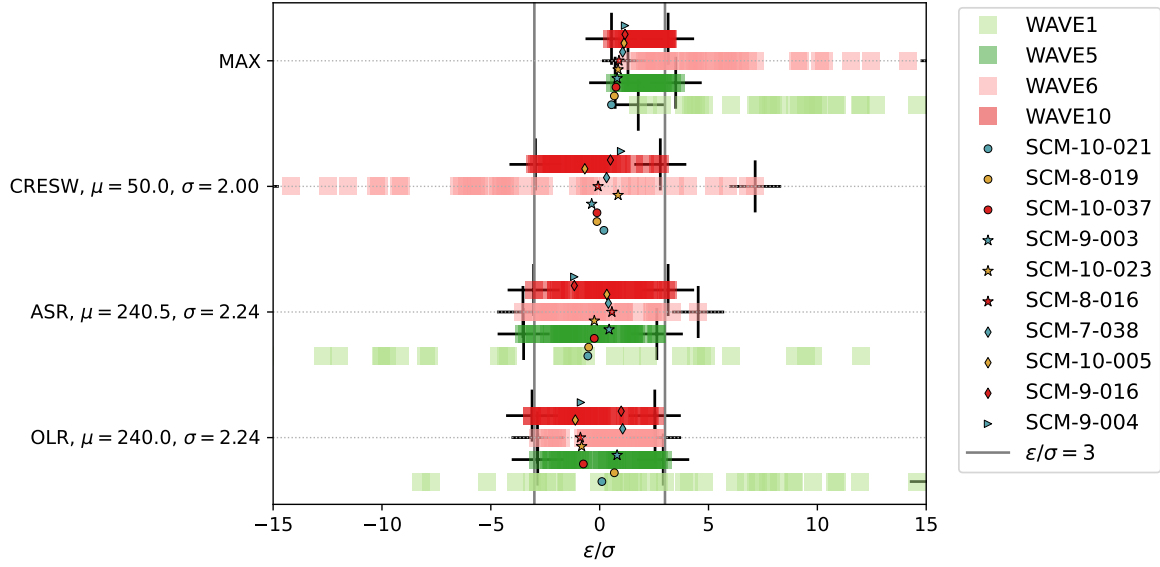


FIGURE III.11 – Scores calculés sur les simulations des vagues 1 (vert clair), 5 (vert foncé), 6 (rouge clair) et 10 (rouge foncé). Les scores sont calculés avec l'Equation III.14 pour chacune des métriques. Le maximum des scores (Equation III.15) est également tracé en haut de la figure. Les minimum et maximum de chaque ensemble sont tracés avec des croix noires. Les dix meilleures simulations, déterminées comme les simulations ayant le plus petit score maximum sur les cinq dernières vagues, sont indiquées avec des marqueurs (ronds, étoiles, losanges, triangle).

l'on se rapproche de l'espace des paramètres recherché. Lors de l'ajout de la nouvelle métrique vague 6, le score maximal sur l'ensemble des métriques est à nouveau très dispersé, à cause de la dispersion des simulations sur la nouvelle métrique. Au bout de quelques vagues, l'expérience a convergé avec les trois métriques, et l'ensemble des scores pour les trois métriques sont compris entre -3 et 3 à la vague 10.

Les meilleures simulations, celles ayant le plus petit score maximum, sont toutes issues des vagues 8 à 10 sauf une. Cela illustre bien qu'au fur et à mesure des itérations, on se rapproche de l'espace des paramètres recherché, qui est plus densément échantillonné donc qui contient avec une plus grande probabilité les meilleures simulations.

Dans le cas où l'expérience n'a pas strictement convergé, ce qui est très courant, on pourra tout de même exploiter les PPEs produits lors de l'expérience de tuning en étudiant par exemple toutes les simulations dont le score est inférieur à un certain seuil, ou alors étudier les n meilleures simulations.

Si on a besoin de figer une configuration pour le modèle, c'est-à-dire de choisir un vecteur de paramètres comme configuration de référence pour le modèle, on utilisera toujours un vecteur de paramètres pour lequel on a effectivement fait tourner une simulation, même si les émulateurs produisent des implausibilités plus faibles pour d'autres vecteurs. On pourra par exemple choisir la simulation qui a le meilleur score, ou rajouter de nouvelles métriques uniquement pour trier les simulations (c'est-à-dire sans refaire tourner de simulation, si on est à cours de temps de calcul par exemple). A noter que dans le cadre de l'history matching, en dessous de l'incertitude de comparaison entre référence et simulation, dire qu'une simulation est meilleure qu'une autre n'a plus de sens.

III.3.3.3 Visualisation des espaces de paramètres libres

Il n'est pas possible de visualiser les NROYs tels quels dès que le nombre de dimensions des espaces N est supérieur ou égal à 3. On dispose cependant d'une visualisation partielle des NROYs en deux dimensions en projetant l'espace sur toutes les combinaisons possibles de paires de paramètres, ce qui donne une matrice de dimension $N \times N$. Cette figure est appelée matrice d'in vraisemblance et est sortie automatiquement par l'outil `htexplo` à chaque itération w_i sous le nom de `InputSpace_wavewi.pdf`. La matrice d'in vraisemblance des cinquième et dixième itérations de l'expérience sont montrées en Figure III.13, mais avant ça, un extrait est présenté en Figure III.12 pour expliquer en détail les informations contenues dans ces matrices.

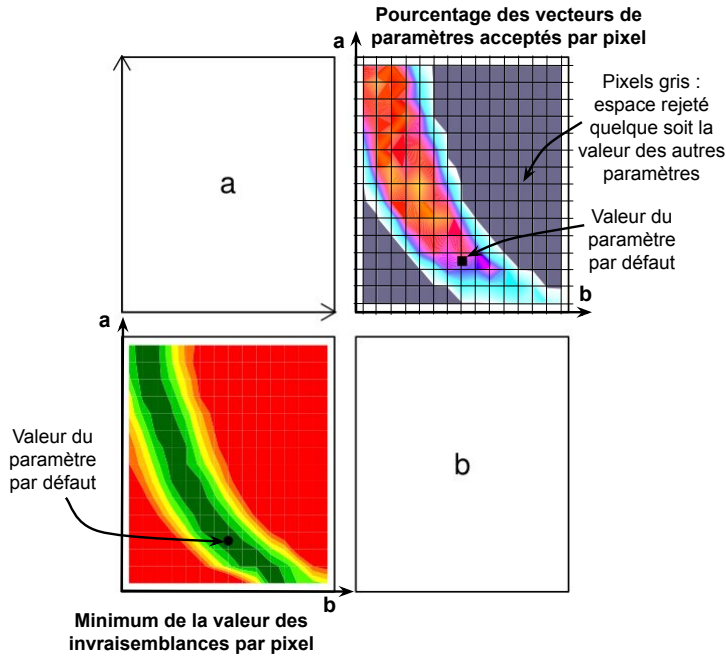


FIGURE III.12 – Zoom sur les 2×2 sous-figures correspondant aux paramètres **a** et **b** de la matrice d'in vraisemblance de la cinquième vague de l'expérience. Les échelles de couleurs sont celles de la Figure III.13 (gauche)

Sur la diagonale de la matrice on retrouve le nom des N paramètres libres que l'on a choisi d'ajuster. La diagonale renseigne les axes des sous-figures du triangle supérieur de la matrice, comme explicité sur la Figure III.12. Les axes des sous-figures du triangle inférieur de la matrice sont les mêmes que celles de leur symétrique dans le triangle supérieur de la matrice. Chaque axe des figures de la matrice est divisé en 15 intervalles ici (ce nombre peut être changé) de même taille, de manière à former une matrice de 15×15 pixels de mêmes aires. Les pixels sont surlignés pour exemple dans la Figure III.12. Les vecteurs de paramètres tirés au début de l'étape d'history matching sont répartis dans les pixels. On note ici que tous les paramètres ont été normalisés entre -1 et 1. Le pixel en bas à gauche des figures d'axe **a** et **b** contient ainsi tous les vecteurs de paramètres dont les valeurs de leurs composantes **a** et **b** sont comprises entre -1,0 et -0,867 (dans l'espace normalisé), peu importe la valeur des autres composantes (ici **c** et **d**). Dans les figures du triangle supérieur de la matrice, on colorie chaque pixel selon l'Equation III.16.

$$\frac{\text{Nombre de } \lambda \text{ ayant une invraisemblance inférieure à } k \text{ dans le pixel}}{\text{Nombre de } \lambda \text{ échantillonnés à valeur dans ce pixel}} \quad (\text{III.16})$$

L'échelle des couleurs est renseignée à droite de la figure. Lorsque cette fraction est nulle, c'est-à-dire lorsqu'aucun vecteur de paramètres dans ce pixel n'appartient au NROY, peu importe la valeur des autres paramètres, le pixel est gris. Cela permet de visualiser rapidement les parties de l'hypercube des paramètres qui sont exclues du NROY, et qui sont donc exclues de l'espace des paramètres \mathcal{A} que l'on cherche à déterminer.

Dans les sous-figures du triangle inférieur de la matrice, les pixels sont colorés selon la valeur du minimum de l'in vraisemblance des vecteurs de paramètres échantillonnés dans ce pixel. L'échelle de couleurs de ces figures est situé à gauche de la matrice d'implausibilité. Les pixels ayant un minimum d'in vraisemblance supérieur au seuil de rejet k , en rouge, correspondent aux pixels gris de la partie triangulaire supérieur de la matrice.

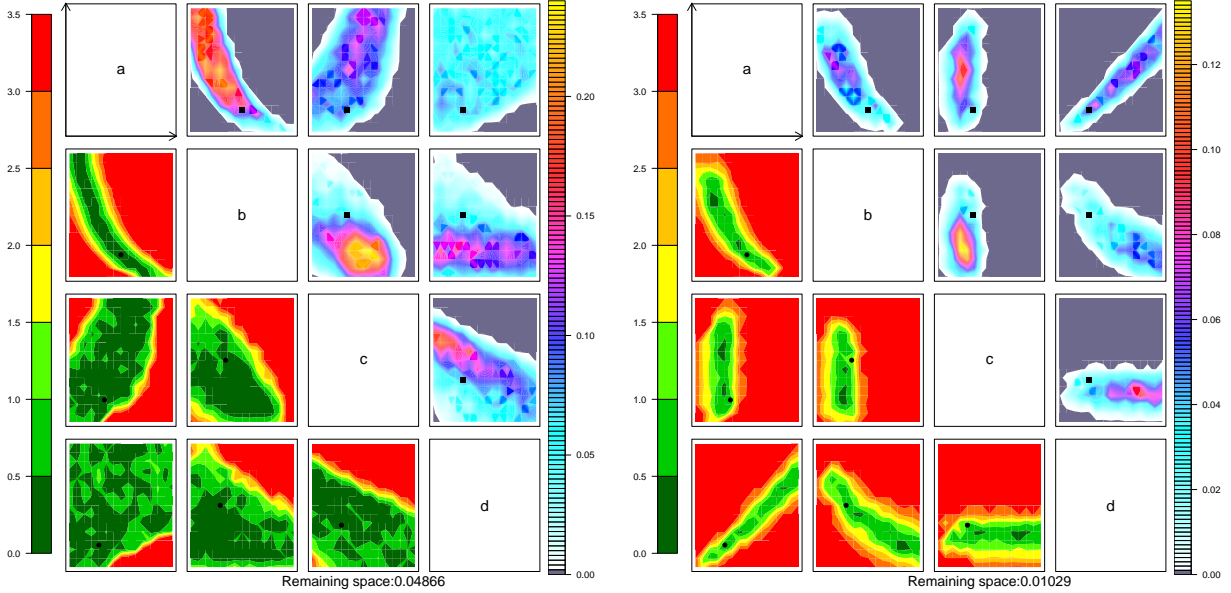


FIGURE III.13 – Matrice d'in vraisemblance de la cinquième itération (gauche) et de la dixième itération (droite) de l'expérience de tuning du modèle conceptuel.

L'interprétation de la matrice d'in vraisemblance est à faire avec précaution. En effet, elle est construite à partir des in vraisemblances calculées à partir des émulateurs. Des valeurs faibles de minimum d'in vraisemblance dans une région de l'espace peuvent ainsi être dues à des grandes incertitudes des prédictions de l'émulateur dans cette région, et n'indiquent donc pas forcément que cette région appartient à l'espace des paramètres recherchés. Lorsque l'expérience n'a pas strictement convergé, c'est-à-dire lorsque les incertitudes des prédictions des émulateurs ne sont pas négligeables devant les tolérances à l'erreur, seules les régions exclues de l'espace des paramètres peuvent être interprétées avec confiance.

Les points noirs tracés sur la matrice (un par sous-figure) correspondent aux valeurs de référence des paramètres, qui sont renseignées par l'utilisateur au début de l'expérience. Ces valeurs ne sont jamais utilisées dans la méthode. Elle sont tracées de manière indicative, pour situer le vecteur de paramètre de référence dans le NROY. Le vecteur de paramètres de référence de l'exemple, évalué Section III.3.3.1 à partir des termes du bilan d'énergie terrestre évalué pour le climat actuel, est inclus dans les NROYs de la cinquième et de la dixième itération (Figure III.13). Cela renforce la confiance à la fois dans le modèle conceptuel construit ici, mais également dans l'estimation des valeurs des paramètres libres. On note que d'autres combinaisons de paramètres que ceux estimés avec les observations conduisent à la bonne simulation des métriques avec notre modèle.

La matrice d'in vraisemblance permet également de mettre en évidence des relations entre paramètres. Dans l'exemple, la relation entre **a** et **b**, qui sort directement de l'Equation III.11, est visible sur les deux matrices d'in vraisemblance. Des relations entre **b** et **d**, et entre **a** et **d**, apparaissent lorsque la métrique sur l'effet radiatif des nuages dans le solaire est ajoutée (Figure III.13 droite). En comparant les matrices correspondant aux deux itérations, on peut également voir l'effet de l'ajout d'une métrique. On voit que le NROY est plus petit, avec plus

de zone exclues (grises ou rouge) dans la matrice de droite. L'ajout de la métrique permet d'éliminer toutes les grandes valeurs de \mathbf{c} : un trop fort albédo associé au ciel clair ne permet pas de simuler correctement les trois métriques f_{OLR} , f_{ASR} et f_{CRE} dans notre modèle ; et ce, peu importe la valeur des paramètres \mathbf{a} , \mathbf{b} et \mathbf{d} parmi leur plages initiales.

A faire :

Naj : rajouter de l'interprétation physique ici, un petit bilan de ce qu'on a appris ? pour cloturer cette section. cf commentaire dans le mattermost **A faire :**

Point de fin de relecture par FH et Maëlle

Remarque lecture : A partir d'ici ça doit être repris

III.4 Tuning de LMDZ avec htexplo

Les métriques élaborées lors de l'ANR HighTune pour le tuning de LMDZ sont présentées dans Hourdin et al. (2021) et dans cette section. Elles sont composées d'une série de métriques portant sur des cas tests uni-colonne de convection peu profonde, dites métriques orientées processus, et de métriques portant sur les flux radiatifs de simulations globales en mode océan forcé. Elles sont depuis régulièrement utilisées pour le tuning de LMDZ, pour retuner le modèle après l'ajout de nouvelles paramétrisations (Touzé-Peiffer, 2021; ?) ou pour l'étude de l'incertitude paramétrique du modèle (Hourdin et al., 2023; ?). Cet ensemble de métriques est régulièrement enrichi, au grès du tuning de nouvelles paramétrisations. C'est par exemple le cas de Touzé-Peiffer (2021) qui introduit une nouvelle paramétrisation du recouvrement vertical des précipitations (voire Section ??) contrainte par trois métriques sur les précipitations globales (voir Section ??) ; ou de Vignon et al. (2024) qui introduit des métriques sur des cas uni-colonne de couche limite stable conjointement à la revisite du schéma de turbulence de LMDZ. Cette pratique tend à se généraliser autour de LMDZ depuis le début de cette thèse, qui coïncide avec la fin de l'ANR HighTune ; et sera abordé plus en détail dans le chapitre ??. **A faire :**

FH : tu veux dire quoi exactement ? Le chapitre meta ? oui

III.4.1 Tuning dans le cadre 1D/LES

Le cadre 1D/LES (appelé aussi SCM/LES pour *Single Column Model*) est un cadre de développement et d'évaluation des GCMs axé sur les processus physiques atmosphériques. Dans ce cadre, une colonne de GCM, appelé SCM, et une simulation LES (*Large Eddy Simulation*) sont réalisées en utilisant les mêmes forçages : conditions aux bords de la colonne du GCM [*au bord = en haut et en bas tu veux dire ?*] et au bord du domaine de la LES, conditions de surface, avec parfois des tendances de vents verticaux ou des profils de taux de chauffages radiatifs imposés dans la colonne du GCM ou dans le domaine de LES. Les résultats de la simulation LES, qui simule explicitement un certain nombre de processus paramétrisés dans les GCMs, sont alors utilisés comme référence pour la simulation SCM associée. Les forçages sont choisis pour être typiques des situations météorologiques dont les processus sont paramétrisés dans le GCM (couche limite stable pour la turbulence, couche limite convective avec ou sans nuage pour les thermiques [*et le couplage turbulence/convection... en fait c'est plutôt différents régimes de fonctionnement des paramétrisations qui sont visées avec ces cas d'études*], convection profonde pour les orages etc). Une bibliothèque de ces cas 1D rassemblant les forçages disponibles dans un format standardisé afin d'être utilisable par tous les modèles, est mise à disposition par le collectif Dephy (2020) (Développement et Evaluation PHYsique des modèles Atmosphérique). Ce cadre de travail permet de développer et d'évaluer les paramétrisations du GCM au niveau des processus physiques, sans interaction avec la dynamique ni avec les autres colonnes du modèle, donc dans un environnement très contrôlé. Ce cadre présente aussi l'avantage d'avoir

un très faible coût de calcul. La simulation LES de référence est coûteuse (des dizaines d’heures de calcul sur des centaines ou milliers de CPU suivant les configurations) mais n’a besoin d’être réalisée qu’une fois, ou quelques fois avec des configurations légèrement différentes les unes des autres afin de documenter l’incertitude propre à la LES. Par comparaison, la simulation d’une colonne de GCM prend typiquement une seconde par jour simulé sur un CPU classique.

La première partie du tuning de LMDZ consiste à contraindre l’espace des paramètres libres avec des métriques sur des cas 1D de convection peu profonde ayant servi au développement du modèle du thermique (Couvreur et al., 2009; Jam et al., 2013; Rio et al., 2010; Hourdin et al., 2019). Quatre cas 1D sont utilisés dans cette première phase de tuning. Le premier cas, IHOP, est un cas de couche limite convective presque sans nuage, observé pendant la campagne de mesure IHOP (*International H2O Project*). Ce cas est issu des observations réalisées le 14 juin 2002 dans les Southern Great Plains des Etats-Unis (Couvreur et al., 2005). Le second cas, ARMCU, est un cas typique du cycle diurne de la convection peu profonde sur continent, avec développement de nuages de cumulus non précipitant dans la journée. Il est issu des observations du 21 juin 1997 du site instrumenté ARM (*Atmospheric Radiation Measurement*) également dans les Southern Great Plains étasuniennes (Brown et al., 2002). Le troisième cas, RICO (*Rain In Cumulus over the Ocean*), est un cas de cumulus précipitant sur océan typique des régions d’alizés (vanZanten et al., 2011). Enfin, le dernier cas, SANDU (Sandu and Stevens, 2011), est un cas composite de transition de stratocumulus à cumulus. Ce cas a été construit en combinant les conditions de grande échelle échantillonnées le long d’un ensemble de trajectoires lagrangiennes individuelles de 3 jours qui se sont produites dans le nord-est du Pacifique pendant les mois d’été de 2006 et 2007. La couche de stratocumulus présente un cycle diurne prononcé qui commence à se désagréger au cours du deuxième jour, lorsque la couche limite s’approfondit [*Naj : c’est le cycle diurne qui se désagrége ou la couche de stcu ? ?*]. Deux variantes de ce cas, correspondant à une transition plus lente (SANDU/SLOW) et plus rapide (SANDU/FAST) du régime de stratocumulus à cumulus, ont été construites de manière similaire et sont utilisées pour le tuning de LMDZ.

Cas	IHOP	ARMCU	RICO	SANDU	SANDU	SANDU
Sous-cas	REF	REF	REF	REF	SLOW	FAST
Période (h)	7-9	7-9	19-25	50-60	50-60	50-60
$\theta_{400-600m}$	X	X				
$qv_{400-600m}$		X				
α_{max}		X	X			
z_{nuage}		X		X		
$z_{max,nuage}$		X		X	X	X

TABLE III.2 – Métriques utilisées sur les cas de convection peu profonde

Les métriques utilisées sur ces cas sont résumées Table III.2. On a tout d’abord des métriques qui portent sur les conditions thermodynamiques moyennes de la couche limite mélangée : la moyenne entre 400 et 600 m de la température potentielle $\theta_{400-600m}$ et de l’humidité $qv_{400-600m}$. Ces métriques sont sensibles à l’efficacité du mélange du schéma des thermiques. Ensuite, des métriques sur les fractions nuageuses sont utilisées. On a une métrique sur le maximum de fraction nuageuse sur la verticale, α_{max} , et deux métriques sur la hauteur des nuages : $z_{nuage} = \frac{\int_0^\infty \alpha z dz}{\int_0^\infty \alpha dz}$ et $z_{max,nuage} = \frac{\int_0^\infty \alpha^4 z dz}{\int_0^\infty \alpha^4 dz}$, qui fait mieux ressortir la hauteur du maximum de fraction de nuageuse. Ces métriques en intégrales sont moins dépendantes de la résolution verticale du modèle que la hauteur maximale des nuages par exemple. Elles permettent d’évaluer le développement vertical de la couche limite, particulièrement important dans les cas de tran-

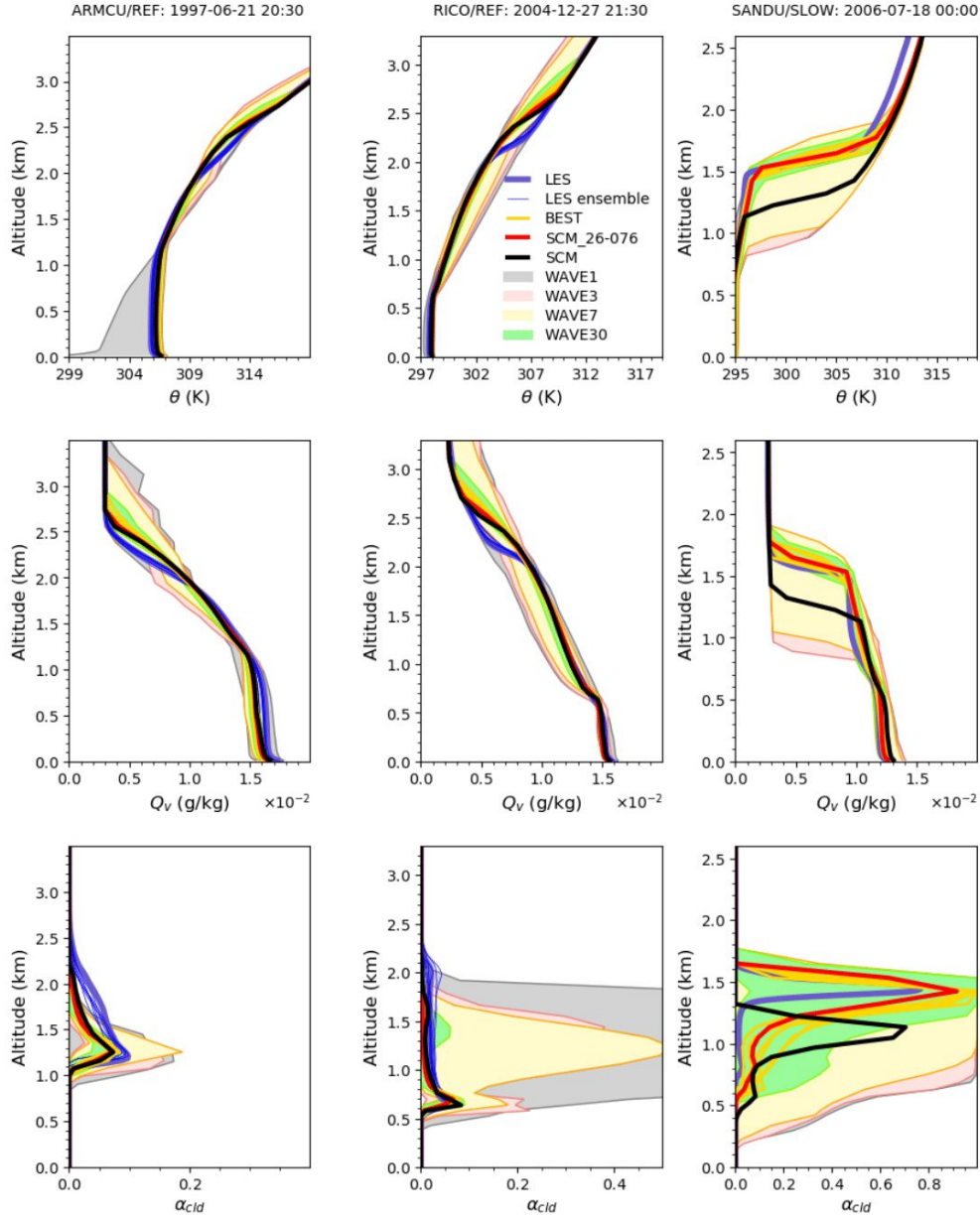


FIGURE III.14 – Figure issue de Hourdin et al. (2021). Elle montre l'effet des métriques 1D (Table ??) sur les profils de température potentielle (première ligne), humidité spécifique (deuxième ligne) et fraction nuageuse (troisième ligne) sur la moyenne des heures **A faire :**

COMPLÉTER LES XXXXXX XXX du cas ARMCU/REF (première colonne), des heures XXX du cas RICO/REF (deuxième colonne) et des heures XXX du cas SANDU/SLOW. Les enveloppes des PPEs de la première itération (gris), troisième itération (rouge), septième itération (jaune) et trentième (dernière) itération (vert) sont tracées, ainsi que l'ensemble des profils LES de référence (bleu). La meilleure simulation au sens du score maximal sur les métriques 1D est tracée en rouge, et les 10 meilleures simulations sur les métriques 1D et globales sont tracées en orange. Toutes les simulations sont faites avec une physique CMIP6 et une grille verticale à 95 niveaux, excepté la simulation tracée en noir qui correspond à la configuration de référence de LMDZ pour CMIP6 à 79 niveaux verticaux. Les métriques utilisées permettent de contraindre correctement l'altitude de l'inversion des trois cas SANDU (seul le cas SANDU/SLOW est montré ici), avec des difficultés qui persistent pour simuler correctement le profil de fraction nuageuse de ces cas. Pour les cas de cumulus (ARMCU et RICO), l'inversion est légèrement trop haute. Les métriques utilisées permettent également de sélectionner des configuration qui simulent correctement le premier pic nuageux de ces cas, avec des difficultés qui persistent sur la simulation de l'étendue verticale des cumulus. On note que la configuration de LMDZ pour CMIP6 (en noir) est très proche de la meilleure simulation sur les métriques 1D, sauf pour les cas de stratocumulus où l'altitude d'inversion et du maximum des nuages est sous-estimée. Cela peut s'expliquer par le fait que ces altitudes sont très sensibles à la discrétisation verticale du modèle.

sition stratocumulus-cumulus. Ces métriques sont moyennées sur quelques heures, pour chacun des cas, afin d'atténuer d'éventuelles oscillations numériques.

	Type	Valeur
$\theta_{400-600m}$	Absolue	0,5 K
$qv_{400-600m}$	Absolue	$5.10^{-4} \text{ kg.kg}^{-1}$
α_{max}	Relative à la fraction nuageuse	$0,25.\alpha$
z_{nuage}	Relative à l'altitude	$0,06.z - 0.2z$
$z_{max,nuage}$	Relative à l'altitude	$0,06.z - 0.2z$

TABLE III.3 – Tolérances à l'erreur associées aux métriques des cas de convection peu profonde

Les sorties des LES sont mises à disposition automatiquement par l'outil `htexplo`. Les simulations LES de référence des cas IHOP, ARMCU et RICO sont réalisées avec le modèle Meso-NH (Lac et al., 2018). L'incertitude autour de la LES de référence est évaluée sur un ensemble de simulations de sensibilité aux résolutions horizontale et verticale, à la taille du domaine, et aux paramétrisations (turbulence, microphysique, flux de surface et rayonnement). Les tolérances à l'erreur sont choisies comme des valeurs absolues pour les métriques de température et d'humidité, et comme des valeurs relatives à la fraction nuageuses ou à l'altitude sur les métriques liées aux nuages (Table III.3). Pour les métriques d'altitude des nuages, ce choix est motivé par le fait que l'épaisseur de la couche dépend presque linéairement de l'altitude. Une erreur relative en terme d'altitude correspond alors à une erreur absolue sur l'épaisseur des couches. Les tolérances à l'erreur sont choisies de manière à être toujours supérieures à l'incertitude de la référence, et à avoir un NROY non vide à la fin de l'expérience. Elle peuvent donc varier suivant la physique de LMDZ utilisée et suivant sa résolution verticale. Pour les trois versions du cas SANDU, nous ne disposons des résultats que d'une seule LES réalisée avec le modèle **A faire** :

C'est quoi les simulations Sandu ?

Entre 20 et 60 itérations sont nécessaires à la convergence de l'expérience avec ces métriques, dites métriques 1D. Ce grand nombre d'itérations est rendu possible grâce au très faible coût de calcul des simulations 1D, et à l'intégration du cadre 1D/LES dans l'outil de tuning `htexplo`. L'enchaînement des itérations est totalement automatique et nécessite de l'ordre d'une journée de calcul (temps humain).

Le résultat de l'expérience de tuning de Hourdin et al. (2021) pour les métriques présentée ici est présenté en Figure ?? pour les profils de température potentielle, d'humidité spécifique et de fraction nuageuse des cas ARMCU/REF, RICO/REF et SANDU/SLOW. Elle permettent de bien contraindre la simulation au niveau de l'inversion de la couche limite dans les cas de stratocumulus, et de bien reproduire l'altitude des nuages de ces cas, bien que l'enveloppe du dernier PPEs de l'expérience reste large pour le profil de fraction nuageuse. L'altitude d'inversion des cas de cumulus reste légèrement surestimée, avec une couverture nuageuse et une altitude du maximum des nuages bien simulées, mais une épaisseur verticale de la couche nuageuse moins bien simulée.

III.4.2 Tuning de LMDZ en global

Dans cette phase, LMDZ est utilisé en mode global en suivant un protocole proche de celui utilisé pour l'exercice d'intercomparaison des GCMs atmosphériques, dit AMIP pour *Atmospheric Model Intercomparison Project*. Il est couplé au modèle de surface continentale ORCHIDEE et forcé par un cycle saisonnier moyen de température de surface océanique et de couverture de glace de mer. Les métriques utilisées ici concernent les flux radiatifs solaires et

infrarouges au sommet de l'atmosphère, moyennés sur la deuxième année de simulation (la première étant considérée comme une phase de mise à l'équilibre statistique du modèle avec ses forçages, couramment appelée phase de spin-up). La référence utilisée est le produit satellitaire CERES-EBAF, dont l'incertitude est de l'ordre de 4 W.m^{-2} (Loeb et al., 2009) [*Naj : pour quels flux ? tous ? les flux net ? au TOA ? solaire ou IR ?*]. Des métriques portant sur les précipitations sont également régulièrement utilisées.

III.4.2.1 Présentation des métriques

Bilan radiatif au sommet de l'atmosphère

A faire :

Je pense qu'on peut alléger et simplifier cette partie mais je vais regarder sur papier. Je pense aussi qu'elle pourrait venir après les autres métriques radiatives. Genre "un traitement spécial est réalisé pour le bilan global ... Mais je vais imprimer ces deux pages pour les relire à tête reposée"

Le bilan radiatif au sommet de l'atmosphère, c'est-à-dire la somme des flux solaires et infrarouges entrant et sortant, moyenné sur tout le globe, est une cible particulière pour le tuning, partagée par de nombreux modèles de climat (Hourdin et al., 2017). Ce bilan doit être nul dans un climat à l'équilibre par conservation de l'énergie du système Terre. Cependant, le climat actuel n'est pas à l'équilibre à cause du changement climatique. Une partie de la chaleur retenue dans le système par l'augmentation des gaz à effet de serre est absorbée par les océans, qui ont des temps d'équilibrage long (une vingtaine d'années pour l'océan de surface, plusieurs centaines d'années pour l'océan profond). Ce déséquilibre provoque un déséquilibre du rayonnement global au sommet de l'atmosphère d'environ $0,5$ à 1 W.m^{-2} , qui est en dessous de l'incertitude d'observation des flux radiatifs au sommet de l'atmosphère (Loeb et al., 2009). Mais ce n'est pas ce chiffre qui est visé, car on sait que pour atteindre des températures de surface océanique à l'équilibre en accord avec les observations, dans le modèle couplé (LMDZ-ORCHIDEA-NEMO), il a fallu viser un déséquilibre énergétique de $2,7 \text{ W.m}^{-2}$ au sommet de l'atmosphère pour CMIP6. [*Naj : 2.7 en mode AMIP tu veux dire ?*]

Cette différence d'équilibrage entre modèle couplé océan-atmosphère et modèle atmosphérique forcé est expliquée par Gregory and Mitchell (1997). Il montre que dans le modèle couplé, un changement de forçage radiatif F provoque un changement de température à la surface du globe T et de bilan radiatif au sommet de l'atmosphère G suivant l'Equation III.17.

$$\Delta G = \Delta F + \eta \Delta T \quad (\text{III.17})$$

Avec η une constante. Appliquée au changement climatique, cette équation définit la sensibilité climatique, c'est-à-dire la température qu'aurait la surface de la Terre après avoir subi un forçage radiatif ΔF et être revenue à l'équilibre ($\Delta G = 0$). Cette équation peut aussi s'appliquer pour comparer deux jeux de simulations.

On prend ici une simulation globale couplée atmosphère-océan de référence dont la température de surface est proche des observations en moyenne globale et une autre simulation dans laquelle le modèle atmosphérique a un peu changé (un changement de vecteur de paramètres libres, ou un nouveau développement physique par exemple). Entre les deux simulations avec océan forcé [*Naj : tu viens de dire qu'on est en couplé la...*], la température de surface sera à peu près la même puisque la température océanique de surface est imposée. La différence de bilan radiatif au sommet de l'atmosphère entre les deux simulations ΔG_a sera alors égale à la différence de forçage radiatif entre les deux simulations ΔF_a . On suppose que la différence de forçages entre les deux simulations couplées et les deux simulations forcées est la même [*Ah*

donc tu as 4 simulés ?]. Entre les deux simulations couplées, on aura donc $\Delta G_c = \Delta F_a + \eta \Delta T_c$. Pour que notre deuxième simulation ait la même température de surface que notre simulation de référence, il faut qu'à l'équilibre, ΔT_c soit nul. Et puisque ΔG est nul à l'équilibre, il en résulte qu'il faut que $\Delta F_a = \Delta G_a$ soit nul. Ainsi, pour que notre deuxième simulation ait la même température de surface que la simulation de référence, il faut que le bilan radiatif global au sommet de l'atmosphère de la simulation forcée soit égal à celui de la simulation de référence, qui vaut ici $2,7 \text{ W.m}^{-2}$. *[Naj : dur à suivre mais nul doute ça va être utile quand ce sera plus clair !! peut être labelliser les simulés : v1-couplée, v1-forcée, v2-couplée, v2-forcée ?]*

La cible de $2,7 \text{ W.m}^{-2}$ pour le bilan radiatif global est donc une valeur empirique. Elle contient l'aspect transitoire du changement climatique ($0,5\text{-}1 \text{ W.m}^{-2}$), les fuites d'énergie du modèle (évaluées autour de $0,7 \text{ W.m}^{-2}$) et probablement des changements dans le climat moyen affectant le bilan global comme la structure des SSTs, la banquise ou la position du courant-jet qui contrôle entre autres les trajectoires des dépressions de moyennes latitudes. Ce courant est plus proche de l'équateur dans les simulations couplées que forcées, ce qui concentre les nuages tropicaux proche de l'équateur, augmentant ainsi leur effet parasol dans le solaire donc diminuant le bilan radiatif global au sommet de l'atmosphère. La valeur cible pour le bilan radiatif global en forcé peut changer si le modèle change trop. On la considère stable lorsque l'on change les vecteurs de paramètres libres du modèle atmosphérique, au cours de l'exercice de tuning du modèle atmosphérique. Cette hypothèse est vérifiée en traçant le bilan radiatif global G_a d'un ensemble aux paramètres perturbés de simulation forcées, en fonction de la température de surface de l'océan SST dans l'ensemble de simulations couplées associé (c'est-à-dire ayant les mêmes vecteurs de paramètres libres pour l'atmosphère). L'incertitude autour de la métrique de bilan radiatif global est prise égale à la dispersion de cet ensemble de simulations autour de la valeur de température océanique de référence, soit égale à $0,2 \text{ W.m}^{-2}$. Le flux radiatif global sortant au sommet de l'atmosphère dans le spectre solaire est également utilisé comme métrique de tuning, avec cette fois une tolérance à l'erreur de 5 W.m^{-2} , c'est-à-dire de l'ordre de l'incertitude du produit de référence. Cela permet de contraindre la répartition solaire/infrarouge dans le bilan radiatif global au sommet de l'atmosphère.

Autres métriques radiatives

D'autres métriques sont définies à partir des flux solaire et infrarouge sortant au sommet de l'atmosphère, moyennés sur différents masques spatiaux. Trois masques sont utilisés pour classer la bande tropicale océanique en régime convectif, intermédiaire (alizés) et subsident. Ces masques sont calculés sur des climatologies de vitesses verticales issues des réanalyses ERAI *[C'est publié quelque part ?]*. Des masques sont également utilisés pour définir des métriques d'anomalies entre des régions. Le premier est un masque entre la région du courant circumpolaire Antarctique au quarantième rugissant et la région tropicale. La métrique définie par ce masque permet de cibler explicitement le biais chaud du courant circumpolaire Antarctique dans le modèle. Le deuxième masque vise une anomalie entre les bords Est des continents tropicaux, caractérisée par la présence de stratocumulus, et le reste de la bande tropicale. Le masque est appelé ETOA pour *East Tropical Ocean Anomaly*. Les métriques définies à l'aide de ce masque permettent de cibler explicitement le biais chaud des températures de surface océanique de ces régions (Hourdin et al., 2015). Ces métriques ont toutes des tolérances à l'erreur de 5 W.m^{-2} .

Métriques sur les précipitations

Touzé-Peiffer (2021) introduit trois métriques sur les précipitations, dont la référence est le produit GPCP (*Global Precipitation Climatology Project*, Adler et al. (2018)). La première métrique porte sur la variabilité des précipitations dans les régions concernées par l'oscillation de Madden-Julian *[océan pacifique ?]*, la seconde métrique porte sur les précipitations moyennes du

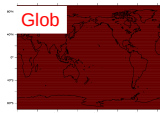
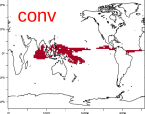
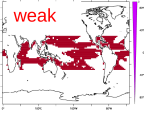
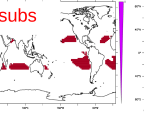
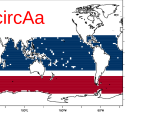
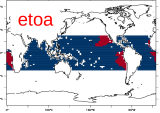
Mask	Variable	Metrics	target W m^{-2}	error W m^{-2}
	Total rad. TOA (rt) Swup TOA (rsut)	glob.rt	2.5	0.2
		glob.rsut	99.6	5
		circAa.rsut	24.0	5
		circAa.rlut	-48.6	5
Convective, intermediate, subsiding, Circum Antact. anomaly	SWup TOA (rsut) LWup TOA (rlut)	subs.rsut	84.9	5
		weak.rsut	81.8	5
		conv.rsut	103.2	5
		subs.rlut	274.6	5
		weak.rlut	264.3	5
		conv.rlut	235.8	5
	SWup TOA (rsut)	etoe.rsut	11.0	5
Eastern Tropical Ocean anomaly				

FIGURE III.15 – Figure issue de Hourdin et al. (2021) montrant les différents masques utilisés pour calculer les métriques (gauche) et la liste des métriques radiatives avec leur valeurs de références et les tolérance à l'erreur classiquement utilisées (droite). Les métriques de différences (circAa et ETOA) sont calculées comme la différence entre les métriques moyennées sur les régions rouges et bleues. Les notations rt, rsut et rlut utilisent la norme des variables des CMIP6 : r pour radiatif, s pour solaire ou l pour infrarouge (*longwave*), u pour pour les flux montant (*up*) et t pour TOA (*Top Of Atmosphere*).

Sahel et la dernière métrique porte sur l'occurrence des précipitations journalières supérieures à 50 mm.j^{-1} , le modèle ayant tendance à surestimer ces précipitations sur le continent maritime et au Nord-Ouest de l'Amérique Latine.

III.4.2.2 Résultats

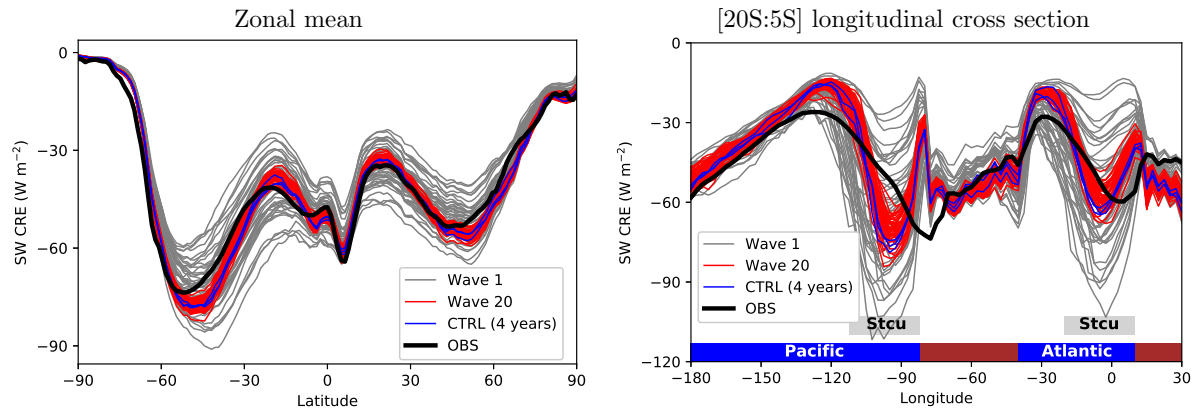


FIGURE III.16 – Figure issue de Hourdin et al. (2021)

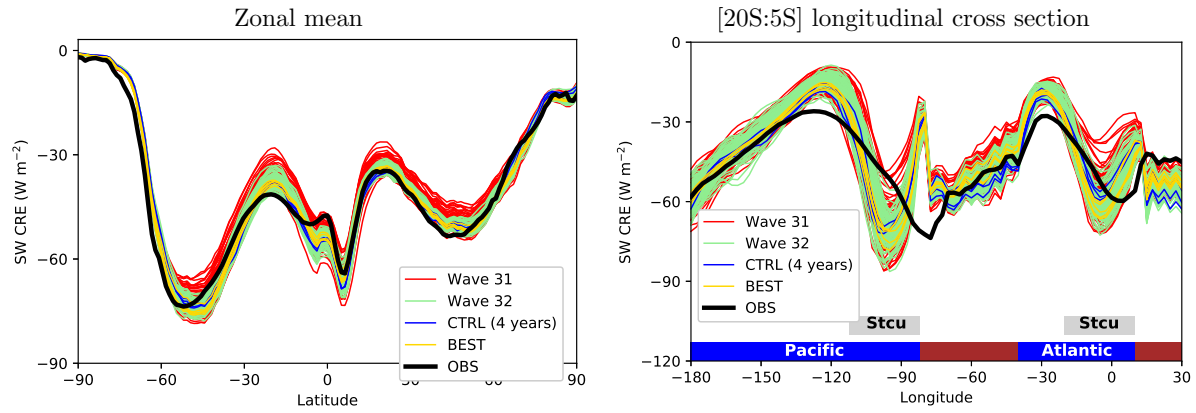


FIGURE III.17 – Figure issue de Hourdin et al. (2021)

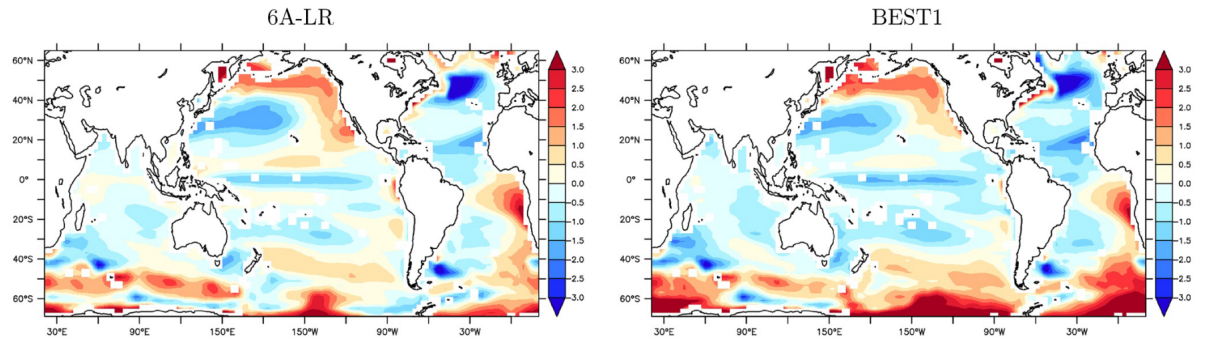


FIGURE III.18 – Figure issue de Hourdin et al. (2021)

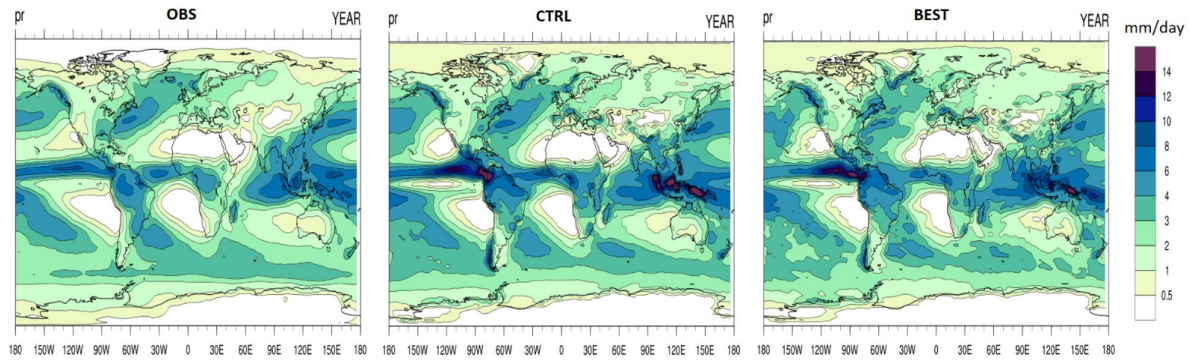


FIGURE III.19 – Surface precipitation in observations (OBS), in the standard version of LMDZ (CTRL) and in the BEST simulation of the new version. As in Fig. 5.10, precipitation observations are from GPCP (Adler et al., 2018). Precipitation in CTRL and BEST is averaged over a year of simulation.

Chapitre IV

Tuning visant les observations du site instrumenté du SIRTa : mise en place d'une configuration de référence

Remarque lecture : Quasiment terminé

IV.1 Introduction

Bien simuler le climat à la surface des continents est un objectif important des modèles globaux de climat. Les projections climatiques produites par ces modèles sont en effet utilisées pour évaluer les impacts du changement climatique sur les installations humaines (ville, énergie solaire, ressources en eau, réseau électrique etc) se situant majoritairement à la surface des continents, ainsi que sur la santé humaine (pollution de l'air, canicules etc), l'agriculture et les écosystèmes terrestres. Les impacts peuvent être évalués directement dans les projections climatiques fournies par les GCMs ou dans des modèles plus fins forcés par les simulations globales des GCMs.

Le climat à la surface des continents est au premier ordre le résultat des équilibres atmosphériques globaux et à plus courtes échelles de temps, de la météorologie. Aux moyennes latitudes, la séquence d'anticyclones et de dépressions va par exemple être déterminante pour la météorologie locale et les variations intra-saisonnière. L'état de la surface des continents (humidité du sol, type de sol, végétation, etc) influence également le climat à la surface des continents ; et ce plus ou moins fortement selon la situation météorologique et les échelles de temps et d'espace. Les échelles régionales et intra-saisonnières sont globalement les plus sensibles aux couplages entre l'atmosphère et la surface. La compréhension des mécanismes qui régissent ces couplages entre atmosphère et surfaces continentales, ainsi que leur réponse au changement climatique sont des sujets de recherche actifs.

Une manière d'améliorer la simulation du climat à la surface des continents est de cibler celle-ci directement dans la procédure de tuning. **Ainsi, la stratégie de tuning proposée ici est élaborée avec l'intention de contraindre le climat de surface au plus proche des processus, pour éviter les compensations d'erreurs et créer des conditions favorables à une meilleure compréhension des interactions entre surface et atmosphère.**

Nous proposons de construire une stratégie de tuning ciblant explicitement le couplage entre atmosphère et surface continentale, afin de garantir que les variables résultants de ce couplage, comme la température et l'humidité de l'air proche de la surface, soient simulées correctement et si possible pour les bonnes raisons. Cibler le couplage est également en théorie un bon moyen

de tuner conjointement les paramètres libres de LMDZ et d'ORCHIDEE, qui sont actuellement tunés séparément.

Les références utilisées ici pour ce tuning sont des observations du site instrumenté du SIRTa, situé à Palaiseau en banlieue parisienne. Les sites instrumentés tel que celui du SIRTa ont été déployés avec l'objectif d'aider à la compréhension des processus atmosphériques et d'observer le climat et le changement climatique. Des moyens importants sont alloués au fonctionnement de ces "super-sites". Le potentiel offert par la synergie entre les multiples observations qui y sont recueillies est cependant parfois considéré comme sous exploité (Chiriaco et al., 2018).

Nous proposons d'explorer dans cette étude le potentiel de ces observations sur super-sites comme référence pour le tuning de modèle de climat. C'est pourquoi nous choisissons de nous focaliser sur un seul site instrumenté. Le site du SIRTa est choisi pour son appartenance à des réseaux de mesures internationaux (FLUXNET pour les flux turbulents de surface et BSRN pour les flux radiatifs), pour sa proximité avec les équipes du LMD et son histoire de comparaison avec LMDZ (Campoy et al., 2013; Cheruy et al., 2013; Coindreau et al., 2007). Des comparaisons LMDZ-SIRTa sont disponibles en temps réel sur le site <https://web.lmd.jussieu.fr/~lmdz/LMDZOPE/LMDZ-SIRTa/meteo/> (LMD/EMC3, 2020)

Enfin, cette étude est motivée par un souci de sobriété numérique. Elle vise en effet à ajouter, dans la stratégie de tuning de LMDZ, une étape intermédiaire pas trop coûteuse en temps de calcul et permettant d'assurer une continuité entre les métriques 1D et les métriques globales. Focalisée sur des flux et variables météorologiques en surface, alors que le tuning 3D effectué jusque là portait sur des flux au sommets, cette étape se veut également plus proche des processus (notamment couche limite, convection et nuages associés) et plus systématique que les quelques cas tests utilisés dans le tuning 1D. Une attention particulière est ainsi portée à l'utilisation d'une configuration légère de LMDZ pour la contrainte du climat de surface.

Cibler des observations locales et orientées processus dans une configuration de LMDZ globale, zoomée à dynamique contraintes est un intermédiaire dont les potentialités pour le tuning de modèle de climat ont été identifiées par Hourdin et al. (2017) et dont nous testons, dans cette étude, la mise en pratique.

IV.1.1 Histoire de l'expérience

La première expérience de tuning des paramètres libres de LMDZ visant les observations du SIRTa, a été un échec. La compréhension de cet échec a été longue. Un travail méthodique et conséquent sur la configuration du modèle a dû être effectué afin de pouvoir mettre en place une expérience de tuning robuste, permettant une comparaison fiable et quantifiée du modèle avec les observations du SIRTa. Ce travail de mise au point de la configuration de LMDZ, au final très instructif, est l'objet du présent chapitre. Le chapitre suivant présente l'expérience de tuning

Dans cette première expérience, l'objectif était de tuner 18 paramètres libres de LMDZ en visant l'effet radiatif des nuages à la surface et les précipitation en moyenne sur plusieurs périodes d'intérêt, en même temps que des discussions étaient menées avec des collègues pour savoir quels paramètres d'ORCHIDEE inclure dans un futur tuning. Les tolérances à l'erreur associées à chacune des métriques étaient choisies supérieures aux incertitudes d'observations, et de manière à avoir un NROY non vide. Dès la deuxième vague de tuning, les résultats de l'expérience étaient déroutants. Le PPE de la deuxième vague était presque aussi large que celui de la première vague, et ne paraissait pas se rapprocher des références, avec une fraction volumique du NROY chutant drastiquement sur les deux premières vagues.

Nous nous sommes alors rendu compte que nous n'avions évalué la variabilité interne du modèle, que l'on pensait faible grâce à un guidage des champs météorologique vers des réanalyses (référence météorologiques, voire Section IV.3.3.2). Ce guidage était en effet sensé enlever la part

d’incertitude liée à la très forte sensibilité des simulations météorologiques à l’état initial, celle-là même qui limite fortement l’horizon de prévisibilité du temps.

Cette variabilité interne s’est avérée être en fait très importante sur les précipitations des périodes de convection profonde, et non négligeable sur toutes les variables étudiées sur les périodes de convection profonde et les semaines suivantes. Sur les métriques de précipitations, la tolérance à l’erreur utilisée dans l’expérience de tuning était inférieure à la variabilité interne du modèle. Un vecteur de paramètres considéré comme bon sur le critère de l’invraisemblance, pourra être considéré comme mauvais dans une simulation où l’on aura fait une petite perturbation ; et inversement. C’est comme si les vecteurs de paramètres retenus comme bon à la fin de la première vague avaient été choisis au hasard. Cela explique que la dispersion des simulations des PPEs ne diminuent pas entre la première et la deuxième itération ; et pourquoi le NROY devenait si petit si rapidement. Augmenter la tolérance à l’erreur sur les métriques concernées ne pouvait pas permettre de contourner ce problème, puisque la variabilité interne était de l’ordre de grandeur de la dispersion des simulations des PPEs sur certaines métriques. Dans ce cas, on ne peut plus discriminer les simulations des PPEs, et considérer qu’une simulation est meilleur qu’une autre.

Une partie significative du travail de thèse a alors consisté à investiguer ce problème et à trouver des stratégies de contournement pour profiter malgré tout des observations sur site pour le tuning des modèles, et de ce fait montrer le potentiel mais aussi les limites de cette proposition.

Une partie du travail a en particulier consisté à étudier la variabilité interne du modèle avec guidage pour la comprendre et pouvoir si possible la diminuer. Tout d’abord, nous avons testé différents types de guidage. Nous montrons Section IV.4.2 que le guidage en vent, humidité et température a permis de diminuer la variabilité interne avant la première période de convection profonde par rapport au guidage en vent uniquement. Nous nous sommes ensuite intéressés à la variabilité interne des précipitations en période convective, que l’on soupçonnait d’être responsable de la forte variabilité interne des autres variables (Section IV.5). Nous avons montré que la variabilité spatiale de ces précipitations était fortement surestimée, ce qui pouvait expliquer leur forte variabilité interne. Là encore, les PPEs générés par les expériences de tuning nous ont donc permis de mettre en évidence un biais structurel du modèle.

Nous avons cependant montré que le fait de moyenner les précipitations sur plusieurs mailles faisait suffisamment diminuer leur variabilité interne pour pouvoir les utiliser comme métrique pour le tuning. La référence utilisée est alors le produit radar Coméphore de Météo France (Tabary et al., 2012), qui fournit des données de précipitations toutes les heures à l’échelle kilométrique sur toute la France.

Cependant, si le fait de moyenner les précipitations spatialement permet de les utiliser comme métrique de tuning, le fait que la précipitation puisse varier du tout au tout d’une maille à l’autre (ou de façon équivalente, en supposant le champ de pluie ergodique, d’une simulation à l’autre sur la maille du SIRTa), ne résout pas le fait que : (1) la colonne de sol au SIRTa va recevoir aléatoirement beaucoup ou très peu de précipitations ; (2) la colonne atmosphérique du SIRTa va en retour “voir” aléatoirement des simulations avec des sols sec ou humide. L’ambition de contraindre les paramètres libres liés à l’hydrologie du sol, qui faisait partie de nos motivations, est donc empêchée par cette variabilité surestimée des précipitations. Il a néanmoins été possible de contourner les difficultés liées à l’impact de la variabilité de l’humidité de sol sur la colonne atmosphérique au SIRTa, en coupant la rétroaction de l’hydrologie d’ORCHIDEE sur l’atmosphère. Ce faisant, nous limitons l’ambition de l’expérience de tuning, puisqu’elle ne pouvait plus concerner le couplage avec l’hydrologie du sol, un des objectifs initiaux de l’exercice. Nous maintenons cependant la possibilité de tuner le modèle au SIRTa, en matière de physique de Nuages, de rayonnement et de couplage avec la conduction dans le sol, mais dans un cadre da-

vantage contraint, où à la fois la succession des séquences météorologiques et la capacité du sol à évaporer étaient contraintes par l'observation. En pratique, la proposition consiste à imposer l'évolution en temps, sur la maille du SIRTa, du facteur d'aridité (appelé aussi facteur d'évaporation) diagnostiqué à partir des observations (Section IV.6). Ce facteur, noté β , est d'ordinaire évalué par ORCHIDEE pour calculer l'évaporation comme produit de ce β par l'évaporation potentielle atmosphérique (capacité de l'atmosphère à recevoir un flux d'évaporation).

Un des objectifs initiaux du tuning au SIRTa, motivant la prise en compte des paramètres libres d'ORCHIDEE dans le tuning (ce qui n'était pas le cas de ce premier tuning exploratoire), était de cibler l'assèchement printanier du sol, bien trop rapide et intense dans les simulations. Cet assèchement printannier s'accompagne d'une sous-estimation de l'effet radiatif des nuages (c'est-à-dire d'un ensoleillement trop important de la surface) et d'une sous-estimation de l'évaporation des sols en fin de printemps et durant l'été. Cette boucle de rétroaction positive se poursuit tout au long de l'été. Ce comportement du modèle est bien connue, et a été mis en évidence avec LMDZ au SIRTa par les études de Campoy et al. (2013); Cheruy et al. (2013) et Coindreau et al. (2007). L'origine de ces biais n'est cependant pas fermement établie : difficulté des modèles d'atmosphère à faire suffisamment d'effet radiatif des nuages ? Difficulté des modèles d'hydrologie de surface à mettre en jeu des constantes de temps longues, nécessaire à une évaporation d'été puisant dans les ressources en eau accumulée dans le sol en hiver par l'eau accumulée en hiver (dans les moyennes latitudes) ? Ce biais fait toutefois partie des raisons invoquées pour expliquer une tendance des modèles de climat à partir en mode caniculaire l'été, sur les continents des moyennes latitudes, motivant son étude. Le fait d'imposer le facteur d'aridité dans le protocole de tuning ne permet donc plus de cibler spécifiquement ce biais, mais permet en revanche de s'affranchir de la boucle de rétro-action positive du manque d'humidité du sol sur l'effet radiatifs des nuages, et donc de savoir si, avec une disponibilité suffisante en eau dans le sol, le modèle d'atmosphère est capable de simuler suffisamment d'effet radiatif des nuages ou non, et d'autres aspects de la météorologie au SIRTa.

Ce protocole expérimental avec β imposé met en évidence un enjeu mieux identifié aujourd'hui qu'à l'époque du début de ma thèse : l'utilisation de l'approche d'HMIR comme un outil d'exploration de questions scientifiques bien identifiées, dans des configurations adaptées. On prônerait peut-être d'emblée aujourd'hui cette stratégie à deux vitesses : 1) dans un premier temps modifier la configuration pour couper la rétroaction avec le sol afin d'aborder la question de la représentation des nuages dans le modèle atmosphérique "pour une hydrologie du sol correcte"; 2) dans un second temps utiliser les configurations ainsi réglées pour attaquer les questions relatives aux constantes de temps longues de l'assèchement au printemps. Le travail nécessité par la mise au point de la configuration n'a pas permis d'aller jusque là mais il a largement pavé le chemin pour un tel travail. [Reprendre ça dans le meta]

Une nouvelle difficulté est alors apparue sur un ensemble de simulations réalisé avec cette configuration. Imposer le facteur d'évaporation sur une seule maille du modèle créait des incohérences dans les variables proches de la surface, probablement parce que l'air venant des mailles adjacentes avait vu un facteur d'évaporation très différent et était donc différemment chargé en humidité de façon irréaliste. Typiquement, on peut penser que l'été, un air asséché dans les mailles adjacentes qui n'avaient plus de capacité d'évaporation, arrivait subitement sur une surface plus humide donc par exemple sans doute nettement plus froide en pleine journée. Pour éviter ce problème, nous avons imposé le facteur d'évaporation diagnostiqué au SIRTa dans un rayon de 100 km autour de la maille du SIRTa.

Mais l'histoire ne s'est pas terminée ici. Dans un grand nombre de simulations réalisées avec ce coefficient β imposé sur un rayon de 100 km autour du site, l'humidité du sol devenait négative lorsque l'on imposait le facteur d'évaporation. Or dans ORCHIDEE, l'inertie thermique du sol, qui contrôle la conduction thermique sous la surface, dépend de l'humidité du sol. Il faut donc

en plus de maintenir une évaporation correcte, maintenir un contenu en eau du sol raisonnable pour bien simuler la température du sol et donc de l'air dans la colonne atmosphérique au-dessus. Ce problème faisait écho à un problème préalablement identifié au SIRTa lors de la thèse d'Aurélien Campoy, à savoir la présence d'une nappe perchée sur le site, qui maintient en permanence un sol humide au SIRTa vers 1 m de profondeur. Pour résoudre ce problème, nous avons alors de plus imposé une humidité à saturation en dessous de 1.3 m en augmentant la discrétisations du sol à 22 couches (contre 11 dans la version de référence), en suivant les choix de Campoy (2013). Le LAI (*Leaf Area Index* ou indice de surface foliaire), qui rentre dans le calcul des hauteurs de rugosité, a également été imposé afin de s'affranchir de sa dépendance à l'hydrologie et de réduire l'incertitude liée à la localité du site.

L'ensemble des développements nécessaires à ce tuning au SIRTa s'est avéré, au bout du compte, représenter une partie significative de mon travail de thèse.

IV.1.2 Plan

Nous avons ainsi choisi de découper cette étude en deux chapitres : un premier chapitre qui expose le travail sur l'évaluation de LMDZ avec les observations du site instrumenté du SIRTa, aboutissant à une configuration de référence du modèle, et un deuxième chapitre qui propose une première expérience de tuning utilisant cette configuration de modèle et visant les observations du SIRTa.

Ce chapitre ci commence par présenter les éléments qui ont motivés cette étude Section IV.2, puis expose Section IV.3 les observations utilisées ainsi qu'une première configuration de référence de LMDZ. La Section IV.4 traitera la question de la variabilité interne du modèle, et la Section IV.5 se focalisera plus précisément sur la variabilité interne des précipitations convectives. Une configuration avec un facteur d'aridité imposé est construite et étudiée Section IV.6. La configuration finalement retenue est résumée Section IV.7, et les conclusions de tout ce travail seront exposées Section IV.8.

IV.2 Motivations

IV.2.1 Importance du couplage entre l'atmosphère et la surface

L'interface entre atmosphère et surface sur continents est modélisée sans dimension verticale dans LMDZ6A. Même avec des arbres pouvant atteindre plusieurs dizaines de mètres, et une première couche du modèle de 20 m d'épaisseur (dans toutes les configurations utilisées dans cette thèse), les obstacles présents à la surface des continents sont considérés comme petits devant l'épaisseur de la première couche d'atmosphère. On prend en compte leur effet sur le freinage du vent ou les échanges d'eau et d'énergie à la surface au travers du concept de hauteur de rugosité.

Les échanges d'eau à la surface sont régis par une équation de bilan de masse :

$$Pr - E - Rs - Inf = 0 \quad (IV.1)$$

On note Pr le flux de précipitation arrivant à la surface, E l'évapotranspiration, Rs le ruissellement et Inf l'infiltration d'eau dans le sol.

Le bilan d'énergie en surface s'écrit quant à lui :

$$SW_{down} - SW_{up} + LW_{down} - LW_{up} - LAT - SENS - TC = 0 \quad (IV.2)$$

où SW_{down} et LW_{down} sont les flux radiatifs solaire et infrarouge arrivant à la surface, SW_{up} et LW_{up} les flux radiatifs solaire et infrarouge montant depuis la surface, LAT est le flux de

chaleur latente, $SENS$ le flux de chaleur sensible et TC le flux de chaleur partant dans le sol. Ce dernier est généralement faible devant les autres termes du bilan d'énergie.

Les bilans d'eau et d'énergies sont liés par le flux de chaleur latente qui est proportionnel à l'évapotranspiration, $LAT = LE$ où $L \simeq 2,45 \cdot 10^6$ J/kg est la chaleur latente de vaporisation.

L'humidité du sol est une variable importante du couplage atmosphère-surface sur continents. Ses mécanismes d'interaction avec le climat et le changement climatique sont complexes et sont le sujet de nombreuses recherches (Seneviratne et al., 2010). Le temps caractéristique de résidence de l'eau dans le sol est de l'ordre de quelques mois pour les aquifères classiques. L'humidité du sol intègre l'histoire météorologique des derniers mois et agit comme une mémoire des bilans d'eau passés. Son couplage avec l'atmosphère dépend du régime d'évaporation. On parle de régime saturé lorsque le sol est proche de la saturation et que l'évapotranspiration ne dépend que de l'énergie disponible, c'est-à-dire au premier ordre du rayonnement solaire arrivant à la surface ; et de régime sec lorsque le sol est très sec et que l'évapotranspiration ne dépend que de l'eau disponible, c'est-à-dire au premier ordre des précipitations. Entre ces deux situations, le régime est dit transitoire ; et c'est le régime où le couplage humidité du sol - atmosphère est le plus fort. Dans la moitié nord de la France, où le SIRTa est situé, le régime d'évaporation est saturé en hiver, et transitoire sinon ; hors période de sécheresse estivale où le régime peut devenir sec. Les climats méditerranéens sont quant à eux caractérisés par une prédominance de régime sec tout au long de l'année.

Une augmentation d'humidité du sol (en régime sec ou transitoire) conduit à une évaporation plus importante donc à une atmosphère plus humide, avec un plus grand potentiel de formation de nuage et de précipitations. Les précipitations agissent directement sur l'humidité du sol et le bilan d'eau à la surface (Equation IV.1). Les nuages influencent quant à eux les flux radiatifs solaires et infrarouges arrivant à la surface, donc le bilan d'énergie (Equation IV.2). Les nuages constituent le mécanisme principal de rétroaction de l'humidité du sol sur le flux radiatif solaire net à la surface. Une rétroaction de l'humidité sur l'albédo du sol est également possible : un sol nu humide aura un albédo un peu moins important qu'un sol sec ; et l'humidité du sol va impacter le couvert végétal qui impacte à son tour l'albédo de la surface. Les rétroactions de l'humidité du sol sur le rayonnement infrarouge sont plus complexes. Le rayonnement infrarouge arrivant à la surface dépend des nuages (couverture nuageuse et température de la base des nuages), de la température et de l'humidité de l'atmosphère, notamment des couches les plus proches de la surface ; sachant que ces trois composantes rétroagissent les unes sur les autres et sont influencées par la surface. Le rayonnement infrarouge émis par la surface est quant à lui directement lié à la température de la surface par la loi de Stefan. L'émissivité de la surface est imposée à 1 dans LMDZ, pour des raisons "structurelles" de transfert radiatifs [*Maelle : Quelqu'un peu confirmer ?*], alors qu'elle est en réalité un peu plus faible en particulier dans les régions désertiques. Enfin, une augmentation de l'humidité du sol augmente l'inertie thermique du sol, ce qui diminue en retour l'amplitude du cycle diurne de la température du sol et des premiers mètres de l'atmosphère, toutes choses étant égales par ailleurs.

IV.2.2 Réduire les biais classiques dans la simulation du climat de surface par les GCMs

Les études d'inter-comparaison de simulations, comme les exercices CMIP, sont une bonne façon de contextualiser les protocoles et améliorations visés dans ce chapitre.

Ces études d'intercomparaison mettent en évidence des biais persistants sur la simulation du climat à la surface dans les modèles de climat. Quand les GCMs atmosphériques, couplés aux modèles des surface continentales, sont forcés par des températures observées de l'océan (protocole AMIP pour *Atmospheric Model Intercomparison Project*), ils tendent généralement à simuler des températures trop chaudes l'été sur les continents de l'hémisphère nord. Ces biais

sont plutôt réduits en moyenne quand les mêmes GCMs sont couplés à un modèle d’océan, du fait d’une compensation d’erreur. Les modèles couplés ont en effet tendance à prédire des températures océaniques trop froides dans les moyennes latitudes, induisant un refroidissement relatif dans ces régions de biais chauds sur continents (Figure IV.1). La dernière version de référence du modèle LMDZ, LMDZ6A, fait partie des modèles qui montrent un biais chaud relativement modéré l’été, bien réduit par rapport aux configurations antérieures. Cette configuration simule en revanche des températures continentales hivernales trop froides (Figure IV.2), et surestime le rayonnement solaire arrivant à la surface des continents, en particulier en hiver (Cheruy et al., 2020).

Les modèles climatiques surestiment en général les précipitations moyennes globales (sans que puisse être totalement écarté un biais systématique dans les jeux d’observations disponibles), avec des écarts régionaux et intermodèles importants (Boucher et al., 2020).

Si la réduction de ces biais est un moteur pour les équipes développant les modèles couplés, il est important de mener des recherches et développer des stratégies pour réduire ces biais “pour de bonnes raisons”. Plus que la réduction elle-même, ce qui importe avant tout c’est la montée en compréhension des processus impliqués et leur meilleure représentation. C’est seulement à ce prix qu’on obtient, sur le long terme, une amélioration des climatologies des modèles pour de bonnes raisons plutôt qu’au prix de compensations d’erreurs.

Tuner en ciblant le couplage à la surface pourra permettre de quantifier l’erreur paramétrique de ces biais, et donc de savoir à quel point ces biais sont dus au choix des valeurs des paramètres ou à des défauts structurels du modèle.

IV.2.3 Exploiter les données sur site

Les sites instrumentés, appelés parfois super sites d’observations, sont des sites de recherche où sont déployés un grand nombre d’instruments pour observer l’atmosphère et parfois le sol. Les observations y sont recueillies à hautes fréquences et de façon pérenne afin d’observer le climat et le changement climatique. Cette synergie d’observations est conçue pour appuyer la compréhension des processus atmosphériques et de couplage avec la surface des continents à différentes échelles de temps. Les observations de ces sites, et en particulier la synergie d’observations qu’ils apportent, sont souvent considérées comme sous-utilisées au vu des moyens déployés pour les développer et les maintenir dans le temps (Chiriaco et al., 2018). Les raisons évoquées de cette sous-utilisation portent principalement sur la complexité de la comparaison d’observations issues de différents instruments, avec des protocoles de calibration et de contrôle qualité différents, accompagnés de post-traitements parfois complexes, des échelles temporelles différentes etc.

Des raisons intrinsèques à l’observation locale peuvent aussi expliquer la sous-utilisation des observations sur site : à quel point les observations d’un site sont-elles représentatives de l’état atmosphérique environnant ? À quel point les tendances observées sur un site sont-elles le témoin d’un changement climatique ? Ces questions sont aujourd’hui des questions de recherches actives. La question de la représentativité des observations des sites instrumentés pour l’étude des couplages surface-atmosphère est actuellement au cœur de l’ANR MOSAI (Modèles et Observations pour les Interactions entre la Surface et l’Atmosphère, Lohou et al. (2025)).

Une façon de tirer profit de la richesse de ces jeux de données est de les utiliser pour évaluer et améliorer les GCMs, en les utilisant par exemple comme référence de tuning comme proposé ici. On peut distinguer deux grands types d’études comparant sorties de GCMs et observations sur sites. Dans la première on utilise un réseau d’observations sur sites ciblant un processus particulier ou une variable d’intérêt. Dans la seconde, on privilégie la synergie des observations sur un site en particulier.

Les réseaux d’observations sur sites sont créés sur des thématiques précises, afin de partager

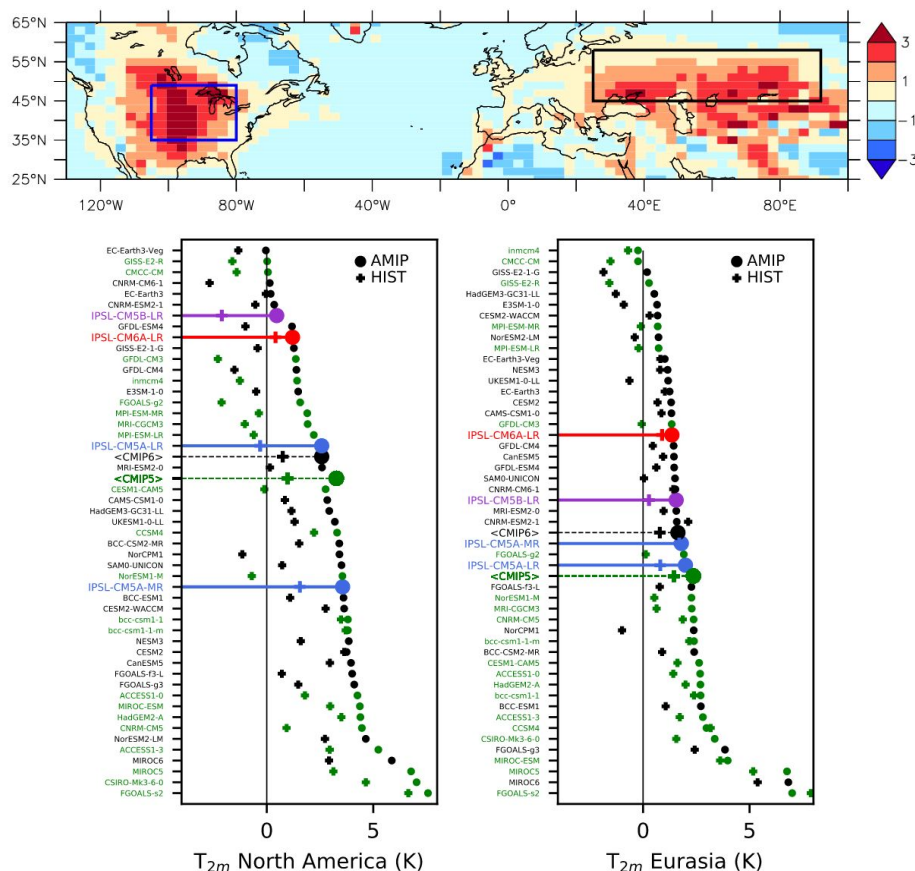


FIGURE IV.1 – Figure issue de Boucher et al. (2020). En haut, la carte montre le biais moyen de température à deux mètres en été dans l'ensemble CMIP6 par rapport aux réanalyses ERA Interim (Dee et al., 2011). Les rectangles tracés sur ces cartes montrent les régions définies pour construire les indices de biais de température nord-américain (105–80°W, 35–50°N) et eurasien (23–92°E, 45–58°N). Ces indices sont tracés sur les deux figures du bas, pour l'ensemble des modèles de CMIP5 en vert et de CMIP6 en noir. Les versions du modèle de l'IPSL de CMIP5 sont tracées en bleu (IPSL-CM5A-LR et IPSL-CM5A-MR) et en violet pour la physique intermédiaire dite 5B (IPSL-CM5B-LR). La version IPSL-CM6A-LR est tracée en rouge. Les moyennes des indices pour les ensembles CMIP5 et CMIP6 sont tracés en pointillés. Les marqueurs rond indiquent les scores des simulations AMIP et les marqueurs en croix indiquent les scores des simulations historiques couplées. La quasi-totalité des simulations AMIP surestiment la température à deux mètres sur ces deux régions. Les simulations historiques associées sont toujours moins chaudes. On voit l'amélioration globales des modèles de CMIP5 à CMIP6, ainsi que l'amélioration du modèle de l'IPSL.

des méthodes de mesure, de post-traitement et de contrôle qualité permettant une comparaison fiable des observations des différents sites. Le réseau FLUXNET (Pastorello et al., 2020) est dédié aux flux turbulents de surface pour l'étude des échanges surface-atmosphère (échanges d'eau, d'énergie et de gaz à effet de serre). En 2020, il rassemblait 212 sites autour du monde. Ces observations sont classiquement utilisées pour l'évaluation et le tuning du modèle ORCHIDEE (Bastrikov et al., 2018; Kuppel et al., 2012; Raoult et al., 2024; Williams et al., 2009). Le réseau BSRN (*Baseline Surface Radiation Network*, Driemel et al. (2018)) regroupe des sites instrumentés mesurant le rayonnement à la surface. Il comptait 59 sites en 2018. Ce réseau est régulièrement utilisé pour la validation de modèles de climat et de données satellites. Utiliser de tels réseaux d'observations comme référence de tuning permet de contraindre le modèle sur une grande diversité de climats, bien que les sites soient majoritairement situés en

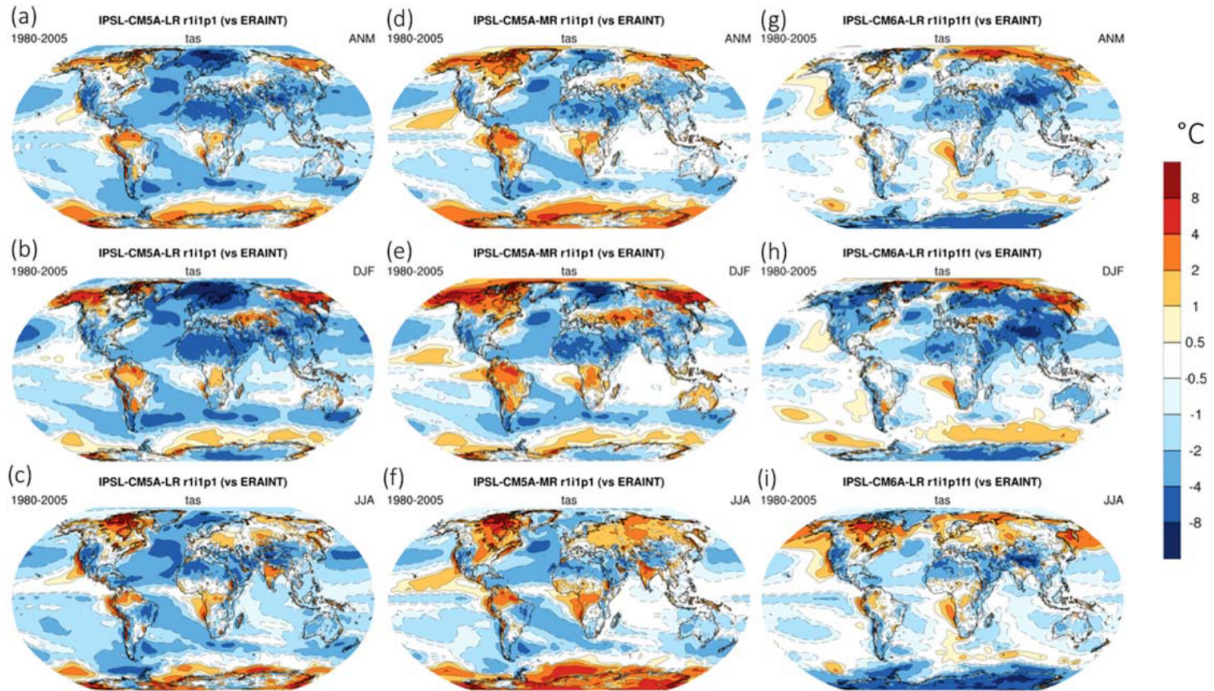


FIGURE IV.2 – Figure issue de Boucher et al. (2020) montre la distribution des biais de température atmosphérique à deux mètres ($^{\circ}\text{C}$) en moyenne annuelle (ANM, a, d, g), en hiver (DJF, b, e et h) et en été (JJA, c, f, i) pour IPSL-CM5A-LR (version pour CMIP5 à basse résolution horizontale), IPSL-CM5-MR (version pour CMIP5 de résolution horizontale identique à IPSL-CM6A-LR) et IPSL-CM6A-LR (version pour CMIP6). Les biais sont tracés par rapport aux réanalyses ERA Interim (Dee et al., 2011)

Europe et en Amérique du Nord. Cependant, le nombre important de sites rend la construction de métriques orientées processus, ainsi que l'estimation de l'incertitude de comparaison simulation-observation, plus difficile.

C'est pourquoi le choix a été fait dans cette thèse de contraindre le modèle en ciblant les observations d'un super-site en particulier ; avant d'utiliser éventuellement des réseaux de mesures pour le tuning du modèle. Se focaliser sur un site d'observation permet d'avoir une approche axée sur la compréhension des processus, en utilisant la synergie d'observations colocalisées mise à disposition par le site.

L'utilisation d'observations de sites instrumentés comme référence pour l'évaluation ou le développement des modèles de climat sur des questions spécifiques est très répandue, que ce soit pour LMDZ ou pour d'autres modèles. Elles peuvent être utilisées pour la création de cas tests 1D/LES, comme par exemple les observations du site ARM des grandes plaines des Etats-Unis qui ont été utilisées pour créer le cas de cumulus continental ARMCU (Brown et al., 2002), ou celles du site de Cabauw pour le cas de couche limite stable GABLS3 (Bosveld et al., 2014). Elles sont fréquemment utilisées pour évaluer la capacité des modèles à simuler des processus spécifiques, sur des périodes plus ou moins longues (par exemple pour les couches limites stables (Baas et al., 2018; Vignon, 2017), les vents catabatiques (?), divers types de nuages (Borella et al., 2025; Coindreau et al., 2007; Morcrette et al., 2012; Wang et al., 2018)), parfois comme appui direct au développement de nouvelles paramétrisations. Elles sont également largement utilisées pour évaluer la simulation des interactions surface-atmosphère (Arjdal et al., 2024; Campoy et al., 2013; Cheruy et al., 2013; Diallo et al., 2017; Phillips et al., 2017). Ces études ciblent souvent l'évaluation d'un modèle en particulier, mais certaines utilisent de telles observations pour évaluer des ensembles CMIP. C'est le cas par exemple de l'étude de Zhang et al. (2018) qui aborde la question du biais chaud continental dans l'ensemble multi-modèles CMIP5, en

utilisant les observations du site instrumenté ARM des grandes plaines Etats-uniennes.

IV.3 Observations du SIRTa et configuration de LMDZ

IV.3.1 Description du site du SIRTa

Le SIRTa (Site Instrumental de Recherche par Télédétection Atmosphérique) est un observatoire de recherche situé au sud-ouest de Paris sur le plateau de Saclay (48,71°N, 2.2°E) dans une zone semi-urbaine. Il a été créé en 1999 et a pour mission d’observer l’atmosphère et le sol à des échelles climatiques, pour le suivi du changement climatique et l’étude des processus physico-chimique de l’atmosphère. Il fait partie des réseaux d’observations internationaux FLUXNET (Pastorello et al., 2020) pour les flux turbulents de surface et BSRN (Driemel et al., 2018) pour les mesures de rayonnement à la surface et accueille régulièrement des campagnes de mesures. Le site s’est légèrement déplacé suite à des travaux d’agrandissement et à la construction d’un bâtiment inauguré en 2021. Une partie des instruments de mesures atmosphériques, qui était placés sur le bâtiment du LMD polytechnique ou sur la pelouse du SIRTa, sont depuis posés sur le toit du bâtiment. Le SIRTa accueille aujourd’hui environ 150 instruments de mesures. La Figure IV.3 montre le site avant et après son déménagement, avec les principaux instruments.

Le produit ReOBS-SIRTa, utilisé dans cette étude, met à disposition un certain nombre de variables observées par les instruments du site du SIRTa, post-traitées de manière cohérente et avec une chaîne de contrôle qualité des données. Ce produit est décrit en détail par Chiriaco et al. (2018).

Le SIRTa est équipé d’un mât de mesure de 30 mètres, remplacé par un mât de 50 mètres depuis 2019. Il est équipé d’un ensemble d’instruments classiques pour observer la météorologie locale : thermomètre, hygromètre, baromètre, anémomètre, girouette, pluviomètre. Thermomètre et hygromètre sont situés à plusieurs niveaux du mât de mesure (1, 2, 5, 10, 20 et 30 m) pour observer les profils verticaux. Le SIRTa est équipé d’instruments permettant d’observer différents types de flux radiatifs à la surface. Les flux radiatifs solaire et infrarouge descendant à la surface sont mesurés avec un pyranomètre et un pyrgéomètre. Le rayonnement solaire direct est également mesuré avec un pyréliomètre. Une estimation des flux solaire et infrarouge ciel-clair est également fournie, permettant de calculer facilement l’effet radiatif des nuages à la surface. Depuis 2007, les flux solaire et infrarouge montant sont également mesurés à 30 mètres au dessus du sol. Les flux de chaleur sensible et latent sont dérivés des mesures de turbulence faites avec un anémomètre sonique et un hygromètre à krypton. La température et l’humidité du sol sont mesurées à 5, 10, 20, 30 et 50 cm de profondeurs depuis 2007 et à 1 mètre de profondeur depuis 2014. Ces mesures ont permis de mettre en évidence la présence d’une nappe phréatique à faible profondeur sur le site du SIRTa (Campoy, 2013), probablement présente sur une grande partie du plateau de Saclay.

IV.3.2 Traitement des données ReOBS et des sorties de simulation

Une chaîne de post-traitement des sorties de simulations a été développée pendant cette thèse. Cette chaîne est pensée pour être adaptée au traitement de gros ensembles de simulations (de l’ordre de 200 par vague de tuning) tout en étant suffisamment souple pour qu’elle puisse être modifiée au besoin. Cette chaîne s’est construite au fur et à mesure de la progression de l’expérience, et en lien avec le développement de scripts de visualisation des ensembles de simulations. Une chaîne de post-traitement du produit ReOBS (Chiriaco et al., 2018), contenant les observations du SIRTa prises comme référence dans ce chapitre, a également été développée. Ces chaînes sont composées de script `python` et `bash`. Les scripts `bash` utilisent les fonctions de `nco` pour la manipulation des fichiers `netcdf` et de `cdo` pour le calcul des moyennes temporelles.



FIGURE IV.3 – (a) Figure issue de Chiriaco et al. (2018) montrant une partie des instruments du SIRTa avant son déménagement. (b) Mât de mesure météorologique actuel de 50 m (IPSL/SIRTa, sd) (c) Etat du site actuel, une partie des instruments sont désormais posés sur le toit du bâtiment (IPSL/SIRTa, sd)

Les deux chaînes de post-traitement sont appliquées à des séries temporelles horaires sur la période de temps désirée (typiquement une année). La chaîne de traitement des sorties de simulations commence ainsi par l'extraction des séries temporelles de la maille contenant le SIRTa. Après le traitement des séries temporelles horaires, les moyennes journalières et mensuelles sont calculées, ainsi que les cycles diurnes mensuels.

Un sous-ensemble de variables présentes dans le produit ReOBS est sélectionné comme variables d'intérêt, sur lequel différents post-traitement vont être appliqué en vue de la comparaison aux sorties du modèle. Tout d'abord, les données horaires de rayonnement solaire arrivant à la surface du produit ReOBS sont mises à zéro lorsqu'elles sont négatives. Ces données négatives correspondent à des moment où le rayonnement solaire est trop rasant pour être mesurés avec précision par le pyranomètre. Quelques variables dérivées sont calculées à partir de ce sous-ensemble du produit ReOBS et des sorties de simulations si nécessaire. On calcule notamment l'effet radiatif des nuages, la somme des flux turbulents, le rayonnement net à la surface, le

rapport de Bowen etc.

Une attention particulière est portée à la gestion des données manquantes. Lors du calcul de nouvelles variables à partir des variables présentes dans le produit ReOBS, les masques des variables utilisées pour le calcul de la nouvelle variable sont toujours concaténés. Pour le calcul de la somme des flux turbulents par exemple, il suffit qu’une valeur soit manquante sur l’un des deux flux pour que la somme des flux soit catégorisée comme une donnée manquante à l’heure considérée. Ensuite, deux masques différents sont appliqués aux sorties horaires des simulations. Le premier masque, appelé “masque-par-variable”, consiste à masquer les sorties horaires des variables d’intérêt dès que la variable correspondante dans le produit ReOBS est manquante. Cela signifie que les séries temporelles des variables d’intérêt n’ont pas le même masque ; ce qui peut poser des problèmes lors de la comparaison de différentes variables ; en particulier pour analyser les bilans d’eau et d’énergie à la surface (Equation IV.1 et Equation IV.2). C’est pourquoi un deuxième masque, appelé “masque-global”, est construit. Il concatène les masques des variables impliquées dans les bilans d’énergie (flux de chaleur sensible et latent, flux radiatifs solaires et infrarouges, descendants et montants, tout ciel et ciel clair) et d’eau (précipitations) ainsi que ceux des variables météorologiques résultants de ces bilans (température de l’atmosphère à deux mètres, humidité spécifique et relative à deux mètres). Il est ensuite appliqué à ces variables, mais également à d’autres variables que l’on souhaite analyser, qu’elles soient observées ou non. Enfin, les sorties d’humidité et de température du sol sont interpolées aux profondeurs où elles sont observées, par une interpolation linéaire classique.

Dans la suite de ce chapitre, il sera toujours précisé si les données utilisées ont été masquées avec le masque “masque-par-variable”, par le masque “masque-global”, où si elles n’ont pas été masquées.

IV.3.3 Une première configuration de LMDZ

Comme expliqué Section IV.1.1, la mise en place de la configuration de LMDZ a été un travail de longue haleine, pavé de nombreux rebondissement. Nous commençons par présenter dans cette section la configuration zoomée (Section IV.3.3.1) et guidée (Section IV.3.3.2), classiquement utilisée pour comparer LMDZ à des observations sur site. Le choix de l’année d’étude et la méthodologie d’équilibrage des modèles sont présentées Section IV.3.3.3. Enfin, la stratégie adoptée pour traiter la question de la représentativité du site est décrite Section IV.3.3.4.

Certains choix effectués sur la configuration de LMDZ à utiliser pour l’expérience de tuning ont varié au cours de la mise en place de cette dernière et de l’étude de la variabilité interne du modèle. Nous avons commencé l’étude en utilisant la dernière “physique”¹ tunée de LMDZ, appelée physique LUDO-L95. Cette physique ajoute à la physique de CMIP6 une paramétrisation de recouvrement verticale de précipitations (voir Section II.3.3.1), et affine la discrétisation verticale du modèle (Figure II.3). Au vu des difficultés rencontrées et du peu de recul dont nous disposons sur cette nouvelle physique, nous avons fait le choix de revenir à la physique standard CMIP6 à 79 niveaux.

Par soucis d’économie de temps humain, d’économie de temps de calcul et de sobriété énergétique, les simulations présentées dans ce chapitre n’ont pas été refaites avec la physique et le guidage finalement choisis. Les simulations présentées dans les différentes sections ne sont donc pas toujours strictement comparables ; sans que ces différences soient trop grandes pour remettre en cause les choix effectués.

1. Par abus de langage, on parle de “la physique” du modèle quand on parle de l’ensemble des paramétrisations physiques ou d’une configuration particulière de ces paramétrisations.

IV.3.3.1 Grille zoomée

LMDZ a la possibilité de fonctionner avec une grille zoomée. C'est d'ailleurs la signification du Z du sigle LMDZ. Cela permet d'avoir des mailles horizontales plus fines sur une région d'intérêt tout en gardant un coût de calcul raisonnable des simulations. Pour ce faire, on définit les coordonnées du centre du zoom, les facteurs de grossissement en latitude et longitude (par rapport à la même grille régulière) et l'extension du zoom (c'est-à-dire la taille de la zone zoomée). La transition entre la zone zoomée et le reste de la grille se fait via une fonction tangente hyperbolique. La raideur de cette transition est également définie par l'utilisateur.

Les pas de temps de calcul doivent être ajustés à la taille de la plus petite maille afin de respecter les conditions CFL (Courant-Friedrichs-Lewy). Ces conditions peuvent être déterminées analytiquement seulement pour des systèmes d'équation dont la solution analytique est connue, ce qui n'est pas le cas des équations primitives utilisées dans LMDZ. Il est néanmoins possible de déterminer l'ordre de grandeur au dessus duquel la discrétisations temporelle n'est plus assez fine compte tenu de la discrétisation spatiale pour garantir la stabilité numérique des schémas utilisés dans la dynamique du modèle. Le critère CFL est alors calculé comme la taille de la plus petite maille divisée par la vitesse de propagation horizontale la plus rapide, pour l'advection des masses d'air et la propagation des ondes. En l'espèce, c'est une onde de gravité équatoriale qui domine ce critère dans LMDZ.

En pratique, lorsque l'on change de grille horizontale, on applique simplement une règle de proportionnalité entre une configuration stable numériquement et celle que l'on souhaite mettre en place; puis l'on teste si le pas de temps choisi ne conduit pas à des instabilités dans la simulation. Ces instabilités sont détectées automatiquement par le modèle lorsque des températures ou des humidités dépassent des seuils considérés comme aberrants; et la simulation est automatiquement arrêtée en cas de dépassement de ces seuils. Plus le zoom est fort et plus la raideur du zoom est importante, plus il est difficile d'avoir des configurations stables numériquement. Des paramètres réglant la dissipation numérique peuvent être modifiés pour stabiliser le modèle. Le guidage, détaillé dans la partie suivante, permet également de stabiliser le modèle et donc d'utiliser un zoom plus fort.

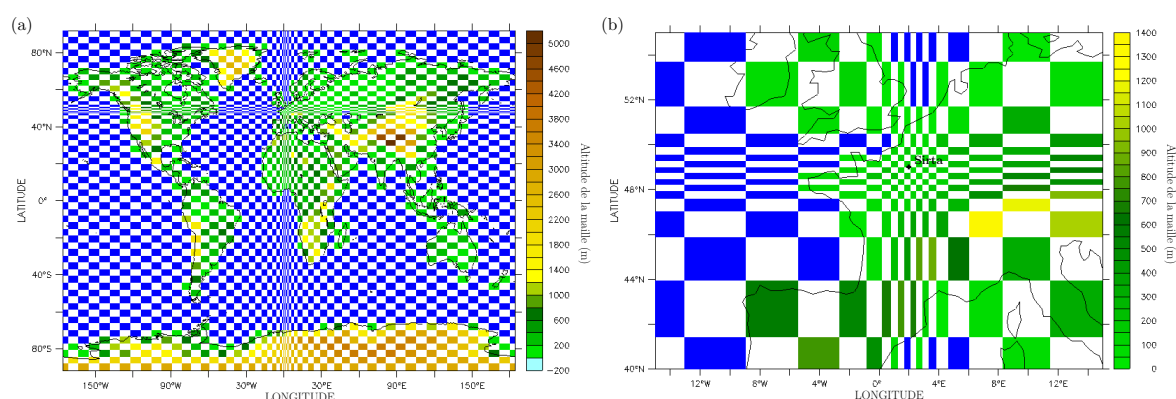


FIGURE IV.4 – Grille utilisée pour les simulations. Une maille sur deux est colorée, soit en bleu pour les mailles océaniques, soit avec une couleur correspondant à l'altitude moyenne de la maille. La figure (a) montre tout le globe, la figure (b) montre le zoom au SIRTa.

Le zoom choisi pour cette étude est tel que le SIRTa soit au centre du zoom, avec une maille de $30 \times 30 \text{ km}^2$ sur le SIRTa et une extension de la région dans le zoom de l'ordre de $300 \times 300 \text{ km}^2$. On souhaite avoir le nombre de point de grille le plus petit possible pour réduire au maximum le coût de calcul des simulations, tout en garantissant la stabilité numérique de la configuration. On souhaite également diminuer la taille de la maille au SIRTa afin de

réduire les incertitudes liées à la localité du site (les tailles des mailles sont de 180x140 km² dans les configurations globales de références actuelles de LMDZ); ceci sans compromettre les hypothèses d'équilibres hydrostatiques utilisées dans les équations de la dynamique, valides jusqu'à typiquement 20x20 km². Afin de garantir la stabilité numérique, la dissipation numérique est augmentée. La grille finale, montrée Figure IV.4, possède 64 points en latitude et en longitude (contre 144x143 dans la version CMIP6 du modèle). Les caractéristiques du zoom sont détaillées Table IV.1.

Nom du paramètre	Description	Valeur
clon	longitude du centre du zoom	2.206
clat	latitude du centre du zoom	48.718
grossismx	facteur de grossissement selon la longitude	13.8181
grossismy	facteur de grossissement selon la latitude	10.472
dzoomx	extension en longitude de la zone du zoom	0.0113077
dzoomy	extension en latitude de la zone du zoom	0.0149208
taux	raideur du zoom en longitude	5.
tauy	raideur du zoom en latitude	5.
daystep	nombre de pas par jour pour la dynamique	4320
tetagdiv	temps de dissipation des plus petites longueurs d'ondes pour u et v (gradient)	400
tetagrot	temps de dissipation des plus petites longueurs d'onde pour u et v (rotationnel)	400
tetatemp	temps de dissipation pour les plus petits longueurs d'onde pour la température	600

TABLE IV.1 – Paramètres du zoom

IV.3.3.2 Guidage

Le guidage consiste à relaxer une ou plusieurs variables d'état du modèle vers une valeur de référence en ajoutant un terme de rappel dans les équations d'évolutions des variables d'état que l'on souhaite guider :

$$\frac{\partial \phi}{\partial t} = \text{Physique}(\phi) + \text{Dynamique}(\phi) - \frac{\phi - \phi_{\text{guide}}}{\tau_{\text{guide}}} \quad (\text{IV.3})$$

avec ϕ la variable d'état guidée, ϕ_{guide} la valeur de référence et τ_{guide} la constante de temps de guidage. Cette dernière indique à quel point le rappel vers la valeur référence est fort. Pour se faire une image de ce qui se passe lors du guidage, on peut résoudre l'Equation IV.3 en ne gardant que la force de rappel, c'est-à-dire résoudre l'équation :

$$\frac{\partial \phi}{\partial t} = - \frac{\phi - \phi_{\text{guide}}}{\tau_{\text{guide}}} \quad (\text{IV.4})$$

La solution de cette équation

$$\phi(t + dt) = \phi(t) e^{-\frac{dt}{\tau_{\text{guide}}}} + \phi_{\text{guide}} (1 - e^{-\frac{dt}{\tau_{\text{guide}}}}) \quad (\text{IV.5})$$

est une moyenne pondérée entre la valeur de la variable d'état au pas de temps précédent et la valeur de référence

Les poids dépendent d’une exponentielle décroissance de temps caractéristique τ_{guide} , fixée par l’utilisateur. Plus τ_{guide} est petit, plus $e^{-\frac{dt}{\tau_{guide}}}$ se rapproche de 0 donc plus la valeur de $\phi(t + dt)$ se rapproche de celle de ϕ_{guide} : on dit que le guidage est fort.

Dans LMDZ, le guidage peut se faire aux choix sur les variables d’états suivantes : les vents zonaux, les vents méridiens, la température et l’humidité relative². Guider les vents zonaux et méridiens du modèle est le guidage le plus classique. Ce dernier permet de reproduire les séquences météorologiques qui se sont effectivement déroulées sur la période d’intérêt, tout en laissant les paramétrisations physiques du modèle s’exprimer. Aux moyennes latitudes, le guidage des vents horizontaux permet par exemple d’imposer la séquence d’anticyclones et de dépressions qui régit une grande partie du temps la variabilité atmosphériques d’échelles dites synoptiques (de temps caractéristiques 4 jours). Le guidage en vents est ainsi largement utilisé pour la comparaison de LMDZ à des observations sur des périodes de temps inférieures aux échelles climatiques, ou pour différencier le biais causé par la dynamique du modèle de celui causé par sa physique. Les constantes de guidage peuvent avoir des valeurs différentes à l’extérieur et à l’intérieur du zoom. Il est classique d’avoir un guidage fort à l’extérieur du zoom, où la résolution est dégradée, et d’avoir un guidage faible voire insignifiant dans la zone zoomée.

Dans la configuration initiale utilisée dans cette étude, seuls les vents horizontaux sont guidés. Leur constante de temps est prise à 3h à l’extérieur du zoom et à 12h à l’intérieur ; ce qui correspond à des guidages forts. Ces valeurs sont à mettre au regard des temps caractéristiques des processus atmosphériques d’intérêt, de l’ordre de quelques heures dans la couche limite atmosphérique. Ces valeurs ont été choisies afin de garantir la stabilité numérique de la configuration. Dans la configuration finalement choisie, les températures et humidités sont également guidées (voir Section IV.4.2), ce qui a permis d’augmenter jusqu’à 100 jours les constantes de guidage des vents horizontaux à l’intérieur de la zone zoomée (Table IV.4) tout en garantissant la stabilité numérique de la configuration.

Le produit de référence utilisé pour guider le modèle est le produit ERA5 (Hersbach et al., 2020). C’est la cinquième génération de réanalyses produites par le Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT, ECMWF en anglais). Les analyses sont les états initiaux utilisés pour faire les prévisions météorologiques. Elles sont actuellement produites au CEPMMT avec des méthodes d’assimilation variationnelle qui corrigent l’état initial des simulations (état final d’une simulation du jour précédent par exemple) pour que la prévision numérique (simulation effectuée à partir de cet état initial) passe au plus près des observations recueillies pendant typiquement la journée écoulée. Les réanalyses sont des analyses faites à posteriori, en utilisant une version de modèle et une méthode d’assimilation fixe pour toute la période réanalysée. Cela permet d’avoir un produit cohérent sur une grande période de temps.

IV.3.3.3 Durée des simulations, spin-up et choix de l’année

Puisque les couplages entre atmosphère et surface des continents sont particulièrement importants pour la simulation des échelles saisonnières, nous choisissons de faire des simulations d’une durée de un an.

Le temps d’équilibrage (ou spin-up) des modèles dépend du temps caractéristique des processus physiques modélisés le plus long. La dynamique atmosphérique du modèle étant guidée et l’humidité et la température initiale du sol étant donnés, le modèle d’atmosphère a besoin de seulement quelques jours pour s’équilibrer. L’hydrologie du sol a quant à elle besoin de plusieurs années pour que ses réservoirs se mettent à l’équilibre, même avec une atmosphère imposée.

2. Pour l’humidité, plus précisément, on lit les humidités relatives et on guide l’humidité spécifique du modèle avec l’humidité à saturation du modèle multipliée par l’humidité relative (clé `guide_hr=true` dans le code) ce qui revient fondamentalement à un guidage en humidité relative.

C'est la composante de l'hydrologie du sol qui détermine donc ce temps de spin-up. Pour les zones tempérées comme celle autour du SIRTa, on considère généralement que trois ans de spin-up suffisent à équilibrer les modèles. En théorie, à chaque changement de configuration (incluant des changements de paramètres libres) trois ans de spin-up doivent être réalisés, ce qui alourdit considérablement le coup de calcul du tuning. C'est pourquoi, un seul spin-up long (de plus de trois ans) est effectué avec la configuration de référence du modèle. Les simulations aux paramètres perturbés sont ensuite réalisées sur un an, en utilisant comme état initial celui créé par la longue simulation équilibrée. Les premiers mois de la simulation sont considérés comme spin-up. On fait débiter ces simulations en janvier car c'est un mois où l'humidité du sol est proche de la saturation au SIRTa, ce qui diminue le temps d'équilibrage nécessaire lorsque les paramètres sont perturbés. Avoir une humidité du sol proche de la saturation aux mois de décembre et janvier est un critère pour choisir l'année de la simulation.

La disponibilité des données d'observation du SIRTa est notre premier critère pour choisir l'année de simulation. Le SIRTa a commencé ses observations en 2003. Cependant, l'humidité du sol a commencé à être observée en 2006, avec un changement de sonde en 2014. Les flux radiatifs montant sont disponibles à partir de 2012 et les flux turbulents à partir de 2015 avec une période de données manquantes de plusieurs mois en 2017. Compte tenu de la disponibilité des observations, l'année est choisie parmi les années 2015, 2016, 2018, 2019 et 2020.

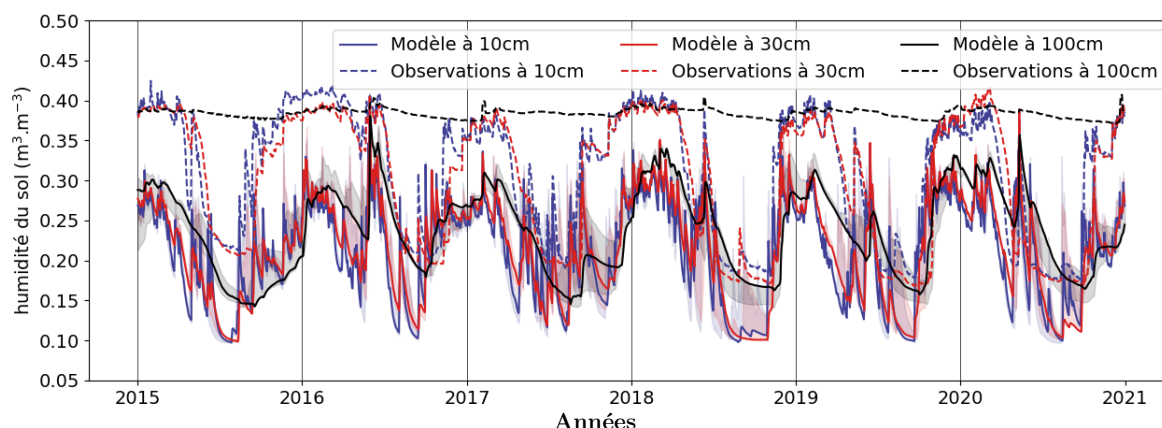


FIGURE IV.5 – Evolution de l'humidité du sol (en m^3 d'eau par m^3 de sol) en moyenne journalière sur la période 2015-2020 pour trois profondeurs correspondant à la profondeur de trois sondes : en bleu à 10 cm, en rouge à 30 cm et en noir à 100 cm. Les observations sont tracées en pointillés et les sorties du modèle, interpolées sur les niveaux des sondes, sont tracées en trait plein. L'humidité à saturation vaut $0,43 \text{ m}^3.\text{m}^{-3}$ dans le modèle et varie entre $0,37$ et $0,43 \text{ m}^3.\text{m}^{-3}$ dans les observations (Campoy, 2013). L'étendue de l'ensemble des 5 meilleures simulations faites avec la physique LUDO-L95 est montrée en clair, avec la meilleure simulation en trait plein. Les simulations sont post-traitées avec le masque-par-variable.

Un ensemble de simulations est réalisé sur la période 2014-2020 afin de choisir l'année de simulation. Cette ensemble est constitué des cinq meilleures simulations issues du tuning de la physique LUDO-L95 (Touzé-Peiffer, 2021). Les humidités du sol de ces simulations sur la période 2015-2020 sont tracées Figure IV.5. Les humidités simulées par le modèle sont systématiquement trop faibles, en particulier l'hiver où le sol est proche de la saturation dans les observations. L'absence de cycle saisonnier de l'humidité du sol observée à 1 m s'explique par la présence d'une nappe perchée au SIRTa, dont la profondeur est estimée par (Campoy et al., 2013) entre 0,5 et 1,3 m de profondeur. Puisqu'elle n'est pas modélisée, il est normal que les humidité du modèle à 1 m soient inférieure à celles observées, avec un cycle saisonnier marqué. Les années 2017, 2019 et 2020 sont rejetées car l'humidité du sol observée n'est pas assez proche de la saturation au mois de décembre ou janvier. L'hiver 2018-2019 présente de plus des variations

d'humidité caractéristiques de périodes d'assèchement, dont la sensibilité aux paramètres libres de l'hydrologie d'ORCHIDEE ont été étudiés, par exemple par Raoult et al. (2021).

Parmi les années restantes, c'est l'année 2018 qui est retenue ; tout d'abord pour sa séquence d'humidité du sol. Elle présente un premier assèchement de printemps, puis une recharge d'eau très rapide fin mai-début juin 2018, causée par une période de forte précipitation, et suivie enfin d'un assèchement du sol exceptionnel et prolongé jusqu'à l'hiver (Météo-France, 2019). Cibler une année avec de fortes variations d'humidité des sols pourrait être intéressant pour évaluer et contraindre cette dynamique saisonnière dans les simulations. L'année 2018 a été une année exceptionnellement orageuse (l'année la plus foudroyée des trente dernières années). Les précipitations du 11 juin 2018 ont atteint le record de la journée la plus pluvieuse des mois de juin de la période 1991-2020 à la station Paris Montsouris (Météo-France, 2025b). La longue période orageuse, de fin mai à mi-juin 2018, pourrait être intéressante pour contraindre les paramètres libres de la paramétrisation de convection profonde de LMDZ, qui ne sont actuellement pas préconditionnés. Cet épisode est suivi par une séquence d'une quinzaine de jours où des cumulus de beau temps se développent en journée, typique des couches limites convectives.

L'année 2018 a également été une année très chaude, la plus chaude du XX^e siècle à cette date pour la France métropolitaine, mais depuis détrônée par les années 2020, 2022, 2023 et ex aequo avec l'année 2024 (Météo-France, 2023, 2025a). Une canicule importante a lieu du 24 juillet au 8 août 2018, comparable en longueur à celle de 2003, mais d'une moindre intensité (Météo-France, 2023). C'est néanmoins la seule année, parmi les années restantes, à avoir eu un épisode caniculaire important. Cibler une année présentant plusieurs extrêmes climatiques, et en particulier des événements extrêmes chaud, est une manière de s'assurer que le modèle est capable de les simuler correctement, ou au moins d'évaluer la capacité du modèle à les simuler. La capacité du modèle à simuler correctement des événements extrêmes est important au vu de leur impact sociétal, et de l'augmentation de leur occurrence et intensité avec le changement climatique.

IV.3.3.4 La question de la représentativité des observations du site du SIRTa

La question de la représentativité d'observations locales consiste à se demander dans quelle mesure une observation réalisée à un endroit donné représente bien ce qui se passe effectivement dans l'environnement de l'observation, et peut être comparé avec fiabilité à des sorties de simulations. La représentativité d'un site est différente suivant la variable qui est comparée, l'échelle spatio-temporelle d'intérêt, et les processus modélisés. Ici, nous comparons les sorties de simulation de LMDZ ayant un pas de temps de 15 min et une maille d'aire 30x30 km² sur le site du SIRTa.

La première stratégie qui a été adoptée pour accroître la fiabilité de la comparaison observation-simulation est de diminuer la taille de la maille que l'on va comparer avec les observations du SIRTa. Il est en effet plus facile d'avoir des observations locales comparables à des variables simulées pour une portion d'atmosphère de 30x30 km², plutôt que pour une portion d'atmosphère de 180x140 km² (taille de la maille du SIRTa de la grille standard).

Plus les observations sont situées dans un environnement homogènes, plus elles sont généralisables. Le site du SIRTa est situé dans un environnement peri-urbain hétérogène, avec une végétation mixte. Les mesures sont prises sur une pelouse bordée d'arbres alignés, ce qui peut significativement affecter le sens et la force des vents, ainsi que la turbulence des premières dizaines à centaines de mètre de l'atmosphère. Les vents, observés à 10 m, ne seront donc pas jugés comme des variables très fiables, et pourront n'être regardées qu'à titre indicatif. Les flux turbulents pourraient également être affectés par cette caractéristique du site.

Un petit lac, non modélisé, est présent à proximité du site, ce qui peut affecter localement l'hydrologie. De plus, comme on l'a déjà mentionné plus haut, le SIRTa est situé au dessus d'une

nappe perchée, probablement présente sur une grande partie du plateau de Saclay (Campoy, 2013). Campoy et al. (2013) montre que le manque d'évaporation des simulations au niveau du SIRTa en été, déjà mis en évidence par Coindreau et al. (2007), pourrait s'expliquer par la présence de cette nappe perchée, qui n'est pas modélisé dans la configuration standard de LMDZ. Les mesures d'humidité du sol et du flux d'évaporation (ou de chaleur latente) du SIRTa ne sont donc pas directement comparables avec les simulations de LMDZ. Le flux de chaleur sensible, très dépendant du flux de chaleur latent, n'est pas non plus considéré comme comparable. De plus, l'influence d'un environnement hétérogène sur les flux turbulents n'est pas clairement quantifiée. Les flux turbulents sont donc écartés des variables que l'on pourra cibler lors du tuning au SIRTa, et devront être analysé en prenant en compte ces éléments.

Les mesures de rayonnement solaire sont également influencées par la localité du site. Le rayonnement solaire descendant est d'une part affecté par l'albédo du site de mesure via la retrodiffusion du rayonnement solaire montant, et d'autre part par la présence d'aérosols dans l'atmosphère, dont la variabilité spatiale peut être très importante. Pour s'affranchir de l'albédo spécifique du site du SIRTa, ce dernier est diagnostiqué à partir des observations du SIRTa puis imposé sur la maille du SIRTa dans LMDZ. Ce travail est présenté Section IV.3.3.5.

Pour les aérosols, le modèle climatique LMDZ est généralement forcé par des "climatologie" précalculées, qui ne rendent pas compte des variations synoptiques du contenu en aérosol, en fonction notamment de l'origine des masses d'air. Par souci de simplicité, le choix est fait ici de ne pas les inclure dans les simulations de LMDZ, et d'instruire la sensibilité du rayonnement aux aérosols afin de la prendre en compte dans les incertitudes de comparaison modèle-observations (voir Section IV.3.3.6).

IV.3.3.5 Diagnostique de l'albédo au SIRTa et forçage de l'albédo dans LMDZ

Pour limiter les problèmes de représentativité de l'albédo de surface, nous avons considéré uniquement le rayonnement solaire descendant arrivant à la surface comme métrique de tuning. Cependant, au second ordre, même ce rayonnement descendant est sensible à l'albédo de surface, car le rayonnement solaire réfléchi par la surface est lui même rétro diffusé vers la surface par le gaz atmosphérique (diffusion Rayleigh) et les particules, aérosols ou condensats liquide ou solide (diffusion de mie pour des particules sphériques homogènes). En présence de nuages notamment, cette partie rétro-diffusée vers la surface peut être significative et il convient donc d'essayer d'utiliser un albédo adapté au site du SIRTa. C'est pourquoi, nous mettons en place méthodologie pour diagnostiquer l'albédo à partir des observations du site du SIRTa, puis nous l'imposons dans la maille correspondant au SIRTa dans LMDZ.

L'albédo α est défini comme le rapport du flux solaire réfléchi par le flux solaire arrivant à la surface :

$$\alpha = \frac{SW_{up}}{SW_{down}} \quad (IV.6)$$

Modélisation de l'albédo dans LMDZ

L'albédo est modélisé dans LMDZ comme une propriété de la surface, avec l'Equation IV.7, pour deux gammes spectrales (visible et proche-infrarouge).

$$\alpha_{modele} = f_{neige} \cdot \alpha_{neige} + (1 - f_{neige}) \cdot (f_{solnu} \cdot \alpha_{solnu} + \sum_{i=1}^{nPFT} f_{veget} \cdot \alpha_{veget}) \quad (IV.7)$$

C'est une moyenne de l'albédo du sol nu α_{solnu} , de la végétation α_{veget} et de la neige α_{neige} pondérée par les fractions de la maille occupées par chacune des trois composantes (f_{neige} ,

f_{solnu} et f_{veget}). Les fractions de sol nu et de végétation évoluent en fonction de la croissance des plantes, via l'indice de surface foliaire (LAI pour *Leaf Area Index*). L'albédo du sol nu est une constante par type de sol, et celui de la végétation est une constante par type de plante (appelées PFT *Plant Functional Type*). L'albédo de la neige dépend quant à lui de l'âge de la neige. L'hypothèse est faite dans le modèle que la neige recouvre uniformément tous les types de surfaces (c'est-à-dire le sol nu et les différentes PFT). Parce que ce qu'il est considéré comme une propriété de la surface, l'albédo modélisé dans LMDZ ne dépend pas de l'angle solaire. La dépendance de l'albédo du sol nu à l'humidité du sol n'est pas active dans la version de LMDZ utilisée dans cette étude.

Diagnostic de l'albédo au SIRTa

On souhaite imposer un albédo sur la maille du SIRTa (1) qui ne dépend pas de l'angle solaire (2) qui garde la rétroaction de la neige. Le diagnostic l'albédo au SIRTa est fait en conséquence.

On utilise l'Equation IV.6 pour diagnostiquer l'albédo au SIRTa à partir des données horaires des flux radiatifs solaires montant et descendant. En l'état, l'albédo présente un cycle diurne, particulièrement marqué les jours de ciel-clair (Figure IV.6); avec un minimum en milieu de journée. Les premières et dernières valeurs de la journées ne sont pas jugées significatives car elles sont le rapport de deux valeurs faibles, et que l'incertitude de mesure augmente fortement avec l'angle solaire.

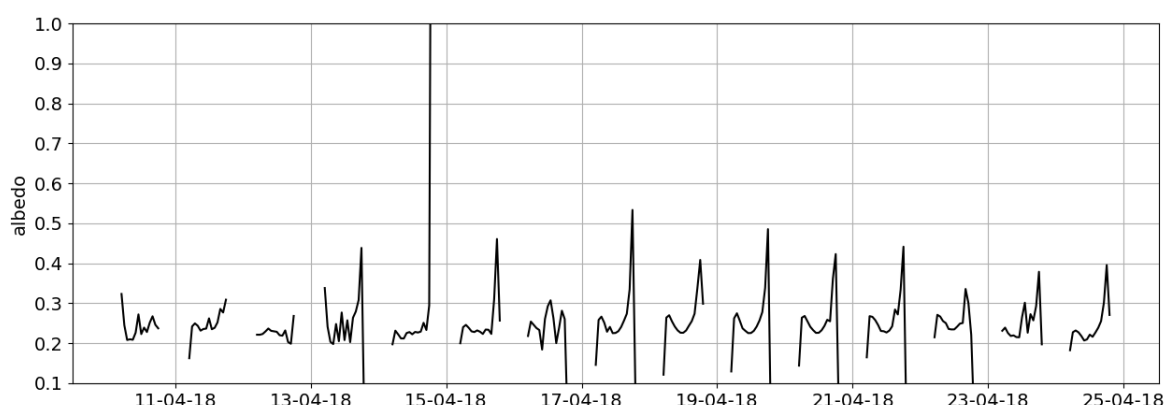


FIGURE IV.6 – Evolution de l'albédo du 10 au 25 avril 2018. L'albédo est diagnostiqué sur les données horaires du SIRTa avec l'Equation IV.6. On note les variations diurnes de l'albédo, avec un cycle diurne particulièrement marqué pendant la période de ciel-clair du 17 au 21 avril.

Pour s'affranchir de ce cycle diurne, un albédo est calculé par jour à partir des albédos horaires de milieu de journée, considéré comme plus représentatif d'un albédo caractéristique de la surface. La Figure IV.7 (a) montre la sensibilité de cet albédo journalier au choix des heures retenues pour le calcul de l'albédo : de 10h à 15h, de 11h à 14h, de 12h à 13h ou sur toute la journée. Les données ont préalablement été filtrées en éliminant toutes les valeurs d'albédo horaires supérieures à 1. C'est la plage horaire 11h-14h qui est retenue. L'augmentation de l'albédo en février (Figure IV.7) est due à un épisode de neige, qui a persisté pendant quelques jours.

On souhaite garder la rétroaction de la neige sur l'albédo. Pour ce faire, on élimine du calcul de l'albédo journalier toutes les valeurs d'albédo horaires au-dessus d'un certain seuil. La Figure IV.7 (b) montre l'albédo mensuel diagnostiqué pour des seuils de 0,4, 0,6, 0,8 et 1. L'albédo du mois de février est très affecté par le seuil choisi sauf pour les seuils 0,4 et 0,6. Cela signifie que le seuil 0,6 suffit pour éliminer les albédos affectés par la présence de neige. Les

différences sur les autres mois de l'année sont minimales. Hors période de neige, les variations de l'albédo au cours de l'année sont assez lentes : imposer une valeur d'albédo par mois suffit.

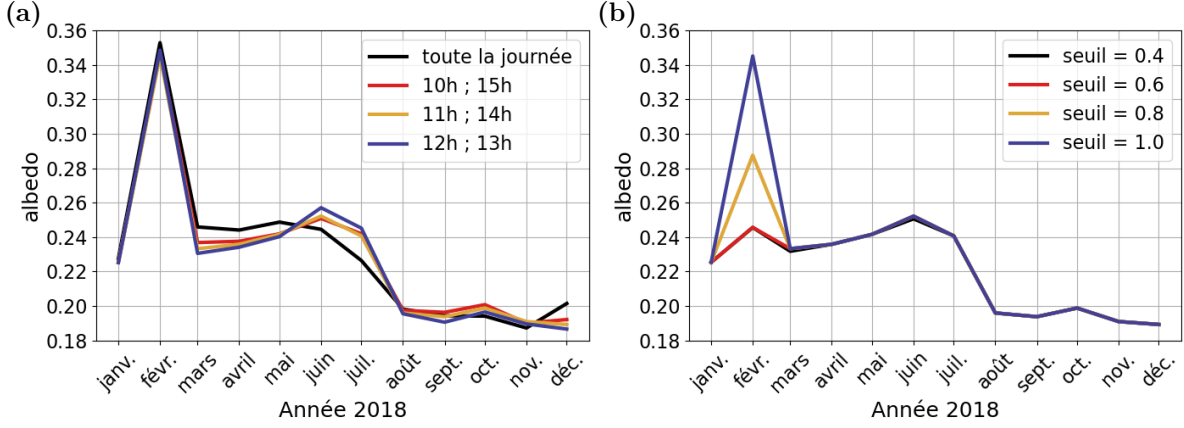


FIGURE IV.7 – Evolution de l'albédo en moyenne mensuelle sur l'année 2018, diagnostiqué à partir des observations du SIRTa (a) Test de sensibilité à la plage horaire (b) Test de sensibilité au seuil à partir duquel l'albédo horaire est exclu du calcul

Simulations avec un albédo forcé sur la maille du SIRTa

L'albédo mensuel diagnostiqué au SIRTa, α_{SIRTa} , est imposé sur la maille du SIRTa dans LMDZ, en gardant la rétroaction de la neige. L'Equation IV.7 devient alors :

$$\alpha_{modele} = f_{neige} \cdot \alpha_{neige} + (1 - f_{neige}) \alpha_{SIRTa} \quad (IV.8)$$

La même valeur de l'albédo diagnostiqué au SIRTa est utilisé pour le calcul de l'albédo dans le visible et le proche infra-rouge. Ainsi, l'albédo simulé pour les gammes spectrales du visible et du proche-infrarouge seront les mêmes, exceptés lors des épisodes de neige.

La Figure IV.8 montre les albedos simulés sur la maille du SIRTa lorsque l'albédo est imposé avec l'Equation IV.8 et lorsqu'il est calculé par le modèle. L'albédo imposé est significativement plus grand que celui calculé par le modèle dans le visible (de 0,15 à 0,2 de différence selon les périodes) et à peu près égal dans le proche IR. On voit l'effet de la neige sur l'albédo en février, mais aussi de manière très brèves à d'autres moment de l'année.

Ce changement d'albédo a une incidence significative sur la simulation des flux radiatifs solaires, comme le montre la Figure IV.9. Comme attendu, le flux solaire montant à la surface augmente dans la simulation où l'albédo est imposé, car celui-ci est systématiquement plus grand que l'albédo simulé par le modèle. Ce flux, qui était sous-estimé avec un albédo simulé, est désormais systématiquement surestimé en moyenne mensuelle par rapport aux observations du SIRTa (Figure IV.9 (a)). Cela s'explique au premier ordre par la surestimation du flux solaire descendant à la surface (Figure IV.9 (b)). Imposer l'albédo du SIRTa conduit à une augmentation du flux solaire descendant (Figure IV.9 (b)), dont une partie s'explique par l'augmentation du flux solaire ciel-clair (non montré), directement lié à la rétrodiffusion du rayonnement solaire montant. L'autre partie est due à la rétroaction du changement de rayonnement sur la simulation de l'état atmosphérique (notamment des nuages) et à la variabilité interne du modèle (discutée Section IV.4). En conséquence, le flux solaire net à la surface diminue jusqu'à 20 W.m^{-2} en juin (Figure IV.9 (c)), et le flux radiatif (solaire + infrarouge) net à la surface diminue jusqu'à une quinzaine de W.m^{-2} (Figure IV.9 (d)); réduisant le biais du modèle par rapport au SIRTa.

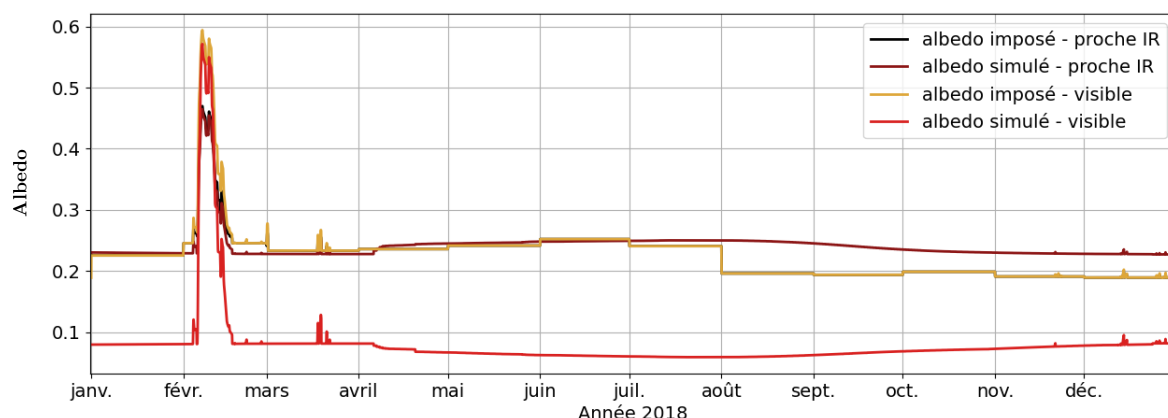


FIGURE IV.8 – Evolution des albédos dans le visible (courbes rouge et jaune) et proche IR (courbes rouge foncée et noire) sur l'année 2018 en fréquence horaire. En rouge et rouge foncée dans une simulation sans albédo imposé et en jaune et noir dans une simulation où l'albédo diagnostiqué au SIRTa est imposé sur la maille du SIRTa. La courbe noire est uniquement visible pendant l'épisode de neige de février, et est superposée à la courbe jaune le reste du temps, car l'albédo imposé est le même dans les bandes spectrales visibles et proche-infrarouge, sauf pour la neige. Les simulations sont faites avec la physique CMIP6-L79, un guidage en vent, température et humidité (voir Section IV.4.2). Les données ne sont pas masquées.

IV.3.3.6 Sensibilité des simulations aux aérosols

Un test de sensibilité aux aérosols est effectué, en faisant une simulation avec des aérosols climatiques, c'est à dire en lisant et en interpolant dans le temps 12 champs 3D (un par mois) de concentrations en différents aérosols. Ces champs, utilisés dans les simulations CMIP6, sont obtenus en moyennes saisonnières sur une décennie à partir de simulations réalisées il y a plusieurs années avec le modèle LMDZ couplé au modèle de chimie-aérosols Inca. Ces aérosols sont pris en compte dans le transfert radiatifs. Ce test est effectué sur la configuration finale du modèle, résumé Section IV.7, utilisant la physique CMIP6-L79 et avec un albédo imposé sur la maille du SIRTa. Un ensemble perturbé aux conditions initiales (dit PISE, voir Section IV.4.1), fait sur la configuration sans aérosol, est également tracé afin d'évaluer la significativité des différences. Si la simulation avec aérosol sort de l'ensemble perturbé aux conditions initiales, la différence est considérée significative.

Les aérosols réduisent significativement le flux solaire ciel-clair descendant à la surface (entre 5 W.m^{-2} et 7.5 W.m^{-2} , Figure IV.10 (a)), ce qui est cohérent avec le fait que les aérosols opacifient l'atmosphère. Cette diminution se repercute sur le flux solaire descendant à la surface et sur l'effet radiatif des nuages dans le solaire. Le changement d'aérosol provoque une diminution significative du flux radiatif net à la surface certains mois, d'au plus 5 W.m^{-2} (Figure IV.10 (b)).

Si la courbe sort de la plage des simulations de sensibilité à l'état initial, elle restent cependant comparables à celles ci en termes de biais par rapport aux observations.

IV.4 Variabilité interne

Les systèmes météorologique et climatique comportent une forte variabilité naturelle à toutes les échelles de temps et d'espace. En forçant les températures de surface de l'océan plutôt que d'utiliser un modèle de circulation océanique pour les calculer, on élimine la variabilité dite "couplée", celle des grandes échelles de temps qui vient principalement de l'océan. Il reste alors

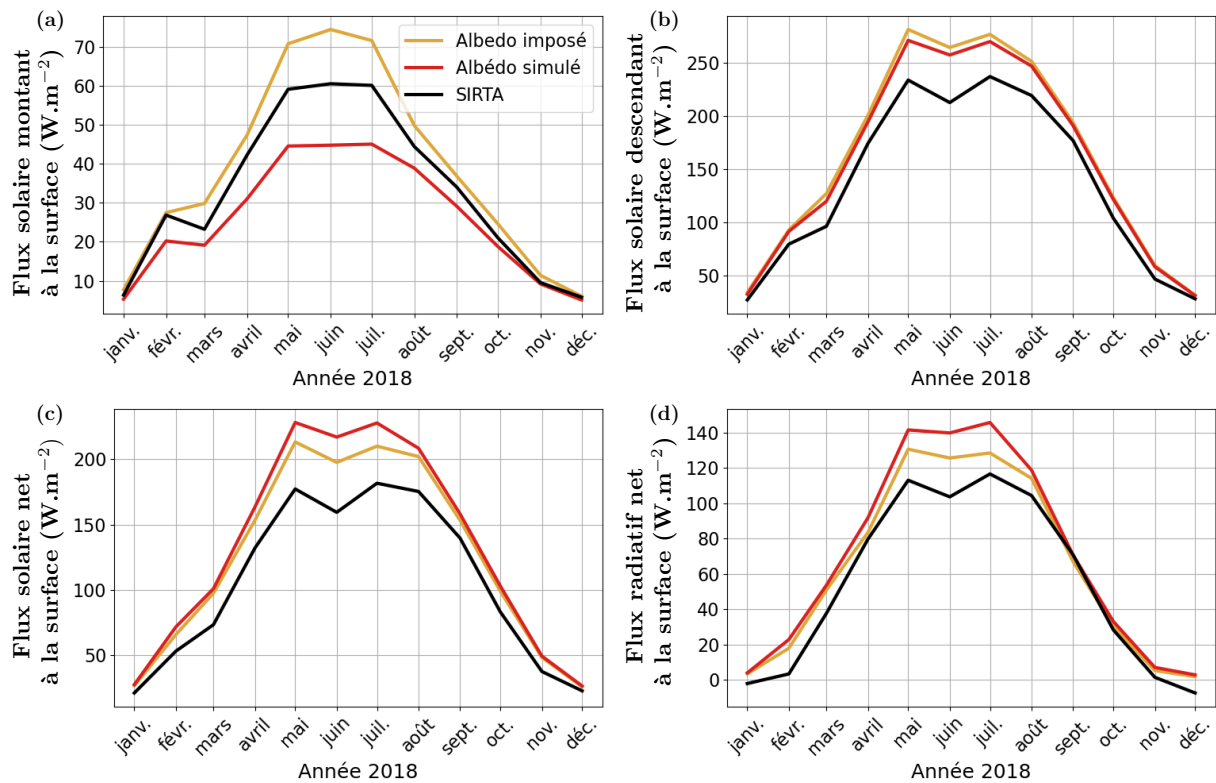


FIGURE IV.9 – Evolution en moyenne mensuelle sur l'année 2018 (a) du flux solaire montant à la surface (b) du flux solaire descendant à la surface (c) du flux solaire net à la surface (d) du flux radiatif totale (solaire + infrarouge) net à la surface. En noir les observations du SIRTa, en rouge la simulation sans albédo imposé et en jaune la configuration où l'albédo est imposé sur la maille du SIRTa. Les simulations sont faites avec la physique CMIP6-L79, un guidage en vent, température et humidité (voir Section IV.4.2). Les données tracées utilisent le masque-par-variable.

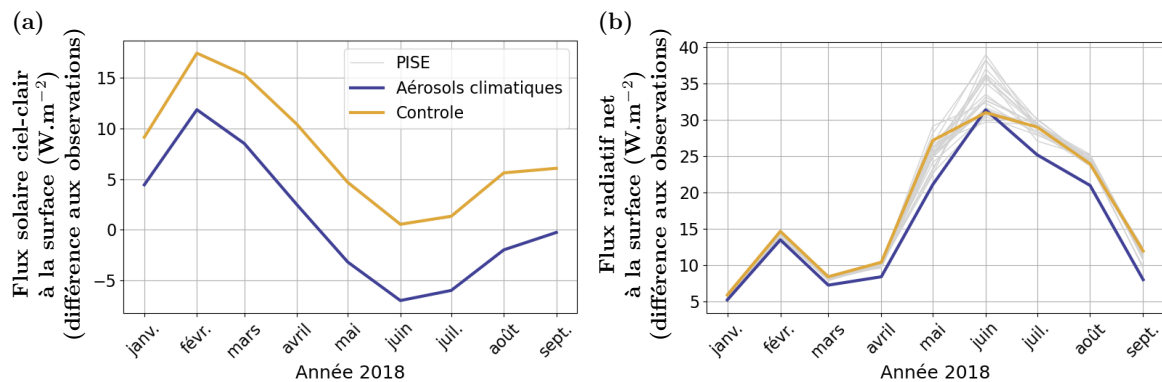


FIGURE IV.10 – Evolution en moyenne mensuelle de janvier à septembre 2018 (a) du flux solaire ciel-clair descendant à la surface (b) du flux radiatif totale (solaire + infrarouge) net à la surface. Les courbes sont des différences avec les observations du SIRTa, en jaune pour la simulation de contrôle avec albédo imposé et sans aerosol, et en bleu pour la même simulation avec des aérosols climatiques. Les courbes gris clair montre l'ensemble aux états initiaux perturbés correspondant à la simulation de contrôle (voir Section IV.4.1). Il n'est pas visible sur la figure (a) car il n'y a pas de variabilité interne sur le flux solaire ciel-clair descendant à la surface. Les simulations sont faites avec la configuration finale du modèle (voir Section IV.7), et les données tracées sont post-traitée avec le masque-par-variable.

une variabilité, dite interne à l’atmosphère, ou en fait au système couplé atmosphère-surfaces continentales, qu’il est beaucoup plus difficile de découpler. Cette variabilité est notamment liée au caractère chaotique de l’écoulement atmosphérique, celui là même qui limite l’horizon temporel des prévisions du temps à une quinzaine de jours. L’importance relative de cette variabilité varie selon les régions et les variables regardées. Par exemple le flux global infra-rouge sortant au sommet de l’atmosphère est beaucoup moins variable que les températures d’été sur la France, ou encore que les pluies de moussons au Sahel au mois de Juillet.

Dans les expériences de tuning avec le modèle climatique global forcé, on évalue la contribution de cette variabilité interne soit en effectuant une simulation climatique longue avec des forçages constants, et en estimant par exemple la variabilité inter-annuelle d’une moyenne annuelle ou mensuelle, soit en réalisant un ensemble de simulations aux conditions initiales perturbées, appelé ici PISE pour *Perturbed Initial State*, et en calculant la dispersion des membres.

En guidant le modèle en vent, on espérait s’affranchir d’une grande part de cette variabilité interne du système atmosphérique. Comme expliqué plus haut, la sensibilité aux conditions initiales des simulations au SIRTa, présentée Section IV.4.1, a en fait été totalement sous évaluée au début de ce travail, et les tests de sensibilité de la variabilité interne au guidage présenté Section IV.4.2, n’ont pas permis de la réduire.

IV.4.1 Ensemble aux états initiaux perturbés

La quantification de la variabilité interne atmosphérique dans la configuration au SIRTa est évaluée dans cette étude avec des ensembles aux états initiaux perturbés (ou PISE). Ils seront également utilisés dans le chapitre suivant pour évaluer la sensibilité des métriques aux états initiaux.

Ces ensembles sont produits en changeant l’état initial sur un ensemble de simulations, toute chose étant égale par ailleurs. En pratique, une première simulation d’une vingtaine d’année est réalisée avec la configuration de contrôle, mais sans l’albédo imposé car les données de rayonnement montant depuis la surface ne sont pas disponible sur une assez grande période. Les états initiaux des premiers janvier de chaque année de cette longue simulation sont utilisés comme état initial pour simuler l’année 2018, sauf ceux des trois premières années car on considère ces années comme un spin-up. On a ici différencié l’état initial du modèle atmosphérique de celui du modèle de surface continentale, afin d’évaluer si certaines variables sont particulièrement sensibles à l’état initial du sol, et pourrait donc être significativement affectée par le fait que le sol n’est pas rééquilibré à chaque changement de vecteur de paramètre libre dans l’expérience de tuning.

Un premier PISE est réalisé avec la physique LUDO-L95, un guidage en vent (Table IV.2) et sans albédo imposé sur la maille du SIRTa. 10 simulations sont réalisées en changeant les états initiaux atmosphériques et 10 autres en changeant les états initiaux du sol. La figure Figure IV.11 montre le résultats de ces PISE pour quelques variables clé.

La faible dispersion des mois de janvier à mars sur la plupart des variables est ce à quoi on s’attendait initialement : les simulations étant guidée par les réanalyses, l’état initial est vite oublié (en quelques heures ou quelques jours), et la séquence des situations météorologiques imposée par le guidage suffit à contraindre les températures, l’humidité, les flux radiatifs, etc.

L’état initial du modèle ORCHIDEE aurait pu avoir une influence à plus longue échéance, au vu des temps caractéristiques d’équilibrage de l’hydrologie du modèle. Effectivement, l’humidité du sol en profondeur met plusieurs mois à s’équilibrer, mais l’humidité du sol de surface (20 premier centimètres du sol) est équilibrée dès le mois de février (Figure IV.11 (d)), et c’est cette dernière qui est le plus fortement couplée avec l’atmosphère. La sensibilité des variables atmosphériques à l’état initial du sol est du même ordre que leur sensibilité à l’état initial de l’atmosphère dès le mois de février. Cela justifie a posteriori le choix de commencer les simulations

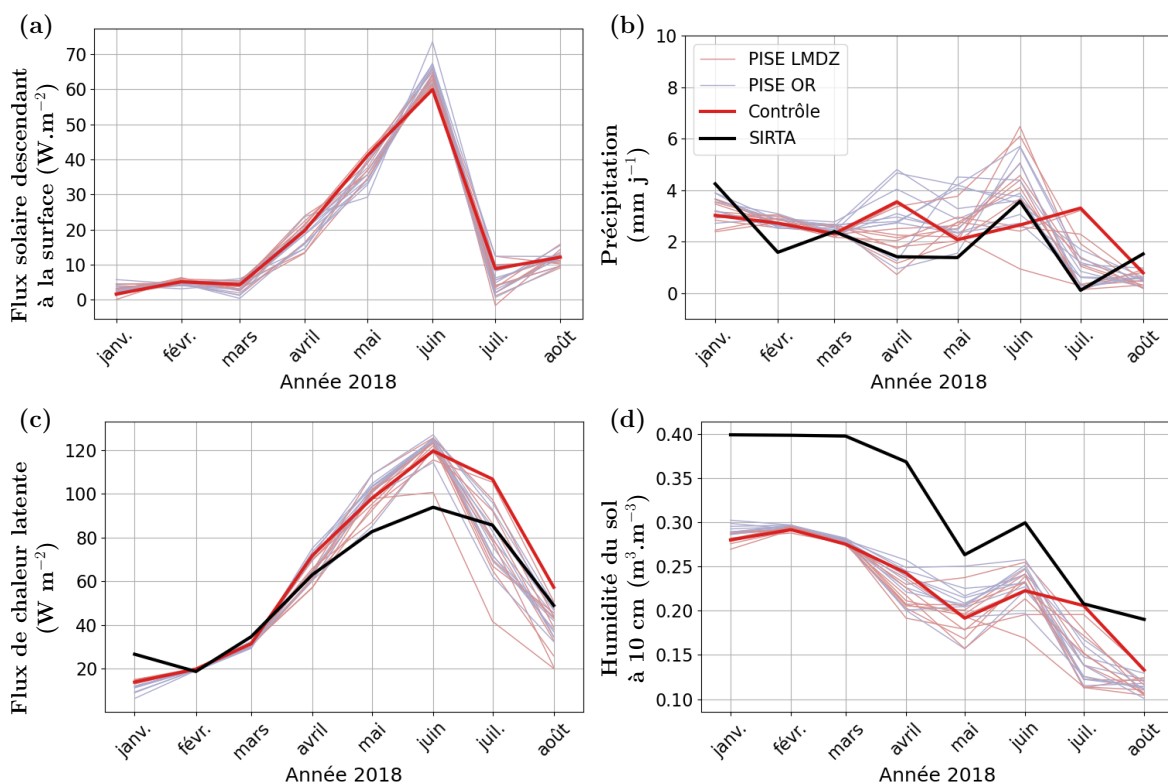


FIGURE IV.11 – Evolution en moyenne mensuelle (a) du flux solaire descendant à la surface (b) des précipitations (c) du flux de chaleur latente (d) de l’humidité du sol à 10 cm de profondeur, sur les mois de janvier à août 2018. La figure (a) montre une différence par rapport aux observations du SIRTa. En noir les observations du SIRTa, en rouge la simulation de contrôle avec physique LUDO-L95, guidage en vent (Table IV.2) et sans albédo imposé. En rouge claire les simulations de l’ensemble où les états initiaux de LMDZ ont été perturbés et en bleu celles où les états initiaux de ORCHIDEE ont été perturbés. Les simulations sont post-traitées avec le masque-par-variable, sauf pour l’humidité du sol qui est post-traitée avec le masque-global.

en janvier, et de ne pas rééquilibrer le modèle à chaque changement de vecteurs de paramètre libre, tant que les variables sensibles à l’état initial du sol ne sont pas directement visées dans le tuning.

En revanche, à partir d’avril, la dispersion des simulations des deux ensembles devient significative si on la compare typiquement aux différences entre simulations et observations, qui va être visée par les exercices de tuning. C’est le cas particulier pour les précipitations et les humidités du sol (Figure IV.11 (b) et (d)). Les variables météorologiques à 2 mètres (température, humidité, humidité relative) ainsi que les flux turbulents (par exemple le flux de chaleur latente (Figure IV.11 (c)) et dans une moindre mesure les flux radiatifs (par exemple le flux radiatif solaire descendant à la surface (Figure IV.11 (a))) ont également une variabilité interne qui augmente à partir d’avril 2018.

Cette augmentation de sensibilité à l’état initial après plusieurs mois de simulation, dans une configuration guidée, est inattendue. Elle est presque toujours indifférente au type d’état initial perturbé (atmosphérique ou continental). La sensibilité des précipitations à l’état initial de la simulation, plusieurs mois après le début de la simulation est particulièrement impressionnante. Cette sensibilité est telle que les précipitations simulées ne peuvent être comparées aux observations du SIRTa sur la période avril-juillet 2018, c’est-à-dire qu’on ne peut pas savoir si les précipitations sont bien simulées par le modèle ou pas. En l’état, elles ne peuvent donc pas

être utilisées comme métrique pour le tuning.

IV.4.2 Sensibilité au guidage

Guider les vents du modèle vers les réanalyses est un choix classique et nécessaire pour que le modèle suive la séquence météorologique qui s'est effectivement déroulée sur la période d'étude. Cela permet de comparer les simulations à des observations sur des échelles de temps inférieures à l'échelle climatique, sans perturber la physique du modèle. Dans certaines conditions, il peut également être intéressant de guider d'autres variables d'état du modèle atmosphérique vers des réanalyses ; tel que la température et/ou l'humidité. C'est le cas des situations où la grille du modèle est zoomée sur une région d'intérêt. Les températures et humidités atmosphériques peuvent être guidées à l'extérieur de la zone zoomée, afin de ne pas cumuler les biais des autres régions, dont la résolution est dégradée, à ceux de la région d'intérêt. La configuration ressemble alors à un modèle à air limité (LAM - *Limited Area Model*) sur la zone zoomée.

On peut espérer également que le fait de guider en humidité et/ou température puisse permettre de diminuer la variabilité interne de la configuration, et/ou de comprendre son origine.

Deux ensembles supplémentaires de simulations aux états initiaux perturbés sont ainsi réalisés dans cette partie, en utilisant les mêmes états initiaux que le PISE de la section précédente. Dans le premier ensemble, les vents et les températures sont guidés vers les réanalyses ERA5. Dans le deuxième, l'humidité relative est en plus guidée, également vers les réanalyses ERA5. Les constantes de temps de guidage sont résumées Table IV.2. Elles sont choisies de manière à avoir un rappel fort en température et humidité à l'extérieur du zoom et très faible dans la zone zoomée (constante de temps de 10 jours). Les vents zonaux sont néanmoins guidés à l'intérieur de la zone zoomée, avec une constante de temps de 12h, suffisamment longue pour ne pas affecter directement les processus de couche limite.

Nom du paramètre	Variables guidée	Constante de temps en jour
<code>tau_min_u</code>	Vent zonal à l'extérieur du zoom	0.125
<code>tau_max_u</code>	Vent zonal à l'intérieur du zoom	0.5
<code>tau_min_v</code>	Vent méridien à l'extérieur du zoom	0.125
<code>tau_max_v</code>	Vent méridien à l'intérieur du zoom	0.5
<code>tau_min_T</code>	Température à l'extérieur du zoom	0.125
<code>tau_max_T</code>	Température à l'intérieur du zoom	10
<code>tau_min_Q</code>	Humidité relative à l'extérieur du zoom	0.125
<code>tau_max_Q</code>	Humidité relative à l'intérieur du zoom	10

TABLE IV.2 – Constantes de temps de guidage utilisées pour guider le modèle en vent (zonaux et méridiens) et/ou température et/ou humidité dans les PISE de cette partie.

Un certain nombre de variables atmosphériques simulées sur la maille du SIRTa sont significativement sensibles au type de guidage. Cela signifie que les variables simulées sur la maille du SIRTa sont sensibles à l'état atmosphérique en dehors de la zone zoomée. La différence est considérée comme significative lorsqu'elle est supérieure à la variabilité interne, c'est-à-dire supérieure à l'étendue des PISE réalisés avec chaque type de guidage. Le guidage des températures entraîne une augmentation systématique de la température atmosphérique à deux mètres et de l'humidité spécifique à deux mètres (Figure IV.12 (e) et (f)), avec une diminution quasi systématique du flux solaire descendant à la surface liée à une augmentation de l'effet radiatif des nuages dans le solaire (Figure IV.12 (a)). L'ajout d'un guidage en humidité ramène à des humidités spécifiques comparables à la simulation de contrôle, guidée uniquement en vent (Figure IV.12 (e)). L'atmosphère est globalement moins humide dans ces simulations, avec moins

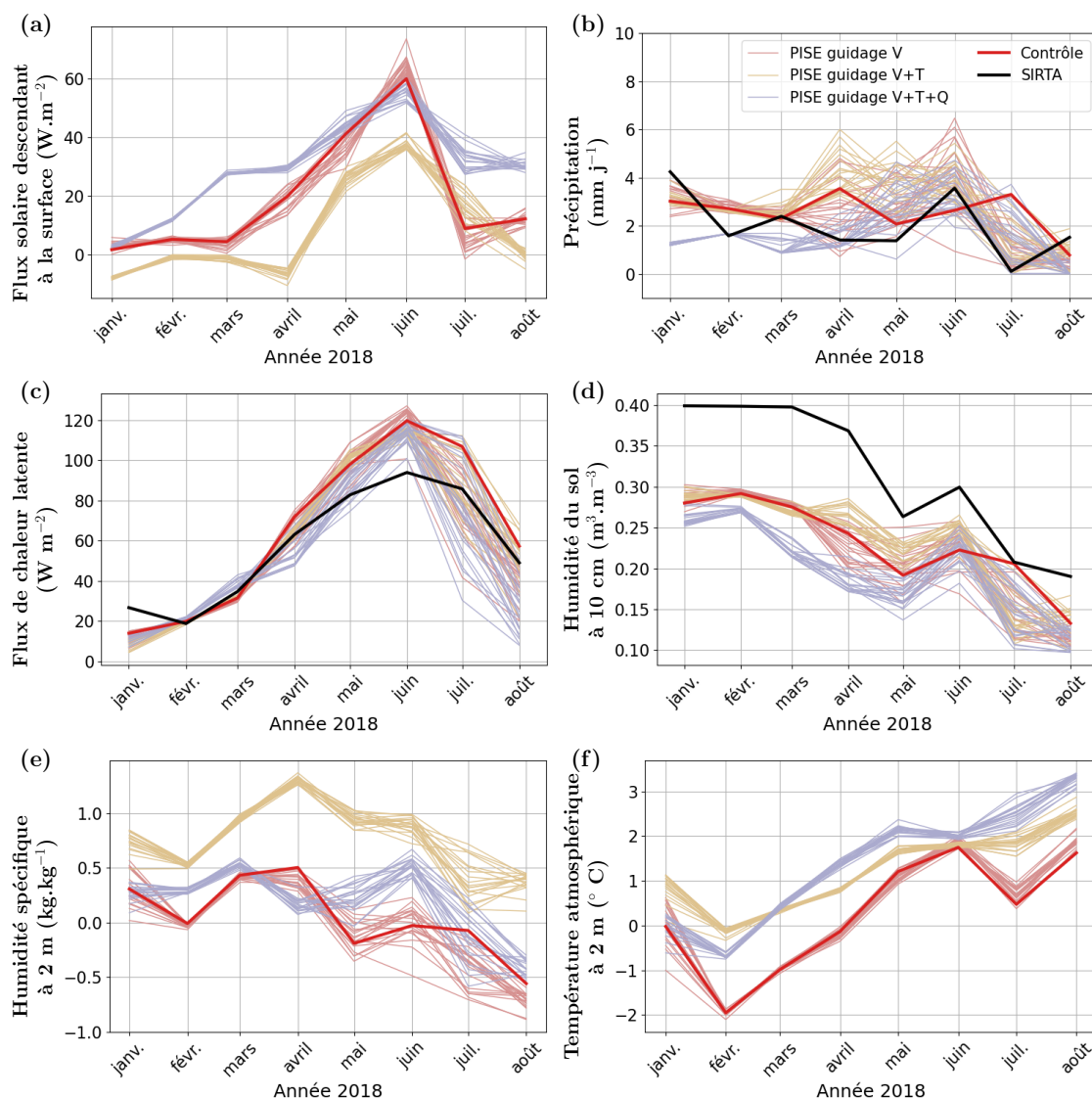


FIGURE IV.12 – Evolution en moyenne mensuelle (a) du flux solaire descendant à la surface (b) des précipitations (c) du flux de chaleur latente (d) de l’humidité du sol à 10 cm de profondeur (e) de l’humidité spécifique à 2 mètres et (f) de la température atmosphérique à 2 mètres, de janvier à août 2018. Les figures (a), (e) et (f) montrent les différences par rapport aux observations du SIRTa. Sur les autres figures, les observations du SIRTa tracées en noir. La configuration de contrôle (physique LUDO-L95, guidage en vent et albédo imposé sur la maille du SIRTa) est tracée en rouge, et le PISE associé en rouge clair. Le PISE avec guidage en vent et température et tracé en jaune clair, et celui avec guidage en vent, température et humidité relative est tracé en bleu. Les simulations sont post-traitées avec le masque-par-variable excepté pour l’humidité du sol qui est post-traitée avec le masque-global.

d’effet radiatif des nuages dans le solaire et l’infra-rouge, et ainsi moins de flux radiatif descendant à la surface (Figure IV.12 (a)). Les température atmosphériques de cet ensemble sont ainsi plus élevées que celles des deux autres PISE (Figure IV.12 (f)).

L’étendue du PISE guidé en température et du PISE guidé en température et humidité est plus faible que celui guidé uniquement en vent sur les mois de février et mars sur un certain nombre de variable (Figure IV.12). La variabilité interne des précipitations est en particulier significativement réduite sur ces premiers mois de l’année, mais continue d’être étonnamment importante à partir du mois d’avril 2018 (Figure IV.12 (b)). Sur ces premiers mois de l’année, la

valeur moyenne des précipitations est significativement affectée par le fait de guider en humidité à l'extérieur de la zone zoomée.

Les précipitations des mois d'hiver en île-de-france sont en effet principalement issues de l'advection d'air humide depuis l'Atlantique. Ces mois correspondent à une période où les précipitations sont simulées par la paramétrisation de condensation de grande échelle ; contrairement aux mois suivants (avril-août) où les précipitations sont dominées par la convection profonde.

La variabilité interne des précipitations des mois d'avril-mai-juin-juillet diminue un peu lorsque à la fois les vent, la température et l'humidité sont guidés, par rapports aux deux autres types de guidage testés. Cette légère diminution de la variabilité interne pourrait venir de la diminution de la variabilité interne des précipitations grande échelle. La variabilité interne des précipitations sur cette période est néanmoins encore trop importante pour pouvoir évaluer les précipitations simulées par LMDZ avec les observations du SIRTa.

La variabilité interne de l'humidité du sol (Figure IV.12 (d)) suit celle des précipitations, avec en plus un effet mémoire qui provoque une dispersion des simulations également sur les mois d'automne et d'hiver (d'autant plus longue que le sol est profond, voire par exemple le PISE Figure IV.19 avec la physique CMIP6-L79). Cet effet mémoire de l'humidité augmente avec la profondeur.

La variabilité interne des flux turbulent (Figure IV.12 (c)) augmente également fortement à partir d'avril ; ce qui peut également être la conséquence de la variabilité interne des précipitations via l'humidité du sol. En effet, dans ces mois de printemps et d'été, l'évaporation est liée à l'eau disponible et à l'énergie disponible (régime d'évaporation transitoire) ou uniquement à l'eau disponible (régime sec, éventuellement présent pendant la sécheresse à partir de juillet). La variabilité interne de l'humidité du sol peut ainsi se repercuter sur le flux de chaleur latente, et sur toutes les variables sensibles au couplage entre atmosphère et surface continentale. L'hypothèse que la variabilité interne des variables sensibles aux couplage est provoquée, au moins en partie, par la variabilité interne des précipitations via l'humidité du sol est testée Section IV.6.2.2.

Au vu de la sensibilité des variables d'intérêt au guidage en température et en humidité à l'extérieur de la zone zoomée, on choisit de guider ces variables d'état dans la suite de l'étude. De plus, on montre Section IV.5.2 que la variabilité interne des précipitations convectives est largement surestimée. Il est très probable que les précipitations convectives simulées à l'extérieur de la zone zoomée ait également une variabilité interne surestimée, et que cette dernière se repercute sur les humidité advectées dans la zone zoomée ; ce que l'on préfère éviter.

IV.5 Variabilité interne des précipitations

Dans la section précédente, nous avons montré via des ensembles perturbés aux conditions initiales (PISE) que la variabilité interne des précipitations était importante à partir du mois d'avril. Cela signifie que, plusieurs mois après le début de la simulation, et malgré un guidage en vent, température et humidité à l'extérieur de la zone zoomée, les précipitations sont très sensibles à l'état initial du modèle.

Pour étudier cette variabilité, deux PISE de 20 membres sont réalisés sur l'année 2018, en suivant la méthodologie présentée Section IV.4.1. Les états initiaux atmosphériques sont perturbés dans le premier ensemble, et ceux du sol sont perturbés dans le deuxième. Ces ensembles sont réalisés avec la physique CMIP6-L79, le guidage en vent, température et humidité de la Section IV.4.2, et l'albédo imposé sur la maille du sirta (Section IV.3.3.5)

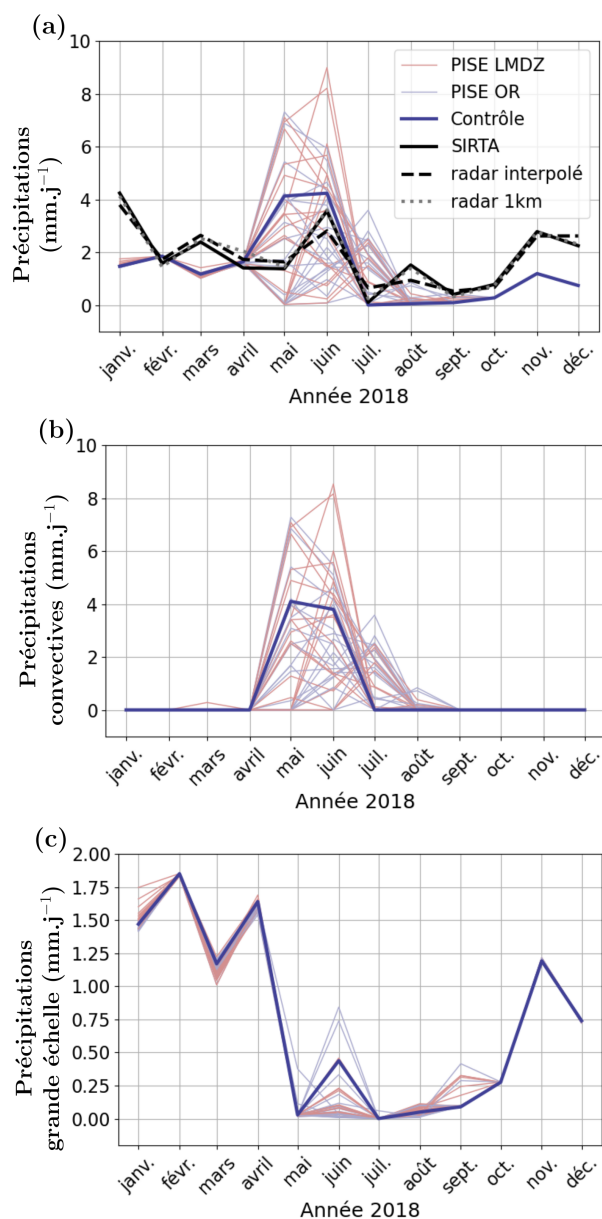


FIGURE IV.13 – Evolution en moyenne mensuelle sur l'année 2018 (a) des précipitations (b) des précipitations issues de la paramétrisation de la convection profonde (c) des précipitations issues de la paramétrisation grande échelle, sur la maille du SIRTa. En noir les observations du pluviomètre du SIRTa ; en noir pointillé les réanalyses Coméphore regrillées sur la grille du GCM, en gris pointillé les réanalyses Coméphore non regrillées (résolution de 1 km²), en bleu foncé la simulation de contrôle, en bleu claire le PISE où les états initiaux atmosphériques sont perturbés, en rouge claire le PISE associé où les états initiaux du sol sont perturbés. Les simulations utilisent la physique CMIP6-L79, l'albédo imposé sur la maille du SIRTa et un guidage en vent, humidité et température (Table IV.2). Données non masquées. [Tentative de figure en colonne. Hackée en copie d'écran]

IV.5.1 Variabilité interne des précipitations convectives

Contrairement aux simulations de la section précédente effectuées avec la physique LUDO-L95, la variabilité interne des précipitations augmentent avec la physique CMIP6-L79 de mai à août 2018 (Figure IV.13 (a)). Cette période correspond aux mois de l'année 2018 où les précipitations viennent principalement de la paramétrisation de la convection profonde (Figure IV.13 (b)), contrairement aux autres mois de l'année où les précipitations viennent de la paramétrisation de condensation grande échelle (Figure IV.13 (c)). On retrouve le fait que la sensibilité des précipitations est la même que l'on change l'état initial du modèle d'atmosphère ou du modèle de sol, excepté sur les précipitations grandes échelles.

Cet effet sur les pluies grande échelles peut se comprendre car un changement d'état initial du sol sur tout le globe implique un changement d'humidité des sols sur toute la zone zoomée (l'humidité atmosphérique étant guidée à l'extérieur du zoom, on peut raisonnablement supposer qu'un changement d'humidité du sol à l'extérieur de la zone zoomée n'impact pas significativement l'humidité de l'air advecté sur la maille du SIRTa). Il a été montré qu'un changement

local d'humidité au SIRTa n'affectait pas les précipitations du SIRTa, mais qu'un changement d'humidité régional pouvait effectivement rétroagir sur les précipitations (Campoy et al., 2013). Il est possible que ce changement ne soit pas visible sur les précipitations convectives du fait de leur trop forte variabilité interne.

IV.5.2 Étude régionale des précipitations autour du SIRTa

Pour mieux comprendre cette variabilité interne des précipitations, essayer d'évaluer si elle est réaliste ou non, et éventuellement trouver un moyen de la diminuer, nous allons étudier le comportement des deux PISE précédents sur une région de quelques mailles autour du SIRTa. Nous utilisons ici comme référence les réanalyses Coméphore (COMbinaison en vue de la Meilleure Estimation de la Précipitation HORaiRE) de Météo-France (Tabary et al., 2012). Cette réanalyse est disponible toutes les heures à 1km de résolution sur toute la France métropolitaine depuis 1997. Elle combine les observations de précipitations faites avec des radars (observations spatialisées) avec des observations faites avec des pluviomètres (observations locales). Le produit est interpolé sur la grille du GCM (on dira regrillée par commodité) utilisée dans cette étude (Figure IV.4) pour permettre sa comparaison avec les précipitations simulées par le modèle.

On vérifie tout d'abord sur la Figure IV.13 (a) que le produit Coméphore sur la maille du SIRTa est en très bon accord avec les observations sur site, signe d'une faible incertitude liée à la localité du site.

On montre sur la Figure IV.14 (a) la distribution géographique des précipitations moyennes du mois de juin 2018 de ces réanalyses. On définit trois régions centrées sur la maille du SIRTa : une région de 3×3 mailles ($90 \times 90 \text{ km}^2$), une de 5×5 mailles ($150 \times 150 \text{ km}^2$) et une de 7×7 mailles ($230 \times 230 \text{ km}^2$), représentées Figure IV.14 (b) par les rectangles bleu, vert et rouge. L'évolution sur l'année 2018 des précipitations moyennes sur ces trois régions est tracées Figure IV.15.

La variabilité interne des précipitations reste faible sur les mois d'automne et d'hiver (de janvier à avril puis de septembre à décembre 2018), avec des précipitations systématiquement sous-estimées sur ces périodes sur les trois régions étudiées (Figure IV.15 (a), (b) et (c)), comme sur la maille du SIRTa (Figure IV.13 (a)). Tout comme sur la maille du SIRTa, la variabilité interne augmente pendant les mois d'été où les précipitations sont dominées par les précipitations issues de la paramétrisation de la convection profonde. Cependant, plus les précipitations sont moyennées sur une grande région, plus leur variabilité interne diminue.

La variabilité interne des précipitations moyennées régionalement à suffisamment diminué les mois d'été pour pouvoir être comparée aux observations, sur les trois régions. On peut ainsi affirmer que les précipitations des mois de juin et juillet 2018 sur les régions de 9, 25 et 49 mailles autour de la maille du SIRTa. Ces précipitations régionales sont également dominées par des précipitations issues de la paramétrisation de la convection profonde.

On calcule également l'écart type spatial des moyennes mensuelles des précipitations sur les trois régions. Ces écart types mesurent la variabilité spatiale des précipitations en moyenne mensuelle sur les trois régions étudiées (Figure IV.16 (a)). Dans le produit radar, cet écart type est un peu plus fort en juin que sur les autres mois de l'année 2018, ce qui peut s'expliquer par la présence d'orage sur la partie nord de la France de fin mai à mi-juin 2018.

On retrouve une augmentation de la variabilité spatiale des précipitations mensuelles sur les mois d'été dans les simulations du PISE (mai, juin et juillet 2018), mais cette dernière est fortement surestimée par rapport à la variabilité spatiale observée. Cette surestimation se retrouve dans la variabilité spatiale des précipitations issu de la paramétrisation de la convection profonde (Figure IV.16 (b)) et est illustrée sur des cartes pour le mois de juin 2018 (Figure IV.14). Les précipitations issues des réanalyses Figure IV.14 (a) sont beaucoup plus lisses spatialement que les précipitations de la simulation de contrôle (b) ou que deux autre membre du PISE (c)

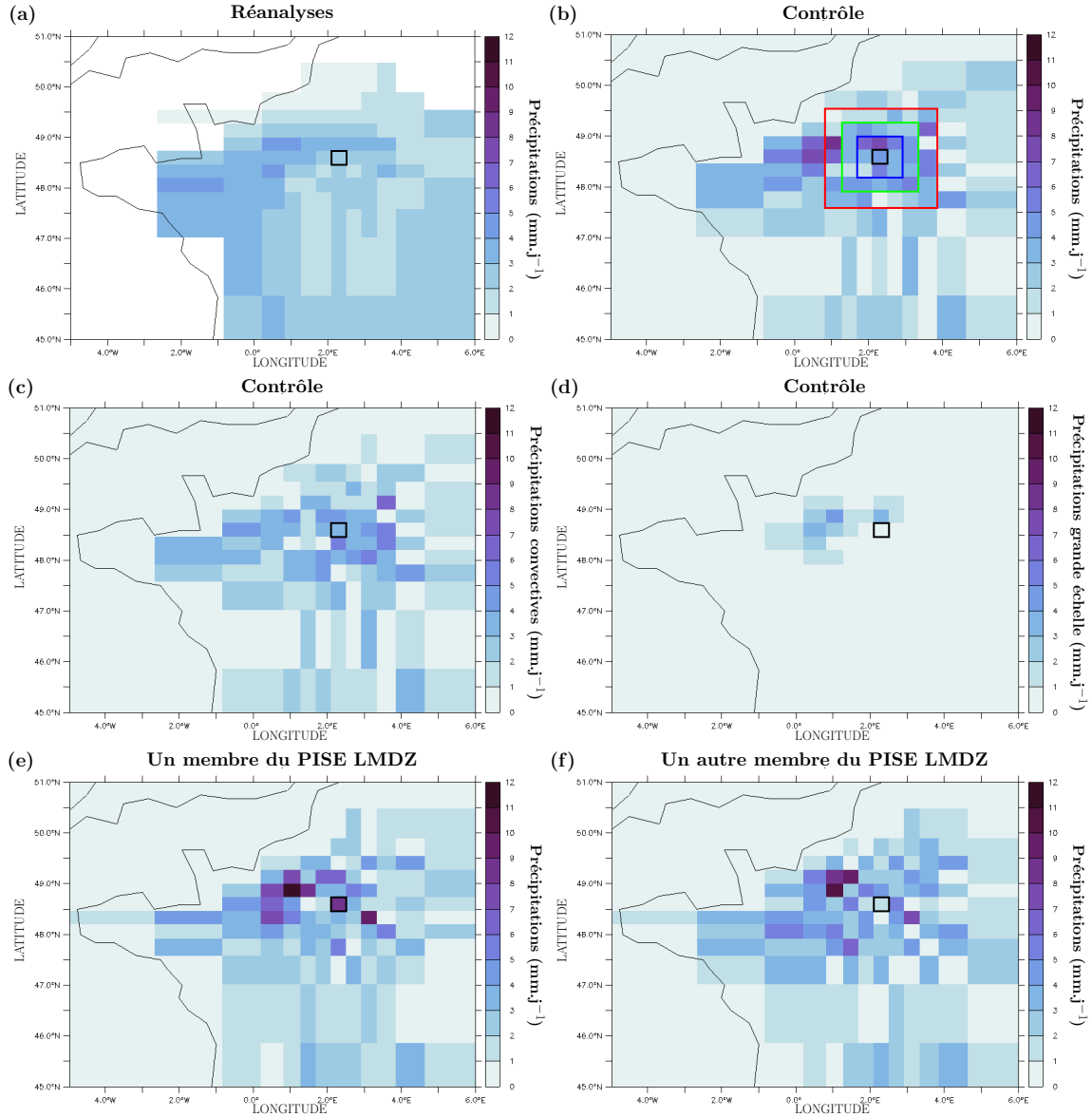


FIGURE IV.14 – Cartes des précipitations moyennes du mois de juin 2018 (a) des réanalyses Coméphore regrillées sur la grille du modèle (b) de la simulation de contrôle (c) des précipitations issues de la paramétrisation de la convection profonde dans la simulation de contrôle (d) des précipitations issues de la paramétrisation de condensation grande échelle de la simulation de contrôle (e) de la simulation du PISE ayant comme état initial atmosphérique celui du 1er janvier 2002 (f) de la simulation du PISE ayant comme état initial atmosphérique celui du 1er janvier 2011. Le rectangle noir indique la maille où se site les observations du SIRTa, les rectangles bleu, vert et rouge indique les régions définies autour de la maille du SIRTa de respectivement 9, 25 et 49 mailles. Les simulations utilisent la physique CMIP6-L79, l'albédo imposé sur la maille su SIRTa et un guidage en vent, température et humidité. Données non masquées.

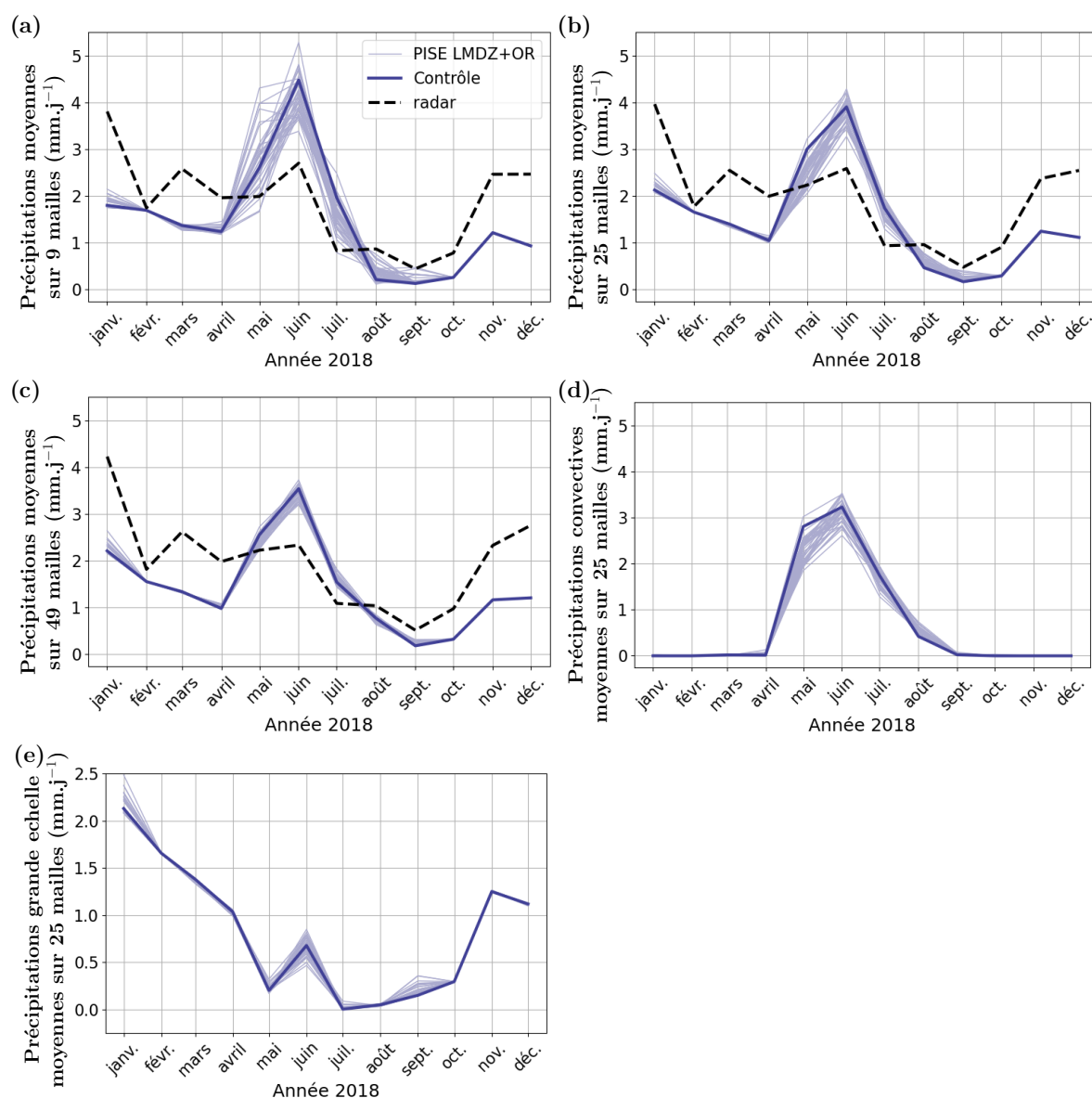


FIGURE IV.15 – Evolution en moyenne mensuelle sur l'année 2018 des précipitations moyennes (a) sur 9 mailles (b) sur 25 mailles (c) sur 49 mailles autour de la maille du SIRTa ; (d) des précipitations issues de la paramétrisation de la convection profonde en moyenne sur 25 mailles (e) des précipitations issue de la paramétrisation des précipitations grandes échelles en moyenne sur 25 mailles. En noir pointillé les réanalyses Coméphore, en bleu foncé la simulation de contrôle, en bleu claire les deux PISE associés à la simulation de contrôle (perturbation des états initiaux du modèle d'atmosphère et du sol sans distinction de couleur). Les simulations utilisent la physique CMIP6-L79, l'albédo imposé sur la maille su SIRTa et un guidage en vent, température et humidité. Données non masquées. [Je pense que ça pourrait être séparé en une colonne de 3 figures avec les différentes zones de lissage et une figure avec juste les deux panneaux de la séparation grande échelle convectif.]

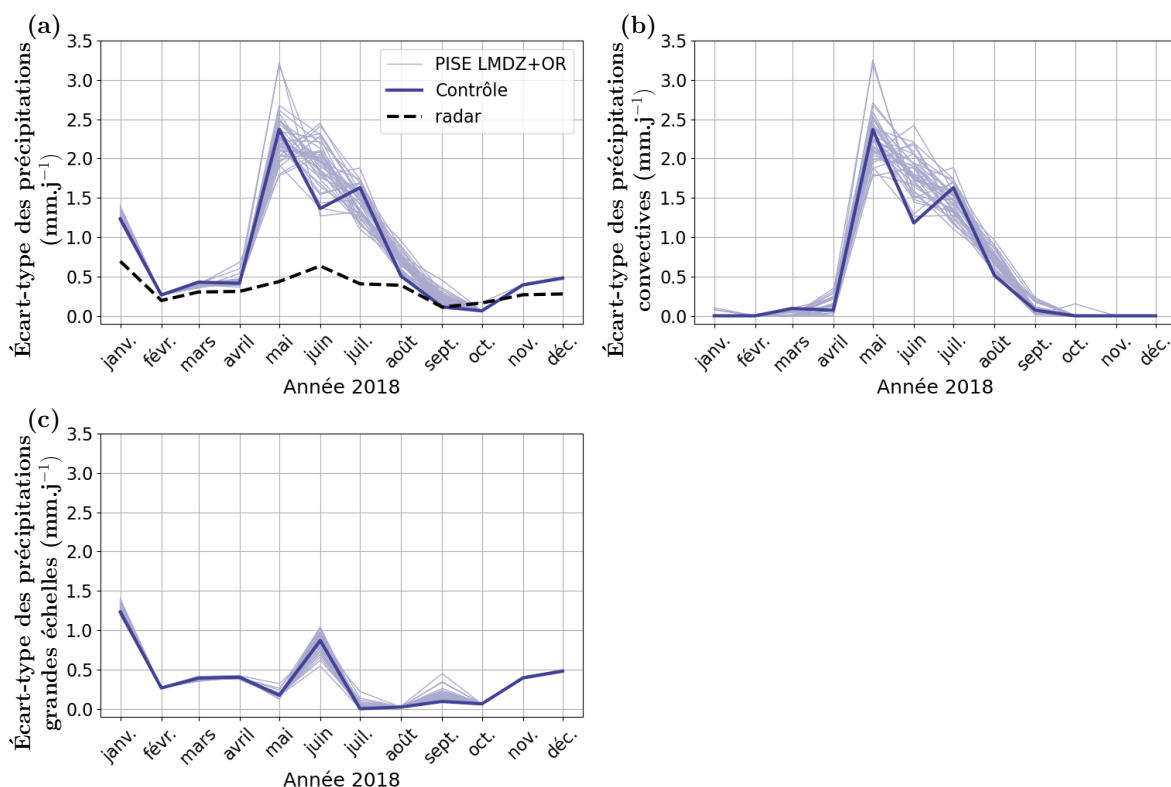


FIGURE IV.16 – Evolution en moyenne mensuelle sur l’année 2018 des écarts-type spatiaux des précipitations sur 25 mailles (a) précipitations totales (b) précipitations issues de la paramétrisation de la convection profonde (c) précipitations issues de la paramétrisation de condensation grande échelle. En noir pointillé les réanalyses Coméphore, en bleu foncé la simulation de contrôle, en bleu claire les deux PISE associés à la simulation de contrôle (perturbation des états initiaux du modèle d’atmosphère et du sol). Les simulations utilisent la physique CMIP6-L79, l’albédo imposé sur la maille du SIRTa et un guidage en vent, température et humidité. Données non masquées.

et (f). En comparant la carte des précipitations de la simulation de contrôle Figure IV.14 (b) avec celles des précipitations issues de la paramétrisation de convection profonde Figure IV.14 (c) et de condensation grande échelle Figure IV.14 (d), on voit que les précipitations et leur variabilité spatiale sont majoritairement issues de la paramétrisation de convection profonde. Les précipitations grande échelle sont étonnamment non négligeable dans cette période de convection profonde. Elles se produisent de manière massives sur peu d’événements courts.

Cela suggère que ces précipitations sont le résultat d’orages point de grille, qui se produisent lorsque la convection profonde est résolue explicitement par le modèle (et non plus simulée par la paramétrisation dédiée). Les orages points de grilles ne sont pas désirables car ils produisent des précipitations irréaliment intense sur des courtes périodes de temps.

Pendant la période de convection profonde, l’emplacement des mailles où il va y avoir plus de précipitations que les autres, dans les régions étudiées, change avec l’état initial de la simulation. Autrement dit, dans le PISE, ce ne sont jamais les mêmes mailles qui ont le plus ou le moins de précipitations. Ainsi, pour une moyenne régionale similaire, les précipitations vont être aléatoirement plus forte, ou moins forte, sur la maille du SIRTa par rapport aux autres mailles de la région. C’est pourquoi on peut penser que si la variabilité spatiale des précipitations était correctement simulée, alors la variabilité interne des précipitations sur la maille du SIRTa serait réduite. Autrement dit, la forte variabilité interne des précipitations sur la maille du SIRTa pendant les périodes orageuses est due à la simulation d’une variabilité spatiale des

précipitations convective surestimées. Ainsi, la variabilité interne des précipitations sur la maille du SIRTa est surestimée pendant les périodes de convection profonde.

Ces résultats soulèvent beaucoup de questions.

Tout d'abord, on peut se demander si la variabilité spatiale des précipitations ne vient pas d'une sensibilité de la paramétrisation de la convection à la résolution du modèle. Les paramétrisations sont pensées sous une hypothèse d'homogénéité statistique. Cette hypothèse est facilement vérifiée pour des phénomènes dont la taille caractéristique est inférieure d'au moins un ordre de grandeur à la taille de la maille, et qui ne sont pas organisés à plus grande échelle, comme pour la turbulence de petite échelle ou la convection peu profonde. La taille des systèmes de convection profonde varie quant à elle de quelques kilomètres pour les orages locaux, à plusieurs dizaines voire centaines de kilomètres pour les cellules et les fronts orageux. Même lorsque les orages sont locaux et non organisés, ils ne sont que quelques uns dans une maille de modèle de $30 \times 30 \text{ km}^2$. Il est donc possible que la variabilité spatiale des précipitations issues de la convection profonde soit réaliste sur la grille native du modèle mais devienne irréaliste pour des mailles de $30 \times 30 \text{ km}^2$. Cette hypothèse mériterait d'être testée dans des PISE de simulations guidées en vent avec la grille standard du modèle. Il est également possible que la diminution de la taille des mailles augmente la présence d'orages point de grille.

L'investigation de l'origine de la surestimation de la variabilité spatiale est une question scientifique à part entière, qui n'a pu être profondément instruite pendant cette thèse. Son lien avec la rétroaction de l'hydrologie du sol est cependant étudiée dans la section suivante, et sa dépendance aux valeurs des paramètres libres est analysée dans le Chapitre V.

IV.6 Couper la rétro-action de l'humidité du sol sur l'atmosphère en imposant le facteur d'aridité

La variabilité interne des précipitations, surestimée en période convection profonde (en mai-juin-juillet 2018), peut se répercuter sur les variables atmosphériques sensibles au couplage à la surface, via son impact sur l'humidité du sol. La variabilité interne des précipitations pourrait ainsi expliquer l'augmentation de variabilité interne d'un certain nombre de variable atmosphériques à partir d'avril 2018 (tel que la température et l'humidité à deux mètres Figure IV.12 (e) et (f), le flux de chaleur latente Figure IV.12 (c)).

Pour tester cette hypothèse, nous coupons la rétroaction de l'humidité du sol sur l'atmosphère, en imposant le facteur d'aridité diagnostiqué à partir des observations du SIRTa, sur la maille du SIRTa du modèle.

IV.6.1 Diagnostic du facteur d'aridité dans les observations du SIRTa

La rétro-action de l'humidité du sol sur l'atmosphère se fait via l'évaporation E , et plus exactement via le facteur d'aridité noté β . En effet, on peut voir l'évaporation comme le produit de deux termes : l'évaporation potentielle, notée E_{pot} , et le facteur d'aridité β tels que

$$E = \beta E_{pot} \quad (\text{IV.9})$$

L'évaporation potentielle correspond à la capacité de l'atmosphère à recevoir de la vapeur d'eau de la surface, vue aussi comme une demande d'évaporation de l'atmosphère. Elle correspond à l'évaporation qu'il y aurait au-dessus d'une surface d'eau libre. Le facteur d'aridité correspond à la capacité du sol à fournir l'eau demandée par l'atmosphère. Il est compris entre 0 et 1, et vaut 1 si le sol peut entièrement répondre à la demande d'évaporation, ce qui est le cas d'une surface d'eau libre.

Dans le modèle, LMDZ calcule l'évaporation potentielle et ORCHIDEE calcule le facteur d'aridité puis renvoie le flux d'évaporation à LMDZ. Le facteur d'aridité dépend en premier lieu de l'humidité du sol, mais aussi du type de surface : un sol nu ne pourra évaporer que l'eau de ses premières couches alors qu'une surface courverte de végétation pourra puiser l'eau plus profondément dans le sol grâce à ses racines.

Le facteur d'aridité est diagnostiqué dans les observations du SIRTa en utilisant l'Equation IV.10, pour laquelle le flux de chaleur latente LAT est observé au SIRTa, et avec L_v la chaleur latente de vaporisation de l'eau qui vaut 2500 kJ.kg^{-1} .

$$\beta = \frac{E}{E_{pot}} = \frac{LAT}{L_v E_{pot}} \quad (\text{IV.10})$$

IV.6.1.1 Diagnostic de l'évaporation potentielle à partir des observations du SIRTa

L'évaporation potentielle n'est pas observée, et il s'agit de la calculer à partir des observations disponibles. L'évaporation potentielle s'écrit

$$E_{pot} = \rho C_h |\vec{V}| (q_{sat}(T_s) - q_{air}) = \frac{\kappa}{C_p} (q_{sat}(T_s) - q_{air}) \quad (\text{IV.11})$$

avec ρ la masse volumique de l'air en kg.m^{-3} , $|\vec{V}|$ le module du vent horizontal dans la couche limite de surface en m.s^{-1} , q_{air} l'humidité spécifique de l'air à un niveau donné au dessus de la surface, en $\text{kg d'eau par kg d'air}$, $q_{sat}(T_s)$ l'humidité spécifique à saturation à la température de la surface T_s et C_h est un coefficient d'échange, dépendant de la hauteur à laquelle sont estimées les variables atmosphériques. Ce coefficient dépend notamment de la stabilité de la couche limite.

Pour simplifier l'écriture dans la suite, on a introduit le coefficient $\kappa = \rho C_h |\vec{V}| C_p$, avec C_p la capacité calorifique de l'air sec, égale à $1004 \text{ J.kg}^{-1}.\text{K}^{-1}$. L'évaporation potentielle est proportionnelle à la différence de l'humidité spécifique de l'air dans la couche qui est en équilibre avec la surface d'eau libre, et l'humidité de l'air à un niveau situé dans la couche limite de surface, qui correspond à la couche atmosphérique où les flux turbulents restent quasiment égaux aux flux de surface (de l'ordre de 10-100 m de haut). L'air en équilibre avec la surface est une couche "imaginaire" de quelques millimètres au-dessus de la surface, qui a la température de la surface T_s et est saturé en humidité, donc a une humidité $q_{sat}(T_s)$. Dans cette formulation de l'évaporation potentielle (Equation IV.11), il est important que toutes les variables atmosphériques (ρ , C_h , $|\vec{V}|$ et q_{air}) soient évaluées à la même altitude, et que cette altitude soit dans la couche limite de surface. Dans LMDZ, ce sont les valeurs de la première couche du modèle qui sont utilisées. Cette couche est située à une altitude d'environ 10 m au dessus de la surface pour les discrétisations verticales standard à 79 et 95 niveaux Figure II.3.

L'Equation IV.11 ne peut pas être utilisé telle quelle pour diagnostiquer l'évaporation potentielle à partir des observations du SIRTa, car les variables C_h et $q_{sat}(T_s)$ ne sont pas observées. Pour dériver E_{pot} des observations, on utilise la formulation du flux de chaleur latente proposé par Hourdin et al. (2015), redémontrée ci dessous.

$$E_{pot} = \frac{\kappa}{C_p} [q_{sat}(T_s) - q_{sat}(T_{atm}) + q_{sat}(T_{atm}) - q_{air}] \quad (IV.12)$$

$$= \frac{\kappa}{C_p} q_{sat}(T_{atm}) \left[\frac{q_{sat}(T_s) - q_{sat}(T_{atm})}{q_{sat}(T_{atm})} + 1 - \frac{q_{air}}{q_{sat}(T_{atm})} \right] \quad (IV.13)$$

$$\approx \frac{\kappa}{C_p} q_{sat}(T_{atm}) \left[\left. \frac{\partial q_{sat}(T)}{\partial T} \right|_{T=T_{atm}} \frac{(T_s - T_{atm})}{q_{sat}(T_{atm})} + 1 - RH \right] \quad (IV.14)$$

$$= \frac{\kappa}{C_p} q_{sat}(T_{atm}) \left[\frac{L_v(T_s - T_{atm})}{R_v T_{atm}^2} + 1 - RH \right] \quad (IV.15)$$

On utilise la loi de Clausius-Clapeyron dans le passage entre IV.14 et IV.15 pour approximer la dérivée de l'humidité à saturation par rapport à la température. On note R_v la constante des gaz parfait pour l'air, qui vaut $287 \text{ J.kg}^{-1}.\text{K}^{-1}$.

La température de surface T_s est calculée à partir des observations de rayonnement infrarouge montant au SIRTa via la loi de Stefan : $T_s = \left(\frac{LW_{up}}{\sigma} \right)^{\frac{1}{4}}$, avec σ la constante de Stefan-Boltzmann.

Pour être cohérent avec la formulation de LMDZ, on souhaite que T_{atm} soit la température de l'atmosphère à 10 m, RH l'humidité relative à 10 m et $q_{sat}(T_{atm})$ l'humidité spécifique à saturation à 10 m. La température et l'humidité relative sont observées sur la mâât météorologique du SIRTa à 10 m. L'humidité spécifique à saturation est le rapport de l'humidité spécifique par l'humidité relative. L'humidité spécifique n'étant pas observée à 10 m mais à 2 m (ainsi que la température et l'humidité relative), nous utilisons un développement limité à l'ordre 1 pour approximer l'humidité à saturation à 10 m. L'Equation IV.18 est ainsi utilisée pour diagnostiquer l'humidité à saturation à 10 m dans les observations du SIRTa.

$$q_{sat}(T_{atm}) \approx q_{sat}(T_{2m}) + \left. \frac{\partial q_{sat}}{\partial T} \right|_{T=T_{2m}} (T_{atm} - T_{2m}) \quad (IV.16)$$

$$\approx q_{sat}(T_{2m}) + \frac{L_v q_{sat}(T_{2m})(T_{atm} - T_{2m})}{R_v T_{2m}^2} \quad (IV.17)$$

$$\approx \frac{q_{2m}}{RH_{2m}} \left[1 + \frac{L_v(T_{atm} - T_{2m})}{R_v T_{2m}^2} \right] \quad (IV.18)$$

Le coefficient $\kappa = \rho C_h |\vec{V}| C_p$, n'est pas directement observée, mais il est apparait dans le calcul du flux sensible :

$$SENS = \rho C_h |\vec{V}| C_p (T_s - \theta_{atm}) \quad (IV.19)$$

en faisant l'hypothèse que le coefficient d'échange, C_h , est le même pour le flux de chaleur sensible que pour l'évaporation potentielle.³

De la même manière que pour l'évaporation potentielle, les variables atmosphériques de l'Equation IV.19 doivent toutes être prises au même niveau. Elles sont prises dans la première couche de l'atmosphère par LMDZ et à 10 m pour le diagnostique dans les observations du SIRTa. On fait de plus l'hypothèse que la température potentielle virtuelle à 10 m, θ_{atm} , est égale à la température de l'air à 10 m, ce qui est justifié par la très faible variation de pression sur 10 m.

On utilise l'équation

$$\kappa = \frac{SENS}{T_s - T_{atm}} \quad (IV.20)$$

3. Bien qu'ils ne soient pas théoriquement égaux, cette hypothèse est assez classique et est faite dans LMDZ

pour diagnostiquer le coefficient κ à la fréquence horaire, qui est la fréquence de disponibilité des observations. Les figures IV.17 (a) et IV.18 (a) montrent l'évolution de κ ainsi diagnostiqué pour les mois de janvier et juin 2018.

Ponctuellement, des valeurs très grandes de $|\kappa|$ apparaissent, allant jusqu'à $\pm 40\,000$. Cela peut se produire lorsque $T_{atm} \approx T_s$ ou lorsque le flux de chaleur sensible est proche de zéro. Afin d'éviter ces situations, on masque toutes les heures où le calcul conduit à des valeurs de κ supérieures à $50\text{ W m}^{-2}\text{ K}^{-1}$ ou inférieures à $0,2$ ⁴

L'évaporation potentielle est calculée à partir des observations du SIRTa en utilisant les équations IV.15 et IV.18, ainsi que le κ diagnostiqué avec l'Equation IV.20. Afin d'évaluer les équations utilisées pour le diagnostic de l'évaporation potentielle, on compare l'évaporation potentielle simulée par le modèle avec l'évaporation potentielle diagnostiquée a posteriori sur la simulation. Ce diagnostic utilise l'Equation IV.20 pour diagnostiquer κ avec le même masque des valeurs inférieur à 0,2 et supérieur à 50, et l'Equation IV.15 pour diagnostiquer l'évaporation potentielle à partir des température, humidité spécifique et humidité relative de la première couche de l'atmosphère. Ces trois évaporation potentielles sont tracées Figure IV.17 (b) et Figure IV.18 (b) pour les mois de janvier et juin 2018.

L'évaporation potentielle diagnostiquée à partir de la simulation est presque systématiquement plus grande que celle directement simulée par le modèle sur les journées de fin février à début Novembre 2018 (Figure IV.18 (b) pour le mois de juin). L'évaporation potentielle simulée par le modèle est quant à elle souvent plus importante que celle dérivée des observations du SIRTa sur les journées d'avril à mi-octobre 2018. Cela est cohérent avec le fait que le modèle simule un climat trop chaud et trop sec [*A vérifier sur les simulations et mettre les figures après*]. De plus, la méthode utilisée pour le diagnostic de l'évaporation potentielle permet le diagnostic d'évaporations potentielles négatives, ce qui n'est pas le cas de l'évaporation potentielle simulée. Les évaporation potentielles négatives sont principalement diagnostiquées lors des nuit d'hiver (Figure IV.17 (b)), mais peuvent également avoir lieu en journée, en particulier lorsque le sol est gelé et/ou que le sol est recouvert de neige.

IV.6.1.2 Diagnostic du facteur d'aridité

Une fois l'évaporation potentielle calculée, on diagnostique le facteur d'aridité avec l'Equation IV.10, puis on applique plusieurs masques afin d'éviter les situations où le calcul du facteur d'aridité n'est pas considéré comme fiable. C'est le cas des situations où le flux sensible est négatif, qui sont donc masquées. Le flux sensible est négatif lorsque l'atmosphère est plus chaude que la surface, ce qui se produit fréquemment en hiver ou la nuit. Dans ces situations, il n'y a pas d'évaporation à la surface et elles sont donc retirées du calcul du facteur d'aridité. Ensuite, les valeurs du facteur d'aridité supérieures à 1 sont mises à 1, car le facteur d'aridité est en théorie borné par 1 (au moins pour un sol nu). Enfin, pour éviter la division par zéro dans l'Equation IV.10, le facteur d'aridité est masqué lorsque la valeur absolue de l'évaporation potentielle est inférieur à $5 \cdot 10^{-6}\text{ kg.s}^{-1}.\text{m}^{-2}$. Le facteur d'aridité ainsi diagnostiqué et masqué est tracé Figure IV.17 (c) et Figure IV.18 (c) pour janvier et juin 2018.

Afin de vérifier la robustesse de la chaîne de calcul menant au diagnostic du facteur d'aridité dans les observations, le flux latent observé au SIRTa est comparé au flux latent reconstruit avec la chaîne de calcul, c'est-à-dire au flux latent diagnostiqué comme : $LAT = L_v \beta E_{pot}$. Ces deux flux latent sont tracés Figure IV.17 (d) et Figure IV.18 (d) pour les mois de janvier et juin 2018. 96 points ont des différences sur ces deux flux latent supérieur à $\pm 10\text{ W.m}^{-2}$ sur l'année 2018, ce qui est considéré comme satisfaisant. Le nombre de valeurs manquantes sur

4. Ces bornes ont été établies comme suffisantes pour reproduire les flux turbulents en surface en mode "débranché" dans une études sur l'impact du climat sur la thermique du bâtiment (Villefranche et al., 2022).

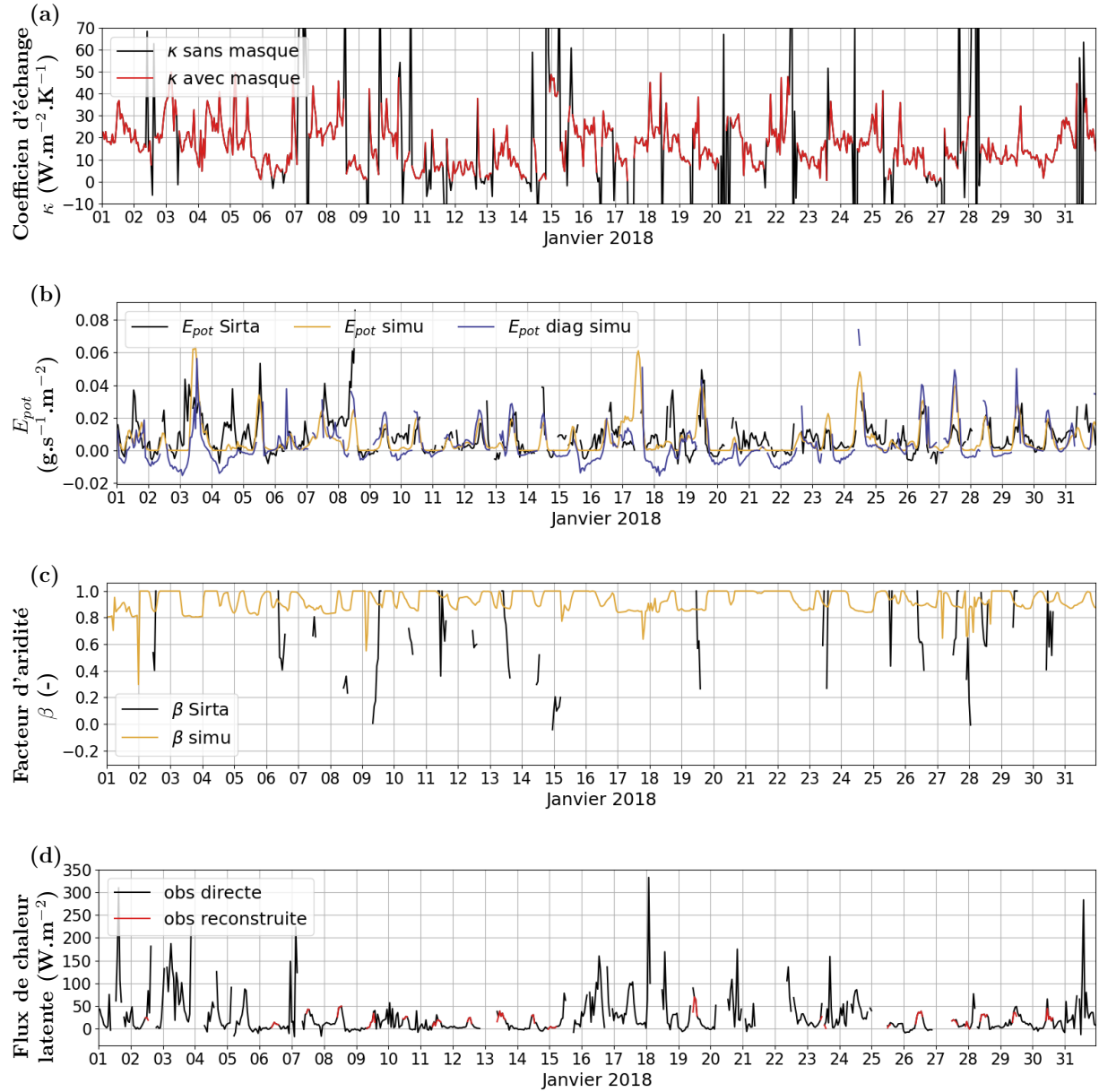


FIGURE IV.17 – Evolution à fréquence horaire sur le mois de janvier 2018 (a) du coefficient d'échange κ diagnostiqué avec l'Equation IV.20 avec masque (rouge) et sans masque (noir) (b) de l'évaporation potentielle diagnostiquée avec l'Equation IV.15 dans les observations du sirta (noir), simulée par le modèle (jaune) et diagnostiquée avec l'Equation IV.15 dans le modèle (bleu) (c) du facteur d'aridité diagnostiqué avec l'Equation IV.10 (noir) et simulé par le modèle (jaune) (d) du flux de chaleur latente observé au Sirta (noir) et reconstruit avec l'Equation IV.10 (rouge). La simulation est réalisée avec la physique CMIP6-L79, l'albédo imposé sur la maille du Sirta, un guidage en vent, température et humidité.

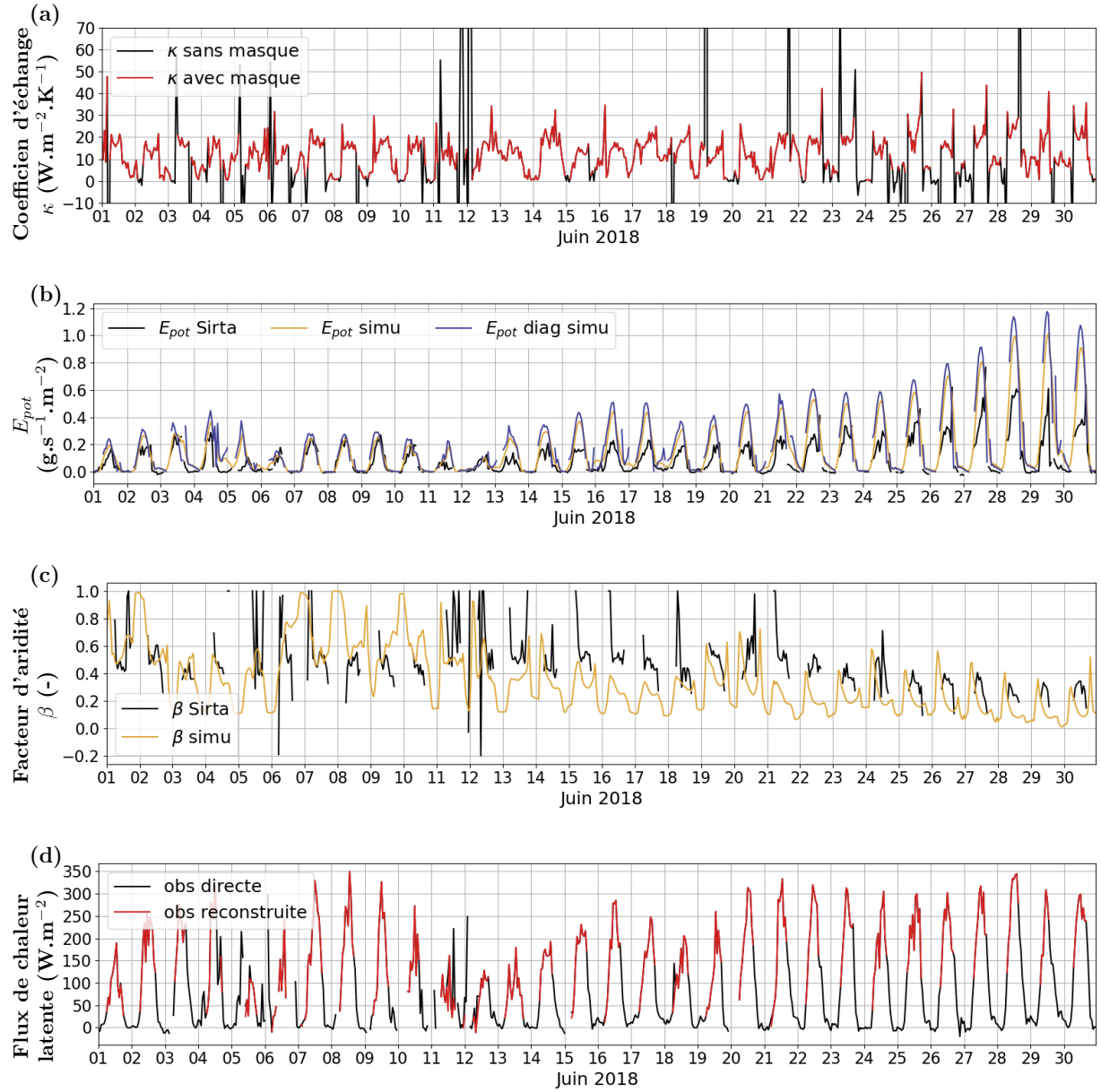


FIGURE IV.18 – Mêmes graphiques que la Figure IV.17 pour le mois de Juin 2018

les mois d'hiver questionnent néanmoins sur le masquage des données, qui a peut-être été trop restrictif et pourrait être revisité.

Le facteur d'aridité ainsi diagnostiqué présente des variations au cours de la journée, représentatives d'un cycle diurne, comme pour les journées du 25 au 30 juin 2018 par exemple. D'autres jours ont des facteurs d'aridité très bruités, comme les 5 et 6 juin et la 11 et 12 juin 2018, du fait de la présence de précipitations. On ne souhaite pas garder la variabilité horaire dans le facteur d'aridité que l'on va imposer sur la maille du SIRTa, ni même la variabilité journalière. En effet, à ces échelles de temps, le facteur d'aridité peut faire apparaître des pics associés aux événements de précipitations dont on sait qu'ils ne vont pas être simulés exactement au même moment dans les simulations. On souhaite donc imposer sur la maille du SIRTa un facteur d'aridité représentatif d'échelles de temps plutôt saisonnières.

Pour ce faire, on commence par calculer un facteur d'aridité par jour, comme moyenne pondérée des facteurs d'aridité horaires par le rayonnement solaire ciel clair descendant à la surface. Cette pondération revient à faire une pondération par angle solaire zénithal, pour privilégier les facteurs d'humidité de milieu de journée, là où l'évaporation est la plus forte. Ce facteur d'aridité journalier, tracé Figure IV.19 (a), présente encore de fortes variations en jour à jour, dont une partie est due à des événements de précipitations. Pour s'affranchir de cette variabilité, on applique une moyenne glissante sur les données journalières du facteur d'aridité. Plusieurs longueurs de moyennes glissantes sont testées (Figure IV.19 (a)) et c'est une moyenne glissante de 21 jours (10 jours avant et 10 jours après) qui est retenue pour le diagnostic final du facteur d'aridité. Cette moyenne permet en effet de s'affranchir de la variabilité journalière. *[Cela nous informe au passage qu'il faut 20 jours pour oublier le signal journalier du beta ?]* *[On en a assez dit. Je pense que c'est plus intéressant de commenter le fait qu'il varie moins que celui du modèle. Mais j'imagine que tu le fais plus loin.]*

Au vu du nombre important de données manquantes les mois d'hiver, ainsi que début 2017 où les flux turbulents n'ont pas pu être observés pendant plusieurs mois, une gestion particulière des données manquantes est appliquée lors du calcul de la moyenne glissante. Pour calculer le facteur d'aridité d'un jour j , on utilise toujours 10 valeurs du facteur d'aridité journalier avant le jour j et 10 valeurs après le jour j . Si une donnée est manquante entre le jour $j + 1$ et le jour $j + 10$, on va alors utiliser la valeur du jour $j + 11$ pour calculer le facteur d'aridité du jour j . Si n données sont manquantes alors on va utiliser les valeurs des facteurs d'aridités des jours $j + 11$ à $j + n + 10$. Et ce jusqu'à avoir 10 valeurs du facteur d'aridité dans l'intervalle $\llbracket j + 1; j + n \rrbracket$. On procède de la même manière pour les valeurs utilisés dans l'intervalle $\llbracket j - n; j - 1 \rrbracket$. Les facteurs d'aridités tracés Figure IV.19 (a) utilisent cette gestion des données manquantes.

Une erreur dans le diagnostic du facteur d'aridité au SIRTa a été détectée après la réalisation des expériences de tuning. Cette erreur porte sur la prise en compte des données masquées dans le calcul du facteur d'aridité journalier pondéré par le flux solaire ciel clair descendant. L'écart entre le facteur d'aridité avec et sans l'erreur est montré Figure IV.19 (a) *[montrer éventuellement l'impact sur la reconstruction du flux latent]*. L'erreur conduit à une sous-estimation systématique du facteur d'aridité pondéré lorsqu'il y a des données masquées dans la journée considérée. Elle est d'autant plus importante que le nombre de données masquées est important. C'est pourquoi elle conduit à une sous-estimation importante des facteurs d'aridité pendant les mois d'hiver, qui comportent beaucoup de données masquées, avec une sous-estimation du facteur d'aridité atteignant les 0,2 pendant le mois de janvier. Par manque de temps, aucune simulation avec le facteur d'aridité corrigé n'a été effectuée. Un test de sensibilité au facteur d'aridité a néanmoins été réalisé avec la configuration finalement choisie (résumée Section IV.7), où le facteur d'aridité est imposé à 0,2 toute l'année. Les différences sur les métriques utilisées dans l'exercice de tuning sont inférieures à trois fois la variabilité interne sur ces métriques, évaluées comme l'écart-type de la dispersion d'un ensemble aux états initiaux atmosphériques perturbés

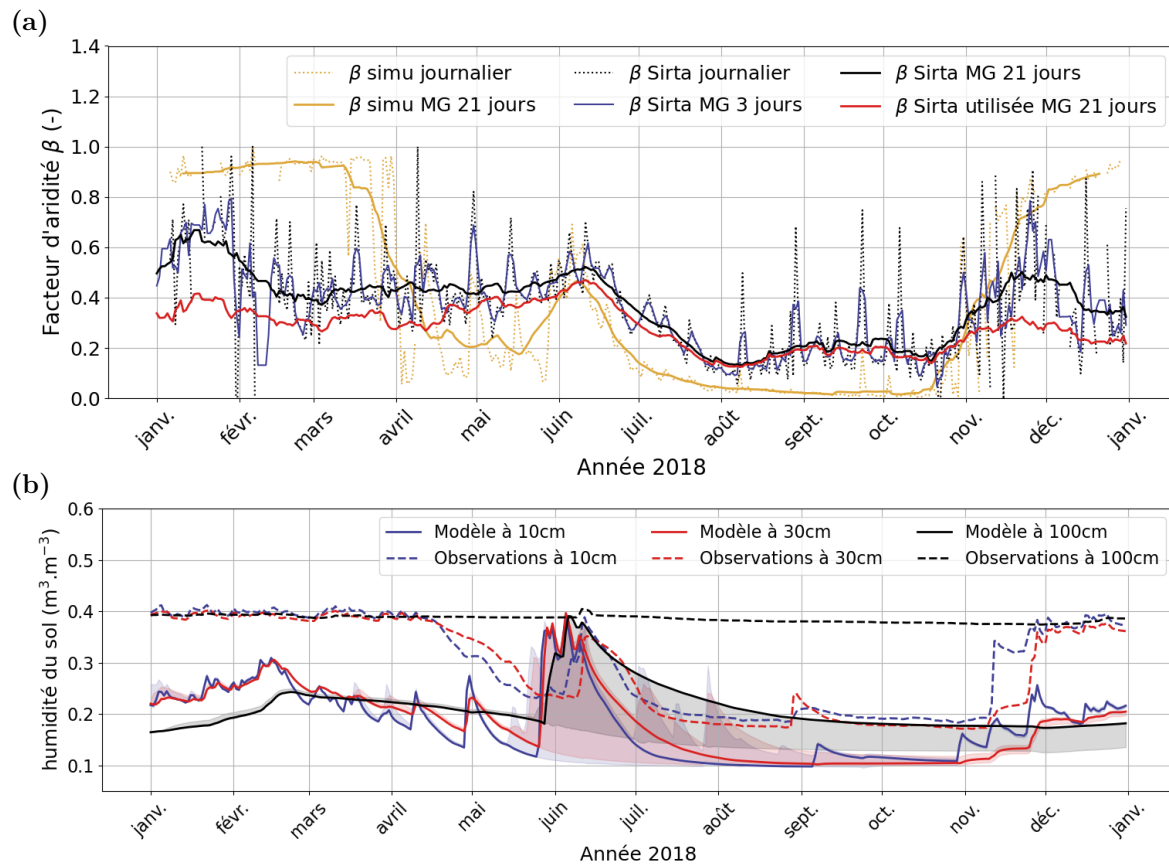


FIGURE IV.19 – Evolution journalière sur l'année 2018 (a) du facteur d'aridité β et (b) de l'humidité du sol. Figure (a) : le facteur d'aridité simulé et post-traité (masqué masque-par-variable et moyenne horaire pondérée par le flux solaire descendant ciel-clair) est tracé en jaune, en pointillé pour les valeurs journalières et en trait plein pour la moyenne glissante à 21 jours. Les autres courbes montrent le facteur d'aridité diagnostiqué au SIRTA, en moyenne pondérée journalière (pointillés noir), moyenne glissante de 3 jours (bleu foncé) et moyenne glissante de 21 jours (noir). Le facteur d'aridité utilisé dans la suite est celui tracé en rouge, dont le calcul comporte une erreur. Figure (b) : humidité du sol (bleu) à 10 cm de profondeur (rouge) à 30 cm de profondeur (noir) à 100 cm de profondeur ; en trait plein dans la simulation de contrôle, et en pointillé dans les observations du SIRTA. L'étendue du PISE où les états initiaux atmosphériques sont perturbés est montré en ombrages bleu, rouge et noir. La simulation de contrôle est réalisée avec la physique CMIP6-L79, l'albédo imposé sur la maille du SIRTA, et un guidage en vent température et humidité. Les humidités du sol de la simulation sont post-traitées avec le masque-par-variable.

[Est-ce que tu saurais mettre en ombrage jaune l'ensemble des facteurs d'aridité simulés, correspondant aux grisées du contenu en eau du sol?]

(Section V.1.3). Les valeurs sont résumées Table IV.3.

IV.6.1.3 Comparaison du facteur d'aridité diagnostiqué et simulé au SIRTA

Le facteur d'aridité diagnostiqué au SIRTA est comparé avec celui simulé dans la configuration de contrôle (physique CMIP6-L79, albédo imposé sur la maille du SIRTA, guidage en vent, température et humidité). Le facteur d'aridité simulé est post traité de manière suivante : le facteur d'aridité horaire est masqué lorsque les observations de flux sensible sont manquantes ou que le flux sensible simulée est négatif. Puis, le facteur d'aridité journalier est calculé comme la moyenne pondéré des facteurs d'aridité horaires (non masqués) par le flux solaire descendant

Métrique	Moyenne du $CRESW_{sfc}$ du 22 mai au 12 juin 2018	Moyenne du $CRESW_{sfc}$ du 13 au 30 juin 2018	Moyenne des précipitations du 22 mai au 12 juin 2018
Valeur de la simulation de contrôle	-77,82 W.m ⁻²	-41,91 W.m ⁻²	8.13 mm.j ⁻¹
Valeur de la simulation $\beta=0.2$	-72,83 W.m ⁻²	-39,46 W.m ⁻²	7,94 mm.j ⁻¹
Variabilité interne	5,06 W.m ⁻²	0,99 W.m ⁻²	0,37 mm.j ⁻¹

TABLE IV.3 – Impact de la valeur du facteur d’aridité imposé dans les simulations, sur les métriques choisies pour le tuning (Chapitre V). La simulation de contrôle est réalisée avec la configuration finale du modèle, décrite Section IV.7, où le facteur d’aridité imposé comporte une erreur. Le même configuration est utilisée pour la simulation de sensibilité, avec une valeur du facteur d’aridité imposé égale à 0,2 sur toute l’année 2018. La variabilité interne est calculée comme l’écart-type de la dispersion d’un ensemble de simulations aux états atmosphériques perturbés réalisé avec la configuration de contrôle.

ciel-clair. La Figure IV.19 (a) montre le facteur d’aridité ainsi post-traité dans la simulation de contrôle, ainsi que ce même facteur d’aridité moyennée avec une moyenne glissante de 21 jours. On vérifie ainsi que dans la simulation, la moyenne glissante de 21 jours permet de s’affranchir de la variabilité du facteur d’aridité due aux événements de pluies.

Le facteur d’aridité simulé, en moyenne glissante sur 21 jours, présente un cycle saisonnier bien plus marqué que celui diagnostiqué dans les observations du SIRTa (Figure IV.19 (a)). Cela se traduit par des valeurs élevées du facteur d’aridité en hiver (autour de 0.9) et très faible en été (avec un minimum à 0.016) ; mais aussi par des réponses fortes et rapides aux conditions atmosphériques. C’est le cas de la diminution du facteur d’aridité, caractéristique du printemps quand le facteur passe de 0,8 mi-mars à presque 0,2 mi-avril. Les événements orageux de fin mai à mi-juin conduisent pour leur part à une augmentation rapide du facteur d’aridité, qui passe de 0,2 à 0,4 en deux semaines, suivie ensuite d’une longue période d’assèchement, avec un facteur d’aridité atteignant 0,12 début juillet et 0,016 à la mi-septembre.

Les valeurs du facteur d’aridité simulées en hiver sont attendues, même si le cycle diurne que présente ce dernier est une surprise (Figure IV.17). Il est en revanche étonnant (1) que l’humidité du sol simulée soit si loin de la saturation en hiver ((Figure IV.19 (b)) malgré un facteur d’aridité proche de 1 ; (2) que le facteur d’aridité diagnostiqué au SIRTa soit si faible en hiver. Les valeurs du facteur d’aridité atteintes de juillet à octobre dans la simulation paraissent très faibles, même si les régions du Nord-Est de la France sont traversées par une sécheresse record pendant cette période (Météo-France, 2019). Ce faible facteur d’aridité est accompagné d’humidité du sol très faible, atteignant la valeur de l’humidité résiduelle dans les trente premiers centimètres (qui correspond à l’humidité minimale qu’un sol peut atteindre), d’un très faible flux latent et d’un rapport de Bowen très fort (caractéristiques des déserts si supérieur à 10 ou des climat semi-aride si entre 2 et 6). Ce déficit de flux de chaleur latent en été au SIRTa ont déjà été pointés par les études de Campoy et al. (2013); Coindreau et al. (2007); Cheruy et al. (2013). Il pourrait s’expliquer par l’hydrologie spécifique du plateau de Saclay, avec une nappe perchée qui pourrait alimenter en eau les couches superficielles du sol et ainsi booster l’évaporation de cette région. Il pourrait également s’expliquer par un modèle qui surestime l’évaporation potentielle en été, en simulant une atmosphère trop chaude et/ou trop sèche. Enfin, ce déficit d’évaporation pourrait être dû à des défauts du modèle d’hydrologie, qui ne serait pas capable de simuler correctement les échelles de temps saisonnières de l’hydrologie du sol, avec des "effets mémoires" de l’humidité du sol trop courts.

IV.6.2 Simulations avec un facteur d'aridité imposé

IV.6.2.1 Derniers ajustements sur la configuration

Imposer le facteur d'aridité sur une région autour du site du SIRTa

Dans les premiers tests effectués en imposant le facteur d'aridité préalablement diagnostiqué au SIRTa sur la maille du modèle contenant le site du SIRTa, certaines incohérences dans les variables proches de la surface apparaissaient dans les simulations. Notamment, l'augmentation de flux de chaleur latentes en été provoqué par le β imposé sur la maille du SIRTa, ne conduisait pas à une augmentation significative de l'humidité spécifique de l'air proche de la surface en moyenne mensuelle, et qu'à une très faible diminution de la température à 2 mètres, donc de l'humidité saturante à 2 mètres, malgré la diminution significative du flux de chaleur sensible et de la température du sol (courbes rouges, Figure IV.20).

Ces incohérences pourraient s'expliquer par le fait que l'air advecté sur la maille du SIRTa vient de mailles ayant un facteur d'évaporation beaucoup trop faible, et en tous cas nettement plus faible que celui imposé sur la maille du SIRTa (au printemps et en été), et contient ainsi beaucoup moins de vapeur d'eau que s'il avait été en contact avec un facteur d'évaporation, donc un flux d'évaporation, plus fort pendant cette période.

Pour contourner ce problème, nous avons alors fait le choix d'imposer le facteur d'aridité diagnostiqué au SIRTa sur les mailles présentent dans un rayon de 100 km autour de celle du SIRTa (Figure IV.21), ce qui est suffisant pour que les masses d'airs s'équilibrent avec la surface. On voit notamment une nette augmentation de l'humidité spécifique de l'air à la surface et à 2 mètres, ainsi qu'une diminution de la température à 2 mètres et de l'humidité saturante à 2 mètres (courbes bleues, Figure IV.20). Cela conduit à une humidité relative à 2 mètres beaucoup plus proche des observations (courbes noires), ainsi qu'à une légère baisse du flux de chaleur latente, consécutive de la légère baisse d'évaporation potentielle due aux changements d'humidité et de température de l'air.

Imposer une humidité du sol à saturation à partir d'une certaine profondeur

L'augmentation d'évaporation provoquée par le fait d'imposer le facteur d'aridité entraîne une diminution de l'humidité du sol, qui atteint des valeurs négatives dans certaines sous-colonnes hydrologique (Figure IV.22 (a)). Cela rétroagit sur l'atmosphère de deux manières. Tout d'abord, l'humidité du sol impacte l'inertie thermique du sol, qui en retour influence l'amplitude du cycle diurne de la température de l'air proche de la surface (Figure IV.23 (a)). Ensuite, le manque d'eau dans le sol affecte la croissance des plantes, diminuant leur indice de surface foliaire (Figure IV.23 (c)), qui rentre dans le calcul de la longueur de rugosité (Figure IV.23r (d)). Cela provoque une diminution du freinage de l'air proche de la surface (Figure IV.23 (b)), et affecte la turbulence atmosphérique.

Pour palier à ce problème, on impose que le sol soit saturé en eau à partir d'une certaine profondeur. La profondeur est choisie suivant l'étude de Campoy et al. (2013), qui montre qu'imposer une humidité du sol saturée à partir de 1,3 mètres de profondeur améliore la simulation de l'humidité du sol sur la période 2002-2009 dans les 50 premiers centimètres du sol, l'humidité n'étant pas observée plus profondément pendant cette période. Cela semble indiquer que la nappe perchée est située autour de 1,3 mètres sur le site du SIRTa. Imposer cette humidité du sol nécessite d'augmenter la discrétisation du sol, car cela crée un fort gradient d'humidité dans une zone où le sol est faiblement discrétisé. On choisit d'augmenter la discrétisation du sol de 11 à 22 couches, en gardant une profondeur de sol de 2 mètres pour l'hydrologie. Cette discrétisation est proche de celle utilisée par Campoy et al. (2013), et raffine la discrétisation à 11 couches à partir de la sixième couche, avec des couches de 12,5 cm d'épaisseurs. On impose

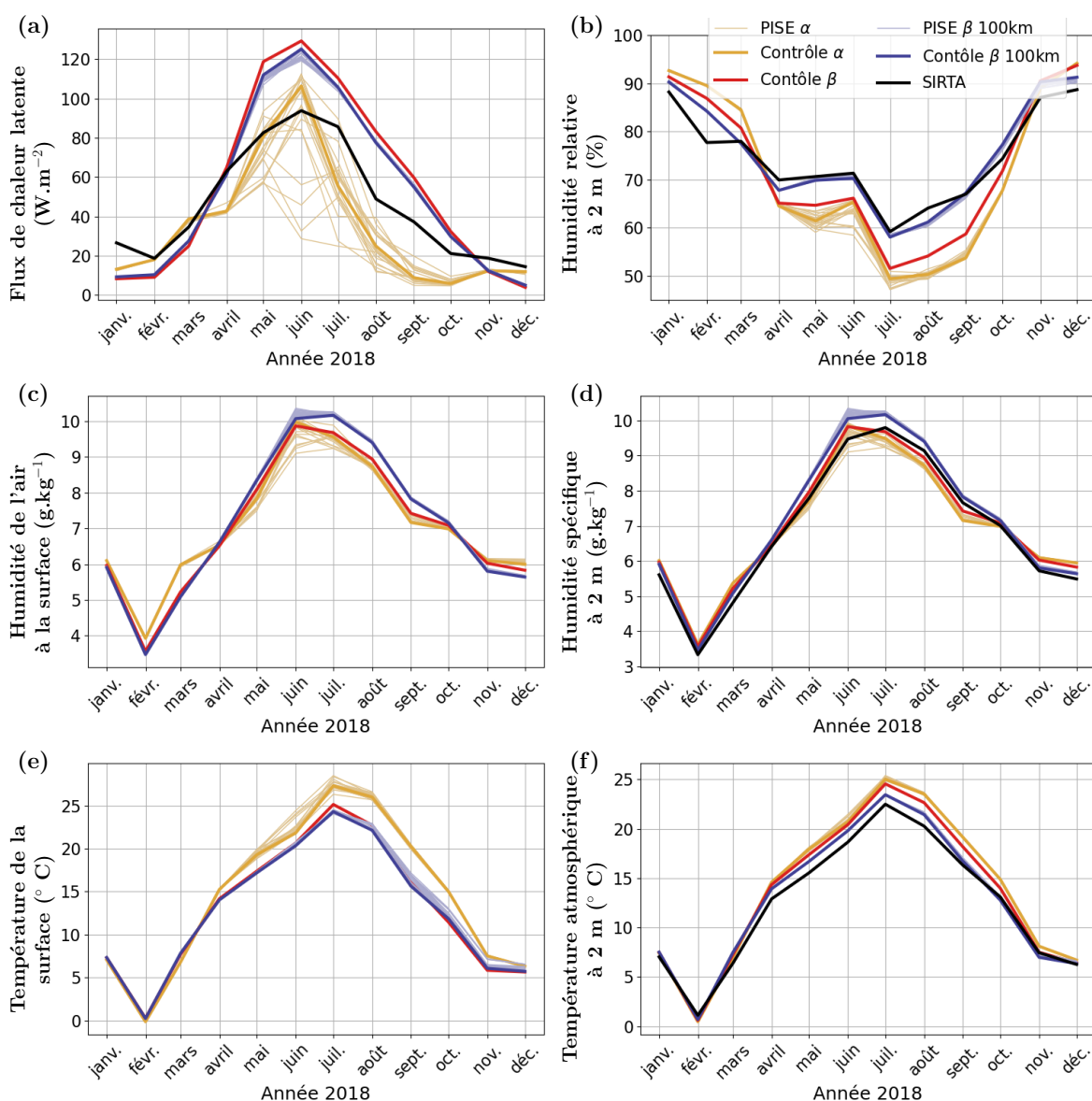


FIGURE IV.20 – Evolution en moyenne mensuelle sur l'année 2018 et sur la maille du SIRTa (a) du flux de chaleur latente (W.m^{-2}) (b) de l'humidité relative à 2 mètres (%) (c) de l'humidité de l'air à la surface (g.kg^{-1}) (d) de l'humidité de l'air à 2 mètres (g.kg^{-1}) (e) de la température de la surface ($^{\circ}\text{C}$) (f) de la température atmosphérique à 2 mètres ($^{\circ}\text{C}$). Les observations du SIRTa, lorsqu'elles sont disponibles, sont tracées en noir. La simulation de contrôle (physique CMIP6-L79, guidage en vent, température, et humidité, albédo imposé sur la maille du SIRTa) et le PISE associé est tracé en jaune. La simulation avec le facteur d'aridité imposé sur la maille du SIRTa est tracé en rouge. La simulation avec le facteur d'aridité imposé dans un rayon de 100 km autour de la maille du SIRTa, et son PISE associé, est tracé en bleu. Les variables pour lesquels des observations sont disponibles sont post-traitées avec le masque-par-variable. Les autres ne sont pas masquées.

de plus l'indice de surface foliaire, avec une climatologie d'indice foliaire interpolée sur chaque jours de l'année. L'humidité du sol de cette nouvelle configuration est tracée Figure IV.22 (b). Il n'y a plus d'humidité négative du sol, bien qu'elle atteigne la valeur de l'humidité résiduelle une bonne partie de l'été. L'amplitude du cycle saisonnier de l'humidité à 1 mètre de profondeur est fortement réduite. L'humidité à cette profondeur reste cependant bien en dessous des observations, ce qui pourrait indiquer que la nappe perchée est un réalité un peu plus haute que

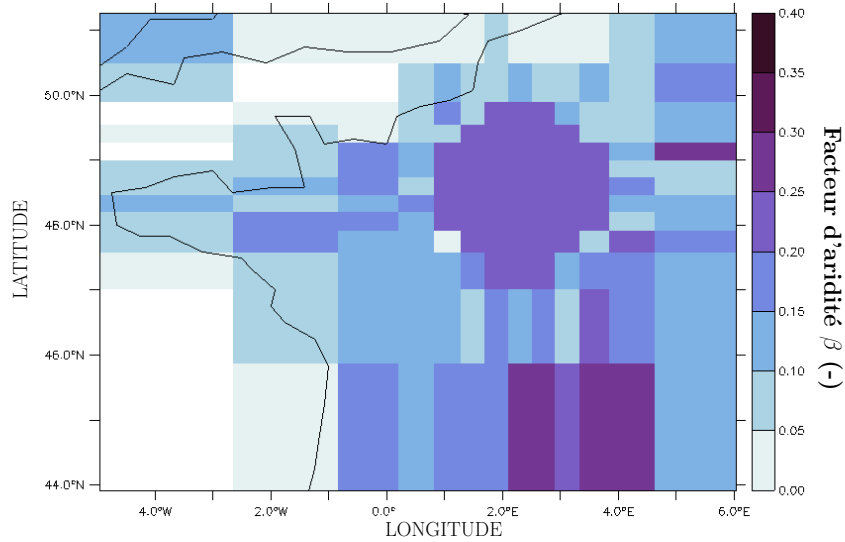


FIGURE IV.21 – Carte du facteur d'aridité, en moyenne mensuelle pour le mois de juillet 2018, dans la simulation où le facteur d'aridité est imposé dans les mailles situées dans un rayon de 100 km autour de la maille du SIRTa. La simulation est effectuée avec la physique CMIP6-L79, un guidage en vent, température et humidité et l'albédo imposé sur la maille du SIRTa.

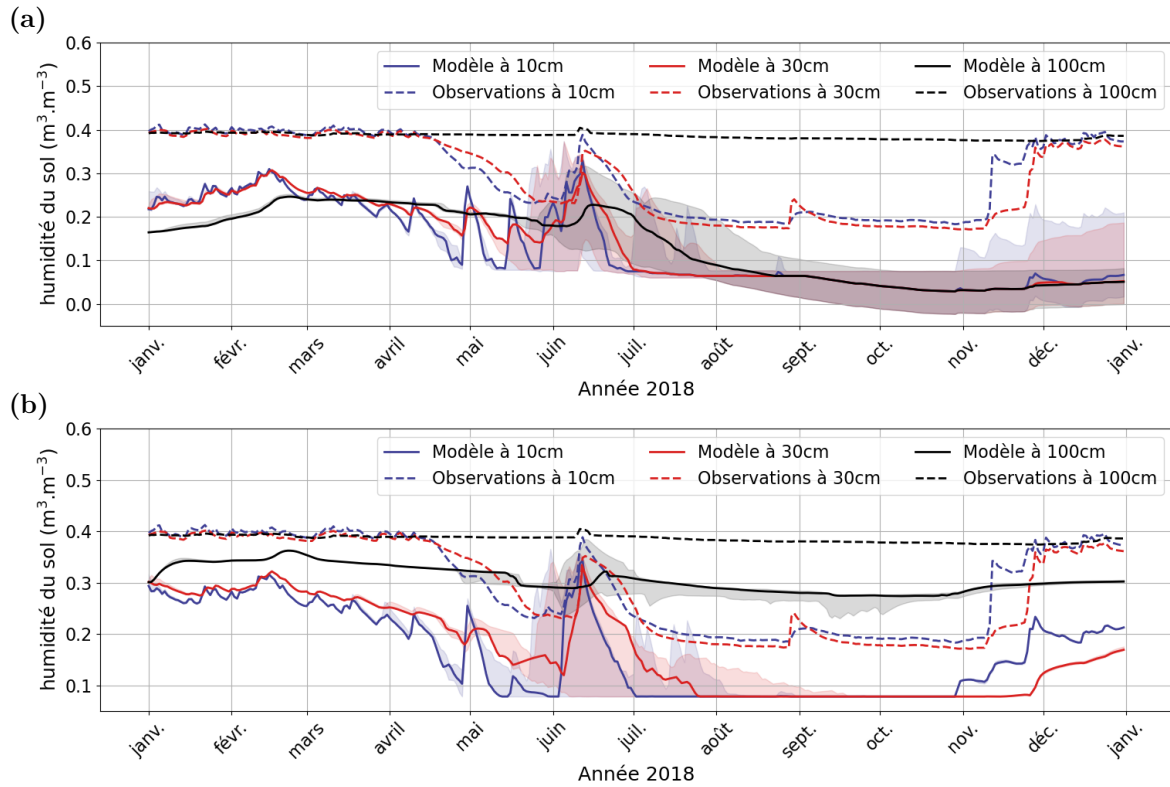


FIGURE IV.22 – Evolution en moyenne journalière de l'humidité du sol sur l'année 2018 à 10 cm (bleu), 30 cm (rouge) et 100 cm (noir) de profondeur, sur la maille du SIRTa : (a) pour la configuration avec β imposé sur les mailles dans un rayon de 100 km autour de la maille du SIRTa ; (b) pour la configuration avec en plus l'humidité du sol imposée à saturation en dessous de 1,3 mètres de profondeur et une discrétisation du sol à 22 couches. Les courbes en pointillé indiquent les observations au SIRTa, les courbes pleines indiquent la simulation de contrôle, avec physique CMIP6-L79, albédo imposé sur la maille du SIRTa et guidage en vent température et humidité (Table IV.2), et les ombrages bleu, rouge et gris indiquent l'étendue du PISE réalisé pour chacune des configurations. L'humidité du sol à saturation est égale à $0,43 \text{ m}^3.\text{m}^{-3}$ sur la maille du SIRTa du modèle. Les simulations sont post-traitées avec le masque-par-variable.

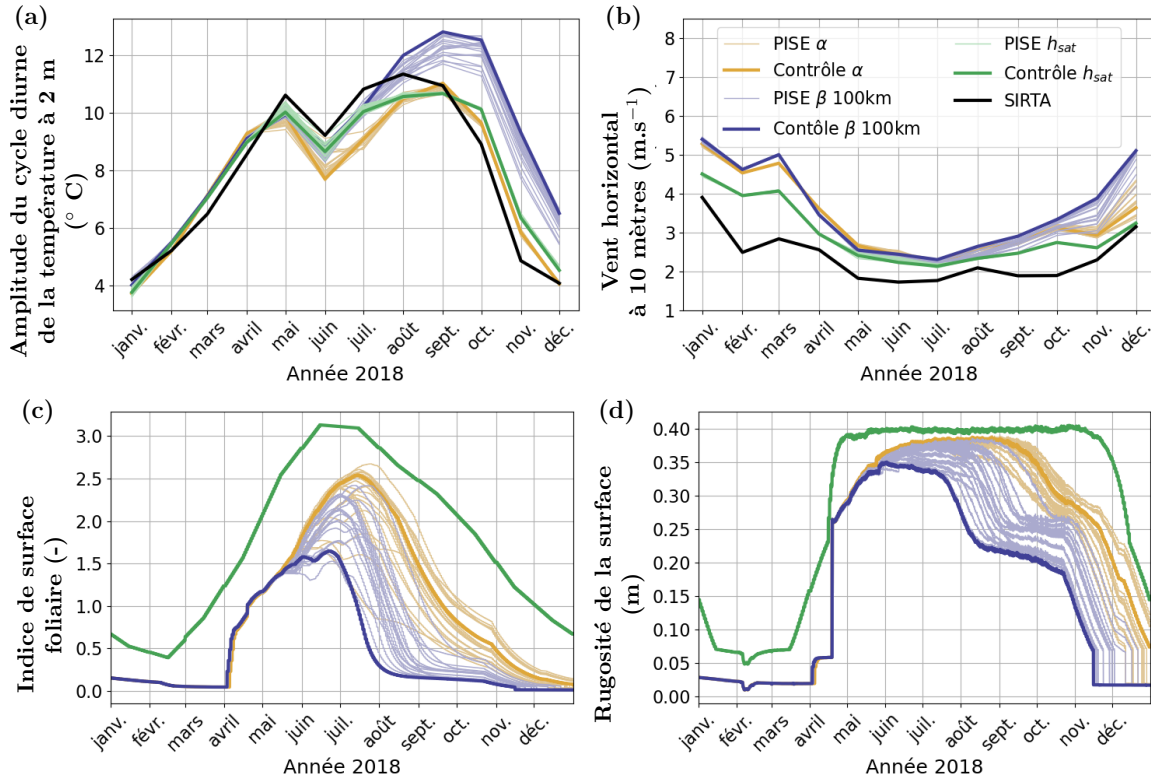


FIGURE IV.23 – Evolution sur l’année 2018 (a) de l’amplitude du cycle diurne de la température atmosphérique à 2 mètres, en moyenne mensuelle ($^{\circ}\text{C}$) (b) du module du vent horizontal à 10 mètres (m.s^{-1}) en moyenne mensuelle (d) de l’indice de surface foliaire moyen sur la maille, en moyenne journalière (sans unité) (e) de la hauteur de rugosité de la surface (m). La simulation de contrôle (physique CMIP6-L79, guidage en vent température et humidité, et albédo imposé sur la maille du SIRTA) et son PISE associé sont tracés en jaune. La simulation ayant en plus un facteur d’aridité imposé sur un rayon de 100 km autour de la maille du SIRTA, avec son PISE associé, sont tracés en bleu. La simulation ayant en plus une humidité du sol à saturation (notée h_{sat}) à partir de 1,3 mètres, un sol discrétisé à 22 couches et un indice de surface foliaire journalier imposé, ainsi que son PISE, sont tracés en vert. Les figures (a) et (b) utilisent le post-traitement masque-par-variable, et les figures (c) et (d) n’utilisent pas de masque.

1,3 mètres sur le site du SIRTA. On note également que la dispersion de PISE disparaît durant l’automne sur le premier mètre de profondeur du sol.

Relâcher les constantes de guidage de la configuration

Enfin, parce le guidage en vent, température et humidité a stabilisé numériquement la configuration par rapport à la configuration guidée en vent uniquement, on relâche les constantes de guidages à l’extérieur de la zone zoomée à 6 heures (au lieu de 3h), et à 100 jours à l’intérieur de la zone zoomée, c’est-à-dire quasiment aucun guidage (Table IV.4). Cela provoque une augmentation de la variabilité interne des variables atmosphériques nuageuses au mois de juin (Figure IV.24 (c) et (d)), ainsi qu’un léger, mais significatif, changement d’état moyen atmosphérique (légèrement plus chaud au plus sec sur les variables atmosphériques à deux mètres, Figure IV.24 (a) et (b)). Avec le recul, et au vue la raideur de la transition du zoom, il est possible qu’un guidage léger du vent, et éventuellement de la température et de l’humidité, soit nécessaire à l’intérieur de la zone zoomée.

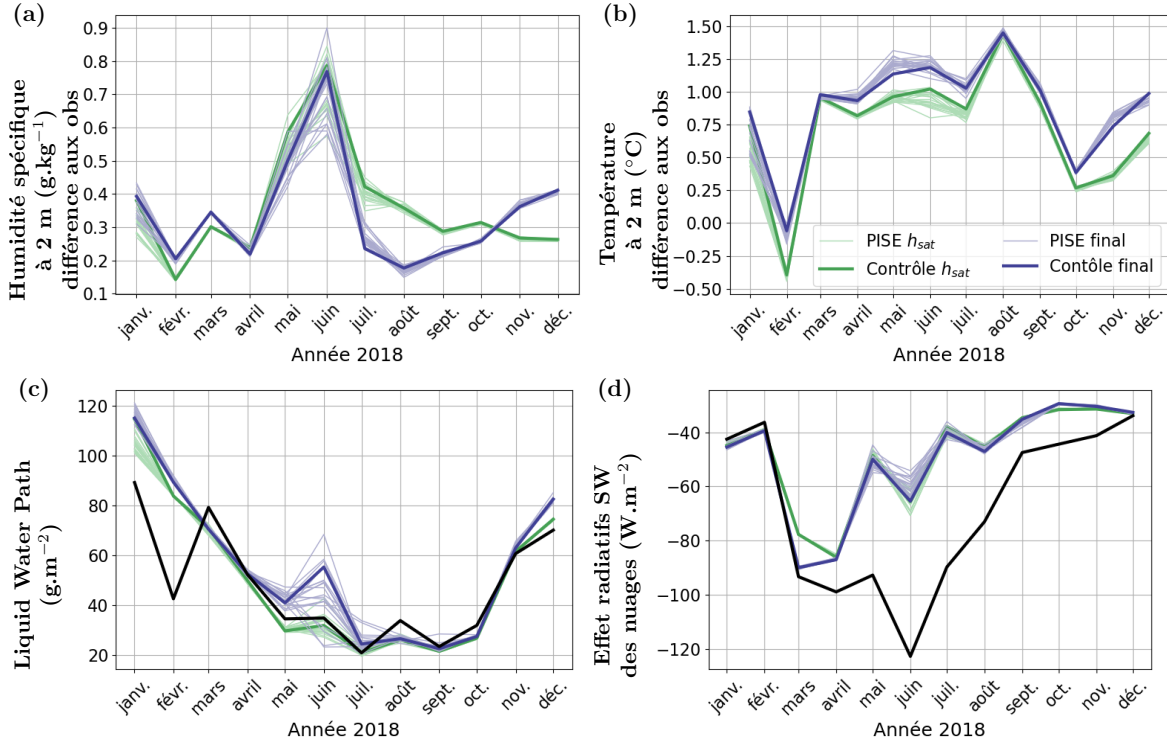


FIGURE IV.24 – Evolution en moyenne mensuelle sur l’année 2018 : (a) du biais (différence aux observations du SIRTa) de l’humidité spécifique à 2 mètres (g.kg^{-1}) ; (b) du biais de température atmosphérique à 2 mètres ($^{\circ}\text{C}$) ; (c) de la quantité d’eau liquide intégrée sur la verticale (*Liquid Water Path*) en g.m^{-2} ; (d) de l’effet radiatif solaire des nuages (W.m^{-2}). Les deux ensembles tracés sont issus de simulations réalisées avec : la physique CMIP6-L79, l’albédo imposée sur la maille du SIRTa, le facteur d’évaporation imposé dans un rayon de 100 km autour de la maille du SIRTa, l’humidité imposée à saturation sous 1,3 mètres du sol, et l’indice de surface foliaire imposé. La simulation avec l’ancien guidage (Table IV.2) et son PISE associé sont tracés en vert, la simulation avec le nouveau guidage (Table IV.4) et son PISE associé sont tracés en bleu.

IV.6.2.2 Impact du découplage de l’hydrologie au SIRTa

Effets du β imposé sur le couplage à la surface

Le fait d’avoir imposé un β saisonnier coupe le couplage rapide de l’humidité du sol de surface avec l’atmosphère, caractérisé par une décroissance exponentielle de l’humidité du sol dans les jours qui suivent un événement de pluie, comme on peut le voir par exemple après les pluies de juin sur la Figure IV.19 (b).

Imposer le facteur d’aridité, que ce soit sur une maille ou sur la région autour de la maille du SIRTa, améliore grandement la simulation du rapport de Bowen, c’est-à-dire du rapport du flux de chaleur sensible sur le flux de chaleur latente. Ce dernier est largement surestimé d’avril à octobre lorsque le facteur d’aridité n’est pas imposé, et présente une très forte variabilité interne, qui disparaît lorsque le facteur d’aridité est imposé (Figure IV.25 (a) et (b)). La variabilité persistante des mois de Janvier à Mars puis de Novembre à Décembre s’explique quant à elle par les valeurs faibles de ces flux, conduisant à des rapports non significatifs, du fait soit d’un des deux flux turbulent négatifs en moyenne mensuelle, soit de la division par un flux proche de zéro voire des deux flux proches de zéro. Le rapport de Bowen est une mesure de la répartition (ou du partitionnement) des flux turbulents de surface en flux de chaleur sensible et flux de chaleur latent. Le fait que ce partitionnement soit proche de celui observé au SIRTa

Nom du paramètre	Variable guidée	Constante de temps
τ_{\min_u} , τ_{\min_v} , τ_{\min_T} , τ_{\min_Q}	Vent zonal et méridien, température et humidité relative à l'extérieur du zoom	0,25 jours = 6 heures
τ_{\max_u} , τ_{\max_v} , τ_{\max_T} , τ_{\max_Q}	Vent zonal et méridien, température et humidité relative à l'intérieur du zoom	100 jours

TABLE IV.4 – Constantes de temps de guidage utilisées dans la configuration finale du modèle

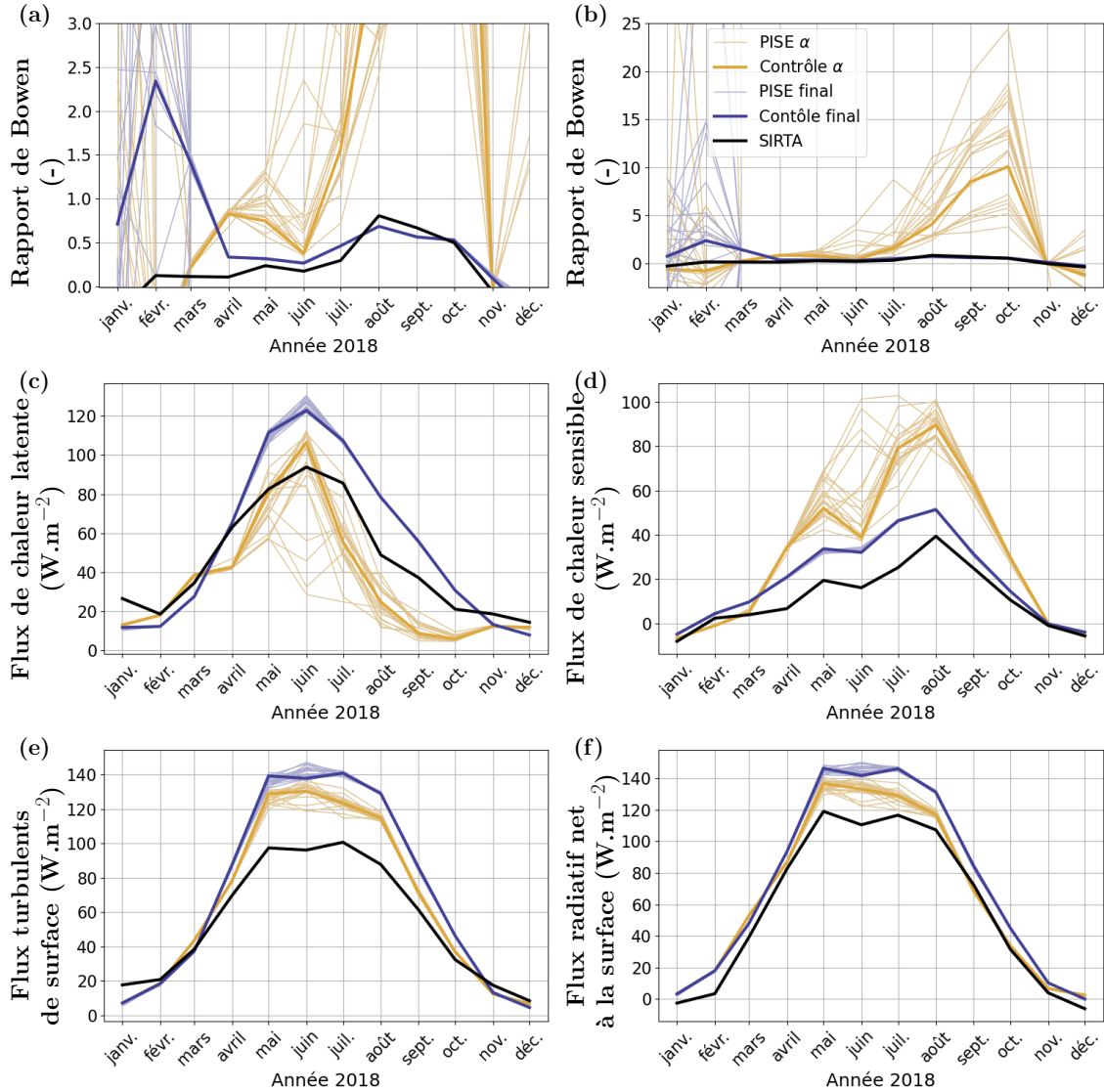


FIGURE IV.25 – Evolution en moyenne mensuelle sur l'année 2018 : (a) et (b) du rapport de Bowen (rapport du flux de chaleur sensible et latent) avec un axe des ordonnées différent ; (c) du flux de chaleur latent (W.m^{-2}) ; (d) du flux de chaleur sensible (W.m^{-2}) ; (e) de la somme des flux turbulents de surface (c'est-à-dire de la somme du flux de chaleur latente et du flux de chaleur sensible, W.m^{-2}) ; (f) du flux radiatif net à la surface (W.m^{-2}). En noir sont tracées les observations du SIRTa, en jaune la simulation de contrôle (physique CMIP6-L79, guidage en vent, température et humidité (Table IV.2) et albédo imposé sur la maille du SIRTa) avec son PISE associé, en bleu foncé la configuration de contrôle finale (tel que récapitulée Section IV.7) avec son PISE associé. Les figures (a), (b), (c) et (d) sont tracées avec de simulations post-traitées avec le masque-par-variable ; les figures (e) et (f) avec le masque-global, et ce pour faciliter les comparaisons dans la décomposition du bilan d'énergie.

d'avril à octobre lorsque le facteur d'aridité est imposé valide, a posteriori, la pertinence de la configuration développée ici. En particulier, cela valide la méthode du diagnostique du facteur d'aridité développée Section IV.6.1 et la manière dont il a été imposé dans la configuration finale (Section IV.6.2.1).

Néanmoins, bien que ce partitionnement soit proche de celui observé au SIRTa, les flux de chaleurs latents et sensibles simulés sont surestimés dans les simulations avec facteur d'aridité imposé. Le flux de chaleur latente est significativement augmenté d'avril à octobre dans ces simulations, en accord avec l'augmentation de facteur d'aridité lorsqu'il est imposé durant cette période (Figure IV.19 (a)). Le flux de chaleur sensible est, quant à lui, significativement diminué sur la même période (Figure IV.25 (c) et (d)). La surestimation du flux turbulent total à la surface (Figure IV.25 (e)) vient contrebalancer une surestimation du rayonnement net à la surface ; en considérant que le flux de chaleur partant dans le sol est négligeable dans le bilan d'énergie à la surface quand on considère des moyennes mensuelles (Equation IV.2). A noter au passage que les flux radiatifs observés au SIRTa l'été semblent excéder les flux turbulents d'une quinzaine de W.m^{-2} , suggérant des biais dans les observations.

En décortiquant les différents termes du bilan radiatif à la surface au SIRTa, tracés en moyenne mensuelle Figure IV.26, on peut conclure que c'est le manque d'effet radiatif dans le solaire qui domine ce biais. En effet, dans le spectre solaire, le biais peut venir soit de l'albédo de la surface, qui est ici imposé, soit du flux descendant ciel-clair, qui n'est que faiblement biaisé, soit de l'effet radiatif des nuages, qui est ici très sous-estimé à partir d'avril 2018. La surestimation du rayonnement solaire net à la surface est légèrement compensée par le rayonnement infra-rouge net à la surface, qui reste trop négatif, bien qu'amélioré dans les simulations avec β imposé. Le rayonnement infra-rouge net à la surface est beaucoup plus dépendant du couplage à la surface que le rayonnement solaire net à la surface. En effet, le rayonnement infra-rouge montant est directement lié à la température de la surface par la loi de Stefan-Boltzmann, et cette dernière résulte directement du bilan d'énergie à la surface. Sa simulation est cependant améliorée lorsque le facteur d'évaporation est imposé. La simulation du rayonnement infra-rouge descendant à la surface est améliorée dans les simulations avec β imposé et est très proche du rayonnement infra-rouge descendant observé au SIRTa. Cependant, la décomposition de ce rayonnement en sa partie ciel-clair et nuageuse fait apparaître que ce rayonnement est bien simulé mais pour les mauvaises raisons : l'effet radiatif des nuages dans l'infrarouge est sous-estimé, et le rayonnement infra-rouge descendant ciel-clair est surestimé, à cause d'une surestimation de la température atmosphérique et de l'humidité spécifiques dans les basses couches atmosphériques au SIRTa. On note que la simulation des nuages est légèrement, mais significativement, impactée par le fait d'avoir imposé un facteur d'aridité sur une région autour du site du SIRTa mais pas lorsque le facteur d'aridité est imposé uniquement sur le site du SIRTa. Ce sont en particulier les nuages bas qui sont impactés, avec une diminution de leur couverture ainsi que de leur contenu en eau au mois de juillet, et une agmentation de leur couverture en automne (Figure IV.27 (a) et (b)).

Enfin, l'humidité relative à 2 mètres est très proche de celle observée au SIRTa, presque toute l'année, dans les simulations réalisées avec la configuration finale, dans laquelle le facteur d'aridité est imposé (Figure IV.27 (e)). La simulation de la température atmosphérique à deux mètres est également améliorée presque toute l'année dans la configuration finale, bien qu'elle reste surestimée d'1 à 1,4°C d'avril à septembre 2018 (Figure IV.27 (c)). L'humidité spécifique à deux mètres augmente, comme attendu au vue de l'augmentation du flux de chaleur latente, et est désormais légèrement surestimée sur toute l'année 2018, entre 0,2 et 0,8 g.kg^{-1} (Figure IV.27 (d)).

Effets du β imposé sur la variabilité interne

Imposer le facteur d'aridité sur la SIRTa, ou sur une région autour du SIRTa, provoque

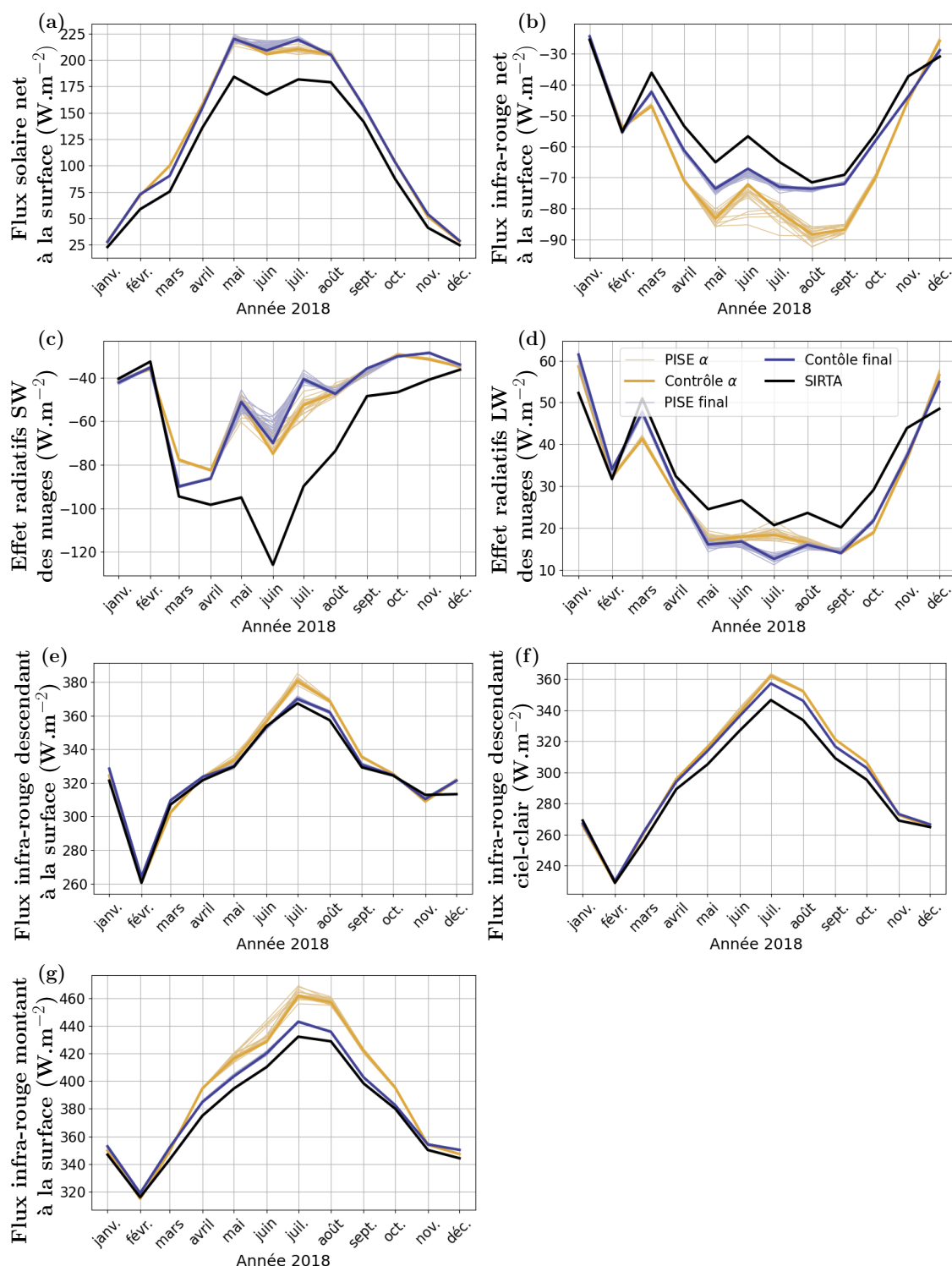


FIGURE IV.26 – Même figure que Figure IV.25 pour : (a) le flux solaire net à la surface (W.m^{-2}); (b) le flux infra-rouge net à la surface (W.m^{-2}); (c) l'effet radiatif solaire des nuages à la surface (W.m^{-2}); (d) l'effet radiatif infra-rouge des nuage à la surface (W.m^{-2}); (e) le flux infra-rouge descendant à la surface (W.m^{-2}); (f) le flux infra-rouge descendant à la surface en ciel clair (W.m^{-2}); (g) le flux infra-rouge montant à la surface (W.m^{-2}). Toutes les simulations sont ici post-traitées avec le masque-global.

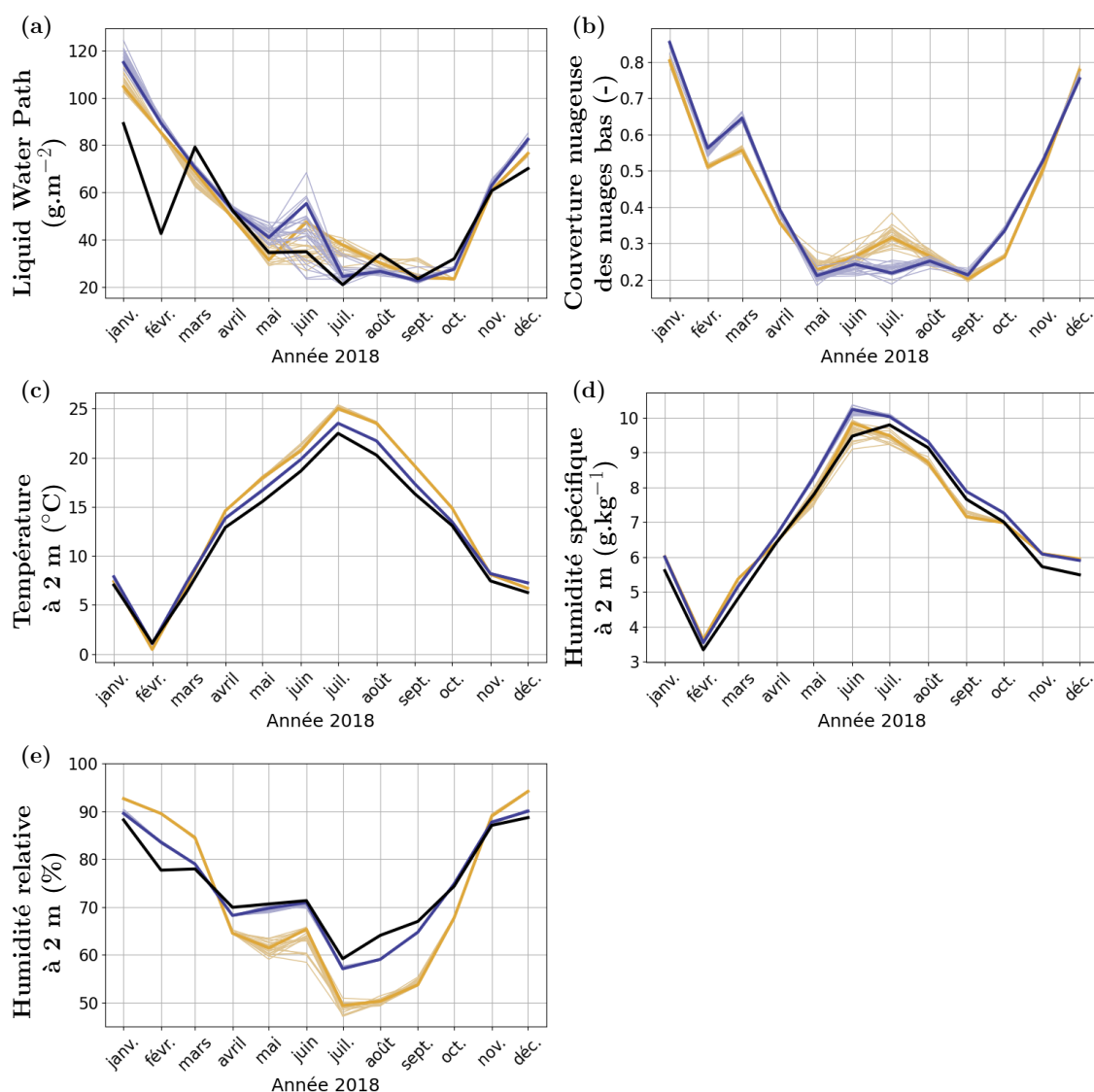


FIGURE IV.27 – Même figure que Figure IV.25 pour : (a) la quantité d’eau liquide intégrée sur la verticale (*Liquid Water Path* en g.m^{-2}) ; (b) la couverture nuageuse des nuages bas ; (c) la température atmosphérique à deux mètres ($^{\circ}\text{C}$) ; (d) l’humidité spécifique à deux mètres (g.kg^{-1}) ; (e) l’humidité relative à deux mètres (%). Toutes les simulations tracées ici sont post-traitées avec le masque-par-variable sauf pour la figure (b) où les données ne sont pas masquées.

une diminution de la variabilité interne d’un certain nombre de variables liés directement ou indirectement au facteur d’aridité. Les variables sensibles à l’humidité du sol et/ou à la rugosité de la surface voient aussi leur variabilité interne diminuer à partir d’août 2018, grâce à la diminution de la variabilité interne de l’humidité du sol (Figure IV.23 (b)) et de la rugosité de la surface et (Figure IV.23 (d)), documenté Figure IV.23 (a) et (b) pour l’amplitude du cycle diurne de la température atmosphérique à 2 mètres et le module du vent horizontal à 10 mètre.

La diminution de la variabilité interne du flux de chaleur latent, directement liée au fait d’avoir imposé le facteur d’aridité, est très importante, tout comme celle du flux de chaleur sensible et du rapport de Bowen (Figure IV.25 (b), (c) et (d)). Les flux radiatifs infra-rouges à la surface, descendant et montant, ciel-clair et nuageux, montrent également une diminution de leur variabilité interne dans la configuration finale du modèle, ce qui se répercute sur la variabilité interne du bilan radiatif à la surface et de la somme des flux turbulents de surface

(Figure IV.26). La variabilité des variables atmosphériques comme la température, l'humidité et l'humidité relative à 2 mètres, et plus généralement dans les basses couches atmosphériques, sont également diminuées (Figure IV.27).

Effet du β imposé sur la variabilité des précipitations

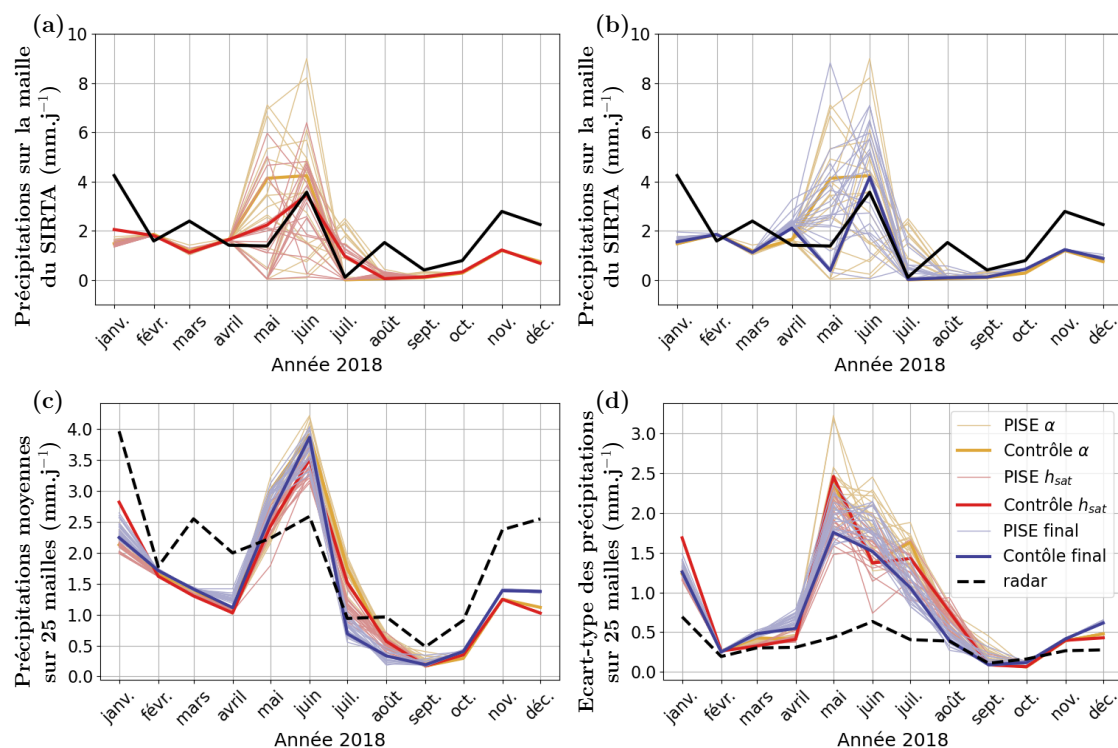


FIGURE IV.28 – Evolution en moyenne mensuelle sur l'année 2018 : (a) et (b) des précipitations sur la maille du SIRTa ; (c) des précipitations moyennes sur une région de 25 mailles autour du SIRTa (mm.j^{-1}) ; (d) de l'écart-type spatial des précipitations sur cette région (mm.j^{-1}). En jaune, la simulation de contrôle (physique CMIP6-L79, guidage en vent, température et humidité (Table IV.2) et albédo imposé sur la maille du SIRTa) et le PISE associé, en rouge la simulation avec β imposé dans un rayon de 100 km autour du site du SIRTa, l'humidité à saturation imposé à partir de 1,3 mètres et l'indice de surface foliaire imposé, et son PISE associé. En bleu foncé, la simulation avec la configuration finale et son PISE associé. En noir les observations du SIRTa, en noir pointillé le produit Coméphore intrpolé sur la grille du modèle. Les données des figures (a) et (b) sont post-traitées avec la masque-par-variable, et celles des figures (b) et (c) ne sont pas masquées.

Imposer le facteur d'évaporation dans un rayon de 100 km autour du site du SIRTa provoque une légère diminution de l'écart-type spatial des précipitations sur la région de 25 mailles autour du SIRTa sur les mois d'été (Figure IV.28 (d)), ce qui implique une légère diminution de leur variabilité interne sur la maille du SIRTa (Figure IV.28 (a) et (b)). A noter que le facteur d'évaporation est imposé sur toute la zone de 25 mailles (voir Figure IV.14 (b)). La configuration finale, dont le guidage est moins fort, voit (1) sa variabilité spatial des précipitations légèrement augmenter par rapport à la configuration ayant un guidage un peu plus fort (2) ses précipitations en moyennes sur 25 mailles légèrement augmenter au moins d'avril et diminuer aux mois de juillet et d'août (Figure IV.28 (c)), ce qui explique l'augmentation (respectivement la diminution) de l'écart-type des précipitations en avril (respectivement en juillet et août). L'écart-type des précipitations sur 25 mailles en mai-juin-juillet 2018 reste très surestimé par rapport à celui diagnostiqués dans les observations Coméphore, et ne permet toujours pas une évaluation des précipitations des mois d'été avec les données du SIRTa, justifiant l'utilisation de données

moyennes sur 25 mailles pour l'évaluation de la pluie.

Au final, la mise en place de la configuration a amélioré beaucoup de points permettant une meilleure exploitation des observations au SIRTa. La variabilité de la pluie à l'échelle du point de grille est réduite, mais seulement légèrement. Ceci suggère que la rétroaction avec l'eau du sol ne contribue que faiblement à cette variabilité et qu'il s'agit au premier ordre d'un problème du modèle atmosphérique lorsqu'on accroît sa résolution. [Mentionner ou montre les tests globaux guidés, là ou ailleurs, montrant une mauvaise climato mais une faible variabilité]

IV.7 Configuration finalement retenue

[Je pense que ce résumé de la config serait très bien en intro du chapitre suivant. Genre rappel de ce qu'on a fait pour préparer le tuning]

La configuration finalement retenue utilise la physique CMIP6-L79 décrite Chapitre II avec les changements suivant :

- la grille définie Section IV.3.3.1
- un guidage en vent, température et humidité avec les constantes de guidage résumées Table IV.4
- l'albedo diagnostiqué au SIRTa imposé sur la maille du SIRTa (Section IV.3.3.5)
- le facteur d'aridité diagnostiqué au SIRTa imposé dans un rayon de 100 km autour du SIRTa (Section IV.6)
- l'humidité du sol imposée saturée à partir de 1,3 m de profondeur, une discrétisation verticale du sol à 22 couches et l'indice de surface foliaires imposé (Section IV.6.2.1)

IV.8 Conclusion

Ce chapitre a permis la mise en place d'une configuration de référence de LMDZ zoomée et guidée comparable aux observations du SIRTa, sauf pour les précipitations qui doivent être moyennées régionalement pour être comparables aux observations de précipitations régionales Coméphore.

La stratégie adoptée pour le traitement des incertitudes a consisté en : (1) une réduction et/ou une quantification de l'incertitude liée à la localité du site ; (2) la quantification, et si possible la réduction, de la variabilité interne des simulations. Nous avons en particulier mis en évidence la nécessité d'imposer l'albédo du site dans la maille du SIRTa pour évaluer les flux radiatifs solaires montant et descendant à la surface. Un certain nombre de détails pourraient encore être peaufinés pour permettre une meilleure comparaison : prise en compte des aérosols et/ou étude de sensibilité à des prises en compte plus fines des aérosols ; étude de sensibilité aux constantes de temps de guidage ; étude et éventuellement réglage de la profondeur du sol à partir de laquelle l'humidité du sol doit être imposée à saturation pour garantir une humidité du sol à 1 mètre proche des observations ; meilleur diagnostic du facteur d'évaporation en hiver.

Au bout du compte, ce travail a rendu possible une comparaison robuste du modèle aux observations pendant les mois de mai et de juin, en particulier en ce qui concerne les précipitations régionales et l'effet radiatif solaire des nuages, utilisés dans le chapitre suivant comme métriques de tuning. Le travail exposé dans ce chapitre souligne la complexité de l'utilisation de la synergie d'observations des super-site pour l'évaluation des modèles de climat, et propose une stratégie pour contourner cette complexité.

Une grande part du travail de ce chapitre a porté sur l'étude de la variabilité interne de la configuration qui, malgré le guidage, est particulièrement forte à partir du premier événement

de précipitations convective de l'année 2018 (avril ou mai suivant la physique utilisée). Cette étude a permis de démontrer que ce sont les précipitations convectives, à cause de la surestimation de leur variabilité spatiale, qui sont responsables de cette surestimation au premier ordre. C'est la rigueur portée sur la quantification des incertitudes simulations-observations, nécessaire au tuning avec la méthode d'HMIR, qui a permis de la détecter. Nous montrons que les précipitations pourront néanmoins être comparées et utilisées comme métrique de tuning, mais en moyenne sur une région autour SIRTA et non plus sur la maille du SIRTA. En effet, et malgré nos efforts, ces dernières ne sont pas comparables aux observations sur une seule maille du fait de leur trop grande variabilité interne. Nous avons pu montrer que la rétroaction de l'hydrologie du sol sur les précipitations n'est pas la cause de la surestimation de leur variabilité spatiale, donc de leur variabilité interne, bien qu'elle soit y contribue légèrement. La cause de cette surestimation reste à investiguer, et une première piste pourrait être d'étudier leur sensibilité à la résolution spatiale du modèle.

Pour s'affranchir de la rétro-action de la variabilité interne des précipitations sur la variabilité interne des variables atmosphériques sensibles au couplage, nous avons diagnostiqué puis imposé le facteur d'évaporation du SIRTA dans un rayon de 100 km autour du site du SIRTA. Cela a permis (1) d'identifier avec certitude les variables sensibles à ce couplage (2) de séparer les biais venant du modèle du sol de ceux venant du modèle atmosphérique. Nous avons ainsi pu mettre en évidence que la sous-estimation de l'effet radiatif des nuages dans le solaire, et dans une moindre mesure dans l'infrarouge, est présent à partir d'avril 2018 peut importe le guidage et même avec un facteur d'évaporation imposé. Diminuer ce biais est donc prioritaire avant de pouvoir poursuivre les investigations ; et nous verrons dans l'expérience de tuning exposée au chapitre suivant, à quel point ce biais peut être diminué grâce à un meilleur choix de paramètres libres. Cela nous amène à penser une stratégie de tuning en deux étape au SIRTA : tuning des paramètres libres du modèle atmosphérique avec une surface imposée ; puis tuning conjoint des paramètres libres de LMDZ et d'ORCHIDEE visant le couplage à la surface.

Faute de temps, cette deuxième étape n'a pu être explorée pendant cette thèse. Finalement, en montrant que l'effet radiatif des nuages dans le solaire et que les précipitations ne sont que très peu sensibles au facteur d'évaporation aux mois de mai et juin, il est possible que le tuning des paramètres libres de LMDZ au SIRTA puisse être fait, sur ces métriques, dans une configuration sans le facteur d'évaporation imposé. Cependant, si l'on souhaite utiliser d'autre variables et/ou d'autres périodes comme métriques de tuning, et en particulier si l'on souhaite tuner les flux radiatifs infrarouge, imposer le facteur d'évaporation peut s'avérer nécessaire.

Chapitre V

Tuning visant les observations du site instrumenté du SIRTa : première expérience de tuning

Remarque lecture : Premier jet, quand même pas mal abouti même si les remarques de Fred n'ont pas encore été intégrées. Il manque une partie de résultat et la conclusion de ce chapitre

Ce chapitre expose la première expérience de tuning réalisée en faisant varier les paramètres libres de LMDZ et en visant les observations du SIRTa pour le rayonnement solaire à la surface, et le produit radar Coméphore pour les précipitations régionales autour du SIRTa. Cette expérience vise à améliorer la simulation du couplage à la surface, en ciblant le biais d'effet radiatif solaire des nuages à la surface, identifié comme un biais de première ordre dans le chapitre précédent, tout en garantissant une simulation correcte des précipitations. Cela est fait avec une configuration de modèle où l'hydrologie du sol est découplé de l'atmosphère, ce qui conduit à éviter de compenser les erreurs du schéma d'hydrologie de surface par l'atmosphère lors de l'ajustement du modèle. Cela a également pour but de distinguer l'origine de ce biais, en quantifiant la part du biais expliqué par un mauvais choix de valeurs de paramètre libre. Ce chapitre est dans la continuité du chapitre précédent : l'expérience mise en place ici est faite avec la configuration développée dans la Chapitre IV et en tenant compte des résultats du Chapitre IV.

Le protocole expérimentale de l'expérience de tuning est décrit Section V.1, et ses résultats sont présentés et analysés Section V.2. La place de ces métriques dans la chaîne de tuning actuelle est étudié Section V.3 ; et les conclusions de ce travail sont présentées Section V.4.

V.1 Protocole expérimental

Comme exposé Chapitre III, concevoir une expérience de tuning avec la méthode d'HMIR consiste à :

- Choisir les paramètres libres que l'on souhaite ajuster, ainsi que leur plage d'exploration (Section V.1.1)
- Choisir les métriques et leur tolérance à l'erreur (Section V.1.3)

L'expérience de ce chapitre a également l'objectif d'être une étape intermédiaire de tuning de LMDZ, c'est-à-dire une étape entre les métriques portant sur les cas uni-colonne (dites aussi métriques 1D, décrites Section III.4.1) et les métriques portant sur le modèle atmosphérique global (Section III.4.2). C'est pourquoi un preconditionnement de l'espace des paramètres libres est réalisé avec les métriques 1D, dont les résultats sont montrés Section V.1.2.

V.1.1 Choix des paramètres libres de l'expérience

24 paramètres libres ont été choisis pour être ajustés dans cette expérience de tuning. Ils sont résumés Table V.1. Lorsque la plage à échantillonner court sur plusieurs ordres de grandeurs, l'échantillonnage est fait de manière logarithmique, ce que signifie qu'une transformation logarithmique de l'intervalle est effectué avant de faire l'échantillonnage de l'espace des paramètres par hypercube latin. Si ce n'est pas le cas, on dit que l'échantillonnage est fait de manière linéaire. La valeur de référence correspond à la valeur du paramètre dans la configuration CMIP6-L79. Les paramètres choisis regroupent les paramètres classiquement variés du modèle du thermiques (décrit Section II.3.2) et de précipitation liquide (**CLC** et **EVAP**, décrit Section II.3.3) pour l'ajustement de LMDZ (Hourdin et al., 2021), auquel on a ajouté les paramètres réglant la largeur de la distribution sous maille de vapeur d'eau sans présence de thermique (**RQSDP**, **RQSP0**, **RQSTOP**, Section II.3.3). Un certains nombre de paramètres de la paramétrisation de la convection profonde, décrite Section II.3.4, ont été ajoutés. Ils agissent sur les paramétrisations de ses précipitations (**OMEPMX**, **ELCV**, **TLCV** et **SIGDZ** pour les descentes précipitantes), de son déclenchement (**STRIG**, **WBSRF**), de son lien avec les processus sous-nuageux (**ALPBLK** pour les thermiques et **ALPWKK** pour les poches froides) et la paramétrisation des poches froides (**WDENSO**). Enfin, le paramètre **FALLV** portant sur les précipitations de glace, et **REI** portant sur la microphysique des nuages, sont ajoutés. Cette liste de 24 paramètres est proche de celle effectuée dans l'expérience de tuning de Hourdin (2024).

Les paramètres ont été choisis parmi les paramètres classiquement variés pour tuner le modèle atmosphérique, de manière à représenter au mieux la diversité des paramétrisations et leur fonctionnement. Un focus est néanmoins fait sur la paramétrisation de la convection profonde et de ses précipitations, pour explorer au mieux ses différents comportements et leur impacts sur les précipitations convectives, dont les problématiques ont été soulevées Section IV.5. Enfin, le nombre de paramètres libre de l'expérience est limité. Tout d'abord parce que les étapes de construction des émulateurs et de prédiction de **htexplo** deviennent trop longues par rapport à nos moyens de calcul à partir d'une trentaine de paramètres libre. Ensuite, parce que le coût de calcul de l'expérience est proportionnel à 10 fois le nombre de paramètres libre. Et enfin, parce qu'un trop grand nombre de paramètre rend difficile l'exploration des NROY.

V.1.2 Préconditionnement avec les métriques uni-colonne

La première étape de l'expérience consiste à contraindre le NROY avec des métriques sur des cas 1D comparés à des LES de référence. Nous suivons la méthode de Hourdin et al. (2021) en utilisant les mêmes métriques sur les cas 1D (Table V.2), mais avec des paramètres libres en plus. *[Je sais pas à quel point c'est un truc que j'aurais expliqué avant]*. Cette étape est une étape de preconditionnement du NROY pour des ajouts de métriques aux vagues ultérieurs. Elle permet de réduire significativement le NROY avec un coût de calcul très petit. Cela permet d'avoir moins de simulations à faire pour explorer et réduire le NROY avec des configurations du modèle plus coûteuse. De plus, elle permet d'exclure les combinaison de paramètres libres conduisant à la mauvaise simulation de ces cas, où seul la physique de la colonne est active. Cela permet de contraindre les paramétrisation de ces cas de manière isolé et ainsi d'éviter une partie des compensations d'erreurs. On s'assure ainsi que les paramétrisations des couche limites de convection peu profonde (sèche pour le cas IHOP, cumulus continental pour ARMCU, cumulus océanique pour RICO et stratocumulus pour SANDU) sont bien réglé de manière isolé avant de passer à des configuration où les paramétrisations vont inter-agir entre elle et avec d'autres composantes du modèle comme la dynamique ou d'autres modèles comme ORCHIDEE.

40 vagues de tuning 1D sont effectuées en faisant varier les paramètre de Table V.1 avec les métriques de la Table ???. A l'issue de ces 40 première vagues, quelques paramètres (**A1**, **B1**,

Paramètre libre	Valeur minimale	Valeur maximale	Valeur de référence	Type d'échantillonnage
A1	0.5 (0.55)	1.2	0.66666667	linéaire
A2	1.5e-3	4.e-3 (3e-3)	2.e-3	linéaire
B1	0.(0.05)	1. (0.4)	0.4 (1.1 ?)	linéaire
BG1	0.4	2. (1.1)	1.1	linéaire
BG2	0.03	0.2	0.09	linéaire
CLC	1e-4	1e-3	6.5e-4	linéaire
DZTH	0.05 (0.135)	0.2	0.07	linéaire
ELCV	0.0001	0.002	0.0003	linéaire
EVAP	5e-5	5e-4	1e-4	logarithmique
FALLV	0.3	2.	0.8	linéaire
OMEPMX	0.0001	0.1	0.001	logarithmique
REI	0.5	1.3	1.	linéaire
RQSDP	5000	23000 (22000)	10000	linéaire
RQSP0	30000	60000	45000	linéaire
RQSTOP	0.05	0.6	0.4	linéaire
TLCV	-65	-35	-55	linéaire
WDENSO	1.e-11	1.e-8	1.e-9	logarithmique
ALPBLK	0.2	0.8	0.5	linéaire
ALPWKK	0.01	0.5	0.25	linéaire
SIGDZ	0.001	0.02	0.003	linéaire
WBSRF	0.05	2.	0.5	linéaire
WBTOP	1.5	6	2.8	linéaire
STRIG	4.0e6	3.6e7	1.2e7	linéaire
CQ	0. (0.007)	0.02 (0.014)	0.012	linéaire

TABLE V.1 – Paramètres libres ajustés pendant l'expérience de tuning, avec leur plage d'exploration initial et leur valeur dans la configuration CMIP6-L79. Les valeurs entre parenthèses sont les valeurs exploré après avoir réduit le NROY avec les métriques 1D (voire Section V.1.2). La signification de chaque paramètres libres est expliquée Section II.3 et est résumé Annexe A.

BG1, **DZTH**, **REI**, **RQSDP** et **CQ**) ont des plages de valeurs pour lesquelles l'implausibilité est toujours au-dessus du cutoff, peut importe la valeur des autres paramètres. Ces plages sont alors retiré de l'exploration afin d'éviter le coût d'échantillonnage de cette portion de l'espace qui ne contient que des vecteurs de paramètres rejeté à l'issu de cette étape de préconditionnement 1D.

Ensuite, 40 vagues de tuning avec les paramètres restreints de la Table V.1 et les métriques de la Table ?? sont effectuées. Deux ordres de grandeur sont gagnés sur la fraction du NROY par rapport au NROY du initial en restreignant les plages des paramètres explorés (Figure V.1). On note que la convergence est atteinte à l'issu de ces 40 vagues, au sens où le NROY n'est plus significativement réduit dans les dernières vagues. De plus, la convergence stricte est atteinte pour les métriques des cas IHOP et ARMCU. Les incertitudes des émulateurs des métriques de RICO et SANDU sont de l'ordre de leur tolérance à l'erreur. Les 4 premières vagues de tuning sont faites avec un seuil de rejet sur l'implausibilité (appelé cutoff) de 3, les 3 suivantes avec un cutoff de 2,5, puis toutes les vagues suivantes avec un cutoff de 2, afin d'accélérer la convergence de l'expérience.

La majorité des paramètres libres associé à la paramétrisation des thermiques sont contraints par les métriques choisies, de manière comparable aux résultats de (Hourdin et al., 2021). Le paramètre **RQSDP**, agissant sur la largeur de la distribution de vapeur d'eau lorsque les

Cas	IHOP	ARMCU	RICO	SANDU	SANDU	SANDU
Sous-cas	REF	REF	REF	REF	SLOW	FAST
Période (h)	7-9	7-9	19-25	50-60	50-60	50-60
$\theta_{400-600m}$	X	X				
$qv_{400-600m}$		X				
α_{max}		X	X			
z_{nuage}		X		X		
$z_{max,nuage}$		X		X	X	X

TABLE V.2 – Métriques 1D

	Type	Valeur
$\theta_{400-600m}$	Absolue	0,5 K
$qv_{400-600m}$	Absolue	$5.10^{-4} \text{ kg.kg}^{-1}$
α_{max}	Relative à la fraction nuageuse	$0,25.\alpha$
z_{nuage}	Relative à l'altitude	$0,12.z$
$z_{max,nuage}$	Relative à l'altitude	$0,12.z$

TABLE V.3 – Tolérances à l'erreur associées aux métriques 1D

thermiques ne sont pas actives, est également contraint par cette expérience.

[est-ce que je relance l'exp de tuning papier Hourdin 9 paramètres en L79 pour avoir le scatter plot des métriques et la bonne matrice d'implausibilité et tout ? Moi je crois que non]

[je commente la figure de l'implausibilité parce qu'elle met du temps à s'afficher]

[est-ce que je mets des figures des cas genre profil nuageux des cas ? le scatter plot des métriques ? Sachant que les figures montrées dans hourdin2021 sont : profils de theta, qv et rneb pour les 6CAS (et aussi des profils de "variables internes" mais ça c'est trop hors sujets)]

V.1.3 Choix des métriques au SIRTa et de leurs tolérances à l'erreur

A l'issu de cette 40ème vague de tuning, 240 vecteurs de paramètres libres appartenant au NROY de la vague 40 sont sélectionnés. 240 simulations avec la configuration du SIRTa sont lancées sur l'année 2018, ce qui constitue l'ensemble de simulation aux paramètres perturbés de la vague 41 appelé PPE41 (PPE pour Perturbed Physical Ensemble) sous la contrainte que ces vecteurs de paramètres conduisent à la bonne simulation des métriques 1D.

Pour choisir les métriques de tuning, on identifie les variables et les périodes où on a (1) des biais que l'on souhaiterait corriger (2) une sensibilité de ce biais aux paramètres de tuning choisies dans cette expérience (3) une comparaison fiable et/ou quantifiable aux observations. La sensibilité du biais aux paramètres libres se caractérise concrètement par une étendu du PPE non nulle et supérieure à l'étendu du PISE.

Le biais le plus marquant et robuste du modèle est son manque quasi-systématique d'effets radiatifs des nuages, qui sont sous-estimés, en LW et en SW, sur une bonne partie de l'année 2018 (Figure V.2). Ce biais peut-être du à un manque de simulation de fraction de nuageuse et de contenu en eau par les paramétrisations physiques de LMDZ, et/ou à des défaut de paramétrisations du transfert radiatif, qui fait des hypothèses d'overlap max-random par exemple donc sous-estime systématiquement la couverture nuageuse à fraction nuageuses égale. Au vu de la valeur du biais, on peut penser que les deux mécanismes sont à l'oeuvre. Puisque le modèle de transfert radiatif utilisé est rrtm, il n'est pas possible dans cette expérience d'évaluer la part de ce biais attribuable aux hypothèses prises en compte des nuages par le modèle de transfert radiatif. Le PPE montre comment lmdz est capable de simuler plus ou moins la fraction nuageuse.

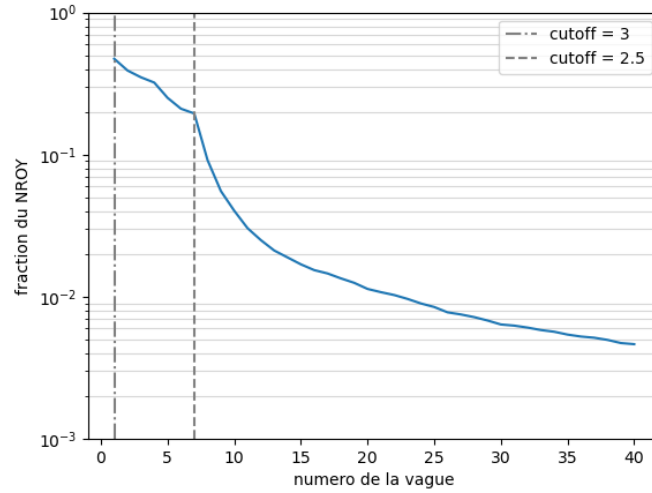


FIGURE V.1 – Evolution de la fraction du NROY restante en fonction du numéro de la vague sur les 40 vagues du préconditionnement 1D. En bleu pour l'expérience avant restriction de l'espace de paramètres initial, et en rouge après restriction.

geuse et/ou de contenu en eau en fonction des valeurs des paramètres libres choisies, via l'effet radiatif des nuages. L'étendu du PPE est bien supérieure à celui du PISE sur ces variables, ce qui signifie que la sélection de meilleurs vecteurs de paramètres pourrait permettre de diminuer ce biais [dont on voit (figure) sa repercussion sur les flux turbulents via le bilan d'énergie à la surface]. [vérifier le discours sur le LW quand j'écrirais la mise en place de la configuration]. L'effets radiatifs des nuages dans le LW n'est finalement pas retenu comme métrique de tuning du fait de sa potentielle trop forte dépendance à la localité du site. Le rayonnement LW descendant à la surface en ciel clair, utilisé pour calculer l'effet radiatif des nuages, est très dépendant des profils de température et d'humidité proche de la surface [ou sur toute la couche limite ?], eux-mêmes très dépendant de la localité du site [j'imagine que j'aurais montré des trucs comme ça dans la mise en place de la configuration]. Cette dépendance est difficilement quantifiable et nous avons jugés préférable de ne "pas prendre ce risque". Les métriques retenues [concernant le rayonnement] sont la moyenne du $CRESW_{sfc}$ sur deux périodes temporelles. La première est la période du 22 mai au 12 juin 2018 incluse, pendant laquelle la convection profonde est active. La deuxième période est celle du 13 au 30 juin 2018 incluse, pendant laquelle des cumulus de beaux temps se développent. L'évolution de l'effet radiatif pendant ces deux périodes est montré Figure V.3. Ces périodes sont idéales pour le tuning car ce sont des périodes longues pendant lesquelles un processus majoritaire est actif (convection profonde dans le premier cas et convection peu profonde dans le deuxième cas).

Les biais sur les précipitations sont également des biais que nous souhaitons réduire et étudier dans cette expérience de tuning. Compte tenu de leur variabilité interne, ce sont les précipitations en moyennes sur 25 mailles autour de la maille du SIRTa qui sont étudiées, avec comme référence le produit de précipitation radar-pluviomètre comephore [Mettre la référence à la section sur la variabilité des précipitations]. La Figure V.2 (c) montre l'évolution de ces précipitations sur l'année 2018. Les périodes où les précipitations sont sous-estimées (de janvier à avril puis d'août à décembre) sont des périodes où les précipitations sont issues des paramétrisations de précipitations dites "de grande échelle" et particulièrement sensible à l'advection d'humidité donc au guidage [à vérifier après avoir écrit la partie correspondante]. L'étendue de leur PPE est assez faible, ce qui signifie qu'elles ont peu de sensibilité aux paramètres libres

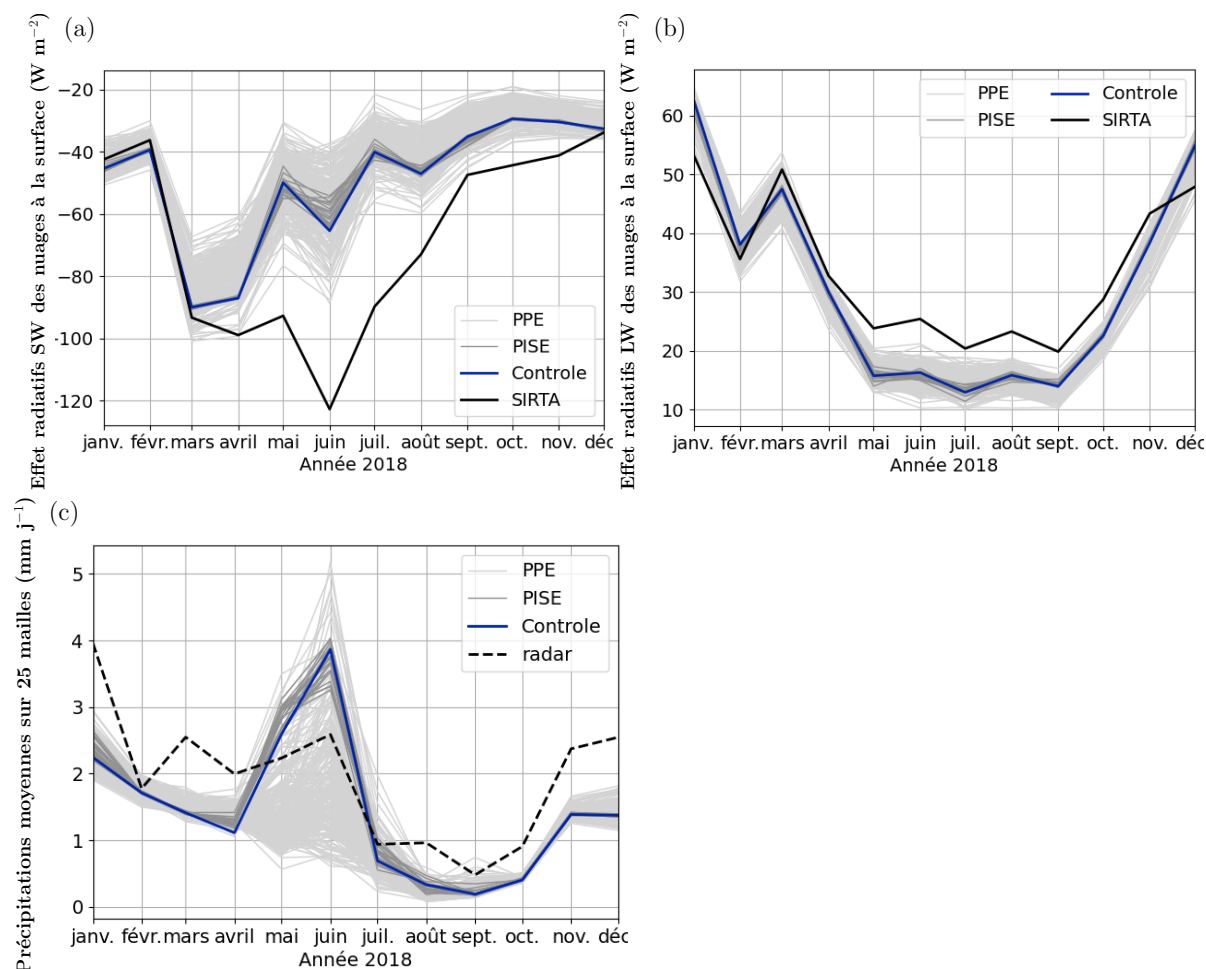


FIGURE V.2 – Evolution en moyenne mensuelle sur l'année 2018 de (a) l'effet radiatifs SW des nuages à la surface (b) effet radiatif LW des nuages à la surface (c) des précipitations moyennes sur 25 mailles autour de la maille du SIRTa. En gris claire l'ensemble aux paramètres perturbées de la vague 41 de l'expérience, en gris foncé l'ensemble aux conditions initiales perturbé, et bleu la simulation de contrôle de l'expérience (configuration SIRTa), en noir les observations du SIRTa et en noir pointillé les observations radar pour les précipitations

variés dans cette expériences de tuning. Les précipitations des mois de mai juin et juillet sont en revanche très sensibles aux paramètres libres variés dans cette expériences, avec une étendue du PPE largement supérieur au PISE. Ces précipitations sont en majorité issue de la paramétrisation de convection profonde. Cette forte sensibilité s'explique en partie par le fait que l'on ait spécifiquement choisi de varier des paramètres libres des paramétrisations de convection profonde et de leur précipitation dans cette expérience. La période orageuse du 22 mai au 12 juin est une période idéale pour aller contraindre ces précipitations (Figure V.3). Il est de plus intéressant de contraindre sur la même période l'effet radiatif des nuage et les précipitations, afin de sélectionner des vecteurs de paramètres libres conduisant à la simulation de bon nuage (vu via leur effet radiatif) qui font les bonnes précipitations.

Les tolérances à l'erreurs sont choisies de manière (1) à être supérieur à la somme [*quadratique*] des incertitudes de l'expérience (2) à assurer l'existence d'un NROY non vide. Les incertitudes quantifiées pour cette expérience sont les incertitudes d'observation et de variabilité interne du modèle. L'incertitude d'observation du rayonnement SW est de $5 W.m^{-2}$ d'après Chiriaco et al. (2018). [*Pour les précipitations j'ai pas fait le travail d'aller chercher l'incerti-*

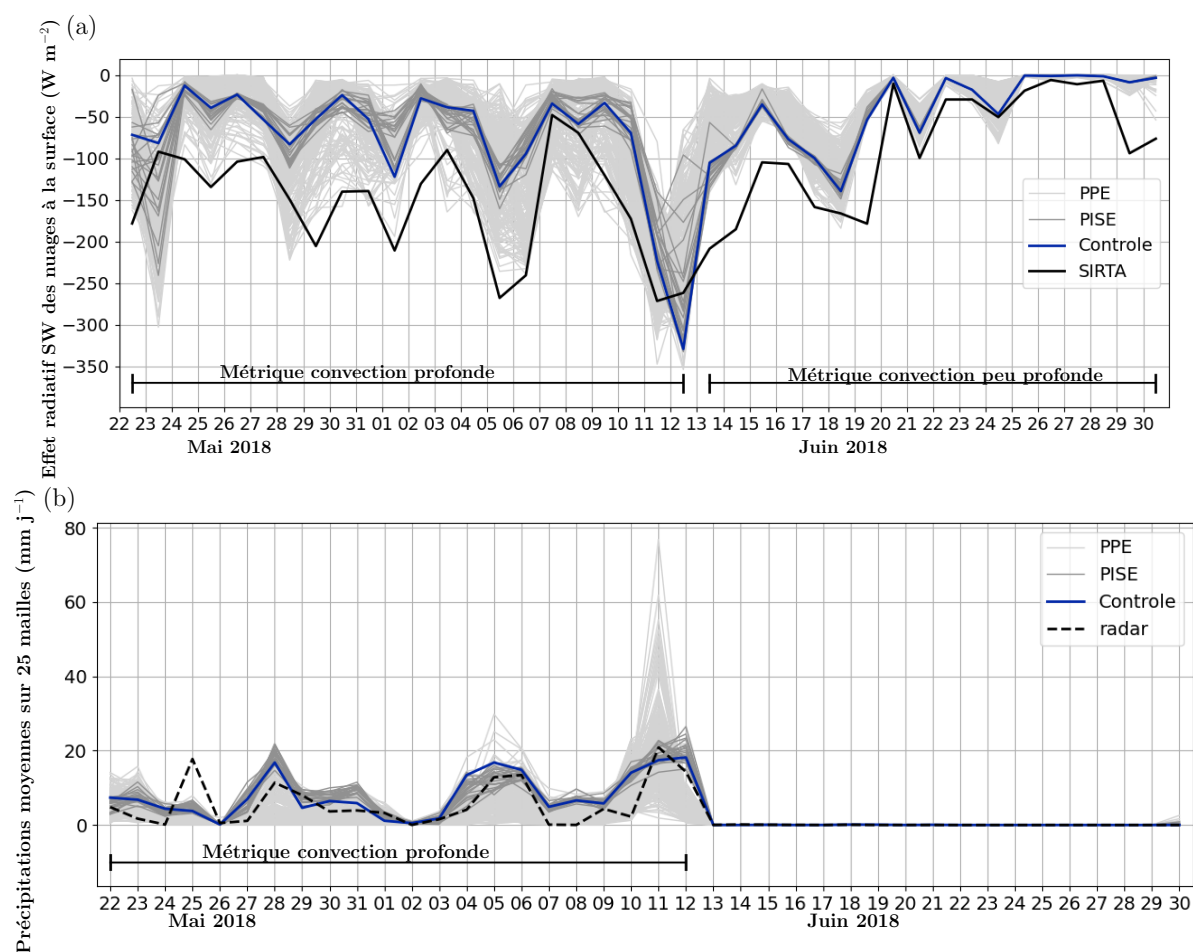


FIGURE V.3 – Evolution en moyenne journalière du 22 au 30 juin 2018 de (a) l'effet radiatif SW des nuages à la surface (b) des précipitations moyennes sur 25 mailles autour de la maille du SIRTa. En gris claire l'ensemble aux paramètres perturbés de la vague 41 de l'expérience, en gris foncé l'ensemble aux conditions initiales perturbé, et bleu la simulation de contrôle de l'expérience (configuration SIRTa), en noir les observations du SIRTa et en noir pointillé les observations radar pour les précipitations

tude du produit comephore. Les références sur le jeux de données sont difficiles à trouver. Il y a une variable ERR dans le produit qui donne "l'incertitude sur le cumul de lame d'eau horaire en 1/10 mm" que je n'ai jamais été regardé : ça pourrait être intéressant de voir ce que ça donne. Pour description des données radars : Tabary et al. (2012)]. L'incertitude liée à la variabilité interne du modèle est quantifiée comme l'écart-type de la valeur de la métrique sur l'ensemble de simulation aux états initiaux perturbés. En faisant ce choix, on fait l'hypothèse que la variabilité interne du modèle n'évolue pas avec les valeurs des paramètres libres [c'est peut-être un test à faire à posteriori ?]. Les valeurs des incertitudes pour les trois métriques sont résumées Table V.4. Pour s'assurer que le NROY ne sera pas vidé avec l'ajout de ces trois métriques, on s'assure qu'il existe au moins une simulation dont le score [mettre l'équation quelque part] est inférieur au cutoff, qui est ici de 2 [c'est ici que j'ai fait une erreur, j'ai cru à l'époque que le cutoff était de 2.5]. En pratique, on augmente progressivement la tolérance à l'erreur des métriques jusqu'à ce que le score d'une simulation soit en dessous du cut-off, en mettant la même tolérance à les erreurs pour les deux métriques sur le $CRESW_{sfc}$. On abouti à des tolérances à l'erreur de 15 $W.m^{-2}$ sur les métriques de $CRESW_{sfc}$ et 0.9 $mm.j^{-1}$ sur la métrique précipitation.

[Discussion avec Fred : ok pour commenter les biais du modèle, ok pour décrire comment

Métrique	Valeur de la référence	Incertitude d'observation	Incertitude de variabilité interne	Tolérance à l'erreur
Moyenne du $CRESW_{sfc}$ du 22 mai au 12 juin 2018	-153,43 W.m ⁻²	5 W.m ⁻²	5,06 W.m ⁻²	15 W.m ⁻²
Moyenne du $CRESW_{sfc}$ du 13 au 30 juin 2018	-85,73 W.m ⁻²	5 W.m ⁻²	0,99 W.m ⁻²	15 W.m ⁻²
Moyenne des précipitations du 22 mai au 12 juin 2018	5,92 mm.j ⁻¹	COUCOU	0,37 mm.j ⁻¹	0,9 mm.j ⁻¹

TABLE V.4 – Incertitudes et tolérances à l'erreur des métriques

on a hoisi les métriques, pas aller trop dans le détail des autres métriques qui aurait pu être choisie parce qu'on ne l'ai pas testé. Ce sera plutôt dans la phase d'évaluation du tuning qu'on pourra aller voir sur d'autres périodes spécifiques thermique et convection profonde si leurs biais s'améliorent aussi ou pas]

On applique l'history matching sur cette 41ème vague de tuning, avec les métriques des cas 1D (Table V.2) et leur tolérances à l'erreur (Table V.3) plus les métriques du SIRTa et leur tolérances à l'erreur (Table V.4).

V.2 Résultats

A l'issue de l'history matching de la vague 41, le NROY très réduit. Sa fraction par rapport à l'espace des paramètre initial est de $8,5 \cdot 10^{-6}$. 240 vecteurs de paramètres sont sélectionnés à l'issu de l'history matching et les 240 simulations correspondantes sont tournés avec la configuration du SIRTa, ce qui constitue le PPE de la vague 42. Les métriques de la Table V.4 sont calculé sur ce nouvel ensemble.

$$S(\boldsymbol{\lambda}) = \max \left\{ \frac{|r_1 - f_1(\boldsymbol{\lambda})|}{T_1}; \dots; \frac{|r_p - f_p(\boldsymbol{\lambda})|}{T_p} \right\} \quad (\text{V.1})$$

La Figure V.4 montre le score $S(\boldsymbol{\lambda})$ classiquement utilisé pour trier les simulations (Equation V.1) et un score calculé pour chacune des métriques i comme $\frac{r_i - f_i(\boldsymbol{\lambda})}{T_i}$. Avec r_i la valeur de la référence pour la i^{me} métrique, $f_i(\boldsymbol{\lambda})$ la valeur de la métrique de la simulation ayant pour vecteur paramètre $\boldsymbol{\lambda}$ et T_i la tolérance à l'erreur associé à la métrique i . Les 10 meilleures simulations sont les 10 simulations ayant le plus petit score. L'étendue des scores des simulations de la vague 42 est toujours plus resserré que celle de la vague 41, ce qui illustre la selection faite par l'outil de tuning. L'history mathcing pour la vague 42 n'a pas pu aboutir car le NROY est trop vide et son temps d'exploration est devenu trop grand. En 20h d'execution (temps humain), l'outil n'a trouvé que deux bons vecteurs de paramètres pour 150 millions de tirages, ce qui correspond à une fraction de NROY de l'ordre de $1 \cdot 10^{-8}$. Pour enchaîner un nouvelle vague, il faut tirer 240 bons vecteurs de paramètres, ce qui prendrait plus d'une centaine de jours. Le NROY₄₂ est devenu très vide car aucune simulation n'a un score en-dessous du cutoff dans cette vague, le score de la meilleure simulation étant de 2,11. Les tolérances à l'erreur ont était choisies trop petites. Lors du choix de ces tolérances, une erreur a été commise : elles ont été choisi de manière à ce qu'une simulation au moins ait un score inférieure à 2,5 et non 2. Si l'expérierie était à refaire, le cutoff serait laissé à 3 pour toutes les vagues de tuning car il est plus prudent de garder un cutoff large lors de l'ajout de nouvelles métriques (donc de nouveaux émulateurs). Ce NROY très vide nous informe que le modèle n'est pas capable de simuler les trois métriques choisies avec une erreur inférieure à deux fois leurs tolérances à l'erreur. Les

biais peuvent tout de même être réduits par un meilleur choix de vecteurs de paramètres mais il ne peuvent être entièrement corrigés. Ce résultat est sensible au choix des paramètres variés dans cette expérience et à leur plage d'exploration. Il reste néanmoins probable que des erreurs plus "profondes" dans la modélisation des processus en jeu dans la simulation des métriques sont responsables des biais observés.

Le fait que le $NROY_{42}$ soit si vide n'empêche pas d'étudier les PPE des vagues 41 et 42 pour essayer de comprendre comment les métriques choisies contraignent le NROY et la simulation d'autres variables.

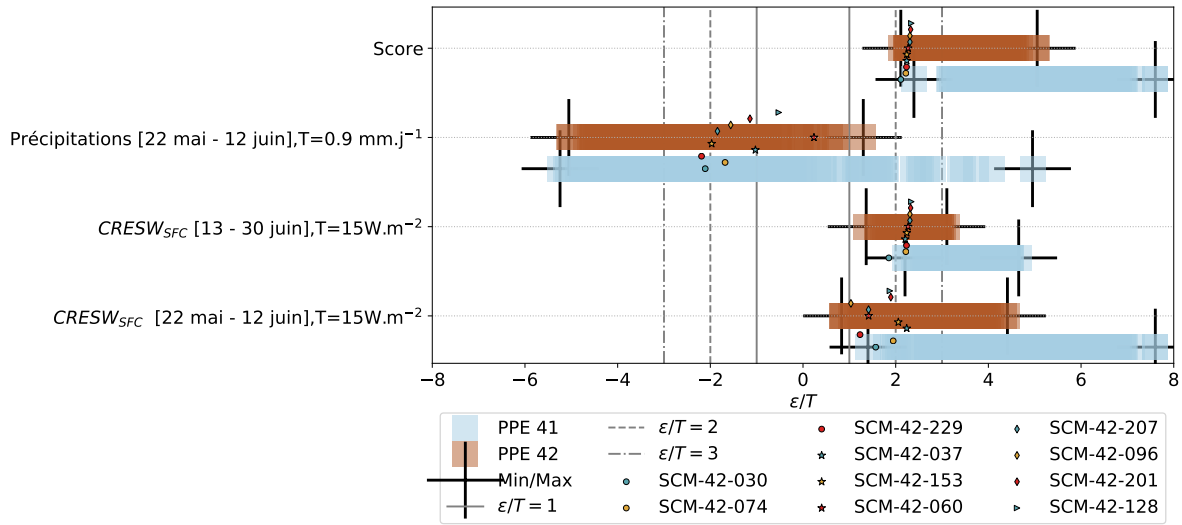


FIGURE V.4 – Score (Equation V.1) et score par métrique pour l'ensemble des simulations du PPE41 en bleu et du PPE42 en marron. Les croix noires marquent les valeurs minimales et maximales des scores des deux PPE. Les valeurs des scores des 10 meilleures simulations sont indiquées avec des symboles.

Les scores des métriques $CRESW_{sfc}$ sont toujours strictement positifs, ce qui correspond à un manque d'effets radiatifs des nuages dans le SW. L'ensemble des scores de la vague 42 est décalé vers 0 par rapport à ceux de la vague 41 sur les deux métriques $CRESW_{sfc}$, sans jamais l'atteindre. Cela signifie que le modèle n'est pas capable de simuler le bon $CRESW_{sfc}$ même en faisant varier les paramètres libres du modèle, parmi ceux sélectionnés pour cette expérience. Le modèle a plus de difficulté à simuler le bon $CRESW_{sfc}$ sur la période du 13 au 30 juin, qui correspond à la période de convection peu profonde, que sur la période du 22 mai au 12 juin, qui correspond à la période de convection peut profonde. Cela peut être du au fait que l'on soit relativement plus exigeant sur cette période, et que les paramètres des paramétrisations de convection peu profonde sont contraint par le préconditionnement 1D mais pas ceux de la convection profonde. On voit Figure V.5 (a) que sur les deux périodes, les simulations de la vague 42 simulent plus d'effets radiatifs des nuages dans le solaire que celles de la vague 41. Cela reste vrai sur toute l'année dans le solaire et aussi l'infrarouge (Figure V.6 (a) et (b)). Cette augmentation d'effet radiatifs de nuage est la conséquence de la simulations de nuages plus chargés en eau et en glace (Figure V.7 (a) et (b)) dans le PPE de la vague 42 par rapport à celui de la vague 41, mais pas significativement plus couvrant (Figure V.7 (c)). Une sélection est tout de même faite sur les couvertures nuageuses, sur les périodes des métriques (Figure V.5 (b)) et sur l'année (Figure V.7 (c)), où les nuages trop peu couvrant d'été sont éliminés du PPE42. Les simulations du PPE42 simulent un peu de moins de couverture nuageuse pour les nuage haut (Figure V.7 (f)) et un peu plus sur les nuages moyens (Figure V.7 (e)).

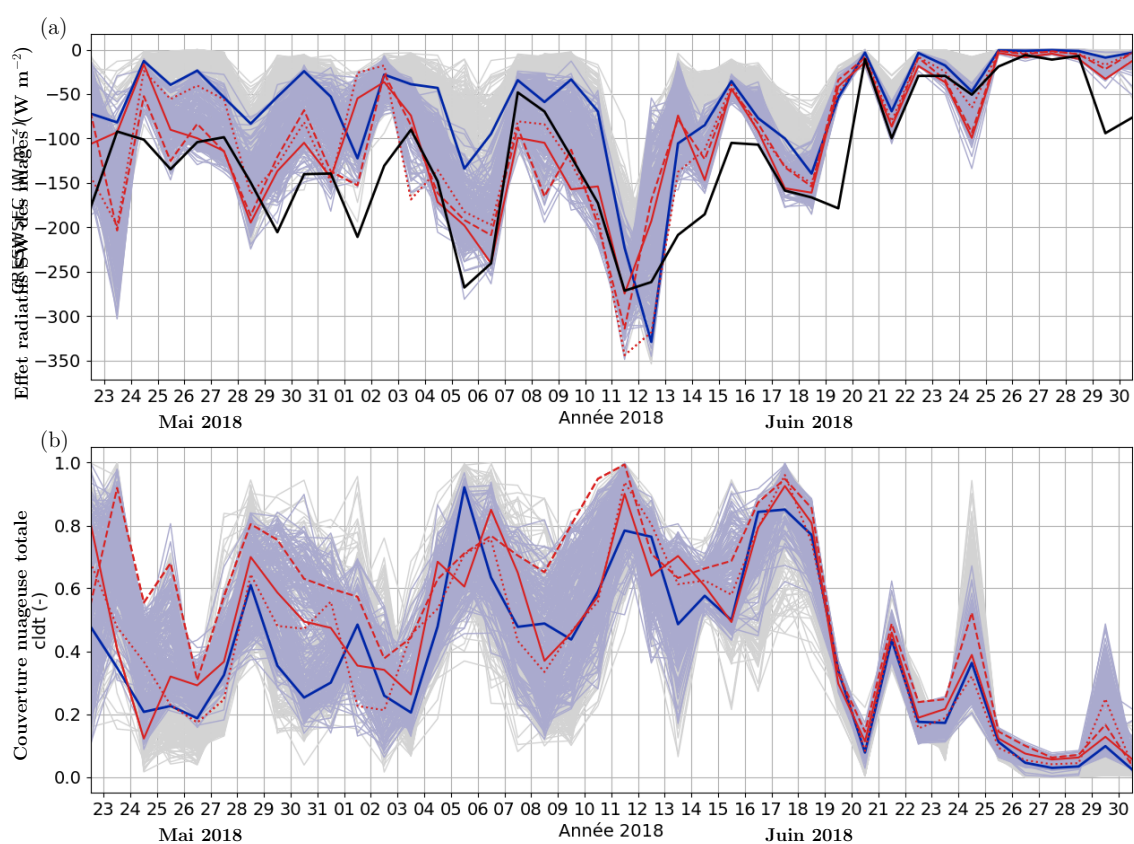


FIGURE V.5 – Evolution en moyenne journalière du 22 mai au 30 juin 2018 (a) de l'effet radiatif SW des nuages à la surface, (b) de la couverture nuageuse intégrée. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa.

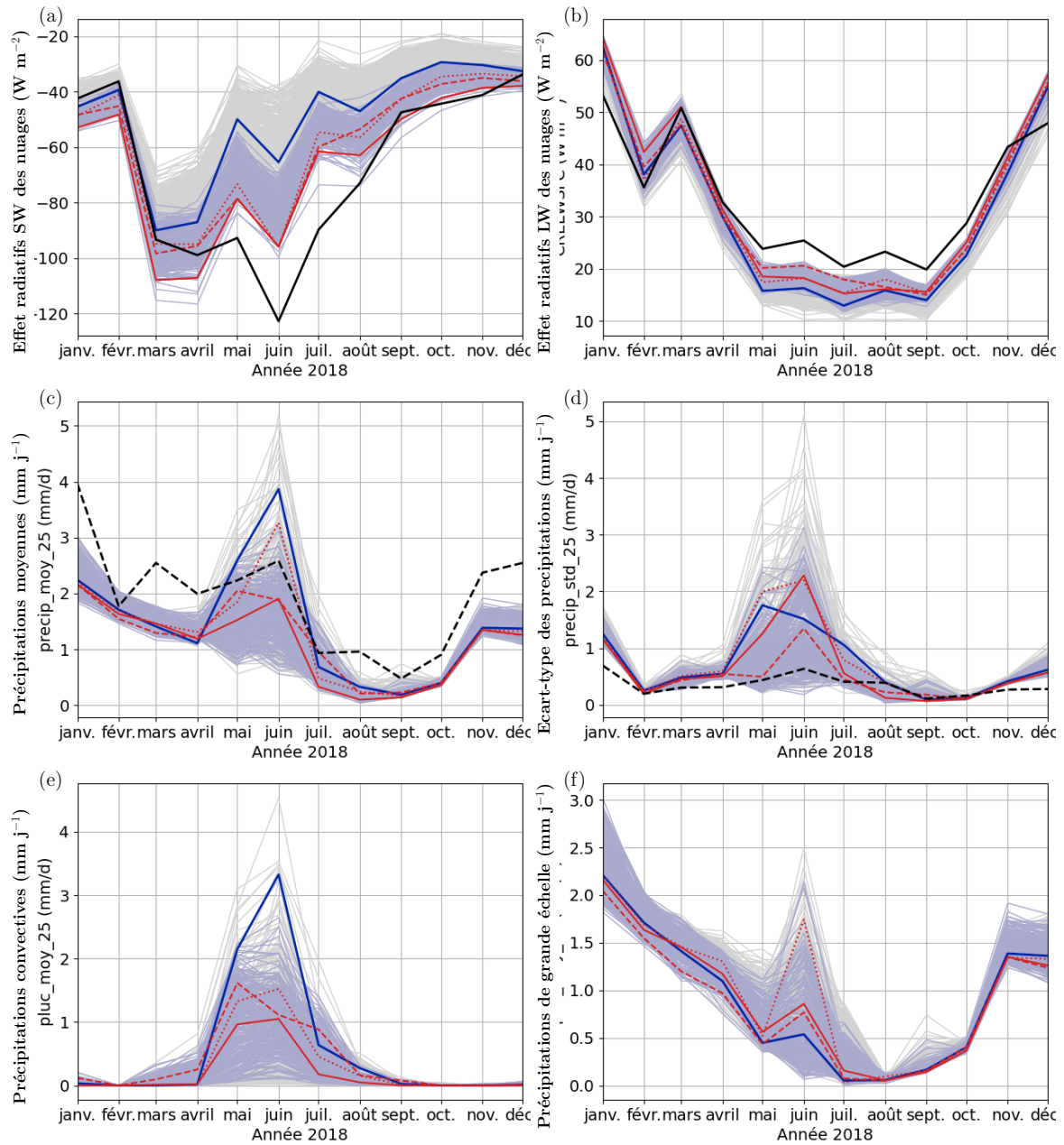


FIGURE V.6 – Evolution en moyenne mensuelle sur l'année 2018 (a) de l'effet radiatif SW des nuages à la surface, (b) de l'effet radiatif LW des nuages à la surface, (c) des précipitations moyennes sur 25 mailles, (d) de l'écart-type spatial des précipitations sur 25 mailles, (e) des précipitations issues de la paramétrisation de convection profonde (moyennes sur 25 mailles), (f) des précipitations issues de la paramétrisation de condensation grande échelle. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTA.

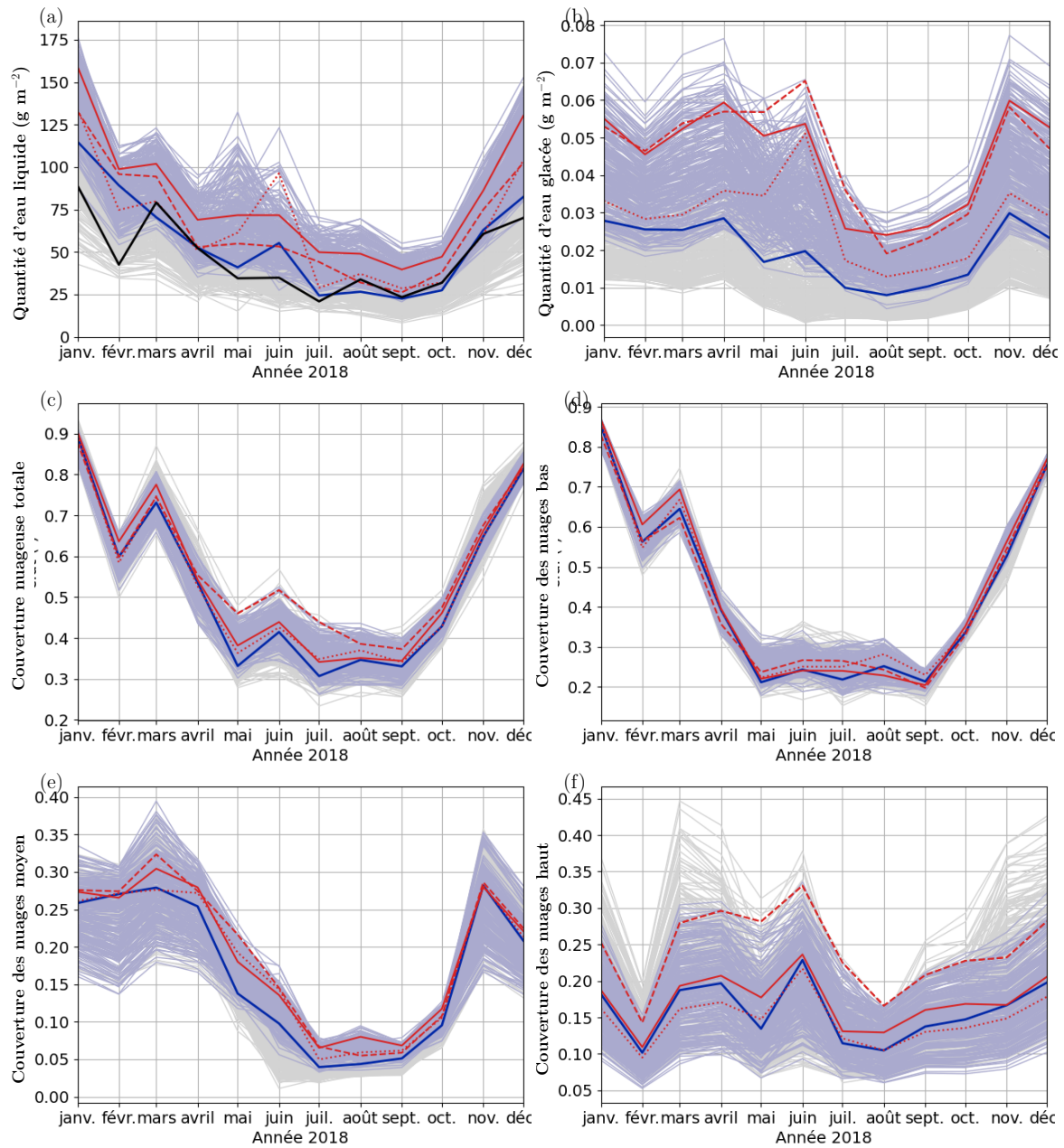


FIGURE V.7 – Evolution en moyenne mensuelle sur l'année 2018 (a) de la quantité d'eau liquide (b) de la quantité d'eau glacée (c) de la couverture nuageuse intégrée totale (d) de la couverture nuageuses des nuages bas (e) de la couverture nuageuse des nuages moyens (f) de la couverture nuageuse des nuages haut. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle.

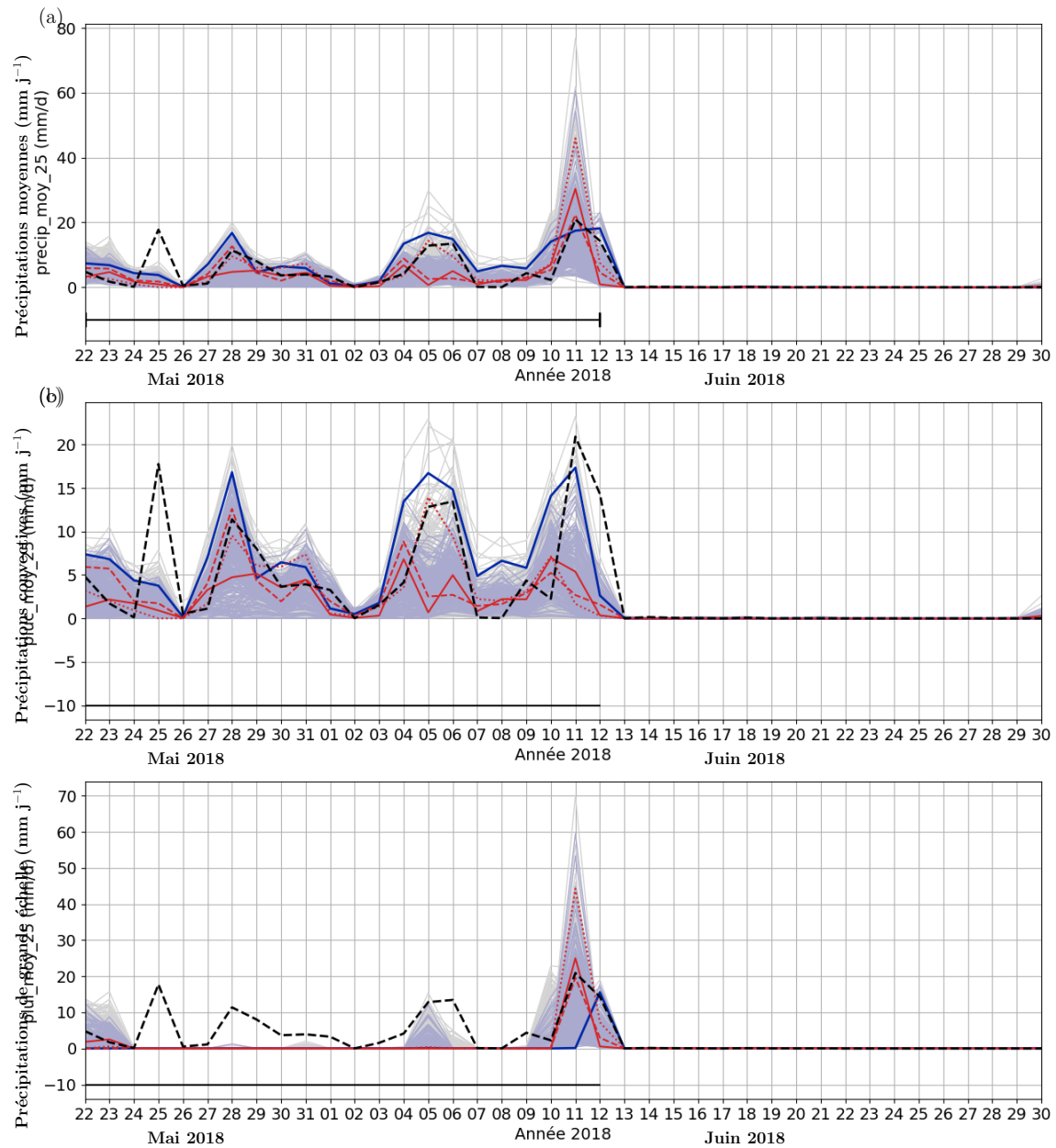


FIGURE V.8 – Evolution en moyenne journalière du 22 mai au 30 juin 2018 (a) des précipitations en moyennes sur 25 mailles, (b) des précipitations issu de la paramétrisation de convection profonde, en moyenne sur 25 mailles, (c) des précipitations issues de la paramétrisation de condensation grande échelle, en moyenne sur 25 mailles. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir pointillées les observations radar.

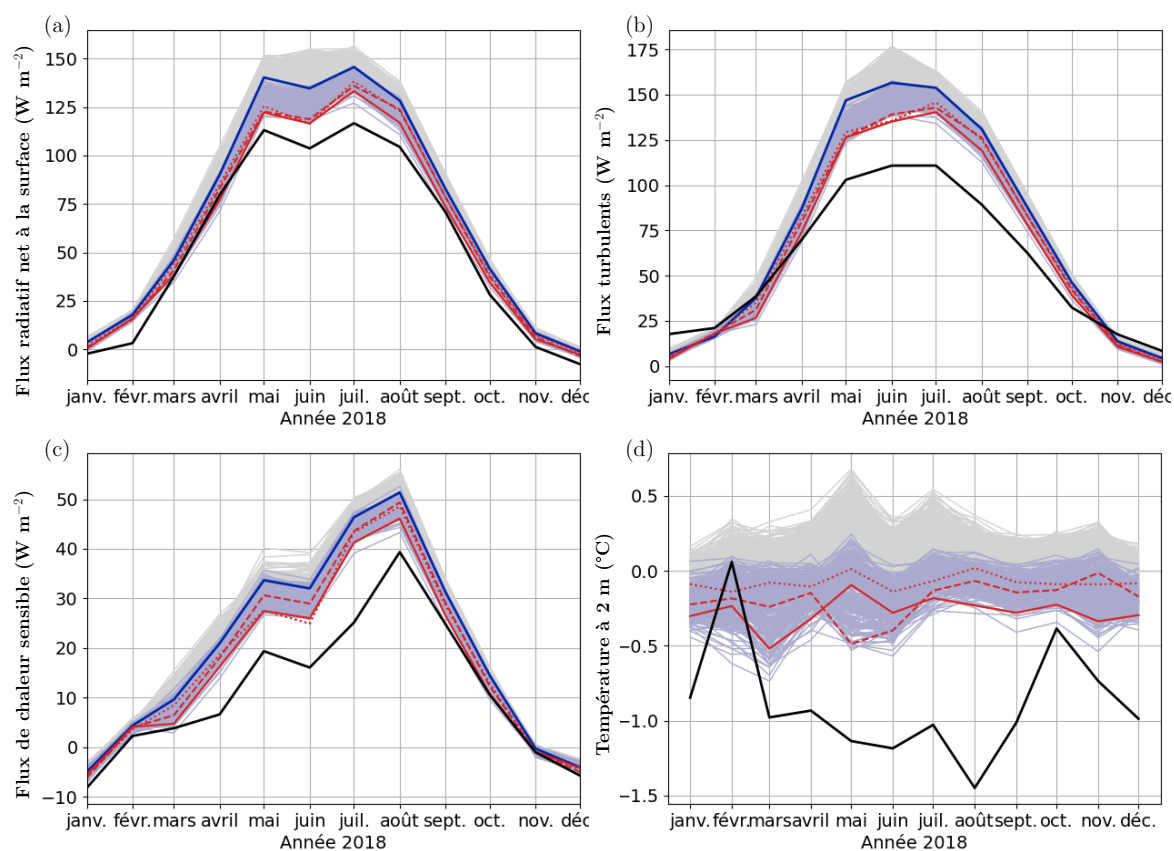


FIGURE V.9 – Evolution en moyenne mensuelle sur l’année 2018 (a) du flux radiatif net à la surface (b) de la somme des flux turbulents de surface (c) du flux de chaleur sensible (d) de la température à 2 mètre. En gris claire le PPE de la vague 41, en bleu claire le PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa

La sélection des vecteurs de paramètres de la vague 42 s’est également faite avec la contrainte de simuler les bonnes précipitations (en moyenne sur 25 mailles) dans la période de convection profonde. L’étendue des score de cette métrique dans la vague 41 va de -5 à +5 (Figure V.4). Le modèle est donc capable de simuler des précipitations très différentes sur cette période, dont des précipitations proche de la référence. Les scores sur les précipitations de la vague 42 vont de -5 à 1,5 : les vecteurs de paramètres conduisant à de trop fortes précipitations ont été exclus du NROY, mais pas ceux conduisant à de trop faibles précipitations. On le voit sur la période convective utilisée comme métrique (Figure V.8 (a)) mais aussi en moyenne mensuelle de mai à août 2018 (Figure V.6 (c)). Cela peut s’expliquer par le fait que de trop forte précipitations conduisent à une diminution de la durée de vie des nuages donc à une diminution de l’effet radiatif des nuages. Ces vecteurs de paramètres ont pu être exclu à cause de la mauvaise simulation des métriques sur l’effet radiatif des nuages. Hors de la période mai-août 2018, les PPE41 et 42 sont aussi étendu sur les précipitations (Figure V.6 (c)), où les précipitations sont très liées à l’advection d’humidité [Je sais plus trop]. Sur la période du 22 au 12 mai, les précipitations sont généralement fait par la paramétrisation de la convection profonde (Figure V.8 (b)), mais dans certains cas, s’ajoute à ces pluies des précipitations venant de la paramétrisation de condensation de grande échelle (Figure V.8 (c)). La journée du 11 juin est un cas où la paramétrisation de condensation de grande échelle produit beaucoup plus de pluies que la paramétrisation de convection profonde ; et en produit beaucoup trop. Cela est caractéristique de ce que l’on appelle des “orages en points de grille”, pendant lesquelles une partie de la convection profonde est résolue explicitement par le modèle (et non plus par sa

paramétrisation), et qui se traduit généralement par des des précipitations très anormalement élevée et "brutales". *[est-ce qu'il y a des papiers qui parlent de ça dans lmdz ?]*. Ce défaut connu du modèle pourrait être exaerbé par la forte résolution du modèle dans notre configuration. La métriques "précipitation" du tuning ne contraint ainsi pas que les précipitations issues de la paramétrisation de la convection profonde comme attendue ; mais aussi les précipitations issues de la grande échelles. L'événement d'orage en point de grille est associés à une variabilité spatiales des pluies très élevés. Hors de ces événements, l'étendue des PPE 41 et 42 sont assez large en ce qui concerne la variabilité spatiale des précipitations : le modèle est donc capable de simuler une variabilité spatiales des précipitations satisfaisantes. Les simulations ayant des variabilité spatiales des précipitations trop élevée sont exclues du PPE42 (Figure V.6 (d)). *[J'ai pas mi les figures qui décompose la variabilité spatiale par type de pluies sur la période et/ou en moyenne mensuelle ; je peut les rajouter si besoin]*.

Les précipitations ne rétroagissent plus sur la modèle via l'humidité du sol, à part sur l'inertie thermiques du sol. En revanche, l'augmentation de l'effets radiatif des nuages des simulations du PPE42 impacte les termes du bilan d'énergie à la surface. Les flux radiatifs solaire descendant et net à la surface sont plus faibles dans le PPE42. Le flux radiatif infrarouge descendant à la surface n'est quand a lui pas impacté : l'augmentation de l'effet radiatif des nuages dans l'IR est contrebalancé par une diminution du flux radiatif descendant en ciel clair. Le flux radiatif LW net à la surface augmente quand à lui de quelques W.m^{-2} dans le PPE42, du à une diminution de la température du sol. Au final, le rayonnement net à la surface diminue jusqu'à une quinzaine de W.m^{-2} en été dans le PPE42 par rapport à la simulation du controle (Figure V.9 (a)). La même diminution est présente sur la somme des flux turbulents (Figure V.9 (b)), dont le partitionnement est fixé dans la configuration *[Mettre une ref à la section qui explique ça]*. La diminution du flux de chaleur sensible est finalement de l'ordre de 5 W.m^{-2} ((Figure V.9 (c)), et celle de la température à 2m de quelques dixièmes de degrés ((Figure V.9 (d)).

V.2.1 Analyse des meilleurs simulations

Les meilleurs simulations nous donne un aperçu de ce vers quoi tendraient les PPE des vagues suivantes si nous avons pu continuer l'expérience. Parmi les 10 meilleurs simulations, on selectionne la simulation ayant le meilleure score sur la métrique $CRESW_{sfc}$ de la période de convection profonde (SCM-42-096), celle ayant le meilleure score $CRESW_{sfc}$ sur la période de convection peu profonde (SCM-42-030) et celle ayant le meilleur score sur les précipitations (en moyenne sur 25 mailles) sur la période de convection profonde (SCM-42-060). La simulation ayant le meilleur score sur la métrique $CRESW_{sfc}$ de la période de convection peu profonde est aussi la meilleure simulation en générale sur les deux PPE. Ces simulations font partie des meilleurs pour simuler le $CRESW_{sfc}$ du mois de juin, ce qui confirme que ces métriques sont plus sélectives que celle sur les précipitations. Elles font toutes un meilleur $CRELW_{sfc}$ que la simulation de controle mais ne font pas forcément partie des meilleurs du PPE42. Les trois simulations ont des contenue en eau liquide et des couverture nuageuses contrastés. On retrouve un effet de compensation de ces deux variables pour augmenter l'effet radiatif des nuages : SCM-42-096 fait des nuages moins chargés en eau que les autres simulations en mai et juin mais plus de couverture nuageuse total, en particulier sur les nuages hauts. Le fait que ces simulations font parties des meilleures pour le $creswsfc$, fait qu'elles font partie des meilleures pour la rayonnement net à la surface et la somme des flux turbulents, mais cela ne se répercute pas sur les flux de chaleur sensible et encore moins sur la température à 2m. Au niveau des précipitations, SCM-42-060 simule en effet de bonnes précipitations sur la période visé, excepté le 25 mai où aucune simulation ne s'approche de la référence, et le 11 juin. Ce jours la, les précipitations de SCM-42-060 sont très majoritairement faite par la paramétrisation de condensation grandes échelles. SCM-42-060 fait partie des simulations du PPE42 qui fait le plus

gros orage en point de grille. Les métriques choisies ne sont donc pas capable de contraindre ce phénomène. SCM-42-060 surestime largement l'écart-type spatial des précipitations, à cause de l'orage en point de grille, mais aussi de manière systématique sur les précipitations convectives. La simulation SCM-42-096 en revanche, arrive à faire des écarts-type spatiaux raisonnable hors orage point de grille.

[
— *best CRESWSFC conv = SCM-42-096 '-' courbe rouge tiretée*
— *best CRESWSFC nonpr = SCM-42-030 = best générale '-' courbe rouge*
— *best precip moy 25 conv = SCM-42-060 ':' courbe rouge pointillé*
]

V.2.2 Evaluation sur d'autres années

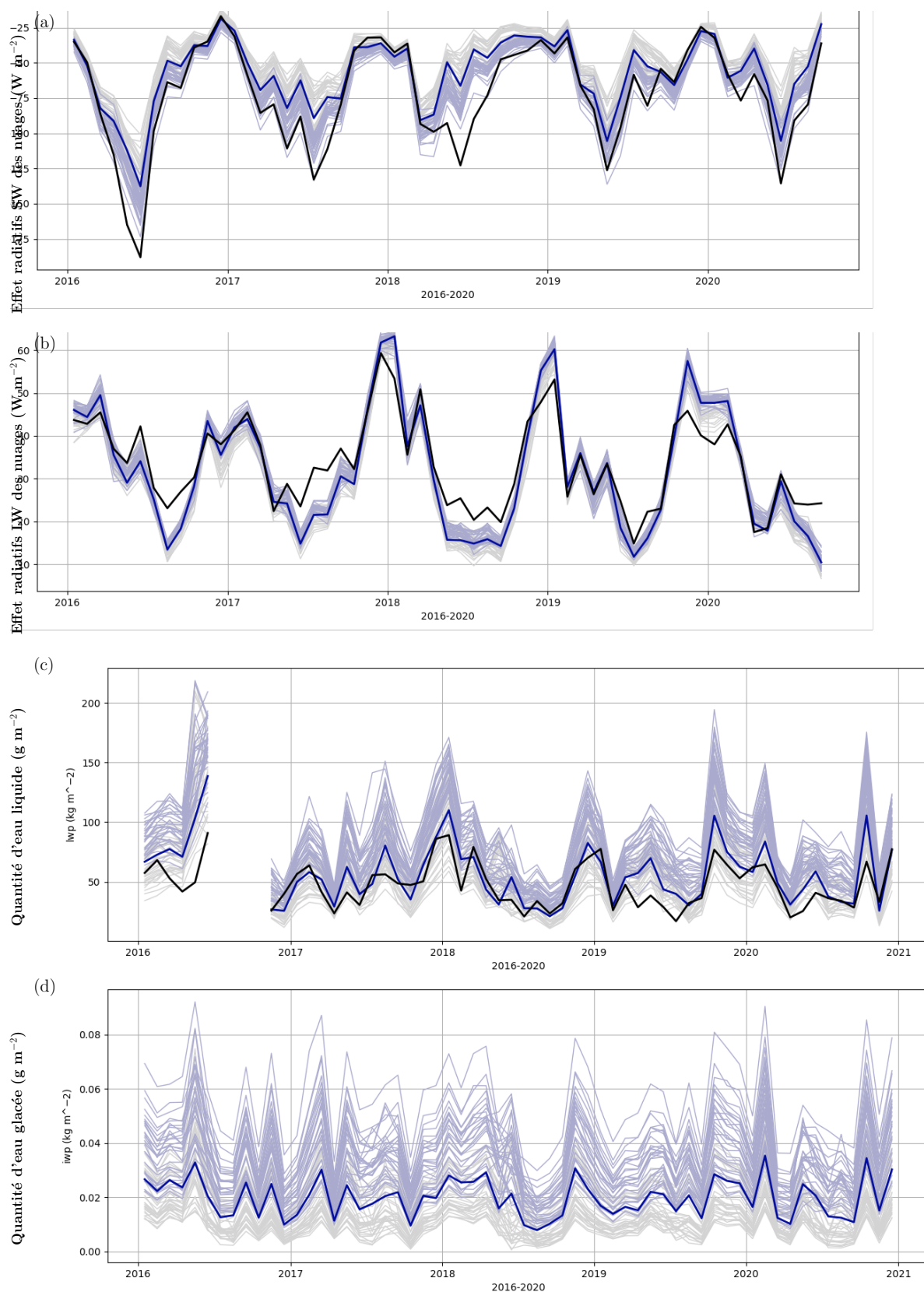


FIGURE V.10 – Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris clair les 30 premiers membre du PPE de la vague 41, en bleu clair les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa

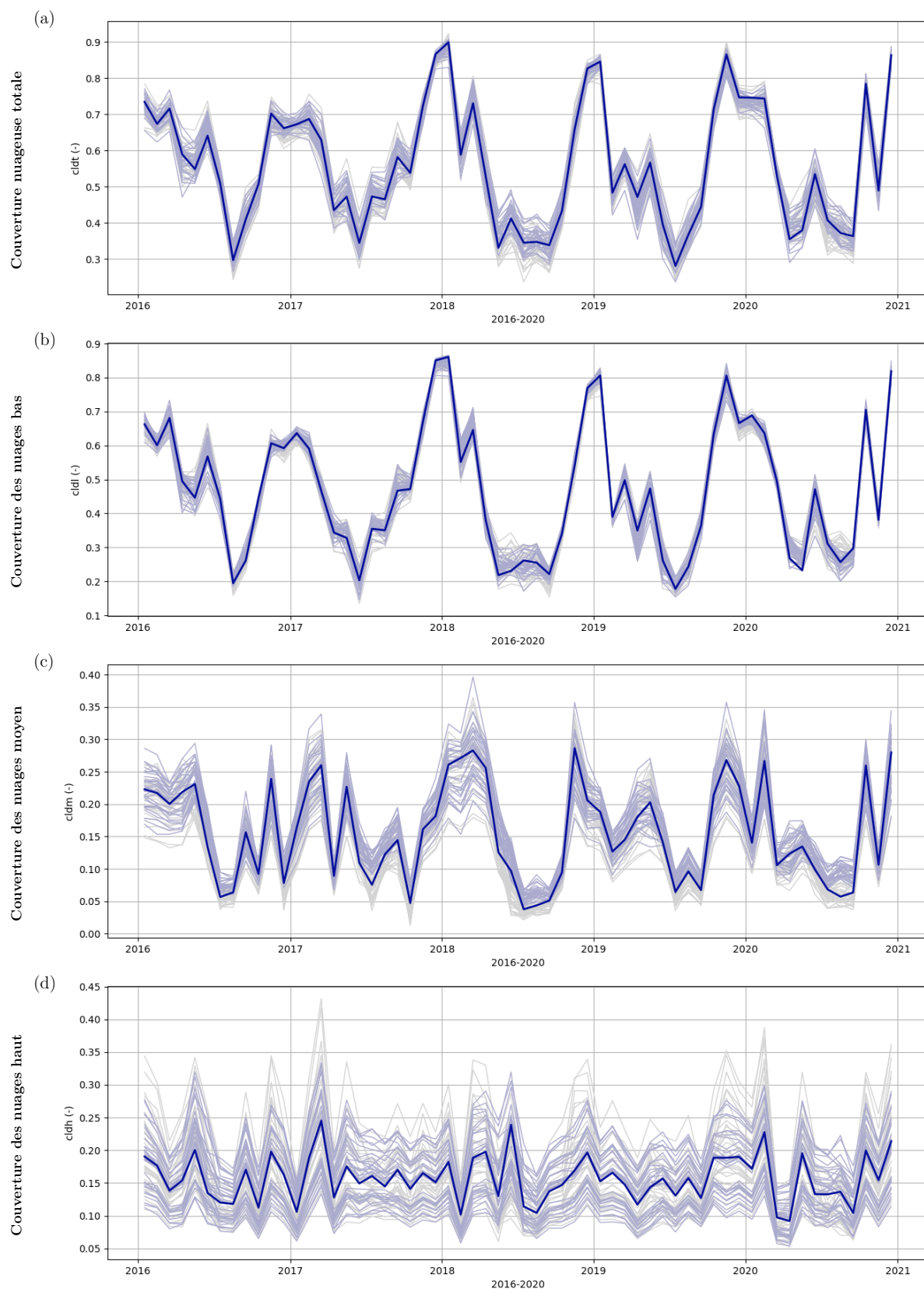


FIGURE V.11 – Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris clair les 30 premiers membre du PPE de la vague 41, en bleu clair les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa

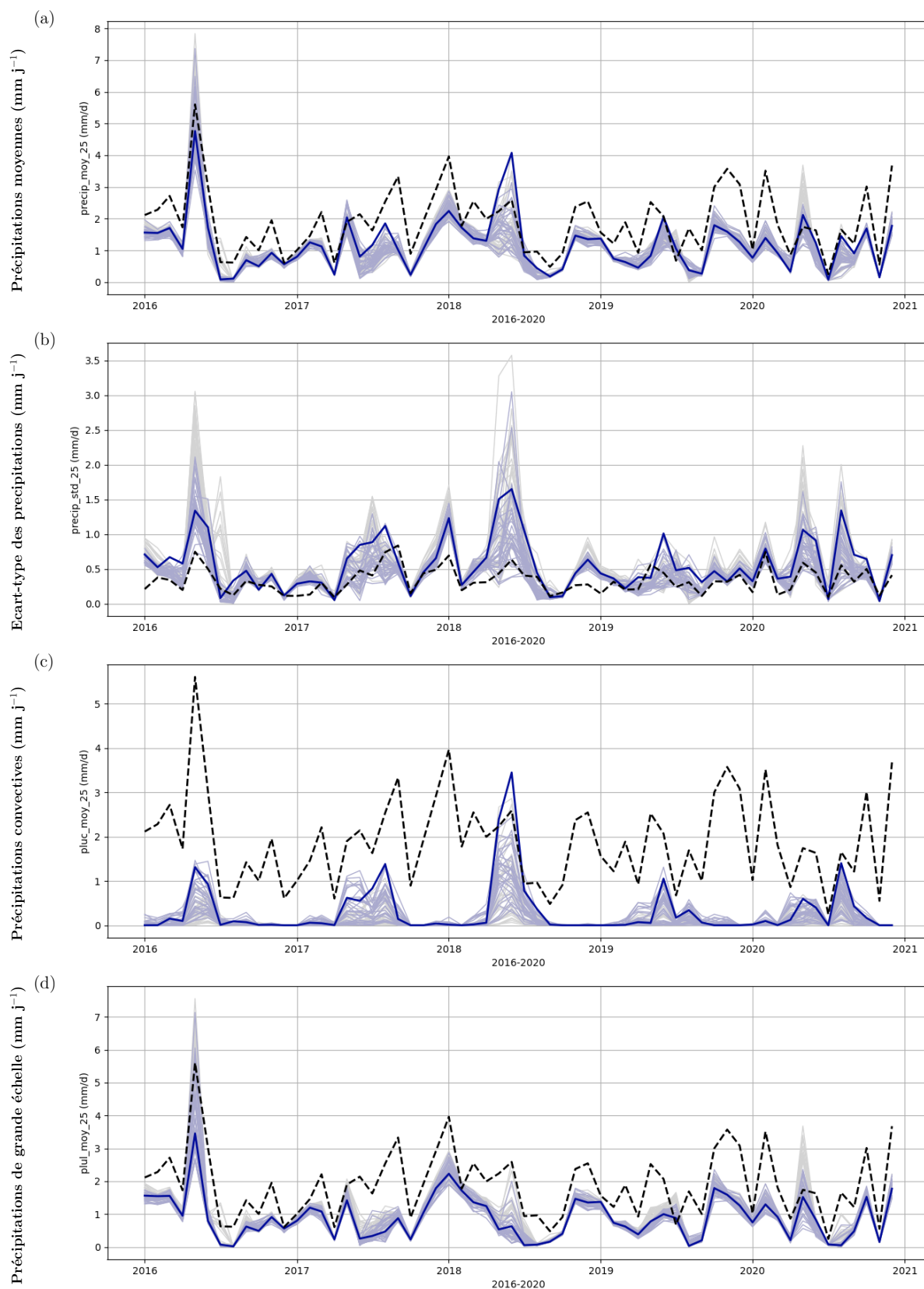


FIGURE V.12 – Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris clair les 30 premiers membre du PPE de la vague 41, en bleu clair les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTA

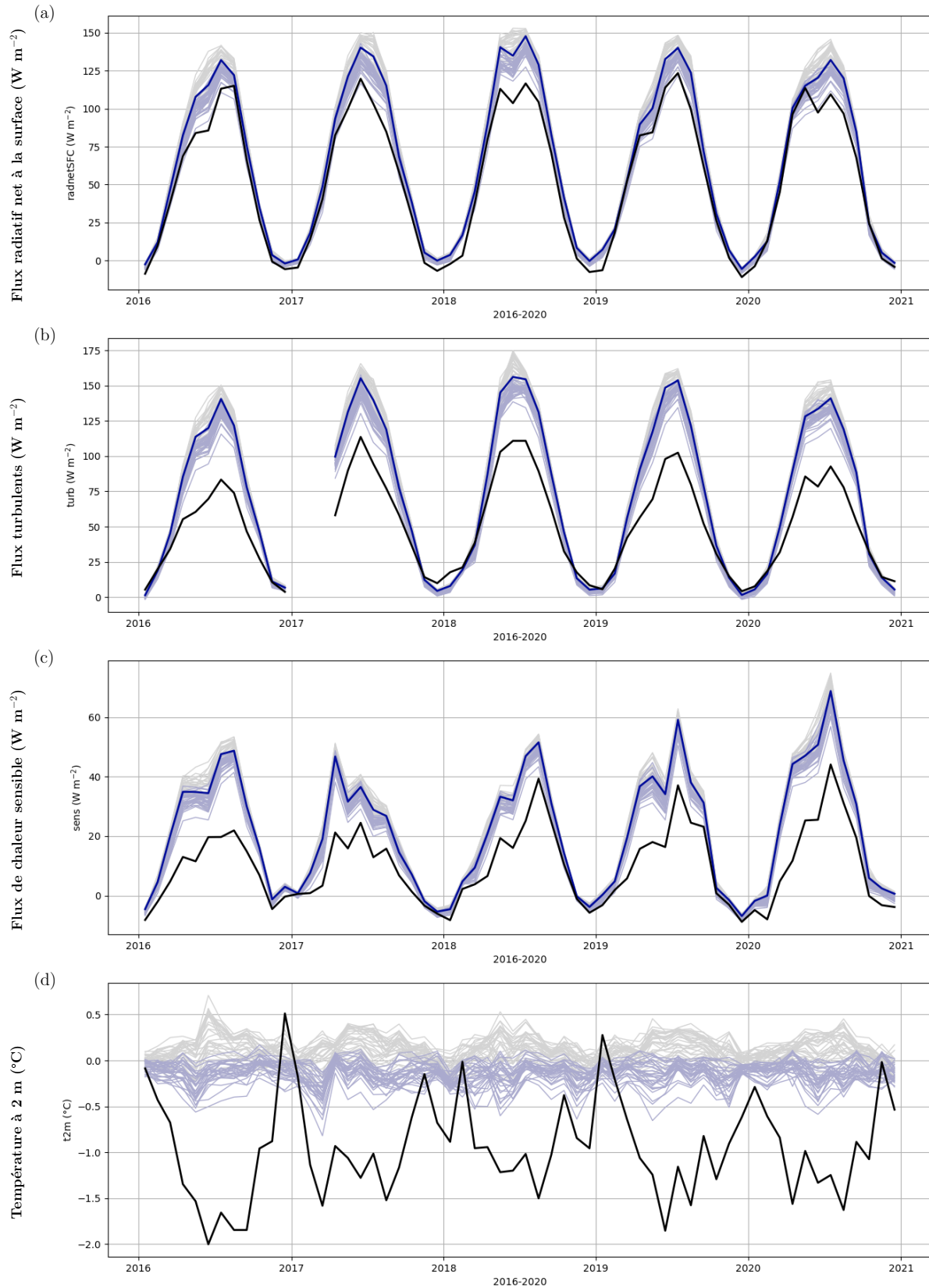


FIGURE V.13 – Evolution en moyenne mensuelle sur la période 2016-2020 de (a) ... En gris clair les 30 premiers membre du PPE de la vague 41, en bleu clair les 30 premiers membre du PPE de la vague 42, en bleu la simulation de contrôle et en noir les observations du SIRTa

Nous souhaitons savoir si les métriques choisies ne sont pas trop spécifiques à la période choisie pour contraindre le modèle. Pour se faire, les 30 premiers membres des PPE41 et 42 sont lancées sur la période 2016-2020. Choisir les 30 premiers membre revient à choisir au hasard 30 membre dans les PPE. *[j'ai aussi fait les trois même bests je sais pas si je les montre ici ; et je ne sais pas ce que je dois montrer comme figure parmi cette sous-sélection :]*

Sur la période 2016-2020, les simulations du PPE42 ont systématiquement plus d'effet radiatif des nuages en SW et en LW, avec des nuages plus chargés en eau liquide et glacée sur toute la période. Cela conduit à une amélioration de la simulation des effets radiatifs sauf sur les mois d'hiver où l'effet radiatif des nuages est peu biaisé dans le solaire et plutôt surestimé dans l'infrarouge. Cela conduit également à l'amélioration du bilan radiatif net en surface tous les étés, avec certains été où des simulations du PPE42 atteignent les valeurs de référence. L'augmentation de la quantité d'eau liquide dans le PPE42 crée en revanche un biais qui n'était pas ou peu présent dans la configuration de référence du modèle. On retrouve également quelques sélections sur les couvertures nuageuses totales, avec des nuages moyens presque systématiquement plus couvrant et des nuages haut légèrement moins couvrant. Contrairement à ce qui était vu sur l'année 2018, certaines périodes (d'été ou d'hiver) ont plus de couverture des nuages bas dans le PPE42, et d'autres en ont moins. Les métriques de cette expérience de tuning restreignent bien les paramètres des paramétrisations des nuages bas avec un effet sur leur couverture nuageuse. La configuration de référence du modèle simule systématiquement trop peu de précipitations (en moyenne sur 25 mailles), même au printemps-été, excepté au mois de juin 2018. La métrique de précipitations a donc été choisie sur un mois exceptionnellement pluvieux. Les différences entre les deux PPE sont plutôt petites. Le PPE42 a tout de même tendance à faire un peu plus de précipitations convectives et un peu moins de précipitations grandes échelles au printemps-été. L'intensité des orages points de grilles est également diminué dans le PPE42, comme pour celui d'avril 2016 et d'avril 2020. L'écart-type spatiale des précipitations est surestimée sur toute la période dans la simulation de référence. Les simulations du PPE42 ont des écart-type spatiaux des précipitations plus faible sur presque toute la période que celles du PPE41.

Finalement, les métriques choisies contraignent de manière similaire l'année 2018 que la période 2016-2020. Cela valide la "robustesse" des métriques choisies et montre la force de choisir des métriques "orientées processus" comme ici. *[bon c'est super mal dit, mais je crois que si c'est bien généralisable c'est qu'on a pas trop mélangé de trucs dans les métriques nan ?]*

V.3 Place dans la chaîne de tuning actuelle

On souhaite évaluer comment se place les métriques du SIRTa dans la chaîne de tuning actuelle du modèle, qui se compose d'une étape dite de préconditionnement 1D décrite Section ?? et d'une étape de tuning avec des métriques globales. Qu'elles contraintes apportent les métriques 1D sur les ensembles de simulations au SIRTa ? Est-ce que les métriques du SIRTa sont compatibles avec les métriques globales voire accélèrent la convergence de ces métriques ?

V.3.1 Effet du préconditionnement 1D sur les simulations au SIRTa

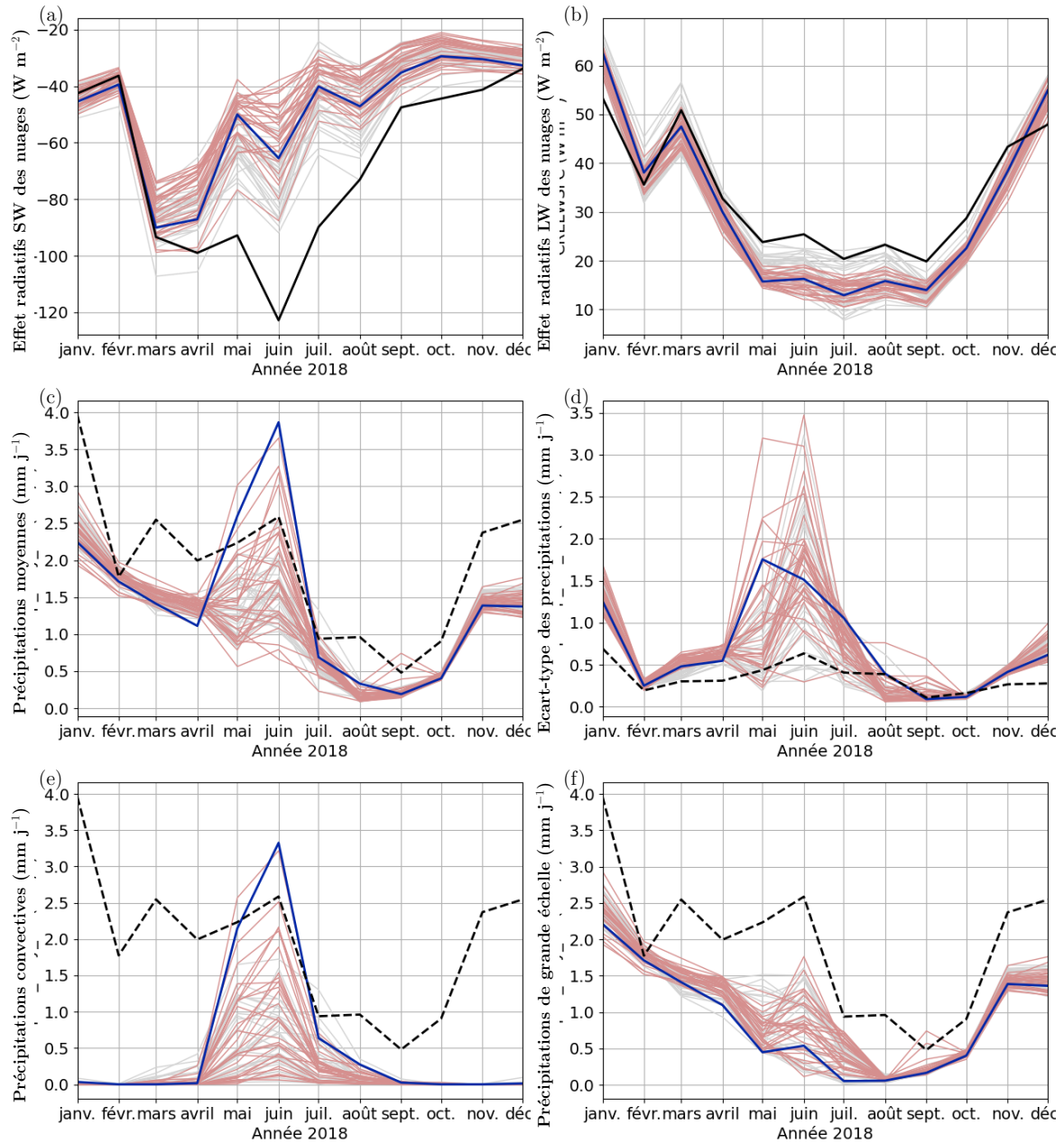


FIGURE V.14 – Evolution en moyenne mensuelle sur l'année 2018 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa

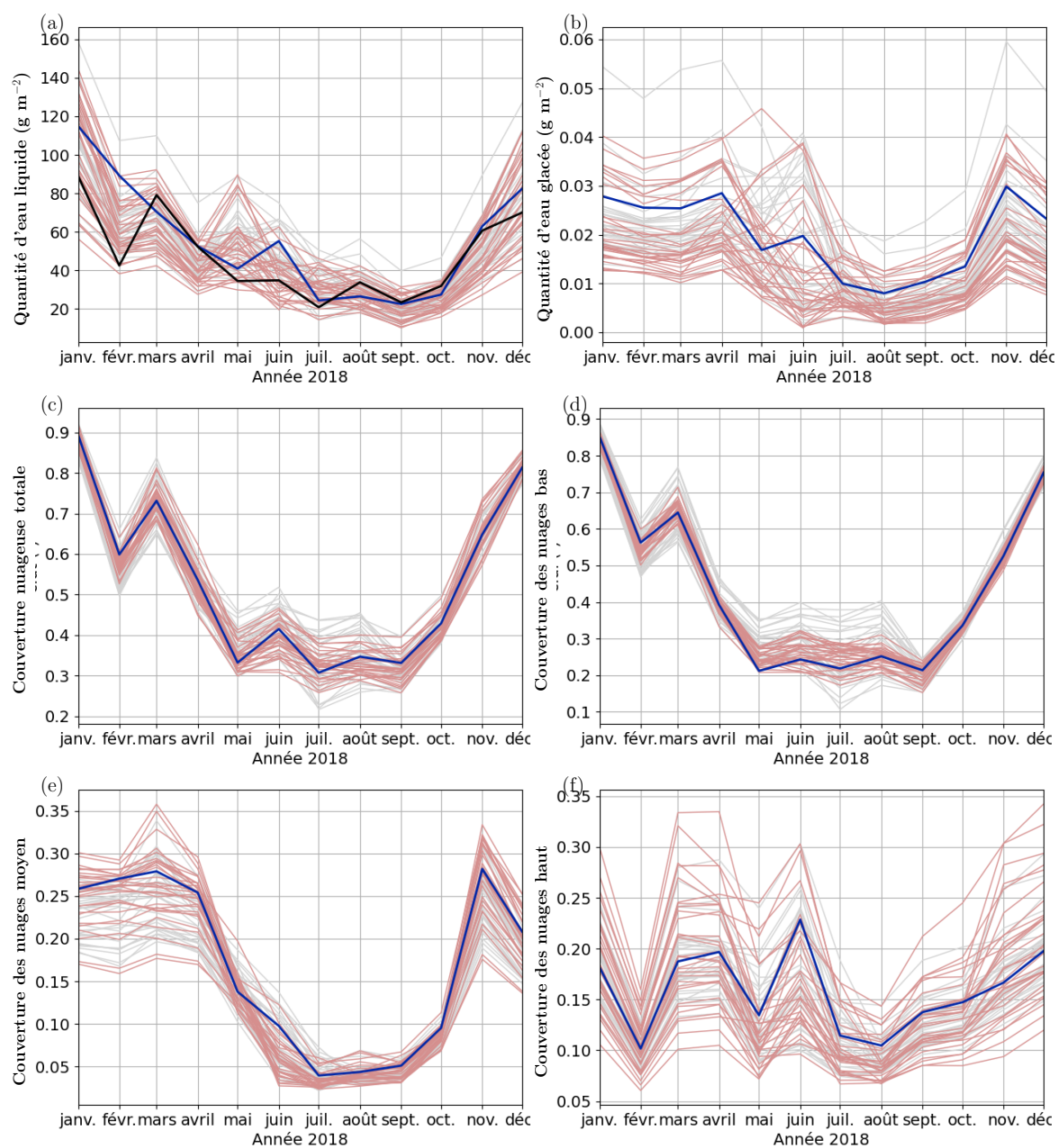


FIGURE V.15 – Evolution en moyenne mensuelle sur l'année 2018 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa

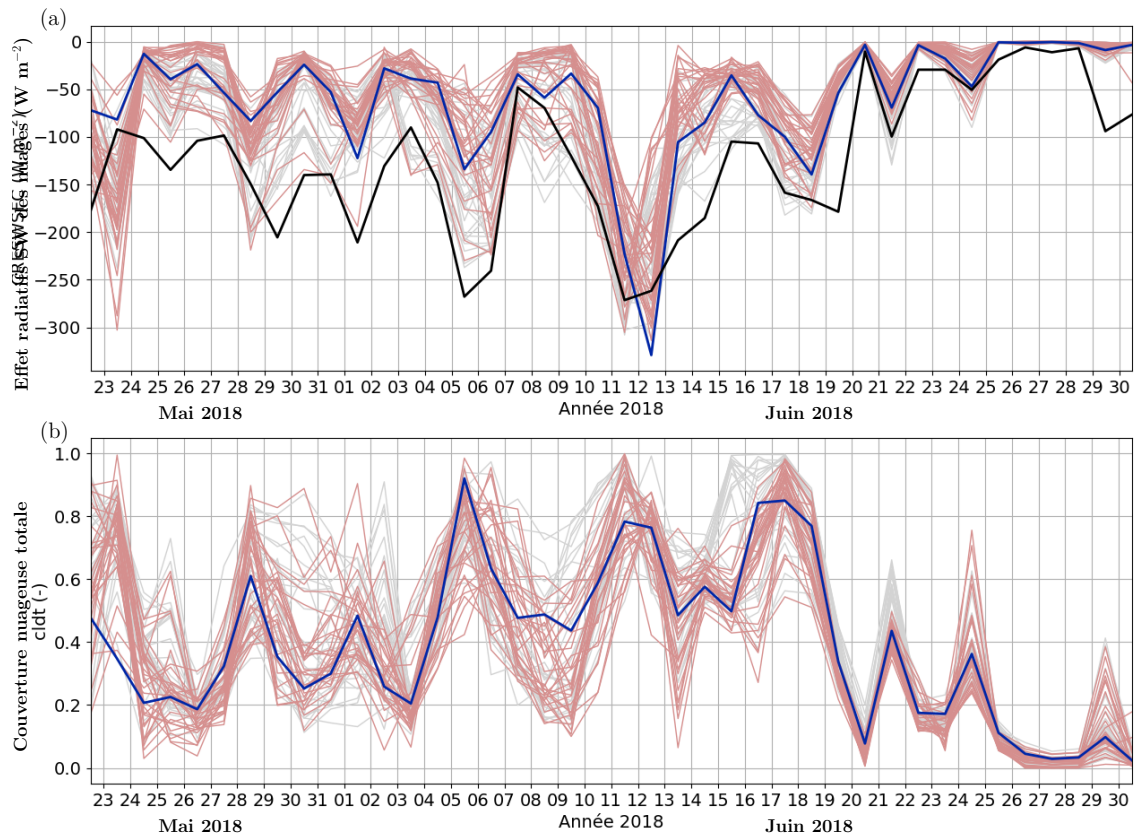


FIGURE V.16 – Evolution en moyenne journalière du 22 mai au 31 juin 2018 de (a) CRESWSFC [c'est plus visible sur le CRELWSFC donc je changerais] (b) /bin/bash : q : command not found En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa

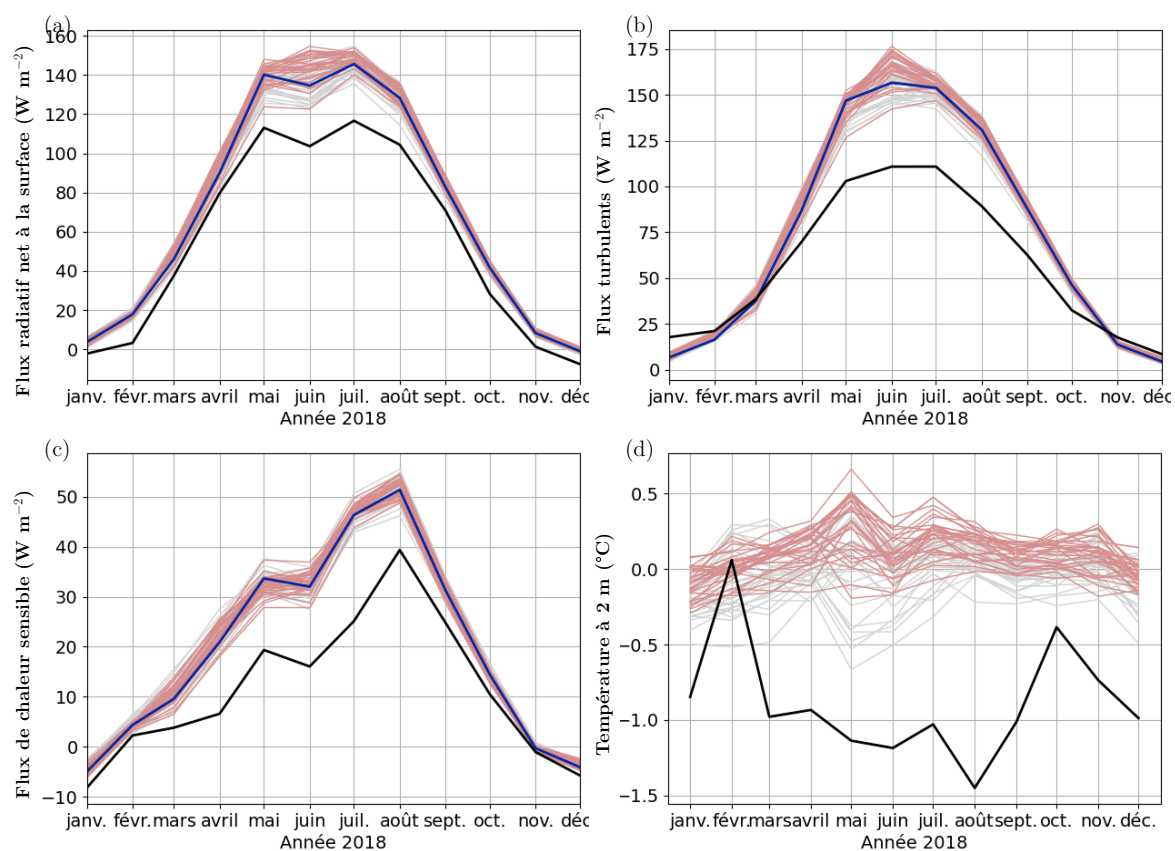


FIGURE V.17 – Evolution en moyenne mensuelle sur l'année 2018 de (a) ... En gris claire les 30 premiers membre du PPE de la vague 1, en rouge claire les 30 premiers membre du PPE de la vague 41, en bleu la simulation de contrôle et en noir les observations du SIRTa

Pour évaluer l'effet du préconditionnement 1D sur les ensembles de simulation de SIRTa, on sélectionne un sous-ensemble de 30 membres issue de la première vague de tuning, que l'on compare à un sous-ensemble de 30 membre issue de la vague 41 du tuning. Les 30 premiers membres des vague 1 et 41 sont choisis, ce qui revient à choisir les membres aléatoirement parmi les NROY_1 et NROY_{41} . L'ensemble de la première vague correspond à un échantillonnage uniforme de l'hypercube formés par les 24 paramètres. L'ensemble de 30 membres de la vague 41 correspond à un ensemble de vecteur ayant été conditionnés uniquement par les métriques 1D (Table V.2) contraignant les couches limites de convection peu profonde.

Le préconditionnement 1D opère une sélection sur les couvertures nuageuse des nuages bas, et en particulier sélectionne des nuages bas moins couvrant en été. Cela se repercute sur la couverture nuageuse totale et sur l'effet radiatif des nuages dans l'IR. Avant préconditionnement, certaines simulations s'approchent de la cible sur l'effet radiatifs des nuages dans l'IR en été, alors que ce n'est plus le cas dans le PPE41. Le biais structurel de l'effet radiatif des nuages dans le solaire est déjà existant dans l'ensemble de simulation avant préconditionnement, et le préconditionnement tend à le renforcer. Une légère sélection s'opère également sur la quantité d'eau nuageuse, qui est plutôt plus faible après le préconditionnement 1D. La période de convection profonde du 22 mai au 12 juin 2018 est également affecté par le préconditionnement 1D, dont les métriques portent uniquement sur des cas de couche limite de convection peu profonde. On voit par exemple un effet radiatif des nuages dans l'IR moins fort après le préconditionnement 1D (Figure ??). Cela s'explique par le fait que (1) les deux paramétrisations sont liées (flux de masse et déclenchement de la convection profonde dépendent de la convectino peu profonde) (2) une modification de la convection peu profonde peu modifier l'état de l'atmosphère donc la convection profonde. La sélection est tout de même moins forte sur les périodes de convection profonde, car ses paramètres n'ont pas encore été contraint.

[Paragraphe sur les conséquences sur la surface]

[Interprétation sur le préconditionnement 1D vs controle CMIP6]

Finalement, le préconditionnement 1D tend à sélectionner des vecteurs de paramètre de convection peu profonde produisant moins d'effet radiatif sur les mois d'été au SIRTa, ce qui renforce les biais du modèle. L'ajout de métriques au SIRTa permet néanmoins de diminuer ces biais, en ajustant les paramètres de convection profonde non contraint en 1D.

[Remarque sur l'interprétation des ensembles : certaines variables comme cldh ont l'air d'avoir plus de variabilité après le préondictionnement qu'avant : effet statistique d'ensemble ? => Aller voir le PIS aussi sur ces variables car c'est peut-être juste ça qu'on voit "en plus"]

V.3.2 3D

V.4 Conclusion

— Ouverture sur expériences de tuning avec les stations BSRN

Chapitre VI

Rayonnement

Remarque lecture : Première partie relue par Najda

[Avant de commencer, redire la place de ce chapitre dans le manuscrit. Je fais une proposition ci dessous ; inspirée des passages que j'ai copié de ton positionnement.]

Nous entrons ici dans le chapitre concernant l'ajustement des flux radiatifs en présence de nuages. *[une phrase pour dire l'enchaînement avec ce qui précède ? Genre "on a vu que les flux radiatifs au SIRTa..."]* Les travaux présentés ici sont motivés par trois éléments, qui structurent le chapitre en trois sections :

1. l'arrivée du nouveau modèle de transfert radiatif ecRad (Hogan and Bozzo, 2018), proposant notamment différentes hypothèses de prise en compte des nuages par le modèle de transfert radiatif, constitue aujourd'hui un contexte très favorable à la revisite de la modélisation des interactions entre nuages et rayonnement dans LMDZ. ecRad est présenté en Section VI.1 ;
2. l'enrichissement du cadre 1D/LES par des calculs de référence radiatifs par Monte Carlo Villefranque et al. (2019) permet de déployer ce cadre pour éclairer la question des compensations d'erreurs entre nuages et rayonnement, ce qui est l'objet d'un article présenté en Section VI.2 ;
3. le développement de la nouvelle version de LMDZ pour l'exercice d'intercomparaison CMIP7, incluant le nouveau modèle de rayonnement ecRad, a permis de réaliser des premiers tests de sensibilité au choix du modèle radiatif en 3D. Des premiers résultats sont présentés en Section VI.3.

VI.1 Transfert radiatif avec ecRad

ecRad est un logiciel de transfert radiatif développé au Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT ou ECMWF en anglais pour *European Center for Medium-Range Weather Forecasts*). Il est utilisé de manière opérationnelle pour les prévisions météorologiques du Centre Européen, effectué avec le modèle global IFS (Integrated Forecasting System), depuis juillet 2017. Son utilisation dans l'IFS est décrite par Hogan and Bozzo (2018). Le développement de ecRad a été motivé par deux grandes raisons : (1) Optimiser le code pour réduire le coût de calcul du transfert radiatif afin de pouvoir l'appeler plus fréquemment et/ou utiliser des modèles de transfert radiatif plus complexes pour le même coût de calcul ; (2) Avoir un modèle de transfert radiatif flexible afin de faciliter son développement, son partage et la recherche qui l'entoure. Cette flexibilité est rendue possible grâce à une organisation modulaire de ecRad en 5 composantes : 4 composantes indépendantes pour calculer les propriétés optiques (de la surface, des gaz, des aerosols et des nuages) et une composante, appelé solveur, combinant les propriétés optiques du ciel-clair, des nuages et de la surface (Figure VI.1). C'est cet

Table 2*Summary of the Main Options Available in ecRad, Each of Which Can Be Selected at Run-Time*

Property	Options
Gas optics model	Monochromatic; RRTM-G (other options are planned)
Aerosol optics model	Generalized (supporting Tegen and CAMS types); optics computed externally
Liquid cloud optics	Slingo (1989) and Lindner and Li (2000); SOCRATES
Ice cloud optics	Fu (1996) and Fu et al. (1998) ; Yi et al. (2013); Baran et al. (2014)
δ -Eddington mode	Particles and gases; Particles only
SW and LW solver	Homogeneous; McICA ; Tripleclouds; SPARTACUS
Longwave scattering	Off; cloud only ; cloud and aerosols
Cloud overlap scheme	EXP-EXP (only available with McICA); MAX-RAN; EXP-RAN
Cloud overlap parameter	alpha (Hogan & Illingworth, 2000) ; beta (Shonk et al., 2010)
Cloud water PDF shape	Gamma ; Lognormal

Note. The settings expected in the next operational cycle (46R1) are shown in bold. The earlier IFS implementation of ecRad in Cycle 43R3 were the same except for longwave scattering being switched off and δ -Eddington scaling being applied to particles and gases. RRTM-G = Rapid Radiative Transfer Model for GCMs; SPARTACUS = Speedy Algorithm for Radiative Transfer through Cloud Sides; SW = shortwave; LW = longwave.

FIGURE VI.2 – Ce tableau est issu de Hogan and Bozzo (2018). Il répertorie les principales options possibles pour les différents modules de ecRad. En gras sont indiqués les options choisies pour la version de ecRad intégrée dans l'IFS.

dans ecRad : Homogeneous, McICA, TripleClouds et SPARTACUS. Le solveur Homogeneous fait l'hypothèse que chaque couche est soit complètement occupée par un nuage homogène, soit totalement ciel-clair. Ce solveur n'est donc pas adapté pour le calcul de transfert radiatif d'un GCM et ne sera pas étudié ici. En plus des tables de propriétés optiques des gaz et des nuages, et des propriétés thermodynamiques de l'atmosphère (température et pression), les solveurs prennent en entrée les profils verticaux de quantité d'eau condensée (liquide et solide), de fraction nuageuse et de taille de gouttes et cristaux. Différentes hypothèses sont formulées afin de pouvoir calculer le transfert radiatif à partir de ces informations.

VI.1.1.1 Hypothèses de prise en compte des nuages disponibles dans ecRad

Recouvrement vertical des nuages

Les solveurs ont besoin de connaître comment les fractions nuageuses fournies en entrées sont positionnées les unes au-dessus des autres. La façon de positionner les fractions nuageuses sur la verticale est appelée modèle de recouvrement vertical. Ce modèle va aussi déduire du profil de fraction nuageuse, la couverture nuageuse, c'est-à-dire la fraction surfacique de la colonne qui, vue du ciel (ou de la surface), est couverte par des nuages. Cette couverture nuageuse influence au premier ordre les flux radiatifs solaires dans la colonne : à fraction nuageuse égale, plus la couverture nuageuse est grande, plus le flux solaire transmis à la surface va être petit. Tous les solveurs n'utilisent pas directement cette information : elle est utilisée explicitement dans MCICA tandis que dans Tripleclouds et Spartacus elle est uniquement diagnostique.

Il existe plusieurs modèles de recouvrement possible, illustrés Figure VI.3. Le modèle de recouvrement aléatoire dispose aléatoirement les fractions nuageuses sur la verticale du modèle : la position du nuage dans une couche donnée ne dépend pas de la position des nuages dans les autres couches. La couverture nuageuse de deux couches nuageuses adjacentes i et j , notée $c_{rand}(i, j)$ est alors donnée par l'Equation VI.1, avec α_i et α_j les fractions nuageuses des couches i et j .

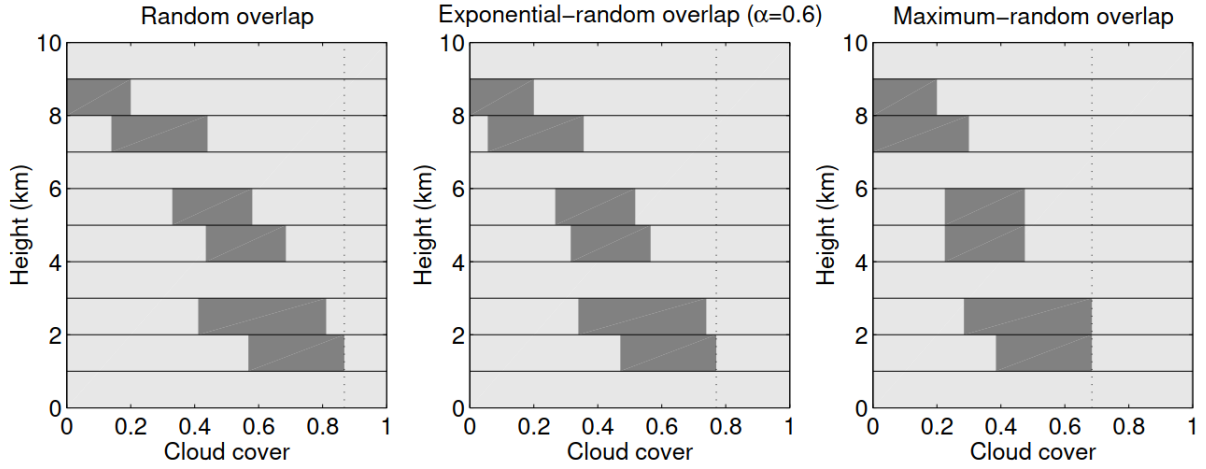


FIGURE VI.3 – Schéma de trois hypothèses possibles de recouvrement des nuages dans les colonnes d'un GCM, issu de Hogan and Shonk (2008). A gauche le recouvrement aléatoire, au milieu le recouvrement exponentiel-aléatoire avec une facteur w égale à 0.6 pour toutes les couches de nuage adjacent et à droite recouvrement maximum-aléatoire. La ligne verticale en pointillé indique la couverture nuageuse totale de la colonne.

$$c_{rand}(i, j) = \alpha_i + \alpha_j - \alpha_i \cdot \alpha_j \quad (\text{VI.1})$$

Parmi les modèles de recouvrement existant, ce modèle maximise la couverture nuageuse totale à profil de fraction nuageuse égale. Il a le défaut d'être très sensible à la résolution verticale du modèle.

Le modèle maximum-aléatoire, appelé "Max-Ran" pour maximum-random dans ecRad, distingue les couches de nuages adjacents, des couches de nuages non adjacents, c'est-à-dire séparés par au moins une couche de ciel-clair. Ce modèle recouvre aléatoirement les couches nuageuses non adjacentes, et maximise le recouvrement des couches nuageuses adjacentes. La couverture nuageuse de deux couches de nuages adjacentes, notée $c_{max}(i, j)$ s'exprime alors avec l'Equation VI.2

$$c_{max}(i, j) = \max(\alpha_i, \alpha_j) \quad (\text{VI.2})$$

Ce modèle minimise la couverture nuageuse à profil de fraction nuageuse égale. Ce modèle de recouvrement est classiquement utilisé dans les GCMs, et c'est celui qui est utilisé dans la version CMIP6 de LMDZ.

Le modèle exponentiel-aléatoire, appelé "Exp-Ran" dans ecRad pour exponential-random, est un modèle de recouvrement intermédiaire aux deux modèles précédents. Il a été développé par Hogan and Illingworth (2000) à partir d'observations radar de nuages. Cette étude montre que lorsque deux couches de nuages sont séparées par du ciel-clair sur la verticale, le recouvrement des nuages de ces couches est bien aléatoire. Par contre, si les couches de nuages ne sont pas séparées par du ciel-clair, leur recouvrement n'est pas forcément maximal comme le suppose le modèle Max-Ran. Ils proposent d'exprimer le recouvrement de ces couches comme une somme pondérée d'un recouvrement maximal et d'un recouvrement aléatoire, selon l'Equation VI.3.

$$c_{exp}(i, j) = w c_{max} + (1 - w) c_{rand} \quad (\text{VI.3})$$

La formulation du poids w accordé au recouvrement maximal est dérivée des observations et montre qu'il peut être exprimé sous forme d'exponentielle décroissante de la distance entre

les couches Δz . On modélise ainsi $w = e^{-\frac{\Delta z}{\ell_0}}$. La longueur caractéristique de cette décorrélation est donnée par ℓ_0 , appelée longueur de décorrélation. Lorsque la distance entre deux couches est très petite devant cette longueur de décorrélation, alors les couches sont encore très corrélées et w est proche de 1 : on retrouve un recouvrement maximal des couches. Si au contraire la distance entre les couches est très grande devant cette longueur de décorrélation, alors les deux couches sont très peu corrélées et w est proche de 0 : on retrouve un recouvrement aléatoire des couches.

Les trois solveurs MCICA, Tripleclouds et Spartacus peuvent au choix utiliser un recouvrement vertical maximum-aléatoire ou exponentiel-aléatoire. Dans ce dernier cas, la longueur de décorrélation ℓ_0 est un paramètre libre de ce modèle de recouvrement vertical des nuages.

Hétérogénéité horizontale de l'eau nuageuse



FIGURE VI.4 – Rendu physiquement réaliste d'un stratocumulus complètement couvrant, à gauche vu du dessous et à droite vu du dessus. On voit des zones plus claires et d'autres plus foncées, traduisant des différences d'épaisseur optique du nuage. Vues du dessous, les zones plus claires correspondent à des zones d'épaisseur optique plus fines. Vu du dessus, c'est l'inverse. Le stratocumulus est issu du cas FIRE simulé avec le modèle LES Meso-NH Lafore et al. (1998); Lac et al. (2018) et la visualisation est rendue possible par le code de transfert radiatif htrdr développé dans la thèse de Villefranche (2019).

Les nuages ont des épaisseurs optiques hétérogènes sur l'horizontale. Cela est particulièrement visible sur les stratocumulus, comme le montre la Figure VI.4. Plus le nuage est hétérogène, plus il transmet du rayonnement solaire et moins son albédo est fort. L'amplitude de ces hétérogénéités peut être modélisée par le paramètre **FSD** (pour *Fractional Standard Deviation*), défini par l'Equation VI.4,

$$\mathbf{FSD} = \frac{\sigma_\tau}{\bar{\tau}} \quad (\text{VI.4})$$

avec $\bar{\tau}$ l'épaisseur optique moyenne du nuage sur l'horizontale et σ_τ son écart-type sur l'horizontale, à l'intérieur d'une couche donnée.

L'épaisseur optique d'une couche d'eau nuageuse est proportionnelle à son contenu en eau divisé par le rayon effectif des condensats. Soit $\tau \propto q_c / r_{eff}$. L'approximation est souvent faite que ce sont les variations horizontales du contenu en eau condensé qui sont les premières responsables des variations horizontales de l'épaisseur optique. Le paramètre d'intensité des hétérogénéités horizontales des épaisseur optique du nuage **FSD** est alors approximé par l'Equation VI.5.

$$\mathbf{FSD} = \frac{\sigma_{q_c}}{\bar{q}_c} \quad (\text{VI.5})$$

avec \bar{q}_c le contenu en eau condensée nuageuse dans la maille, en kg d'eau condensé par kg d'air nuageux, et σ_{q_c} son écart-type sur l'horizontale.

Le paramètre libre **FSD** est utilisé dans les trois solveurs de ecRad pour quantifier l'intensité des hétérogénéités horizontales des épaisseur optique des nuages. Dans la modèle de transfert

radiatif de LMDZ version CMIP6, les épaisseur optiques des nuages sont considérées comme homogènes horizontalement ($\mathbf{FSD} = 0$).

Recouvrement vertical des hétérogénéités horizontales

Le recouvrement vertical des hétérogénéités horizontales a également un impact sur la transmissivité du rayonnement. En effet, si toutes les zones optiquement fines du nuage se recouvrent sur la verticale, et toutes les zones optiquement épaisses aussi, alors le rayonnement solaire transmis sera moins important que si les hétérogénéités horizontales se recouvrent aléatoirement, car le nuage sera au total plus homogène (intégré sur la verticale). Si le modèle de recouvrement choisi est le modèle maximum-aléatoire, alors les hétérogénéités vont également suivre un modèle de recouvrement maximum-aléatoire. Si le modèle vertical de recouvrement choisi est un modèle exponentiel-random, une deuxième longueur de décorrélation est utilisée pour le recouvrement des hétérogénéités, que l'on note ℓ_h .

Effets radiatifs 3D

Dans un champ nuageux 3D, les effets radiatifs 3D sont définis comme la différence entre la résolution des équations du transfert radiatif dans les trois dimensions (le rayonnement se propage dans tout le domaine) et une résolution verticale uniquement (où le rayonnement ne se propage pas horizontalement). Ils quantifient donc ce qui est négligé lorsque le transfert radiatif est résolu sous hypothèse 1D. Aux échelles d'une LES, les effets 3D qui comptent sont "inter-colonnes", c'est à dire que l'erreur vient du fait que l'hypothèse 1D néglige les transferts de rayonnement entre les colonnes. Aux échelles d'un GCM, les effets 3D qui comptent sont "sous-maille", c'est à dire que l'erreur vient du fait que l'hypothèse 1D néglige le transfert horizontal au sein de la colonne (à travers les bords des nuages qui sont eux-mêmes sous-maille, notamment).

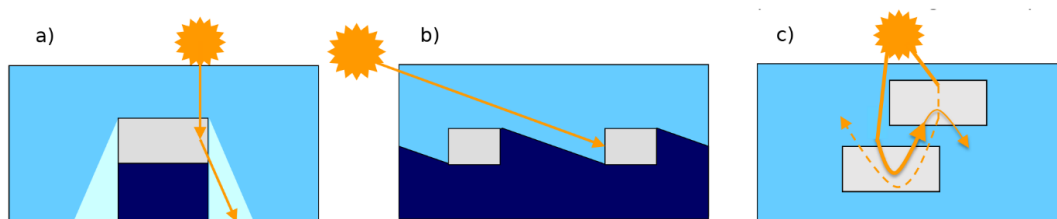


FIGURE VI.5 – Schéma issu de Schäfer et al. (2016), représentant les effets radiatifs 3D associés aux nuages dans le solaire. (a) *side leakage or escape* (b) *side illumination* ou effet d'ombre (c) *entrapment* ou piégeage

Les effets radiatifs 3D dans l'atmosphère sont majoritairement dus à la présence de nuage.

Dans le solaire, les effets radiatifs 3D dus à la présence de nuage sont généralement catégorisés en 4 processus :

- *side leakage or escape* : les photons à l'intérieur du nuage peuvent s'en échapper par les bords latéraux du nuage et atteindre la surface. Ce processus augmente la transmissivité de la scène nuageuse en créant une surillumination de la surface, autour des ombres des nuages (Figure VI.5 (a)). Pour les cumulus, ce processus est majoritaire lorsque le soleil est proche du zénith, et ajoute environ 25% de flux à la surface [citations ?].
- *side illumination* ou effet d'ombre : lorsque le soleil n'est pas parfaitement au zénith, les photons arrivant sur le nuage sont également interceptés par les bords latéraux des nuages. Ce processus augmente la réflectivité et l'absorptivité des nuages. C'est ce qui fait que les nuages ont des plus grandes ombres lorsque le soleil est proche de l'horizon qu'à midi (Figure VI.5 (b)). Pour les cumulus, ce processus est majoritaire lorsque le soleil est proche de l'horizon, et peut doubler le flux réfléchi au sommet de l'atmosphère [citation ?]

- *channeling* : les photons émis par les zones optiquement épaisses du nuage vont rencontrer des zones optiquement plus fines, leur permettant d'atteindre plus facilement la surface, ce qui augmente la transmissivité du nuage. Ce processus se rapproche du side leakage mais a lieu à l'intérieur du nuage, plutôt qu'entre nuage et ciel clair.
- *entrapment* ou piégeage : lorsqu'ils repartent vers l'espace, les photons réfléchis ou émis par la surface ou par un nuage bas peuvent être interceptés par un nuage situé à une altitude plus élevée qui ne serait pas directement au dessus du nuage bas, et redirigés vers la surface (Figure VI.5 (c)). Ce processus domine dans les scènes où il y a plusieurs couches nuageuses, ou en cas de fort albédo de surface (Hogan et al., 2019; ?).

Dans le solaire, les effets 3D peuvent être négatifs ou positifs sur le flux incident en surface, suivant le type de nuage et l'angle solaire (positif pour les angles solaires proches du zenith et négatif pour les angles solaires proches de l'horizon).

Les processus de *side escape*, *channeling* et piégeage, sont fortement liés à la présence de diffusion multiple. Or, dans le domaine infrarouge, les nuages sont moins diffusants que dans le solaire. Ces phénomènes sont donc présents mais les effets 3D sont dominés par l'émission de rayonnement infrarouge par les bords des nuages vers le haut et vers le bas. Les effets radiatifs 3D associés aux nuages sur le flux radiatif arrivant à la surface sont toujours positifs.

Seul le solveur SPARTACUS propose une paramétrisation des effets radiatifs 3D associés aux nuages. Les autres solveurs ainsi que le modèle de transfert radiatif de LMDZ version CMIP6 ne les prennent pas en compte.

VI.1.1.2 Traitement des hypothèses par les solveurs

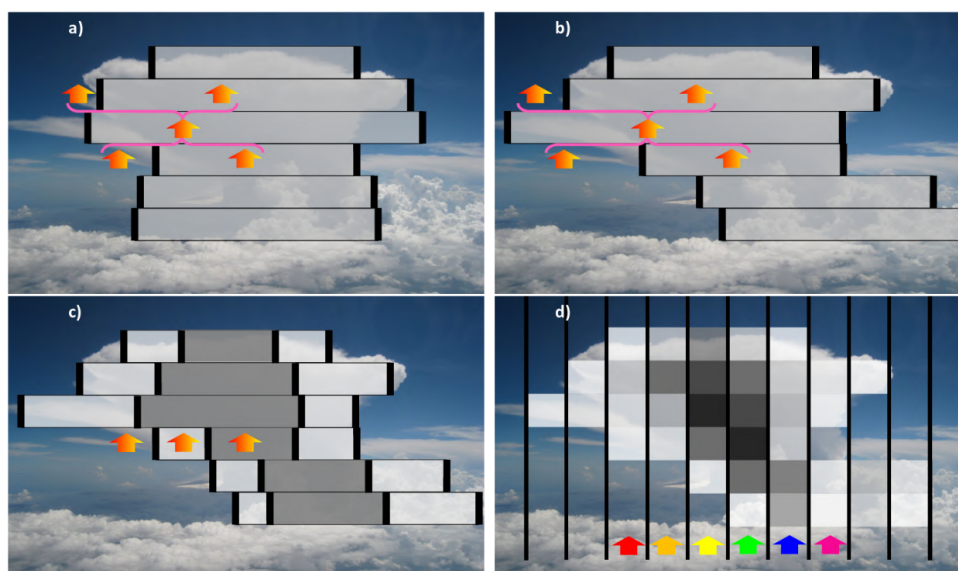


FIGURE VI.6 – Schéma issu de Villefranque (2019) (a) nuage homogène avec recouvrement maximal, ce qui correspond aux hypothèses du schéma de transfert radiatif de LMDZ dans sa version CMIP6, (b) nuage homogène avec recouvrement exponentiel (c) nuage hétérogène illustrant le solveur TripleClouds avec un recouvrement exponentiel (d) nuage hétérogène illustrant la méthode McICA avec un recouvrement exponentiel

McICA

Le solveur McICA (pour Monte Carlo Independent Column Approximation) est décrit en détail dans Hogan and Bozzo (2018). Il combine une approche statistique (Monte Carlo) avec

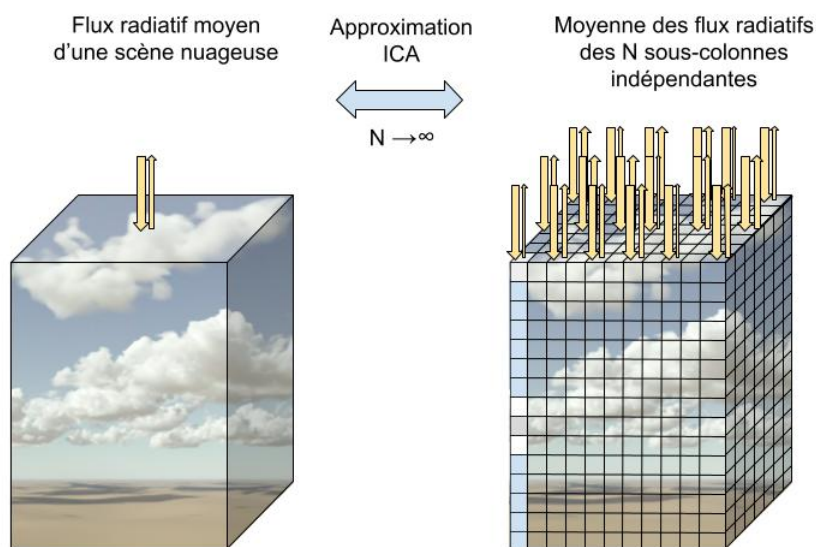


FIGURE VI.7 – Schéma représentant l'approximation des colonnes indépendantes (ICA). Cette approximation stipule que le flux radiatif moyen d'une colonne atmosphérique est égal à la moyenne des flux radiatifs calculés indépendamment dans les N sous-colonnes découpant la colonne initiale (où chaque maille de chaque sous-colonne est optiquement homogène) ; lorsque le nombre de sous-colonnes tend vers l'infini. L'image des nuages est un rendu physiquement réaliste effectué avec htrdr (Villefranche et al., 2019) de la 8ème heure de la simulation du cas ARMCU effectuée par le modèle LES Meso-NH.

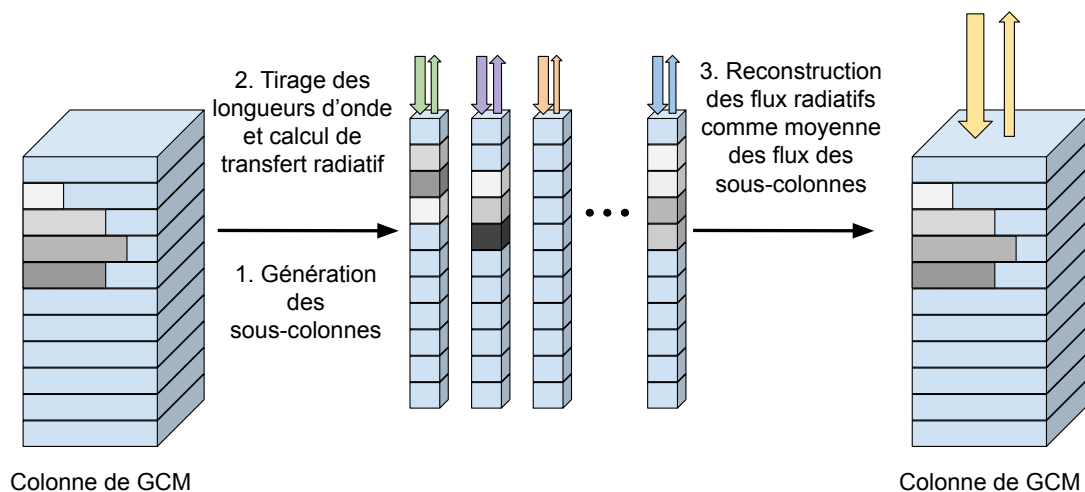


FIGURE VI.8 – Schéma représentant la méthode de McICA pour le calcul du transfert radiatif. Etape 1 : Les profils d'entrées, fournis par le GCM, associés aux hypothèses de recouvrement vertical et d'hétérogénéité horizontale des nuages, sont utilisés pour générer N sous-colonnes atmosphériques. Etape 2 : Un tirage aléatoire est réalisé pour déterminer la longueur d'onde associée à chaque sous-colonne, puis le transfert radiatif monochromatique est résolu dans chaque sous-colonne avec un modèle à deux flux. Etape 3. Les profils de flux radiatifs dans les domaines solaires (ou IR) finaux sont calculés comme la moyenne des profils des sous-colonnes.

l'approximation des colonnes indépendantes (ICA pour *Independent Column Approximation*), qui permet de prendre en compte des modèles de recouvrement verticaux des nuages exponentiel-random ainsi que l'hétérogénéité horizontale des épaisseurs optiques nuageuses.

L'approximation des colonnes indépendantes dit que le flux radiatif moyen d'une colonne

atmosphérique, dont la description spatiale du milieu est parfaitement connue, est égal à la moyenne des flux radiatifs calculés de manière indépendante dans chacune des sous-colonne dont les propriétés optiques sont homogènes dans chaque maille, lorsque le nombre de sous-colonne tend vers l'infini (Figure VI.7). Cette approximation néglige les effets radiatifs 3D.

Cette approximation est utilisée pour calculer le transfert radiatif dans des colonnes atmosphériques hétérogènes typiques des GCMs. Un certain nombre de sous-colonne vont être générées, dans lesquels les propriétés optiques vont être homogènes dans chaque maille. Ces colonnes sont générées de manière à représenter correctement les nuages, c'est-à-dire à représenter correctement leur profils d'eau condensée et de fraction nuageuse fournis par le GCM, mais aussi leurs recouvrement vertical et hétérogénéité horizontale selon les hypothèses souhaitées. La méthode ICA ainsi décrite n'est pas utilisable dans les GCMs à cause de son coût de calcul trop élevé, égale au nombre de colonnes générées multiplié par le coût du transfert radiatif dans chaque colonne. Pour le réduire, Pincus et al. (2003) proposent de ne calculer le transfert radiatif que pour une des fréquences du spectre dans chacune des colonnes. L'échantillonnage des sous-colonne s'accompagne alors d'un échantillonnage des fréquences du spectre. Le fait que la moyenne des flux radiatifs monochromatiques des sous-colonnes tende vers le flux radiatif de la colonne intégré sur le spectre (solaire ou IR) lorsque le nombre de sous-colonnes générées est grand repose sur les principes des méthodes d'estimation de Monte Carlo. Cette méthode introduit un bruit statistique en sous-échantillonnant les distributions, ce qui peut être problématique dans un modèle de Prévision Numérique du Temps par exemple.

TripleClouds

Le solveur TripleClouds propose une extension des modèles de transfert radiatif à deux flux classiques, pour prendre en compte les hétérogénéités horizontales des épaisseurs optiques des nuages. La proposition faite par Shonk and Hogan (2008) est de séparer la partie de la maille nuageuse en deux zones : une zone de nuage optiquement fin homogène et une zone de nuage optiquement épais homogène, comme illustré Figure VI.6 (c).

La définition de ces deux zones est pensée dans Shonk and Hogan (2008) à partir de la densité de probabilité (PDF) de l'eau nuageuse sous-maille, notée p_{Q_c} , et est illustrée Figure VI.9. Dans la proposition initiale, la médiane de la PDF est choisie pour séparer les deux zones. Les fractions de nuage occupées par le nuage optiquement fin α_{fin} et le nuage optiquement épais α_{epais} sont déterminées par les équations VI.6. Par définition, si l'on choisit la médiane pour séparer les deux régions alors les deux fractions sont égales à la moitié de la fraction nuageuse α .

$$\alpha_{fin} = \int_0^{q_c^{lim}} p_{Q_c}(q_c) dq_c \text{ et } \alpha_{epais} = \int_{q_c^{lim}}^{+\infty} p_{Q_c}(q_c) dq_c \quad (\text{VI.6})$$

Puis, le contenu en eau condensée de chacune des deux zones est donné par une valeur unique, q_c^{fin} pour la zone de nuage optiquement fin et q_c^{epais} pour la zone de nuage optiquement épais. Le contenu en eau condensée de la zone de nuage fin est choisi pour correspondre au 16^{ième} percentile de la distribution. Le contenu en eau du nuage épais est ensuite déterminé de manière à ce que la moyenne des deux contenus en eau corresponde au contenu en eau liquide du nuage totale, qui est une donnée d'entrée du modèle.

Dans l'implémentation de TripleClouds, ce sont les expressions de $f^{fin} = q_c^{fin}/\bar{q}_c$ et $f^{epais} = q_c^{epais}/\bar{q}_c$ qui sont paramétrisées puis utilisées comme facteur sur l'épaisseur optique moyenne du nuage afin de calculer l'épaisseur optique du nuage fin et du nuage épais. Si l'on remplace la PDF de l'eau condensé par la PDF des épaisseurs optiques dans les raisonnements de Shonk and Hogan (2008) et sur la Figure VI.9 alors on a bien caractérisé les hétérogénéités horizontales des épaisseurs optiques du nuage et on a : $f^{fin} = \tau^{fin}/\bar{\tau}$ et $f^{epais} = \tau^{epais}/\bar{\tau}$.

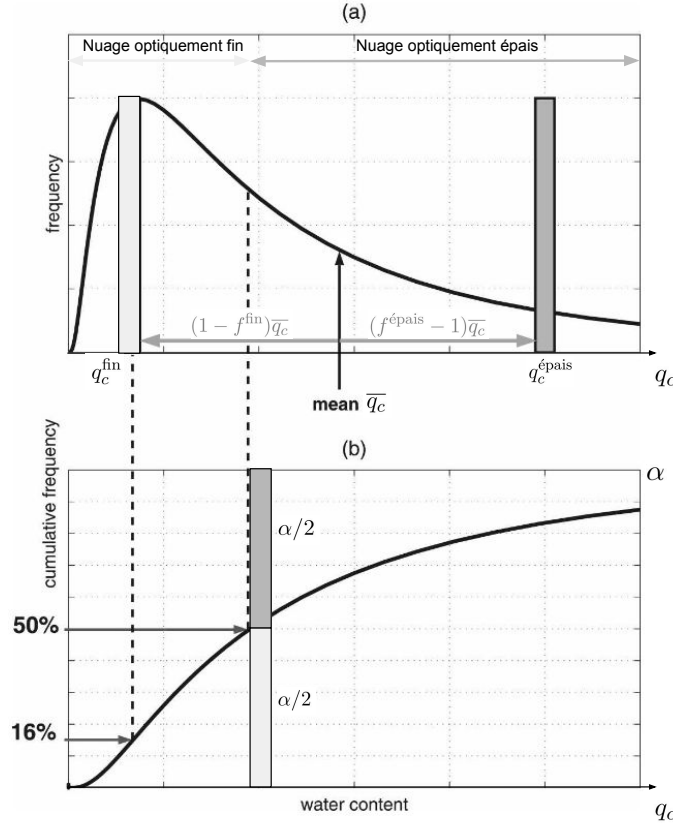


FIGURE VI.9 – Schéma adapté de Shonk and Hogan (2008) (a) fonction de densité de probabilité p_Q du contenu en eau condensée q_c , de moyenne \bar{q}_c (b) fonction de répartition associée. q_c^{fin} (respectivement $q_c^{épais}$) est la valeur du contenu en eau condensée du nuage optiquement fin (respectivement épais), situé à la distance $(1 - f^{fin})\bar{q}_c$ (respectivement $(f^{épais} - 1)\bar{q}_c$) de la moyenne. La valeur asymptotique de la fonction de répartition de l'eau condensée est la fraction nuageuse α .

f^{fin} est une fonction de **FSD**, qui est le rapport entre l'écart-type de la distribution d'épaisseur optique et sa moyenne (Equation VI.4). La fonction dépend de la loi de distribution choisie pour caractériser les hétérogénéités des épaisseurs optiques. C'est sur ce point que les choix ont évolué depuis Shonk and Hogan (2008). propose d'utiliser une loi lognormale puis Hogan et al. (2019) un loi gamma; dont les expressions de f^{fin} , c'est-à-dire de leur 16^{ième} percentile normalisé, sont données Equation VI.7.

$$f^{fin} = \begin{cases} e^{-\mathbf{FSD}[1+0.5\mathbf{FSD}(1+0.5\mathbf{FSD})]} & (\text{gamma}) \\ \frac{e^{-\sqrt{\log(\mathbf{FSD}^2+1)}}}{\sqrt{\mathbf{FSD}^2+1}} & (\text{lognormale}) \end{cases} \quad (\text{VI.7})$$

Hogan et al. (2019) montre qu'utiliser la loi gamma donne de meilleurs résultats qu'avec la loi lognormale, avec toutefois des erreurs qui deviennent importantes pour des grandes valeurs de **FSD** (supérieure à 2). Une correction est alors appliquée aux formulations de f^{fin} . Tout d'abord, f^{fin} est borné pour les valeurs basses par 0.025, tel que : $f_{corr}^{fin} = 0.975f^{fin} + 0.025$. Ensuite, pour des valeurs de **FSD** comprise entre 1,5 et 3,75, la fraction occupé par la région de nuage optiquement fin augmente linéairement de la fraction initiale $0,5\alpha$ à $0,9\alpha$. Puis, pour des **FSD** supérieurs à 3,75, la fraction occupée par la région de nuage optiquement fin est égale à $0,9\alpha$, avec α la fraction nuageuse totale de la maille.

Le recouvrement des hétérogénéités est pris en compte par un paramètre libre f_ℓ tel que la

longueur de décorrélation des hétérogénéités est égale à $f_{\ell} \ell_0$.

SPARTACUS

Le solveur SPARTACUS (pour SPeedy Algorithm for Radiative TrAnsfer through CloUd Side, Hogan et al. (2016), Schäfer et al. (2016) et Hogan et al. (2019)) est un prolongement du solveur TripleClouds qui modélise les effets radiatifs 3D des nuages dans le solaire et l'infrarouge. Cette modélisation consiste plus précisément à modéliser les effets des transferts radiatifs horizontaux au sein de la colonne atmosphérique, sur les flux radiatif verticaux. Cette modélisation est séparée en deux : une paramétrisation des effets des transports radiatifs horizontaux à travers les bords des nuages, au sein d'une même maille, sur les flux radiatifs verticaux montant et descendant aux interfaces de la maille (qui comprend les processus de *side illumination*, *side escape*, *channeling* et émission tridimensionnelle pour l'IR), et une paramétrisation du piégeage qui implique un changement de région d'une partie du flux (ciel-clair vers nuage optiquement fin ou épais) d'une couche à une autre, mais qui n'aurait pas eu lieu s'il s'était propagé uniquement verticalement (c'est à dire que le modèle de recouvrement nuageux utilisé dans un cadre 1D ne suffit pas à expliquer ce transfert).

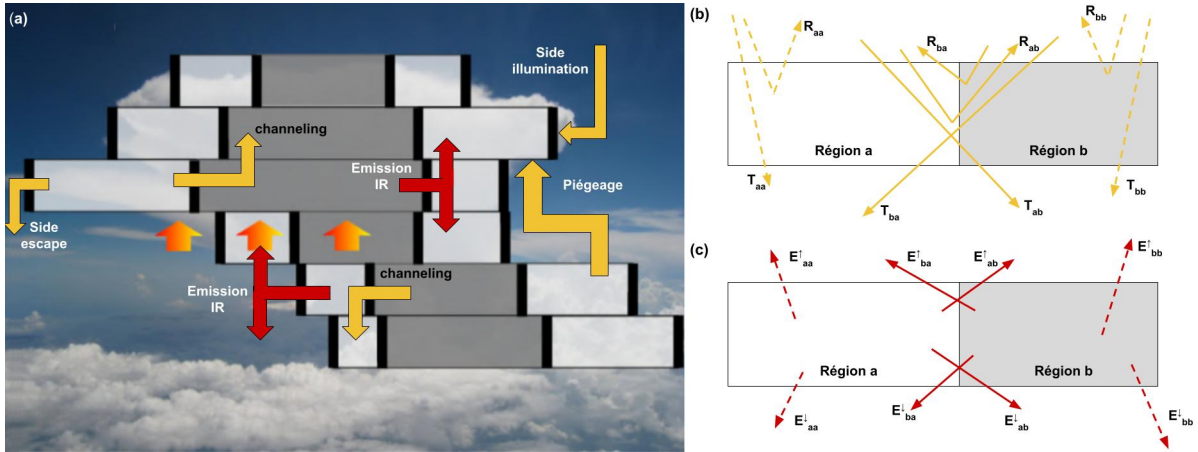


FIGURE VI.10 – Schémas représentant les flux radiatifs paramétrisés par le solveur SPARTACUS. (a) Lien entre processus et flux radiatifs paramétrisés dans SPARTACUS. Les trois flèches oranges représentent les flux radiatifs déjà paramétrisés par le solveur TripleClouds. Les flèches jaunes et rouges représentent les flux radiatifs liés aux effets 3D, les flèches rouges pour les flux radiatifs spécifiques à l'IR. (b) et (c) Schéma adapté de Hogan et al. (2016) représentant les échanges radiatifs modélisés dans SPARTACUS entre deux régions a et b, qui peuvent être du ciel-clair et/ou des nuages optiquement fin et/ou épais. Les flèches pointillées schématisent les échanges déjà modélisés dans TripleClouds, celles en traits pleins les échanges ajoutés dans Spartacus (b) Le flux radiatif arrivant en haut d'une région par diffusion peut être soit réfléchi (R) soit transmis (T) dans la même région (R_{aa} , R_{bb} , T_{aa} , T_{bb}) ou dans une autre région (R_{ab} , R_{ba} , T_{ab} , T_{ba}) (c) l'émission radiative dans une région dans l'IR (E) contribue au flux radiatif ascendant (E^\uparrow) et descendant (E^\downarrow) de la même région (E_{aa} , E_{bb}) ou d'une autre région (E_{ab} , E_{ba}).

Une première paramétrisation modélise l'effet des transferts radiatifs horizontaux à l'intérieur d'une couche entre les régions ciel-clair et nuage optiquement fin, et entre les régions de nuage optiquement fin et optiquement épais. Le transport au sein d'une couche entre nuage optiquement épais et ciel-clair n'est pas autorisé. Les échanges pris en compte entre deux régions d'une même maille sont schématisés Figure VI.10 (b) et (c). Le traitement du recouvrement vertical des nuages et de leurs hétérogénéités est pris en compte avec un modèle exponentiel-aléatoire. Tous ces transports horizontaux sont proportionnels à la surface de l'interface entre les couches. Les nuages sont considérés remplir toute la hauteur de la couche, comme sur la Figure VI.10. Les transports horizontaux sont alors proportionnels à la hauteur de la couche

et à la longueur de l'interface entre les régions ramenée à la surface de la maille. Cette densité de longueur des interfaces entre le ciel-clair et le nuage optiquement fin correspond à la densité de bord des nuages, ou périmètre des nuages, noté p , en m^{-1} . Lorsque la fraction nuageuse est supérieure à 0.5, on parlera plutôt de densité de périmètre des trous.

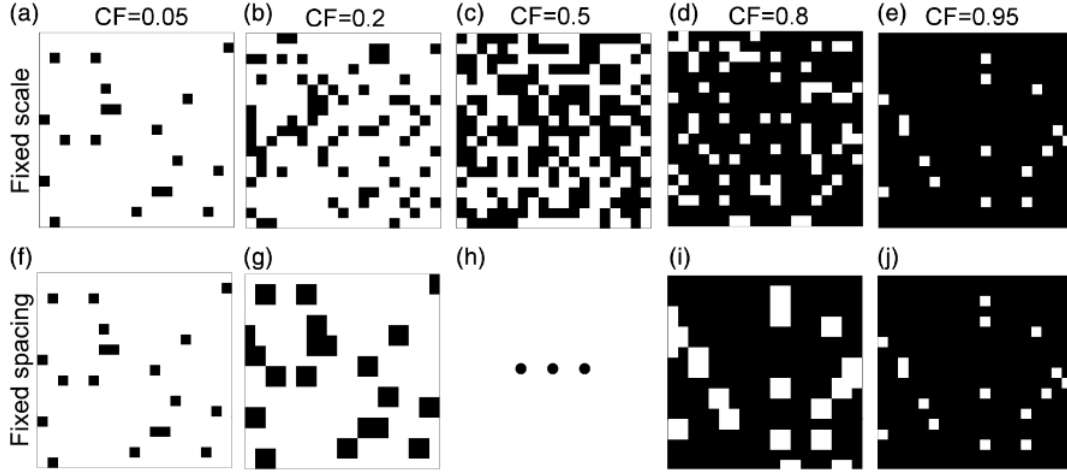


FIGURE VI.11 – Schéma issu de Fielding et al. (2020) montrant la disposition des nuages (en noir) dans des pixels de $1 \times 1 \text{ km}^2$ pour différentes fraction nuageuses (colonne) (a-e) en utilisant la formulation VI.8 avec une taille effective des nuages $C_s = 1 \text{ km}$ et (f-j) en utilisant la formulation VI.9 avec un espacement effectif des nuages $C_x = 4.6 \text{ km}$. On voit le nombre de nuages augmenter avec la fraction nuageuse (a-c) puis le nombre de trous diminuer avec la fraction nuageuse (d-e) lorsque l'on utilise la formulation avec une taille effective des nuages fixe. Lorsque l'on utilise la formulation avec un espace effectif fixe, on voit une augmentation de la taille des nuages avec la fraction nuageuse (f-g) puis une diminution de la taille des trou avec l'augmentation de la fraction nuageuse (i-j)

Plusieurs formulation de cette grandeur ont été proposées, afin de la relier à un paramètre libre, idéalement indépendant de la fraction nuageuse. Schäfer et al. (2016) propose de formuler ce périmètre surfacique par l'Equation VI.8

$$p = \frac{4\alpha(1 - \alpha)}{C_s} \quad (\text{VI.8})$$

Avec α la fraction nuageuse et C_s un paramètre libre représentant la taille effective des nuages, en mètre. Cette formulation est illustrée sur la première ligne de la Figure VI.11, où l'on voit que le nombre de nuages augmente avec la fraction nuageuse, jusqu'à une fraction nuageuse de 0.5, puis le nombre de trou diminue lorsque la fraction nuageuse augmente.

Fielding et al. (2020) propose une formulation de la densité de longueur d'interface p en fonction du paramètre libre C_x , représentant l'espacement effectif des nuages (en mètre), selon l'Equation VI.9, avec α la fraction nuageuse dans la maille.

$$p = \frac{4\sqrt{\alpha(1 - \alpha)}}{C_x} \quad (\text{VI.9})$$

Pour des fractions nuageuses inférieures à 0.5, la surface des nuages augmente lorsque la fraction nuageuse de la maille augmente, pour un même nombre de nuage. Pour des fractions nuageuses supérieures à 0.5, la surface des trous diminue lorsque la fraction nuageuse augmente, pour un même nombre de trous (Figure VI.11 (f-j)).

De même, la densité de périmètre de l'interface entre nuage optiquement fin et nuage optiquement épais, notée p_{cl} , est déterminée en utilisant soit l'équation VI.8 soit l'équation VI.9,

selon l'option choisie pour calculer p , en remplaçant α la fraction nuageuse α par la fraction nuageuse du nuage épais $\alpha_{\text{épais}}$.

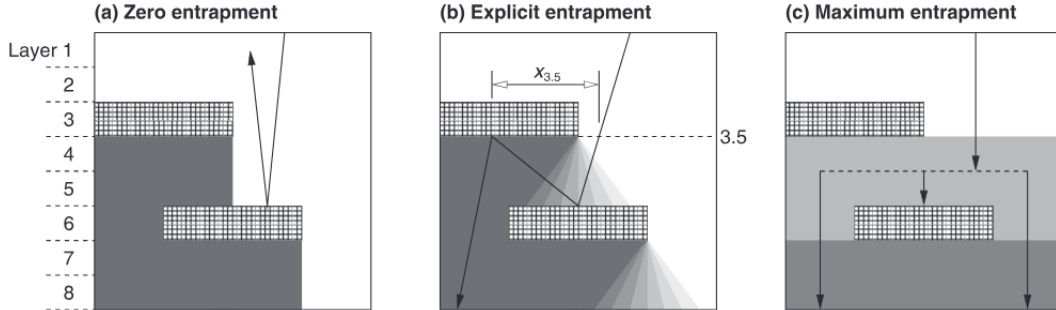


FIGURE VI.12 – Schéma issu de Hogan et al. (2019) représentant les différents traitements possibles pour le piégeage dans SPARTACUS. Deux nuages de fraction 0,5 et épais d'une couche sont situés à la couche 3 et à la couche 6 de la colonne atmosphérique. Leur recouvrement est aléatoire. L'atténuation du rayonnement solaire descendant par les nuages est montrée avec les différents niveaux de gris. (a) cas sans piégeage (b) avec traitement explicite du piégeage et (c) avec traitement maximum du piégeage.

Le processus de piégeage prend en compte le fait que les rayons peuvent changer de région au cours de leur propagation à cause de leur direction de propagation qui n'est pas totalement verticale. Par exemple, un rayon réfléchi par le haut d'un nuage repart toujours vers l'espace s'il se propage uniquement sur la verticale (Figure VI.12 (a)) alors qu'il peut en fait être intercepté par un nuage situé plus haut (Figure VI.12 (b)), y compris un morceau du même nuage (Figure VI.10 (a)). Ce processus de piégeage existe même sans modélisation des bords des nuages, contrairement aux autres effets radiatifs 3D modélisés dans SPARTACUS. Il est modélisé explicitement dans SPARTACUS (Hogan et al. (2019), Figure VI.12 (b)) via la paramétrisation de la distance moyenne horizontale parcourue par les rayons qui sont interceptés par un nuage (la distance $x_{3,5}$ sur la Figure VI.12 (b)). Ce processus est important dans les scènes où existent plusieurs couches de nuage (cumulus recouverts de cirrus par exemple). La modélisation de ce processus conduit toujours à une diminution de la réflectivité des scènes nuageuses (Hogan et al., 2019).

Les effets radiatifs 3D moyens dans le solaire, évalués sur 65 scènes nuageuses différentes par Hogan et al. (2019) ou par Barker et al. (2015), ont tendance à réduire la réflectivité des scènes nuageuses et ont donc un effet réchauffant sur l'atmosphère.

VI.1.2 ecRad et le cadre SCM/LES

Les paramétrisations physiques de LMDZ sont traditionnellement développées et évaluées dans le cadre SCM/LES, ou 1D/LES. Dans ce cadre, une version uni-colonne du GCM, appelée SCM pour Single Column Model, et une simulation LES sont tournées en utilisant les mêmes forçages (profils initiaux, conditions de surface, avec parfois des tendances supplémentaires (sources ou puits) de vents, d'humidité et de chauffage imposés dans la colonne du GCM et dans le domaine de LES). Les résultats de la simulation LES, dont la résolution horizontale est de l'ordre de la dizaine ou centaine de mètres, sont alors utilisés comme référence pour la colonne du GCM associée. Les forçages sont choisis pour être typiques des situations météorologiques dont les processus sont paramétrisés dans le GCM (couche limite stable pour la turbulence, couche limite convective avec ou sans nuage, convection profonde etc). Une bibliothèque de ces cas 1D rassemblant les forçages disponibles dans un format standardisé afin d'être utilisable par tous les modèles, est mise à disposition par Dephy (2020) (Développement et Evaluation PHYsique des modèles Atmosphériques). Ce cadre de travail permet de développer et d'évaluer les para-

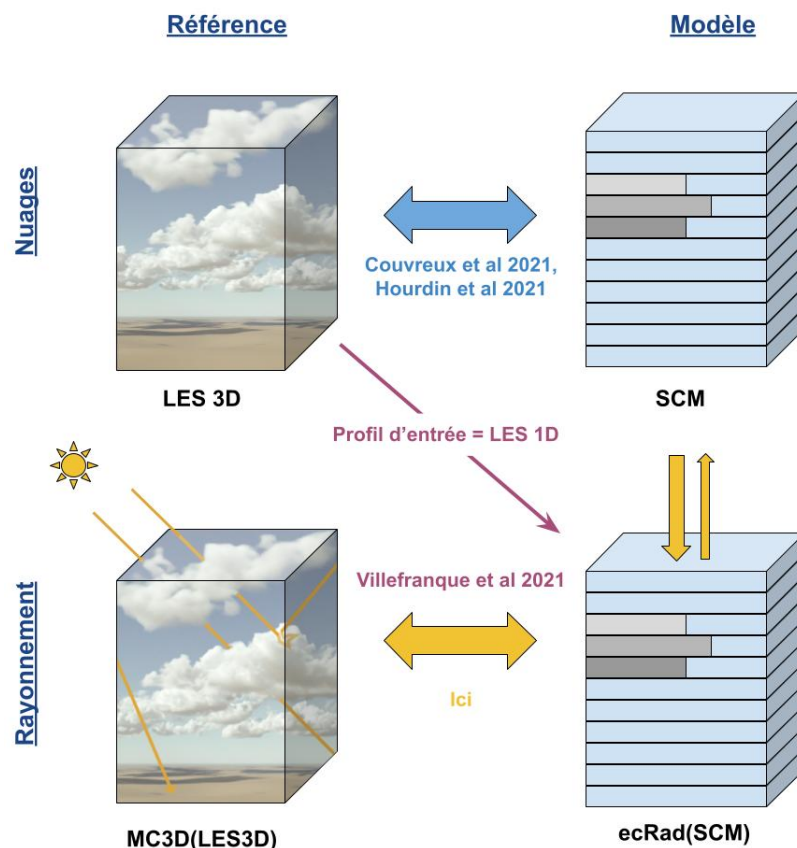


FIGURE VI.13 – Schéma illustrant le cadre de travail SCM/LES (ou 1D/LES), pour l'évaluation, le développement et le tuning des GCMs. Les sorties de simulation d'une colonne de GCM peuvent être comparées aux sorties de simulations LES tridimensionnelles qui servent de références pour les variables de physique atmosphériques et nuageuses, et pour le tuning des paramètres libres du GCM (Couvreur et al., 2021; Hourdin et al., 2021). Les champs de LES convertis en propriétés optiques sont utilisés pour calculer des références radiatives par Monte Carlo (MC3D), qui peuvent ensuite être comparées à un calcul de transfert radiatif à deux flux utilisant comme profil d'entrée le profil moyen de la LES 3D (Villefranke et al., 2021) ou le profil du SCM (ici).

métrisations du GCM au niveau des processus et sans interaction avec la dynamique, dans un environnement très contrôlé et à très faible coût de calcul.

Ce cadre de travail a été élargi au tuning des paramètres libres des modèles atmosphériques (LMDZ à l'IPSL, Meso-NH, AROME, ARPEGE à Météo-France) sur des cas de convection peu profonde dans le cadre de l'ANR HighTune, comme expliqué en détail section [mettre ref section]. Un des objectifs de l'ANR HighTune était de pouvoir tuner conjointement les propriétés physiques et radiatives des nuages à l'échelle des processus. Cependant, les modèles de LES utilisent des paramétrisations du transfert radiatif basées sur les mêmes hypothèses que dans les modèles de climat (approximation plan parallèle, colonnes indépendantes donc effets radiatifs 3D négligés). Les flux radiatifs simulés par la LES ne peuvent donc pas être considérés comme une référence. La première étape pour élargir le cadre 1D/LES au rayonnement a été de développer des outils de calcul de référence de transfert radiatif sur les scènes nuageuses simulées par les LES, décrit section VI.1.2.1. Ces outils ont été développés lors de l'ANR HighTune et dans le cadre de la thèse de Villefranke (2019). Les calculs ainsi réalisés ont servi comme référence pour le tuning des paramètres libres du solveur SPARTACUS de ecRad dans l'étude de Villefranke et al. (2021), dans lequel ecRad était exécuté en mode diagnostique sur les profils atmosphériques issues de la même simulation LES ayant servi pour le calcul de référence du

rayonnement (section VI.1.2.2). Enfin, dans la présente thèse, le cadre 1D/LES a été élargi au tuning des paramètres libres de ecRad et/ou de LMDZ, dans lequel ecRad est tourné en mode diagnostique et prenant en entrée le profil issu du SCM; et en utilisant comme référence les simulations LES et leurs références radiatives associées. Les choix permettant cette comparaison sont détaillés section VI.1.2.2.

VI.1.2.1 Calculs radiatifs de référence par Monte Carlo sur des LES

Les calculs de référence de transfert radiatif sur des scènes nuageuses issues de simulations LES ont été développés dans la thèse de Villefranke (2019). Ces calculs de transfert radiatif sont basés sur une méthode de Monte Carlo en trois dimensions sur des champs de LES décrivant le milieu de propagation du rayonnement (l’atmosphère nuageuse) également en trois dimensions.

Les méthodes de Monte Carlo sont des méthodes probabilistes, permettant d’estimer l’espérance d’une variable aléatoire X en réalisant un grand nombre de variables aléatoires indépendantes et identiquement distribuées à X . L’espérance d’une variable aléatoire X , noté $E[X]$, s’exprime pour les cas continus comme : $E[X] = \int_{x \in \Omega} xp_X(x)dx$, avec Ω le domaine des réalisations possibles de X et p_X la fonction densité de probabilité associée à X . L’espérance de X peut s’estimer à partir d’un échantillon de l’estimateur \bar{X} , construit comme la moyenne de n variables aléatoires i.i.d $\{X_i\}_{i \in [1;n]}$, soit : $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Le théorème central limite assure que la loi de \bar{X} tend vers une loi normale d’espérance $E[X]$ et d’écart type σ_X/\sqrt{n} , lorsque n tend vers l’infini. Autrement dit, pour un n assez grand, une réalisation de \bar{X} , obtenue en réalisant les n variables aléatoires i.i.d X_i , a une probabilité de 95% de tomber dans l’intervalle de confiance $[E[X] - 2\sigma_X/\sqrt{n}, E[X] + 2\sigma_X/\sqrt{n}]$. Les méthodes de Monte Carlo sont très utiles pour estimer des intégrales de fonction dont on ne connaît pas la primitive. Si on note \mathcal{I} l’intégrale d’une fonction quelconque f , on peut réécrire cette intégrale comme l’espérance de la variable aléatoire $f(U)$ avec U une variable aléatoire que l’on sait échantillonner (par exemple de loi uniforme), et estimer cette espérance à partir de la réalisation d’un échantillon de variables U_i i.i.d à U , $\{f(u_i)\}_{i \in [1;n]}$. Les méthodes de Monte Carlo sont particulièrement efficaces pour estimer des intégrales multiples, où un tirage successif peut être fait à chaque nouvelle intégrale, grâce à la propriété de linéarité des espérances, aussi appelée ”double randomisation” (Villefranke et al., 2022).

L’équation du transfert radiatif est très difficile à résoudre avec des méthodes déterministes en milieu hétérogène, à cause de la diffusion multiple et de sa haute dimensionalité (3 dimensions spatiales, 2 dimensions directionnelles). Cependant, il existe une formulation intégrale de cette équation, qui peut être développée sous forme d’intégrales multiples grâce à la méthode du noyau itéré. Cette écriture est favorable à une lecture probabiliste, et à l’utilisation des méthodes de Monte Carlo. Les méthodes de Monte Carlo peuvent ainsi être utilisées pour estimer la luminance en un point, une direction et une longueur d’onde, ou des intégrales de cette luminance (sur une partie du spectre ou sur les angles solides par exemple).

Dans le cadre de la comparaison avec ecRad, on souhaite estimer le flux radiatif solaire descendant (ou montant) à une certaine altitude z_0 par méthode de Monte Carlo. C’est-à-dire estimer l’intégrale sur le spectre solaire, sur les directions venant du demi-hémisphère haut (ou bas pour le flux montant) et sur la surface horizontale d’altitude z_0 , de la luminance par Monte Carlo. Le problème va se modéliser avec :

- Une source d’émission de photon unique, qui est le soleil, dont on connaît la luminance en fonction de la longueur d’onde et qui a une position fixe déterminée par ses angles zenithal et azimutal. C’est une condition aux limites de l’équation de transfert : la luminance au sommet de l’atmosphère est connue : dans la direction du soleil, c’est la luminance solaire, dans les autres directions, c’est zéro.

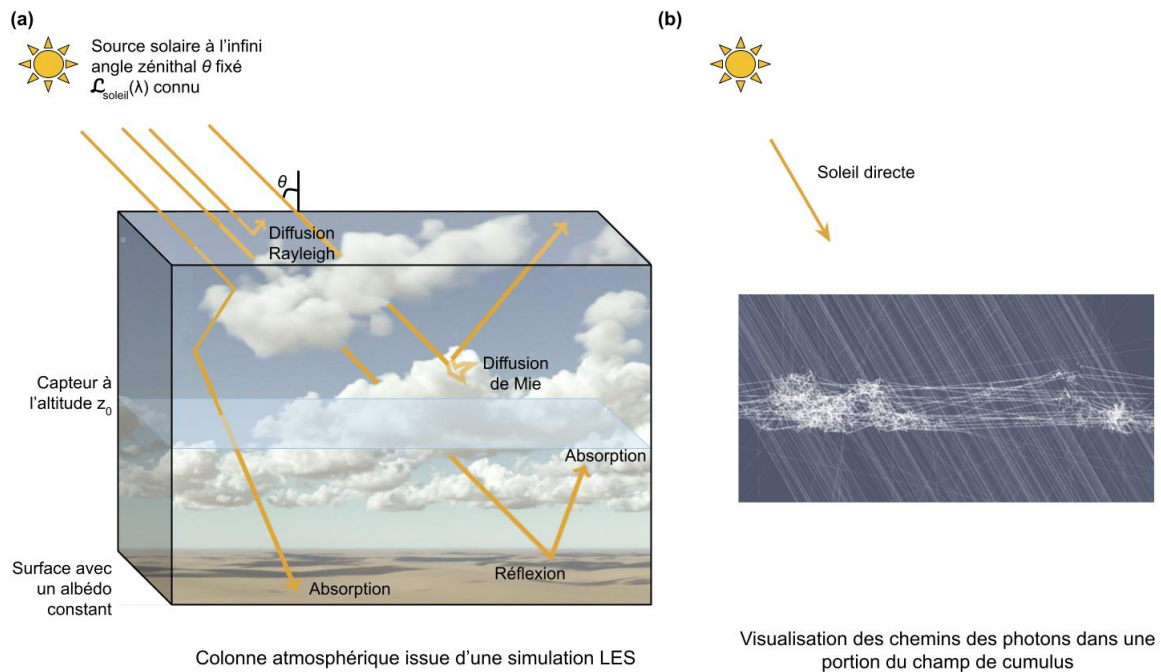


FIGURE VI.14 – Schéma illustrant l'algorithme analogue de Monte Carlo. (a) Quelques chemins parcourus par les photons dans la colonne atmosphérique sont schématisés ici. Un capteur transparent chargé de compter les photons arrivant sur sa surface (depuis le haut ou le bas) est placé à l'altitude z_0 . L'image utilisée est un rendu photo réaliste produit par htrdr sur une scène de la simulation LES du cas ARMCU par Mésio-NH. (b) Visualisation des chemins des photons (en blanc et vue de face) effectivement parcourus lors d'un calcul de transfert radiatif par Monte Carlo avec htrdr dans une portion de champ de cumulus. Les nuages sont repérables aux endroits où les photons sont fortement diffusés (diffusion de Mie) et changent ainsi souvent de direction sur de courtes distances.

- Une colonne atmosphérique avec un certain maillage contenant la description des propriétés optiques du milieu, dérivées de la simulation LES
- La surface, caractérisée dans ce calcul par un albédo constant dans l'espace, sur le spectre et les directions, et une fonction de réflectivité lambertienne
- Un capteur virtuel placé à la position où l'on souhaite calculer le flux (ici à l'altitude z_0), normal à la demi-sphère sur laquelle on veut intégrer le flux, (ici vers le haut).

Un algorithme de Monte Carlo analogue à la formulation intégrale du problème pourra être le suivant :

- Etape 1 : Un photon émis par le soleil arrive en haut de la colonne atmosphérique. La longueur d'onde λ du photon est tirée suivant la loi choisie pour décrire la luminance du soleil en fonction de la longueur d'onde (fonction de Planck à la température du soleil, spectre solaire discrétisé en fines bandes...), sa position d'entrée dans la colonne est tirée uniformément sur la surface de la colonne, et sa direction de pénétration dans l'atmosphère est connue et correspond à la position du soleil
- Etape 2 : On tire la distance l parcourue par le photon avant qu'il n'interagisse avec l'atmosphère, suivant la loi de Beer-Lambert. Si pendant son trajet, le photon intercepte :
 - le sommet de l'atmosphère : c'est qu'il est en train de monter et s'échappe dans l'espace. On arrête l'algorithme, le photon ne retouchera plus jamais le capteur
 - un côté du domaine : il est transposé de l'autre côté du domaine (conditions aux limites périodiques sur les bords du domaine)

- la surface : on tire un nombre aléatoire uniforme pour savoir s'il va être réfléchi ou absorbé par la surface, la probabilité de réflexion étant égale à l'albédo de la surface. S'il est absorbé on arrête l'algorithme, s'il est réfléchi on continue en changeant sa direction de propagation, que l'on tire aléatoirement suivant la fonction de distribution des directions réfléchies par la surface (BRDF)
- le capteur virtuel, en arrivant du haut : on compte ce photon (on ajoute 1 au compteur) et on continue l'algorithme
- Etape 3 : Le photon rencontre une particule (molécule de gaz ou gouttelette d'eau nuageuse). Un tirage est effectué pour savoir si le photon sera diffusé par la particule ou absorbé, connaissant la probabilité d'être absorbé (resp., diffusé) qui est au pro-rata du coefficient d'absorption (resp., de diffusion) sur le coefficient d'extinction total.
 - Si le photon est absorbé alors son chemin s'arrête ici et on arrête l'algorithme
 - Si le photon est diffusé, alors on effectue un tirage pour savoir quel type de particule le photon a rencontré, en fonction de leur concentration relative locale (à l'endroit de la diffusion). Puis la direction dans laquelle il est diffusé est tirée suivant la loi de probabilité angulaire ou fonction de phase, qui dépend du type de particule rencontré (diffusion de Rayleigh pour les gaz, diffusion de Mie si c'est une goutte d'eau liquide), de la longueur d'onde du photon et de l'angle incident du photon sur la particule. Puis l'algorithme est repris à partir de l'étape 2, avec cette nouvelle direction.

Cet algorithme décrit le chemin parcouru par un photon dans le milieu atmosphérique, depuis sa source jusqu'à son capteur. Cet algorithme est répété n fois, décrivant le chemin de n photons et comptant à chaque fois qu'un photon passant par le capteur (un même photon peut passer plusieurs fois). Le flux radiatif hémisphérique est alors égal à

$$\text{Constante solaire} \times \frac{\text{Nombre de passage de photon par le capteur au total}}{\text{Nombre de photons simulés} = n}$$

. L'incertitude de l'estimation décroît en \sqrt{n} . On voit que si l'atmosphère est transparente, alors tous les photons passent exactement une fois par le capteur, et le flux radiatif hémisphérique descendant sur le capteur sera alors exactement la constante solaire.

Les images physiquement réalistes de LES sont réalisées avec le code htrdr (Villefranche et al., 2019), dont le principe est le même que celui des calculs de références mais dont l'implémentation est différente. htrdr peut faire des calculs Monte Carlo dans le spectre solaire et dans le spectre infrarouge. La recherche concernant les méthodes de Monte Carlo concerne à la fois la formulation intégrale des équations qui permet de travailler formellement sur les algorithmes (en modifiant par exemple les lois de tirages pour améliorer la convergence des estimateurs), mais aussi toutes les questions d'implémentation permettant de réduire le coût de calcul. La nature très hétérogène de l'atmosphère en présence de nuages a nécessité des développements spécifiques pour permettre aux calculs d'être réalisables même sur des données de champs nuageux extrêmement détaillées (de l'ordre du milliard de mailles pour décrire un champ). Le résultat de ce travail est que le temps associé au suivi des photons est insensible à la complexité (détail, raffinement) de la donnée d'entrée.

L'algorithme de Monte Carlo fournit donc une estimation non biaisée (étant donné le modèle et les données décrivant le système) de la luminance ou du flux radiatif dont l'incertitude est connue grâce au théorème central limite. Typiquement, avec cent mille photons simulés, on a une estimation du flux solaire incident en surface sous des cumulus avec une incertitude autour de un pour mille (0.1%). Pour les calculs de référence, on simule plutôt 10 millions de trajectoires, pour avoir une incertitude autour de un pour dix mille (0.01%). L'incertitude est donc très faible relativement au flux total (plusieurs centaines de $\text{W}\cdot\text{m}^{-2}$), mais quand on regarde par exemple

des différences entre flux 3D et 1D estimés par Monte Carlo, la différence des deux peut être petite, de quelques $\text{W}\cdot\text{m}^{-2}$ seulement, mais les incertitudes des estimateurs se cumulent donc il est nécessaire que chaque estimateur soit bien convergé pour avoir des estimations fiables de leurs différences.

VI.1.2.2 Utiliser les calculs de référence Monte Carlo pour évaluer et tuner les modèles de transfert radiatif de ecRad

Villefranche et al. (2021) propose une expérience de tuning des paramètres libres du solveur SPARTACUS de ecRad en visant les références radiatives Monte Carlo sur une ensemble de scènes issues de simulations LES de cumulus continentaux (cas ARMCU et SCMs) et océanique (cas RICO et BOMEX). Cette expérience repose sur la comparaison des flux radiatifs calculés par SPARTACUS, tourné en mode diagnostique sur un profil vertical issu des scènes tridimensionnelles simulées par la LES, et d'un calcul par Monte Carlo réalisé directement sur le champ tridimensionnel de la LES avec 10 millions de tirage. Les flux radiatifs sont comparés en ayant vu le même profil moyen nuageux. Cette expérience ajuste les paramètres libres de SPARTACUS de manière à ce qu'il soit capable de simuler les bons flux radiatifs si le profil nuageux qui lui est fourni est correct. La mise en place de cette comparaison à nécessité de traiter la question de la conversion des champs et profils de LES en propriétés optiques pour le rayonnement ; et celle de la conversion des champs tridimensionnels de LES en profils verticaux équivalents.

Dériver les propriétés optiques des champs issus de simulations LES

Dans les deux cas, le milieu atmosphérique simulé par la LES doit être converti en propriété optiques, c'est-à-dire en champs ou en profil spectraux de coefficient d'absorption, de diffusion et de fonction de phase. Pour la partie ciel-clair, ce sont les k-distributions du modèle RRTMG-IFS (Iacono et al., 2008) inclus dans ecRad qui sont utilisées, à la fois pour le calcul Monte Carlo et pour le calcul avec ecRad. La question des aérosols n'est pas abordée : toutes les simulations sont faites sans aérosols (bien que les différents modèles radiatifs que l'on utilise soient prêts à prendre en compte les aérosols). Pour la partie nuageuses, seule la question des nuages d'eau liquide est traitée car les cas de cumulus étudiés ne sont constitués que d'eau liquide.

Pour le calcul de référence par Monte Carlo, une description des propriétés optiques du milieu peut être déduite de la théorie de Mie. On fait l'hypothèse que les gouttes sont sphériques et uniquement constituées d'eau pure. Connaissant la distribution en taille des rayons des gouttes, les sections efficaces pour l'absorption et pour la diffusion peuvent être calculées précisément par la théorie de Mie en fonction de la longueur d'onde. Ces sections efficaces sont multipliées par la masse d'eau fournie par la LES pour ensuite calculer les coefficients d'absorption et de diffusion. Notons que le calcul des sections efficaces nécessite de connaître la distribution en taille des rayons des gouttelettes nuageuses, une donnée qui n'est pas fournie par les LES utilisées. Une première hypothèse doit donc être formulée ici. On tabule donc des données de Mie pour une hypothèse de distribution de taille lognormale, paramétrée par un rayon effectif des gouttes (le rapport des moments d'ordres 3 et 2 des distributions de rayon des gouttes d'eau liquide, représentant un rayon effectif des gouttes pour le transfert radiatif) égale à $10\text{ }\mu\text{m}$.

Dans ecRad, on a deux options pour calculer les propriétés optiques d'un profil atmosphérique nuageux : soit à partir des données optiques du modèle SOCRATES (Manners et al., 2017), assez grossièrement discrétisées sur le spectre ; soit à partir de tables de Mie plus fines spectralement. Les sections efficaces pré-tabulées dans ces fichiers, en fonction du spectre et des rayons effectifs, sont de nouveau multipliés par les rapport de mélange en eau liquide fournies en entrée d'ecRad. De la même manière que pour le calcul de référence, une hypothèse sur la distribution de taille des gouttes est sous-jacente à la création des tables de données optiques. Ici, on ne contrôle pas l'hypothèse qui a été faite pour créer ces données.

Ensuite, dans les modèles à deux flux, deux approximations sont faites sur la diffusion du rayonnement solaire par l'eau liquide. Tout d'abord, la fonction de phase n'est pas décrite dans son entier, mais c'est son facteur d'asymétrie qui est utilisé pour la caractériser et en déduire la proportion du flux rétrodiffusé par la couche. Ensuite, l'approximation de δ -Eddington (Joseph et al., 1976) est utilisée afin de corriger l'approximation précédente en rajoutant la partie du rayonnement diffusé dans un petit angle solide autour de la direction du flux directe dans le flux solaire direct (donc transporté vers l'avant). Ces deux approximations sont inhérentes au calcul du transfert radiatif par méthode à deux flux.

Toutes les approximations classiquement faites par ecRad pour passer d'un profil atmosphérique à un profil de propriétés optiques peuvent être utilisées pour transformer le champ 3D d'une LES en propriétés optiques pour le calcul du transfert radiatif de référence par Monte Carlo. La cohérence des approximations faites pour les deux transformations dépend de ce que l'on souhaite comparer, évaluer et/ou tuner. Les choix faits dans l'étude sont les suivants :

- le calcul de référence utilise une description détaillée de la fonction de phase, alors que le calcul fait avec ecRad utilise une description utilisant un facteur d'asymétrie et l'approximation de δ -Eddington. Ces hypothèses sont considérées comme intrinsèques aux schémas radiatifs à deux flux, et on souhaite les évaluer lors de la comparaison et du tuning.
- le rayon effectif des gouttes est fixé à 10 μm partout dans les deux calculs, ce qui est vu comme une donnée d'entrée pour les modèles de transfert et n'ont donc pas de raison d'être différents d'un modèle à l'autre. Le modèle utilisé pour dériver les propriétés optiques de gouttes d'eau liquide est celui de Mie dans les deux cas, mais les données utilisées ne sont pas exactement les mêmes. La différence entre les deux a un impact très faible sur les flux calculés, et fait partie de l'incertitude inhérente aux paramétrisations radiatives qui sont souvent livrées avec leurs données de propriétés optiques dont on ne contrôle pas toujours la précision en tant qu'utilisateur. Le tuning de ecRad fournira donc un modèle simulant des flux radiatifs corrects à condition que les rayons effectifs soient bons, mais ne sera pas sur-calibré de façon à ce que si les hypothèses de forme de distribution de taille de goutte changent, on trouve encore des configurations du modèle qui soient acceptables.

L'influence de ces choix sur les flux radiatifs simulés dans des nuages plan parallèle homogènes sont documenté dans l'annexe de Villefranque et al. (2021).

Création des profils verticaux à partir de LES tridimensionnelles

Les profils verticaux atmosphériques utilisés sont des profils de température, pression, vapeur d'eau, eau liquide et fractions nuageuses issus des moyennes horizontales des champs simulés par la LES à chacun de ses niveaux verticaux, excepté pour les fractions nuageuses qui sont calculées comme la fraction des mailles contenant de l'eau liquide à chaque niveau. L'atmosphère n'est simulé que jusqu'à 4 km par la LES. Les profils atmosphériques utilisés au dessus du domaine de la LES sont issus des profils atmosphériques moyens des cumulus de moyenne latitude d'été de I3RC (Cahalan et al., 2005) distribués notamment avec le logiciel ecRad, dans lesquels il n'y a pas de nuage au dessus de 4 km. Ce raccordement est également fait pour le calcul des références radiatives par Monte Carlo.

Choix des configurations radiatives pour les comparaisons SCM / LES

Jusque là, on a comparé des flux radiatifs simulés par ecRad sur des profils nuageux issus de LES — on note ces flux ecRad(LES) ; avec les flux radiatifs de référence calculés sur la LES de référence, notés MC(LES). On souhaite désormais comparer les flux radiatifs simulés par ecRad sur les nuages issus de la version unicolonne du GCM, notés ecRad(SCM), avec les flux de référence MC(LES). Ce faisant, on élargit le cadre de comparaison 1D/LES existant à la

comparaison des flux radiatifs solaires avec ecRad tourné en mode diagnostique sur les profils atmosphériques simulés par le GCM. Pour l’instant, les rétro-actions du transfert radiatif sur la physique atmosphériques restent coupées. On remplace seulement les profils nuageux des LES par ceux simulés par le SCM.

Les questions de propriétés optiques soulevées dans le travail de Villefranche et al. (2021) se posent de la même manière ici, et ses choix ont été suivis. La question du traitement de l’atmosphère ciel-clair se pose également dans la comparaison ecRad(SCM) vs. MC(LES). Notre objectif ici est de traiter de la question des flux radiatif solaires associés aux nuages. Le choix est alors fait de remplacer les profils atmosphériques ciel-clair du SCM par ceux de la simulation LES utilisée pour faire le calcul de référence Monte Carlo. Dans cette comparaison, on enlève donc les différences dues aux profils de pression, température et vapeur d’eau sur les flux radiatifs. Dans le solaire, ces effets sont faibles comptent pour quelques W.m^{-2} au plus. En pratique, les profils de température, pression et vapeur d’eau issus de la LES utilisée pour les calculs de référence radiatives sont interpolés sur la grille du SCM jusqu’à 4km. Au-dessus du domaine de la LES, les profils issus de I3RC (Cahalan et al., 2005) représentant les profils ciel-clair typiques de nuages de cumulus de moyennes latitudes d’été sont utilisés à la fois pour le calcul de Monte Carlo et pour le ciel-clair de la SCM. Ainsi, les différences entre les flux MC(LES) et ecRad(SCM) viennent essentiellement des profils de contenu en eau liquide et fraction nuageuses simulées par le SCM, et de la modélisation du transfert radiatif lui-même (2 flux dans ecRad vs. Monte Carlo).

La question du traitement du ciel-clair dans l’infrarouge est plus sensible car les flux dans l’infrarouge dépendent au premier ordre de la température de l’atmosphère et des lieux d’émission et d’absorption, qui dépendent eux même fortement de la concentration en vapeur d’eau, en plus des nuages. Le découpage choisi pour le solaire ne paraît pas pertinent pour l’infrarouge. Cette question n’est pas abordée dans cette thèse et pourra faire l’objet de travaux futurs.

La question de la moyenne temporelle utilisée pour la comparaison se pose également dans la comparaison. En effet, le pas de temps de calcul de la LES est de une seconde et peut raisonnablement être considérée comme instantannée, contrairement à celui de la physique des GCM qui est autour de 15 min. On peut sortir des profils instantannés des simulations SCM mais il ne paraît pas évident que l’on veuille contraindre les flux radiatifs (et donc les nuages) du SCM à être en phase à la seconde près avec ceux du LES. On fera donc plutôt des comparaisons de flux moyennés sur une heure pour ne pas sur-contraindre le SCM.

Ce cadre de comparaison pour les flux radiatifs solaires à été mis à disposition dans l’outil de tuning htexplo. Il est désormais possible de tuner les paramètres libres de ecRad et/ou du GCM en visant les flux radiatifs solaires des calculs de référence et/ou les sorties classiques de la LES (variables météorologiques et nuageuses).

Des premières expériences exploitant ce nouveau cadre expérimental ont été réalisées pour éclairer la question des compensations d’erreurs au sein du modèle radiatif et entre les paramétrisations de rayonnement et de nuages. Cette étude a fait l’objet d’un article soumis à JAMES et présenté en Section [VI.2](#)

VI.2 Experience compensation d'erreur rayonnement - nuage

Investigating cloud–radiation error compensations stemming from climate model calibration

M. Coulon-Decorzens¹, F. Hourdin¹, N. Villefranque²

¹Laboratoire de Meteorologie Dynamique, Sorbonne Université/IPSL/CNRS, Paris, France

²Centre National de Recherches Météorologiques, Météo-France, CNRS, Toulouse, France

Key Points:

- Ensembles of perturbed-parameter simulations of single column models are used to investigate cloud–radiation compensating errors
- We find that structural errors in radiative transfer models due to lack of 3D effects are compensated by overestimated cloud fractions
- Radiative transfer models with and without 3D effects should not be used with the same overlap and heterogeneity parameters

Corresponding author: M. Coulon-Decorzens, maelle.coulon-decorzens@lmd.ipsl.fr

Abstract

Compensating errors are an obstacle to the development of climate models. We wonder if systematic errors in simulated cloud properties might result from forced error compensation due to targetting top-of-atmosphere radiative fluxes in the tuning process while using an inaccurate radiative transfer parameterization. Here, we investigate parametric and structural errors in state-of-the-art radiative transfer models and ask if these errors might be compensated by errors in cloud properties, to what extent and in which circumstances. Convection and cloud parameters of two versions of a Single-Column version of a climate Model (SCM), with and without a parameterization of 3D radiative effects of clouds, are tuned targetting reference solar fluxes calculated on Large-Eddy Simulation cloud fields. When 3D effects are neglected, accurate fluxes are obtained only at the expense of overestimated cloud fractions, compensating underestimated cloud reflectivity at low sun. Targetting mean solar-angle fluxes instead of detailed ones removes this mechanism entirely.

Plain Language Summary

Global climate models are used to understand the climate system and anticipate the consequences of global warming. They are based on a set of equations that represent the various aspects of the system. Among these, cloud physics and radiative transfer (light propagation) play a major role. The equations, called parameterizations, are based on our understanding of the physics, complexity–accuracy compromises, and include parameters that need to be calibrated against observations. Satellite-observed radiative fluxes are often targeted in this calibration process, with the risk of obtaining the right fluxes (close to the observed ones) for the wrong reasons, and in particular for the wrong cloud properties. In this study, we show that the radiative transfer models currently used in climate models present structural errors that could, through the calibration process, be balanced by introducing additional errors in cloud properties. This implies that persistent systematic biases in clouds simulated by global models could be partly due to structural errors in the radiative transfer model, rather than to inadequate modeling of cloud physics. We use a History Matching approach based on machine learning to shade light on these questions and study error compensations in details, albeit in an idealized framework.

1 Introduction

General circulation models (GCM) used for climate projections, are, like any model, imperfect representations of the climate system. Their behaviour depends on free parameters that need to be adjusted, which is achieved through calibration.

When calibrating numerical models as complex as GCMs, it is very difficult, if not impossible, to remove compensating errors. The issue of reducing compensating errors and finding ways to better characterize and control them remains a major challenge in climate modeling, one that we hope to address more effectively thanks to increased computing power and machine learning algorithms. This issue is crucial for the reliability of climate change projections.

Hourdin et al. (2017) report that a common practice in calibrating (tuning) climate models is to target observed top-of-atmosphere (TOA) radiative fluxes by adjusting parameters associated with the most uncertain processes controlling these fluxes: those related to clouds. In so doing, accurate TOA fluxes are often obtained at the expense of cloud-related compensating errors: between cloud properties and e.g. surface albedo or jet position (Hourdin et al., 2013), between low-, middle- and high-level clouds (Webb, Senior, Bony, & Morcrette, 2001; Nam, Bony, Dufresne, & Chepfer, 2012) or even be-

tween physical, optical and radiative properties of a given cloud regime (Konsta et al., 2022).

To understand these errors, and in particular those involving cloud physics and radiative transfer (RT), it is necessary to disentangle model errors stemming from each parameterization.

Cloud parameterizations typically provide vertical profiles of cloud fraction and water condensate in each atmospheric column. They result from complex non linear interactions between the various processes taken into account in the GCM. Having the “right clouds” in a GCM means simulating these profiles accurately. Additional assumptions concerning clouds can or must be made to compute RT from these profiles, such as their vertical overlap, horizontal inhomogeneity or size. These assumptions are usually made inside the RT scheme and their first-order effects on TOA fluxes are quite well known (see e.g., McKee and Cox (1974); Barker, Stephens, and Fu (1999); Várnai and Davies (1999); Shonk, Hogan, Edwards, and Mace (2010); Hogan, Fielding, Barker, Villefranque, and Schäfer (2019)): The widely used maximum overlap assumption for low clouds tends to underestimate total (vertically integrated) cloud cover, and consequently, TOA fluxes. Neglecting in-cloud optical-depth inhomogeneity systematically leads to overly reflective clouds. Assuming infinite cloud size means neglecting 3D radiative effects, which produces either too-bright or too-dim clouds depending on solar zenith angle. To remove these errors, radiative transfer models that are able to treat “cloud geometry” (overlap, heterogeneity and size) in a more sophisticated way than maximally-overlapped homogeneous slabs have been developed in the last decades (e.g., Shonk and Hogan (2008); Hogan and Shonk (2013); Hogan et al. (2019); Pincus, Barker, and Morcrette (2003)). The development of the ecRad code at ECMWF (Hogan & Bozzo, 2018), which implements these new radiative transfer models (solvers) in the same software, has remarkably facilitated the investigation of cloud-geometry effects in radiative transfer models.

A now standard way to improve the representation of atmospheric physics in GCMs is to isolate one column, working on physics parameterizations without interaction with large scale dynamics. In this approach, Large Eddy Simulations (LES, i. e. 3D simulations with resolution of a few tens of meters on domains of a few tens of km) of a cloud scene are used as a reference for evaluation of simulations run with the Single Column Model (SCM) version of a GCM. Radiation is still often poorly represented in LES, but reference fluxes associated with the 3D cloud scenes can now be computed using offline 3D radiative codes based for instance on Monte Carlo methods (Villefranque et al., 2019). In recent years, the SCM/LES approach has also been empowered by machine learning and automatic tuning procedures (Couvreur et al., 2021) based on the history matching approach proposed by Williamson et al. (2013). This approach performs global sensitivity experiments to separate parametric from structural errors in the model, thus providing new perspectives to the long-lasting issue of compensating errors. This framework has led to significant advances in the parameterizations of boundary layer convection and associated cumulus and stratocumulus clouds (see e.g. Hourdin et al. (2019)). It is at the heart of the hierarchical tuning process promoted by Couvreur et al. (2021), Hourdin et al. (2021) and Villefranque et al. (2021), upon which the present work builds.

In Couvreur et al. (2021) and Hourdin et al. (2021), parameters of an SCM boundary-layer parameterization are tuned targetting LES cloud properties. In Villefranque et al. (2021), cloud-geometry parameters of an RT scheme are tuned by running offline radiation upon mean LES vertical profiles, targetting reference solar fluxes obtained from 3D Monte Carlo simulations on the 3D LES clouds. Here, we begin with an extension of the Villefranque et al. (2021) study, and we then present a new type of tuning experiments, specifically designed to study cloud-radiation compensations.

This work was originally motivated by the practical need to tune atmospheric radiation and clouds after the introduction of ecRad in the LMDZ GCM (Hourdin et al.,

2020), the atmospheric component of the IPSL coupled model IPSL-CM, used in particular for CMIP exercises (Boucher et al., 2020). One particular question was to activate or not the **Spartacus** solver of ecRad, which is numerically more expensive than its 1D counterpart **Tripleclouds**, but allows for the first time to account for 3D radiative effects of clouds in the atmosphere. With this question in mind, we designed cheap, idealized, column experiments, to mimic what may occur when tuning cloud parameterizations in a climate model while targeting TOA radiative fluxes observed by satellites (as is the common practice reported by Hourdin et al. (2017)). In these experiments, we focus on two kinds of compensating errors: those that occur within the RT scheme itself, and those that occur between cloud parameterizations and the RT scheme. Under the pretext of addressing our own specific questions, we also mean to show how History Matching can be used to guide model development, and in particular to evidence and investigate compensating errors.

The first part of the study follows that of Villefranque et al. (2021). History Matching is applied to the calibration of cloud-geometry parameters of an RT scheme. PPEs are produced by running the RT scheme on horizontally-averaged cloud profiles output from LES (“perfect cloud” experiments). The reference targets in the calibration exercise are 3D solar fluxes computed using Monte Carlo methods on the 3D LES cloud fields. In this approach, ecRad and MC computations thus “see” the same average cloud scenes. Three ecRad parameters relative to cloud geometry are varied in these experiments: overlap decorrelation length, degree of in-cloud inhomogeneity, and cloud size. In Villefranque et al. (2021), ecRad was run using the **Spartacus** solver. Ranges of parameters were identified that give a very good agreement with MC computations, including a good representation of the dependency of reflected radiation to solar zenith angle, which requires a representation of 3D effects. A more systematic investigation is done here, comparing tuning experiments with the **Spartacus** and **Tripleclouds** solvers.

The second part of the study investigates error compensations between radiative transfer (including the representation of cloud geometry) and cloud physics (which provides vertical profiles of temperature, cloud fraction and liquid water content). To that end, cloud-physics parameters of the SCM are tuned, as in Hourdin et al. (2021), but targeting reference radiative metrics computed on the LES rather than targeting directly the cloud fraction and thermodynamic profiles of the LES. That is, clouds are adjusted to get the right radiative fluxes, for a given version of the RT model. In these experiments, ecRad uses either **Tripleclouds** or **Spartacus**, with their free parameters set to values identified as “best” values in the “perfect cloud” ecRad tuning experiments. We find that when radiative fluxes at various solar zenith angles are simultaneously targeted in the tuning process, experiments using **Tripleclouds** produce wrong cloud profiles to compensate for errors in the RT scheme. Then, using more appropriate ecRad parameters for **Tripleclouds** (instead of best **Spartacus**), and targeting one SZA-averaged flux (instead of three SZA-dependent fluxes), we show that both clouds and radiative fluxes can be reasonably well simulated in the experiments using **Tripleclouds**.

The rest of the paper is organized as follows: Section 2 describes the data, tools and models that are used in the study; Section 3 investigates compensations that are internal to the radiation scheme; Section 4 investigates compensations that arise between clouds and radiation schemes; Section 5 presents the conclusions of the study, and points out that, beyond these specific results, this work also shows how History Matching can be used to go deeper into the understanding of climate models.

2 Data, models and tools

To implement tuning experiments, three ingredients are needed: reference data (simulated or observed) that will serve as targets for the tuning; a model, which includes free

Table 1. Cloud properties in four scenes extracted from ARMCU and RICO LES.

Case, hour	ARMCU, 8h	ARMCU, 12h	RICO, 5h	RICO, 12h
Cloud cover (%)	26.07	6.99	13.43	20.12
Maximum cloud fraction (%)	11.11	3.09	5.45	5.96
Cloud layer depth (km)	1.625	1.20	1.25	1.675
Cloud cov. / max. cloud frac.	2.346	2.26	2.462	3.379

parameters that will be adjusted in the tuning process; a tuning tool (methodology and software) that will make part of the protocol objective and automatic.

In this study, two different kinds of tuning experiments are implemented. Both use the same tuning tool, described in Section 2.4, and the same reference data: radiative fluxes computed using a 3D Monte Carlo code, run on 3D cloud fields output from LES. These are described in Section 2.1. The difference between the two types of experiments is the model that is tuned. In the first type of experiments, the models are the **Tripleclouds** and **Spartacus** solvers of the radiative transfer scheme ecRad, run on horizontally-averaged LES cloud profiles. In the second type of experiments, the model is the SCM version of LMDZ. They are respectively described in Section 2.2 and Section 2.3. All experiments are made using idealized cumulus cases, following Villefranche et al. (2019).

2.1 Reference data

Two typical cumulus cloud cases are used from the set of idealized cases that are distributed in a standardized format by Dephy (Développement et Evaluation PHYsiques des modèles atmosphériques, Dephy (2020)). The ARMCU case (Brown et al., 2002) is typical of the development of boundary-layer clouds over continent during the day, while the RICO case (vanZanten et al., 2011) is typical of trade-wind cumulus developing over a stationary ocean. LES of these two cases are run with the Meso-NH model (Lafore et al., 1998; Lac et al., 2018) at 25 metres horizontal and vertical resolutions on a $12.8 \times 12.8 \times 4 \text{ km}^3$ domain. Large scale dynamics, radiative heating and surface conditions are imposed. These simulations provide reference values for the thermodynamic and cloud variables, and their uncertainties are quantified running sensitivity experiments to numerical and physics options as described in Couvreux et al. (2021). The reference simulations are presented in Figure 1. Four scenes from these two simulations will be used as constraints in the tuning experiments. They are documented in Table 1.

Reference solar fluxes are computed using a 3D Monte Carlo (MC) code run on 3D cloud fields extracted every hour from the LES, as described in Villefranche et al. (2019). 3D fields of liquid water content are taken from the LES and cloud-droplet effective radius is homogeneously set to $10 \mu\text{m}$. Cloud optical properties are obtained from Mie theory. Gas optical properties are calculated using the k-distribution model RRTMG-IFS included in the ecRad radiation scheme (Hogan & Bozzo, 2018), for temperature, pressure and humidity profiles corresponding to the LES horizontal mean below 4 km, and to Standard Mid-Latitude Summer profile above. The solar constant is set to 1368 W.m^{-2} and the surface albedo to 0.08. For each 3D cloud field, additional MC calculations are made under the Independent Columns Assumption (ICA), by computing 3D radiation independently in distinct columns of the 3D LES cloud field and taking the average. This assumes independent columns, as in the MCICA model of Pincus et al. (2003), which removes 3D radiative effects from the calculation. Differences between 3D and ICA MC fluxes yield estimates of 3D radiative effects.

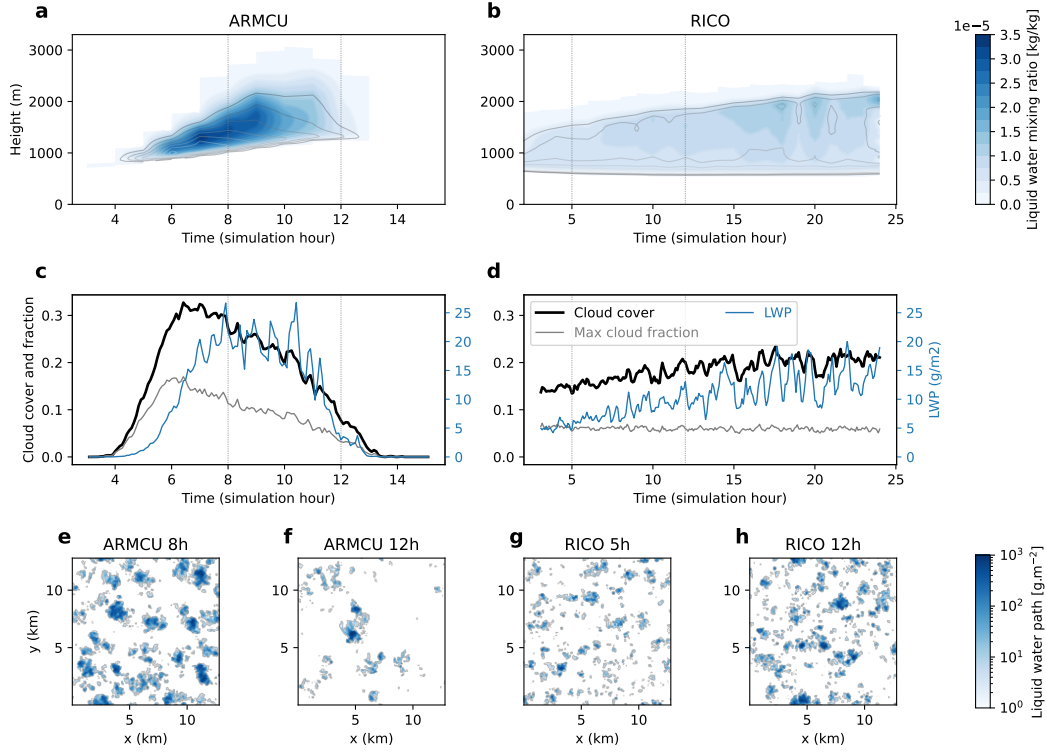


Figure 1. Reference simulations of the ARM CU (left) and RICO (right) cases used in this study. All variables are computed from LES 3D cloud fields. a-b: time evolution of vertical profiles of cloud fraction (contour lines, 0 to 0.16 every 0.02 in ARM CU, 0 to 0.07 every 0.01 in RICO) and cloud water mixing ratio (blue shading). c-d: time evolution of total (vertically integrated) cloud cover, maximum cloud fraction and domain-mean liquid water path. e-h: liquid water path maps at the two hours retained for tuning (ARM CU 8th and 12th hours, RICO 5th and 12th hours, identified by vertical lines in a-d plots).

2.2 Radiation parameterization

The radiative models under investigation in this study are the **Tripleclouds** and **Spartacus** solvers implemented in ecRad, the radiative transfer model developed at European Centre for Medium-Range Weather Forecasts (Hogan & Bozzo, 2018). ecRad provides a flexible interface that allows users to configure various aspects of the radiation model. Cloud droplet effective radius, gas optics, clear-sky profiles (gas concentrations, temperature and pressure) and radiative boundary conditions are set as in the MC simulations so that they are excluded from causes of possible differences between parameterized and reference fluxes. In “perfect clouds” experiments, input liquid water content and cloud fraction profiles are taken from horizontally averaged LES 3D fields and are hence also excluded from potential causes of differences between parameterized and reference fluxes. In SCM experiments, liquid water content and cloud fraction profiles input to ecRad are taken from SCM simulation outputs.

Cloud optics are interpolated from a Mie look-up table provided with ecRad, similar to, but slightly different from, the one used in the Monte Carlo simulations. The difference between optical properties taken from the two tables are typically less than a change of $1\ \mu\text{m}$ in the effective radius of clouds (not shown).

The RT models at the heart of **Tripleclouds** and **Spartacus** are modified versions of the two-stream model (Meador & Weaver, 1980), which directly integrate the effects of cloud geometry on radiation transport through assumptions on vertical overlap, horizontal heterogeneity and, in **Spartacus** only, cloud size.

In our configurations of both **Tripleclouds** and **Spartacus**, vertical overlap is represented using the exponential-random model parameterized by its decorrelation length ℓ (Hogan & Illingworth, 2000), and a two-region cloud representation (the **Tripleclouds** model of Shonk and Hogan (2008)) is used to account for in-cloud water sub-grid heterogeneity, whereby layer-wise optical depths in thin-cloud and thick-cloud regions are calculated according to the fractional standard deviation (*FSD*) parameter. In the **Tripleclouds** solver, no 3D effects are taken into account, whereas in **Spartacus** (Hogan, Schäfer, Klinger, Chiu, & Mayer, 2016; Schäfer, Hogan, Klinger, Chiu, & Mayer, 2016; Hogan et al., 2019), intensity of 3D effects is proportional to cloud-side perimeter length (Hogan & Shonk, 2013), itself a function of cloud fraction and cloud effective scale (C_s). It was shown previously that 3D effects in cumulus clouds remove around $10\ \text{W}\cdot\text{m}^{-2}$ from TOA solar upward (reflected) flux when the sun is high (solar zenith angle close to zero), and account for about $10\ \text{W}\cdot\text{m}^{-2}$ at low suns, compared to 1D radiation (see e.g. Villefranque et al. (2019)).

2.3 LMDZ Single-Column Model

LMDZ-6A (Hourdin et al., 2020) is the atmospheric component of the IPSL-6A General Circulation Model, which participated in the sixth phase of the Coupled Model Intercomparison Project (CMIP6). Here, its single-column version is used with a refined 95-level grid as in Hourdin et al. (2019, 2021) to simulate ARMCU and RICO cases. The same large-scale dynamics, radiative trends and surface conditions are imposed as in the LES so that physical parameterizations are the only active part of the model.

More specifically, the parameterizations that are active here are the boundary-layer transport and cloud schemes. The parameterization of vertical sub-grid transport is based on an Eddy-Diffusivity and Mass-Flux approach. The Eddy-Diffusivity model parameterizes the effects of small-scale turbulence on the mean state using the Turbulent Kinetic Energy prognostic equation formulated by Yamada (1983) with a 1.5-order closure. The Mass-Flux model parameterizes the effects of organized convective cells or rolls on the mean state using an effective thermal plume model. The plume transports air and state variables from the surface to the boundary-layer top. Exchanges with the environ-

ment are modeled through lateral entrainment and detrainment formulations (Hourdin et al., 2019). Water condensate and cloud fraction profiles are computed using a bi-Gaussian probability density function of the saturation deficit, with one mode accounting for saturation deficit in the thermal plume and one mode in the environment (Jam, Hourdin, Rio, & Couvreur, 2013). This combination of Eddy-Diffusivity-Mass-Flux scheme with a bi-Gaussian cloud scheme provides a unified framework that has been shown to accurately represent both dry and cloudy convective boundary layers with cloud regimes ranging from cumulus to stratocumulus (Hourdin et al., 2019). The conversion of cloud water into precipitation and the evaporation of precipitation are detailed in Madeleine et al. (2020).

As for radiation, ecRad was recently implemented in LMDZ and will be part of the forthcoming versions of the GCM. However, in this study, it is the offline version of ecRad that is run on the SCM output profiles. It calculates the radiative fluxes associated with the SCM cloud profiles and ecRad cloud-geometry parameters. The offline version was used to make more robust comparison, focused on cloud radiative effects. In particular it allows to use the same clear-sky profiles as in the MC simulation. This was done in practice by replacing the clear-sky profiles in the SCM outputs by the profiles of the LES averaged horizontally. Note that in the two cumulus cases that are used as constraints, surface fluxes and radiative cooling are prescribed as forcings: radiation is not interactive in the simulations and thus does not affect the clouds (consistently in the LES and SCM).

2.4 The High-Tune:Explorer tuning tool

High-Tune:Explorer is a tuning tool based on History Matching with iterative re-focusing (Vernon, Goldstein, & Bower, 2010; Williamson et al., 2013). It aims at finding the subspace of model free parameters that match a set of constraints. To that end, the parameter space hypercube, $[\lambda_{min}^1, \lambda_{max}^1] \times [\lambda_{min}^2, \lambda_{max}^2] \times \dots \times [\lambda_{min}^N, \lambda_{max}^N]$, with $\lambda^1, \dots, \lambda^N$ the N free parameters to tune, is iteratively reduced by ruling out parameter vectors for which the model's predictions, for a set of user-defined metrics, do not match reference values within the range of user-defined tolerance to error.

To accelerate the exploration of the hypercube, emulators based on Gaussian Processes are built for each metric. These emulators are trained on metrics computed from an ensemble of model runs (with typically $10 \times N$ members), to then provide rapid predictions of metric values for huge sets of free parameter vectors¹. Prediction uncertainty is added to tolerance-to-error in order to avoid ruling out parameter vectors that might be far from the target because of bad emulators but actually acceptable configurations of the model. To this end, implausibility is defined for any parameter vector λ as

$$I(\lambda) = \max \left\{ \frac{|r_1 - \mu_1(\lambda)|}{\sqrt{\sigma_1^2(\lambda) + T_1^2}}, \dots, \frac{|r_p - \mu_p(\lambda)|}{\sqrt{\sigma_p^2(\lambda) + T_p^2}} \right\}, \quad (1)$$

with r_i the reference (target) value and T_i its tolerance to error.

The parameter vector λ is ruled out if its implausibility $I(\lambda)$ is greater than an arbitrary value Γ , which represents the size of the confidence interval (reference $\pm \Gamma$ times

¹ Training an emulator means characterizing the law of the Gaussian Process $G \sim \mathcal{GP}(m, k(.,.))$, that is, specifying its expectation function m and variance kernel $k(.,.)$. This is done by fitting standard analytic functions to the simulated data for metric i , $\{f_i^j\}_{j \in 1, \dots, 10N}$ obtained from model runs at $10N$ points $\{\lambda^j\}_{j \in 1, \dots, 10N}$ sampled in parameter space. Then, for any new λ^* in parameter space, the i -th emulator provides a prediction of the i -th metric as the expectation μ_i and standard deviation σ_i of the conditional random variable $G_i^* = G_i(\lambda^*) | \{G_i(\lambda^j) = f_i^j, \forall j \in [1, 10N]\}$; and $G_i^* \sim \mathcal{N}(\mu_i, \sigma_i)$.

uncertainty), typically chosen between 2 and 3. At the end of each iteration, the new Not-Ruled-Out-Yet (NROY) space of parameters is determined using this implausibility condition. The next iteration starts by sampling a set of parameter vectors in the NROY space of the previous iteration. Then a new ensemble is run, metrics are evaluated, emulators are built, etc.

As the iterative process progresses, the NROY space narrows down, mostly because emulators uncertainty decreases, which is due to denser information being collected for training (same amount of points in a smaller NROY space). The tuning experiment is considered to have strictly converged when emulator uncertainties are significantly smaller than tolerances to error for every metrics. In that case, the final NROY space is exactly the subspace of free parameters that matches the user-defined constraints and emulators can be considered reliable models for the metrics (inside the final NROY space).

In practice, emulator uncertainties rarely fall one order of magnitude under tolerances to error for all metrics and hence the experiment rarely strictly converges. We therefore find it useful to define weak convergence: The experiment is considered to have weakly converged when adding new iterations does no longer significantly reduce NROY space. In that case, the final NROY space is larger than the sought parameter subspace.

Importantly, during the tuning experiment, many simulations will have been run, constituting a Perturbed-Parameter Ensemble (PPE). For each metric i and each simulation j , its distance to reference value can be calculated as $\frac{|r_i - f_i^j|}{T_i}$, where $f_i^j = f_i(\boldsymbol{\lambda}^j)$ is the actual model output for metric i and parameter vector $\boldsymbol{\lambda}^j$. The score of a simulation j is then defined as its worst-metric score:

$$S(\boldsymbol{\lambda}^j) = \max \left\{ \frac{|r_1 - f_1(\boldsymbol{\lambda}^j)|}{T_1}, \dots, \frac{|r_p - f_p(\boldsymbol{\lambda}^j)|}{T_p} \right\}, \quad (2)$$

where emulator predictions and uncertainties no longer appear, contrary to the implausibility defined in Equation (1). Keeping only simulations whose scores are under a given threshold provides, independently from the emulators convergence status, constrained PPEs that can be studied to learn about the quality of the model. Of course, if the emulators convergence is bad, the NROY space might remain too large and model configurations sample might not be dense enough, so low-probability, acceptable model behaviours might be absent from the PPEs.

3 Internal compensations between parts of the radiative transfer model

First, we examine ecRad PPEs in a “perfect cloud” framework. “Perfect clouds” means that ecRad is run on horizontally-averaged LES cloud profiles. The 3D cloud fields from the same LES were used as inputs for the reference Monte Carlo radiative simulations. Errors in ecRad fluxes can thus not be attributed to wrong cloud fractions or mean liquid contents, unlike in cases where cloud fraction and liquid content are taken from boundary-layer parameterizations of a climate model, as discussed in the following section. We use these PPEs to investigate compensating effects between various aspects of cloud geometry. In **Tripleclouds** and **Spartacus**, the effect of cloud geometry on radiation (i.e. the fact that clouds are not horizontally infinite homogeneous slabs isolated from their environment) is taken into account by adding state variables and source terms to the two-stream model of Meador and Weaver (1980). Cloud geometry includes vertical overlap, heterogeneity of in-cloud water, and cloud size. Free parameters associated with these new terms are overlap decorrelation length ℓ , fractional standard deviation FSD , and cloud effective scale C_s . In the present “perfect cloud” experiment, they are the only parameters that are calibrated in order to obtain fluxes as close as possible as those computed using Monte Carlo methods on the 3D cloud fields.

Here, we look for the parameter values that provide an accurate description of ARMCU and RICO cumulus cloud geometry *from a radiative transfer model perspective*. This approach is fundamentally different from trying to extract estimates of these parameter values directly from the simulated 3D cloud fields. Here, we accept that cloud geometry parameters are *effective parameters* that make sense *in the context of a particular model* (here the RT solver), rather than in the absolute.

3.1 Where we find that Tripleclouds and Spartacus should not use the same cloud-geometry parameter values

First, we look at two experiments: one where **Tripleclouds** targets Monte Carlo 1D (ICA) fluxes and one where **Spartacus** targets Monte Carlo 3D fluxes. Using 1D MC fluxes as targets for **Tripleclouds** avoids forcing the model to compensate for a structural error that we know is present in the solver, namely the absence of 3D radiative effects. Comparing 1D MC computation is a classical way of evaluating two-stream RT models that do not take 3D effects into account (e.g., Barker et al. (1999)).

The MC fluxes targeted as tuning metrics are taken at three solar zenith angles. Simulation ensembles are produced varying the three radiative parameters (ℓ , FSD and C_s , only the first two being active in **Tripleclouds**) under the constraint that targets are reached within three times the tolerance to error, which is set to $3.5 \text{ W}\cdot\text{m}^{-2}$.

For each of these two experiments, ten waves (iterations) of history matching are produced, each consisting in 101 simulations. In total, for each experiment, 1010 configurations of the solver (either **Tripleclouds** or **Spartacus** depending on the considered experiment) are sampled and run. Among these configurations, the 300 simulations with the lowest scores are selected. Figure 2 shows the location of these 300 best configurations per experiment in parameter space. The first three lines and columns of the subplot matrix are relative to parameter values, while the last line and column is relative to the score $S(\lambda)$ of each configuration, as per Equation (2). Since the RAD_CS parameter is not used in **Tripleclouds**, the red points corresponding to **Tripleclouds** configurations spread randomly along the RAD_CS axis.

First, we note that all configurations in the **Tripleclouds** experiment have scores below 1, while all those of **Spartacus** have scores above 1. Of course it does not mean that **Tripleclouds** is better than **Spartacus** in an absolute sense, since the scores for **Tripleclouds** are computed against referenced intentionally biased by not accounting for 3D effects. It just means that **Tripleclouds** is closer to reference 1D fluxes than **Spartacus** is close to reference 3D fluxes. It might not be too surprising for a modified two-stream model to perform better at simulating 1D radiation than 3D, especially when considering the many years spent on the 1D problem (i.e., how to account for vertical cloud structure and horizontal heterogeneity in vertical light propagation) compared to the relatively short time dedicated to modelling horizontal transport within a the two-stream framework.

The next thing one can notice is that the 10 best configurations of each experiment are located in distinct regions of the $\ell \times FSD$ space. The best **Tripleclouds** simulations, highlighted by large red symbols in Figure 2, have small ℓ and large FSD , while the best **Spartacus** simulations (large blue symbols) correspond to larger ℓ values, for approximately the same range of FSD values. The best **Tripleclouds** simulations exhibit a strong negative correlation between ℓ and FSD .

In 1D radiation, this is easy enough to understand: with a smaller decorrelation length, clouds are more randomly (or less maximally) overlapped, leading to a larger cloud cover and smaller cloud optical thickness in regions covered by clouds. The radiative effect of increasing cloud cover dominates that of decreasing cloud optical thickness, be-

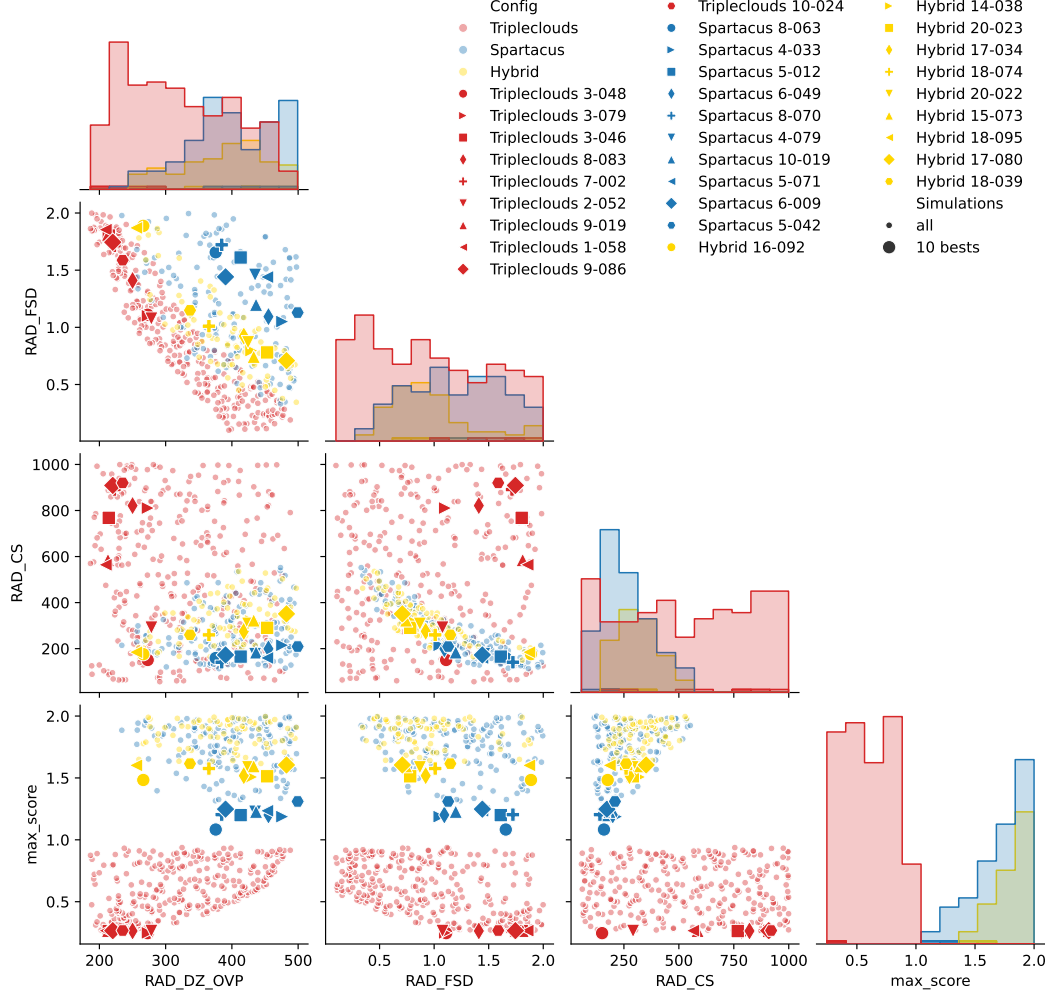


Figure 2. Parameters of 300 best simulations for three “perfect-cloud” ecRad tuning experiments: Tripleclouds (red), Spartacus (blue) and Hybrid (gold). RAD_DZ_OVP is the vertical overlap decorrelation length parameter ℓ , RAD_FSD is the relative in-cloud inhomogeneity parameter FSD , RAD_CS is the cloud size parameter C_s . The max.score is $S(\lambda)$. Each column corresponds to one parameter. The top subplot of each column shows histograms of the parameter values for the 300 best configurations, for each experiment. Then, each line corresponds to a second parameter, and subplots show the location of the configurations in the 2D space of the column \times line parameters.

cause reflectivity is an exponential function of optical thickness τ ,

$$R \sim 1 - \exp(-\tau).$$

The optical thickness at a given location on the horizontal is the sum of the optical thicknesses if the cloudy layers above. Distributing the individual optical thicknesses of the cloud layers horizontally on a wider area (large cloud cover with clouds of moderate optical thickness) leads to a larger reflectivity than when cloudy layers overlap, because of the saturation of the exponential term.

The same reasoning can be applied to in-cloud water horizontal heterogeneity: because of the exponential function, horizontally homogeneous clouds are more reflective than heterogeneous clouds that have the same horizontally-averaged optical depth. Thus, decreasing ℓ leads to more reflective clouds, and increasing FSD leads to less reflective clouds. This explains why the tuning algorithm selects smaller values of the FSD to compensate for larger values of ℓ in order to simulate the right fluxes.

In **Spartacus**, the relationship between ℓ and FSD also seems to exist, but the 10 best configurations of the **Spartacus** experiment show more dispersion than the 10 best **Tripleclouds** ones. Our first guess was that horizontal transport, driven by the third parameter C_s , modulates reflectivity in **Spartacus** while it has no effect in **Tripleclouds**, which means that, in **Spartacus**, a given value of ℓ can be associated with various values of FSD and still produce the same fluxes. For instance a larger FSD could be compensated by more intense 3D effects.

However, C_s does not vary much between these 10 best configurations so this explanation might not be sufficient. Moreover, 3D effects tend to decrease reflectivity when the sun is close to zenith and increase it otherwise; so we need to consider our three metrics, which correspond to three different solar zenith angles, in order to understand if significantly larger FSD can indeed be compensated by slightly smaller clouds.

All things considered, the internal dispersion of the best **Spartacus** configurations might be a detail compared to the fact that best **Spartacus** configurations can be found in the “large ℓ , large FSD ” corner of the parameter space, where none of the best **Tripleclouds** simulations can be found. This is an important result, as it means that finding the best **Spartacus** configuration and removing 3D effects (e.g., using **Tripleclouds** with the same (ℓ, FSD) values, or setting C_s to infinity in **Spartacus**), will produce wrong fluxes compared to the 1D Monte Carlo reference. Indeed, none of these “large ℓ , large FSD ” configurations leads to 1D fluxes closer to the reference than $2 \times 3.5 = 7 \text{ W}\cdot\text{m}^{-2}$ (figure not shown).

The **Spartacus** solver was built upon the **Tripleclouds** solver by adding terms to represent 3D effects. So one could wish **Spartacus** to simulate both 3D radiation and 1D radiation accurately when 3D effects are removed. In this view, the results shown here suggest that compensation errors are at work between the **Tripleclouds** solver and 3D effects in **Spartacus**. However, when running **Spartacus**, there is no real distinction between non-3D and 3D effects. It is the full radiation which is computed and affected by the parameterization of 3D effects. Another, less stringent but probably more sensible view is to accept that free parameters are only *effective parameters* and hence that there is no objective reason to set them to (often only loosely equivalent) observed values.

To illustrate this point further, let us look at cloud covers calculated by the exponential-random overlap model, for the ten decorrelation length values considered as “best” values in **Tripleclouds** and **Spartacus** experiments. The upper panel of Figure 3 presents a comparison of true cloud covers diagnosed in the 3D fields of ARMCU and RICO LES simulations, with those calculated by the overlap scheme. Note that covers calculated in ecRad, when using **Tripleclouds** or **Spartacus** solvers, do not enter directly in the flux equations. They are only computed as a diagnosis. Rather, the overlap model is used

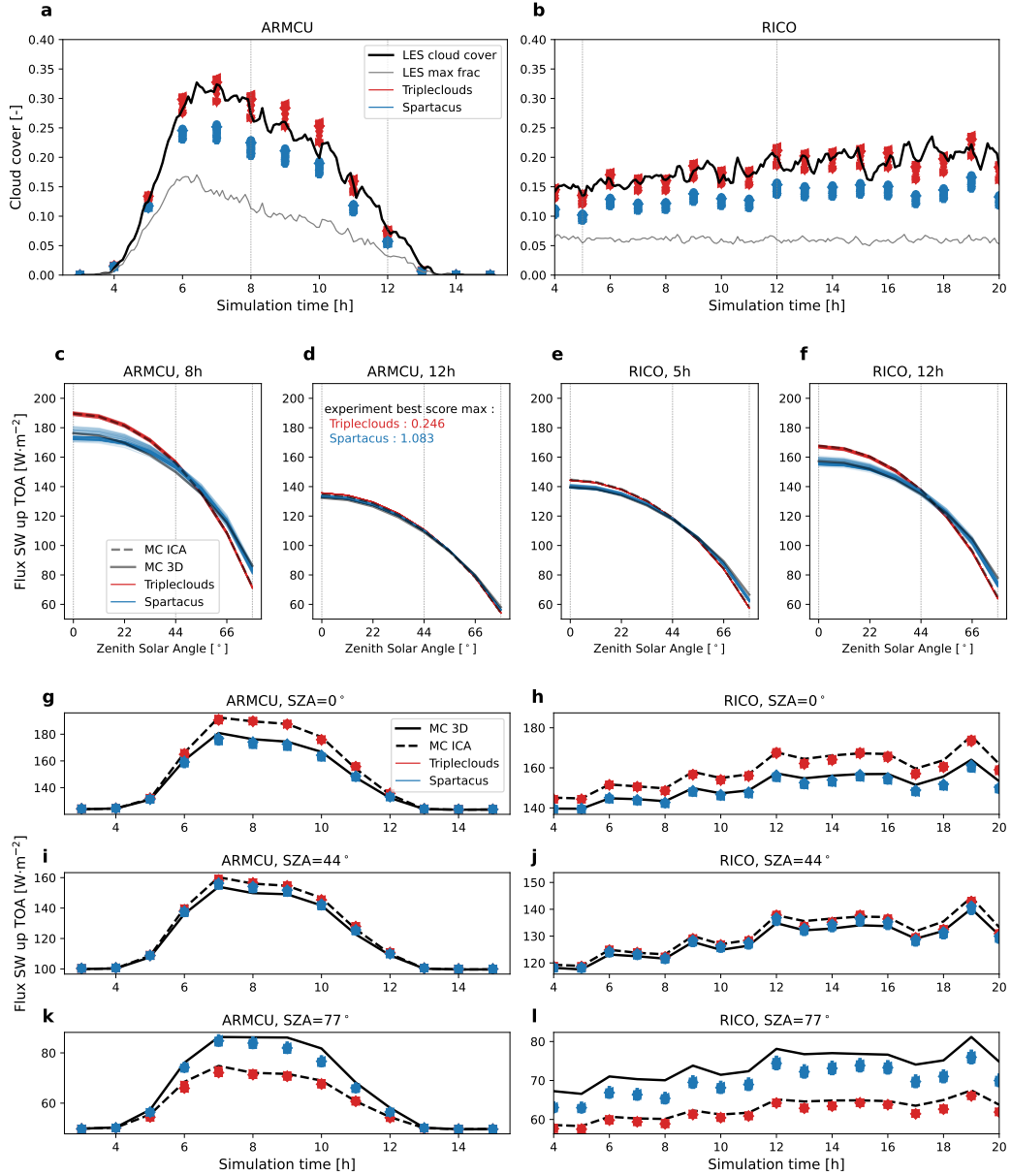


Figure 3. Cloud and radiation variables for the 10 best simulations of two perfect-cloud experiments (Tripleclouds in red, Spartacus in blue), compared to reference. Left column: ARMCU. Right column: RICO. a-b: cloud cover (black) and maximum cloud fraction (gray) in LES fields, and cloud cover computed by ecRad (color points) using input cloud fraction profiles calculated from LES cloud fields, and the exponential-random overlap model parameterized by various decorrelation lengths (RAD_DZ_OVP values in Figure 2). c-l: Fluxes simulated by Monte Carlo 3D (black full lines) and 1D (black dashed lines), Tripleclouds (red) and Spartacus (blue). c-f: upward TOA fluxes as a function of solar zenith angle, for the 4 cloud scenes used as constraints in the tuning experiments: ARMCU 8th (c) and 12th (d) hours, RICO 5th (e) and 12th (f) hours. g-l: upward TOA fluxes, calculated at every hour of the two LES simulations, each time for three solar angles. g-h: SA=0°, i-j: SA=44°, k-l: SA=77°.

locally, at each interface between layers, to distribute fluxes from cloudy and clear regions of a given layer, to cloudy and clear regions of the next layer. These plots show that covers corresponding to the best **Tripleclouds** simulations are very close to those of the LES. It is tempting to conclude that **Tripleclouds** fluxes are right *for the right reasons*, i.e., the right cloud geometry. In comparison, cloud covers corresponding to the best **Spartacus** simulations are systematically smaller than those of the LES (in agreement with the fact that ℓ values are larger). Yet, **Spartacus** fluxes are globally accurate. Is it for wrong reasons, i.e. the wrong cloud geometry? Does this fall under the “compensating error” category? Do we want to eliminate this behaviour or shall we accept, and even expect, that “cloud cover” (most often, a vertically integrated quantity), does not have the same meaning, and does not play the same role, in 1D vs. 3D radiative transfer models? These are open question that the modelers should have in mind when using, evaluating and tuning their models.

What fundamentally shows the improvement of the **Spartacus** solver compared to **Tripleclouds** is its ability to well capture the angular dependency of the “true” reflected radiation (3D MC reference) as clearly seen in the second line of panels in Figure 3. The **Tripleclouds** is able to well simulate 1D radiative reference computations but it will always underestimate the reflected radiation at high solar angle compared to zenith.

3.2 Where we look for a compromise between **Tripleclouds** and **Spartacus** preferred regions of cloud-geometry parameter space

We rather focus here on finding a configuration that is acceptable for both **Tripleclouds** and **Spartacus** models. Because best **Spartacus** configurations are bad choices for **Tripleclouds**, it implies that **Spartacus** simulations, set with a compromise configuration, will be less accurate than the previously examined best **Spartacus** simulations. To that end, we design a Hybrid tuning experiment, in which we target both 1D MC fluxes with the **Tripleclouds** solver and 3D MC fluxes with **Spartacus** at the same time. This means that for every sampled parameter vector (ℓ, FSD, C_s) , we run both **Tripleclouds** and **Spartacus** on the four clouds scenes (two hours from ARMCU et two hours from RICO) and three solar zenith angles (0, 44 and 77 degrees). The score associated with each configuration is the worst (metric-wise) score over all the **Tripleclouds** and **Spartacus** metrics.

The 300 best configurations are presented in Figure 2 (yellow points). The scores obtained are generally worse than for previous experiments (mostly, above 1.5). The location of the best configurations in the (ℓ, FSD) space best illustrates the compromise that was found: for a given choice of FSD , the 10 best simulations of the Hybrid experiment correspond to ℓ values between best **Tripleclouds** and best **Spartacus** values. Note that relatively small values of FSD of around 0.75, associated with relatively large C_s of around 400 m, yield scores that are among the ten smallest ones, which was not the case in any of the previous experiments.

The very best configuration in the Hybrid experiment corresponds to large FSD and small ℓ values, close to the best **Tripleclouds** simulations. Accordingly, in the upper panel of Figure 4, this best Hybrid configuration corresponds to the points that are closest to the 3D LES cloud cover. Comparing fluxes in Figure 4 with those of Figure 3, we can see that **Tripleclouds** fluxes, which were previously extremely accurate, are now less so: most of the ten best configurations from the Hybrid experiment lead to smaller cloud cover than in the 3D LES field, which negatively impacts **Tripleclouds** performances. As for **Spartacus**, it seems to perform better or equally at $SZA=0$, and to be slightly worse than before at large solar zenith angles. If we look at the “large ℓ , small FSD ” configurations of the Hybrid experiment, we note that, for a given ℓ value, FSD is smaller than in the pure **Spartacus** experiment, which should lead to more reflective clouds than before, if nothing else changed. However, these configurations are also as-

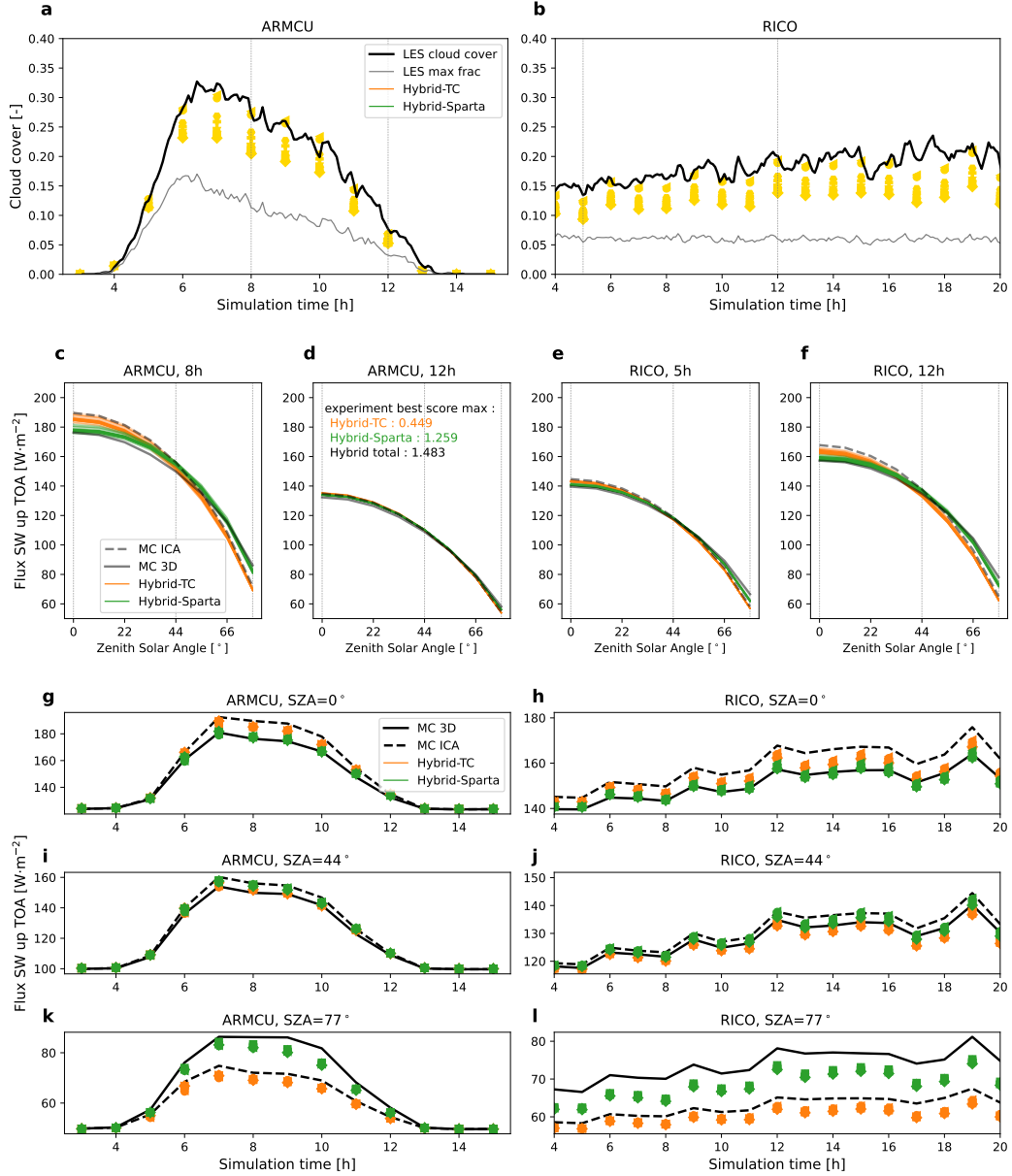


Figure 4. Cloud and radiation variables for the 10 best configurations of the Hybrid experiment, compared to reference. As in Figure 3 except that: in a-b, decorrelation length parameters used to compute cloud cover are those of the 10 best simulations from the Hybrid experiment; in c-l, *Spartacus* (green) and *Tripleclouds* (orange) are both configured with the same 10 best Hybrid configurations (instead of their own best configurations in Figure 3).

sociated with larger C_s values than the ten best of the pure **Spartacus** experiment, which means less intense 3D effects, which means more reflective at small SZA and less reflective at large SZA. The fact that smaller C_s values would not be chosen for these “large ℓ , small FSD ” configurations, despite the fact that it would systematically improve fluxes at SZA=77, seems to indicate that fluxes at SZA=0 would be too large (more intense 3D effects would decrease reflectivity at SZA=0) and would become the limiting metrics.

Finally, we find that the very best configuration of the Hybrid experiment, that is, simulation number 16-092, is a good candidate for being our “unique” ecRad configuration in LMDZ, as far as cumulus cloud geometry is concerned.

4 Compensations between radiative transfer and cloud models

Here, we examine PPEs run with the single column version of LMDZ for the same two cumulus cases, under radiative constraints. Using both **Tripleclouds** and **Spartacus** solvers, with configurations from the previous section, and testing two sets of radiative metrics, we investigate the conditions that lead to cloud-radiation compensating errors.

4.1 Experimental design

We design experiments to mimic the tuning protocol often followed in GCMs when targeting top-of-the-atmosphere (TOA) radiative metrics from satellite observations. In our idealized 1D version of the protocol, simulations with both **Spartacus** and **Tripleclouds** target the same MC 3D radiative computations (“true” radiative fluxes). Tuning is performed using either **Tripleclouds** or **Spartacus**, in order to investigate in particular how convection and cloud physics could compensate for the structural error consisting in not accounting for 3D radiative transfer in **Tripleclouds**. The metrics are hourly-mean radiative fluxes at fixed solar angles. The reference ones are averaged from MC fluxes computed every 5 min on the LES clouds. In the SCM, they are the fluxes calculated on hourly-averaged cloud profiles.

Thirteen parameters that control boundary-layer and cloud parameterizations in LMDZ are varied as in Hourdin et al. (2021) and Hourdin et al. (2023) (see details in Table S1 of Supporting Information).

Each tuning experiment consists in 30 iterations. At each iteration, 130 free-parameter vectors are sampled in the NROY space and ARMCO and RICO are simulated using these 130 configurations of LMDZ. ecRad is then run offline for the two chosen times of the two cases and each of the 130 LMDZ configurations, to compute solar reflected fluxes at three solar zenith angles each; in total, 1560 ecRad runs per iteration. Then one emulator is built for each of the 12 metrics, using their 130 evaluations as a learning database. Finally, implausible free-parameter vectors are ruled out using a threshold that varies from 3 in the first 5 iterations to 2.5 in iterations 6 to 10, to 2 in iterations 11 to 30. The NROY space is thus efficiently narrowed down.

First, the 30-iterations **Spartacus** experiment is repeated following a trial and error process that seeks the smallest tolerance-to-error value yielding non-empty NROY space without falling below $3.3 \text{ W}\cdot\text{m}^{-2}$, which is the minimum estimated structural error of **Spartacus** (see Supporting Information Text S1). The resulting tolerance to error is $3.5 \text{ W}\cdot\text{m}^{-2}$. Then, the **Tripleclouds** experiment is run once using the same tolerance to error. As this tolerance-to-error value is smaller than **Tripleclouds** structural error on the same metrics, this experiment stops after 5 iterations with empty NROY space. The tolerance to error for **Tripleclouds** is then rescaled by the minimum score found in the aborted experiment. The resulting tolerance to error is $9 \text{ W}\cdot\text{m}^{-2}$. A 30-iterations experiment is run with this new value.

Table 2. Maximum and average absolute errors in **Spartacus** (Sparta) and **Tripleclouds** (TC) fluxes for the various tuning experiments presented in Section 4

Experiment	Shown in	Max. abs. err. ($\text{W}\cdot\text{m}^{-2}$)	Ave. abs. err. ($\text{W}\cdot\text{m}^{-2}$)
Sparta-08-3ang	Fig. 5	3.6480	2.4931
TC-08-3ang	Fig. 5	14.2253	8.2945
Sparta-16-3ang	Fig. 6	3.7890	2.8047
TC-16-3ang	Fig. 6	14.2117	8.3226
Sparta-16-1ave	Fig. 7	0.4360	0.2976
TC-16-1ave	Fig. 7	1.2210	0.7972
TC-16-2ang	Fig. S1	12.6082	8.7031
TC-best	Fig. S1	11.2147	7.4093

4.2 Where we suspect compensations between cloud fraction and 3D effects

We start by comparing experiments using either **Spartacus** or **Tripleclouds**, with ecRad cloud-geometry parameters set to the values of the best simulation in the tuning of **Spartacus**, i.e. configuration 8-063 in Figure 2. The idea in this experiment is to use the best possible version of **Spartacus** as a perfect model, and to introduce a structural error by removing 3D effects, using **Tripleclouds** instead of **Spartacus**.

For each experiment, maximum and average absolute errors (metrics-wise) of the best-score simulation are retained as a measure of the accuracy of the model. They are presented in Table 2. Note that **Spartacus** fluxes are as accurate as when **Spartacus** was run on reference cloud profiles. On the other hand, flux errors in the **Tripleclouds** experiment are about four times larger than those of **Spartacus**. This means that even with the possibility of compensating radiation error with wrong clouds, **Tripleclouds** could not produce fluxes as accurate as **Spartacus** in these experiments. Note that **Tripleclouds** scores are much larger than in perfect cloud experiments but this is partly because in previous experiments it was compared to 1D MC fluxes, whereas it is now compared to 3D MC fluxes.

Features of the 30 best simulations of each tuning experiment are analyzed in Figure 5. The top panel presents temporal evolutions of layer-wise maximum cloud fraction for ARMCU and RICO test cases. The bottom panel shows, for cloud scenes when radiative constraints were applied (8th and 12th hour of ARMCU, 5th and 12th hour of RICO), errors on the TOA upwelling fluxes as a function of solar zenith angles. Errors due to neglecting 3D effects in the MC calculations, and errors made by **Spartacus** and **Tripleclouds** on LES mean cloud profiles are also represented on these plots. It shows that although **Tripleclouds** simulations were constrained to be as close as possible to 3D MC fluxes, they still closely follow the 1D MC fluxes. Fluxes computed in the **Spartacus** experiment however match the MC reference for almost every angles, including those that were not constrained (not shown). In the upper plots, we see that, in addition to better simulating solar radiation, the 30 best **Spartacus** simulations exhibit cloud fraction evolutions that are compatible with reference values for ARMCU, and close to the LES ones for RICO. Conversely, the 30 best **Tripleclouds** simulations persistently overestimate maximum cloud fraction for both cases.

We interpret this as a sign of compensating errors. Plots in the lower row of Figure 5 show that the **Tripleclouds** simulation run on LES mean cloud profiles (red stars) systematically underestimate reflected solar radiation at $\text{SZA}=77^\circ$. The absolute error is the largest for the 8th hour of ARMCU where it reaches $-20.5 \text{ W}\cdot\text{m}^{-2}$. The tuning pro-

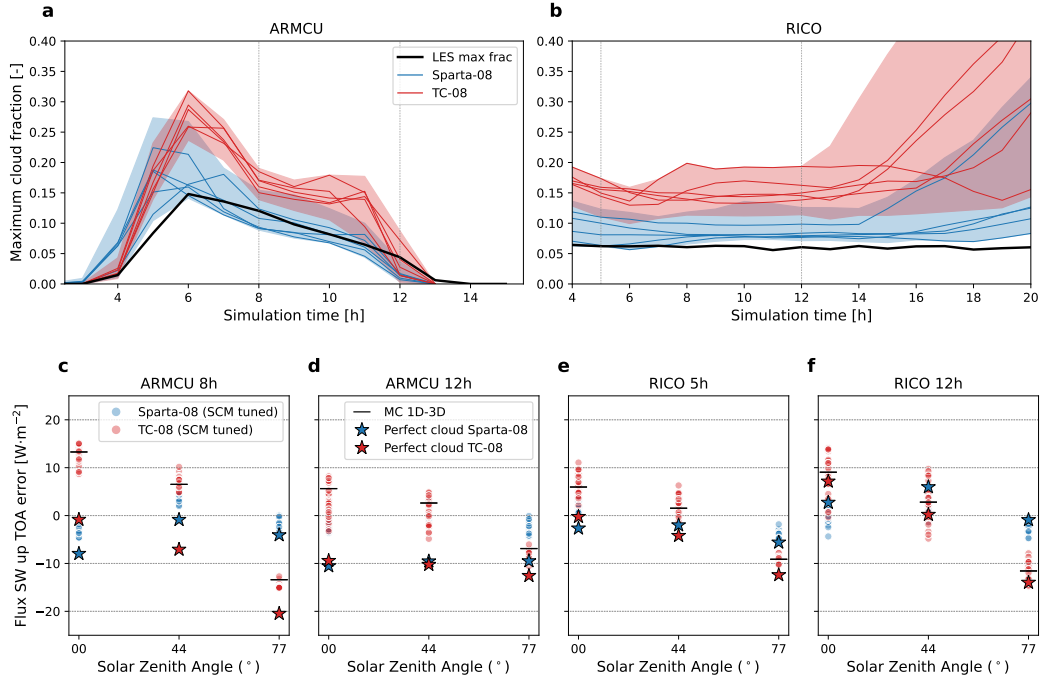


Figure 5. Clouds and radiation for the 30 best simulations for two experiments calibrating LMDZ parameters under radiative constraints using either **Tripleclouds** (red) or **Spartacus** (blue). Both solvers use parameters of the best **Spartacus** configuration from Section 3.1 (**Spartacus** 8-063). Left: ARMCU. Right: RICO. a-b: maximum cloud fraction (shadings represent the 30 best envelope, lines show the 5 best). c-f: error in upwelling TOA fluxes as a function of solar zenith angle, compared to MC 3D, for the four cloud scenes: ARMCU 8h (c), ARMCU 12h (d), RICO 5h (e), RICO 12h (f).

tol, which rejects simulations when the metric-wise maximum error is too large, systematically selects simulations where the error at $\text{SZA}=77^\circ$ is smaller than for “perfect clouds”, thanks to an increase of cloud fraction that compensates lack of 3D effects. This increase of cloud fraction is limited by the enhancement of the reflected radiation at $\text{SZA}=0^\circ$. The tuning finally selects simulations for which the bias shows the same absolute value for the maximum underestimation (about $-15 \text{ W}\cdot\text{m}^{-2}$ for $\text{SZA}=77^\circ$ at the 8th hour of ARMCU) and maximum overestimation (about $15 \text{ W}\cdot\text{m}^{-2}$ at $\text{SZA}=0^\circ$, especially for the 8th hour of ARMCU and 12th hour of RICO).

Overestimated cloud fraction probably also compensates in part for the fact that the radiative parameters are not optimal for **Tripleclouds**. Indeed, we saw in Figure 3 that for this configuration of ecRad, the overlap parameter leads to underestimating cloud cover (blue points) even with perfect vertical profiles of cloud fraction.

For **Spartacus**, angular dependence of reflected radiation is much better represented, avoiding compromises between errors at high and low zenith angles. Furthermore, 3D effects partially compensate for underestimated effective cloud cover. Overall, there is no sign of systematic cloud–radiation compensating errors with **Spartacus**. The best simulations obtained with a radiation-based tuning produce, at the same time, (i) radiation that is on average as good as in the perfect-cloud experiments, and (ii) maximum cloud fractions that correspond reasonably well to those of the LES.

4.3 Where we remove cloud fraction vs. 3D effects compensating errors

To verify our hypothesis, that part of the radiative error that is compensated by clouds in TC-08 experiment is related to 3D effects, we do two more pairs of experiments. First, we change the ecRad configuration, to use the best Hybrid one instead of the best **Spartacus** one of Section 3. Indeed, the best Hybrid configuration used on LES cloud profiles led to cloud cover estimates closer to the LES than the best **Spartacus** one, which is more favourable to **Tripleclouds**. We expect the remaining **Tripleclouds** flux errors to be smaller than before and hence, if our hypothesis is right, cloud fraction overestimation to be less pronounced.

This is verified in Figure 6, which is the same as Figure 5 but with both solvers using the best Hybrid configuration of ecRad instead of best **Spartacus**. The score of the **Spartacus** experiment is a little bit larger than before, suggesting that the degradation introduced in changing ecRad configuration is not compensated by clouds. The time evolution of maximum cloud fractions in the **Spartacus** experiments resembles those of Figure 5, which confirms that changing cloud-geometry parameters to a slightly less good configuration does not impact **Spartacus** too much. In the **Tripleclouds** experiment, the radiative score was not improved by much. Yet, the evolution of maximum cloud fraction is closer to LES and **Spartacus** ones than before: To reach the same radiative accuracy with **Tripleclouds**, in this experiment, cloud fractions did not have to be as large as before.

Maximum cloud fraction is still overestimated in the TC-16 experiment, compared to Sparta-16. It might be for two reasons: in the best Hybrid configuration, overlap parameter is similar to best **Tripleclouds** values, but the associated FSD is larger (for the same ℓ) which means that clouds are less reflective; again, in **Spartacus**, this overlap-heterogeneity combination is balanced by 3D effects, but in **Tripleclouds** a large FSD leads to an underestimated cloud reflectivity. Hence, maximum cloud fraction might be overestimated in compensation. The second reason is that, independently of heterogeneity, accurate 3D fluxes cannot be obtained simultaneously for all solar angles, for the right clouds, if 3D effects are not represented in the model. In TC-16 experiment, cloud fraction is not overestimated as much as in TC-08. Our interpretation is that the error in cloud reflectivity that is compensated by increasing cloud fraction is smaller: perfect-

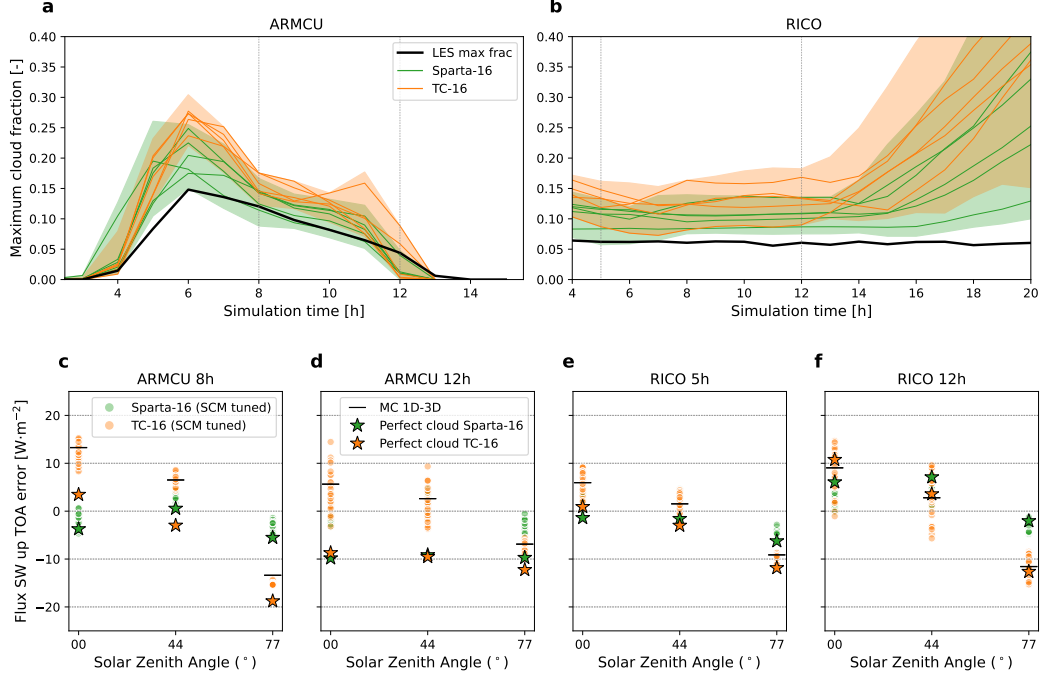


Figure 6. Clouds and radiation for the 30 best simulations for two experiments calibrating LMDZ parameters under radiative constraints using either **Tripleclouds** (orange) or **Spartacus** (green). Same as Figure 5 except that both solvers now use parameters of the best Hybrid configuration from Section 3.2 (Hybrid 16-092).

cloud **Tripleclouds** errors at SZA 77° are smaller with this ecRad configuration, around $-18.8 \text{ W}\cdot\text{m}^{-2}$ (against -20.5) at the 8th hour of ARMCU.

A final test is done in which we change the target metrics to an average over eight solar zenith angles, from 0° to 77° with step 11° . As 3D effects change sign when sun zenith angle increases, they partially cancel each other out when averaged. If lack of 3D effects is the error that is compensated by clouds in the **Tripleclouds** experiment, it should be less pregnant when targetting an average flux in which the expression of 3D effects is almost null. It is the unweighted arithmetic average that is considered here, as the simplest way to test our assumption; more complex averages that better represent solar angle distributions on Earth might be more relevant in a real GCM tuning exercise.

Results are presented in Figure 7. With the average flux metric, cloud fraction evolution is the same in the two experiments, using **Spartacus** or **Tripleclouds**. It means that **Tripleclouds** errors (*FSD* too large or lack of 3D effects) are no longer compensated by cloud fractions. Looking at the detailed flux errors, we see that **Tripleclouds** flux at 77° are more negatively biased compared to previous experiments, in agreement with the fact that this particular value is no longer constrained. It confirms that maximum cloud fractions were previously overestimated to compensate for lack of 3D effects in particular at this large solar zenith angle; because increasing cloud fraction also increases reflected flux at all other angles, in **Tripleclouds**, flux errors could not be entirely compensated by clouds: increasing cloud fraction too much would have led to even larger flux overestimation at $\text{SZA}=0^\circ$, which was not acceptable given the tolerance-to-error set in these experiments. To support this statement, a complementary experiment

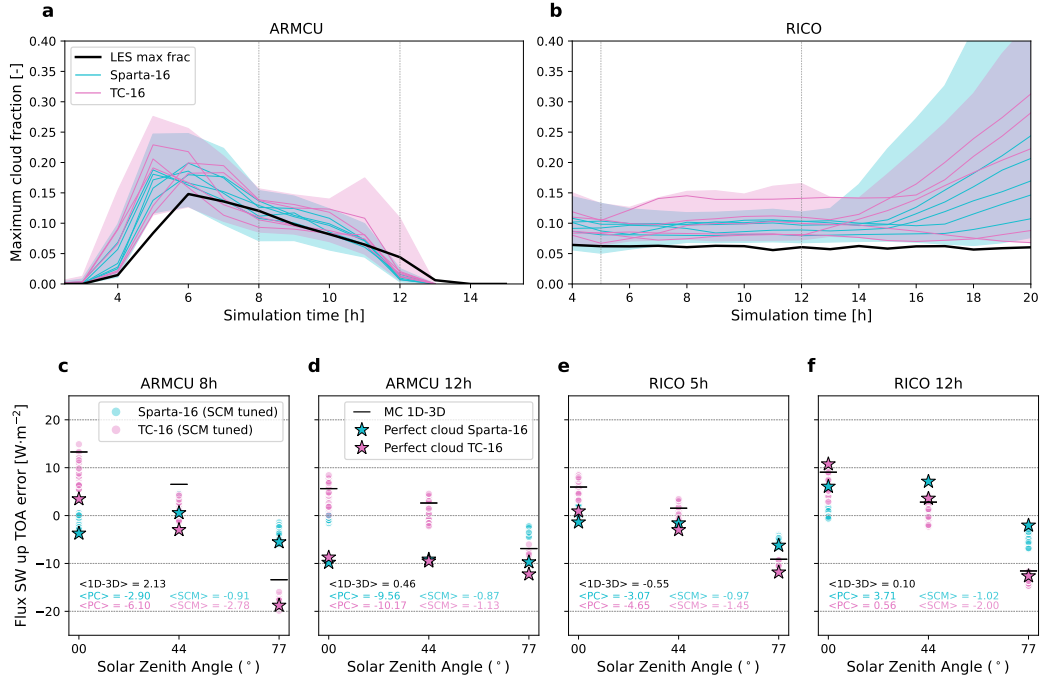


Figure 7. Clouds and radiation for the 30 best simulations for two experiments calibrating LMDZ parameters under radiative constraints using either **Tripleclouds** (pink) or **Spartacus** (cyan). Same as Figure 6 except that the radiative constraint is now the flux averaged over solar zenith angles, instead of flux values taken at three different SZA. c-f: In addition to flux errors as a function of SZA (markers), SZA-mean flux errors are written: $\langle 1D-3D \rangle$ is the average 3D effects, $\langle PC \rangle$ is for Perfect Cloud experiments (associated with star markers), $\langle SCM \rangle$ is for SCM-tuned experiments (associated with bullet markers).

where **Tripleclouds** was tuned using only constraints at $\text{SZA}=77^\circ$ is presented in Supporting Information (Figure S1).

This last experiment, targetting average flux, is an important result for climate modeling: it confirms that 3D radiative effects could, for a large part, self compensate when averaged over diurnal cycle, seasons or latitudes. It also proposes a simple compromise to tune a model with a 1D RT code, against LES/MC reference simulations. However, these results also suggest the possibility of compensating errors at high latitudes, where cloud fraction might be overestimated to compensate for lack of 3D radiative effects, if the model was tuned targetting high-latitude metrics.

5 Conclusion

We have shown that structural errors in radiative transfer models can indeed be compensated by errors in cloud properties when TOA radiative fluxes are targeted in a tuning process. Here, maximum cloud fractions are overestimated to compensate for underestimated cloud reflectivity at large zenith angles, stemming from the lack of 3D effects in the radiative model. This result provides a novel argument in favor of modelling 3D radiative effects in climate models: even if they were small on average and had a weak feedback on circulations and climate, we have shown that systematic errors in radiative transfer can generate systematic errors in other components of the model through tuning. A better radiative transfer model might remove the need for compensating errors and result in better clouds.

Here the demonstration was made in an idealized configuration, and our results should not be directly extrapolated to 3D coupled climate models. Indeed, in the SCM setup considered here, only shallow convection and cloud parameterizations can compensate structural radiative errors, whereas much more processes are at work in a 3D GCM, which can result in other compensating errors. Also, radiative fluxes were the only constraints, but Couvreux et al. (2021); Hourdin et al. (2021) claim that compensating errors can be prevented or at least limited by process-based tuning in SCM mode before tuning the full GCM. With this strategy, constraints can be set directly on cloud properties to rule out model configurations that yield wrong cloud fractions. Note however that tuning towards radiative targets while preventing clouds from compensating radiation errors might generate compensating errors elsewhere in the system.

Our work goes beyond the question of radiative transfer and clouds: We propose to see tuning as a way to investigate compensating errors and guide model development. Through tuning we explore parameter space, that is, model configurations and resulting climates, under a set of chosen constraints. This allows us to disentangle parametric from structural errors. Notably, when no set of parameters can be found for which all simulated metrics comply with user requirements, it indicates that structural errors are larger than tolerated errors, and hence that the model is incomplete. This is a powerful way to guide its development and accelerate its improvement. When simulated metrics do comply with prescribed requirements, resulting perturbed parameter ensembles of simulations (PPE) can be used to investigate compensating errors, better understand the model and its physics through global sensitivity studies, and quantify parametric uncertainty on various aspects of climate.

Finally, the tuning tool used here, High-Tune:Explorer, is based on machine learning techniques: predictive Gaussian Processes are trained on a small amount of simulated data and are then able to emulate the model's response much faster than the actual model. Thanks to this approach, the model's high-dimensional parameter space can be explored and shrunk efficiently. Machine learning is here at the service of physics; it helps saving computing time but not at the expense of the physical consistency of the model. This consistency is crucial for our confidence in climate projections and to keep

using models as tools to better understand climate. In the same spirit, we believe that research that aims at better understanding climate models and the act of modeling itself is a crucial aspect of climate sciences.

Open Research Section

High-Tune Explorer (htexplo) and LMDZ are available through the open source version control system “subversion” (svn). htexplo is distributed under the GPL-v3 license, and LMDZ is distributed under the CeCILL version 2 license. The htexplo release used in the study can be downloaded through `svn checkout http://svn.lmd.jussieu.fr/HighTune -r 568`. The LMDZ release used in the study can be configured and installed directly on Linux machines with an installation bash script `https://lmdz.lmd.jussieu.fr/pub/install_lmdz.sh` run with as `bash install_lmdz.sh -SCM -v 20250337.trunk`. The ecRad offline package is freely available under the terms of the Apache License Version 2.0. The release used in this study corresponds to commit fa642e, which is based on version v1.6-beta. A tar file of the htexplo, LMDZ and ecRad codes used as well as the data that supports this research, the results of the SCM simulations, as well as the scripts for visualization WILL BE MADE AVAILABLE ON A DOI IF THE PAPER IS ACCEPTED FOR PUBLICATION. The corresponding DOIs will be provided during galley proofs by placeholder “IPSL data catalog.”

Acknowledgments

The PhD scholarship of the first author was funded by Institut Pierre-Simon Laplace. This work was supported by the CNRS and the GDR DEPHY. The authors wish to thank the LMDZ team and the htexplo team for their investment in the development and maintenance of community tools.

References

- Barker, H. W., Stephens, G. L., & Fu, Q. (1999). The sensitivity of domain-averaged solar fluxes to assumptions about cloud geometry. *Quarterly Journal of the Royal Meteorological Society*, 125(558), 2127–2152. Retrieved from <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712555810> doi: 10.1002/qj.49712555810
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., ... Vuichard, N. (2020, July). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. , 12(7), e02010. doi: 10.1029/2019MS002010
- Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M., ... Stevens, B. (2002). Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, 128(582), 1075–1093. doi: 10.1256/003590002320373210
- Couvreur, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque, N., ... Xu, W. (2021). Process-Based Climate Model Development Harnessing Machine Learning: I. A Calibration Tool for Parameterization Improvement. *Journal of Advances in Modeling Earth Systems*, 13(3), e2020MS002217. doi: 10.1029/2020MS002217
- Dephy. (2020). *Dephy-scm: Single-column model standards and case drivers*. Retrieved from <https://github.com/gdr-dephy/dephy-scm>
- Hogan, R. J., & Bozzo, A. (2018). A Flexible and Efficient Radiation Scheme for the ECMWF Model. *Journal of Advances in Modeling Earth Systems*, 10(8), 1990–2008. doi: 10.1029/2018MS001364
- Hogan, R. J., Fielding, M. D., Barker, H. W., Villefranque, N., & Schäfer, S. A. K. (2019). Entrapment: An Important Mechanism to Explain the Shortwave

- 3D Radiative Effect of Clouds. *Journal of the Atmospheric Sciences*, 76(7), 2123–2141. doi: 10.1175/JAS-D-18-0366.1
- Hogan, R. J., & Illingworth, A. J. (2000, October). Deriving cloud overlap statistics from radar. *Quarterly Journal of the Royal Meteorological Society*, 126(569), 2903–2909. doi: 10.1002/qj.49712656914
- Hogan, R. J., Schäfer, S. A. K., Klinger, C., Chiu, J. C., & Mayer, B. (2016). Representing 3-D cloud radiation effects in two-stream schemes: 2. Matrix formulation and broadband evaluation. *Journal of Geophysical Research: Atmospheres*, 121(14), 8583–8599. doi: 10.1002/2016JD024875
- Hogan, R. J., & Shonk, J. K. P. (2013, February). Incorporating the Effects of 3D Radiative Transfer in the Presence of Clouds into Two-Stream Multilayer Radiation Schemes. *Journal of the Atmospheric Sciences*, 70(2), 708–724. doi: 10.1175/JAS-D-12-041.1
- Hourdin, F., Ferster, B., Deshayes, J., Mignot, J., Musat, I., & Williamson, D. (2023, July). Toward machine-assisted tuning avoiding the underestimation of uncertainty in climate change projections. *Science Advances*, 9(29), eadf2758. doi: 10.1126/sciadv.adf2758
- Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., ... Roehrig, R. (2013, May). LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, 40(9-10), 2193–2222. doi: 10.1007/s00382-012-1343-y
- Hourdin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M.-P., ... Idelkadi, A. (2019). Unified Parameterization of Convective Boundary Layer Transport and Clouds With the Thermal Plume Model. *Journal of Advances in Modeling Earth Systems*, 11(9), 2910–2933. doi: 10.1029/2019MS001666
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., ... Williamson, D. (2017, March). The Art and Science of Climate Model Tuning. *Bulletin of the American Meteorological Society*, 98(3), 589–602. doi: 10.1175/BAMS-D-15-00135.1
- Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., ... Ghattas, J. (2020). LMDZ6A: The Atmospheric Component of the IPSL Climate Model With Improved and Better Tuned Physics. *Journal of Advances in Modeling Earth Systems*, 12(7), e2019MS001892. doi: 10.1029/2019MS001892
- Hourdin, F., Williamson, D., Rio, C., Couvreux, F., Roehrig, R., Villefranche, N., ... Volodina, V. (2021). Process-Based Climate Model Development Harnessing Machine Learning: II. Model Calibration From Single Column to Global. *Journal of Advances in Modeling Earth Systems*, 13(6), e2020MS002225. doi: 10.1029/2020MS002225
- Jam, A., Hourdin, F., Rio, C., & Couvreux, F. (2013, June). Resolved Versus Parametrized Boundary-Layer Plumes. Part III: Derivation of a Statistical Scheme for Cumulus Clouds. *Boundary-Layer Meteorology*, 147(3), 421–441. doi: 10.1007/s10546-012-9789-3
- Konsta, D., Dufresne, J.-L., Chepfer, H., Vial, J., Koshiro, T., Kawai, H., ... Ogura, T. (2022). Low-Level Marine Tropical Clouds in Six CMIP6 Models Are Too Few, Too Bright but Also Too Compact and Too Homogeneous. *Geophysical Research Letters*, 49(11), e2021GL097593. doi: 10.1029/2021GL097593
- Lac, C., Chaboureaud, J.-P., Masson, V., Pinty, J.-P., Tulet, P., Escobar, J., ... Wautelet, P. (2018, May). Overview of the Meso-NH model version 5.4 and its applications. *Geoscientific Model Development*, 11(5), 1929–1969. doi: 10.5194/gmd-11-1929-2018
- Lafore, J. P., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., ... Vilà-Guerau de Arellano, J. (1998, January). The Meso-NH Atmospheric Simulation System. Part I: Adiabatic formulation and control simulations. *Annales Geophysicae*, 16(1), 90–109. doi: 10.1007/s00585-997-0090-6

- Madeleine, J.-B., Hourdin, F., Grandpeix, J.-Y., Rio, C., Dufresne, J.-L., Vignon, E., ... Bonazzola, M. (2020). Improved Representation of Clouds in the Atmospheric Component LMDZ6A of the IPSL-CM6A Earth System Model. *Journal of Advances in Modeling Earth Systems*, 12(10), e2020MS002046. doi: 10.1029/2020MS002046
- McKee, T. B., & Cox, S. K. (1974). Scattering of Visible Radiation by Finite Clouds. *Journal of the Atmospheric Sciences*, 31(7), 1885–1892. doi: 10.1175/1520-0469(1974)031<1885:SOVRBF>2.0.CO;2
- Meador, W. E., & Weaver, W. R. (1980, March). Two-Stream Approximations to Radiative Transfer in Planetary Atmospheres: A Unified Description of Existing Methods and a New Improvement. , 37(3), 630–643. doi: 10.1175/1520-0469(1980)037<0630:TSATRT>2.0.CO;2
- Nam, C., Bony, S., Dufresne, J.-L., & Chepfer, H. (2012). The ‘too few, too bright’ tropical low-cloud problem in CMIP5 models. *Geophysical Research Letters*, 39(21). doi: 10.1029/2012GL053421
- Pincus, R., Barker, H. W., & Morcrette, J.-J. (2003, July). A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *Journal of Geophysical Research: Atmospheres*, 108(D13), n/a–n/a. doi: 10.1029/2002JD003322
- Schäfer, S. A. K., Hogan, R. J., Klinger, C., Chiu, J. C., & Mayer, B. (2016). Representing 3-D cloud radiation effects in two-stream schemes: 1. Longwave considerations and effective cloud edge length. *Journal of Geophysical Research: Atmospheres*, 121(14), 8567–8582. doi: 10.1002/2016JD024876
- Shonk, J. K. P., & Hogan, R. J. (2008, June). Tripleclouds: An Efficient Method for Representing Horizontal Cloud Inhomogeneity in 1D Radiation Schemes by Using Three Regions at Each Height. *Journal of Climate*, 21(11), 2352–2370. doi: 10.1175/2007JCLI1940.1
- Shonk, J. K. P., Hogan, R. J., Edwards, J. M., & Mace, G. G. (2010, July). Effect of improving representation of horizontal and vertical cloud structure on the Earth’s global radiation budget. Part I: review and parametrization. *Quarterly Journal of the Royal Meteorological Society*, n/a–n/a. doi: 10.1002/qj.647
- vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., ... Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *Journal of Advances in Modeling Earth Systems*, 3(2). doi: 10.1029/2011MS000056
- Várnai, T., & Davies, R. (1999). Effects of cloud heterogeneities on shortwave radiation: Comparison of cloud-top variability and internal heterogeneity. *Journal of the Atmospheric Sciences*, 56(24), 4206–4224.
- Vernon, I., Goldstein, M., & Bower, R. G. (2010, December). Galaxy Formation: A Bayesian Uncertainty Analysis. *Bayesian Analysis*, 05(04). doi: 10.1214/10-ba524
- Villefranche, N., Blanco, S., Couvreur, F., Fournier, R., Gautrais, J., Hogan, R. J., ... Williamson, D. (2021). Process-Based Climate Model Development Harnessing Machine Learning: III. The Representation of Cumulus Geometry and Their 3D Radiative Effects. *Journal of Advances in Modeling Earth Systems*, 13(4), e2020MS002423. doi: 10.1029/2020MS002423
- Villefranche, N., Fournier, R., Couvreur, F., Blanco, S., Cornet, C., Eymet, V., ... Tregan, J.-M. (2019). A Path-Tracing Monte Carlo Library for 3-D Radiative Transfer in Highly Resolved Cloudy Atmospheres. *Journal of Advances in Modeling Earth Systems*, 11(8), 2449–2473. doi: 10.1029/2018MS001602
- Webb, M., Senior, C., Bony, S., & Morcrette, J. J. (2001, September). Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dynamics*, 17(12), 905–922. (WOS:000171263400001) doi: 10.1007/s003820100157
- Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L.,

- & Yamazaki, K. (2013, October). History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. *Climate Dynamics*, *41*(7-8), 1703–1729. doi: 10.1007/s00382-013-1896-4
- Yamada, T. (1983, January). Simulations of Nocturnal Drainage Flows by a q2l Turbulence Closure Model. *Journal of the Atmospheric Sciences*, *40*(1), 91–106. doi: 10.1175/1520-0469(1983)040<0091:SONDFB>2.0.CO;2

VI.3 Rayonnement 3D

Chapitre VII

Synthèse sur la pratique du tuning avec htexplo pour la modélisation du climat

Remarque lecture : Remarques en vrac, rien de figé ni rédigé Un truc de l'intro qui va plutôt aller ici : A noter que des compensations d'erreurs peuvent parfois être désirables. La question de savoir où, à quel point et pour quelles raisons est-il acceptable que des compensations d'erreur soient à l'oeuvre dans les modèles climatiques a été peu instruite jusqu'à aujourd'hui ; probablement par manque de méthode quantitative accessible en temps de calcul. Se doter d'une stratégie permettant d'éviter autant que possible les compensations d'erreur lors de l'ajustement de LMDZ est un parti pris par le groupe de modélisation de l'IPSL pour élaborer la stratégie d'ajustement de son modèle de climat. Les réflexions portent alors sur une stratégie de tuning

VII.1 Discussion du 25 septembre

- Sur le papier rayonnement, mettre dans cette partie la réflexion et ce que l'on a appris des tolérances à l'erreur
- discussion sur la notion de paramètres libres effectifs vs diagnostiqués à partir des observations. Mettre les paramètres libres dans leur contexte

VII.2 Reflexions 19 juin 2025

Mail de Fred de juin 2025 : AU PASSAGE : j'ai réfléchi à ton chapitre Meta. Il pourrait commencer par une liste des autres activités dans lesquelles tu t'es impliquée en tant que tuneuse. Ça permettrait en plus de placer les autres papiers. Ça va juste être trop facile à mon avis si tu fais ça. Tu pourras ramener les ateliers, le tuning du couplé global, les bourrasques de Lamine ...

Liste des ateliers dans lesquels je me suis impliquées de prêt ou de loin sur le tuning :

- PPE ou tuning avec htexplo pour le développement de paramétrisations
 - PPE des papiers de Lea et Valentin → aller voir leur papiers pour voir ce qu'ils en disent
 - Param développée par Audran : je sais qu'il s'est posé la question du tuning mais je n'étais pas impliquée, je ne sais pas ce qu'il a fait au final. Je ne sais pas à quel point il a utilisé le travail de Sidiki.

- Travail avec Lamine : j’imagine que ce n’est pas publié. Il faudrait que j’en rediscute avec Fred et/ou avec Lamine car mes souvenirs sont flous. Je sais qu’il y avait l’idée du tuning LES/LES qui est ”nouvelle”
- on avait fait un peu de tuning dans le dvp de la paramétrisation des
- Travail avec Sidiki : publié + co-autrice
- tuning tke : publié + co-autrice downdrafts
- PPE dynamico vs lon-lat
- PPE couplé : on en a fait quoi au final????
- c’est plus un détail, mais on avait bossé avec un gars du cstb à roffiac qui développait un modèle de corps humain : ça à donné quoi ce truc ?
- Ateliers stratégie de tuning pour LMDZ :
 - PPE pour vérifier le portage du modèle (par Saloua) ; aussi fait par Fred sur Adastra
 -
- Les différents tuning de LMDZ portés par Ionela. Mais je ne suis plus trop cette dynamique ces derniers temps, je ne sais pas trop où on en est.

VII.3 Reflexions 12 Novembre 2024

Remarque : J’ai pas envie que ça fasse un truc du type synthèse des bonnes pratiques où je ne sais quoi ; chose que j’avais pas mal tendance à faire je crois.

Je me suis dit que cette partie devrait répondre aux questions suivantes :

- Qu’a-t-on appris sur le tuning avec ces expériences ?
- Quels messages souhaite-t-on adresser à la communauté ?

Mais en fait j’arrive pas à y répondre comme ça alors peut être il faut que je fasse une liste un peu en vrac.

- Tuner c’est toujours choisir des vecteurs de paramètres libres qui permettent de matcher un certain nombre de contraintes (compte tenu d’une tolérance à l’erreur). En général, quand on dit qu’on tune le modèle, on cherche à trouver les vecteurs de paramètres qui nous permettent d’avoir des simulations aux plus proches du climat actuel. Il faut garder en tête qu’on peut utiliser ce cadre pour d’autre chose.
- On peut avoir envie d’ajuster le modèle pour des utilisations particulières (comme de simuler le climat d’une certaine région en particulier) ; pas forcément pour le climat global.
- Et pour répondre à des questions spécifiques

Le tuning comme aide et accélérateur du développement des GCMs :

- Recalibration rapide
- Identification de besoin de paramétrisation
- Mieux comprendre le modèle pour guider les choix de modélisation

Qu’est ce qui me paraît important à dire (en vrac) :

- Le fait d’objectiver le problème du tuning (=choisir des vecteurs de paramètres libres qui permettent de matcher des contraintes) permet de :
 - documenter les protocoles de tuning choisis pour les modèles donc de partager nos connaissances et pratiques sur le sujet

- Tester différentes manières de tuner notre modèle : choix des paramètres libres, des métriques, des tolérances à l'erreurs, des configurations utilisée pour le tuning...
- D'apprendre sur le modèle quand on l'ajuste (même si on fait juste un tuning "classique") : on apprend sur l'erreur structurelle du modèle, sur sa "dispersion paramétrique". On a un moyen objectif de dire si l'ajout d'une paramétrisation a amélioré le modèle. On sait très vite quand on se plante aussi.
- D'avoir un problème clair avec une réponse clair
- L'outil reste un outil. Il ne fait que répondre à notre question. Il nécessite autant d'expertise que lorsque le tuning était fait à la main (choix des paramètres et des métriques)
- L'outil nous pousse à quantifier les incertitudes de nos configurations (point positif)
- On a toujours tuné les GCMs. Tous modèle [ayant des paramètres incertains ?] nécessite d'être tuné. Donc le tuning est forcément là, autant en faire quelque chose d'instructif
- Tuning et quantification de l'erreur paramétrique → vers une meilleurs quantification de l'incertitude de nos projections climatiques ?
- Tuning multi-configuration et compensations d'erreurs
- Il y a un côté investissement dans le tuning. Déjà il faut se mettre a un nouvel outil. Ensuite, si on veut tuner des configurations lourdes alors c'est lourd, ca prend beaucoup de ressources (temps de calcul et temps humain) ; donc ce n'est pas adapté pour l'aide au développement. Ce qui est cool, c'est de développer des configurations légères (type SCM, LAM, zoomé guidé etc), adapté à la situation que l'on regarde et au développement que l'on veut faire. C'est sur que c'est un gros travail (qui nous apprend aussi beaucoup). Mais une fois qu'on a ça le tuning vient vraiment s'insérer dans la pratique de développement de paramétrisation et la faciliter. C'est ce qu'il s'est fait avec le modèle des thermiques de LMDZ ; qui est donc d'autant plus facile à étudier et développer.
- La problématique du coût de calcul est toujours la

Chapitre VIII

Conclusion

Remarque lecture : Remarques en vrac, rien n'est figé ni rédigé

Si je devais continuer ma thèse aujourd'hui :

- enrichir la stratégie de tuning :
 - Se donner les moyens de tuner "à pas cher" = pré-tuning sur deux mois
 - Ajuster le rayonnement en surface en utilisant le réseau BSRN avec des configurations guidées en vent. C'est aussi l'occasion d'instruire la différence guidée vs non-guidée dans des ensembles paramétriques. C'est-à-dire de quantifier l'erreur qui vient des paramétrisations physiques (à dynamique "fixée") vs erreur du modèle (dynamique + physique couplée). AUssi pour avancer sur la connaissance des incertitudes structurales du modèle et sur les compensations d'erreurs
 - attaquer les biais systématiques des modèles : le biais froid avec préconditionnement 1D sur des cas de convection profonde et côté rayonnement
 - ajouter des métriques sur des cas 1D de convection profonde au tuning des modèles. Lien avec les métriques globales ?
 - ajouter des métriques sur des configurations intermédiaires types zoomé guidé ou LAM. Gros travail sur le SIRTa qui montre qu'on ne peut pas régler conjointement les paramètres des modèles de surface et ceux d'atmosphère.
 - Choix des paramètres pour le tuning : comment "monter en niveau" sur ces choix ? On est limité à 30 paramètres, avec une limite sur la taille d'exploration du NROY. Mais même sans cette limitation, on ne tunera jamais tous les paramètres libres en même temps. Comment on fait pour choisir quel paramètre on fixe et quels paramètres on explore ? Et comment on fait pour fixer ces paramètres et instruire ce que l'on perd en faisant ça ? (typiquement tuning 1D orienté processus avec pleins de paramètres puis on en fixe un certain nombre)
- Rayonnement :
 - lien processus-global en ajoutant des métriques radiatives sur des cas 1D
 - faire des choix sur la configuration de ecRad à utiliser, affiner notre compréhension des impacts des effets 3D sur le climat
 - Paramétriser les paramètres libres de ecRad, pour ajouter de la cohérence entre rayonnement et physique
- outil :
 - Participer à la vie et au développement de htexplo
 - Collaborer avec Daniel pour comprendre et incorporer ses avancées récentes, en particulier en ce qui concerne l'exploration d'espace petit (organiser une visite en Angleterre pour relancer cette collaboration ?)

- Apprendre à utiliser les émulateurs pour mieux comprendre le lien entre métriques et paramètres. Les émulateurs pourraient nous permettre d'isoler la sensibilité d'un seul paramètre en fixant les autres

Bibliographie

- Adler, R. F., Sapiano, M. R. P., Huffman, G. J., Wang, J.-J., Gu, G., Bolvin, D., Chiu, L., Schneider, U., Becker, A., Nelkin, E., Xie, P., Ferraro, R., and Shin, D.-B. (2018). The Global Precipitation Climatology Project (GPCP) Monthly Analysis (New Version 2.3) and a Review of 2017 Global Precipitation. *Atmosphere*, 9(4) :138.
- AERIS, P., IPSL, and SEDOO (Data service OMP) (2019). Strateole 2 – Long-duration balloon flights at the tropical tropopause. <https://strateole2.aeris-data.fr/>.
- Arakawa, A. and Schubert, W. H. (1974). Interaction of a Cumulus Cloud Ensemble with the Large-Scale Environment, Part I. *Journal of the Atmospheric Sciences*, 31(3) :674–701.
- Arjdal, K., Vignon, É., Driouech, F., Chérut, F., Er-Raki, S., Sima, A., Chehbouni, A., and Drobinski, P. (2024). Modeling Land–Atmosphere Interactions over Semiarid Plains in Morocco : In-Depth Assessment of GCM Stretched-Grid Simulations Using In Situ Data. *Journal of Applied Meteorology and Climatology*, 63(3) :369–386.
- Baas, P., van de Wiel, B. J. H., van der Linden, S. J. A., and Bosveld, F. C. (2018). From Near-Neutral to Strongly Stratified : Adequately Modelling the Clear-Sky Nocturnal Boundary Layer at Cabauw. *Boundary-Layer Meteorology*, 166(2) :217–238.
- Barker, H. W., Cole, J. N. S., Li, J., Yi, B., and Yang, P. (2015). Estimation of Errors in Two-Stream Approximations of the Solar Radiative Transfer Equation for Cloudy-Sky Conditions. *Journal of the Atmospheric Sciences*, 72(11) :4053–4074.
- Bastrikov, V., MacBean, N., Bacour, C., Santaren, D., Kuppel, S., and Peylin, P. (2018). Land surface model parameter optimisation using in situ flux data : Comparison of gradient-based versus random search algorithms (a case study using ORCHIDEE v1.9.5.2). *Geoscientific Model Development*, 11(12) :4739–4754.
- Bellprat, O., Kotlarski, S., Lüthi, D., Elía, R. D., Frigon, A., Laprise, R., and Schär, C. (2016). Objective Calibration of Regional Climate Models : Application over Europe and North America. *Journal of Climate*, 29(2) :819–838.
- Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C. (2012). Objective calibration of regional climate models. *Journal of Geophysical Research : Atmospheres*, 117(D23).
- Bernstein, D. N. and Neelin, J. D. (2016). Identifying sensitive ranges in global warming precipitation change dependence on convective parameters. *Geophysical Research Letters*, 43(11) :5841–5850.
- Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., and John, V. O. (2011). COSP : Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, 92(8) :1023–1043.
- Bonnet, P., Pastori, L., Schwabe, M., Giorgetta, M., Iglesias-Suarez, F., and Eyring, V. (2025). Tuning the ICON-A 2.6.4 climate model with machine-learning-based emulators and history matching. *Geoscientific Model Development*, 18(12) :3681–3706.

- Borella, A., Vignon, É., Boucher, O., Meurdesoif, Y., and Fairhead, L. (2025). A New Prognostic Parameterization of Subgrid Ice Supersaturation and Cirrus Clouds in the ICOLMDZ AGCM. *Journal of Advances in Modeling Earth Systems*, 17(8) :e2024MS004918.
- Bosveld, F. C., Baas, P., van Meijgaard, E., de Bruijn, E. I. F., Steeneveld, G.-J., and Holtslag, A. A. M. (2014). The Third GABLS Intercomparison Case for Evaluation Studies of Boundary-Layer Models. Part A : Case Selection and Set-Up. *Boundary-Layer Meteorology*, 152(2) :133–156.
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D’Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, Lionel, E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, 12(7) :e2019MS002010.
- Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P. G., Golaz, J.-C., Khairoutdinov, M., Lewellen, D. C., Lock, A. P., MacVean, M. K., Moeng, C.-H., Neggers, R. a. J., Siebesma, A. P., and Stevens, B. (2002). Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, 128(582) :1075–1093.
- Cahalan, R. F., Oreopoulos, L., Marshak, A., Evans, K. F., Davis, A. B., Pincus, R., Yetzer, K. H., Mayer, B., Davies, R., Ackerman, T. P., Barker, H. W., Clothiaux, E. E., Ellingson, R. G., Garay, M. J., Kassianov, E., Kinne, S., Macke, A., O’hirok, W., Partain, P. T., Prigarin, S. M., Rublev, A. N., Stephens, G. L., Szczap, F., Takara, E. E., Várnai, T., Wen, G., and Zhuravleva, T. B. (2005). THE I3RC : Bringing Together the Most Advanced Radiative Transfer Tools for Cloudy Atmospheres. *Bulletin of the American Meteorological Society*, 86(9) :1275–1294.
- Campoy, A. (2013). *Influence de l’hydrologie souterraine sur la modélisation du climat à l’échelle régionale et globale*. PhD thesis, Université Pierre et Marie Curie - Paris VI.
- Campoy, A., Ducharne, A., Cheruy, F., Hourdin, F., Polcher, J., and Dupont, J. C. (2013). Response of land surface fluxes and precipitation to different soil bottom hydrological conditions in a general circulation model. *Journal of Geophysical Research : Atmospheres*, 118(19) :10,725–10,739.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, 76 :1–32.
- Cheruy, F., Campoy, A., Dupont, J.-C., Ducharne, A., Hourdin, F., Haeffelin, M., Chiriaco, M., and Idelkadi, A. (2013). Combined influence of atmospheric physics and soil hydrology on the simulated meteorology at the SIRTa atmospheric observatory. *Climate Dynamics*, 40(9-10) :2251–2269.
- Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., Bastrikov, V., Vuichard, N., Diallo, B., Dufresne, J.-L., Ghattas, J., Grandpeix, J.-Y., Idelkadi, A., Mellul,

- L., Maignan, F., Ménégoz, M., Ottlé, C., Peylin, P., Servonnat, J., Wang, F., and Zhao, Y. (2020). Improved Near-Surface Continental Climate in IPSL-CM6A-LR by Combined Evolutions of Atmospheric and Land Surface Physics. *Journal of Advances in Modeling Earth Systems*, 12(10) :e2019MS002005.
- Chiriaco, M., Dupont, J.-C., Bastin, S., Badosa, J., Lopez, J., Haefelin, M., Chepfer, H., and Guzman, R. (2018). ReOBS : A new approach to synthesize long-term multi-variable dataset and application to the SIRTAs supersite. *Earth System Science Data*, 10(2) :919–940.
- Coindreau, O., Hourdin, F., Haefelin, M., Mathieu, A., and Rio, C. (2007). Assessment of Physical Parameterizations Using a Global Climate Model with Stretchable Grid and Nudging. *Monthly Weather Review*, 135(4) :1474–1489.
- Collins, M., Booth, B. B. B., Harris, G. R., Murphy, J. M., Sexton, D. M. H., and Webb, M. J. (2006). Towards quantifying uncertainty in transient climate change. *Climate Dynamics*, 27(2) :127–147.
- Courtier, P., Thépaut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519) :1367–1387.
- Couvreur, F., Guichard, F., Redelsperger, J.-L., Kiemle, C., Masson, V., Lafore, J.-P., and Flamant, C. (2005). Water-vapour variability within a convective boundary-layer assessed by large-eddy simulations and IHOP_2002 observations. *Quarterly Journal of the Royal Meteorological Society*, 131(611) :2665–2693.
- Couvreur, F., Hourdin, F., and Rio, C. (2009). Resolved Versus Parametrized Boundary-Layer Plumes. Part I : A Parametrization-Oriented Conditional Sampling in Large-Eddy Simulations.
- Couvreur, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque, N., Rio, C., Audouin, O., Salter, J., Bazile, E., Brient, F., Favot, F., Honnert, R., Lefebvre, M.-P., Madeleine, J.-B., Rodier, Q., and Xu, W. (2021). Process-Based Climate Model Development Harnessing Machine Learning : I. A Calibration Tool for Parameterization Improvement. *Journal of Advances in Modeling Earth Systems*, 13(3) :e2020MS002217.
- Craig, P., Goldstein, M., Seheult, A. H., and Smith, J. A. (1996). Bayes linear strategies for matching hydrocarbon reservoir history.
- Cumming, J. A. and Goldstein, M. (2010). Bayes linear Uncertainty Analysis for Oil Reservoirs Based on Multiscale Computer Experiments. In *The Oxford Handbook of Applied Bayesian Analysis*, number 10 in Oxford Handbooks Series, pages 241–270. Oxford University Press.
- de Finetti, B. (2017). *Theory of Probability : A Critical Introductory Treatment*. John Wiley & Sons.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The ERA-Interim reanalysis : Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656) :553–597.
- Dephy (2020). Dephy-scm : Single-column model standards and case drivers.

- Diallo, F. B., Hourdin, F., Rio, C., Traore, A.-K., Mellul, L., Guichard, F., and Kergoat, L. (2017). The Surface Energy Budget Computed at the Grid-Scale of a Climate Model Challenged by Station Data in West Africa. *Journal of Advances in Modeling Earth Systems*, 9(7) :2710–2738.
- Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F. M., Duprat, T., Fukuda, M., Grobe, H., Haeffelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C. N., Longenecker, D., Lupi, A., Maturilli, M., Mimouni, M., Ntsangwane, L., Ogihara, H., Olano, X., Olefs, M., Omori, M., Passamani, L., Pereira, E. B., Schmithüsen, H., Schumacher, S., Sieger, R., Tamlyn, J., Vogt, R., Vuilleumier, L., Xia, X., Ohmura, A., and König-Langlo, G. (2018). Baseline Surface Radiation Network (BSRN) : Structure and data description (1992–2017). *Earth System Science Data*, 10(3) :1491–1501.
- Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J.-P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J.-Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M.-P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N. (2013). Climate change projections using the IPSL-CM5 Earth System Model : From CMIP3 to CMIP5. *Climate Dynamics*, 40(9) :2123–2165.
- Edwards, N. R., Cameron, D., and Rougier, J. (2011). Precalibrating an intermediate complexity climate model. *Climate Dynamics*, 37(7) :1469–1482.
- Emanuel, K. A. (1991). A scheme for representing cumulus convection in large-scale models. *J. Atmos. Sci.*, 48 :2313–2335.
- Emanuel, K. A. (1993). A cumulus representation based on the episodic mixing model : the importance of mixing and microphysics in predicting humidity. *A.M.S. Meteorol. Monographs*, 24(46) :185–192.
- Fielding, M. D., Schäfer, S. A. K., Hogan, R. J., and Forbes, R. M. (2020). Parametrizing cloud geometry and its application in a subgrid cloud-edge erosion scheme. *Quarterly Journal of the Royal Meteorological Society*, 146(729) :1651–1667.
- Forget, F., Guerlet, S., Hourdin, F., Lebonnois, S., Millour, E., Spiga, A., Talagrand, O., and Turbet, M. (2024). Planétologie comparée et atmosphères extraterrestres au LMD. *Histoire de la recherche contemporaine*, 13(Tome XIII n°1).
- Goldstein, M. (1986). Prevision. In Kotz, S. and Johnson, L. N., editors, *Encyclopedia of Statistical Sciences*, volume 7, pages 175–176. Wiley, New York.
- Goldstein, M. and Wooff, D. (2007). *Bayes Linear Statistics : Theory and Methods*. John Wiley & Sons.
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A. (2000). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, 16(2) :147–168.
- Grailet, J.-F., Hogan, R. J., Ghilain, N., Bolsée, D., Fettweis, X., and Grégoire, M. (2025). Inclusion of the ECMWF ecRad radiation scheme (v1.5.0) in the MAR (v3.14), regional evaluation for Belgium, and assessment of surface shortwave spectral fluxes at Uccle. *Geoscientific Model Development*, 18(6) :1965–1988.

- Grandpeix, J. and Lafore, J. (2010). A Density Current Parameterization Coupled with Emanuel's Convection Scheme. Part I : The Models. *Journal of Atmospheric Sciences*, 67 :881–897.
- Grandpeix, J., Lafore, J., and Cheruy, F. (2010). A Density Current Parameterization Coupled with Emanuel's Convection Scheme. Part II : 1D Simulations. *Journal of Atmospheric Sciences*, 67 :898–922.
- Grandpeix, J. Y., Phillips, V., and Tailleux, R. (2004). Improved mixing representation in Emanuel's convection scheme. *Q. J. R. Meteorol. Soc.*, 130 :3207–3222.
- Gregory, J. M. and Mitchell, J. F. B. (1997). The climate response to CO₂ of the Hadley Centre coupled AOGCM with and without flux adjustment. *Geophysical Research Letters*, 24(15) :1943–1946.
- Guillemot, H. and Hourdin, F. (2021). Modélisation du Climat. *Encyclopædia Universalis*.
- Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M.-A., Walters, S., Lamarque, J.-F., and Holland, E. A. (2004). Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model : Description and background tropospheric chemistry evaluation. *Journal of Geophysical Research : Atmospheres*, 109(D4) :2003JD003957.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730) :1999–2049.
- Heymsfield, A. J. (1977). Precipitation Development in Stratiform Ice Clouds : A Microphysical and Dynamical Study. *Journal of the Atmospheric Sciences*, 34(2) :367–381.
- Heymsfield, A. J. and Donner, L. J. (1990). A Scheme for Parameterizing Ice-Cloud Water Content in General Circulation Models. *Journal of the Atmospheric Sciences*, 47(15) :1865–1877.
- Hogan, R. J. and Bozzo, A. (2018). A Flexible and Efficient Radiation Scheme for the ECMWF Model. *Journal of Advances in Modeling Earth Systems*, 10(8) :1990–2008.
- Hogan, R. J., Fielding, M. D., Barker, H. W., Villefranque, N., and Schäfer, S. A. K. (2019). Entrapment : An Important Mechanism to Explain the Shortwave 3D Radiative Effect of Clouds. *Journal of the Atmospheric Sciences*, 76(7) :2123–2141.
- Hogan, R. J. and Illingworth, A. J. (2000). Deriving cloud overlap statistics from radar. *Quarterly Journal of the Royal Meteorological Society*, 126(569) :2903–2909.
- Hogan, R. J., Schäfer, S. A. K., Klinger, C., Chiu, J. C., and Mayer, B. (2016). Representing 3-D cloud radiation effects in two-stream schemes : 2. Matrix formulation and broadband evaluation. *Journal of Geophysical Research : Atmospheres*, 121(14) :8583–8599.
- Hogan, R. J. and Shonk, J. K. P. (2008). Radiation parametrization and clouds.
- Hourdin, F. (2005). *Représentation du transport direct et inverse dans les modèles globaux de climat et étude des couplages entre composition et dynamique atmosphérique sur Titan*. Habilitation to direct research (HDR), Université Paris-6.
- Hourdin, F. (2024). LMDZ : récit d'un physicien à la barre d'un modèle global dans la tempête climatique. *Histoire de la recherche contemporaine*, 13(Tome XIII n°1).

- Hourdin, F., Couvreux, F., and Menut, L. (2002). Parameterization of the Dry Convective Boundary Layer Based on a Mass Flux Representation of Thermals. *Journal of the Atmospheric Sciences*, 59(6) :1105–1123.
- Hourdin, F., Ferster, B., Deshayes, J., Mignot, J., Musat, I., and Williamson, D. (2023). Toward machine-assisted tuning avoiding the underestimation of uncertainty in climate change projections. *Science Advances*, 9(29) :eadf2758.
- Hourdin, F., Foujols, M.-A., Codron, F., Guemas, V., Dufresne, J.-L., Bony, S., Denvil, S., Guez, L., Lott, F., Ghattas, J., Braconnot, P., Marti, O., Meurdesoif, Y., and Bopp, L. (2013). Impact of the LMDZ atmospheric grid configuration on the climate and sensitivity of the IPSL-CM5A coupled model. *Climate Dynamics*, 40(9) :2167–2192.
- Hourdin, F., Găinusă-Bogdan, A., Braconnot, P., Dufresne, J.-L., Traore, A.-K., and Rio, C. (2015). Air moisture control on ocean surface temperature, hidden key to the warm bias enigma. *Geophysical Research Letters*, 42(24) :10,885–10,893.
- Hourdin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M.-P., Brient, F., and Idelkadi, A. (2019). Unified Parameterization of Convective Boundary Layer Transport and Clouds With the Thermal Plume Model. *Journal of Advances in Modeling Earth Systems*, 11(9) :2910–2933.
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M., and Williamson, D. (2017). The Art and Science of Climate Model Tuning. *Bulletin of the American Meteorological Society*, 98(3) :589–602.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F. (2006). The LMDZ4 general circulation model : Climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dynamics*, 27(7) :787–813.
- Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., Jam, A., Musat, I., Idelkadi, A., Fairhead, L., Foujols, M.-A., Mellul, L., Traore, A.-K., Dufresne, J.-L., Boucher, O., Lefebvre, M.-P., Millour, E., Vignon, E., Jouhaud, J., Diallo, F. B., Lott, F., Gastineau, G., Caubel, A., Meurdesoif, Y., and Ghattas, J. (2020). LMDZ6A : The Atmospheric Component of the IPSL Climate Model With Improved and Better Tuned Physics. *Journal of Advances in Modeling Earth Systems*, 12(7) :e2019MS001892.
- Hourdin, F., Williamson, D., Rio, C., Couvreux, F., Roehrig, R., Villefranque, N., Musat, I., Fairhead, L., Diallo, F. B., and Volodina, V. (2021). Process-Based Climate Model Development Harnessing Machine Learning : II. Model Calibration From Single Column to Global. *Journal of Advances in Modeling Earth Systems*, 13(6) :e2020MS002225.
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases : Calculations with the AER radiative transfer models. *Journal of Geophysical Research : Atmospheres*, 113(D13).
- IPSL, CNRS, CEA, EPFL, and Polytechnique, E. (2023). Site de AWACA - Accueil. <https://awaca.ipsl.fr/>.
- IPSL/SIRTA (s.d.). Galerie photo du SIRTA. <https://sirta.ipsl.fr/>.
- Jackson, C. S., Sen, M. K., Huerta, G., Deng, Y., and Bowman, K. P. (2008). Error Reduction and Convergence in Climate Prediction. *Journal of Climate*, 21(24) :6698–6709.
- Jahangir, E. (2022). *Amélioration de La Paramétrisation Des Propriétés Optiques Des Nuages d’eau Liquide Dans Le Spectre Solaire*. These de doctorat, Toulouse 3.

- Jakob, C. and Klein, S. A. (2000). A parametrization of the effects of cloud and precipitation overlap for use in general-circulation models. *Quarterly Journal of the Royal Meteorological Society*, 126(568) :2525–2544.
- Jam, A., Hourdin, F., Rio, C., and Couvreux, F. (2013). Resolved Versus Parametrized Boundary-Layer Plumes. Part III : Derivation of a Statistical Scheme for Cumulus Clouds. *Boundary-Layer Meteorology*, 147(3) :421–441.
- Joseph, J. H., Wiscombe, W. J., and Weinman, J. A. (1976). The Delta-Eddington Approximation for Radiative Flux Transfer. *Journal of the Atmospheric Sciences*, 33(12) :2452–2459.
- Jourdain, L., Bekki, S., Lott, F., and Lefèvre, F. (2008). The coupled chemistry-climate model LMDz-REPROBUS : Description and evaluation of a transient simulation of the period 1980–1999. *Annales Geophysicae*, 26(6) :1391–1413.
- Klemp, J. B. and Wilhelmson, R. B. (1978). The Simulation of Three-Dimensional Convective Storm Dynamics. *Journal of the Atmospheric Sciences*, 35(6) :1070–1096.
- Konsta, D., Dufresne, J.-L., Chepfer, H., Vial, J., Koshiro, T., Kawai, H., Bodas-Salcedo, A., Roehrig, R., Watanabe, M., and Ogura, T. (2022). Low-Level Marine Tropical Clouds in Six CMIP6 Models Are Too Few, Too Bright but Also Too Compact and Too Homogeneous. *Geophysical Research Letters*, 49(11) :e2021GL097593.
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19(1).
- Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., and Richardson, A. D. (2012). Constraining a global ecosystem model with multi-site eddy-covariance data. *Biogeosciences*, 9(10) :3757–3776.
- Lac, C., Chaboureau, J.-P., Masson, V., Pinty, J.-P., Tulet, P., Escobar, J., Leriche, M., Barthe, C., Aouizerats, B., Augros, C., Aumond, P., Auguste, F., Bechtold, P., Berthet, S., Bielli, S., Bosseur, F., Caumont, O., Cohard, J.-M., Colin, J., Couvreux, F., Cuxart, J., Delautier, G., Dauhut, T., Ducrocq, V., Filippi, J.-B., Gazen, D., Geoffroy, O., Gheusi, F., Honnert, R., Lafore, J.-P., Lebeaupin Brossier, C., Libois, Q., Lunet, T., Mari, C., Maric, T., Mascart, P., Mogé, M., Molinié, G., Nuissier, O., Pantillon, F., Peyrillé, P., Pergaud, J., Perraud, E., Pianezze, J., Redelsperger, J.-L., Ricard, D., Richard, E., Riette, S., Rodier, Q., Schoetter, R., Seyfried, L., Stein, J., Suhre, K., Taufour, M., Thouron, O., Turner, S., Verrelle, A., Vié, B., Visentin, F., Vionnet, V., and Wautelet, P. (2018). Overview of the Meso-NH model version 5.4 and its applications. *Geoscientific Model Development*, 11(5) :1929–1969.
- Lafore, J. P., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., Fischer, C., Hèreil, P., Mascart, P., Masson, V., Pinty, J. P., Redelsperger, J. L., Richard, E., and Vilà-Guerau de Arellano, J. (1998). The Meso-NH Atmospheric Simulation System. Part I : Adiabatic formulation and control simulations. *Annales Geophysicae*, 16(1) :90–109.
- Lenderink, G., Siebesma, A. P., Cheinet, S., Irons, S., Jones, C. G., Marquet, P., üLLER, F. M., Olmeda, D., Calvo, J., Sánchez, E., and Soares, P. M. M. (2004). The diurnal cycle of shallow cumulus clouds over land : A single-column model intercomparison study. *Quarterly Journal of the Royal Meteorological Society*, 130(604) :3339–3364.
- Lenhard, J. and Winsberg, E. (2010). Holism, entrenchment, and the future of climate model pluralism. *Studies in History and Philosophy of Science Part B : Studies in History and Philosophy of Modern Physics*, 41(3) :253–262.
- LMD/EMC3 (2020). Gallery of LMDZ meteograms for SIRTa. <https://web.lmd.jussieu.fr/~lmdz/LMDZOPE/LMDZ-SIRTa/meteo/>.

- Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., Manalo-Smith, N., and Wong, T. (2009). Toward Optimal Closure of the Earth's Top-of-Atmosphere Radiation Budget. *Journal of Climate*, 22(3) :748–766.
- Lohou, F., Lothon, M., Bastin, S., Brut, A., Canut, G., Cohard, J. M., Cheruy, F., Couvreux, F., Dupont, S., Lafont, S., Roehrig, R., Román-Cascón, C., Yagüe, C., Bernard, E., Boone, A., Dupont, J. C., Feofilov, A., Fernandes, R., Grulois, M., Jomé, M., Maison, A., and Zouzoua, M. (2025). Model and Observation for surface–atmosphere interactions over heterogeneous landscape : MOSAI project. *Journal of the European Meteorological Society*, 3 :100019.
- Madec, G. (2008). NEMO ocean engine, Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL). Technical Report 27, IPSL, France.
- Madeleine, J.-B., Hourdin, F., Grandpeix, J.-Y., Rio, C., Dufresne, J.-L., Vignon, E., Boucher, O., Konsta, D., Cheruy, F., Musat, I., Idelkadi, A., Fairhead, L., Millour, E., Lefebvre, M.-P., Mellul, L., Rochetin, N., Lemonnier, F., Touzé-Peiffer, L., and Bonazzola, M. (2020). Improved Representation of Clouds in the Atmospheric Component LMDZ6A of the IPSL-CM6A Earth System Model. *Journal of Advances in Modeling Earth Systems*, 12(10) :e2020MS002046.
- Manabe, S. and Wetherald, R. T. (1975). The Effects of Doubling the CO₂ Concentration on the climate of a General Circulation Model. *Journal of the Atmospheric Sciences*, 32(1) :3–15.
- Manners, J., Edwards, J. M., Hill, P., and Thelen, J.-C. (2017). SOCRATES Technical Guide Suite Of Community RAdiative Transfer codes based on Edwards and Slingo. Technical report.
- Météo-France (2019). Bilan climatique de l'année 2018. Technical report, Meteo-France, Paris.
- Météo-France (2023). 2023 sur le podium des années les plus chaudes en France. <https://meteofrance.com/actualites-et-dossiers/actualites/2023-sur-le-podium-des-annees-les-plus-chaudes-en-france>.
- Météo-France (2025a). Bilan climatique, année 2024, France hexagonale et Corse. Technical report, Meteo-France, Paris.
- Météo-France (2025b). Fiche climatologique Paris-Montsouris (75). Technical report, Météo-France, Toulouse.
- Mignot, J., Hourdin, F., Deshayes, J., Boucher, O., Gastineau, G., Musat, I., Vancoppenolle, M., Servonnat, J., Caubel, A., Chérut, F., Denvil, S., Dufresne, J.-L., Ethé, C., Fairhead, L., Foujols, M.-A., Grandpeix, J.-Y., Levavasseur, G., Marti, O., Menary, M., Rio, C., Rousset, C., and Silvy, Y. (2021). The Tuning Strategy of IPSL-CM6A-LR. *Journal of Advances in Modeling Earth Systems*, 13(5) :e2020MS002340.
- Monin, A. S. and Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci. USSR*, 151(163) :187.
- Morcrette, C. J., O'Connor, E. J., and Petch, J. C. (2012). Evaluation of two cloud parametrization schemes using ARM and Cloud-Net observations. *Quarterly Journal of the Royal Meteorological Society*, 138(665) :964–979.
- Nam, C., Bony, S., Dufresne, J.-L., and Chepfer, H. (2012). The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophysical Research Letters*, 39(21).
- Neelin, J. D., Bracco, A., Luo, H., McWilliams, J. C., and Meyerson, J. E. (2010). Considerations for parameter optimization and sensitivity in climate models. *Proceedings of the National Academy of Sciences of the United States of America*, 107(50) :21349–21354.
- Oueslati, B. and Bellon, G. (2013). Tropical precipitation regimes and mechanisms of regime transitions : Contrasting two aquaplanet general circulation models. *Climate Dynamics*, 40(9) :2345–2358.

- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., Humphrey, M., Isaac, P., Polidori, D., Reichstein, M., Ribeca, A., van Ingen, C., Vuichard, N., Zhang, L., Amiro, B., Ammann, C., Arain, M. A., Ardö, J., Arkebauer, T., Arndt, S. K., Arriga, N., Aubinet, M., Aurela, M., Baldocchi, D., Barr, A., Beamesderfer, E., Marchesini, L. B., Bergeron, O., Beringer, J., Bernhofer, C., Berveiller, D., Billesbach, D., Black, T. A., Blanken, P. D., Bohrer, G., Boike, J., Bolstad, P. V., Bonal, D., Bonnefond, J.-M., Bowling, D. R., Bracho, R., Brodeur, J., Brümmer, C., Buchmann, N., Burban, B., Burns, S. P., Buysse, P., Cale, P., Cavagna, M., Cellier, P., Chen, S., Chini, I., Christensen, T. R., Cleverly, J., Collalti, A., Consalvo, C., Cook, B. D., Cook, D., Coursolle, C., Cremonese, E., Curtis, P. S., D'Andrea, E., da Rocha, H., Dai, X., Davis, K. J., Cinti, B. D., de Grandcourt, A., Ligne, A. D., De Oliveira, R. C., Delpierre, N., Desai, A. R., Di Bella, C. M., di Tommasi, P., Dolman, H., Domingo, F., Dong, G., Dore, S., Duce, P., Dufrêne, E., Dunn, A., Dušek, J., Eamus, D., Eichelmann, U., ElKhidir, H. A. M., Eugster, W., Ewenz, C. M., Ewers, B., Famulari, D., Fares, S., Feigenwinter, I., Feitz, A., Fensholt, R., Filippa, G., Fischer, M., Frank, J., Galvagno, M., Gharun, M., Gianelle, D., Gielen, B., Gioli, B., Gitelson, A., Goded, I., Goeckede, M., Goldstein, A. H., Gough, C. M., Goulden, M. L., Graf, A., Griebel, A., Gruening, C., Grünwald, T., Hammerle, A., Han, S., Han, X., Hansen, B. U., Hanson, C., Hatakka, J., He, Y., Hehn, M., Heinesch, B., Hinko-Najera, N., Hörtnagl, L., Hutley, L., Ibrom, A., Ikawa, H., Jackowicz-Korczynski, M., Janouš, D., Jans, W., Jassal, R., Jiang, S., Kato, T., Khomik, M., Klatt, J., Knohl, A., Knox, S., Kobayashi, H., Koerber, G., Kolle, O., Kosugi, Y., Kotani, A., Kowalski, A., Kruijt, B., Kurbatova, J., Kutsch, W. L., Kwon, H., Launiainen, S., Laurila, T., Law, B., Leuning, R., Li, Y., Liddell, M., Limousin, J.-M., Lion, M., Liska, A. J., Lohila, A., López-Ballesteros, A., López-Blanco, E., Loubet, B., Loustau, D., Lucas-Moffat, A., Lüers, J., Ma, S., Macfarlane, C., Magliulo, V., Maier, R., Mammarella, I., Manca, G., Marcolla, B., Margolis, H. A., Marras, S., Massman, W., Mastepanov, M., Matamala, R., Matthes, J. H., Mazzenga, F., McCaughey, H., McHugh, I., McMillan, A. M. S., Merbold, L., Meyer, W., Meyers, T., Miller, S. D., Minerbi, S., Moderow, U., Monson, R. K., Montagnani, L., Moore, C. E., Moors, E., Moreaux, V., Moureaux, C., Munger, J. W., Nakai, T., Neirynck, J., Nesic, Z., Nicolini, G., Noormets, A., Northwood, M., Nosetto, M., Nouvellon, Y., Novick, K., Oechel, W., Olesen, J. E., Ourcival, J.-M., Papuga, S. A., Parmentier, F.-J., Paul-Limoges, E., Pavelka, M., Peichl, M., Pendall, E., Phillips, R. P., Pilegaard, K., Pirk, N., Posse, G., Powell, T., Prasse, H., Prober, S. M., Rambal, S., Rannik, Ü., Raz-Yaseef, N., Rebmann, C., Reed, D., de Dios, V. R., Restrepo-Coupe, N., Reverter, B. R., Roland, M., Sabbatini, S., Sachs, T., Saleska, S. R., Sánchez-Cañete, E. P., Sanchez-Mejia, Z. M., Schmid, H. P., Schmidt, M., Schneider, K., Schrader, F., Schroder, I., Scott, R. L., Sedláč, P., Serrano-Ortíz, P., Shao, C., Shi, P., Shironya, I., Siebicke, L., Šigut, L., Silberstein, R., Sirca, C., Spano, D., Steinbrecher, R., Stevens, R. M., Sturtevant, C., Suyker, A., Tagesson, T., Takanashi, S., Tang, Y., Tapper, N., Thom, J., Tomassucci, M., Tuovinen, J.-P., Urbanski, S., Valentini, R., van der Molen, M., van Gorsel, E., van Huissteden, K., Varlagin, A., Verfaillie, J., Vesala, T., Vincke, C., Vitale, D., Vygodskaya, N., Walker, J. P., Walter-Shea, E., Wang, H., Weber, R., Westermann, S., Wille, C., Wofsy, S., Wohlfahrt, G., Wolf, S., Woodgate, W., Li, Y., Zampedri, R., Zhang, J., Zhou, G., Zona, D., Agarwal, D., Biraud, S., Torn, M., and Papale, D. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, 7(1) :225.
- Peatier, S., Sanderson, B. M., Terray, L., and Roehrig, R. (2022). Investigating Parametric Dependence of Climate Feedbacks in the Atmospheric Component of CNRM-CM6-1. *Geophysical Research Letters*, 49(9) :e2021GL095084.
- Phillips, T. J., Klein, S. A., Ma, H.-Y., Tang, Q., Xie, S., Williams, I. N., Santanello, J. A.,

- Cook, D. R., and Torn, M. S. (2017). Using ARM Observations to Evaluate Climate Model Simulations of Land-Atmosphere Coupling on the U.S. Southern Great Plains. *Journal of Geophysical Research : Atmospheres*, 122(21) :11,524–11,548.
- Pincus, R., Barker, H. W., and Morcrette, J.-J. (2003). A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *Journal of Geophysical Research : Atmospheres*, 108(D13).
- Pope, V. D., Gallani, M. L., Rowntree, P. R., and Stratton, R. A. (2000). The impact of new physical parametrizations in the Hadley Centre climate model : HadAM3. *Climate Dynamics*, 16(2) :123–146.
- Pukelsheim, F. (1994). The Three Sigma Rule. *The American Statistician*, 48(2) :88–91.
- Raoult, N., Beylat, S., Salter, J. M., Hourdin, F., Bastrikov, V., Ottlé, C., and Peylin, P. (2024). Exploring the potential of history matching for land surface model calibration. *Geoscientific Model Development*, 17(15) :5779–5801.
- Raoult, N., Ottlé, C., Peylin, P., Bastrikov, V., and Maugis, P. (2021). Evaluating and Optimizing Surface Soil Moisture Drydowns in the ORCHIDEE Land Surface Model at In Situ Locations. *Journal of Hydrometeorology*, 22(4) :1025–1043.
- Rio, C. and Hourdin, F. (2008). A Thermal Plume Model for the Convective Boundary Layer : Representation of Cumulus Clouds. *Journal of the Atmospheric Sciences*, 65(2) :407–425.
- Rio, C., Hourdin, F., Couvreux, F., and Jam, A. (2010). Resolved Versus Parametrized Boundary-Layer Plumes. Part II : Continuous Formulations of Mixing Rates for Mass-Flux Schemes. *Boundary-Layer Meteorology*, 135(3) :469–483.
- Rochetin, N., Couvreux, F., Grandpeix, J.-Y., and Rio, C. (2014). Deep Convection Triggering by Boundary Layer Thermals. Part I : LES Analysis and Stochastic Triggering Formulation. *Journal of the Atmospheric Sciences*, 71(2) :496–514.
- Rougier, J. (2007). Probabilistic Inference for Future Climate Using an Ensemble of Climate Model Evaluations. *Climatic Change*, 81(3) :247–264.
- Rougier, J., Sexton, D. M. H., Murphy, J. M., and Stainforth, D. (2009). Analyzing the Climate Sensitivity of the HadSM3 Climate Model Using Ensembles from Different but Related Experiments. *Journal of Climate*, 22(13) :3540–3557.
- Rousset, C., Vancoppenolle, M., Madec, G., Fichet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Levy, C., Masson, S., and Vivier, F. (2015). The Louvain-La-Neuve sea ice model LIM3.6 : Global and regional capabilities. *Geoscientific Model Development*, 8(10) :2991–3005.
- Sandu, I. and Stevens, B. (2011). On the Factors Modulating the Stratocumulus to Cumulus Transitions. *Journal of the Atmospheric Sciences*, 68(9) :1865–1881.
- Schäfer, S. A. K., Hogan, R. J., Klinger, C., Chiu, J. C., and Mayer, B. (2016). Representing 3-D cloud radiation effects in two-stream schemes : 1. Longwave considerations and effective cloud edge length. *Journal of Geophysical Research : Atmospheres*, 121(14) :8567–8582.
- Schlesinger, M. E., Oh, J.-H., and Rosenfeld, D. (1988). A Parameterization of the Evaporation of Rainfall. *Monthly Weather Review*, 116(10) :1887–1895.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate : A review. *Earth-Science Reviews*, 99(3) :125–161.
- Sexton, D. M. H., McSweeney, C. F., Rostron, J. W., Yamazaki, K., Booth, B. B. B., Murphy, J. M., Regayre, L., Johnson, J. S., and Karmalkar, A. V. (2021). A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections : Part 1 : Selecting the parameter combinations. *Climate Dynamics*, 56(11) :3395–3436.

- Shonk, J. K. P. and Hogan, R. J. (2008). Tripleclouds : An Efficient Method for Representing Horizontal Cloud Inhomogeneity in 1D Radiation Schemes by Using Three Regions at Each Height. *Journal of Climate*, 21(11) :2352–2370.
- Sundqvist, H. (1978). A parameterization scheme for non-convective condensation including prediction of cloud water content. *Quarterly Journal of the Royal Meteorological Society*, 104(441) :677–690.
- Tabary, P., Dupuy, P., L’henaff, G., Gueguen, C., Moulin, L., and Laurantin, O. (2012). A 10-year (1997-2006) reanalysis of Quantitative Precipitation Estimation over France : methodology and first results. In *IAHS-AISH publication*, number 351, pages 255–260.
- Talagrand, O. and Courtier, P. (1987). Variational Assimilation of Meteorological Observations With the Adjoint Vorticity Equation. I : Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478) :1311–1328.
- Tiedtke, M. (1989). A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Monthly Weather Review*, 117(8) :1779–1800.
- Touzé-Peiffer, L. (2021). *Parameterization of Atmospheric Convection in Numerical Climate Models - Practices and Epistemological Challenges*. PhD thesis, Sorbonne Université.
- van Dalum, C. T., van de Berg, W. J., Gadde, S. N., van Tiggelen, M., van der Drift, T., van Meijgaard, E., van Ulft, L. H., and van den Broeke, M. R. (2024). First results of the polar regional climate model RACMO2.4. *The Cryosphere*, 18(9) :4065–4088.
- vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., Cheng, A., Couvreux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D. C., Mechem, D., Nakamura, K., Noda, A., Shipway, B. J., Slawinska, J., Wang, S., and Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *Journal of Advances in Modeling Earth Systems*, 3(2).
- Vernon, I., Goldstein, M., and Bower, R. G. (2010). Galaxy Formation : A Bayesian Uncertainty Analysis. *Bayesian Analysis*, 05(04).
- Vignon, E. (2017). *The Extreme Atmospheric Boundary Layer over the Antarctic Plateau and Its Representation in Climate Models*. Theses, Université Grenoble Alpes.
- Vignon, É., Arjdal, K., Cheruy, F., Coulon-Decorzans, M., Dehondt, C., Dubos, T., Fromang, S., Hourdin, F., Lange, L., Raillard, L., Rivière, G., Roehrig, R., Sima, A., Spiga, A., and Tiengou, P. (2024). Designing a Fully-Tunable and Versatile TKE-l Turbulence Parameterization for the Simulation of Stable Boundary Layers. *Journal of Advances in Modeling Earth Systems*, 16(10) :e2024MS004400.
- Vignon, E., Hourdin, F., Genthon, C., Gallée, H., Bazile, E., Lefebvre, M.-P., Madeleine, J.-B., and Van de Wiel, B. J. H. (2017). Antarctic boundary layer parametrization in a general circulation model : 1-D simulations facing summer observations at Dome C. *Journal of Geophysical Research : Atmospheres*, 122(13) :6818–6843.
- Villefranche, N. (2019). *Les Effets Radiatifs 3D Des Nuages de Couche Limite : De Leur Simulation Explicite à Leur Paramétrisation*. These de doctorat, Toulouse 3.
- Villefranche, N., Blanco, S., Couvreux, F., Fournier, R., Gautrais, J., Hogan, R. J., Hourdin, F., Volodina, V., and Williamson, D. (2021). Process-Based Climate Model Development Harnessing Machine Learning : III. The Representation of Cumulus Geometry and Their 3D Radiative Effects. *Journal of Advances in Modeling Earth Systems*, 13(4) :e2020MS002423.
- Villefranche, N., Fournier, R., Couvreux, F., Blanco, S., Cornet, C., Eymet, V., Forest, V., and Tregan, J.-M. (2019). A Path-Tracing Monte Carlo Library for 3-D Radiative Transfer in Highly Resolved Cloudy Atmospheres. *Journal of Advances in Modeling Earth Systems*, 11(8) :2449–2473.

- Villefranque, N., Hourdin, F., d'Alençon, L., Blanco, S., Boucher, O., Caliot, C., Coustet, C., Dauchet, J., El Hafi, M., Eymet, V., Farges, O., Forest, V., Fournier, R., Gautrais, J., Masson, V., Piaud, B., and Schoetter, R. (2022). The “teapot in a city” : A paradigm shift in urban climate modeling. *Science Advances*, 8(27) :eabp8934.
- Volodina, V. and Challenor, P. (2021). The importance of uncertainty quantification in model reproducibility. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 379(2197) :20200071.
- Walker, G. K., Sud, Y. C., and Atlas, R. (1995). Impact of the Ongoing Amazonian Deforestation on Local Precipitation : A GCM Simulation Study. *Bulletin of the American Meteorological Society*, 76(3) :346–362.
- Wang, H., Burleyson, C. D., Ma, P.-L., Fast, J. D., and Rasch, P. J. (2018). Using the Atmospheric Radiation Measurement (ARM) Datasets to Evaluate Climate Models in Simulating Diurnal and Seasonal Variations of Tropical Clouds. *Journal of Climate*, 31(8) :3301–3325.
- WCRP (2017). Workshop on model tuning. <https://www.wcrp-climate.org/sessions/374-model-tuning>.
- Whittle, P. (2000). *Probability via Expectation*. Springer Science & Business Media.
- Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y.-P. (2009). Improving land surface models with FLUXNET data. *Biogeosciences*, 6(7) :1341–1359.
- Williamson, D., Blaker, A. T., Hampton, C., and Salter, J. (2015). Identifying and removing structural biases in climate models with history matching. *Climate Dynamics*, 45(5) :1299–1324.
- Williamson, D., Blaker, A. T., and Sinha, B. (2017). Tuning without over-tuning : Parametric uncertainty quantification for the NEMO ocean model. *Geoscientific Model Development*, 10(4) :1789–1816.
- Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L., and Yamazaki, K. (2013). History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. *Climate Dynamics*, 41(7) :1703–1729.
- Yamada, T. (1983). Simulations of Nocturnal Drainage Flows by a q2l Turbulence Closure Model. *Journal of the Atmospheric Sciences*, 40(1) :91–106.
- Yamazaki, K., Sexton, D. M. H., Rostron, J. W., McSweeney, C. F., Murphy, J. M., and Harris, G. R. (2021). A perturbed parameter ensemble of HadGEM3-GC3.05 coupled model projections : Part 2 : Global performance and future changes. *Climate Dynamics*, 56(11) :3437–3471.
- Yang, B., Qian, Y., Lin, G., Leung, L. R., Rasch, P. J., Zhang, G. J., McFarlane, S. A., Zhao, C., Zhang, Y., Wang, H., Wang, M., and Liu, X. (2013). Uncertainty quantification and parameter tuning in the CAM5 Zhang-McFarlane convection scheme and impact of improved convection on the global circulation and climate. *Journal of Geophysical Research : Atmospheres*, 118(2) :395–415.
- Zhang, C., Xie, S., Klein, S. A., Ma, H.-y., Tang, S., Van Weverberg, K., Morcrette, C. J., and Petch, J. (2018). CAUSES : Diagnosis of the Summertime Warm Bias in CMIP5 Climate Models at the ARM Southern Great Plains Site. *Journal of Geophysical Research : Atmospheres*, 123(6) :2968–2992.

- Zhang, T., Li, L., Lin, Y., Xue, W., Xie, F., Xu, H., and Huang, X. (2015). An automatic and effective parameter optimization method for model tuning. *Geoscientific Model Development*, 8(11) :3579–3591.
- Zou, L., Qian, Y., Zhou, T., and Yang, B. (2014). Parameter Tuning and Calibration of RegCM3 with MIT–Emanuel Cumulus Parameterization Scheme over CORDEX East Asia Domain. *Journal of Climate*, 27(20) :7687–7701.

Annexe A

Résumé des paramètres libres

Remarque lecture : Annexe pas encore relu par Fred et Najda, pas trop de raisons que ça bouge
Paramètres libres dans les expériences de tuning :

Paramètre	Plage par défaut	Valeur CMIP6	Paramétrisation	Description
A1 thermals_afac	[0,5 ; 1,2]	0,667	Thermiques	Facteur sur la contribution de la flotabilité dans l'accélération du panache (sans unité). Equation II.27 et II.28.
A2 thermals_fact_epsilon	[1,5.10 ⁻³ ; 4.10 ⁻³]	2.10 ⁻³	Thermiques	Facteur sur le freinage du panache par frottement (sans unité). Equation II.27.
B1 thermals_betalpha	[0 ; 1]	0,95	Thermiques	Facteur sur l'entraînement et le déentraînement (sans unité). Plus il est grand, plus le panache se mélange avec son environnement. Equation II.27 et II.28.
CQ thermals_detr_q_coef	[0 ; 0,02]	0,12	Thermiques	Facteur sur l'influence du contraste d'humidité entre le panache et l'environnement sur le déentraînement (sans unité). Plus CQ est grand, plus le panache entraîne. Equation II.28.
DZTH thermals_ed_dz	[0,05 ; 0,2]	0,07	Thermiques	Facteur sur la distance à laquelle le déentraînement du panache voit la température au dessus de lui (sans unité). Plus cette distance est grande, plus le panache voit l'inversion de température en avance et plus il va entraîner bas. Equation II.29.
BG1 cloudth_sigma1s_factor	[0,4 ; 2]	1,1	Thermiques (PDF d'eau)	Facteur sur la largeur de la gaussienne de l'environnement (sans unité). Equation II.33.
BG2 cloudth_sigma2s_factor	[0,03 ; 0,2]	0,09	Thermiques (PDF d'eau)	Facteur sur la largeur de la gaussienne des thermiques (sans unité). Equation II.34.

Paramètre	Plage par défaut	Valeur CMIP6	Paramétrisation	Description
CLC cld_lc_lsc cld_lc_con	$[1.10^{-4} ; 1.10^{-3}]$	$6,5.10^{-4} \text{ kg.kg}^{-1}$	Condensation grande échelle (précipitations)	Seuil à partir duquel la conversion de l'eau liquide en précipitation est significative, (kg d'eau liquide par kg d'air sec). Equation II.35 .
CLTAU cld_tau_lsc cld_tau_con	[900,18000]	900 s	Condensation grande échelle (precipitations)	Temps caractéristique de conversion de l'eau liquide en précipitation (secondes). Plus il est grand, plus il faut du temps pour convertir l'eau nuageuse en précipitation, donc plus les nuages ont une grande durée de vie. Equation II.35 .
EVAP coef_eva	$[5.10^{-5} ; 5.10^{-4}]$	1.10^{-4}	Condensation grande échelle (precipitations)	Efficacité de réévaporation des pluies (sans unité). Equation II.30 .
REI (16.REI= rei_min et 61,29.REI= rei_max)	[0,5 ; 1,3]	1	Microphysique (et rayonnement ?)	Facteur sur les rayons des particules de nuages
RQSTOP ratqshaut	[0,05 ; 0,6]	0,4	Condensation grande échelle (PDF d'eau)	Valeur du facteur de l'écart-type de la distribution d'eau sous maille en haut de l'atmosphère (sans unité). Figure II.5 (b).
RQSDP ratqsdp	[500 ; 30 000]	10 000 Pa	Condensation grande échelle (PDF d'eau)	Largeur de la transition de la tangente hypobolique pour le facteur de l'écart-type de la distribution d'eau sous maille (Pa). Figure II.5 (b).
RQSP0 ratqsp0	[30 000 ; 60 000]	45 000 Pa	Condensation grande échelle (PDF d'eau)	Hauteur de la transition de la tangente pour le facteur de l'écart-type de la distribution d'eau sous maille (Pa). Figure II.5 (b).

Paramètre	Plage par défaut	Valeur CMIP6	Paramétrisation	Description
FALLV ffallv_lsc ffallv_con	[0,3 ; 2]	0,8	Convection profonde et Condensation grande échelle	Facteur sur la vitesse de chute des cristaux de glace (sans unité). Plus les cristaux chutes vites, moins les nuages de glaces ont une vie longue. Equation II.37 .
OMEPMX (1- epmax)	[1.10^{-4} ; 0.1]	1.10^{-3}	Convection profonde (precipitations)	1-efficacité maximale de la conversion de l'eau en précipitation. Plus OMEPMX est grand, plus le seuil maximal est petit, donc moins efficace va être la conversation de l'eau en pluie (quand elle est déjà très efficace donc quand il fait très froid)
ELCV elcrit	[1.10^{-4} ; 2.10^{-3}] <i>kg/kg</i>	3.10^{-4}	Convection profonde (precipitations)	Plus ELCV est grand, moins la conversaion de l'eau nuageuse en précipitation est efficace
TLCV tlcrit	[-65 ; -35] degrés celcius!!!	-55	Convection profonde (precipitations)	Température à partir de laquelle la conversation de l'eau en précipitation est maximale. Plus elle est grande (chaude) plus le seuil maximale est atteint, plus on a de chance de précipiter efficacement
WDENSO wdens_ref_o	[1.10^{-11} ; 1.10^{-8}]	1.10^{-9}	Convection profonde (poches froides)	Nombre de poche froide par m ² sur l'océan (= densité de poche froide sur l'océan)
ALPBLK alp_bl_k	[0,2 ; 0,8]	0,5	Convection profonde	Facteur sur la puissance transmise par les thermiques à la convection profonde. Plus il est grand, plus la transmission de puissance est efficace, plus la convection est intense

Paramètre	Plage par défaut	Valeur CMIP6	Paramétrisation	Description
ALPWKK alpk	[0,01 ; 0,5]	0,25	Convection profonde	Facteur sur la puissance transmise par les poches froides à la convection profonde. Plus il est grand, plus la transmission de puissance est efficace, plus la convection est intense
SIGDZ sigdz	[0,001 ; 0,02]	0,003	Convection profonde	Facteur sur "fractionnal area of precipitating downdraughts"
WBSRF flag_wb	[0,05 ; 2] $m.s^{-1}$	0,5	Convection profonde (déclanchement)	Vitesse verticale à la base du nuage si celui-ci était à la surface. Plus WBSRF est grand plus c'est difficile de déclancher la convection profonde.
WBTOP wbmax	[1,5 ; 6]	2,8	Convection profonde (déclanchement)	Facteur sur la fonction de variation de la vitesse verticale à la base du nuage en fonction de la pression. Plus WBTOP est grand, plus c'est difficile de déclancher la convection profonde.
STRIG s_trig	[4.10^6 ; $3,6.10^7$]	$1,2.10^7$	Convection profonde (déclanchement)	Seuil pour le déclanchement stochastique de la convection profonde, qui correspond a une taille effective de thermique à la base des nuages. Plus STRIG est grand, plus c'est difficile de déclancher (moins on a de chance d'avoir un nuage qui dépasse ce seuil)

Annexe B

Emulation par processus gaussiens

Remarque lecture : Annexe bien abouti - j'ai commencé la relecture des relectures

On présente ici les mathématiques qui fondent les choix et l'utilisation des processus gaussiens dans la méthode d'History Matching with Iterative Refocusing (HMIR). Connaître ces mathématiques n'est pas indispensable pour utiliser l'outil mais aide à mieux comprendre le fonctionnement de la méthode et à gagner en confiance sur la robustesse de son utilisation.

Le choix d'émulateurs basés sur des Processus Gaussiens est double. Non seulement les processus gaussiens permettent de travailler avec une base d'apprentissage petite, typiquement les valeurs d'une métrique calculée sur une centaine de simulations ; mais encore, ils fournissent l'incertitude de la prévision statistique faite par l'émulateur.

Cette annexe ne représente pas à proprement parler un résultat de la thèse puisque qu'elle ne fait que décrire une approche développée dans la communauté de la quantification des incertitudes (Uncertainty Quantification ou UQ en anglais), approche utilisée dans le cadre de `htexplo`. En revanche, l'appropriation, par l'équipe de modélisation du climat, des principes de la construction des émulateurs, a été travaillée au cours de ma thèse dans le cadre d'ateliers avec une poignée de collaboratrices.

B.1 Cadre théorique des Processus Gaussiens

Avant de voir comment ils sont utilisés pour l'HMIR dans `htexplo`, voyons d'abord ce que sont les Processus Gaussiens : un type de processus stochastique particulier.

B.1.1 Définition d'un processus stochastique

Un processus stochastique est une variable aléatoire de type fonction, appelé aussi fonction aléatoire ou processus aléatoire. Cela signifie qu'une réalisation d'un processus stochastique est une fonction.

Les processus stochastiques F se définissent de manière équivalente par toute famille ou collection $\{F(x_i)\}_{i \in \mathcal{I}}$, où chaque $F(x_i)$ est une variable aléatoire et \mathcal{I} un ensemble d'indices qui peut être fini ou infini (par exemple l'ensemble des entiers naturels \mathbb{N}). Le processus stochastique F est ainsi défini sur toute collection de points $\{x_i\}_{i \in \mathcal{I}}$. La description de la distribution de la collection de variables aléatoires $\{F(x_i)\}_{i \in \mathcal{I}}$ dans son ensemble se fait par une densité de probabilité multidimensionnelle appelée densité de probabilité jointes. Les lois de probabilités multidimensionnelles qui peuvent être suivies par une collection de variable aléatoire sont parfois appelées lois multivariées. La distribution d'une variable aléatoire unique de la collection est appelée distribution marginale. Dans le cas où les variables aléatoires $F(x_i)$ de la collection $\{F(x_i)\}_{i \in \mathcal{I}}$ sont toutes indépendantes les unes des autres, leur densité de probabilité jointe est

égale au produit des distributions marginales de chacune des variables aléatoires de la collection. Dans le cas général, les variables aléatoires $F(x_i)$ de la collection ne sont pas indépendantes les unes des autres.

Si la loi du processus est connue, on peut l'échantillonner, c'est à dire qu'on peut produire une fonction qui est une réalisation du processus. En pratique, l'expérience permettant de construire une réalisation d'un processus stochastique consiste à échantillonner le processus pour une collection de points $\{x_i\}_{i \in \mathcal{I}_f}$. Le résultat (ou la réalisation) de cet échantillonnage est une collection finie notée $\{f(x_i)\}_{i \in \mathcal{I}_f}$, où \mathcal{I}_f est un ensemble d'indices, de taille finie, inclus dans \mathcal{I} . La fonction réalisée, f , s'appelle aussi une trajectoire. Celle-ci ne peut être en pratique connue que partiellement, sur une collection finie de points.

Par exemple, considérons une fourmi que l'on observe, et dont on modélise la trajectoire, c'est-à-dire sa position au cours du temps, comme la réalisation d'un processus stochastique noté F . La collection de points $\{x_i\}_{i \in \mathcal{I}}$ correspond alors à un ensemble de temps classés dans l'ordre croissant, auxquels on a observé la position de la fourmi. $F(x_i)$ est une variable aléatoire qui traduit la position de la fourmi au temps x_i . $f(x_i)$ est la réalisation du processus stochastique au temps x_i , c'est-à-dire la position de la fourmi au temps x_i . On voit bien dans ce processus que la position de la fourmi au temps x_{i+1} dépend de la position de la fourmi aux temps précédents, donc que les différentes $F(x_i)$ de la collection $\{F(x_i)\}_{i \in \mathcal{I}}$ ne sont pas indépendantes.

B.1.2 Définition d'un processus gaussien

Un processus stochastique G est un processus gaussien si et seulement si la loi jointe associée à toute collection finie de points issus de ce processus est la loi normale multivariée. Cette loi est la généralisation de la loi normale à une collection de variables aléatoires. Les réalisations d'un processus gaussien sont toujours des fonctions à valeur dans \mathbb{R} .

Soit $Z = [G(x_i)]_{i \in \llbracket 1; N \rrbracket}$ un vecteur regroupant les membres d'une collection de taille N . On note qu'on parlera indifféremment de la loi jointe des membres de la collection $\{G(x_i)\}_{i \in N}$ ou de la loi jointes des composantes du vecteur Z . Soit z une réalisation de Z . Si G est un processus gaussien alors Z suit une loi normale multivariée, de densité de probabilité :

$$p_Z(z) = \frac{\exp^{-\frac{1}{2}(z-\mu)^T \Sigma^{-1}(z-\mu)}}{\sqrt{(2\pi)^N \det(\Sigma)}} \quad (\text{B.1})$$

où μ , l'espérance de la loi normale multivariée, est un vecteur de taille N , et Σ , la matrice de variance covariance de la loi normale multivariée, est inversible et de taille $N \times N$. On note la loi normale multivariée $\mathcal{N}_{ND}(\mu, \Sigma)$, par analogie avec la loi normale classique $\mathcal{N}(\mu, \sigma)$, avec pour différence que σ est l'écart type de la loi normale et Σ la généralisation de la variance pour la loi normale multivariée.

Une des propriétés des lois normales multivariées est que si Z suit une loi normale multivariée $\mathcal{N}_{ND}(\mu, \Sigma)$, alors chacune de ses composantes $Z_i = G(x_i)$ suit une loi normale $\mathcal{N}(\mu_i, \sigma_i)$, avec μ_i la $i^{\text{ème}}$ composante du vecteur moyenne μ et σ_i la racine carré du $i^{\text{ème}}$ élément diagonal de la matrice de variance covariance Σ .

Pour que le processus gaussien soit bien défini pour n'importe quelle collection de points où réaliser la fonction, il faut que le vecteur d'espérance μ et la matrice de covariance Σ soient définies à partir de fonctions. La fonction espérance, notée m , associe à chaque point de la collection la valeur de son espérance, avec $\mu = [\mu_i]_{i \in \llbracket 1; N \rrbracket} = [m(x_i)]_{i \in \llbracket 1; N \rrbracket}$. La fonction de covariance, appelée aussi noyau ou opérateur de covariance (*kernel* ou *kernel covariance operator* en anglais), et notée k , associe à chaque couple de points (x, x') de $\{x_i\}_{i \in N}$ la valeur de leur covariance. Soit $k(x, x') = \mathbb{E}[(f(x) - \mu(x))(f(x') - \mu(x')))]$ et $\Sigma = [k(x_i, x_j)]_{(i,j) \in \llbracket 1; N \rrbracket \times \llbracket 1; N \rrbracket}$. Définir la loi du processus gaussien revient donc à définir :

- Une fonction espérance $m(x)$
- Un noyau de covariance $k(x, x')$

La loi du processus gaussien est souvent notée $\mathcal{GP}(m(\cdot), k(\cdot, \cdot))$ dans la littérature. Finalement, toute collection finie de points issus de ce processus suit une loi normale multivariée $\mathcal{N}_{ND}(\mu, \Sigma)$ où $\mu_i = m(x_i)$ et $\Sigma_{i,j} = k(x_i, x_j)$.

B.1.3 Apprentissage

L'étape d'apprentissage d'un processus gaussien consiste à estimer sa fonction moyenne $m(\cdot)$ et son noyau $k(\cdot, \cdot)$ à partir d'un échantillon connu. Cet échantillon est couramment appelé base d'apprentissage. Il est constitué d'un ensemble de points $\{x_i\}_{i \in \mathcal{I}_\ell}$, dit aussi points sondes ou points sondés, où les réalisations de ce que l'on souhaite apprendre avec les processus gaussiens sont connus et ont pour valeurs $\{g(x_i)\}_{i \in \mathcal{I}_\ell}$. Dans l'utilisation des processus gaussiens pour l'émulation, l'étape d'apprentissage correspond à l'étape de construction des émulateurs, qui peuvent être vu comme des interpolateurs passant par (ou proche) des points sondés.

La méthode d'HMIR construit des émulateurs des métriques du modèle, fonction des paramètres libres $\lambda^1, \dots, \lambda^P$. La base d'apprentissage de ces émulateurs est constituée d'un ensemble de simulations réalisées avec les vecteurs de paramètres de l'ensemble $\{\lambda_i\}_{i \in \mathcal{I}}$, sur lesquelles la métrique, ou l'observable, y_i est calculée. On dit aussi que le modèle est sondé aux points $\{\lambda_i\}_{i \in \mathcal{I}}$. Notons qu'aucune observation (ou cible) n'entre dans cette phase. Il s'agit uniquement de construire un interpolateur à partir des métriques simulées, qui servira à estimer cette même métrique en un nouveau point, non sondé auparavant.

Les points sondes vont ainsi être utilisés à deux reprises :

- lors de l'apprentissage, pour estimer la fonction moyenne $m(\cdot)$ et le noyau $k(\cdot, \cdot)$
- lors de la prédiction, qui implique la production d'un émulateur qui passe par (ou proche) des points sondés.

Pour estimer la fonction espérance, on ne dispose que des valeurs aux points sondés. Si on prend une fonction qui interpole parfaitement les points, et en supposant que le modèle est suffisamment complexe (ou bruité) pour qu'on ne puisse pas le représenter exactement par une combinaison linéaire de fonctions simples, on introduit à coup sûr un surajustement. Ce surajustement ne pourra pas être "corrigé" par l'émulateur, qui se base justement sur un apprentissage des statistiques des résidus (écarts entre points sondes et espérance). Puisque ces résidus seront nuls par construction, leur covariance sera parfaite et toutes les prédictions de l'émulateur seront exactement égales à la fonction espérance, avec une incertitude nulle. La probabilité que l'émulateur ait raison (que la fonction à émuler soit exactement l'espérance du processus) est extrêmement faible, mais on aura aucun moyen de détecter que l'émulateur est faux. On voit donc qu'avoir une interpolation parfaite des points sondes est entièrement contre-productif du point de vue de l'utilisation des processus gaussiens pour faire des prédictions fiables.

Deux solutions sont classiquement pratiquées dans la communauté UQ, suscitant de vifs débats entre les tenants de l'une et de l'autre (Daniel Williamson, communication personnelles). Soit on suppose qu'on ne connaît que la moyenne des y_i qu'on utilise pour spécifier $m(\cdot)$ comme une fonction constante. Ou alors, on commence par effectuer une régression classique des valeurs de y_i en fonction de x_i , en faisant attention à ne pas surcontraindre cette régression. C'est cette seconde approche qui est utilisée dans `htexplo` comme on l'explique un peu plus loin.

On choisit ensuite une forme de la fonction noyau de covariance, qui doit typiquement décroître quand les points pour lesquelles elle donne la covariance s'éloignent l'un de l'autre. On choisit dans `htexplo` la forme la plus classique à savoir un noyau exponentiel, où la corrélation

entre deux points de la collection décroît exponentiellement en fonction du carré de la distance entre les deux points, normalisée par une longueur de décorrélation. Apprendre l'émulateur consiste alors en apprendre les longueurs de décorrélation dans chacune des dimensions de x (chacun des paramètres du modèle constituant le vecteur x pour éviter les confusions).

Un exemple classique d'utilisation de cette approche est le krigeage, du nom du statisticien sud-africain D.-G. Krige. Cette approche, développée à l'origine pour la prospection minière, est également utilisée pour reconstituer des champs météorologiques, notamment des champs de pluie. Dans ces applications, l'espace des paramètres est l'espace physique 2D horizontal, sur lequel on dispose de sondages du sol en un certain nombre de points, ou d'un réseau de pluviomètres sur une surface de quelques dizaines de kilomètres de côté par exemple. Les collections sont alors constituées de point x_i à deux dimensions (a, b) , les deux dimensions d'espace jouant donc ici le rôle des P paramètres d'un modèle de climat. Cette méthode utilise, pour reconstituer le champ de pluie, des processus gaussiens d'espérance nulle et d'opérateur de covariance k défini par

$$k(x, x') = \sigma^2 \exp \left(-\frac{(a-a')^2 + (b-b')^2}{2l^2} \right) \quad (\text{B.2})$$

où σ^2 et l sont des hyper paramètres à déterminer. l correspond dans cette formule à une longueur de décorrélation : plus l est petit, plus $k(x, x')$ tend vers 0 pour $x \neq x'$ et moins x et x' sont corrélés. De la même manière dans cette formule, plus x et x' sont éloignés relativement à la distance l , plus $k(x, x')$ est petit donc moins les points x et x' sont corrélés. Lorsque $x = x'$, $k(x, x) = \sigma^2$, ce qui correspond à la variance de la loi normale suivie par $G(x)$.

B.1.4 Prédiction et incertitude de prédiction

Une fois ce cadre posé, on peut estimer la prédiction et l'incertitude de la prédiction, à savoir l'espérance et la variance de l'émulateur en un point quelconque de l'espace des paramètres, x^* .

Supposons que l'on souhaite prédire, en toute généralité, la réalisation d'un processus gaussien suivant la loi $\mathcal{GP}(m(\cdot), k(\cdot, \cdot))$ sur une collection de points $\mathbf{x}_2 = \{x_j\}_{j \in \mathcal{I}_2}$, à laquelle est associée la variable aléatoire $\mathbf{Z}_2 = [G(x_i)]_{i \in \mathcal{I}_2}$. Nous connaissons la réalisation du processus gaussien sur une collection de points $\mathbf{x}_1 = \{x_i\}_{i \in \mathcal{I}_1}$, à laquelle est associée la variable aléatoire $\mathbf{Z}_1 = [G(x_i)]_{i \in \mathcal{I}_1}$ dont une réalisation est $\mathbf{S} = [y_i]_{i \in \mathcal{I}_1}$. Les valeurs de \mathbf{S} nous ont déjà servi de base d'apprentissage des hyperparamètres de la fonction espérance m et de l'opérateur de covariance k , qui sont donc entièrement connus à ce stade. Faire cette prédiction signifie calculer, pour tout point x_j de \mathbf{x}_2 , l'espérance μ^* et la variance Σ^* de la loi d'une nouvelle variable aléatoire, qui correspond au processus pris en x_j , conditionné par les valeurs du processus réalisées sur l'ensemble des points sondes, soit

$$\forall x_j \in \mathbf{x}_2, \quad \mathbf{Z}_j^* = G(x_j) | \{G(x_i) = S_i, \forall i \in \mathcal{I}_1\}.$$

L'ensemble de ces variables aléatoires conditionnelles constitue les composantes de la variable aléatoire \mathbf{Z}^* .

On note \mathbf{x} la concaténation des collections \mathbf{x}_1 et \mathbf{x}_2 , à laquelle on associe la variable aléatoire $\mathbf{Z} = [G(x_i)]_{i \in \mathcal{I}}$, qui est la concaténation des vecteurs \mathbf{Z}_1 et \mathbf{Z}_2 . \mathbf{Z} est ainsi une nouvelle variable aléatoire vecteur, issue du même processus gaussien. Elle suit donc également une loi normale multivariée, d'espérance notée μ et de matrice de variance covariance notée Σ .

Si μ_1 et Σ_1 (respectivement μ_2 et Σ_2) sont l'espérance et la matrice de variance-covariance de \mathbf{Z}_1 (respectivement \mathbf{Z}_2), alors l'espérance et la matrice de variance-covariance de \mathbf{Z} s'écrivent :

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad (\text{B.3})$$

et

$$\Sigma = \begin{pmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{pmatrix} \quad (\text{B.4})$$

On cherche donc à déterminer la probabilité de la variable aléatoire “ \mathbf{Z}_2 sachant \mathbf{Z}_1 ”, $\mathbf{Z}^* = \mathbf{Z}_2 | (\mathbf{Z}_1 = \mathbf{S})$. Pour ce faire, on utilise le théorème de Bayes, à savoir que pour deux variables aléatoires A et B , la probabilité de A sachant B , c’est à dire, la probabilité que l’évènement $A = a$ se réalise sachant que l’évènement $B = b$ a été réalisé, et couramment notée $P(A|B = b)$, est donnée par :

$$P(A|B = b) = \frac{P(A \cap B)}{P(B)} \quad (\text{B.5})$$

Notons que la notation $A \cap B$ est couramment utilisée pour signifier l’intersection des deux évènements : $A = a$ et $B = b$. C’est donc la probabilité de l’évènement “les deux évènements $A = a$ et $B = b$ sont réalisés simultanément” qui apparait au numérateur. Le théorème de Bayes se généralise aux fonctions de densité de probabilité. On applique l’Equation B.5 sur les densités de probabilités avec $A = \mathbf{Z}_2$, $B = \mathbf{Z}_1$, b la réalisation de \mathbf{Z}_1 égale \mathbf{S} et $A \cap B = \mathbf{Z}$, qui désigne l’évènement “ $\mathbf{Z}_2 = \mathbf{z}_2$ et $\mathbf{Z}_1 = \mathbf{S}$ ”.

$$p_{\mathbf{Z}^*}(\mathbf{z}^*) = p_{\mathbf{Z}_2 | (\mathbf{Z}_1 = \mathbf{S})}(\mathbf{z}_2) = \frac{p_{\mathbf{Z}}(\mathbf{z})}{p_{\mathbf{Z}_1}(\mathbf{S})} \quad (\text{B.6})$$

où \mathbf{z} est une réalisation de \mathbf{Z} , soit la concaténation des vecteurs \mathbf{S} et \mathbf{z}_2 . Les densités de probabilité des lois normales multivariées de \mathbf{Z} et \mathbf{Z}_1 qui apparaissent dans l’Equation B.6 sont données par l’Equation B.1, soit plus précisément :

$$p_{\mathbf{Z}}(\mathbf{z}) = \frac{\exp^{-\frac{1}{2}(\mathbf{z}-\mu)^T \Sigma^{-1}(\mathbf{z}-\mu)}}{\sqrt{(2\pi)^N \det(\Sigma)}} \text{ et } p_{\mathbf{Z}_1}(\mathbf{S}) = \frac{\exp^{-\frac{1}{2}(\mathbf{S}-\mu_1)^T \Sigma_1^{-1}(\mathbf{S}-\mu_1)}}{\sqrt{(2\pi)^{N_1} \det(\Sigma_1)}} \quad (\text{B.7})$$

On substitue $p_{\mathbf{Z}}(\mathbf{z})$ et $p_{\mathbf{Z}_1}(\mathbf{S})$ par leurs expressions de l’Equation B.7 dans l’Equation B.6 et on aboutit à

$$p_{\mathbf{Z}^*}(\mathbf{z}^*) = \frac{\exp^{-\frac{1}{2}[(\mathbf{z}-\mu)^T \Sigma^{-1}(\mathbf{z}-\mu) - (\mathbf{S}-\mu_1)^T \Sigma_1^{-1}(\mathbf{S}-\mu_1)]}}{\sqrt{(2\pi)^{N_2} \det(\Sigma) / \det(\Sigma_1)}} \quad (\text{B.8})$$

Une dérivation mathématique relativement simple permet de réécrire cette distribution de probabilité comme

$$p_{\mathbf{Z}^*}(\mathbf{z}^*) = \frac{\exp^{-\frac{1}{2}(\mathbf{z}^* - \mu^*)^T \Sigma^{*-1}(\mathbf{z}^* - \mu^*)}}{\sqrt{(2\pi)^N \det(\Sigma^*)}} \quad (\text{B.9})$$

avec :

$$\mu^* = \mu_2 + \Sigma_{21} \Sigma_1^{-1}(\mathbf{S} - \mu_1) \quad (\text{B.10})$$

et

$$\Sigma^* = \Sigma_2 - \Sigma_{21} \Sigma_1^{-1} \Sigma_{12} \quad (\text{B.11})$$

ce qui montre que \mathbf{Z}^* suit également une loi normale multivariée d’espérance μ^* et de matrice de variance covariance Σ^* définies respectivement par l’Equation B.10 et l’Equation B.11, et dont une réalisation est notée \mathbf{z}^* . μ^* est ainsi la prédiction ou l’émulation sur la collection de point \mathbf{x}_2 et Σ^* la matrice de variance covariance associée à cette prédiction. *[Naj : mettre en gras les vecteurs variance et matrice var-covar ?] [FH : je ne crois pas. Tout va être en gras sinon.]*

B.2 Mise en œuvre dans htexplo

Dans la méthode d’History Matching, on travaille métrique par métrique. On note g cette métrique définie sur un espace Ω de vecteur de paramètres λ . Pour chaque métrique, on connaît un ensemble de N valeurs (N nombres réels) calculées sur les résultats de N simulations réalisées avec le modèle de climat sur N vecteurs λ tirés aléatoirement dans Ω . On prend en général $N = 10 \times P$ où P est le nombre de paramètres. Dans le processus d’HMIR, Ω est l’hyper-cube initial Λ ou l’espace NROY de la vague précédente.

Pour chaque métrique, on construit indépendamment un émulateur par processus gaussien sur l’espace hypercubique des paramètres de tuning Λ . On apprend pour cette métrique les hyperparamètres du processus gaussien $G = \mathcal{GP}(m(\cdot), k(\cdot, \cdot))$ sur l’ensemble des N valeurs sondées pour la métrique, prises comme une réalisation partielle (une collection) du processus gaussien, appelée \mathbf{S} (étape 2). Toute la subtilité dans la définition des processus gaussiens réside dans la définition de la fonction espérance et de l’opérateur de covariance, et dans le choix de leurs hyperparamètres (étape d’apprentissage). Les choix faits pour la méthode d’History Matching utilisée dans l’outil `htexplo` sont résumés dans Couvreur et al. (2021) et détaillés entre autre dans Williamson et al. (2017), Williamson et al. (2013) et Volodina and Challenor (2021).

B.2.1 Apprentissage

L’estimation de l’espérance du processus est obtenue par régression sur l’ensemble d’apprentissage, et rend compte des relations entre les paramètres et la métrique sur tout l’espace des paramètres. Elle est définie comme une combinaison linéaire de fonctions simples des vecteurs de paramètres, regroupées dans un vecteur noté $\mathbf{h}(\lambda)$. Les hyperparamètres sont les poids de la combinaison linéaire, regroupés dans un vecteur noté β . La fonction espérance s’exprime ainsi selon l’Equation B.12.

$$m(\lambda, \beta) = \beta^T \mathbf{h}(\lambda) \quad (\text{B.12})$$

Les fonctions composant $\mathbf{h}(\lambda)$ sont par exemple des monômes, des fonctions de la base de Fourier et des fonctions d’interactions entre les paramètres [Quelqu’un : je ne sais pas exactement ce qu’on utilise dans `htexplo`. Naj : c’est une fonction classique de R qui est appelée pour la régression je crois]. Pour déterminer lesquelles de ces fonctions auront un poids non nul dans l’espérance du processus, une méthode de sélection itérative (“stepwise selection method”) est utilisée, consistant en des allers-retours où les fonctions sont successivement ajoutées et supprimées de la combinaison linéaire (“forwards selection and backwards elimination”), afin de mesurer leur capacité à réduire le résidu total (écart des points sondes à la régression) et à expliquer la variabilité [Quelqu’un : je crois que c’est mal dit. Naj : ouais pas ouf mais j’ai pas mieux : D j’ai pas compris “expliquer la variabilité”?]. La méthode de sélection des fonctions est détaillée dans Williamson et al. (2013).

Une fois l’espérance m du processus déterminée, on travaille à partir des valeurs résiduelles, c’est à dire la distance entre les valeurs mesurées (ou simulées) aux points sondes, et la fonction espérance évaluée à ces mêmes points sondes.

Pour l’opérateur de covariance k , on utilise une fonction exponentielle dont les hyperparamètres sont appris sur les résidus. Ce noyau quantifie donc la corrélation des résidus entre deux points λ et λ' de l’espace des paramètres :

$$k(\lambda, \lambda') \equiv k(\lambda, \lambda'; \delta, \sigma, \nu) = \sigma^2 \left[\nu \mathbb{1}_{\lambda=\lambda'} + (1 - \nu) \exp \left\{ - \sum_{i=1}^N \left(\frac{\lambda_i - \lambda'_i}{\delta_i} \right)^2 \right\} \right] \quad (\text{B.13})$$

L'hyperparamètre ν est appelé pépite ou *nugget* en anglais. Il permet d'ajouter à l'opérateur de variance-covariance une part d'incertitude irréductible, qui tient compte du fait qu'une observable simulée en deux points infiniment proches de l'espace des paramètres peut prendre des valeurs différentes. Cela peut être soit parce que la fonction que l'on cherche à représenter avec le processus est elle-même aléatoire : auquel cas plusieurs sondages même réalisés strictement au même point donneraient des observations différentes. Soit parce que la fonction que l'on cherche à représenter a une variabilité non-nulle aux très petites échelles (voire des points de singularité), qui sera toujours sous-estimée lorsque l'on apprend les longueurs de décorrélation à partir des points sondés. Dans nos applications climatiques, on ne s'attend plutôt pas à rencontrer des discontinuités ou des signaux de très hautes fréquences dans l'espace des paramètres. On utilise en revanche un nugget non-nul pour tenir compte du fait que le modèle (la fonction que l'on cherche à émuler) est bruité. Par exemple, la variabilité interne du système climatique fait que deux simulations réalisées strictement au même point de l'espace des paramètres peuvent aboutir à des valeurs de métriques différentes si on change l'état initial de la simulation. On verra qu'avec les processus gaussiens, si on réalise une prédiction strictement au même point que là où on a sondé le modèle, la prédiction sera la même indépendamment de la valeur du nugget (nul ou non-nul). Le nugget ne permet donc pas, en théorie, de tenir compte du fait que le modèle est bruité. Toutefois, en pratique, quand on utilise les émulateurs pour faire des prédictions, c'est toujours pour des points que l'on a échantillonné dans l'espace des paramètres. On a donc une probabilité strictement nulle de tomber sur un point sondé. On peut par contre être très proche d'un point sondé, et utiliser une valeur de nugget non nulle assure que même dans ce cas (infiniment proche d'un point sonde), il reste une part d'incertitude dans la prédiction.

Le deuxième terme est exprimé comme fonction exponentielle puissance (ici puissance 2) de la distance entre les deux paramètres λ et λ' , dépendant de l'hyperparamètre δ . δ_i est la longueur de décorrélation dans la direction i , c'est-à-dire la distance à partir de laquelle les résidus aux points λ et λ' ne sont plus corrélés qu'à 37% ($\exp(-1)$) pour la $i^{\text{ème}}$ dimension de l'hypercube. *[On est sûr que c'est d^2/l^2 et pas sur $2 l^2$? Si jamais c'est sur 2 , alors au bout de la distance l on est corrélé à $\exp(-1/2) = 60\%$]* On fait l'hypothèse dans cette formulation que la décorrélation des résidus ne dépend que de la distance entre les vecteurs de paramètres, et non pas de leur position absolue dans l'espace des paramètres. On dit dans ce cas que l'opérateur de covariance est stationnaire.

La construction des émulateurs consiste donc à apprendre les fonctions $h(\lambda)$ et les hyperparamètres β , δ , σ et ν . Ces hyperparamètres sont appris en utilisant l'algorithme de Monte Carlo hamiltonien implémenté dans le programme Stan Carpenter et al. (2017).

B.2.2 Prédiction

Une fois les émulateurs construits, on souhaite prédire la valeur de la métrique en un nouveau point λ^* . On utilise alors les équations B.10 et B.11 avec les modèles d'espérance et de covariance définis par les équations B.12 et B.13. On considère une collection de $N+1$ vecteurs de paramètres, où les N premiers vecteurs sont les points sondés et le dernier est le vecteur de paramètres auquel on souhaite estimer la métrique.

La matrice Σ_1 est donc de dimension $N \times N$:

$$\Sigma_1 = \begin{pmatrix} k(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_1) & k(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) & \cdot & \cdot & \cdot & k(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_N) \\ k(\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_1) & k(\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_2) & & & & k(\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_N) \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ k(\boldsymbol{\lambda}_N, \boldsymbol{\lambda}_1) & k(\boldsymbol{\lambda}_N, \boldsymbol{\lambda}_2) & \cdot & \cdot & \cdot & k(\boldsymbol{\lambda}_N, \boldsymbol{\lambda}_N) \end{pmatrix} \quad (\text{B.14})$$

A noter que les éléments de Σ_1 sont fixés une fois pour toute après la phase d'apprentissage des hyper-paramètres. Les termes non diagonaux de la matrice Σ sont ici des vecteurs

$$\Sigma_{12}(\boldsymbol{\lambda}^*) = \Sigma_{21}(\boldsymbol{\lambda}^*)^T = \begin{pmatrix} k(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}^*) \\ k(\boldsymbol{\lambda}_2, \boldsymbol{\lambda}^*) \\ \cdot \\ \cdot \\ \cdot \\ k(\boldsymbol{\lambda}_N, \boldsymbol{\lambda}^*) \end{pmatrix} \quad (\text{B.15})$$

et le dernier élément de la matrice est

$$\Sigma_2 = k(\boldsymbol{\lambda}^*, \boldsymbol{\lambda}^*) = \sigma^2 \quad (\text{B.16})$$

La prédiction (espérance de la variable aléatoire conditionnelle) au point $\boldsymbol{\lambda}^*$ est notée μ^* et l'incertitude σ^* avec :

$$\mu^*(\boldsymbol{\lambda}^*) = m^*(\boldsymbol{\lambda}^*, \boldsymbol{\beta}) = m(\boldsymbol{\lambda}^*, \boldsymbol{\beta}) + \Sigma_{21}(\boldsymbol{\lambda}^*) \Sigma_1^{-1} \begin{pmatrix} S_1 - m(\boldsymbol{\lambda}_1, \boldsymbol{\beta}) \\ \dots \\ S_N - m(\boldsymbol{\lambda}_N, \boldsymbol{\beta}) \end{pmatrix} \quad (\text{B.17})$$

et

$$\sigma^{*2}(\boldsymbol{\lambda}^*) = \sigma^2 - \Sigma_{21} \Sigma_1^{-1} \Sigma_{12} \quad (\text{B.18})$$

On peut regarder ces formules pour trois cas limites.

B.2.2.1 Si le point auquel on réalise la prédiction est strictement un des points sondés

c'est à dire si $\boldsymbol{\lambda}^* = \boldsymbol{\lambda}_j$. Alors Σ_{12} est la colonne j de Σ_1 , c'est à dire que $\Sigma_{12} = \Sigma_1 \mathbf{e}_j$ où \mathbf{e}_j est le vecteur unitaire (0 partout sauf 1 à la ligne j). On a donc $\Sigma_1^{-1} \Sigma_{12} = \Sigma_1^{-1} \Sigma_1 \mathbf{e}_j = \mathbf{e}_j$. Et comme Σ_{21} est la ligne j de Σ_1 , on a

$$\Sigma_{21} \Sigma_1^{-1} \Sigma_{12} = \Sigma_{21} \mathbf{e}_j = \mathbf{k}(\boldsymbol{\lambda}_j, \boldsymbol{\lambda}_j).$$

De la même façon, on a $\Sigma_{21} \Sigma_1^{-1} = \Sigma_{12}^T \Sigma_1^{-1} = \mathbf{e}_j^T \Sigma_1^T \Sigma_1^{-1} = \mathbf{e}_j^T$ par symétrie de la matrice Σ_1 .

Dans ce cas, comme les distances sont nulles dans l'exponentielle, le second terme de l'Equation B.13 est égal à $(1 - \nu)$, et comme on est au point l'indicatrice du premier terme vaut 1, donc au total, indépendamment de la valeur du nugget, on a $k(\boldsymbol{\lambda}_j, \boldsymbol{\lambda}_j) = \sigma^2$, et donc

$$\mu^*(\boldsymbol{\lambda}^* = \boldsymbol{\lambda}_j) = S_j \quad (\text{B.19})$$

et

$$\sigma^{*2}(\boldsymbol{\lambda}^* = \boldsymbol{\lambda}_j) = 0 \quad (\text{B.20})$$

On retrouve bien que le nugget ne permet pas, en théorie, de représenter un modèle bruité : les prédictions passent strictement par les points sondes, avec une incertitude nulle.

B.2.2.2 Si le point auquel on réalise la prédiction est infiniment proche de l'un des points sondes

c'est à dire quand $\boldsymbol{\lambda}^* \rightarrow \boldsymbol{\lambda}_j$. Alors, on voit que les vecteurs Σ_{12} et Σ_{21} tendent vers la colonne (resp. ligne) j de la matrice Σ_1 , puisque les exponentielles sont des fonctions continues. On peut donc réutiliser le raisonnement précédent pour dire qu'à la limite où $\boldsymbol{\lambda}^* \rightarrow \boldsymbol{\lambda}_j$, $\Sigma_{21}\Sigma_1^{-1}\Sigma_{12} \rightarrow k(\boldsymbol{\lambda}_j, \boldsymbol{\lambda}_j)$. Par contre, comme on n'est pas strictement au point, cette fois-ci l'indicateur dans le premier terme de l'Equation B.13 est nul. Il reste donc un terme faisant apparaître le nugget dans l'expression de $k(\boldsymbol{\lambda}^*, \boldsymbol{\lambda}_j)$, qui fait que $k(\boldsymbol{\lambda}^*, \boldsymbol{\lambda}_j) \rightarrow \sigma^2(1 - \nu)$. On retrouve donc que, en l'absence de nugget, i. e. $\nu = 0$, $k(\boldsymbol{\lambda}^*, \boldsymbol{\lambda}_j) \rightarrow \sigma^2$ donc $\sigma^* \rightarrow 0$. Et si la valeur du nugget n'est pas nulle, $\sigma^* \rightarrow \nu\sigma^2$. On voit donc qu'en pratique, grâce au nugget, quand on fera des prédictions même très proches de points sondes, l'incertitude de la prédiction ne sera jamais strictement nulle. Le nugget représente alors la part de variabilité de l'échantillon sondé qui est due à du bruit (par exemple de la variabilité interne du système si on a perturbé les conditions initiales en même temps que les valeurs de paramètres).

B.2.2.3 Si le point auquel on réalise la prédiction est très loin de tous les points sondés

relativement aux longueurs de décorrélation (hyperparamètres δ_i), alors la prédiction tend vers l'espérance du processus gaussien (prédiction "à priori", les points sondes n'apportent aucune information), et l'incertitude associée tend vers la variance du processus σ^2 .

B.2.3 Vérification

A la fin de l'étape de construction des émulateurs, le test standard Leave One Out (LOO) est réalisé afin de vérifier les bonnes capacités de prédiction des émulateurs (voir par exemple Rougier et al. (2009)). Ce test consiste à retirer un point $\boldsymbol{\lambda}_j$ de l'ensemble d'apprentissage et à évaluer en ce point la prédiction $\mu^*(\boldsymbol{\lambda}_j)$ et l'incertitude autour de sa prédiction $\sigma^*(\boldsymbol{\lambda}_j)$.

La version du test LOO utilisé dans l'outil `htexplo` réalise un test partiel : les hyperparamètres du processus gaussien ne sont pas réappris. Seules les équations B.17 et B.18 sont réévaluées sans le point $\boldsymbol{\lambda}_j$. On ne teste donc pas la robustesse de l'apprentissage (est ce que les hyperparamètres optimaux sont différents si on modifie un peu la base d'apprentissage) mais la qualité des prédictions. On les teste là où on a des références pour évaluer ces prédictions, c'est à dire là où on a des points sondes, plutôt que de réaliser des vérifications sur une base de test indépendante qui n'aurait jamais été vue par le processus. Cela est dû au coût potentiellement très élevé des sondages.

Puisque la variable aléatoire conditionnelle (dont l'espérance est la prédiction) suit une loi normale, on peut construire un intervalle de confiance à 2σ autour de la prédiction, indiquant que la probabilité que la fonction que l'on cherche à émuler, évaluée en $\boldsymbol{\lambda}_j$ a 95% de chance d'être dans l'intervalle $[\mu^*(\boldsymbol{\lambda}_j) - 2\sigma^*(\boldsymbol{\lambda}_j), \mu^*(\boldsymbol{\lambda}_j) + 2\sigma^*(\boldsymbol{\lambda}_j)]$. Pour tester le bon comportement du modèle, un calcule les statistiques pour chaque point de l'ensemble d'apprentissage, en enlevant à chaque fois ce seul point de l'ensemble. Si les émulateurs ont bien été construits et ont des bonnes capacités prédictives, alors environ 5% des points vont se retrouver en dehors de l'intervalle $[\mu^*(\boldsymbol{\lambda}_j) - 2\sigma^*(\boldsymbol{\lambda}_j), \mu^*(\boldsymbol{\lambda}_j) + 2\sigma^*(\boldsymbol{\lambda}_j)]$. Un exemple de figures LOO sorties automatiquement par l'outil `htexplo` est montrée Figure III.8.

