



HAL
open science

Modélisation des couches limites stables en Antarctique : calibration et analyse de sensibilité de la paramétrisation de la turbulence du modèle ARPEGE-Climat

Olivier Audouin

► **To cite this version:**

Olivier Audouin. Modélisation des couches limites stables en Antarctique : calibration et analyse de sensibilité de la paramétrisation de la turbulence du modèle ARPEGE-Climat. Sciences de la Terre. Institut National Polytechnique de Toulouse - INPT, 2021. Français. NNT : 2021INPT0056 . tel-04170839

HAL Id: tel-04170839

<https://theses.hal.science/tel-04170839v1>

Submitted on 25 Jul 2023

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

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



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (Toulouse INP)

Discipline ou spécialité :

Océan, atmosphère, climat

Présentée et soutenue par :

M. OLIVIER AUDOUIN

le lundi 31 mai 2021

Titre :

Modélisation des couches limites stables en Antarctique: calibration et analyse de sensibilité de la paramétrisation de la turbulence du modèle ARPEGE-Climat.

Ecole doctorale :

Sciences de l'Univers de l'Environnement et de l'Espace (SDU2E)

Unité de recherche :

Groupe de Météorologie Expérimentale et Instrumentale (CNRM-GMEI)

Directeur(s) de Thèse :

MME FLEUR COUVREUX

M. ROMAIN ROEHRIG

Rapporteurs :

M. FREDERIQUE CHERUY, CNRS PARIS

M. HUBERT GALLEE, UNIVERSITE GRENOBLE ALPES

Membre(s) du jury :

M. SERGE CHAUZY, UNIVERSITE TOULOUSE 3, Président

M. AURELIEN RIBES, METEO FRANCE TOULOUSE, Invité(e)

M. ETIENNE VIGNON, CNRS PARIS, Membre

M. JEAN-LUC REDELSPERGER, CNRS, Membre

MME FLEUR COUVREUX, METEO FRANCE TOULOUSE, Membre

M. ROMAIN ROEHRIG, METEO FRANCE TOULOUSE, Membre

Remerciements

Je remercie chaleureusement Romain Roehrig et Fleur Couvreur pour avoir accepté de diriger ces travaux de thèse et pour avoir réussi l'exploit de me pousser à les mener à leur terme, ce qui n'était pas une mince affaire.

Je tiens également à remercier Aurélien Ribes qui, avec Romain, est à l'initiative de cette de thèse, même si le projet initial a été quelque peu modifié. Merci aussi à Hervé Douville pour son accueil dans feu l'équipe AMACS.

Merci à Nadia Fourrié et Jean-François Mahfouf pour leur accueil dans l'équipe OBS et pour m'avoir permis de finir de rédiger ce manuscrit. Merci en particulier à Jean-François qui a pris le temps de le relire et pour ses remarques pertinentes.

Merci à Thomas et Matthias, avec qui j'ai largement partagé mes états d'âme d'apprenti-chercheur.

Merci à Vincent qui a été d'un grand secours pendant les derniers mois de rédaction et m'a facilité la vie à bien des égards et à Hélène pour la traditionnelle séance de grimpe du jeudi après-midi qui a été une soupape de sécurité indispensable pendant ces trois années et demi.

De façon générale, un grand merci aux copains historiques Jo & Julie, Lucie & Adrien, et aux petits jeunes Axel, Marine, Alice, aux pongistes du midi Hugo, Lucas, Marc, Thibault, Damien, Mayeul et j'en oublie ...

Un merci particulier à mes co-bureaux successifs Josué et Yoann pour leur aide et leur amitié.

Enfin, *last but not least* un énorme merci à Noam et à Maud pour tout ce qu'ils m'apportent jour après jour, et à mes parents sans qui je n'en serais pas là aujourd'hui (qu'ils ne sous-estiment pas leur importance dans mon long parcours d'éternel étudiant).

À Delphine

Table des matières

Table des sigles et acronymes	xiii
1 Introduction	5
2 L'Antarctique et les couches limites stables	9
2.1 L'Antarctique, description et importance dans le système climatique	9
2.1.1 Description géophysique	9
2.1.2 Le climat Antarctique	11
2.1.3 Le climat à Dôme C	15
2.1.4 Rôle dans le système climatique planétaire	17
2.1.5 Changement climatique et Antarctique	19
2.2 Les couches limites stables	22
2.2.1 La couche limite atmosphérique	22
2.2.2 Les couches limites stables	26
3 Outils de modélisation	31
3.1 Modélisation du climat	31
3.2 Le modèle ARPEGE-Climat	32
3.2.1 Présentation du modèle	32
3.2.2 La paramétrisation de la turbulence dans ARPEGE-Climat 6.3	34
3.2.3 Autres paramétrisations	37
3.3 Représentation de la surface : la plateforme SURFEX	37
3.3.1 La représentation du manteau neigeux dans SURFEX	38
3.3.2 Flux turbulents sur les surfaces continentales	40
3.4 Différents types de simulations	42
3.4.1 SCM	42

3.4.2	Simulations 3D	42
3.5	Données de validation	44
3.5.1	Les jeux de données de référence	44
3.5.2	Simulations LES	45
3.6	Conclusion	47
4	Évaluation de ARPEGE-Climat 6.3 sur l’Antarctique	49
4.1	Modéliser le climat antarctique	50
4.1.1	Observations et réanalyses sur l’Antarctique	50
4.1.2	Les difficultés des modèles de climat pour représenter la région antarctique	51
4.2	Évaluation de ARPEGE-Climat 6.3 en configuration AMIP	52
4.2.1	Été austral (DJFM)	52
4.2.2	Hiver austral (JJAS)	57
4.3	Simulation guidée en vent sur l’Antarctique	59
4.3.1	Évaluation de la simulation guidée en vent	59
4.3.2	Sensibilité de la simulation guidée en vent à l’albédo	62
4.4	Évaluation sur Dôme C	65
4.5	Conclusion	69
5	Calibration d’un modèle de climat - Calibration statistique	71
5.1	Calibration de la physique d’un modèle	72
5.1.1	Généralités	72
5.1.2	Problèmes et stratégies	74
5.1.3	La calibration de ARPEGE-Climat 6.3	77
5.2	Vers une nouvelle stratégie. Outils statistiques pour une calibration « orientée processus »	78
5.2.1	La comparaison SCM/LES	80
5.2.2	<i>History Matching</i> et calibration d’un SCM	81
5.3	Conclusions	86

6	Modéliser la couche limite très stable de GABLS4 avec la physique de ARPEGE-Climat	89
6.1	Le problème de la représentation des CLS dans les modèles numériques	90
6.2	Résumé	91
6.3	L'article	93
6.4	Compléments	131
6.4.1	Améliorer la visualisation de l'espace NROY	131
6.4.2	Impact des nouveaux réglages sur le cycle diurne de GABLS4	131
6.5	Conclusions	137
7	Extension de la calibration au cas GABLS1	139
7.1	Analyse de la calibration d'ARPEGE-Climat sur le cas GABLS1	139
7.1.1	Présentation du cas GABLS1	140
7.1.2	Évaluation sur le cas GABLS1	140
7.2	Nouvelle calibration du modèle combinant les cas GABLS4 et GABLS1	142
7.3	Analyse de la nouvelle calibration sur le cas GABLS1	143
7.4	Conclusion	146
8	Conclusions et perspectives	149
8.1	L'Antarctique	149
8.2	Calibration de ARPEGE-Climat sur un cas de couche limite stable	150
8.3	Évaluation de la méthode de calibration par <i>History Matching</i>	152
8.4	Stratégie de calibration d'ARPEGE-Climat	153
	Bibliographie	155

Table des figures

2.1	Carte de l'Antarctique présentant les principales régions géographiques du continent.	10
2.2	Différentes estimations du bilan d'énergie moyen de l'Antarctique sur la période 2001-2010 : $F_{\text{SFC :NET}}$ représente le bilan d'énergie à la surface, $F_{\text{TOA :NET}}$ le bilan d'énergie au sommet de l'atmosphère et F_{WALL} est le flux horizontal d'énergie intégré verticalement sur toute la hauteur de l'atmosphère. (Previdi et al., 2013). Les valeurs positives du bilan correspondent à un gain d'énergie pour l'atmosphère.	12
2.3	Température de surface en moyenne annuelle a), en moyenne hivernale b) et en moyenne estivale c). Ces données de températures sont estimées à partir de données satellites MODIS (Fréville et al., 2014)	13
2.4	Lignes de flux moyennes à a) 100 m et b) 500 m au dessus du sol sur la période juin 2003-mai 2004 au dessus de l'Antarctique, simulées par le modèle de prévision numérique <i>Antarctic Mesoscale Prediction System</i> (Parish and Bromwich, 1987).	14
2.5	L'indice d'Oscillation Antarctique (Gong and Wang, 1999). a) Première EOF du champ de pression au niveau de la mer. b) Coefficients de corrélation entre les moyennes zonales du champ de pression. c) Pourcentage de la variance expliquée du champ de pression par l'AOI	15
2.6	Corrélations entre l'indice SAM et la température de l'atmosphère à 2 m en Antarctique pour la période 1957–2004 : (a) automne ; (b) hiver ; (c) printemps et (d) été. Marshall (2007)	16
2.7	Série temporelle de la température mensuelle climatologique à Dôme C, de 2001 à 2010 à environ 180 cm au-dessus de la surface, mesurée par la station automatique AMRC située à 3 km de la tour instrumentée (Genthon et al., 2013) 18	18
2.8	Cycles diurnes moyens mesurés à Dôme C entre 2011 et 2016 de température, de vent et de flux radiatifs (Vignon, 2017)	18
2.9	Changement de température à 2 m en Antarctique dû à la tendance de l'indice SAM pour la période 1957–2004 : (a) automne ; (b) hiver ; (c) printemps et (d) été. Marshall (2007)	21
2.10	Cycle diurne dans la couche limite (Stull, 1988). Cycle diurne et TKE. . . .	24
2.11	Évolution typique de la structure de la couche limite atmosphérique sur un cycle diurne (Stull, 1988).	24

2.12	Profils verticaux moyens de température, de température potentielle, de rapport de mélange et de vent le jour vers 15 heures locales et la nuit vers 3 heures. G représente la valeur du vent géostrophique et h représente la hauteur de la couche limite. Rodier (2017) d'après Stull (1988).	25
3.1	Résolution verticale du modèle ARPEGE-Climat 6.3 (en bleu) et de sa version précédente, ARPEGE-climat 5.1 (en rouge).	33
3.2	Le modèle SURFEX	38
3.3	Profil vertical du coefficient de nudging en h^{-1}	44
4.1	Températures à 2 m moyennes en été austral (DJFM) dans la référence BEST (à gauche), ARPEGE-Climat 6.3 (au centre) avec le biais de ARPEGE-Climat 6.3 par rapport à BEST à droite. La première ligne présente la température moyenne, la deuxième ligne, la température maximale, et la troisième ligne, la température minimale. Le point noir sur les différentes cartes présentées dans ce chapitre représente la position de Dôme C)	53
4.2	Températures de surface (en ciel clair) moyennes en été austral (DJFM) sur la période 2000-2016 avec de gauche à droite, la référence (MODIS), ARPEGE-Climat 6.3 et le biais de ARPEGE-Climat 6.3 par rapport à MODIS.	54
4.3	Biais de la fraction nuageuse totale en été austral (DJFM) dans ARPEGE-Climat 6.3 par rapport à CALIPSO. À noter que du fait de l'inclinaison de l'orbite du satellite, aucune mesure de la fraction nuageuse n'est disponible au sud de 82° S	55
4.4	Biais de l'effet radiatif des nuages à la surface en été austral (DJFM) dans ARPEGE-Climat 6.3 par rapport à CERES avec de gauche à droite l'effet radiatif des nuages dans le spectre solaire, l'effet radiatif des nuages dans le spectre infrarouge, et l'effet radiatif net (solaire+infrarouge) des nuages.	56
4.5	De gauche à droite : biais en été austral (DJFM) du flux solaire descendant à la surface, du flux solaire net à la surface et de l'albédo de surface de ARPEGE-Climat 6.3 par rapport à CERES.	56
4.6	Comparaison du vent à 10 m en été austral (DJFM) dans ARPEGE-Climat 6.3 (au centre) avec le vent à 10 m dans ERA5 (à gauche). À droite le biais de ARPEGE-Climat 6.3 par rapport à ERA5. Sur la première ligne est représenté le vent à 10 m, sur la deuxième ligne le module du vent à 10 m.	57
4.7	Biais de ARPEGE-Climat 6.3 en hiver austral (JJAS) pour la température à 2 m (par rapport à BEST – en haut à gauche), la fraction nuageuse totale (par rapport à CALIPSO – en haut à droite), les effets radiatifs infrarouge des nuages en surface (par rapport à CERES – en bas à gauche) et le flux infrarouge net en surface (par rapport à CERES – en bas à droite).	58

4.8	Comparaison du vent à 10 m en hiver austral (JJAS) dans ARPEGE-Climat 6.3 (au centre) avec le vent à 10 m dans ERA5 (à gauche). À droite le biais de ARPEGE-Climat 6.3 par rapport à ERA5. Sur la première ligne est représenté le vent à 10 m, et sur la deuxième ligne le module du vent.	59
4.9	Comparaison des biais de température des simulations AMIP (à gauche) et guidée en vent (à droite) de ARPEGE-Climat 6.3 en été austral (DJFM). . .	61
4.10	Comparaison des biais de vent des simulations AMIP (à gauche) et guidée en vent (à droite) de ARPEGE-Climat 6.3 en été austral (DJFM).	62
4.11	Comparaison du biais des albédos de surface avant correction (à gauche) et après correction (à droite) par rapport à l'albédo CERES	63
4.12	Impact de la correction de l'albédo sur les champs de température des simulations guidées en vent de ARPEGE-Climat 6.3 en été austral (DJFM), avec à gauche les résultats avant correction et à droite les résultats après correction.	64
4.13	Impact de la correction de l'albédo sur les champs de vent des simulations guidées en vent de ARPEGE-Climat 6.3 en été austral (DJFM), avec à gauche les résultats avant correction et à droite les résultats après correction.	65
4.14	Comparaison des différentes simulations avec les observations <i>in situ</i> sur la période allant du 15 au 30 Janvier 2015.	66
4.15	Cycle diurne moyen de la température de surface sur le mois de janvier 2015	67
4.16	Cycle diurne moyen des différents éléments du bilan radiatif de la surface sur le mois de janvier 2015, avec le rayonnement solaire sur la première ligne et infrarouge sur la deuxième ligne. La première colonne correspond au rayonnement descendant et la deuxième au rayonnement montant	68
4.17	Profils verticaux moyens pour janvier 2015 de vent (à droite) et de température (à gauche) à 1100LT (en haut) et à 2300LT (en bas).	70
5.1	Exemple d'hypercube latin de $n = 5$ points en dimension $d = 2$, extrait de (Franco, 2008)	83
6.1	Représentation de l'espace des paramètres à l'issue de la neuvième itération de la méthode de tuning statistique pour une gamme initiale des paramètres L MIN et K OZMIN réduite. Les noms des paramètres apparaissant sur la figure sont les noms des variables dans le code : AKN = CM, ALD = CE, ALMAVE = LMIN, ALPHAЕ = AE, ALPHAT = AT, KROVMIN = KOZMIN, KROVZMAX = ZMAX	132

6.2	Évolution de la température potentielle (colonne de gauche et du module du vent (colonne de droite) sur de haut en bas les trois premiers niveaux du modèle (respectivement ~ 8 m, ~ 29 m et ~ 55 m). Les lignes oranges correspondent aux simulations issues de la calibration et les lignes rouges à la simulation CNRM-CM6.	134
6.3	Évolution du flux de chaleur sensible à la surface (à gauche) et de la vitesse de friction (à droite). Les lignes oranges correspondent aux simulations issues de la calibration et les lignes rouges à la simulation CNRM-CM6.	135
6.4	Comparaison des tendances de température (première colonne), du vent zonal (au centre) et du vent méridien (à droite) sur GABLS4-Stage-3. La première ligne présente les tendance au premier niveau du modèle (~ 8 m), la deuxième ligne les tendances au deuxième niveau du modèle (~ 29 m) et la troisième ligne les tendances au troisième niveau du modèle (~ 55 m). Les tendances dues à la turbulence sont en bleu clair et celles dues à la convection sont en orange pour les cents simulations issues de la calibration du modèle. Les lignes bleues (resp. rouges) représentent les tendances dues à la turbulence (resp. à la convection) de CNRM-CM6.	136
7.1	Profils verticaux de température potentielle et de vent après 8 h de simulation sur le cas GABLS1. Les simulations sont issues de l'espace NROY de la vague 8 de la calibration de GABLS4. Les LES ayant une résolution verticale de 2 m, leurs profils verticaux ont été projetés (par une interpolation linéaire) sur la grille verticale d'ARPEGE-Climat pour être comparés aux résultats des SCMs.	141
7.2	À gauche, espace NROY à l'issue de la vague 8 sur le cas GABLS4, forcé en température de surface. À droite, espace NROY à l'issue de la vague 10, avec ajout des métriques calculées sur le cas GABLS1 sur les deux dernières itérations. Les noms des paramètres apparaissant sur la figure sont les nomms des variables dans le code : AKN = CM, ALD = CE, ALMAVE = LMIN, ALPHAE = AE, ALPHAT = AT, KROVMIN = KOZMIN, KROVZMAX = ZMAX	142
7.3	Évolution de la température potentielle θ sur les niveaux 1 et 5 du modèle (respectivement 9.2 m et 154.4 m), et profil vertical après 8 heures de simulation. Les lignes pointillées verticales et horizontales représentent respectivement l'heure et les altitudes auxquelles les métriques ont été calculées.	144
7.4	Profils verticaux à 8 h du module du vent et de ses composantes zonale et méridienne. Les lignes pointillées indiquent les niveaux du modèle sur lesquels les métriques ont été calculées	145
7.5	a), c) et e) représentent le module du vent sur les niveaux 3, 4 et 5 du modèle (respectivement 60, 110 et 154.4 m). b), d) et f) représentent la rotation du vent sur les niveaux 3, 4 et 5 du modèle	147
7.6	Évolution du profil d'énergie cinétique turbulente sur le cas GABLS1 après 1h (a), 4h (b) et 8h(c) de simulations.	148

Liste des tableaux

7.1	Principales caractéristiques des cas GABLS1 et GABLS4 stage 3-10hr.	140
-----	---	-----

Table des sigles et acronymes

CNRM	Centre National de Recherches Météorologiques
ECMWF	European Center for Medium-range Weather Forecast
Cerfacs	Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
OMM	Organisation météorologique mondiale
NOAA	National Oceanic and Atmospheric Administration
IPCC/GIEC	Intergovernmental Panel on Climate Change/Groupe d'Experts Intergouvernemental sur l'Évolution du climat
CMIP6	Coupled Model Intercomparison Project Phase 6
BSRN	Baseline Surface Radiation Network
AMRC	Antarctic Meteorological Research Center
ARPEGE	Action de Recherche Petite Echelle Grande Echelle
SURFEX	SURFace EXternalisée
NWP/PNT	Numerical Weather Prediction/Prévision Numérique du Temps
LES	Large Eddy Simulation
SCM	Single Column Model
GCM	General Circulation Model
ESM	Earth System Model
CLA	Couche Limite Atmosphérique
CLS	Couche Limite atmosphérique Stable
AMO	Atlantic Multi-decadal Oscillation
NAO	North-Atlantic Oscillation
NPO	North-Pacific Oscillation
SO	Southern Oscillation
SAM	Southern Annular Mode
PA	Péninsule Antarctique
AE	Antarctique Est
AO	Antarctique Ouest

Résumé

La modélisation numérique du climat est un outil essentiel pour comprendre le fonctionnement du système climatique et pour réaliser des projections de son évolution future. Pourtant les modèles numériques de climat comportent de multiples sources d'incertitude. Parmi celles-ci, la représentation des phénomènes de couches limites stables reste un point sur lequel les modèles numériques doivent progresser. Les couches limites stables les plus extrêmes sont observées sur le plateau Antarctique. Une bonne modélisation de cette région dans les modèles numériques repose sur une bonne représentation des échanges d'énergie au sein de la couche limite et en particulier des flux turbulents. Or dans un modèle de climat, les processus turbulents, ainsi que tout phénomène non résolu par la composante dynamique du modèle reposent sur un ensemble de « sous-modèles » appelés paramétrisations physiques. Ces paramétrisations, en simplifiant les processus physiques en jeu, introduisent des paramètres dont les valeurs sont plus ou moins bien documentées et qui peuvent être considérés comme les leviers de réglage du modèle. L'étape cruciale de la calibration du modèle vise à choisir les valeurs de ces paramètres. L'approche classique consiste à chercher un réglage du modèle optimal sur la base d'un jeu de métriques. C'est un travail long et fastidieux, dans lequel on fait varier un ou deux paramètres à la fois et qui ne permet pas d'explorer l'ensemble des possibilités de réglage du modèle. De plus les sources d'incertitudes ne sont pas toujours prises en compte et l'ensemble de la procédure est peu documentée, peu formalisée, et peu reproductible. Une approche inspirée de l'*History Matching* a été récemment proposée pour calibrer la physique des modèles atmosphériques. Cette approche propose non pas de chercher un réglage optimal des paramètres, mais de déterminer une région de l'espace des paramètres dans lequel le modèle se « comporte bien ». Associée à l'utilisation d'émulateurs statistiques simulant le comportement du modèle de climat, elle permet de répondre à une partie des problèmes posés par une approche plus classique.

Le travail réalisé au cours de cette thèse se propose de documenter à travers une hiérarchie de configurations le comportement du modèle numérique ARPEGE-Climat sur l'Antarctique. Ces simulations mettent en évidence un biais froid sur le plateau continental. Une modification empirique du calcul de l'albédo de la neige corrige partiellement ce biais durant la période de l'été austral, soulignant l'importance de sa paramétrisation pour la représentation du climat antarctique. Une étude plus détaillée sur Dôme C indique que le modèle ARPEGE-Climat peine à représenter les couches limites stables. Le modèle reproduit ce type d'erreur sur le cas de couche limite stable GABLS4 et dans une moindre mesure sur le cas GABLS1. Une approche de type *History Matching* est appliquée à une version 1D du modèle pour comprendre dans quelle mesure ces erreurs relèvent d'un choix de réglage du modèle ou de limites intrinsèques de sa physique. On montre ainsi que le modèle ARPEGE-Climat contient la physique suffisante pour représenter toute une gamme de couches limites stables. De plus la force de cette méthode est de pouvoir utiliser l'information obtenue dans le cadre 1D pour calibrer le modèle 3D, tout en garantissant un comportement satisfaisant sur les couches limites stables. Ce travail a également été l'occasion de contribuer au développement d'un outil basé sur ces techniques d'*History Matching* et mis à la disposition de la communauté de la modélisation du climat, et de discuter la pertinence de ces nouvelles approches.

Abstract

Climate modeling is a key tool for understanding the climate system and for making projections of its future evolution. Yet numerical climate models have multiple sources of uncertainty. Among these, the representation of stable boundary layer processes remains one of the main points on which numerical models must make progress. The most extreme stable boundary layers are observed on the Antarctic plateau. A good modeling of Antarctica in numerical models is based on a good representation of energy exchanges within the boundary layer and in particular of turbulent fluxes. However, in a climate model, turbulent processes, as well as any other small-scale processes (not resolved by the dynamical part of the model), or processes not related to fluid mechanics (e.g. radiation), are based on a set of sub-models called physical parameterizations. These parameterizations introduce a certain number of parameters whose values are more or less well documented and which can be considered as the adjustment levers of the model. The model tuning step is the choice of the values of these different parameters and is considered as a crucial step in the development of the models. The classical approach is to look for an optimal model setting based on a set of metrics. This is a long and tedious work, in which one or two parameters are varied one at a time and which does not allow the exploration of all the possibilities of model adjustment. Moreover, the sources of uncertainty are not always taken into account and the whole procedure is not very reproducible. An approach inspired by History Matching has recently been proposed to calibrate the physics of atmospheric models. This approach proposes not to look for a potential optimal setting of the parameters, but to identify a region of the parameter space in which the model « performs well ». Associated with the use of statistical emulators mimicking the climate model behaviour, it allows to answer part of the problems posed by a more classical approach (exhaustive exploration of the parameter space, taking into account uncertainties, reproducibility).

The work carried out during this thesis proposes to document through a hierarchy of configurations the performance of the numerical model ARPEGE-Climate over Antarctica. These simulations highlight a cold bias on the continental shelf. An empirical modification of the albedo computation is proposed that partially corrects the temperature bias during the austral summer period. A focus on Dome C (East Antarctica) shows biases regularly highlighted in the literature on the representation of stable boundary layers in numerical models. A History Matching approach is then applied to understand to what extent the biases highlighted on Dome C are due to a choice of model tuning or to intrinsic limits of the model physics. This approach, applied to a 1D version of the model, is based on an extreme stable boundary layer case (GABLS4) and a moderate boundary layer case (GABLS1). It is thus shown that the ARPEGE-Climat model contains sufficient physics to represent a range of stable boundary layers. This work was also an opportunity to contribute to the development of a tool based on these History Matching techniques and made available to the climate modeling community and to discuss the relevance of these new approaches.

Introduction

Le continent Antarctique est le continent le plus méridional de la planète. Il est presque entièrement situé au-delà du cercle polaire austral. Cette situation géographique, associée au fort albédo des surfaces glacées, induit un bilan radiatif à la surface déficitaire : le rayonnement solaire incident à la surface est nul durant l'hiver austral (nuit polaire), et reste relativement faible durant l'été austral du fait de l'incidence importante des rayons du soleil. Ce bilan radiatif, associé à l'altitude élevée du plateau continental (> 2000 m sur l'Antarctique Est) fait de l'Antarctique le continent le plus froid de la planète. C'est également un élément clé du climat de la Terre. En tant que puits de chaleur de l'hémisphère sud, il joue un rôle important dans la machine thermodynamique qu'est le système climatique et sur la circulation atmosphérique générale de l'hémisphère sud. À la différence de l'Arctique qui connaît un réchauffement sans précédent, la réponse de l'Antarctique au changement climatique est plus incertaine. En particulier, cette réponse n'est pas homogène à l'échelle du continent Antarctique : aucun changement significatif n'est observé sur l'Antarctique Est, alors que certaines régions de l'Antarctique Ouest ont connu un réchauffement très important. Ces résultats doivent cependant être pris avec précaution étant donnée la rareté des observations *in situ*. La variabilité interannuelle et multi-décennale de ces régions, insuffisamment comprise, complique encore l'interprétation de ces résultats. Pour améliorer notre compréhension du fonctionnement du climat de cette région, et pour anticiper sa réponse future au changement climatique, les modèles numériques de climat tels que les modèles de circulation générale sont des outils précieux. Pourtant ces modèles numériques peinent à représenter correctement cette région. À titre d'exemple, Bracegirdle et al. (2013) mettait en évidence un biais dans l'ensemble des modèles de climat qu'il évalue sur la position du maximum des vents d'ouest dans l'hémisphère sud, repoussant celui-ci trop loin vers l'équateur. Ce biais est toujours présent dans la plupart des modèles même si il a été très sensiblement réduit (Bracegirdle et al., 2020). La représentation de la glace de mer, et en particulier de sa saisonnalité, est également souvent peu cohérente avec les observations (Mahlstein et al., 2013, Turner et al., 2013). La température de surface sur le plateau continental est également souvent mal représentée. La réanalyse ERA-Interim du Centre Européen de Prévision du Temps (CEPMMT), considérée comme une référence jusqu'à l'arrivée récente de sa remplaçante ERA5, présentait des biais importants de température de surface sur l'ensemble du continent Antarctique, en comparaison à celle estimée par satellite (Fréville et al., 2014). Ce biais est maintenant corrigé et provenait de lacunes dans le schéma de surface et des propriétés de la neige. Un des objectifs de cette thèse est de documenter le comportement du modèle de climat de Météo-France (ARPEGE-Climat) sur l'Antarctique.

Du fait du fort refroidissement nocturne ou hivernal, les couches limites stables sont une

composante importante du climat antarctique (Vignon et al., 2017b). Elles ne sont pas nécessairement spécifiques de l'Antarctique, mais elles y sont particulièrement extrêmes avec des gradients verticaux de température pouvant dépasser 1 Km^{-1} (Genthon et al., 2013). Nombre d'études ont montré que les modèles numériques avaient du mal à représenter les couches limites stables (e.g. Sandu et al., 2013, Holtslag et al., 2013, King et al., 2001), ce qui est probablement exacerbé sur l'Antarctique et explique au moins en partie la difficulté des modèles à représenter le climat Antarctique. En particulier Sandu et al. (2013), Holtslag et al. (2013) et Mahrt (2014) établissent un lien entre la façon dont le mélange turbulent est représenté dans les modèles numériques et leur difficulté à représenter des couches limites très stratifiées. La représentation de ces flux turbulents passe par l'utilisation d'une paramétrisation physique. En effet un modèle de climat a une résolution de l'ordre de la centaine de kilomètres. A cette échelle, le modèle résout les équations de la mécanique des fluides. Mais certains phénomènes de petite échelle mais pertinents pour les échelles résolues (tels que les phénomènes turbulents) ne peuvent pas être décrits explicitement. Pour les représenter, on a recourt à des paramétrisations physiques qui décrivent l'effet moyen de ces phénomènes à travers des représentations simplifiées. Ces paramétrisations, qui représentent des phénomènes complexes, reposent sur un certain nombre d'approximations. De ces approximations émergent des grandeurs, appelés paramètres libres, dont la valeur n'est pas connue a priori mais qu'il convient de régler pour effectuer une simulation. On en connaît en général juste un domaine de variations estimé par exemple lors de campagnes de mesures ou sur la base de considérations théoriques. Une certaine latitude est donc laissée aux modélisateurs pour le choix des valeurs à donner à ces paramètres, et la calibration est l'étape dans le développement des modèles qui consiste à les régler en fonction des qualités attendues du modèle.

Cette étape cruciale dans le développement des modèles est pourtant restée longtemps dans l'ombre et jusqu'à récemment a été peu discutée dans la littérature scientifique ou dans la documentation technique des modèles. Elle n'est pas spécifique à la modélisation du climat et est en fait un problème général qui se retrouve dans le développement de la plupart des grands codes numériques. Depuis une dizaine d'années, l'importance des stratégies de calibration et la nécessité de les décrire dans les articles de présentation des modèles de climat sont soulevées. Hourdin et al. (2017) par exemple proposent trois étapes dans la calibration d'un modèle : une première étape, au plus proche des processus, passant par la réalisation de simulations unidimensionnelles (1D) et de leur comparaison avec des simulations de référence à très fine résolution, puis une calibration utilisant des simulations atmosphériques tridimensionnelles (3D), non couplées avec l'océan, afin de garantir, entre autre, l'équilibre radiatif du modèle, et enfin une calibration du modèle atmosphérique couplé avec un modèle d'océan. Le cœur du travail réalisé dans cette thèse se place dans le cadre de la première étape décrite par Hourdin et al. (2017), *i.e.* la calibration d'une version 1D du modèle. L'approche classique pour la calibration d'un modèle de climat est une approche de type essai/erreur et représente généralement un travail long et fastidieux, reflet d'une formalisation insatisfaisante du problème. Par nature une approche de type essai/erreur nécessite de nombreuses étapes et des retours en arrière lorsque les modifications apportées ne sont pas satisfaisante. Chacune de ces étapes reposera sur l'expertise et les choix du modélisateur. L'ensemble de la méthode est ainsi délicat à documenter et donc peu reproductible. Différentes approches ont été proposées pour répondre aux difficultés posées par la calibration et en particulier afin de rendre cette étape importante du développement d'un modèle plus systématique et plus

objective. On notera en particulier des approches basées sur des techniques d'optimisation qui vise à exhiber un réglage optimal du modèle (e.g. Severijns and Hazeleger, 2005, Zhang et al., 2015). L'approche suivie dans le cadre de cette thèse est différente et vise non pas à exhiber un réglage unique et optimal du modèle, mais à définir les régions dans lesquelles le comportement du modèle est « raisonnable », ce qui doit servir de guide pour la calibration 3D du modèle. Elle repose sur l'utilisation d'outils issus de la communauté de quantification de l'incertitude, et plus précisément sur les techniques dites d'*History Matching*. On suit en cela Williamson et al. (2013) qui propose une application de la méthode d'*History Matching* pour calibrer le modèle atmosphérique du Hadley Center en Angleterre. Une collaboration de la communauté française de modélisation du climat (CNRM, IPSL) avec des chercheurs en méthodes statistiques de l'université d'Exeter a abouti au développement d'un outil commun de calibration des modèles de climat et auquel ces travaux de thèse ont contribué (Couvreur et al., 2020b). L'ambition de cet outil est de mieux formaliser le problème de la calibration 1D et son articulation avec la calibration 3D tout en proposant une approche plus objective. Il s'appuie sur l'utilisation de modèles statistiques appelés émulateurs qui simulent le comportement du modèle atmosphérique, caractérisé à l'aide de métriques définies *a priori*, quand on fait varier ses paramètres libres. L'utilisation de ces émulateurs, du fait de leur faible coût numérique, permet une exploration systématique de l'espace des paramètres libres. La méthode proposée est itérative et permet, à chaque itération, d'améliorer la performance des émulateurs en réduisant leur incertitude et d'éliminer des régions de l'espace des paramètres qui ne sont pas compatibles avec les observations.

Dans ce travail de thèse cette méthode est appliquée à une version 1D de ARPEGE-Climat sur le cas d'étude GABLS4 qui propose un cas de couche limite très stable, basé sur une campagne de mesures réalisée pendant l'été austral 2009 sur le site de Dôme C (Antarctique Est). L'exploration systématique par *History Matching* des performances du modèle lorsque l'on fait varier ses paramètres libres permet de questionner la capacité de la physique du modèle ARPEGE-Climat à simuler des couches limites très stratifiées. Elle permet ainsi de déterminer si d'éventuels biais du modèle sont liés à des limites intrinsèques des paramétrisations physiques (en particulier la paramétrisation de la turbulence), ou bien à des choix de calibration de ces paramétrisations. La capacité du schéma de turbulence de ARPEGE-Climat à simuler des couches limites de stabilité plus modérée (en particulier celle du cas GABLS1) a déjà été démontrée (Cuxart et al., 2006). Les réglages obtenus à partir de la calibration sur GABLS4 sont évalués sur le cas GABLS1 afin de vérifier si ceux-ci restent compatibles avec la représentation de couches limites stables plus modérées.

Organisation du manuscrit Le manuscrit est organisé en sept chapitres :

1. **L'Antarctique et les couches limites stables.** Ce chapitre présente brièvement le continent Antarctique et sa climatologie. Les couches limites stables sont décrites.
2. **Outils de modélisation.** Le modèle atmosphérique ARPEGE-Climat est présenté brièvement, décrivant plus en détail les paramétrisations pertinentes pour la représentation de la couche limite atmosphérique. Quelques éléments de description du modèle de surface SURFEX sont donnés. Les différents jeux de données et simulations de référence utilisés pour évaluer le comportement du modèle ARPEGE-Climat sont passés en revue. Les cas d'étude GABLS, en particulier les cas GABLS1 et GABLS4, largement

utilisés dans ce travail, sont présentés.

3. **Représentation du climat antarctique par ARPEGE-Climat.** Les performances du modèle sur l'Antarctique sont documentées en examinant plus précisément son comportement à la station de Dôme C.
4. **Calibration d'un modèle de climat.** La problématique de la calibration d'un modèle numérique de climat est posée, ainsi que les stratégies de calibration couramment utilisées. L'approche *History Matching* utilisée dans ce travail est présentée, ainsi que l'outil de calibration développé par la communauté française de modélisation du climat en collaboration avec l'université d'Exeter. Ce chapitre, ainsi que les deux suivants représentent le cœur de ces travaux de thèse
5. **Application des outils d'*History Matching* au cas GABLS4.** Ce chapitre est central et inclut un article publié dans le *Journal of Advances in Modeling Earth System* qui décrit l'application des outils d'*History Matching* présentés précédemment au cas de couche limite très stable GABLS4. Quelques compléments aux résultats exposés dans l'article sont proposés.
6. **Extension sur le cas GABLS1.** Les résultats de la calibration de ARPEGE-Climat sur le cas très stratifié GABLS4 sont évalués et étendus sur un autre cas d'étude (GABLS1) de stabilité modérée.
7. **Conclusions et perspectives.** Ce dernier chapitre résume les résultats de ce travail de thèse et présente quelques perspectives.

L'Antarctique et les couches limites stables

Sommaire

2.1	L'Antarctique, description et importance dans le système climatique	9
2.1.1	Description géophysique	9
2.1.2	Le climat Antarctique	11
2.1.3	Le climat à Dôme C	15
2.1.4	Rôle dans le système climatique planétaire	17
2.1.5	Changement climatique et Antarctique	19
2.2	Les couches limites stables	22
2.2.1	La couche limite atmosphérique	22
2.2.2	Les couches limites stables	26

2.1 L'Antarctique, description et importance dans le système climatique

2.1.1 Description géophysique

Le continent Antarctique, d'une superficie de 14.2 millions km², est situé presque entièrement au-delà du cercle polaire austral (66° 33' S) et représente 10% des terres émergées de la Terre. Il se compose de trois parties distinctes : l'Antarctique Est, l'Antarctique Ouest et la Péninsule Antarctique (fig. 2.1). L'Antarctique Est, la plus étendue de ces trois régions, est recouverte par une calotte glaciaire appelée *inlandsis*. Excepté près des côtes, l'altitude de l'*inlandsis* de l'Antarctique Est est partout supérieure à 2 000 m et dépasse même 4 000 m dans sa partie centrale. Pour autant le relief est peu marqué et l'Antarctique Est présente une surface essentiellement plate avec des pentes très faibles. L'*inlandsis* est dominée par des dômes de glace, *i.e.* des maxima locaux de la calotte, reliés par des lignes de partage glaciaire (*Ice Divide* en anglais) séparant les directions d'écoulement opposées de la glace. Ces lignes sont les analogues des lignes de partage des eaux (*Water Divide* en anglais), utilisées en hydrologie pour délimiter les bassins versants. L'écoulement de la glace s'effectue depuis ces lignes vers les côtes de l'Océan Austral. Les sommets de la calotte que sont les dômes sont quasiment plats et leur relief est indiscernable à l'oeil nu. Le plus haut d'entre-eux, le Dôme A (ou Dôme Argus) culmine à une altitude de 4 093 m. Le Dôme C (ou Dôme Circé ou Charlie),

situé plus à l'est à 3 233 m d'altitude domine le plateau Antarctique au sud de Terre-Adélie et abrite la station de recherche franco-italienne Concordia. La géographie de ces dômes favorise le développement de couches limites très stratifiées et le cas d'étude GABLS4 utilisé dans cette thèse est basé sur une campagne de mesures qui a eu lieu à Dôme C. L'Antarctique Ouest, séparé de l'Antarctique Est par la chaîne Transantarctique, est également dominé par une calotte glaciaire mais présente une altitude moyenne plus faible (850 m). Enfin la Péninsule Antarctique est une étroite chaîne de montagnes qui s'avance dans l'Océan Glacial Antarctique et présente une largeur moyenne de 70 km, pour une hauteur moyenne de 3000 m.

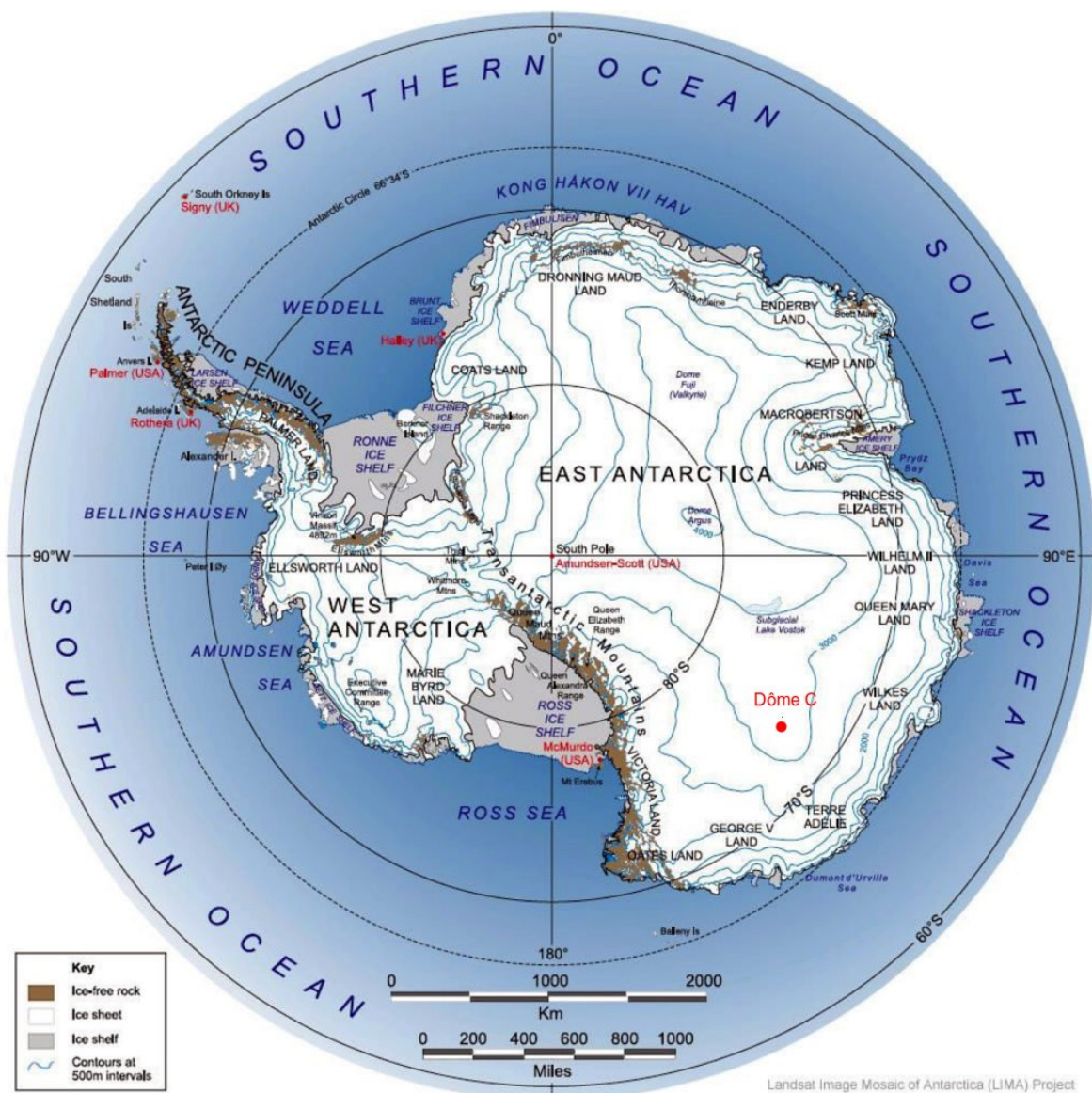


FIGURE 2.1 – Carte de l'Antarctique présentant les principales régions géographiques du continent.

La cryosphère antarctique La quasi totalité (98%) du continent est recouverte d'une calotte de glace. Celle-ci représente environ 70% de l'eau douce de la planète (King and Turner, 1997). La profondeur de cette couche de glace est très variable et atteint plus de 4700 m dans sa partie la plus épaisse (4776 m à 400 km au sud de la base française Dumont-d'Urville en Terre Adélie). Une partie de la calotte glaciaire repose donc sur un sol situé sous le niveau de la mer. La fosse subglaciale de Bentley avec une profondeur de 2555 m sous le niveau de la mer a longtemps été considérée comme le point le plus bas de la Terre qui ne soit pas recouvert par les océans. Ce record est désormais la propriété du canyon du glacier Denman, très récemment cartographié, avec une profondeur de 3500 m sous le niveau de la mer (Morlighem et al., 2020). Les glaciers constituant l'*inlandsis* s'écoulent vers les bords du continent et se prolongent sur l'Océan Austral en plates-formes de glace flottante (*iceshelf* en anglais) dont l'épaisseur dépasse souvent la centaine de mètres et peut même atteindre 1000 m (Griggs and Bamber, 2011). Les plates-formes de Ronne et de Ross, qui débouchent sur les mers de Weddell et de Ross (fig. 2.1) et d'une superficie respective de 430 000 et 500 000 km², sont les plus importantes et sont alimentées par les glaciers de l'Antarctique Ouest. De surface plate et souvent très homogène (faible rugosité), ces régions, qui représente 44% du trait de côte (King and Turner, 1997), sont des lieux privilégiés pour la formation des couches limites stable (Derbyshire, 1995, Mahrt, 1999). Ces vastes étendues de glace flottante ne doivent pas être confondues avec la banquise qui est constituée de glace de mer et résulte directement du gel de l'eau de mer. La banquise Antarctique se forme tout autour du continent Antarctique pendant l'hiver et peut s'étendre vers le nord jusqu'à 55°S. Contrairement à l'Arctique, elle disparaît presque complètement en été et est donc essentiellement constituée de glace jeune, d'une épaisseur beaucoup plus modeste que les plates-formes de glace.

2.1.2 Le climat Antarctique

L'Antarctique est le continent des extrêmes météorologiques. C'est le continent le plus froid, le plus sec et le plus venteux de la planète. Cette section décrit les principales caractéristiques du climat antarctique.

Bilan radiatif En reprenant les notations de Previdi et al. (2013), le bilan d'énergie $\frac{\partial E}{\partial t}$ de l'atmosphère antarctique peut s'écrire

$$\frac{\partial E}{\partial t} = F_{\text{TOA}:\text{NET}} + F_{\text{SFC}:\text{NET}} + F_{\text{WALL}} \quad (2.1)$$

où $F_{\text{TOA}:\text{NET}}$ est le flux radiatif net (solaire + infrarouge) au sommet de l'atmosphère, $F_{\text{SFC}:\text{NET}}$ est le flux d'énergie net à la surface et F_{WALL} est le flux horizontal d'énergie intégré verticalement sur toute la hauteur de l'atmosphère. La figure 2.2 présente le cycle annuel moyen des différents termes de l'équation 2.1 pour l'Antarctique. On constate que le flux radiatif net au sommet de l'atmosphère est négatif toute l'année, avec une moyenne annuelle de -102 W m^{-2} (données CERES, Previdi et al., 2013). $F_{\text{TOA}:\text{NET}}$ est minimal ($\simeq -140 \text{ W m}^{-2}$) entre avril et août alors que le flux solaire est quasi-nul. Durant l'été austral, $F_{\text{TOA}:\text{NET}}$ reste négatif, le flux solaire net, du fait d'un angle d'incidence important et de

l'albédo de la surface, ne compensant pas le refroidissement radiatif. En surface, $F_{\text{SFC:NET}}$ est calculé comme la somme des flux radiatifs nets (solaire et infrarouge) et des flux de chaleur latente et sensible. Pendant la période hivernal, la surface se refroidit, avec $F_{\text{SFC:NET}} > 0$ de mars à octobre. Le refroidissement radiatif de l'atmosphère est compensé par les flux méridiens (F_{WALL}) sous forme d'énergie potentielle et de chaleur sensible (Malardel, 2005).

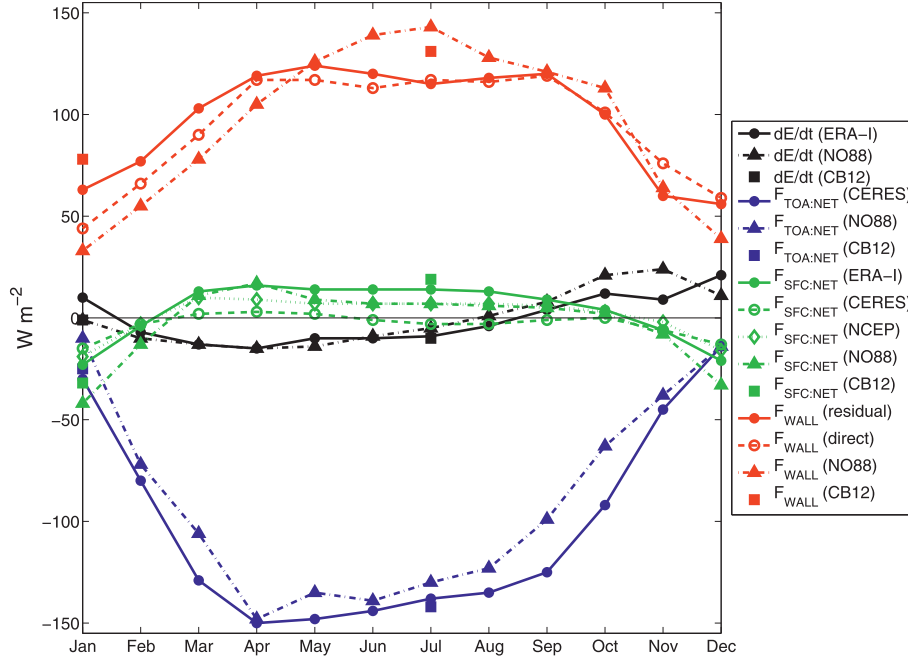


FIGURE 2.2 – Différentes estimations du bilan d'énergie moyen de l'Antarctique sur la période 2001-2010 : $F_{\text{SFC:NET}}$ représente le bilan d'énergie à la surface, $F_{\text{TOA:NET}}$ le bilan d'énergie au sommet de l'atmosphère et F_{WALL} est le flux horizontal d'énergie intégré verticalement sur toute la hauteur de l'atmosphère. (Previdi et al., 2013). Les valeurs positives du bilan correspondent à un gain d'énergie pour l'atmosphère.

Températures En conséquence du bilan d'énergie et de l'altitude élevée du plateau continental, les températures de surface restent largement négatives toute l'année. La figure 2.3a montre la température de surface moyenne annuelle sur le continent. Exceptée sur une étroite bande côtière, celle-ci est inférieure à -20°C sur l'ensemble du continent. La partie centrale de l'inlandsis de l'Antarctique Est présente même des températures annuelles moyennes inférieures à -50°C . Durant l'hiver austral (figure 2.3b), une large zone centrale présente des températures inférieures à -65°C . Si l'été est plus doux, la température reste très largement négative ($< -36^{\circ}\text{C}$ sur la partie centrale). Le plateau continental présente les températures à 2 m les plus froides jamais mesurées à la surface de la Terre. On y trouve le record de la température la plus froide officiellement mesurée à la surface de la planète avec -82°C , mesurée à la station Vostock le 21 juillet 1983 (Budretsky, 1984). Récemment, des températures autour -98°C ont été estimées dans des dépressions très peu profondes (de l'ordre de 2 m, et quasiment indiscernables à l'oeil nu) au sud de Vostok (Scambos et al., 2018). Ces températures n'ont toutefois pas été mesurées *in situ* mais estimées à l'aide de méthodes statistiques à partir de données de rayonnement mesurées par satellite. Le record du 21 juillet 1983 tient

donc toujours.

Du fait que l'émissivité de la neige supérieure à celle de l'air (Hudson and Brandt, 2005), la surface se refroidit plus vite que l'atmosphère. Une inversion de température se forme alors près de la surface avec des gradients verticaux de température particulièrement intenses et qui peut durer plusieurs semaines. Des perturbations d'origine océanique peuvent toutefois pénétrer suffisamment loin sur le plateau continental et générer des épisodes de réchauffement brusque détruisant momentanément l'inversion (Genthon et al., 2013).

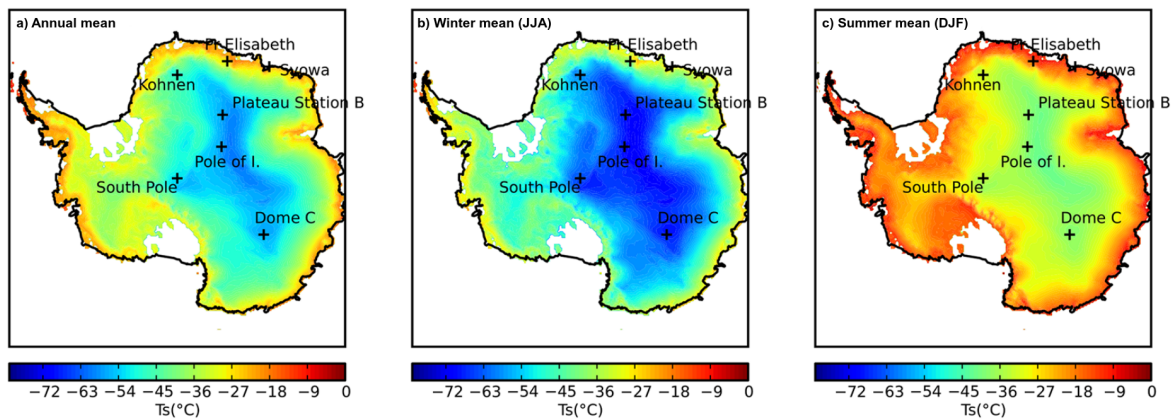


FIGURE 2.3 – Température de surface en moyenne annuelle a), en moyenne hivernale b) et en moyenne estivale c). Ces données de températures sont estimées à partir de données satellites MODIS (Fréville et al., 2014)

Vents catabatiques L'augmentation de la densité de la masse d'air sur l'*inlandsis* liée au fort refroidissement radiatif génère un écoulement lié aux forces de flottabilité depuis le sommet de la calotte glaciaire vers les côtes. Les vents correspondants à cet écoulement, appelés vents catabatiques, accélèrent progressivement à mesure qu'ils se rapprochent des côtes et que la pente devient plus prononcée. Les vents les plus forts soufflent dans les 150 premiers km à l'intérieur des terres, là où les pentes dépassent 1/100 (Parish and Bromwich, 1987). Les vents catabatiques sont grossièrement orientés le long des lignes de pentes (l'angle dont ils s'écartent de la ligne de pente est principalement dû à la force de Coriolis – e.g. Parish and Bromwich, 1987). En surface, le vent moyen résulte d'un équilibre entre l'effet des forces de pression et celui des forces de flottabilité. Ainsi à 100 m au dessus du sol la direction moyenne du vent fait un angle marqué avec la direction du vent à 500 m (Figure 2.4), altitude à laquelle la circulation est anticyclonique et suit au premier ordre les lignes de niveau de la calotte glaciaire. Les vent catabatiques prédominent durant l'hiver austral et ce d'autant plus que l'on se rapproche des régions côtières et de leur pentes prononcées (van den Broeke et al., 2002, van den Broeke and van Lipzig, 2003).

Précipitations La distribution des précipitation est très hétérogène (Palermé et al., 2014) et se produisent presque exclusivement sous forme de neige (les valeurs des précipitations annuelles seront données dans la suite en mm d'équivalent en eau). Elles sont importantes sur

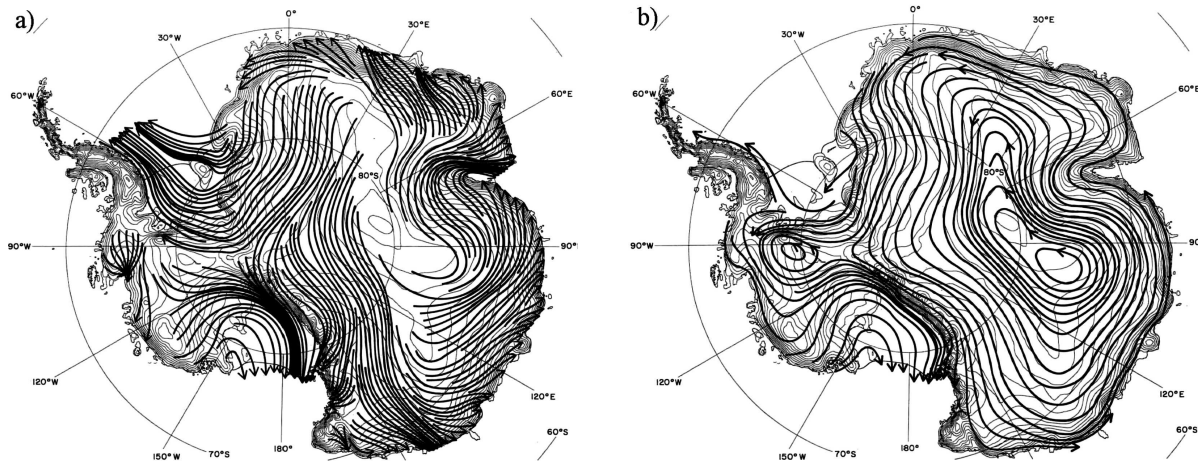


FIGURE 2.4 – Lignes de flux moyennes à a) 100 m et b) 500 m au dessus du sol sur la période juin 2003-mai 2004 au dessus de l'Antarctique, simulées par le modèle de prévision numérique *Antarctic Mesoscale Prediction System* (Parish and Bromwich, 1987).

les zones côtières avec des valeurs allant de 400 à 1000 mm.an⁻¹ et pouvant même dépasser les 1500 mm.an⁻¹ sur la Péninsule Antarctique (Bromwich et al., 2004). Sur le plateau continental de l'Antarctique Est, au contraire, les précipitations sont très rares avec des moyennes annuelles inférieures à 50 mm.an⁻¹.

Oscillation Antarctique Une zone de vents d'ouest forts pilotée par un système de basses pressions entre 40° et 60° de latitude sud ceinture le continent Antarctique. Ce système de basses pressions se déplace selon la direction Nord-Sud et la variabilité interannuelle de ce déplacement est principalement pilotée par l'Oscillation Antarctique (OA) ou Southern Annular Mode (SAM – e.g. Gong and Wang, 1999). Définie de façon similaire aux grands modes de variabilité de l'atmosphère identifié par Walker (1928) comme l'Oscillation Nord-Atlantique (NAO), l'oscillation Nord-Pacifique Nord (NPO) ou l'oscillation australe (SO), l'OA décrit une oscillation de grande échelle du champ de pression entre les moyennes et les hautes latitudes de l'hémisphère sud. Walker (1928) soupçonnait déjà son existence, mais le manque d'observations sur cette région du globe a retardé sa mise en évidence. Gong and Wang (1999) proposent une définition d'un indice de l'OA (ou AOI pour Antarctic Oscillation Index) afin de quantifier sa variabilité et d'établir les éventuelles connexions avec des situations météorologiques particulières ou d'autres modes de variabilité de l'atmosphère. À partir d'une analyse en composantes principales de la pression au niveau de la mer (Figure 2.5a, b) ils proposent la formulation suivante $AOI = P_{40^{\circ}S}^* - P_{65^{\circ}S}^*$ où $P_{40^{\circ}S}^*$ et $P_{65^{\circ}S}^*$ correspondent à la moyenne zonale de la pression au niveau de la mer aux latitudes 40°S et 65°S respectivement. L'AOI explique une large partie de la variabilité du champ de pression atmosphérique sur l'Antarctique (Figure 2.5c) allant jusqu'à 50% sur les calottes glacières de l'Antarctique Est et Ouest. Les valeurs positives de l'AOI correspondent à des vents d'ouest plus forts que la moyenne aux hautes latitudes. Les valeurs négatives correspondent au contraire à des vents d'ouest plus faibles que la moyenne. Marshall (2007) montre une relation forte entre l'AOI et les températures de l'air à la surface du continent. La figure 2.6 illustre cette corrélation.

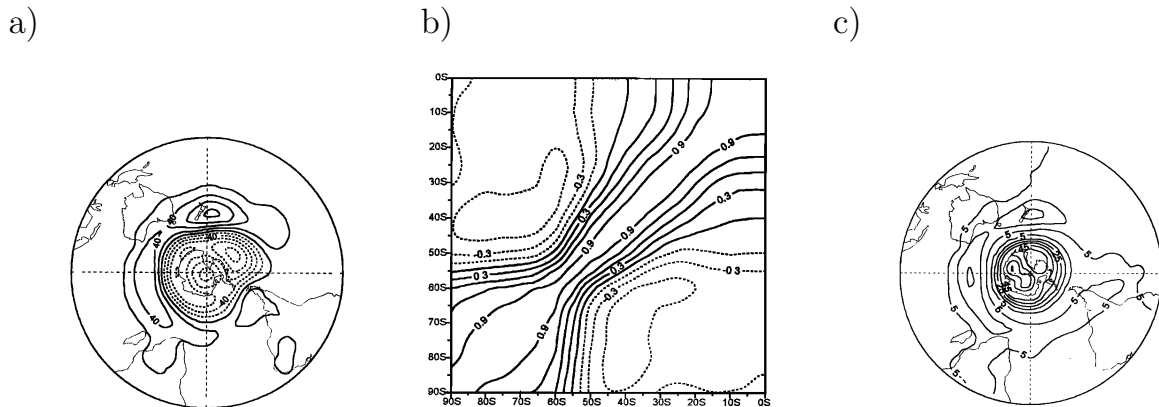


FIGURE 2.5 – L'indice d'Oscillation Antarctique (Gong and Wang, 1999). a) Première EOF du champ de pression au niveau de la mer. b) Coefficients de corrélation entre les moyennes zonales du champ de pression. c) Pourcentage de la variance expliquée du champ de pression par l'AOI

2.1.3 Le climat à Dôme C

Le cas d'étude de couche limite stable GABLS4 qui sera utilisé dans la suite de ce travail pour la calibration du modèle ARPEGE-Climat a été construit à partir d'observations à Dôme C dont on décrit ici le climat. Le Dôme C est situé dans la partie orientale du haut plateau Antarctique ($75^{\circ}06' \text{ S}$, $123^{\circ}20' \text{ E}$) à 3 233 m au dessus du niveau de la mer. Le paysage y semble parfaitement plat et la pente autour de Dôme C est inférieure à 1/1000 (Cavitt et al., 2018). Le sol est recouvert de structures de neiges érodées par le vent appelées sastrugi d'une taille variant de 5 à 30 cm. Ces structures ont servi pour estimer la direction du vent dominant dans des régions où il n'y avait pas de mesures (Parish and Bromwich, 1987). Dôme C abrite la station de recherche franco-italienne Concordia à proximité de laquelle est installée une tour instrumentée de 45 m depuis 2008. Cette tour possède six niveaux sur lesquels sont déployés des instruments de mesures météorologiques. Un des objectifs de cette tour est d'étudier les couches limites stables de très faible épaisseur qui se développent à Dôme C. Une station automatique du programme Antarctic Meteorological Research Center (AMRC) est également installée à proximité.

Le vent moyen à 10 m est modéré (autour de 3 ms^{-1} en moyenne annuelle) et peut atteindre de façon exceptionnelle 15 ms^{-1} . Ces faibles valeurs s'expliquent par la topographie très plate du site qui empêche la formation de vents catabatiques locaux (Aristidi, E. et al., 2005, King and Turner, 1997) ce qui fait de Dôme C un des points les moins venteux du continent Antarctique (van den Broeke and van Lipzig, 2003). Le vent est généralement un vent de sud, apportant un air sec et froid. Le ciel à Dôme C est ainsi le plus souvent clair. La vitesse du vent ne présente pas de cycle annuel prononcé (Argentini et al., 2014).

Ces conditions (vent modéré, ciel clair) font de Dôme C un lieu propice au développement de couches limites très stables (Vignon et al., 2017b). En hiver, de fortes inversions de surface se développent qui génèrent les vents catabatiques observés dans certaines zones de confluence

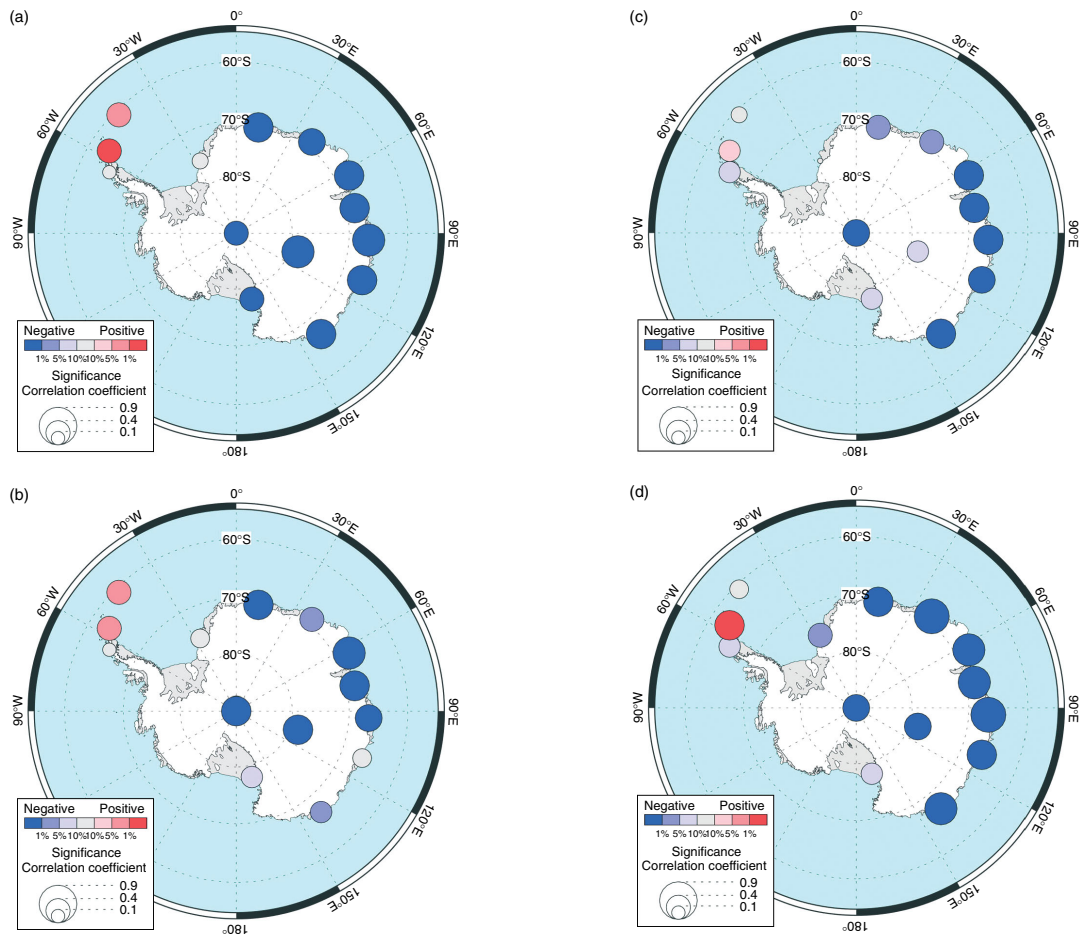


FIGURE 2.6 – Corrélations entre l'indice SAM et la température de l'atmosphère à 2 m en Antarctique pour la période 1957–2004 : (a) automne ; (b) hiver ; (c) printemps et (d) été. Marshall (2007)

le long de la côte est de l'Antarctique. La température moyenne mensuelle à 2 m possède un cycle annuel très marqué et s'échelonne de -25°C pendant l'été austral à -70°C pendant la nuit polaire. La figure 2.7, reprise de Genthon et al. (2013) illustre le cycle annuel de température mesurée par la station automatique AMRC installée à Dôme C. On remarque en particulier que la variabilité interannuelle est importante en hiver avec des écarts de moyennes mensuelles en juillet pouvant excéder 15 K (voir les années 2009 et 2010). Cette variabilité se retrouve également en moyenne troposphère et peut être reliée aux principaux modes de variabilité de l'atmosphère dans cette région (AOI, SOI, ENSO) (Genthon et al., 2003, 2013). L'été, un cycle diurne marqué est observé malgré un soleil présent au dessus de l'horizon en permanence (Genthon et al., 2010). Pendant la journée, le rayonnement solaire est suffisant pour réchauffer la surface. Cet excès de chaleur est transmis à l'atmosphère par le flux de chaleur sensible. Une couche limite convective (Genthon et al., 2010) se met alors en place qui peut atteindre une épaisseur de 300 m. Dès que le flux solaire incident devient plus faible que le rayonnement thermique de la surface, la couche limite commence à se refroidir (voir la figure 2.8 présentant les cycles diurnes moyens sur Dôme C à partir de mesures *in situ*). Le refroidissement peut être suffisamment intense pour qu'une couche limite très stable et très peu épaisse (< 50 m) se forme à Dôme C, avec un fort refroidissement de la surface et un gradient de température marqué.

Les précipitations sont faibles et l'accumulation de neige est inférieure à 10 cm de neige par an (Genthon et al., 2016)). L'essentiel des chutes de neige est dû à des advections d'air chaud et humide d'origine océanique (Genthon et al., 2016). Des précipitations sous forme de poudrin de glace peuvent également avoir lieu par temps clair et pourrait constituer une part significative des précipitations totales à Dôme C (Schlosser et al., 2016). Les phénomènes de transport de la neige par le vent ont un rôle important dans le bilan de masse sur le plateau antarctique (Das et al., 2013). On distingue deux modes de transport de la neige par le vent : la saltation (ou drift) correspond à un transport sur de courtes distance dans lequel les grains de neige rebondissent au sol, et le transport de neige soufflée dans lequel les grains de neige sont en suspension. Si les phénomènes de saltation sont fréquents à Dôme C, ils ne concernent que des transports de neige au niveau local. Les épisodes de transport de neige soufflée sont rares (Libois et al., 2014, Genthon et al., 2016).

2.1.4 Rôle dans le système climatique planétaire

L'Antarctique et la circulation atmosphérique générale Conséquence du positionnement de la quasi totalité de l'Antarctique au delà du cercle polaire et du fort albédo de la neige et de la glace, l'Antarctique joue le rôle d'un puits d'énergie pour l'hémisphère sud. En prenant un point de vue purement thermodynamique, on peut considérer l'Antarctique comme la source froide d'une machine thermique (dont la source chaude est la zone intertropicale). Son rôle est donc prépondérant dans la circulation atmosphérique générale de l'hémisphère sud. Par conservation de la masse, les vents catabatiques créés par l'inversion de température génèrent une subsidence de grande échelle à la surface du plateau continental. Toujours par conservation de la masse un flux convergent en altitude transporte de l'énergie des moyennes latitudes vers l'Antarctique (voir section 2.1.2 et figure 2.2). L'orographie de l'Antarctique et en particulier de la Péninsule Antarctique qui constitue une barrière perpendiculaire aux flux zonaux atmosphériques joue également un rôle dans la circulation générale dans l'hémi-

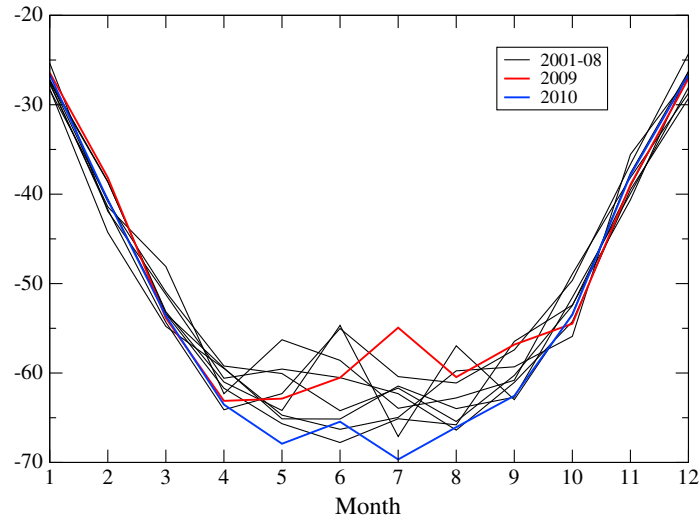


FIGURE 2.7 – Série temporelle de la température mensuelle climatologique à Dôme C, de 2001 à 2010 à environ 180 cm au-dessus de la surface, mesurée par la station automatique AMRC située à 3 km de la tour instrumentée (Genthon et al., 2013)

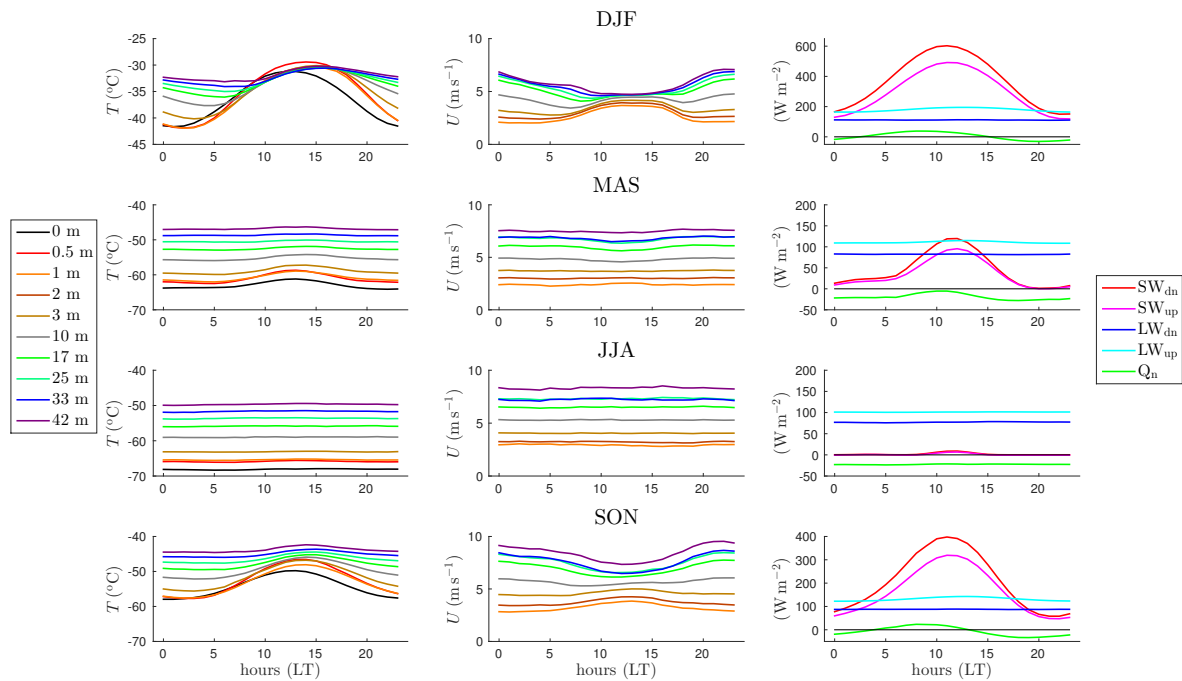


FIGURE 2.8 – Cycles diurnes moyens mesurés à Dôme C entre 2011 et 2016 de température, de vent et de flux radiatifs (Vignon, 2017)

sphère sud en générant des ondes de Rossby (King and Turner, 1997), élément majeur de la météorologie des hautes latitudes mais pouvant se propager dans toute l'hémisphère.

Glace de mer La banquise qui se forme durant l'hiver austral disparaissant presque totalement pendant l'été, on observe une forte variabilité annuelle de la concentration en glace de mer dans l'océan entourant le continent. Celle-ci connaît également une forte variabilité naturelle interannuelle et si celle-ci est restée longtemps largement inconnue (King and Turner, 1997), des progrès récents ont été effectués et le lecteur intéressé pourra par exemple se reporter à Roach et al. (2019). Or le rôle de la glace de mer sur la circulation générale dans les hautes latitudes est essentiel. Son albédo est très largement supérieur à celui de l'océan (90% contre 15%) et modifie donc considérablement le bilan radiatif à la surface. La glace de mer, par son rôle isolant, empêche les échanges de chaleur et d'humidité entre l'océan et l'atmosphère. Maykut (1978) montre ainsi que les échanges de chaleur sont de deux ordres de grandeur plus faibles pour un océan couvert d'une glace jeune. Ces modifications de l'albédo et des échanges avec la surface océanique vont affecter la stabilité de l'atmosphère et la formation des nuages et donc le régime de précipitations aux hautes latitudes. Les vents catabatiques ont un rôle important dans la formation et surtout dans la débâcle de la banquise. Ils jouent en particulier un rôle important dans la formation des polynies côtières (zones libres de glace au sein de la banquise). De façon générale, les échanges entre surface et atmosphère sont compliqués à représenter dans les modèles numériques, en particulier à travers le rôle complexe des polynies (King and Turner, 1997).

Antarctique et circulation océanique L'Antarctique est une des principales zones de formation d'eau profonde (Foster, 1995, Broecker et al., 1998). Sous l'effet de l'air froid en provenance de la calotte glaciaire, la surface océanique dans les zones de polynies se refroidit. Associé à ce refroidissement, lors de la formation de glace de mer en surface, le sel est relâché et la concentration en sel des eaux environnantes augmente. L'augmentation de la densité de ces eaux froides et très salées les fait plonger. Ce mouvement de convection permet ainsi la formation des eaux profondes antarctiques. Cette formation est saisonnière (entre avril et octobre), principalement en mer de Weddell et en mer de Ross (e.g. pour la mer de Ross Jacobs et al., 1970). Ainsi les processus atmosphériques en Antarctique, à travers la génération des vents catabatiques et de leur influence sur les polynies, jouent un rôle important sur la formation des eaux profondes et par là sur la circulation océanique à l'échelle planétaire.

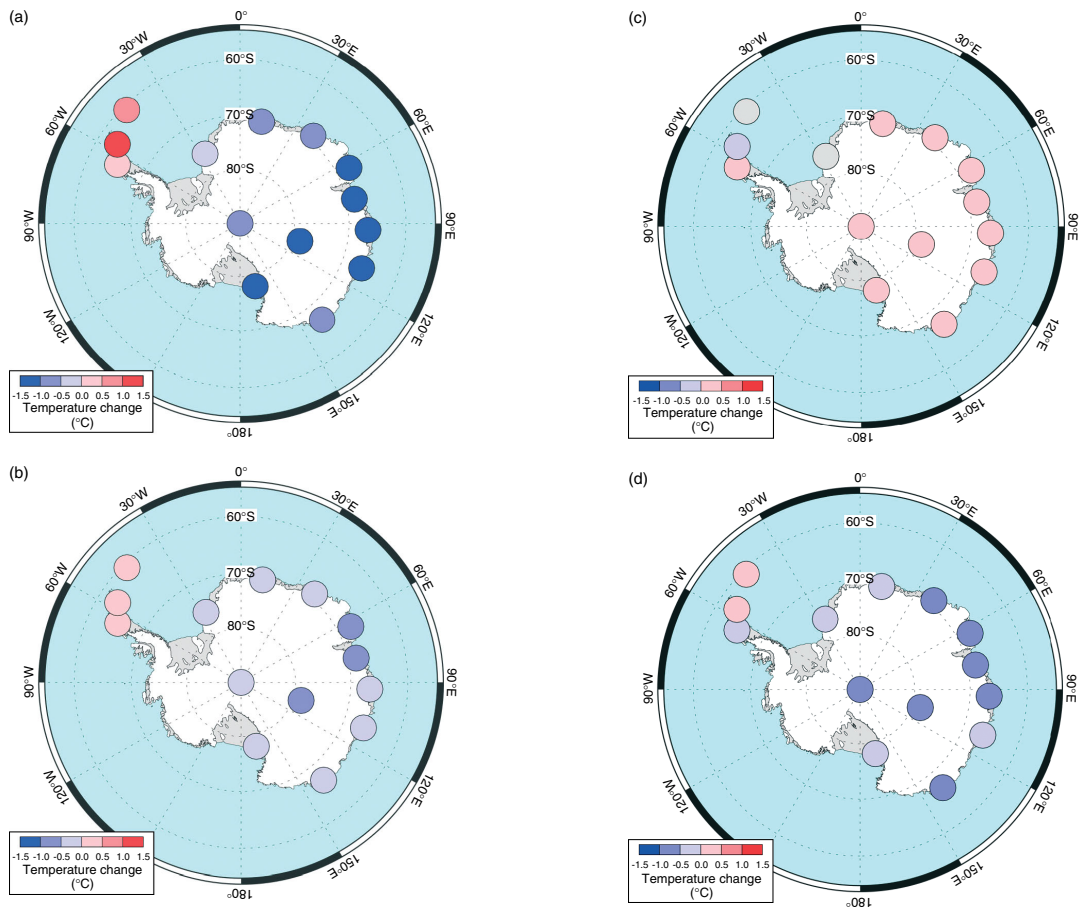
2.1.5 Changement climatique et Antarctique

Amplification polaire La réponse d'une région ou d'un système au changement climatique dépend fortement de l'existence de rétroaction, qui peut être positive ou négative. Une rétroaction positive à un phénomène donné correspond à une réponse du système qui tend à amplifier ce phénomène alors qu'une rétroaction négative tend à ramener le système vers un état d'équilibre. L'amplification polaire, qui est le phénomène par lequel un changement dans le forçage radiatif (écart entre le rayonnement solaire reçu par la Terre et le rayonnement infrarouge qu'elle émet sous l'effet de facteurs d'évolution du climat, calculé au sommet de l'atmosphère) induit une réponse plus prononcée dans les régions polaires que sur l'ensemble

de la planète en est un exemple bien connu. Formellement, elle est calculée comme le quotient entre le réchauffement mesuré dans une région polaire (arctique ou antarctique) et le réchauffement mesuré à l'échelle de la planète. Ce phénomène est lié à rétroaction positive entre température et albédo de surface pour les régions qui sont recouvertes de neige ou de glace de façon saisonnière : l'augmentation du forçage radiatif tend à une diminution de la couverture de neige ou de glace et donc à une diminution de l'albédo qui entraîne une augmentation du forçage radiatif à la surface... Le comportement de l'Arctique et de l'Antarctique vis-à-vis de cette rétroaction est très différent et s'explique par la géographie respective des deux régions polaires. L'Arctique est essentiellement un océan entouré de terres émergées, alors que l'Antarctique est un continent entouré d'océan. L'Arctique comporte ainsi de nombreuses terres émergées avec une couverture de neige qui varie saisonnièrement. L'augmentation des températures de surface dans cette région tend naturellement à restreindre la couverture de neige et donc à diminuer l'albédo de la région. Le même phénomène s'observe avec la glace de mer. Pendant l'été, le réchauffement de la température superficielle réduit l'extension de la glace de mer et par là-même l'albédo de l'Océan Arctique. Cette diminution de l'albédo de la région Arctique pendant l'été induit donc une rétroaction positive au changement climatique et explique pour partie l'amplification polaire arctique observée. La situation est différente en Antarctique. La calotte glaciaire Antarctique couvre de façon permanente la quasi totalité du continent ce qui empêche la rétroaction température/albédo de se mettre en place. La rétroaction due à la glace de mer existe comme dans l'Arctique, mais du fait de la disparition presque complète de la banquise antarctique durant l'été, celle-ci est plus limitée. Les mécanismes à l'origine de l'amplification polaire ne sont pas complètement compris et restent sujet à débat (IPCC, 2019). On peut citer tout de même le rôle de l'AOI (voir plus bas) et une connection avec les températures de surface de la mer dans le Pacifique (Ding et al., 2011, Bromwich et al., 2013). L'ordre de grandeur de l'amplification polaire Antarctique est estimé autour de 1.2 (Collins et al., 2013, Masson-Delmotte et al., 2006).

Le rôle de l'AOI dans le changement climatique en Antarctique Le renforcement du vortex polaire sous l'effet de la dégradation de la couche d'ozone au dessus du pôle sud a eu comme conséquence une augmentation de l'AOI au cours des dernières décennies (Arblaster and Meehl, 2006, Polvani et al., 2011, Smith et al., 2013). La relation entre l'AOI et la température explique ainsi en partie la variabilité spatiale du changement climatique observé sur le continent Antarctique avec un réchauffement sur la Péninsule antarctique et un refroidissement sur l'Antarctique Est pendant les saisons automnales et dans une moindre mesure hivernales (Marshall, 2007). La figure 2.9 illustre l'influence de la tendance de l'indice SAM (similaire à l'indice AOI) sur le changement de température pour la période 1957-2004. Le lien entre une tendance positive de l'AOI et le réchauffement sur une station donnée peut s'expliquer des effets de foehn renforcés par l'intensification des vents d'ouest.

Bilan de masse et contribution à l'élévation du niveau de la mer La question de la contribution de l'Antarctique à la montée du niveau de la mer est cruciale : une fonte totale de la calotte Antarctique correspondrait à une élévation du niveau de la mer de près de soixante mètres, avec des contributions respectives de l'Antarctique Ouest, l'Antarctique Est et la Péninsule Antarctique de 4.3, 53.3 et 0.2 mètres (Fretwell et al., 2013). La perte de masse de l'inlandsis contribue dorénavant et déjà à l'élévation du niveau de la mer (IPCC,



estimated change in Antarctic near-surface temperatures for 1957–2004 caused by trends in the SAM: (a) autumn (b) winter (c) spring and (d) summer.

FIGURE 2.9 – Changement de température à 2 m en Antarctique dû à la tendance de l'indice SAM pour la période 1957–2004 : (a) automne ; (b) hiver ; (c) printemps et (d) été. Marshall (2007)

2019). Le bilan de masse est le résultat d'un équilibre fin entre précipitations et décharge des glaciers. Ces deux phénomènes dépendent potentiellement du changement climatique. En Antarctique, la perte de masse de la calotte a été de $-199 \pm 26 \text{ Gt an}^{-1}$ sur la période 2012-2016, ce qui représente une accélération significative par rapport aux décennies précédentes ($-82 \pm 27 \text{ Gt an}^{-1}$ sur la période 2002 à 2011 et $-51 \pm 73 \text{ Gt an}^{-1}$ sur la période 1992 à 2001). Cette perte de glace est dominée par l'accélération et l'amincissement rapide des principaux glaciers de sortie de l'inlandsis de l'Antarctique occidental (WAIS), entraînés par la fonte des plates-formes glaciaires par les eaux plus chaudes de l'océan.

2.2 Les couches limites stables

2.2.1 La couche limite atmosphérique

2.2.1.1 Définition et importance de la couche limite atmosphérique

La couche limite atmosphérique (CLA) est la partie de l'atmosphère qui subit l'influence de la surface avec un temps de réponse relativement court (de l'ordre d'une heure au plus suivant la définition donnée dans Stull (1988)). L'atmosphère libre, qui la surplombe, réagit également à la surface, mais avec des temps de réponse sensiblement plus longs. La couche limite atmosphérique est le lieu des échanges d'énergie (chaleur sensible, quantité de mouvement) et de matière (vapeur d'eau, aérosols) avec la surface. C'est également la partie de l'atmosphère dans laquelle se déroule la quasi-totalité des activités humaines. Une bonne représentation des différents phénomènes et processus au sein de la CLA est ainsi essentielle pour réaliser des prévisions correctes de température, de brouillard ou d'épisodes de pollution. Le problème de la représentation de la CLA est également crucial dans la modélisation numérique du climat d'une part pour l'interaction avec la surface (océanique et continentale), mais également pour ses interactions avec les phénomènes convectifs (et plus généralement avec la représentation de nuages), ces dernières restant, selon le rapport IPCC (2013), une source majeur d'incertitude dans les projections climatiques.

2.2.1.2 Caractéristiques de la couche limite

Turbulence La couche limite atmosphérique est le lieu de mouvements turbulents. Ceux-ci ont des origines diverses (instabilité de l'atmosphère du fait des forces de flottabilité, cisaillement vertical de vent). Ces mélanges turbulents sont réalisés sur une grande variété d'échelles qui engendrent des structures allant de quelques centimètres à quelques kilomètres, voir quelques dizaines de kilomètres (phénomène de rues de cumulus par vent d'Autan par exemple). L'intensité du mélange turbulent dans l'atmosphère est mesurée par l'énergie cinétique turbulente (ou TKE pour Turbulent Kinetic Energy).

Cycle diurne et hauteur de la couche limite Au-dessus des surfaces continentales et par temps de ciel clair, la CLA peut être vue comme la partie de l'atmosphère pour laquelle

un cycle diurne marqué est observé à la différence de l'atmosphère libre (voir figure 2.10a). Une autre façon de visualiser le cycle diurne dans la couche limite est de regarder l'évolution de la TKE. La figure 2.10b représente cette évolution sur 48h et sur les 1500 premiers mètres d'altitude. On constate qu'un mélange turbulent intense se met en place entre 12h et 18h sur plus de 1000 m de hauteur, et que celui-ci s'atténue fortement durant la nuit et est alors limité à quelques centaines de mètres.

En journée, la surface se réchauffe sous l'effet du rayonnement solaire, devient plus chaude que l'atmosphère. Le gradient de température potentielle près de la surface devient négatif (atmosphère instable), déclenchant des mouvements de convection. La couche limite est alors dite convective et présente une structure mélangée sur toute sa hauteur avec de très faibles gradients verticaux de température potentielle, de vent ou d'humidité, du fait du brassage de l'atmosphère par les mouvements convectifs (matérialisés sur la figure 2.10b par une énergie cinétique turbulente importante). La hauteur d'une telle couche limite est relativement bien définie par la hauteur sur laquelle les gradients de température potentielle, vent ou humidité sont quasi-nuls (voir figure 2.12). La hauteur typique d'une couche limite convective au dessus des continents est de l'ordre de 1 à 2 km mais peut atteindre 4 à 5 km au dessus du Sahara par exemple. Sur la plateau continental antarctique, durant l'été austral, un cycle diurne se met en place et une couche limite convective peut se développer. Mais du fait de l'angle d'incidence important du rayonnement solaire direct et de l'albédo élevé de la surface qui limitent le réchauffement de celle-ci, la couche limite diurne convective n'atteint que quelques centaines de mètres (Pietroni et al., 2012, Casasanta et al., 2014).

Durant la nuit, par temps clair, les pertes d'énergie de la surface par rayonnement thermique ne sont plus compensées par le rayonnement solaire, la surface se refroidit et devient plus froide que la couche d'atmosphère qui la surplombe. Le gradient de température potentielle devient positif et la CLA devient stable (*i.e.* une particule d'air qui subit un déplacement vers le haut ou vers le bas est rappelée à sa position d'origine sous l'effet des forces de flottabilité). La hauteur typique de la CLA nocturne est entre 100 et 300 m. La hauteur d'une couche limite nocturne très stratifiée comme celles rencontrées pendant l'été sur le plateau antarctique est réduite à quelques dizaines de mètres, voire moins (Pietroni et al., 2012, Petenko et al., 2016).

Écoulement Dans l'atmosphère libre, l'écoulement vérifie l'équilibre géostrophique (*i.e.* équilibre entre les forces de pression et la force de Coriolis). Le vent est alors grossièrement parallèle aux lignes isobares. Dans la CLA, des forces de frottement avec la surface apparaissent et se met alors en place un équilibre à trois forces (pression, Coriolis et frottement) qui s'éloigne sensiblement de l'équilibre géostrophique (Malardel, 2005). Svensson and Holtslag (2009), Lindvall and Svensson (2019) montrent en particulier la dépendance de la rotation du vent dans le cas d'une CLA stable à l'intensité de stratification (voir Rysman et al. (2016) pour le cas de Dôme C).

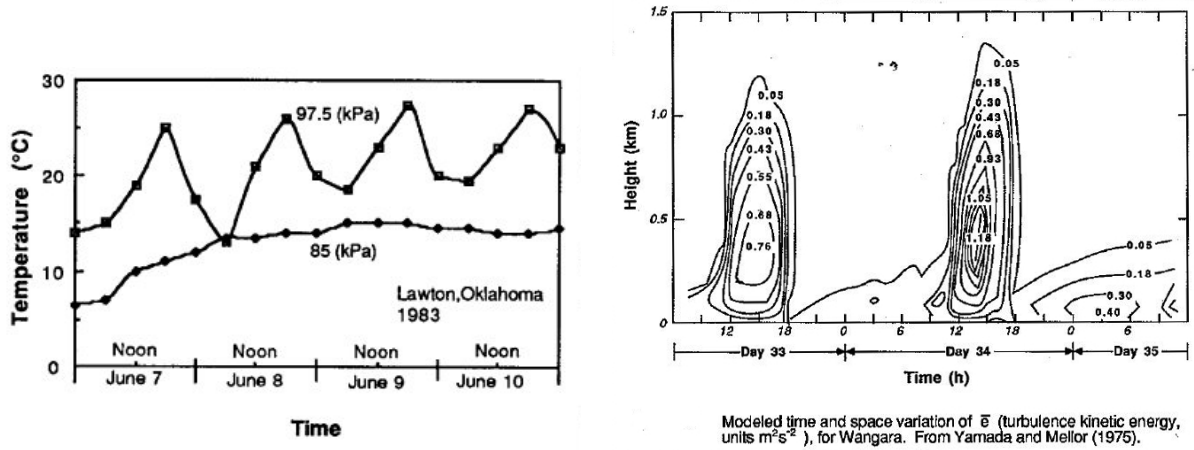


FIGURE 2.10 – Cycle diurne dans la couche limite (Stull, 1988). Cycle diurne et TKE.

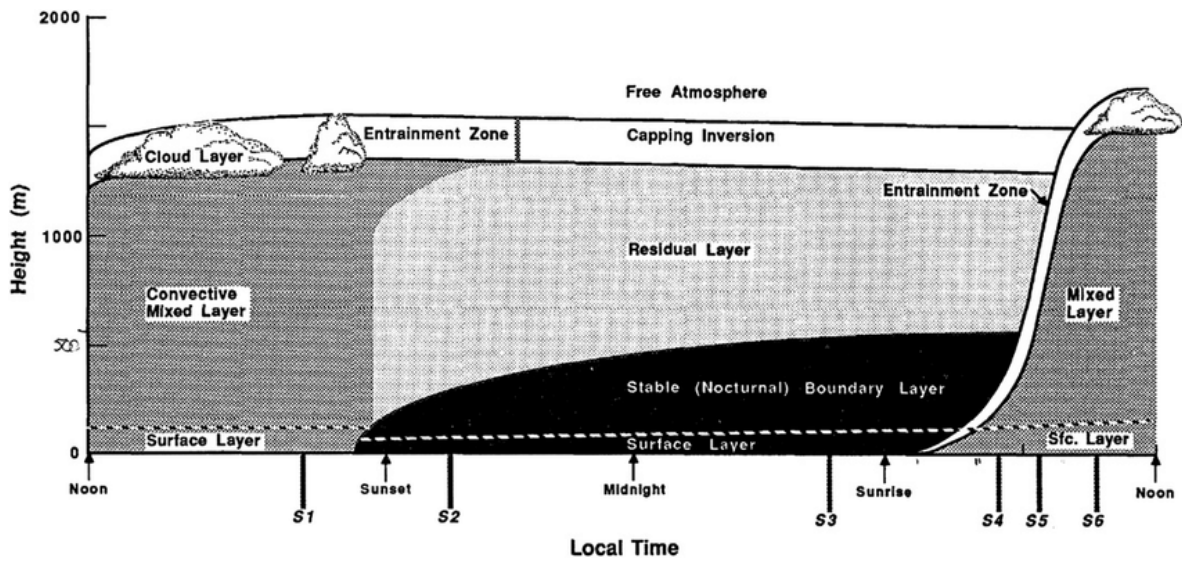


FIGURE 2.11 – Évolution typique de la structure de la couche limite atmosphérique sur un cycle diurne (Stull, 1988).

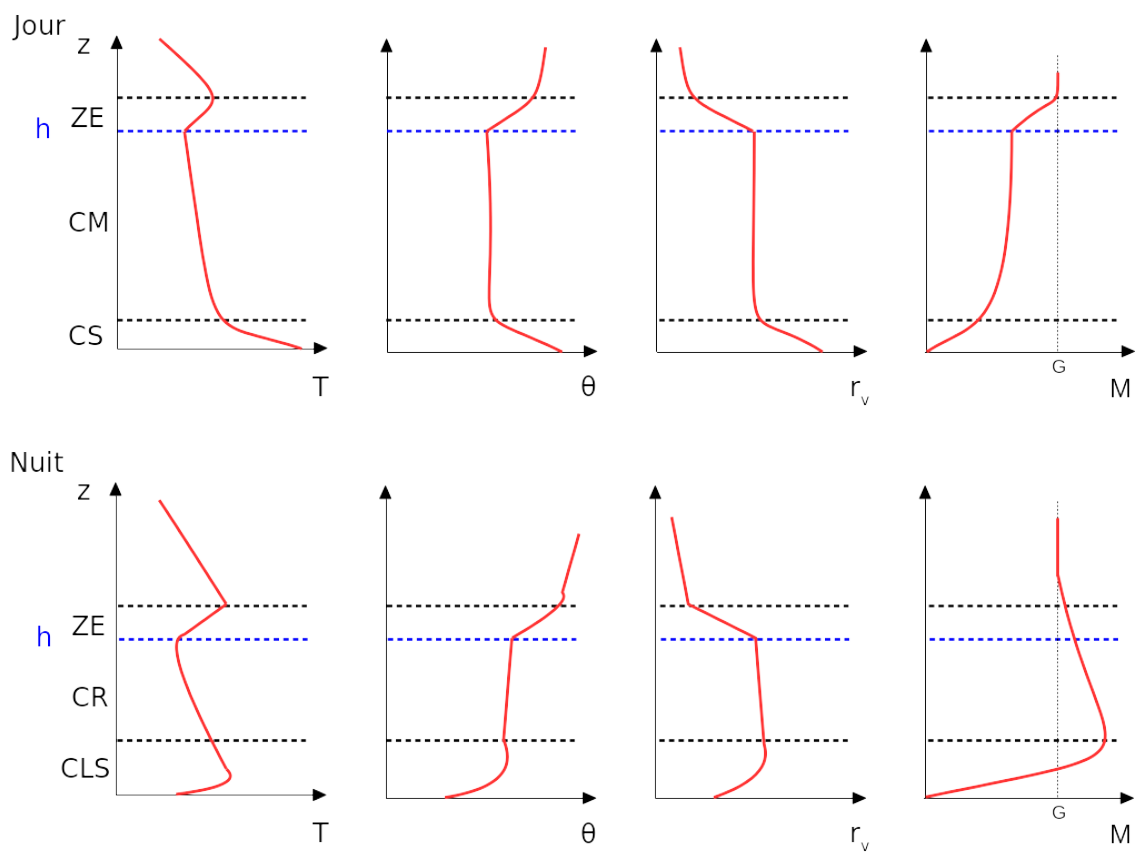


FIGURE 2.12 – Profils verticaux moyens de température, de température potentielle, de rapport de mélange et de vent le jour vers 15 heures locales et la nuit vers 3 heures. G représente la valeur du vent géostrophique et h représente la hauteur de la couche limite. Rodier (2017) d’après Stull (1988).

2.2.2 Les couches limites stables

2.2.2.1 Définition et conditions favorables au développement de couches limites stables

Lorsque la surface est plus froide que la couche d'air qui la surplombe, le flux de chaleur sensible dans les premières couches de l'atmosphère devient négatif, soit un flux de chaleur de l'atmosphère vers le sol, et une stratification thermique se développe dans la couche limite. Le gradient de température est positif (augmentation de la température avec l'altitude) et on parle alors d'inversion. Cette stratification thermique inhibe ainsi les déplacements verticaux au sein de la CLA et on parle de couche limite stable (CLS). L'intensité de l'inversion est en général définie comme la différence de température entre le point de mesure au sol (en général 2 m, mais 1.5 m dans Connolley (1996)) et le maximum de température dans la troposphère.

Les CLS se rencontrent le plus souvent sur continent pendant la nuit, par ciel clair et vent faible. Dans ces conditions, la surface se refroidit, son rayonnement infrarouge n'étant plus compensé ni par rayonnement solaire, ni par le rayonnement thermique des nuages, ni par des flux turbulents de chaleur sensible. Les conditions anticycloniques en hiver favorisent la persistance des CLS : le vent est faible (mélange turbulent très faible même le jour), le ciel est clair (peu de rayonnement infrarouge de l'atmosphère). Ces conditions, associées à une subsidence d'air froid, et un sol couvert de neige (albédo important) favorisent la persistance de CLS sur plusieurs jours. Les régions polaires sont particulièrement propices au développement de CLS, et du fait de la nuit polaire, celles-ci peuvent perdurer plusieurs mois.

2.2.2.2 Classification des CLS

Les CLS sont en général classées selon l'intensité de la stratification thermique : plus le gradient vertical de température potentielle est élevé, plus la CLS est stratifiée et plus elle est stable. On peut définir plusieurs régimes de CLS en fonction de l'intensité de la stabilité. La classification de ces différents régimes varie selon les auteurs (Mahrt, 1998, Sun et al., 2012, Mahrt, 2014). Une revue des différentes méthodes de classification sort du cadre de ce travail. On reprend dans la suite la classification proposée par Mahrt (2014) en deux régimes : un régime faiblement stable et un régime très stable.

CLS faiblement stratifiées Ce type de CLS se développent lorsque les conditions météorologiques limitent le refroidissement radiatif de la surface (e.g., des nuages qui émettent un rayonnement thermique) ou lorsque le vent est suffisant pour maintenir un flux turbulent de chaleur sensible relativement important. Ce type de CLS est plutôt bien compris et les gradients verticaux de quantité de mouvement, de température potentielle ou d'humidité spécifique sont correctement décrits par la théorie des similitudes de Monin-Obukhov. La hauteur des CLS faiblement stratifiées est bien définie et son ordre de grandeur est de 100 m. Ces CLS sont caractérisées par une turbulence homogène, stationnaire et relativement intense.

CLS très stratifiées Elles se développent en général par ciel clair et vent faible. Le gradient vertical de température potentielle est important. La turbulence est faible et peut devenir intermittente (Mahrt, 1999, 2014). On peut observer un découplage de l’atmosphère avec la surface. La théorie de Monin-Obukhov ne s’applique plus (Ha et al., 2007). La hauteur d’une telle CLS devient difficile à définir et son ordre de grandeur autour de la dizaine de mètres. Ces CLS se caractérisent en particulier par un fort refroidissement radiatif de la surface ($> -1 \text{ K h}^{-1}$).

Transition entre les deux régimes Par ciel clair, le critère le plus déterminant pour passer d’un régime à l’autre est la vitesse du vent. Van de Wiel et al. (2012) propose un seuil sur le module du vent qui sépare de façon assez abrupte les deux régimes de CLS. Vignon et al. (2017b) revisite la proposition de Van de Wiel et al. (2012) et l’applique à la couche limite stable du cas d’étude GABLS4 sur le plateau continental antarctique.

2.2.2.3 Processus en œuvre dans les couches limites stables

Refroidissement radiatif Les couches limites stables se développent lorsque la température potentielle de la surface est plus froide que celle de l’atmosphère. On détaille ici un peu plus le processus de refroidissement de la surface et de l’atmosphère.

Durant la nuit, le bilan radiatif de la surface est négatif. En effet, le rayonnement thermique émis par la surface est plus important que le rayonnement thermique reçu émis par l’atmosphère. Le bilan radiatif de la surface pendant une nuit claire est de l’ordre de -100 W m^{-2} (Stull, 1988). Les flux de chaleur sensible (par conduction) au sein du sol compensent partiellement la perte d’énergie de la surface. Le refroidissement se propage alors sous forme de flux turbulent de chaleur sensible de la surface à l’atmosphère.

Dans l’atmosphère, l’équation d’évolution de la température potentielle s’écrit :

$$\frac{\partial \bar{\theta}}{\partial t} = -\frac{\partial}{\partial z} \overline{w'\theta'} - \frac{\partial \bar{R}}{\partial z} - \bar{w} \frac{\partial \bar{\theta}}{\partial z} - \bar{u} \frac{\partial \bar{\theta}}{\partial x} - \bar{v} \frac{\partial \bar{\theta}}{\partial y} \quad (2.2)$$

Le premier terme correspond à la divergence du flux turbulent de chaleur sensible au sein de la couche limite. Le deuxième terme correspond à la divergence du flux radiatif infrarouge. Son ordre de grandeur est en général de -0.1 K h^{-1} (Stull, 1988). Aux hautes latitudes et pendant la nuit polaire, ce terme peut être bien plus élevé. Hoch et al. (2007) montre qu’au Groenland, dans les premières dizaines de mètres de l’atmosphère, une valeur typique du refroidissement dû à la divergence du flux radiatif est plutôt de l’ordre de -10 K j^{-1} et peut aller jusqu’à -35 K j^{-1} . Le terme d’advection verticale $-\bar{w} \frac{\partial \bar{\theta}}{\partial z}$ peut être une source de biais dans les modèles numériques. En particulier dans les anticyclones, la subsidence ($\bar{w} < 0$) peut être sous-estimée, et ainsi conduire à un biais chaud dans les premières couches du modèle (Holtslag et al., 2013, Sandu et al., 2013, Vignon et al., 2018). Les termes d’advection horizontale $-\bar{u} \frac{\partial \bar{\theta}}{\partial x}$ et $-\bar{v} \frac{\partial \bar{\theta}}{\partial y}$ dépendent bien sûr de la situation météorologique. Sur le plateau continental antarctique, l’absence de flux synoptique marqué n’implique pas que les termes

d'advection soient négligeables et peuvent résulter alors d'écoulements catabatiques. Dans le cas de CLS très stratifiées, les différents termes de l'équation 2.2 sont du même ordre de grandeur et aucun ne peut être négligé, ce qui est une des sources de la difficulté à simuler de telles situations avec un modèle de circulation générale.

Hétérogénéités de surface Dans les CLS les structures turbulentes peuvent être influencées par des hétérogénéités de la surface (Derbyshire, 1995, Mahrt, 1999). Les effets de ces hétérogénéités peuvent être mal représentées dans les modèles numériques (Holtslag et al., 2013, Sandu et al., 2013, Vignon et al., 2018). L'introduction d'une turbulence augmentée est une façon de palier cette lacune.

Turbulence intermittente Dans les CLS très stratifiées ($Ri_g \gg 1$), la turbulence devient intermittente, *i.e.* en un point donné, l'intensité de la turbulence varie de façon importante, pouvant quasiment s'annuler avant d'augmenter à nouveau (Van de Wiel et al., 2003). Ce phénomène peut être dû à des phénomènes externes de sous-méso échelle (Vercauteren et al., 2019, Mahrt, 2014, Mortarini et al., 2018) mais également apparaître spontanément et résulter de l'interaction le cisaillement du vent et la stratification de la couche.

Interaction avec les ondes de gravité Les interactions entre la turbulence et les ondes sont complexes et sont l'objet d'une littérature importante (Mahrt, 2014, Sun et al., 2015). On notera surtout que les CLS fortement stratifiées canalise les ondes et favorisent leur propagation. Ces dernières en retour peuvent tout aussi bien augmenter que inhiber la turbulence dans les CLS (voir section précédente sur la turbulence intermittente).

2.2.2.4 Le jet nocturne de basse couche

Le jet nocturne de basse couche est une zone de vent fort qui se forme durant la nuit. Il est mis en évidence pour la première fois dans les grandes plaines américaines (Blackadar, 1957). Il se forme généralement dans des conditions qui favorisent un fort refroidissement radiatif (ciel clair, atmosphère sèche). Il commence alors à se former dès la tombée de la nuit et atteint son maximum d'intensité en deuxième partie de nuit. Il disparaît peu après le lever du jour lorsque la surface se réchauffe et que la convection se met en place mélangeant le gradient vertical de vent. Son intensité peut être supérieure à celle du vent géostrophique (on parle alors de vent supergéostrophique). Le jet se situe en général dans les 500 premiers mètres de la couche limite.

La théorie inertielle du jet nocturne proposée par Blackadar (1957) propose un modèle conceptuel simple en se plaçant dans le cas d'une couche limite homogène horizontalement (gradients horizontaux nuls). Les équations de mouvement s'écrivent alors :

$$\frac{\partial \bar{u}}{\partial t} = f(\bar{v} - v_g) - \frac{\partial}{\partial z} \overline{u'w'} \quad (2.3)$$

$$\frac{\partial \bar{v}}{\partial t} = -f(\bar{u} - u_g) - \frac{\partial}{\partial z} \overline{v'w'} \quad (2.4)$$

où u_g et v_g sont les composantes zonales et méridiennes du vent géostrophique et $f = 2\Omega \sin \phi$ est le paramètre de Coriolis (ϕ étant la latitude du point considéré).

Lorsque l'atmosphère se stratifie en cours de nuit, les terme de stress turbulent $\overline{u'w'}$ et $\overline{v'w'}$ disparaissent et en posant $U = \bar{u} + i\bar{v}$ et $U_g = \bar{u}_g + i\bar{v}_g$, une simple intégration donne la solution suivante :

$$U(t) - U_g = [U(t=0) - U_g] e^{-ift} \quad (2.5)$$

où t_o est l'instant où le bilan radiatif à la surface devient négatif (soit en première approximation le coucher du soleil). L'équation 2.5 montre que l'intensité du jet est d'autant plus grande que l'écart initial au vent géostrophique est grand et que le jet a une période de rotation de $T = \frac{2\pi}{f}$ soit d'un ordre de grandeur de 18 heures pour les moyennes latitudes. Pour Dôme C (75°06'S), le paramètre de Coriolis vaut $f = 1.4 \cdot 10^{-4} \text{s}^{-1}$, et correspond à une période $T \simeq 12.5$ h. Cette théorie très simple explique donc que le maximum d'intensité du jet ait lieu 9 heures environ après le coucher du soleil.

D'autres travaux plus récents ont amélioré et complété le modèle inertiel de Blackadar (1957), en prenant en compte effets de pente et forces de friction (Shapiro and Fedorovich, 2010, Van de Wiel et al., 2010, Shapiro et al., 2016). Mais le modèle initial est suffisant pour notre propos et pour comprendre la phénoménologie du jet nocturne de basses couches.

2.2.2.5 Les CLS en Antarctique

L'Antarctique est un terrain propice pour la formation de CLS intenses en particulier sur la calotte glaciaire de l'Antarctique Est. Par exemple Connolley (1996) met en évidence à Vostok (79°S, 107°E, 3420 m) une inversion moyenne durant l'hiver austral (mai à octobre) de 25 K. Les inversions y sont de plus d'une profondeur réduite (typiquement quelques dizaines de mètres). Vignon et al. (2017b) propose une étude détaillée des couches limites stables sur le site de Dôme C et met en évidence un seuil pour le vent à 10 m en dessous duquel l'atmosphère est découplée de la surface. La turbulence devient intermittente et l'inversion de température se renforce, avec des gradients moyens de 2.5 m^{-1} sur les dix premiers mètres au dessus de la surface. La modélisation de telles couches limites restent aujourd'hui un défi pour les modèles de climat. L'étude du GABLS4 servira de banc d'essai pour tester la capacité de la physique du modèle ARPGE-Climat à représenter ces couches limites extrêmes.

Outils de modélisation

Sommaire

3.1	Modélisation du climat	31
3.2	Le modèle ARPEGE-Climat	32
3.2.1	Présentation du modèle	32
3.2.2	La paramétrisation de la turbulence dans ARPEGE-Climat 6.3	34
3.2.3	Autres paramétrisations	37
3.3	Représentation de la surface : la plateforme SURFEX	37
3.3.1	La représentation du manteau neigeux dans SURFEX	38
3.3.2	Flux turbulents sur les surfaces continentales	40
3.4	Différents types de simulations	42
3.4.1	SCM	42
3.4.2	Simulations 3D	42
3.5	Données de validation	44
3.5.1	Les jeux de données de référence	44
3.5.2	Simulations LES	45
3.6	Conclusion	47

3.1 Modélisation du climat

Un modèle de circulation générale vise à simuler les écoulements géophysiques de l’atmosphère globale. Ce type de modèle permet de faire des prévisions météorologiques – dans ce cas, on parle de modèle de prévision numérique du temps (PNT) – ou de faire des simulations climatiques couvrant de longues périodes (quelques années à plusieurs millénaires). Dans ce dernier cas, du fait de la longueur des périodes simulées, il devient nécessaire de prendre en compte les composantes plus lentes du système climatique (e.g., l’océan, la glace de mer, les cycles biogéochimiques). L’assemblage de ces composantes permet de construire des modèles couplés océan-atmosphère, voire des modèles plus complets appelés Earth System Models (ESM).

Un modèle de circulation générale atmosphérique se compose essentiellement de deux composantes, une partie dynamique et une partie physique. La partie dynamique, appelée noyau dynamique, a pour rôle de résoudre les équations de la dynamique des fluides qui régissent le comportement de l’atmosphère. Plus précisément, le noyau dynamique résout une version discrétisée de ces équations sur la grille du modèle à l’aide de méthodes numériques

(e.g., différences finies, approches spectrales). La partie physique est composée d'un ensemble de paramétrisations qui modélisent et calculent l'impact des processus d'échelles plus petites sur la partie résolue par le noyau dynamique. On peut ainsi résumer un modèle de climat par l'équation suivante :

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{D}(\mathbf{x}) + \sum_p \mathcal{P}_p(\mathbf{x}, \boldsymbol{\lambda}_p) \quad (3.1)$$

où \mathbf{x} est la variable d'état du modèle atmosphérique (contenant par exemple les composantes 3D du vent, la température, l'humidité spécifique, les concentrations des différentes espèces de l'eau condensée), \mathcal{D} représente le noyau dynamique, et \mathcal{P}_p représente le terme source associé à la paramétrisation physique p . Chaque paramétrisation physique est représentée ici comme une fonction de la variable d'état du modèle, ainsi que d'un ensemble $\boldsymbol{\lambda}_p$ de paramètres libres.

3.2 Le modèle ARPEGE-Climat

3.2.1 Présentation du modèle

ARPEGE (Action de Recherche Petite Échelle Grande Échelle) est un modèle atmosphérique global développé par le Centre National de Recherches Météorologiques (CNRM) utilisé pour des applications PNT et climat. Le code de ARPEGE est développé conjointement avec le modèle IFS (Integrated Forecasting System) du Centre Européen pour les Prévision Météorologiques à Moyen Terme (CEPMMT). La version PNT de ARPEGE évolue régulièrement, avec des cycles de convergence tous les 6 à 12 mois avec le modèle IFS.

La première version de ARPEGE-Climat utilisée pour les applications climatiques date des années 90 (Déqué et al., 1994). La version utilisée dans ce travail, ARPEGE-Climat 6.3 (Roehrig et al., 2020), est la composante atmosphérique du modèle couplé océan-atmosphère CNRM-CM6-1 (Voldoire et al., 2019) et du modèle Système Terre CNRM-ESM2-1 (Séférian et al., 2019). Ces deux modèles ont été développés par le CNRM et le Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS) pour participer à la sixième phase du projet d'intercomparaison de modèles couplés (Coupled Model Intercomparison Project Phase 6 – CMIP6).

Les échéances pour la production de simulations climatiques et le temps nécessaire pour effectuer les centaines d'années de simulations nécessaires pour étudier le changement climatique ne permettent pas des évolutions trop rapides du code du modèle. ARPEGE-Climat utilise donc une version du code "gelée" pour plusieurs années. La version ARPEGE-Climat 6.3 est basée sur le cycle 37 de ARPEGE/IFS (2010). Elle diffère également de la version PNT par sa résolution horizontale plus grossière (pour des raisons de coût de calcul,) et par l'ajout de différents couplages avec les composantes plus lentes du système climatique (e.g., océan, glace de mer). Ces couplages, généralement moins pertinents pour la PNT du fait de l'inertie des composantes du système climatique mises en jeu, sont cependant cruciaux pour des applications climatiques.

ARPEGE-Climat 6.3 est décrit dans Roehrig et al. (2020). Par rapport au cycle 37, ont été introduit des développements de paramétrisations physiques plus récents. Certaines paramétrisations sont issues directement de la version PNT de ARPEGE (turbulence et microphysique), dans un effort de convergence entre les modèles et dans le cadre d'une stratégie *seamless* ("sans couture") dans laquelle les prévisions aux différentes échéances (prévision du temps, prévision saisonnière, projections climatiques) doivent être compatibles aux points de vue de leur conception de base, de leur contenu et de leur format (Météo-France, 2014). La paramétrisation de la convection utilisée (Piriou et al., 2018) a été développée initialement pour les versions PNT et Climat mais n'a finalement été intégrée que dans la version Climat.

Le noyau dynamique est présenté succinctement dans le paragraphe suivant. Les paramétrisations sont présentées brièvement dans les sections qui suivent, en focalisant sur celles qui sont le plus pertinentes dans le cadre de cette thèse.

Le noyau dynamique de ARPEGE-Climat Basé sur le cycle 37 de ARPEGE/IFS, le code résout les équations dites primitives. Celles-ci sont obtenues à partir des équations fondamentales de la mécanique des fluides en faisant un certain nombre de simplifications (hypothèse hydrostatique, hypothèse de la couche mince). C'est un modèle spectral utilisant une troncature triangulaire de 127 (T127), ce qui correspond à une résolution horizontale d'à peu près 150 km. C'est un modèle "high-top", ce qui signifie que la stratosphère est résolue avec un dernier niveau (top) dans la mésosphère, vers 80 km. La grille verticale comprend 91 niveaux verticaux selon une grille progressive en coordonnées hybrides σ -pression, avec 12 niveaux dans le premier kilomètre (Fig. 3.1). Le modèle utilise un schéma temporel semi-implicite et un schéma d'advection semi-Lagrangien, lui permettant d'utiliser un grand pas de temps de 15 minutes. Le lecteur pourra se reporter à Roehrig et al. (2020) pour plus de détails.

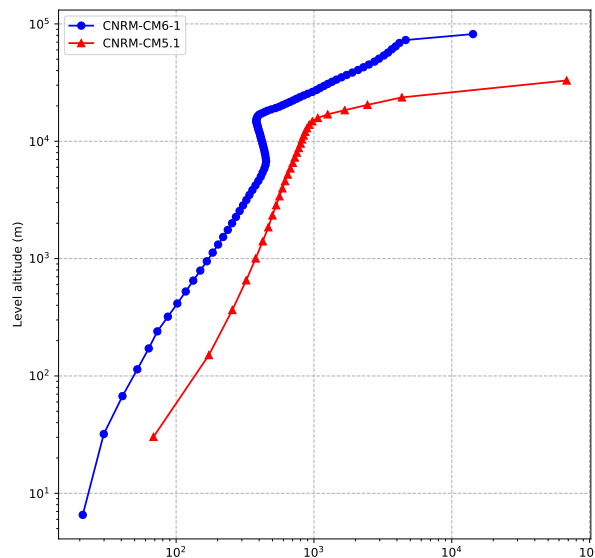


FIGURE 3.1 – Résolution verticale du modèle ARPEGE-Climat 6.3 (en bleu) et de sa version précédente, ARPEGE-climat 5.1 (en rouge).

Les paramétrisations physique dans ARPEGE-Climat La physique du modèle permet de fournir l'effet sur la grande échelle de phénomène de taille inférieure à la grille du modèle ou ne relevant pas de la mécanique des fluides (microphysique, rayonnement). Elle comprend entre autres une paramétrisation de la turbulence, de la microphysique, de la convection, des effets de l'orographie sous-maille ainsi qu'un schéma de transfert radiatif. La paramétrisation de la turbulence (clé pour la représentation des couches limites stables) est présentée en détail dans la suite. Des éléments d'information sont fournis sur quelques paramétrisations également pertinentes pour le travail présenté dans ce manuscrit. Notons que les différentes paramétrisations sont appelées à chaque pas de temps, et calculent une tendance à partir d'un même vecteur d'état. Les tendances des différentes paramétrisations sont additionnées à la fin du pas de temps pour incrémenter le vecteur d'état en une seule fois.

3.2.2 La paramétrisation de la turbulence dans ARPEGE-Climat 6.3

Le schéma de turbulence a pour objectif de calculer les flux turbulents de différentes grandeurs (e.g., quantité de mouvement, chaleur sensible, eau sous ses différentes phases, concentrations des différents aérosols traités par le modèle). La résolution horizontale relativement grossière d'un modèle de climat justifie de ne tenir compte que de la composante verticale des flux turbulents (schéma 1D). Le schéma de turbulence Cuxart, Bougeault, Redelsperger (CBR), utilisé dans ARPEGE-Climat 6.3 suit les travaux de Redelsperger and Sommeria (1982), Redelsperger and Sommeria (1986), et Cuxart et al. (2000). Il repose sur une équation prognostique de l'énergie cinétique turbulente (ECT) \bar{e} et sur le calcul diagnostique de flux turbulents.

Pour chaque variable ψ sensible au mélange turbulent on calcule un flux turbulent correspondant $\overline{w'\psi'}$, où le prime indique une perturbation locale par rapport à une valeur moyenne (dans l'espace et le temps) indiqué par les barres horizontales sur les variables. Les flux turbulents d'une variable ψ donnée sont calculés en faisant l'hypothèse qu'ils sont proportionnels aux gradients verticaux de $\bar{\psi}$ 3.3. Les coefficients de proportionnalité sont appelés coefficients d'échange turbulent (ou coefficients de diffusion turbulente). Ce calcul est effectué pour les composantes moyennes zonale et méridienne du vent \bar{u} et \bar{v} , l'eau liquide totale moyenne $\bar{q}_t = \bar{q}_v + \bar{q}_c$ avec q_v le contenu en vapeur d'eau et q_c le contenu en eau condensée (liquide et solide), la température potentielle liquide θ_l et l'énergie cinétique turbulente moyenne \bar{e} . Pour tenir compte des effets de l'humidité, on utilise la variable conservative θ_l (Betts, 1973) :

$$\bar{\theta}_l = \bar{\theta} \left(1 - \frac{L_v \bar{q}_l + L_f \bar{q}_i}{c_p T} \right) \quad (3.2)$$

où L_v et L_f sont respectivement les chaleurs latentes de vaporisation et de sublimation, \bar{q}_l et \bar{q}_i les contenus moyens en eau liquide et eau solide respectivement. Les flux turbulents s'écrivent :

$$\begin{aligned}
\overline{w'u'} &= -K_m \frac{\partial \bar{u}}{\partial z} & ; & \quad K_m = C_m L_m \sqrt{\bar{e}} \\
\overline{w'v'} &= -K_m \frac{\partial \bar{v}}{\partial z} & ; & \quad K_m = C_m L_m \sqrt{\bar{e}} \\
\overline{w'e'} &= -K_e \frac{\partial \bar{e}}{\partial z} & ; & \quad K_e = C_m L_m \sqrt{\bar{e}} \\
\overline{w'\theta'_l} &= -K_T \frac{\partial \bar{\theta}_l}{\partial z} & ; & \quad K_{\theta_l} = \alpha_T C_m L_m \sqrt{\bar{e}} \phi_T \\
\overline{w'q'_t} &= -K_{qt} \frac{\partial \bar{q}_t}{\partial z} & ; & \quad K_{qt} = \alpha_q C_m L_m \sqrt{\bar{e}} \phi_q
\end{aligned} \tag{3.3}$$

où C_m , α_T et α_q sont des paramètres libres. Dans le schéma CBR, on prend $\alpha_T = \alpha_q$. L_m est la longueur de mélange et est interprétée comme la taille caractéristique des tourbillons les plus énergétiques. ϕ_T et ϕ_q sont des fonctions de stabilité qui permettent de tenir compte des effets de flottabilité sur les flux turbulents de la température potentielle liquide et de l'eau totale. Dans le schéma CBR les fonctions ϕ_T et ϕ_q sont prises égales entre elles et sont calculées comme suit :

$$\phi_T = \phi_q = \frac{1}{1 + C \frac{g}{\theta_{vl}} \frac{L_m^2}{\bar{e}} \frac{\partial \bar{\theta}_{vl}}{\partial z}} \quad \text{où } C \text{ est un paramètre libre.} \tag{3.4}$$

où $\bar{\theta}_{vl} = \bar{\theta}(1 + 0.608\bar{q}_v - \bar{q}_c)$ est la température potentielle virtuelle liquide de Lilly (Lilly, 1968).

Le calcul de la longueur de mélange suit une formulation non locale proposée dans Bougeault and Lacarrère (1989) :

$$L_m^{\text{BL89}} = \left[\frac{1}{2} \left((L_{up})^{-2/3} + (L_{down})^{-2/3} \right) \right]^{-3/2} \tag{3.5}$$

où L_{up} et L_{down} sont respectivement les déplacements maximaux vers le haut et vers le bas qu'une particule pourrait effectuer étant données son énergie cinétique turbulente et la stratification thermique de l'atmosphère environnante, si on ne considère que le travail des forces de flottabilité.

Le calcul des coefficients de diffusion turbulente K_ψ suppose connue l'énergie cinétique turbulente moyenne \bar{e} . Son évolution temporelle est gouvernée par l'équation suivante :

$$\frac{\partial \bar{e}}{\partial t} = -\frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} e' w') - \left(\overline{u'w'} \frac{\partial \bar{u}}{\partial z} + \overline{v'w'} \frac{\partial \bar{v}}{\partial z} \right) + \frac{g}{\bar{\theta}} \overline{w'\theta'_{vl}} - \frac{\bar{e}\sqrt{\bar{e}}}{L_\epsilon} \tag{3.6}$$

où ρ est la densité de l'air, g est l'accélération de la pesanteur, et L_ϵ est la longueur de dissipation. Elle est supposée proportionnelle à la longueur de mélange L_m : $L_\epsilon = C_e \times L_m$ (la formulation utilisée ici diffère de celle utilisée dans Cuxart et al. (2000) et Roehrig et al. (2020)). Le coefficient de proportionnalité C_e est donc un autre paramètre libre du modèle.

Le flux turbulent $\overline{w'\theta'_{vl}}$ est calculé suivant Bechtold et al. (1995) à partir des flux turbulents des variables de Betts $w'\theta'_l$ et $w'q'_t$.

Introduction de limites maintenant un mélange turbulent minimal Au cours de l'implémentation du schéma CBR et de son utilisation dans ARPEGE, différents ingrédients empiriques ont été introduits pour tenter de maintenir un minimum de mélange turbulent. Des bornes inférieures sont appliquées sur trois grandeurs : l'énergie cinétique turbulente, la longueur de mélange, et le mélange turbulent lui-même.

- La valeur minimale pour l'énergie cinétique turbulente est définie par un paramètre e_{\min} . Dans CNRM-CM6, sa valeur est fixée à $10^{-14} \text{ m}^2\text{s}^{-2}$ essentiellement pour des raisons numériques.
- La valeur minimale pour la longueur de mélange, notée L_m^{\min} est un paramètre clé du schéma de turbulence, et de façon générale du modèle ARPEGE, pour la représentation des CLS et de la troposphère libre qui est essentiellement stable. Dans la version PNT de ARPEGE, sa valeur est fixée à zéro alors que dans CNRM-CM6 elle est historiquement fixée à 10 m. La différence de traitement de ce paramètre dans les deux versions de ARPEGE reflète des enjeux différents et la priorité donnée en PNT à la représentation des couches limites. Près de la surface, la taille des tourbillons est contrainte par la distance au sol z et est supposée au moins égale à κz , κ étant la constante de Von Kármán. En tenant compte des ces deux limites, la longueur de mélange se calcule selon la formulation suivante :

$$L_m = \max \left[L_m^{\text{BL89}}, \min(L_m^{\min}, \kappa z) \right] \quad (3.7)$$

- Le mélange turbulent pour θ_l et q_t , u et v est limité directement via la formule suivante :

$$\overline{w'\psi'} = \max \left[-K_\psi \frac{\Delta\bar{\psi}}{\Delta z}, \alpha_\psi \left(K_\Delta \left(1 - \frac{z}{Z_\Delta} \right) \right) \Delta\bar{\psi} \right] \quad (3.8)$$

où $\alpha_\psi = \alpha_T$ pour $\psi = \theta_l$ ou q_t et $\alpha_\psi = 1$ sinon (pour u et v), K_Δ et Z_Δ sont deux paramètres libres, $\Delta\bar{\psi}$ la différence de la grandeur ψ entre deux niveaux verticaux consécutifs du modèle et Δz la différence d'altitude entre ces deux niveaux. Cette limitation n'est pas active au dessus de Z_Δ . Ce paramètre étant fixé dans CNRM-CM6 à 300 m, cette limite ne joue donc un rôle que dans des couches limites stables et peu profondes. Elle a été introduite explicitement historiquement par E. Brun et J.-F. Guérémy pour éviter des boucles de rétroaction qui entraîneraient un fort refroidissement en surface (*runaway cooling*), en particulier sur les surfaces enneigées ou couvertes de glace. Notons que de par sa formulation, cette limitation du mélange turbulent dépend de la résolution verticale du modèle et que tout changement dans cette dernière exige a priori une revisite des valeurs de K_Δ et Z_Δ . Remarquons également que cette limite est identique pour les flux turbulents de température potentielle et de vent. Cette unicité est discutable et pourrait être revue, notamment lors de la calibration du modèle. Toutefois, dans les travaux qui suivent, cette unicité sera conservée.

3.2.3 Autres paramétrisations

On donne ici quelques informations sur la paramétrisation de la convection et sur le schéma de transfert radiatif dans ARPEGE-Climat. Même si ces paramétrisations ne seront pas utilisées dans la suite pour la calibration de ARPEGE-Climat, elles jouent un rôle important sur le cas de couche limite GABLS4 qui sera étudié par la suite.

Paramétrisation de la convection Son rôle est de représenter l'effet de l'ensemble des phénomènes de convection (profonde et peu profonde) sur la grande échelle. La paramétrisation de la convection dans ARPEGE-Climat 6.3 est réalisée par le schéma *Prognostic Condensates Microphysics and Transport* (PCMT), basé sur les travaux de Piriou et al. (2007) et Guérémy (2011). Ce schéma en flux de masse a une fermeture en CAPE (Convective Available Potential Energy, énergie potentielle convective disponible en français). L'évolution de l'ascendance convective dépend de sa flottabilité et de son mélange avec l'environnement par les processus d'entraînement et déentraînement. Ce schéma comprend une microphysique prognostique, cohérente avec celle utilisée pour la composante stratiforme. Le lecteur pourra se référer à Roehrig et al. (2020) pour plus de détail.

Schéma de transfert radiatif Le rôle du schéma de transfert radiatif est de calculer les flux radiatifs en tenant compte des phénomènes de réflexion, de diffusion et d'absorption du rayonnement solaire, ainsi que les phénomènes d'absorption et d'émission des rayonnements infrarouges par l'atmosphère et la surface. Dans ARPEGE-Climat, les deux spectres (solaire et infrarouge) sont traités de manière distincte par deux schémas différents. La partie traitant du rayonnement solaire résout l'équation de transfert radiatif pour la partie du spectre solaire comprise entre 0.2 et $4\ \mu\text{m}$ et est basée sur les travaux de Fouquart and Bonnel (1980). Le code utilisé se base sur une approximation à deux flux (montant et descendant). La partie traitant du rayonnement infrarouge est basée sur le schéma RRTM (*Rapid Radiation Transfer Model*) utilisé par le modèle IFS du CEPMMT et résout l'équation de transfert radiatif également par une approximation à deux flux. Le spectre est divisé en seize bandes spectrales et rend compte de l'absorption du rayonnement par un certain nombre de gaz (H_2O , CO_2 , O_3 ...). Le schéma RRTM repose sur les travaux de Mlawer et al. (1997).

Notons que le schéma de transfert radiatif utilise en entrée les valeurs d'albédo et d'émissivité de la surface calculées par le schéma de surface (voir 3.3).

3.3 Représentation de la surface : la plateforme SURFEX

Généralités Lors d'une simulation réalisée avec un modèle couplant atmosphère et surface, il est nécessaire d'explicitement l'influence que va exercer l'atmosphère sur la surface (via la température, l'humidité spécifique, le vent, la pression, les précipitations, les flux radiatifs) et en retour la façon dont la surface va rétroagir sur l'atmosphère. ARPEGE-Climat ne traitant que de la partie atmosphérique, l'évolution de l'état de la surface ainsi que les échanges de matière et d'énergie entre surface et atmosphère sont traités dans un modèle à part appelé SURFEX

(SURFace EXternalisée). Dans sa version 6.3, ARPEGE-Climat est couplé à SURFEX 8.0 (Lemoigne et al., 2018). SURFEX 8.0 simule les flux moyens entre surface et atmosphère de quantité de mouvement, de chaleur sensible et de chaleur latente, ainsi que de diverses espèces chimiques (voir fig 3.2). SURFEX fournit également au modèle d’atmosphère les propriétés radiatives de la surface (entre autres albédo et émissivité). Ces différentes sorties de SURFEX servent de conditions aux limites pour le modèle atmosphérique auquel il est couplé, en particulier, pour les paramétrisations de la turbulence et pour le schéma de rayonnement.

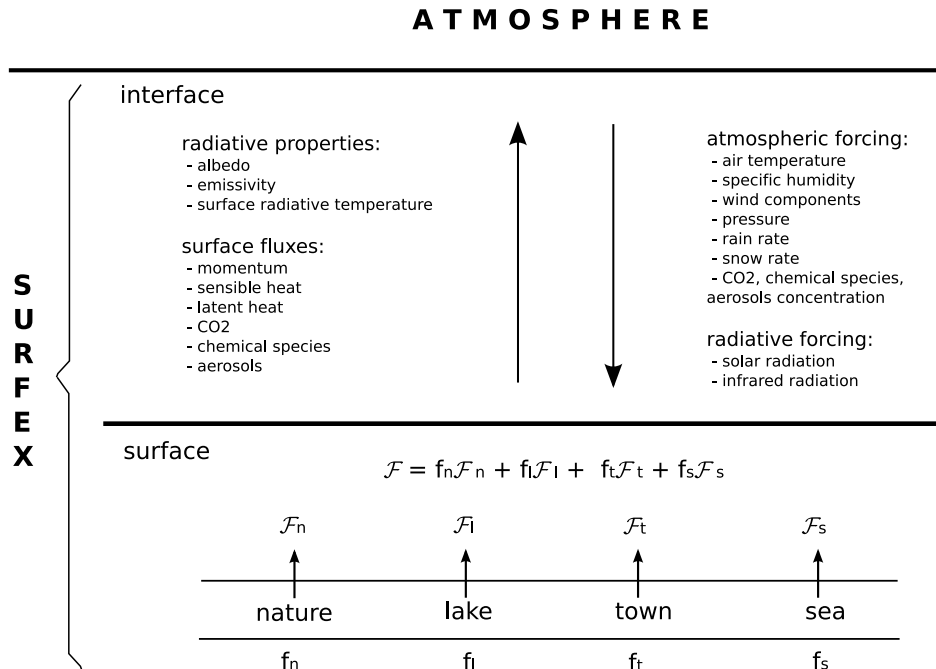


FIGURE 3.2 – Le modèle SURFEX

Les différentes composantes de la plateforme SURFEX Selon le type de surface (en l’occurrence mer/océan, surfaces continentales, lacs), différents sous-modèles de SURFEX sont utilisés. Les lacs sont traités par le modèle Flake (Le Moigne et al., 2016). Les surfaces continentales sont décrites par le modèle ISBA (Interaction Sol-Biosphère-Atmosphère) (Noilhan and Planton, 1989, Noilhan and Mahfouf, 1996). À nouveau, au sein d’ISBA, différents modules sont appelés en fonction du type de surface avec en particulier un module dédié aux surfaces couvertes de neige.

3.3.1 La représentation du manteau neigeux dans SURFEX

La surface du continent antarctique est modélisée dans ARPEGE-CLimat 6.3 est modélisée par des surfaces de neige permanente. On détaille donc ici la façon dont est représenté le manteau neigeux dans le modèle.

Différents niveaux de complexité de représentation du manteau neigeux sont proposés, allant du plus simple (un modèle à une seule couche (Douvillè et al., 1995a,b)) qui a l’avantage

d'être peu gourmand en temps de calcul, au plus compliqué, avec une implémentation du schéma Crocus (Brun et al., 1989, 1992, Vionnet et al., 2012) qui décrit de façon détaillée les processus physiques gouvernant l'évolution du manteau neigeux. Un schéma de complexité intermédiaire, composé de douze couches, a récemment été introduit (Decharme et al., 2016). Il propose entre autre une simplification du schéma CROCUS pour le calcul de l'albédo et de la densification de la neige. Le calcul de l'albédo est détaillé dans le paragraphe suivant.

Calcul de l'albédo de la neige L'albédo caractérise la capacité d'une surface à réfléchir les rayonnements électromagnétiques (en général, et ce sera le cas ici, on ne considère que les rayonnements allant du proche ultraviolet au proche infrarouge). L'albédo de la neige α_n , est calculé sur trois bandes spectrales différentes (Brun et al., 1992). La première bande ($[0.3, 0.8] \mu\text{m}$) représente le domaine de l'ultraviolet et du visible. Les deux autres (respectivement $[0.8, 1.5]$ et $[1.5, 2.8] \mu\text{m}$) représentent deux domaines du proche infrarouge. L'albédo de la neige, α_n , est fonction du diamètre optique de la neige et de l'âge de la première couche du manteau neigeux. La dépendance de l'âge est limitée à la première bande ($[0.3, 0.8] \mu\text{m}$) et vise à représenter de manière très simple la diminution de l'albédo de la neige par les impuretés provenant des dépôts. En effet, des traces d'impuretés absorbant la lumière peuvent réduire de manière significative l'albédo de la neige dans le domaine visible mais n'ont aucun effet dans le domaine du proche infrarouge Warren (1982). Dans chaque bande (numérotée de 1 à 3), l'albédo de la neige est calculé comme suit Brun et al. (1992) :

$$\alpha_n^1 = \max \left[0.6, \min \left(0.92, 0.96 - 1.58\sqrt{d_{\text{opt}}} \right) - \min \left(1, \max \left(\frac{1}{2}, \frac{P_a}{P_{\text{ref}}} \right) \right) \times 0.2 \frac{A_n}{A_{\text{ref}}} \right] \quad (3.9)$$

$$\alpha_n^2 = \max \left[0.3, 0.9 - 1.45\sqrt{d_{\text{opt}}} \right] \quad (3.10)$$

$$\alpha_n^3 = \min [0.0023, d_{\text{opt}}] \quad (3.11)$$

où A_n est l'âge de la première couche de neige exprimé en jours, A_{ref} un âge de référence (fixé à 60 jours de manière générale et à 900 jours pour les zones de neige/glace permanente supposées moins sujettes à la pollution) qui module la diminution de l'albédo de la neige due aux impuretés, P_a (hPa) est la pression atmosphérique près de la surface, et P_{ref} une pression de référence égale à 870 hPa et $d_{\text{opt}} = \min(d_{\text{max}}, g_1 + g_2 \times \rho_{sn}^4)$ est le diamètre optique de la neige dans la première couche du schéma, avec $d_{\text{max}} = 2.796 \times 10^{-3}$ m est le diamètre maximal et g_1, g_2 des coefficients (Anderson, 1976). L'albédo varie également en fonction de l'âge de la neige. L'albédo total (*broadband*), utilisé dans le modèle est alors une pondération de ces 3 albédos spectraux :

$$\alpha_n = 0.71 \times \alpha_n^1 + 0.21 \times \alpha_n^2 + 0.08 \times \alpha_n^3 \quad (3.12)$$

3.3.2 Flux turbulents sur les surfaces continentales

Le calcul des flux turbulents de chaleur, d'humidité et de quantité de mouvement à la surface est traité par SURFEX. Dans le cadre unicolonne (Section 3.4.1), on a le choix entre une configuration avec température de surface prescrite (formulation des flux identique au modèle 3D), et une configuration avec flux de chaleur sensible et flux de vapeur d'eau prescrits. Nous détaillons ici ce calcul pour les configurations utilisées au cours des travaux présentés ici.

Notons que pour les différentes configurations utilisées, les longueurs de rugosité seront toujours prescrites, que ce soit pour la quantité de mouvement ou pour la chaleur (égale à celle de la vapeur d'eau). Ces longueurs de rugosité seront notées par la suite z_0 et z_{0h} respectivement. Sur la neige elles valent respectivement $z_0 = 10^{-3}$ m et $z_{0h} = 10^{-4}$ m.

Configuration avec flux de chaleur sensible et de vapeur d'eau prescrits Les flux de chaleur sensible et d'humidité étant prescrits, on cherche à calculer le flux de quantité de mouvement F_u .

La vitesse de friction u_* est calculée d'après Paulson (1970) :

$$u_* = \frac{\kappa \bar{U}_1}{\ln\left(\frac{z_1}{z_0}\right) - \varphi\left(\frac{z_1}{L_{MO}}\right) + \varphi\left(\frac{z_0}{L_{MO}}\right)} \quad (3.13)$$

où \bar{U}_1 est le module du vent au premier niveau du modèle dont l'altitude est notée z_1 et φ la fonction de similitude donnée par Paulson (1970).

L_{MO} est la longueur de Monin-Obukhov qui représente la hauteur en dessous de laquelle l'effet du cisaillement de vent domine les effets de flottabilité sur les flux turbulents. Elle est calculée suivant :

$$L_{MO} = \frac{-u_*^3}{\kappa \frac{g}{\theta_v} (\overline{w'\theta'_v})_s} \quad (3.14)$$

L_{MO} dépendant elle-même de la vitesse de friction, le calcul de u_* et de L_{MO} est réalisé de manière itérative (10 itérations) en partant d'une stabilité neutre, *i.e.* $L_{MO} = \infty$ (en pratique $= 10^{20}$ dans le code).

Le flux de quantité de mouvement à la surface se calcule alors par

$$F_u = \bar{\rho} u_*^2 \quad (3.15)$$

Configuration avec température de surface prescrite C'est le cas général, qui correspond également à ce qui est utilisé dans le modèle 3D lorsque la température de surface est

interactive. Le calcul des différents flux (quantité de mouvement, chaleur sensible et humidité spécifique) font intervenir les coefficients d'échange C_d , C_h et C_q , et par leur intermédiaire, le nombre de Richardson dit bulk :

$$\mathbf{Ri}_b^0 = \frac{\frac{g}{\theta_v} \frac{d\theta_v}{dz}}{\left| \frac{d\vec{u}}{dz} \right|^2} \quad (3.16)$$

Dans SURFEX, \mathbf{Ri}_b^0 est calculé comme suit :

$$\mathbf{Ri}_b^0 = \frac{gz_1(\overline{\theta_{v1}} - \theta_{vs})}{\frac{1}{2}(\overline{\theta_{v1}} + \theta_{vs})\overline{U}_1^2} \quad (3.17)$$

où $\overline{\theta_{v1}}$ et θ_{vs} sont respectivement les températures potentielles virtuelles au premier niveau du modèle et à la surface. Un nombre de Richardson critique \mathbf{Ri}_c est introduit comme borne minimale pour \mathbf{Ri}_b^0 : $\mathbf{Ri}_b = \min(\mathbf{Ri}_b^0, \mathbf{Ri}_c)$. Il est pris égal à 0.1 sur les surfaces couvertes de glace ou de neige en permanence, et 0.2 sinon.

Les coefficients d'échange sont alors calculés suivant Mascart et al. (1995) et Noilhan and Mahfouf (1996) avec l'hypothèse $C_q = C_h$. Selon que l'on est en cas stable ($\mathbf{Ri}_b > 0$) ou instable ($\mathbf{Ri}_b < 0$), deux formulations différentes sont utilisées. Nous donnons ci-dessous le calcul de C_d et C_h pour le cas stable (le lecteur intéressé pourra se référer à Noilhan and Mahfouf (1996) pour le calcul des coefficients d'échange dans le cas instable).

$$C_d = \frac{\kappa^2}{\left(\ln\left(\frac{z}{z_0}\right)\right)^2} \frac{1}{1 + \frac{10\mathbf{Ri}_b}{\sqrt{1+5\mathbf{Ri}_b}}} \quad (3.18)$$

et

$$C_h = \frac{\kappa^2}{\left(\ln\left(\frac{z}{z_0}\right)\right)^2} \left(\frac{\ln(z/z_0)}{\ln(z/z_{0h})}\right) \left(\frac{1}{1 + 15\mathbf{Ri}_b\sqrt{1 + 5\mathbf{Ri}_b}}\right) \quad (3.19)$$

On peut noter que les différents coefficients entiers qui interviennent dans le calcul de C_d et C_h sont issus d'un travail d'ajustement visant à éviter une résolution itérative coûteuse dans le modèle. Même si il y a a priori une certaine incertitude sur ces coefficients, ils ne seront pas pris en compte dans la calibration du modèle dans les chapitres suivants.

Les flux à la surface sont alors calculés suivant :

$$\begin{aligned} F_u &= \bar{\rho} C_d \overline{U}_1^2 \\ F_s &= \bar{\rho} C_p C_h \overline{U}_1 (\theta_s - \overline{\theta}_1) \\ F_q &= \bar{\rho} C_q \overline{U}_1 (q_{v,s} - \overline{q}_{v,1}) \end{aligned} \quad (3.20)$$

où C_p est la capacité thermique de l'air à pression constante, U_1 est le module du vent au

premier niveau du modèle, θ_s et θ_1 sont respectivement la température potentielle à la surface et au premier niveau du modèle, de même pour les humidités spécifiques $q_{v,s}$ (supposée égale à la saturation pour la température et pression de surface) et $q_{v,1}$.

3.4 Différents types de simulations

3.4.1 SCM

Dans un modèle de climat, et en particulier dans ARPEGE-Climat, les paramétrisations physiques agissent uniquement sur la verticale. On peut donc évaluer leurs performances en réalisant des simulations unicolonnes ou SCM (pour Single Column Model). Une telle simulation se fait en extrayant simplement une colonne et en lui appliquant des forçages de grande échelle (issus par exemple du modèle GCM). Autrement dit, la partie dynamique du GCM est généralement désactivée, si bien que les advections de température et d'humidité sont prescrites. A noter qu'une variante est souvent utilisée dans laquelle c'est la vitesse verticale qui est prescrite (et non l'advection verticale), si bien que le modèle calcule lui son advection verticale. Un modèle SCM peut alors s'écrire sous la forme de l'équation suivante :

$$\frac{\partial \mathbf{x}_c}{\partial t} = \sum_p \mathcal{P}_p(\mathbf{x}_c, \boldsymbol{\lambda}_p) + \mathcal{F}_c(t, z, \mathbf{x}_c) \quad (3.21)$$

où \mathbf{x}_c est un vecteur contenant les variables d'état dans la colonne extraite et \mathcal{F}_c est un terme de forçage de grande échelle, dépendant du temps t , de l'altitude z et potentiellement de l'état du modèle (e.g., dans le cas où la vitesse verticale est prescrite), représentant la partie dynamique du modèle. Le terme $\mathcal{P}_p(\mathbf{x}_c, \boldsymbol{\lambda}_p)$ représente la contribution de la p^{ieme} paramétrisation. On doit également définir des conditions initiales $\mathbf{x}_c(t = 0)$.

À noter que la formulation utilisée ici présente les paramétrisations comme des fonctions des variables d'état du modèle uniquement. En réalité certaines paramétrisations (le schéma radiatif par exemple) prennent également en entrée des informations issues d'autres paramétrisations. Au cours des travaux présentés ici, différentes configurations de forçage en surface ont été utilisés (voir chapitre 6). Pour des configurations forcées en flux, le calcul des ces derniers est réalisé par SURFEX et a été rappelé dans la partie 3.3.2.

3.4.2 Simulations 3D

Un GCM peut être utilisé dans de nombreuses configurations. Au cours de ce travail, deux types de configurations ont été utilisées et sont présentées succinctement ici.

3.4.2.1 Configuration AMIP

Le terme AMIP fait référence à Atmospheric Model Intercomparison Project, un protocole d'intercomparaison pour des modèles GCMs. Ces simulations ne comparent que la partie atmosphérique d'un modèle de climat et n'inclut donc pas de couplage avec l'océan. Le protocole est assez simple et demande que le GCM soit forcé par des températures de surface de la mer réalistes. En pratique la température de surface de la mer et la concentration en glace de mer qui sont prescrites sont issues du jeu de données Program for Climate Model Diagnosis and Intercomparison AMIP Version 1.1.3 (Durack and Taylor, 2017). Par abus de langage, une simulation AMIP est simplement une simulation réalisée par un GCM forcé par la température de surface de la mer et la concentration en glace de mer. Le lecteur intéressé pourra se reporter à Roehrig et al. (2020) pour plus de détails des différents facteurs externes (e.g. gaz à effet de serre, aérosols) sur les simulations AMIP utilisées dans ce travail.

3.4.2.2 Simulations guidées

Les simulations AMIP, si elles permettent de dresser une climatologie du modèle, ne respectent pas la chronologie des événements météorologiques fournie par des observations, ce qui limite la comparaison directe entre modèle et observations. Pour pallier cette difficulté, on utilise des simulations dont on rappelle la dynamique vers une référence, en pratique vers la dynamique d'une réanalyse. On utilise pour cela la technique de nudging (ou relaxation newtonienne) qui consiste à rajouter un terme de rappel dans le modèle de façon à rapprocher celui-ci de la référence. Le modèle « nudgé » peut alors s'écrire de la façon suivante :

$$\frac{\partial \mathbf{y}}{\partial t} = \mathcal{M}(\mathbf{x}) + K(\mathbf{y}_{\text{ref}} - \mathbf{y}) \quad (3.22)$$

où \mathbf{y} est le vecteur des variables qui vont être rappelé vers la référence, et \mathcal{M} la partie du modèle qui gouverne l'évolution de \mathbf{y} , et qui prend en entrée le vecteur des variables d'état complet \mathbf{x} . \mathbf{y}_{ref} est la référence vers laquelle le modèle est rappelé. La « force » du rappel est proportionnel à l'écart du modèle à la référence, K étant le coefficient de pondération. Celui-ci est donc inversement proportionnel à un temps qui peut aller de quelques heures à quelques dizaines d'heures (voir la figure 3.3 qui donne le profil vertical de ce coefficient sur les 1000 premiers mètres). Le coefficient de nudging évolue avec l'altitude et tend vers 0 quand on se rapproche de la surface. On veut ainsi permettre à la couche limite la possibilité de s'ajuster à sa condition de surface, et d'éviter de potentielles instabilités numériques (le modèle tirant dans un sens, et la relaxation dans un autre). On veut aussi que les paramétrisations puissent s'exprimer.

Dans les simulations guidées en vent réalisées pour ce travail, seules les composantes du vent ont été rappelées vers la réanalyse ERA-Interim, produite par le CEPMMT (Dee et al., 2011).

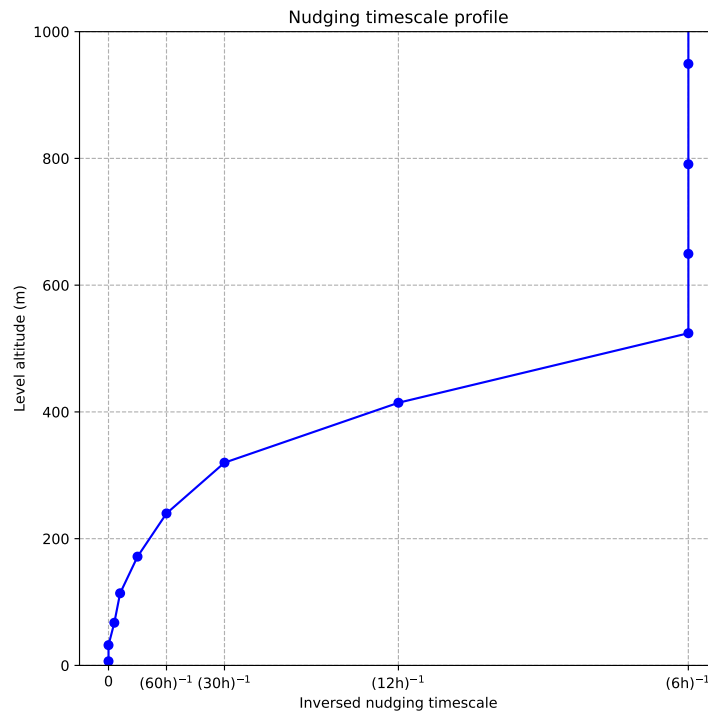


FIGURE 3.3 – Profil vertical du coefficient de nudging en h^{-1}

3.5 Données de validation

3.5.1 Les jeux de données de référence

Plusieurs jeux de données ont été utilisés pour évaluer le modèle ARPEGE-Climat sur l’Antarctique. Pour chacun de ces jeux de données, une interpolation conservative sur la grille du modèle a été effectuée.

Température à 2 m Les données de référence utilisées sont les moyennes mensuelles du jeu de données Berkeley Earth Surface Temperature (BEST), disponibles sur la période 1979-2008 à la résolution $1^\circ \times 1^\circ$ (Rohde et al., 2013a,b). Ce jeu de données est une reconstruction d’un champ 2D à partir de mesures ponctuelles par une méthode de krigeage. Par construction, l’incertitude est importante dans les régions où les stations sont rares, ce qui est le cas du plateau antarctique.

Température de surface Les données utilisées sont les températures de surface par ciel clair mesurées par MOderate-Resolution Imaging Spectroradiometer (MODIS – Wan et al., 2015).

Rayonnement Les données utilisées pour tous les effets radiatifs de nuages, les flux solaires et infrarouges, à la surface et au sommet de l’atmosphère sont tirées du jeu de données Clouds

and Earth's Radiant Energy System (CERES), disponibles sur la période 2001-2018 à la résolution $1^\circ \times 1^\circ$ (Loeb et al., 2018, Kato et al., 2018).

Nuages Pour les fractions nuageuses, le jeu de données GOCCP (pour GCM Oriented CALIPSO Cloud Product) fourni des moyennes mensuelles sur la période 2007-2017 à partir de données fournies par le lidar Cloud Aerosol Lidar with Orthogonal Polarization (CALIOP), embarqué sur le satellite CALIPSO (Winker et al., 2007, Chepfer et al., 2010)). Ces données sont comparées aux sorties du modèle ARPEGE sur lesquelles le simulateur Lidar du package COSP (Bodas-Salcedo et al., 2011) a été appliqué.

Vent et pression au niveau de la mer Pour évaluer les champs de vent (composantes zonale et méridienne) et de pression de surface, les données de réanalyse ECMWF ReAnalysis Version 5 (ERA5 – Hersbach et al., 2019).

Mesures à Dôme C À Dôme C une tour instrumentée de 45 m est installée près de la station de recherche Concordia. Cette tour permet de mesurer la température et le vent à différentes hauteurs. On se reportera à Genthon et al. (2010) pour plus de détails sur les instruments et leur hauteur. Une station du réseau Baseline Station Radiation Network (BSRN) est également installée à proximité et permet de mesurer les flux solaires et infrarouges à la surface. Une température de surface est estimée à partir du flux infrarouge ascendant en supposant une émissivité de la neige de 0.99. La station BSRN mesure également la température à 2 m

3.5.2 Simulations LES

Afin d'évaluer les performances des simulations SCM, on utilise classiquement des simulations LES (pour Large-Eddy Simulation) comme référence. Ces simulations, comme leur nom l'indique, permettent de résoudre les structures turbulentes les plus grandes. On considère en général qu'une simulation est une LES si au moins 70 à 90 % des moments d'ordre 2 sont résolus. Ces simulations sont réalisées sur un domaine à aire limitée de l'ordre du km à quelques dizaines de km. Leur résolution doit être suffisamment fine pour résoudre les structures turbulentes les plus grandes, soit des résolutions de l'ordre de quelques dizaines de mètres pour des couches limites convectives et des résolutions de l'ordre du mètre pour des couches limites stables.

3.5.2.1 Les exercices GABLS

Sous l'égide du projet GEWEX (pour Global Energy and Water Cycle Experiment), quatre exercices d'intercomparaison appelés GABLS (pour GEWEX Atmospheric Boundary Layer Study) ont étudié toute une gamme de CLS afin d'améliorer leur représentation dans des modèles numériques de prévision du temps ou de climat (Holtslag, 2006, Holtslag et al., 2012).

GABLS1 C'est un cas idéalisé de CLS nocturne faiblement stratifié, construit à partir d'observations réalisées en Arctique. L'atmosphère est forcée par la température de surface à laquelle un refroidissement constant de 0.25 K h^{-1} est appliqué (Beare et al., 2006). Cet exercice a permis de comparer différents modèles LES à différentes résolutions verticales. Beare et al. (2006) montrent que les résultats convergent pour des résolutions spatiales inférieures à 2 m. Ces simulations LES ont alors été utilisées comme référence pour évaluer des versions unicolonne de modèle de circulation générale (Cuxart et al., 2006). Un des résultats importants de cette étude est que les paramétrisations de la turbulence utilisant une équation prognostique pour la TKE parviennent mieux à reproduire les résultats des simulations LES.

GABLS2 Ce cas présente un cycle diurne complet (Svensson et al., 2011), basé sur la campagne de mesure Cooperative Atmospheric surface Exchange Study (CASES-99). Comme pour GABLS1, les modèles sont forcés par la température de surface. Des modèles unicolonne sont intercomparés. Cette étude montre une forte sous-estimation du cycle diurne. Les références utilisées pour évaluer les modèles combinent observations et simulations LES.

GABLS3 C'est un cas plus réaliste basé sur des observations réalisées sur un mât instrumenté de 213 m de hauteur installé à Cabauw (Pays-Bas). Il s'agit d'un cas de cycle diurne par ciel clair avec le développement d'un jet nocturne de basses couches. À la différence de GABLS1 et GABLS2, l'atmosphère est couplée avec la surface et le rayonnement n'est pas négligé (Bosveld et al., 2014b,a). Les modèles évalués présentent une couche limite nocturne trop épaisse. Le jet de basses couches n'est pas assez marqué. Du fait de la hauteur trop importante de la couche limite nocturne, le maximum d'intensité du jet est trop haut. À noter que Pour ce cas, des diagnostics caractérisant les interactions entre surface/rayonnement/turbulence ont été mis au point.

GABLS4 À la différence des trois premiers exercices GABLS, ce dernier s'attaque à une couche limite très stable. Le cas GABLS4 présente un cycle diurne complet et est basé sur une campagne d'observations réalisée à Dôme C sur le plateau antarctique pendant le mois de décembre 2009. Ce dernier cas pousse les différents modèles (LES et SCM) dans leur retranchement en leur demandant de représenter une couche limite convective de 300 m suivie d'une couche limite nocturne très stratifiée de moins de 30 m d'épaisseur (gradient de température supérieur à 0.7 K m^{-1}). GABLS4 est un exercice d'intercomparaison de différents types de modèles (modèles de surface, SCM, LES), sur un cycle diurne de couche limite sèche et stable. Les LES ont été initialement lancées avec une résolution spatiale (horizontale et verticale) de 2 m. Elles ont montré un bon accord entre elles sur la première partie du cycle diurne pendant laquelle se développe une couche limite convective très peu profonde. Sur la deuxième partie du cycle diurne, une couche limite très stable se met en place sur une épaisseur inférieure à 30 m et les résultats des différentes LES présentaient des différences importantes. Une nouvelle série de simulations a donc été lancée avec une résolution plus fine (1 m) et une initialisation à 18H00. Elles ne simulent ainsi que la phase stable et ont ainsi pu réduire l'extension verticale de leur domaine (la phase convective et sa couche limite de 300 m nécessitant une extension verticale nettement plus importante). Le fait de ne pas avoir à représenter la phase convective permettait de réduire la taille du domaine des simulations,

ce qui permettait de réduire le coût de calcul. Les résultats ont alors montré une très bonne convergence des résultats. Ce sont ces simulations qui ont été utilisées pour évaluer et calibrer une version SCM de ARPEGE-Climat (Chapitre 6). Pour plus de détail sur les LES réalisées dans le cadre de GABLS4, on pourra se référer à Couvreur et al. (2020a).

Apport des exercices GABLS Les exercices GABLS montrent que les modèles opérationnels ont de moins bonnes performances sur les couches limites stables que les modèles de recherche (Cuxart et al., 2006, Beare et al., 2006, Svensson and Holtslag, 2009, Svensson et al., 2011) avec une couche limite trop profonde, un jet trop peu marqué ainsi qu’une rotation du vent sur la verticale trop faible. Souvent le mélange turbulent est artificiellement augmenté dans les modèles de PNT pour d’une part éviter les phénomènes de runaway cooling, d’autre part pour éviter un creusement trop fort des dépressions (Sandu et al., 2013). Dans la plupart des modèles de climat participant à CMIP5, un mélange turbulent trop fort en condition nocturne était également observé (Svensson and Lindvall, 2015). Ces exercices montrent la nécessité de mieux comprendre les processus de mélange proche de la surface, mais également du rôle de la paramétrisation sous-maille de ces échanges dans le cas des couches très stables comme c’est le cas en Antarctique, (Sandu et al., 2013, Holtslag et al., 2013, Mahrt, 2014).

3.6 Conclusion

L’objectif de ce chapitre était de présenter le modèle ARPEGE-Climat 6.3. Le modèle sera évalué dans la suite de ce travail sur l’Antarctique à travers des simulations AMIP et des simulations guidées en vent. La paramétrisation de la turbulence, élément clé dans la représentation des CLS, a été présentée en détail. Dans la suite de ce travail, une approche novatrice de la calibration des modèles sera appliquée à cette paramétrisation à l’aide d’une version SCM de ARPEGE-climat afin de tester sa capacité à représenter la couche limite stable extrême du cas GABLS4 et de proposer une calibration plus adaptée. Le cas GABLS1 sera alors utilisé pour évaluer cette calibration.

Les différents jeux de références (observations, réanalyses, LES) qui seront utilisés pour évaluer les différentes simulations 1D et D ont été également présentés brièvement.

Évaluation de ARPEGE-Climat 6.3 sur l'Antarctique

Sommaire

4.1	Modéliser le climat antarctique	50
4.1.1	Observations et réanalyses sur l'Antarctique	50
4.1.2	Les difficultés des modèles de climat pour représenter la région antarctique	51
4.2	Évaluation de ARPEGE-Climat 6.3 en configuration AMIP	52
4.2.1	Été austral (DJFM)	52
4.2.2	Hiver austral (JJAS)	57
4.3	Simulation guidée en vent sur l'Antarctique	59
4.3.1	Évaluation de la simulation guidée en vent	59
4.3.2	Sensibilité de la simulation guidée en vent à l'albédo	62
4.4	Évaluation sur Dôme C	65
4.5	Conclusion	69

Ce chapitre présente un rapide état des lieux du comportement de ARPEGE-Climat 6.3 sur le continent Antarctique. Après une première partie qui rappelle les problèmes récurrents des modèles de climat dans leur représentation du climat antarctique, une simulation AMIP du modèle ARPEGE-Climat 6.3 d'une durée de 17 ans est utilisée afin de documenter la climatologie du modèle. Cette dernière se caractérise principalement par un biais froid sur une grande partie du continent (à l'exception des zones côtières et des plates-formes de glaces). On analyse plus particulièrement la saison de l'été austral (les mois de décembre, janvier, février et mars) afin de relier le biais froid à des défauts de représentation du bilan radiatif en surface. Une simulation guidée en vent plus courte, sur la période 2014-2016, est alors comparée à la simulation AMIP afin de vérifier que la structure spatiale des biais et leurs ordres de grandeur sont équivalents. La sensibilité à l'albédo de surface du modèle nudgé est testée. Enfin, un focus est fait sur le comportement du modèle nudgé sur Dôme C dans la dernière partie.

4.1 Modéliser le climat antarctique

Dans cette section, on présente les principales difficultés rencontrées dans la modélisation du climat antarctique. On rappelle dans une première partie la rareté des séries d'observations de qualité dans cette région et l'impact sur la fiabilité des réanalyses, avant de présenter dans une deuxième partie les biais récurrents des modèles de climat.

4.1.1 Observations et réanalyses sur l'Antarctique

Évaluer le comportement d'un modèle sur une région donnée nécessite d'avoir à disposition des observations. Ce point pose le problème de la mesure des paramètres météorologiques en Antarctique. Du fait du climat extrême et de l'isolement du continent il n'existe que peu de stations météorologiques (automatiques ou non) et la plupart d'entre elles sont situées près des côtes. Il n'y a donc que très peu de mesures régulières dans les régions centrales du plateau continental. On ne dispose ainsi que de peu de séries continues d'observations *in situ* (et quasiment aucune avant 1957, l'année géophysique internationale, qui a marqué un tournant dans la connaissance du climat et de la météorologie de l'Antarctique). La mesure elle-même est difficile. Du fait des conditions extrêmes de température et de vent, les instruments sont souvent soumis à des dépôts de glace et à des phénomènes de givrage (Genthon et al., 2011, Genthon et al., 2013). Le problème de la fiabilité des mesures se pose donc, en particulier pour les stations automatiques. La mesure des précipitations en particulier est délicate (Genthon et al., 2009). Celles-ci sont faibles sur le plateau continental et nécessitent d'avoir des instruments de mesure particulièrement sensibles. Par ailleurs, le vent fort peut rendre difficile de discriminer transport de neige par le vent et précipitations.

Depuis 1979, l'essor des satellites d'observation météorologiques permet de palier une partie du manque d'observations. Tout n'est cependant pas parfait. Les satellites permettent par exemple une estimation de la température de surface sur l'Antarctique par satellite (e.g. MODIS, Fréville et al., 2014), mais par ciel clair uniquement. De façon générale, les mesures spatiales sont indirectes par nature et nécessitent un traitement complexe, et qui introduit de nombreuses sources d'incertitude dans les estimations fournies. De plus, du fait de l'inclinaison de l'orbite de certains satellites, on ne dispose pas de mesures de température au delà de 82° S.

Les réanalyses sont un outil essentiel pour étudier le climat et pour évaluer la qualité de modèles numériques. Mais la qualité d'une réanalyse dépend fortement de la qualité des observations utilisés pour la construire. ERAI par exemple assimile des données issues de stations de surface, de radiosondages, de vols commerciaux et des données satellites. Ainsi du fait de la rareté des observations au sol, il n'existe pas de réanalyses fiables sur l'Antarctique avant l'ère des mesures par satellites. Différentes réanalyses ont été évaluées par Bracegirdle and Marshall (2012). Grâce aux données satellites, celles-ci sont plutôt fiables en moyenne troposphère pour les champs de géopotentiel et de température. Proche de la surface, zone moins accessible à la mesure satellite, des biais importants apparaissent. Ainsi les champs de pression à la surface proposés par les différentes réanalyses évaluées sont discutables et un biais chaud de température de surface est systématiquement constaté. Ce biais est souvent relié à un mélange turbulent trop fort en conditions stables (Bracegirdle and Marshall, 2012,

Fréville et al., 2014) mais également à des propriétés du manteau neigeux erronées dans les schémas de surface (Dutra et al., 2015).

4.1.2 Les difficultés des modèles de climat pour représenter la région antarctique

Biais chaud à la surface de l’Antarctique et paramétrisation de la turbulence Le plateau continental antarctique présente souvent des couches limites stables voir très stables, en particulier durant l’hiver austral. La représentation de ce type de couches limites repose principalement sur la paramétrisation de la turbulence dans les modèles numériques. Or cette dernière est en général calibrée de façon à améliorer l’échelle synoptique, ce qui se traduit souvent par un mélange turbulent artificiellement augmenté de façon à tenir compte de processus mal représentés dans le modèle (Steenefeld et al., 2008, Sandu et al., 2013). Cette « turbulence augmentée » est connue pour être détrimentale pour les couches limites stables (e.g., Holtslag et al., 2013) comme pour les réanalyses, et est à l’origine de biais chaud dans plusieurs modèles de climat (Beaumont, 2018). Dans les modèles présentant ce biais chaud dû à une turbulence augmentée, l’intensité de l’inversion qui se développe sur le plateau continental est sous-estimée, ce qui entraîne un biais dans la représentation des vents catabatiques (King et al., 2001, Cassano et al., 2001). La circulation générale est également impactée : le régime des vents catabatiques induit une divergence à la surface du plateau antarctique et donc par conservation de la masse une convergence en altitude (Parish and Bromwich, 2007).

Représentation des nuages Les caractéristiques de la couche nuageuse au dessus de l’Antarctique (fraction nuageuse, épaisseur optique) sont souvent mal représentées dans les modèles de climat (Bromwich et al., 2012, Kay et al., 2016) avec des conséquences sur le bilan radiatif à la surface. L’effet radiatif des nuages est souvent sous-estimé, ce qui conduit à une surestimation du flux solaire incident et à une sous-estimation du rayonnement infrarouge descendant. Ces biais dans la représentation des nuages sont en général reliés à des lacunes dans les schémas microphysiques avec en particulier l’absence de sursaturation par rapport à la glace dans les modèles (van Wessem et al., 2014, Genthon et al., 2017).

Précipitations La répartition des précipitations est très hétérogène avec des cumuls annuels importants sur les côtes alors que l’intérieur du plateau continental est beaucoup plus sec (Palermo et al., 2014). Palermo et al. (2017) comparent les précipitations dans les modèles de climat ayant participé à l’exercice CMIP5 avec des produits issus de mesures spatiales (CloudSat, Wood et al. (2013, 2014)) et la réanalyse ERAI (qui présentent un bon accord). La plupart des modèles sur-estiment (parfois largement) les précipitations sur l’Antarctique. Estimer correctement les précipitations est nécessaire pour calculer le bilan de masse de surface et faire des projections sur la contribution de l’Antarctique au niveau des océans dans le futur. Palermo et al. (2017) met en évidence une divergence importante du taux de précipitation dans les modèles CMIP5. Les différents modèles tendent à surestimer les précipitations à l’intérieur du continent. Les précipitations sur les régions côtières sont plus réalistes. Une étude plus récente (Roussel et al., 2020) donne des résultats similaires en évaluant les modèles CMIP6 et la réanalyse ERA5 contre les produits CloudSat.

4.2 Évaluation de ARPEGE-Climat 6.3 en configuration AMIP

Une simulation AMIP (voir section 3.4.2.1) couvrant la période 2000-2016 a été réalisée avec ARPEGE-Climat 6.3 afin de documenter les biais climatologiques du modèle sur l'Antarctique. Une telle durée est sans doute trop courte pour capturer toute la variabilité interne du système climatique aux hautes latitudes. Cependant les erreurs que l'on met en évidence (e.g., -4.6 K en température maximale sur l'été austral) sont suffisamment importantes pour être nettement supérieures à l'amplitude de la variabilité interne.

4.2.1 Été austral (DJFM)

Températures de l'air en surface La figure 4.1 présente la température de l'air à 2 m et compare la simulation AMIP réalisée avec ARPEGE-Climat 6.3 avec les moyennes mensuelles du jeu de données BEST. L'incertitude sur ces données reste importante sur la région antarctique (voir section 3.5.1). Il convient donc de rester prudent dans l'utilisation de ce jeu de données pour évaluer le modèle. De plus la température à 2 m n'est pas une variable pronostique du modèle ; elle est calculée par une interpolation entre la température de surface et la température au premier niveau du modèle (≈ 8.5 m), prenant en compte la stabilité à la surface selon Geleyn (1988). ARPEGE-Climat 6.3 présente un biais froid (-2.6 K) pendant l'été austral en moyenne sur tout le continent Antarctique à l'exception notable de la plate-forme de Ross pour lequel on constate un biais chaud (> 1 K). Sur l'Antarctique Est, région sur laquelle on va se focaliser, le biais froid est général avec des valeurs comprises entre -1 et -5 K. Une analyse rapide montre que ce biais est principalement dû à la température maximale (-4.64 K sur l'ensemble du continent) et non à la température minimale (-0.94 K). Autrement dit l'atmosphère proche de la surface ne se réchauffe pas suffisamment pendant la journée. Le biais froid sur les températures minimales pourrait être lié au biais froid sur les températures maximales qui perdure durant la nuit.

Notons qu'on peut mettre en évidence une corrélation sur la plate-forme de Ross entre le biais chaud de température minimale et un biais positif dans le contenu en vapeur d'eau intégré sur la verticale (non montré). Sachant que l'essentiel du contenu en vapeur d'eau est dans les couches les plus basses de l'atmosphère, un tel lien entre les deux grandeurs est attendu : une atmosphère plus chaude peut contenir plus de vapeur d'eau selon la loi de Clausius-Clapeyron (7% de contenu en vapeur d'eau supplémentaire par Kelvin) qui en retour va émettre vers la surface plus de rayonnement infrarouge et contribuer au biais chaud observé. Cette corrélation illustre les interactions complexes entre les différents processus et la difficulté qu'il y a à démêler l'écheveau des erreurs du modèle. Le comportement particulier du modèle sur la plate-forme de Ross demanderait une analyse plus poussée non réalisée ici. Nous nous concentrerons dans la suite sur le comportement du modèle sur l'Antarctique Est (entre les longitudes 30° O et 150° E).

Température moyenne de surface en ciel clair La température moyenne de surface dans le modèle ARPEGE-Climat 6.3 est comparée avec la température de surface mesurée par MODIS (voir section 3.5.1) en conditions de ciel clair. Le biais froid observé sur la température

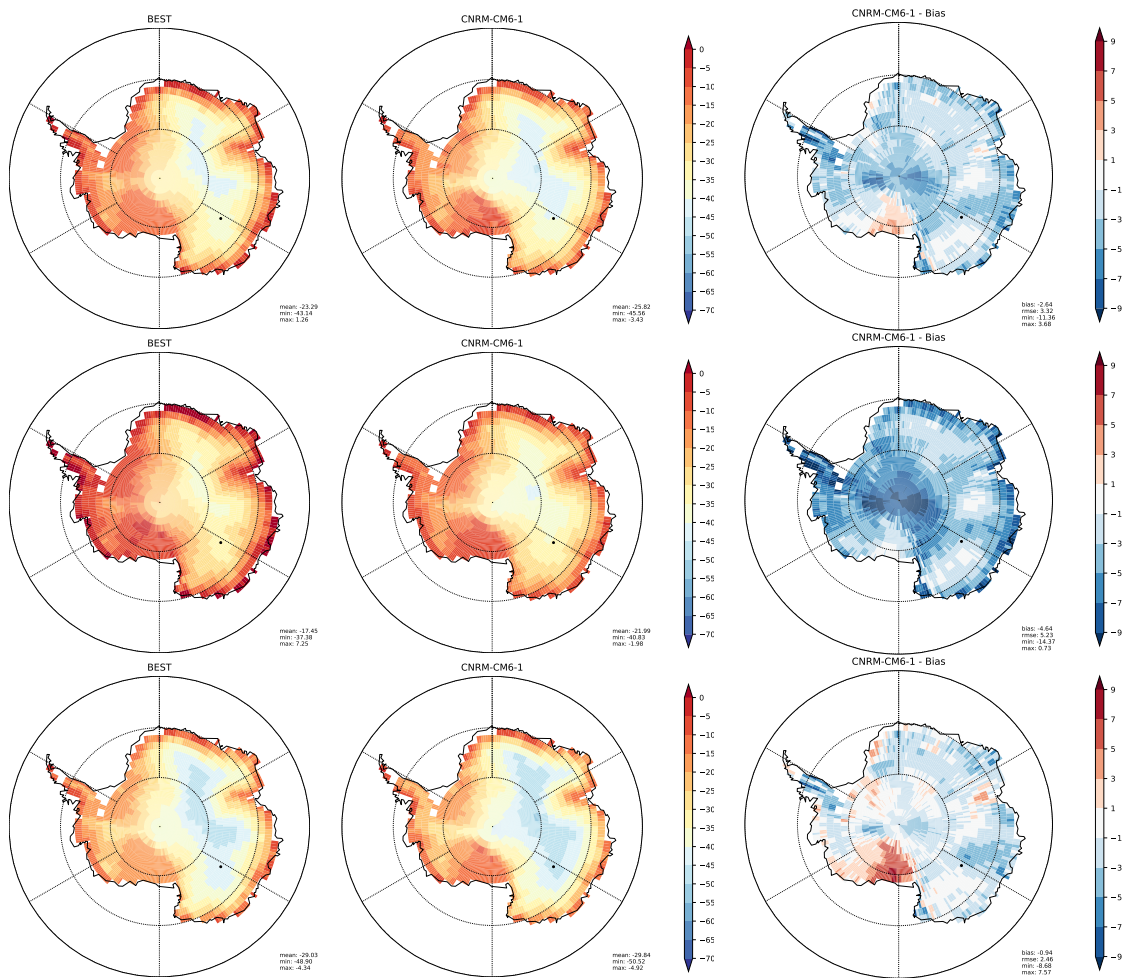


FIGURE 4.1 – Températures à 2 m moyennes en été austral (DJFM) dans la référence BEST (à gauche), ARPEGE-Climat 6.3 (au centre) avec le biais de ARPEGE-Climat 6.3 par rapport à BEST à droite. La première ligne présente la température moyenne, la deuxième ligne, la température maximale, et la troisième ligne, la température minimale. Le point noir sur les différentes cartes présentées dans ce chapitre représente la position de Dôme C)

à 2 m se retrouve dans la température de la surface sur une large partie du plateau continental (figure 4.2). Il est cependant moins important (-0.97 K contre -2.64 K pour la température à 2 m). Des différences dans la spatialisation de ce biais apparaissent également. En particulier sur l'Antarctique Est, ce biais froid est plus accentué sur les régions avec les altitudes les plus élevées. Près des côtes, on constate au contraire un biais chaud de la température de surface que l'on ne retrouve pas sur la température à 2 m. Ce biais chaud dans les zones côtières ne se retrouve pas sur la température à 2 m où l'on constate au contraire un biais froid accentué. Il faut noter que les références pour la température de surface et pour la température à 2 m sont différentes (respectivement MODIS et BEST), et que la température de surface n'est estimée que par ciel clair, or l'occurrence de nuage est plus importante sur les côtes. Ces différences peuvent expliquer une partie de la différence de comportement sur les régions côtières avec notamment un masque terre/mer différent.

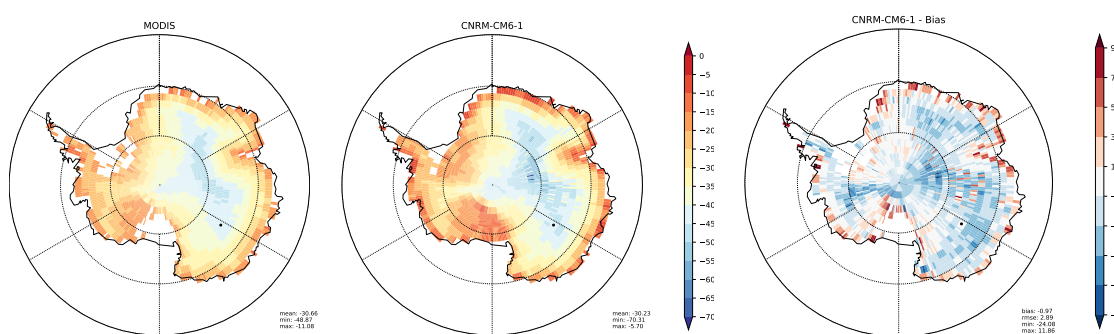


FIGURE 4.2 – Températures de surface (en ciel clair) moyennes en été austral (DJFM) sur la période 2000-2016 avec de gauche à droite, la référence (MODIS), ARPEGE-Climat 6.3 et le biais de ARPEGE-Climat 6.3 par rapport à MODIS.

Représentation des nuages Les nuages sont potentiellement une source d'erreur majeure pour les températures dans la couche limite, car influant le bilan radiatif à la surface. Pendant la journée, ils bloquent une partie du rayonnement solaire et une surestimation de la couverture nuageuse et de ses effets radiatifs peut conduire à un biais froid de la température maximale à 2 m. Pendant la nuit au contraire, ils émettent un rayonnement infrarouge qui limite le refroidissement nocturne en surface et une surestimation des nuages et de leur effet radiatif conduit à un biais chaud de la température minimale.

La figure 4.3 compare la fractions nuageuse totale de ARPEGE-Climat 6.3 avec celle estimée par CALIPSO (voir section 3.5.1). La fraction nuageuse est largement surestimée sur tout l'intérieur de plateau de l'Antarctique Est par ARPEGE-Climat 6.3, alors qu'elle est sous-estimée sur les côtes. Cette structure est dominée par la représentation des nuages de l'étage moyen. Notons ici que du fait de l'altitude élevée du plateau continental ($< 2000\text{ m}$), les nuages de l'étage moyen représente en fait les nuages bas dans le modèle. Une origine possible de cette surestimation de la fraction nuageuse pourrait être la non prise en compte de la sursaturation par rapport à la glace dans la paramétrisation microphysique de ARPEGE-Climat 6.3, alors que ce processus est particulièrement important sur le plateau antarctique (Genthon et al., 2013, 2017). Cette surestimation de la couche nuageuse aurait pu être soupçonnée d'être responsable du biais froid (voir la remarque du paragraphe précédent sur l'effet d'une telle surestimation sur la température à 2 m). Cependant une comparaison de l'effet

radiatif des nuages dans le spectre solaire dans le modèle avec les données CERES (figure 4.4a) montre que celui-ci est en fait sous-estimé dans ARPEGE-Climat 6.3 (avec un biais moyen de 5.65 W m^{-2}). Dans le domaine de l'infrarouge, les résultats sont plus contrastés : l'effet radiatif des nuages dans ARPEGE-Climat 6.3 présente un biais négatif sur la majeure partie de l'Antarctique, *i.e.* les nuages ne limitent pas assez le refroidissement nocturne de la surface. Notons tout de même que sur la région de Dôme C ce biais s'inverse et devient légèrement positif. L'effet radiatif total (solaire + infrarouge) des nuages est positif sur tout le continent, dans le modèle comme dans CERES. Les nuages en Antarctique ont donc tendance à réchauffer la surface. Cependant cet effet est surestimé dans ARPEGE-Climat 6.3 sur la partie centrale de l'Antarctique Est avec un biais positif de 10 W m^{-2} . Ainsi, si la fraction nuageuse dans le modèle est surestimée, les nuages du modèle sont trop ténus et ne peuvent expliquer le biais froid dans la température maximale à 2 m. Le biais froid observé sur les températures maximales n'est donc pas lié à la représentation des nuages dans le modèle.

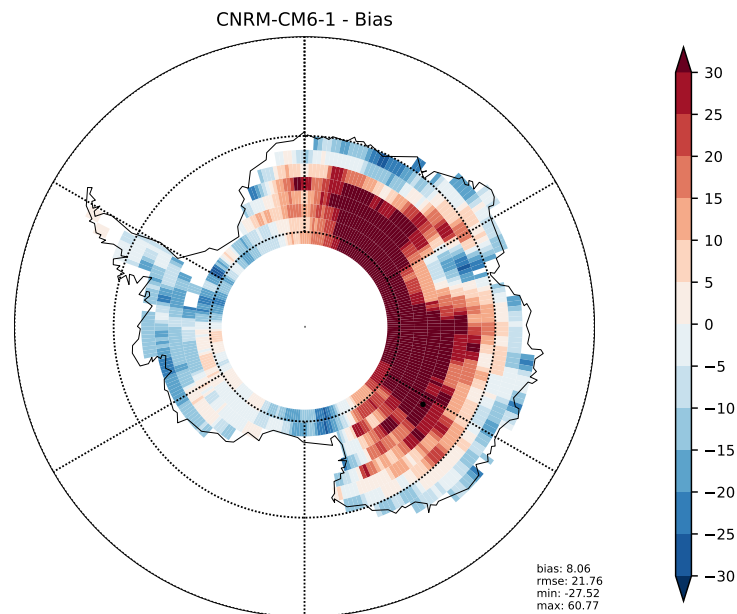


FIGURE 4.3 – Biais de la fraction nuageuse totale en été austral (DJFM) dans ARPEGE-Climat 6.3 par rapport à CALIPSO. À noter que du fait de l'inclinaison de l'orbite du satellite, aucune mesure de la fraction nuageuse n'est disponible au sud de 82° S

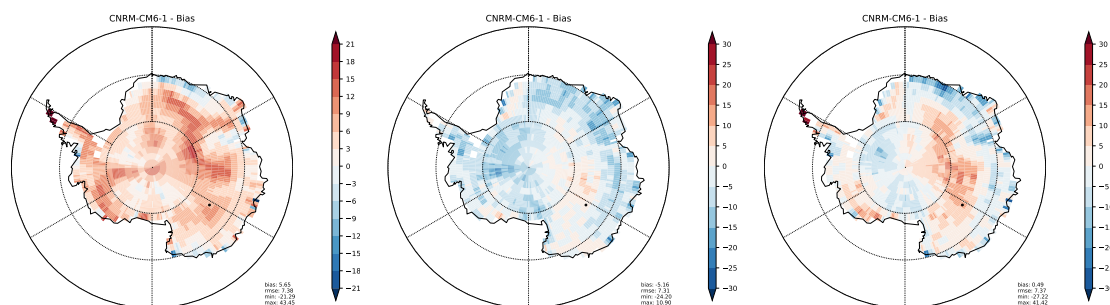


FIGURE 4.4 – Biais de l'effet radiatif des nuages à la surface en été austral (DJFM) dans ARPEGE-Climat 6.3 par rapport à CERES avec de gauche à droite l'effet radiatif des nuages dans le spectre solaire, l'effet radiatif des nuages dans le spectre infrarouge, et l'effet radiatif net (solaire+infrarouge) des nuages.

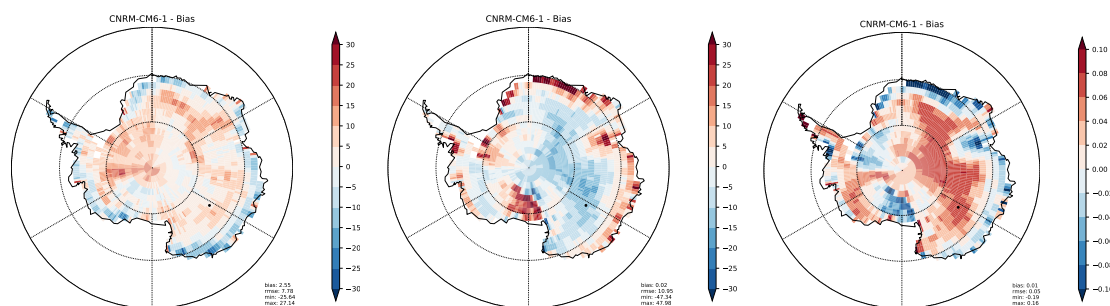


FIGURE 4.5 – De gauche à droite : biais en été austral (DJFM) du flux solaire descendant à la surface, du flux solaire net à la surface et de l'albédo de surface de ARPEGE-Climat 6.3 par rapport à CERES.

Bilan radiatif et albédo La figure 4.5a) présente le flux solaire descendant à la surface. Le flux est surestimé sur le plateau antarctique est. Ce résultat est une conséquence de la sous-estimation de l'effet radiatif des nuages dans le solaire mis en évidence dans le paragraphe précédent. Pourtant le bilan net du rayonnement solaire à la surface, présenté sur la figure 4.5b) met en évidence un biais négatif sur cette même région. Ce résultat s'explique par un albédo de la surface assez largement surestimé dans ARPEGE-Climat 6.3 (avec un biais > 0.06 sur les parties centrales du plateau par rapport aux estimations utilisées dans le produit CERES (voir section 3.5.1). Notons ici que l'albédo de surface produit par CERES est estimé à partir de mesures de radiances satellites et utilise un schéma de transfert radiatif. Il conviendrait de faire une évaluation de l'incertitude de ce produit sur l'Antarctique. Le biais sur l'albédo s'inverse quand on se rapproche des côtes ainsi que sur la plate-forme de Ross et est cohérent avec le biais chaud en température de surface mentionné plus haut.

Représentation du vent à 10 m Sur l'Antarctique Est, la structure du vent est dominée par le régime des vents catabatiques avec une zone de vents faibles sur la partie haute et plate du centre du plateau et des vents qui s'accroissent lorsque les pentes s'accroissent à l'approche des côtes. La première ligne de la figure 4.6 met bien en évidence cette structure dans la réanalyse ERA5, mais également dans ARPEGE-Climat. La deuxième ligne montre cependant de nombreuses régions près des côtes pour lesquelles le module du vent dans ARPEGE-Climat

6.3 est sous-estimé. L'analyse des biais pour chacune des composantes du vent horizontal montre que c'est la partie méridienne (*i.e.* l'écoulement du pôle vers les côtes) du vent qui est alors sous-estimée (non montré).

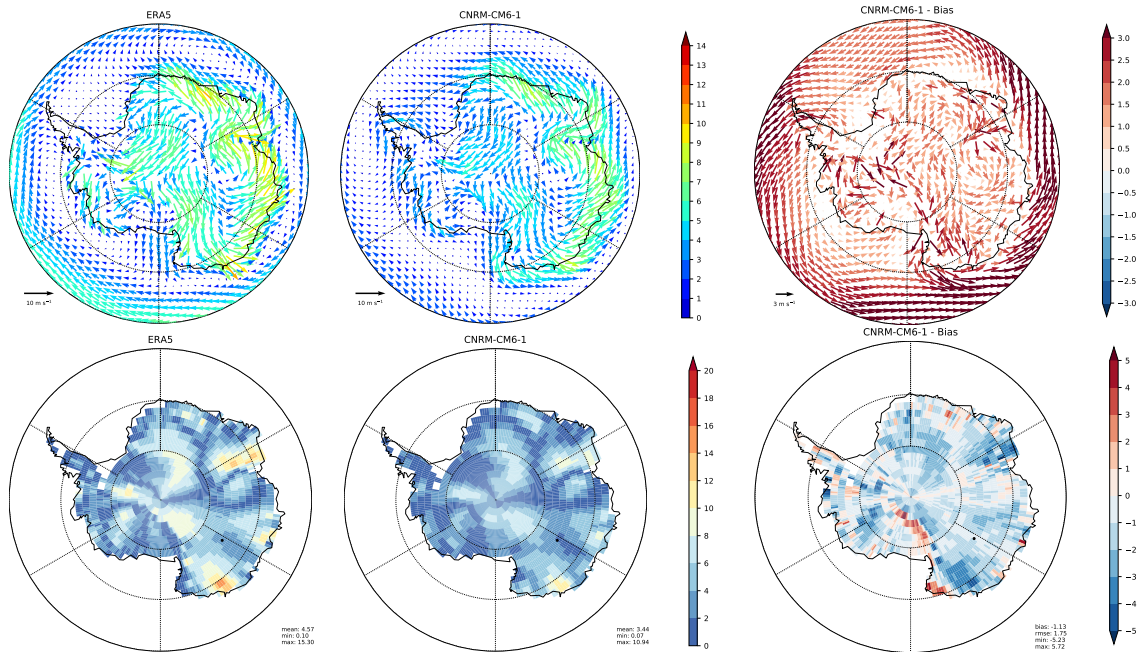


FIGURE 4.6 – Comparaison du vent à 10 m en été austral (DJFM) dans ARPEGE-Climat 6.3 (au centre) avec le vent à 10 m dans ERA5 (à gauche). À droite le biais de ARPEGE-Climat 6.3 par rapport à ERA5. Sur la première ligne est représenté le vent à 10 m, sur la deuxième ligne le module du vent à 10 m.

4.2.2 Hiver austral (JJAS)

Température à 2 m En hiver austral (JJAS) ARPEGE-Climat 6.3 présente également un biais froid quasi généralisé sur l'ensemble de l'Antarctique (Figure 4.7a). Ce biais est plus prononcé près des côtes que sur les parties centrales du plateau. Les figures 4.7b,c,d présentent respectivement les biais sur la fraction nuageuse totale, l'effet radiatif des nuages et le bilan radiatif dans l'infrarouge. Comme en été austral, ARPEGE-Climat 6.3 produit trop de nuages, spécialement sur l'Antarctique Est. Leur effet radiatif en surface dans l'infrarouge est cependant nettement sous-estimé en comparaison aux estimations fournies par CERES. Le biais négatif du bilan radiatif présente une structure spatiale similaire à celle du biais de l'effet radiatif des nuages et peut donc y être relié en grande partie. À la différence de l'été austral, le biais froid à la surface peut être relié (au moins en partie) avec la représentation des nuages dans le modèle.

Représentation du vent à 10 m Comme pendant la saison estivale, les vents en Antarctique sont dominés par les vents catabatiques et ARPEGE-Climat 6.3 représente relativement

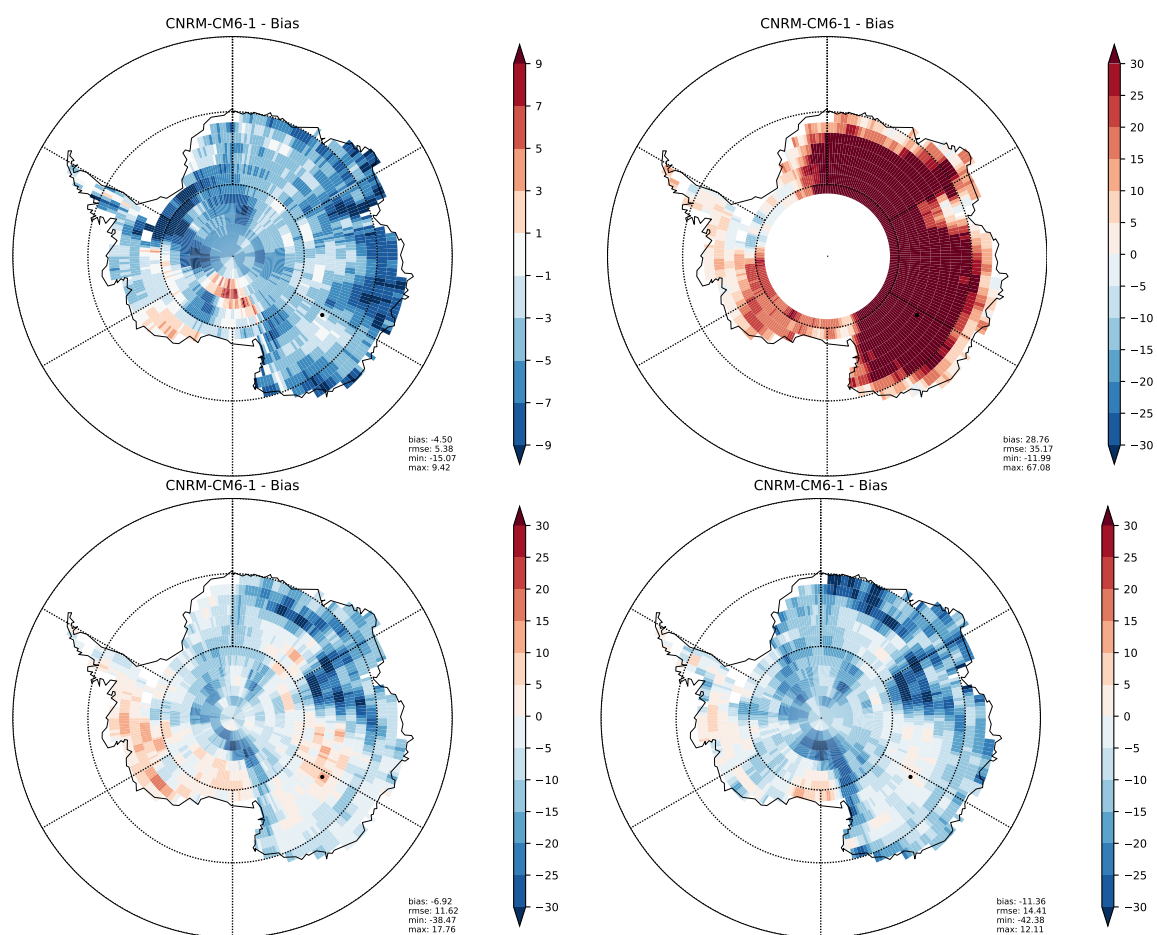


FIGURE 4.7 – Biais de ARPEGE-Climat 6.3 en hiver austral (JJAS) pour la température à 2 m (par rapport à BEST – en haut à gauche), la fraction nuageuse totale (par rapport à CALIPSO – en haut à droite), les effets radiatifs infrarouge des nuages en surface (par rapport à CERES – en bas à gauche) et le flux infrarouge net en surface (par rapport à CERES – en bas à droite).

Interim (voir section 3.4.2.2 pour plus de détails). On compare ses résultats sur l'Antarctique à ceux de la simulation AMIP examinée dans la section 4.2 sur la période estivale. L'objectif est de vérifier que les biais du modèle sont similaires, au moins au premier ordre, dans les deux simulations. On pourra alors étudier le comportement du modèle sur un site particulier (Dôme C) en comparant cette simulation guidée en vent à des séries d'observations réalisées *in-situ*.

Comme l'étude de la simulation guidée en vent sur Dôme C sera uniquement réalisée sur l'été austral, on ne montre ici que les biais sur cette saison. La figure 4.9 montre la structure spatiale des biais des différents types de température de la simulation guidée en vent comparée à celle des biais de la simulation AMIP. On constate que quelque soit le champ regardé (température de surface en ciel clair ou température à 2 m), ces structures sont similaires. On peut également noter que le biais froid est plus prononcé dans la simulation guidée en vent, traduisant une légère compensation d'erreur entre physique et dynamique dans la simulation AMIP.

La figure 4.10 compare les biais des champs de vent (vitesse et direction) à 10 m de la simulation guidée en vent avec ceux de la simulation AMIP, sur la période estivale. Concernant la vitesse du vent (fig. 4.10a), on ne montre le biais qu'à la surface du continent, étant donné que le guidage corrige parfaitement le modèle au-dessus de la couche limite. Comme pour la température, la structure spatiale du biais de vitesse du vent de la simulation guidée en vent est très proche de la simulation AMIP. L'ordre de grandeur est également sensiblement le même avec un biais de -1.13 ms^{-1} et de -1.17 ms^{-1} respectivement. Au-dessus du continent, la structure spatiale du biais de direction de vent est aussi similaire pour les deux simulations (Figure 4.10b). Au-dessus de l'océan, le biais de vent de la simulation guidée en vent est quasi nul : le vent dans la troposphère libre contraint donc presque totalement le vent proche de la surface océanique.

Notons que sur la période hivernal (JJAS) le comportement de la simulation guidée en vent est très proche là encore de la simulation AMIP (non montré). L'ensemble des résultats présentés ici montrent ainsi que la simulation guidée en vent est pertinente pour comprendre la simulation AMIP libre. Le guidage nous permet de comparer la simulation avec des séries d'observations *in situ* et justifie l'étude sur le site de Dôme C (voir la section 4.4).

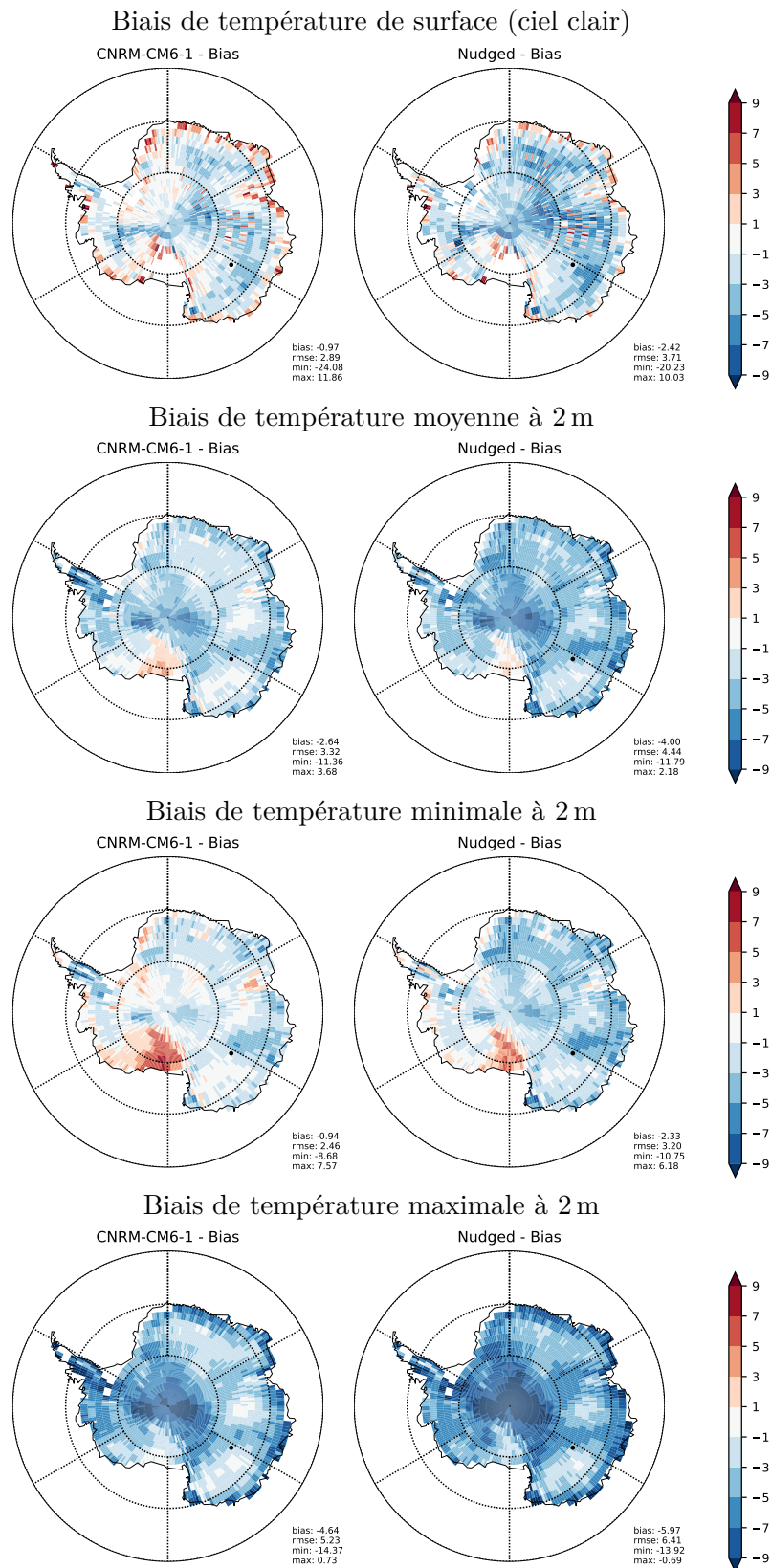


FIGURE 4.9 – Comparaison des biais de température des simulations AMIP (à gauche) et guidée en vent (à droite) de ARPEGE-Climat 6.3 en été austral (DJFM).

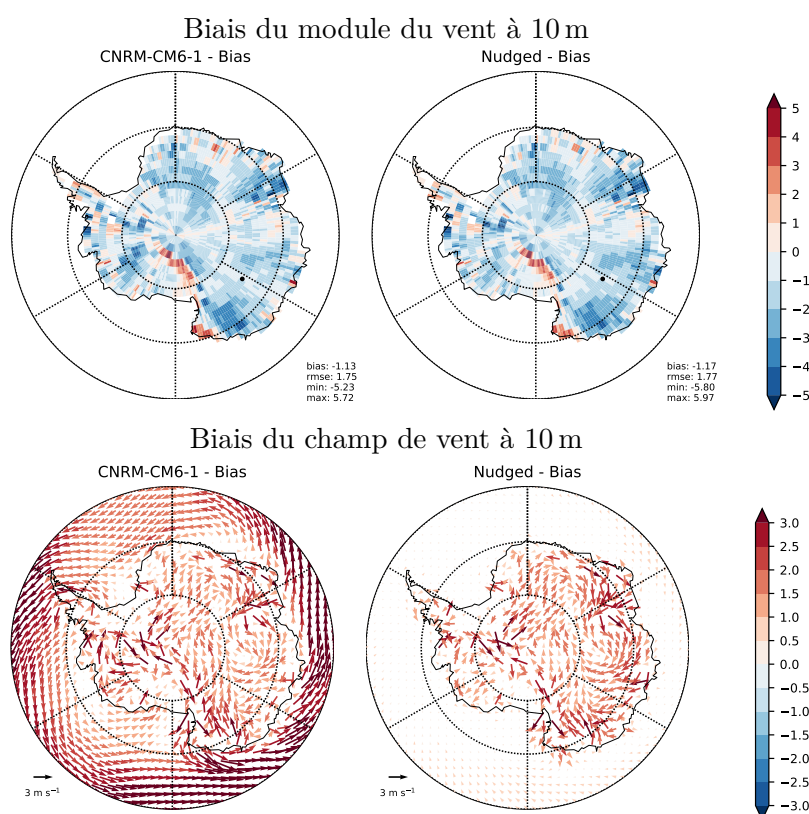


FIGURE 4.10 – Comparaison des biais de vent des simulations AMIP (à gauche) et guidée en vent (à droite) de ARPEGE-Climat 6.3 en été austral (DJFM).

4.3.2 Sensibilité de la simulation guidée en vent à l'albédo

Une correction empirique du calcul de l'albédo Le biais froid de ARPEGE-Climat 6.3 en été austral sur le plateau antarctique semble fortement lié à une surestimation de l'albédo de surface. Une correction empirique du calcul de l'albédo a donc été testée pour analyser la sensibilité du modèle à ce paramètre et valider (et quantifier) cette possible origine du biais froid. La modification ne concerne que le calcul de l'albédo de la première bande spectrale ($[0.3, 0.8] \mu\text{m}$, soit le proche ultraviolet et le visible – cf. Section 3.3.1) qui représente l'essentiel de l'énergie solaire arrivant à la surface (Decharme et al., 2019). L'équation 3.11 est modifiée comme suit :

$$\alpha_n(1) = \max \left[0.6, \min \left(0.92, \mathbf{0.85} - 1.58 \sqrt{d_{\text{opt}}} \right) \right] \quad (4.1)$$

Le terme **0.85** était initialement fixé à 0.96.

La modification du paramètre en rouge dans la formule 4.1 est purement empirique. Elle a été effectuée « à la main » de façon à obtenir une carte d'albédo sur l'intérieur de l'Antarctique Est proche de celle proposée par CERES (voir figure 4.11), et ainsi tester notre hypothèse que ce biais est important pour comprendre les biais de température sur le continent. Le cas

échétant, ce test imposera, au-delà de cette thèse, de revisiter la paramétrisation de l'albédo de la neige dans ARPEGE-Climat 6.3. Ce calcul fait intervenir de nombreux paramètres (via notamment le calcul du diamètre optique de la neige) dont les valeurs ont été calibrées dans des contextes différents (Decharme et al., 2016). Notons que l'âge de la neige n'est plus pris en compte dans ce nouveau calcul. Sur les zones côtières et les plates-formes de glace, pour lesquelles un albédo trop faible était déjà présent, le biais négatif est évidemment accentué. On évalue naturellement l'impact de cette modification sur le plateau antarctique pendant la période de l'été austral.

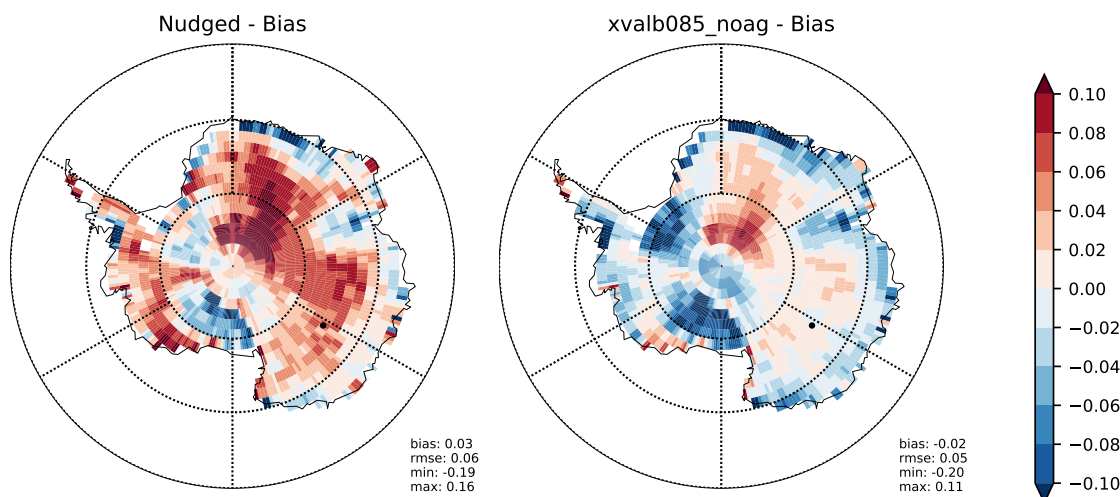


FIGURE 4.11 – Comparaison du biais des albédos de surface avant correction (à gauche) et après correction (à droite) par rapport à l'albédo CERES

Impact de la modification de l'albédo sur les biais du modèle nudgé en été austral

La figure 4.12 montre que la correction proposée du calcul de l'albédo permet de réduire de façon importante le biais de la température moyenne à 2 m, passant ainsi de -3.76 K à -0.96 K. Cette réduction du biais de température s'observe également sur le champ de température maximale, la modification de l'albédo modifiant le bilan radiatif solaire pendant la journée. Sur la région de Dôme C, le biais qui était légèrement négatif s'inverse et on observe alors un biais chaud, en particulier sur la température minimale. Ce point sera revu plus en détail dans la section 4.4 suivante.

La modification de l'albédo se retrouve également sur le champ de température de surface en ciel clair. On passe ainsi sur l'ensemble du continent d'un biais moyen froid de -2.71 K à un biais moyen chaud de 0.71 K. Ce biais chaud est particulièrement fort sur la région de Dôme C pouvant atteindre des valeurs supérieures à 3 K.

Les modifications du champ de température à la surface ont un impact direct sur les écoulements catabatiques. Ceux-ci sont moins intenses sur toute la zone côtière de l'Antarctique Est. Le biais existant déjà dans le modèle est accentué (Figure 4.13). Il convient toutefois de rester prudent sur ce résultat, la dynamique du modèle étant rappelée vers celle de ERAI (qui présente également un biais chaud), alors que la simulation est évaluée par rapport à la réanalyse ERA5.

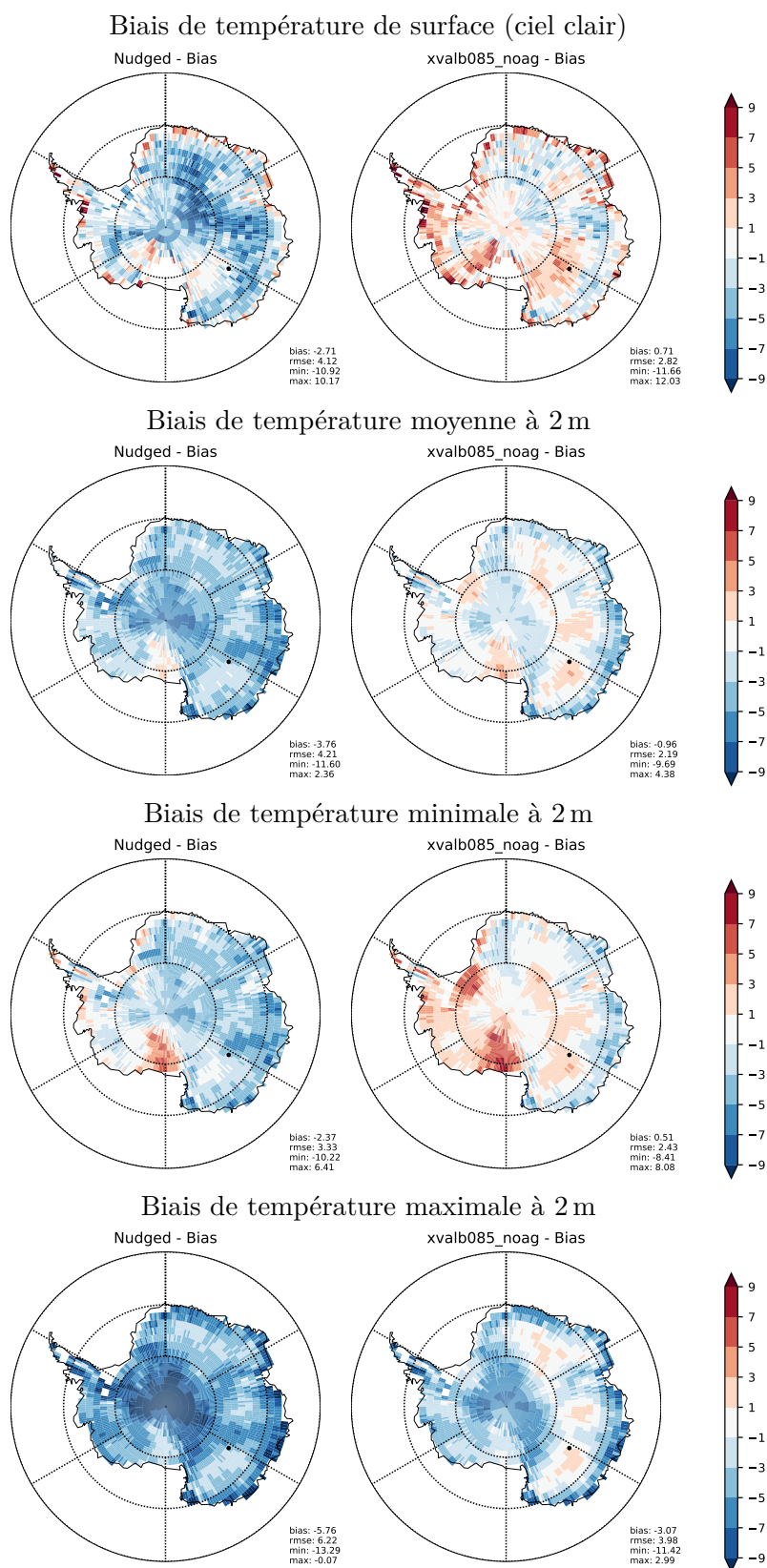


FIGURE 4.12 – Impact de la correction de l'albédo sur les champs de température des simulations guidées en vent de ARPEGE-Climat 6.3 en été austral (DJFM), avec à gauche les résultats avant correction et à droite les résultats après correction.

La réduction du biais de température en été austral sur l'Antarctique par la correction de l'albédo ne perdure pas pendant l'hiver austral. Le biais froid mis en évidence en hiver sur l'Antarctique ne provient donc pas a priori du biais de température estival (même si la faible durée de la simulation – 3 années – ne permet pas de conclure complètement et d'exclure un effet d'inertie du manteau neigeux).

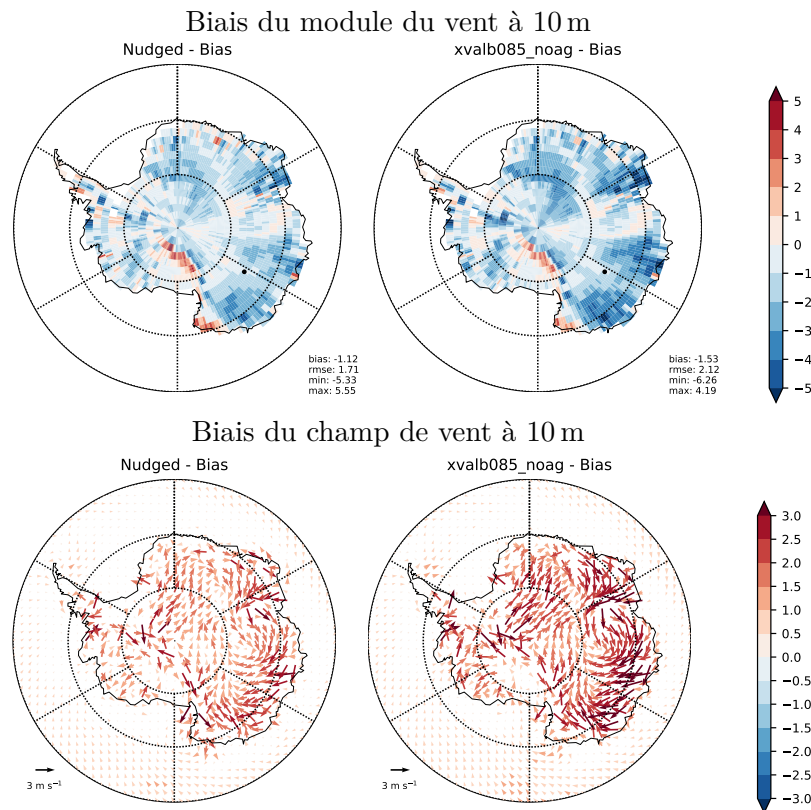


FIGURE 4.13 – Impact de la correction de l'albédo sur les champs de vent des simulations guidées en vent de ARPEGE-Climat 6.3 en été austral (DJFM), avec à gauche les résultats avant correction et à droite les résultats après correction.

4.4 Évaluation sur Dôme C

Évaluation sur le mois de janvier 2015 On évalue ici les deux versions (avec ou sans albédo corrigé) de ARPEGE-Climat guidées en vent sur le site de Dôme C. Ce site abrite une station de recherche sur laquelle sont installées une tour instrumentée de 45 m fournissant des mesures de vent et de température à différentes hauteurs et une station du réseau Baseline Station Radiation Network (BSRN) (Lanconelli et al., 2011) qui fournit des données de rayonnement desquelles on peut déduire une température de surface. La simulation guidée en vent s'étend sur les années 2014-2016 et l'évaluation ci-dessous se concentre sur le mois de janvier 2015 pour lequel les diverses séries temporelles de vent, de température et de rayonnement sont les plus complètes.

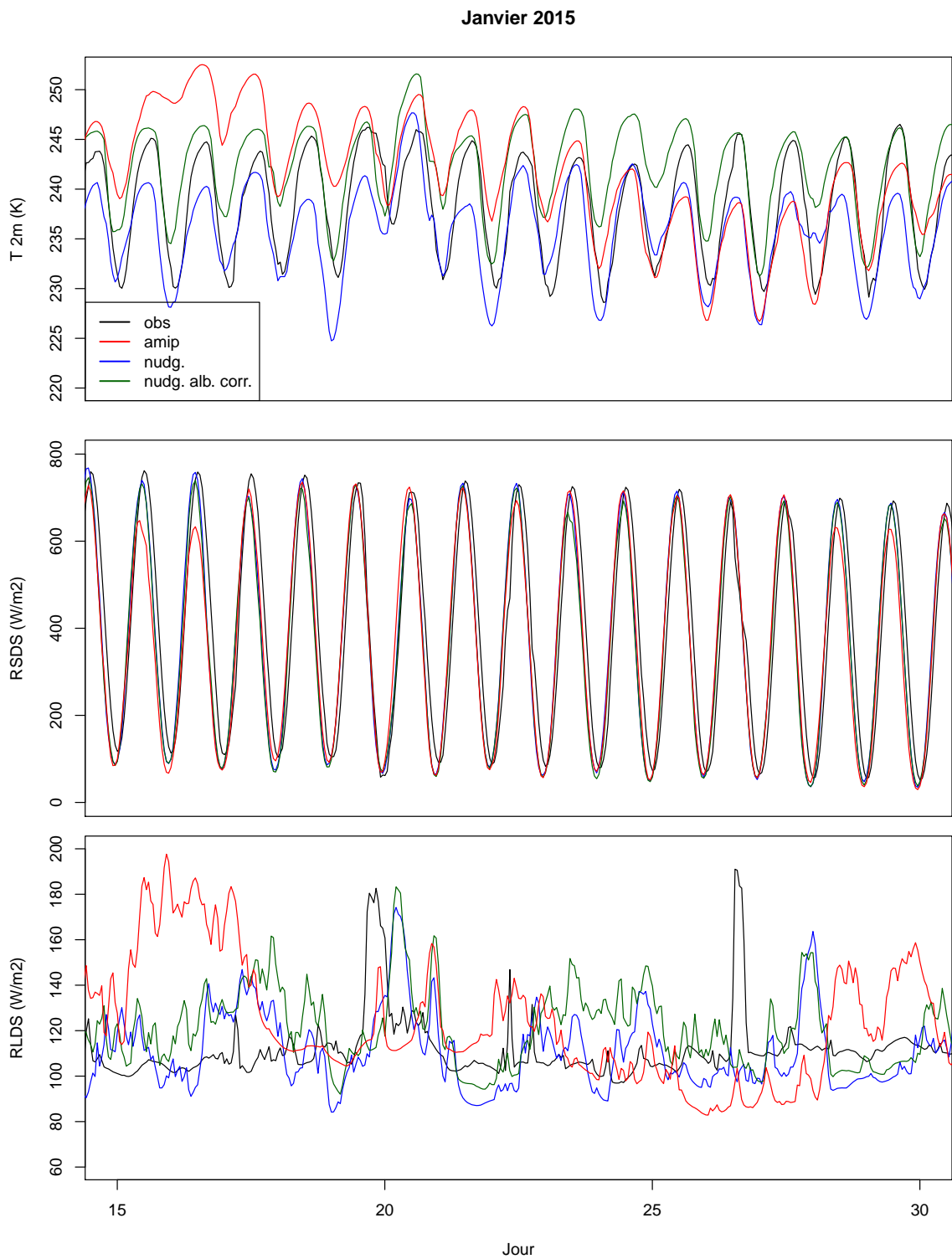


FIGURE 4.14 – Comparaison des différentes simulations avec les observations *in situ* sur la période allant du 15 au 30 Janvier 2015.

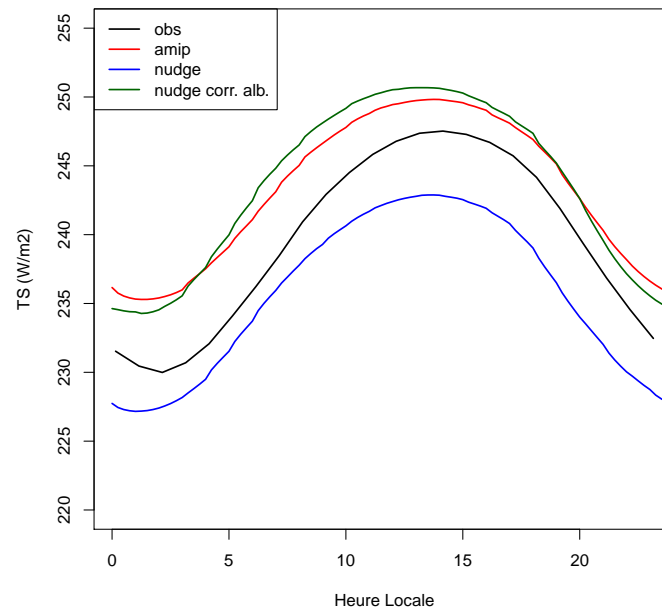


FIGURE 4.15 – Cycle diurne moyen de la température de surface sur le mois de janvier 2015

La figure 4.15 montre l'évolution temporelle de la température à 2 m dans les différentes simulations réalisées ainsi que dans les observations *in-situ* (station BSRN), sur la période allant du 15 au 30 janvier 2015. Les résultats de la simulation guidée en vent présente un biais froid sur toute la période. La correction de l'albédo inverse ce biais et permet d'avoir des températures maximales qui se rapprochent des températures observées. Les températures minimales sont régulièrement sur-estimées. La sous-estimation de l'amplitude du cycle diurne par le modèle est un des problèmes souvent évoqué dans la représentation des couches limites stables dans les modèles numériques dû à une surestimation du mélange turbulent pendant la nuit. L'origine de ces biais peut également être due à la représentation des nuages dans le modèle ou à l'inertie thermique de la neige.

Les figures 4.14b et 4.14c montrent les flux à la surface (respectivement solaire descendant et infrarouge descendant). Les flux des deux simulations guidées en vent sont comparables dans leur chronologie mais des différences d'intensité peuvent apparaître (visibles principalement dans l'infrarouge). Les différences de ces flux avec ceux mesurés par la station BSRN peuvent s'expliquer par des différences dans les scènes nuageuses et être à l'origine d'une partie des biais de température à 2 m.

La figure 4.16 présente le cycle diurne de la température de surface des différentes simulations et les compare avec la température mesurée par la station BSRN, moyenné sur le mois de janvier 2015. Le biais froid de température de la simulation sans correction de l'albédo est retrouvé et est plus marqué en milieu de journée (vers 15h00). La correction de l'albédo inverse ce biais. L'amplitude du cycle diurne est légèrement augmentée (17.5 K contre 16.5 K pour la simulation sans correction de l'albédo) et est comparable avec l'amplitude moyenne

mesurée 18.0 K.

Le cycle diurne des différents flux à la surface est montré sur la figure 4.16. La correction de l'albédo permet d'obtenir un flux solaire montant similaire à celui mesuré. Les flux infrarouges montant et descendant reproduisent les biais de température (négatif pour la simulation sans correction et positif pour la simulation avec correction).

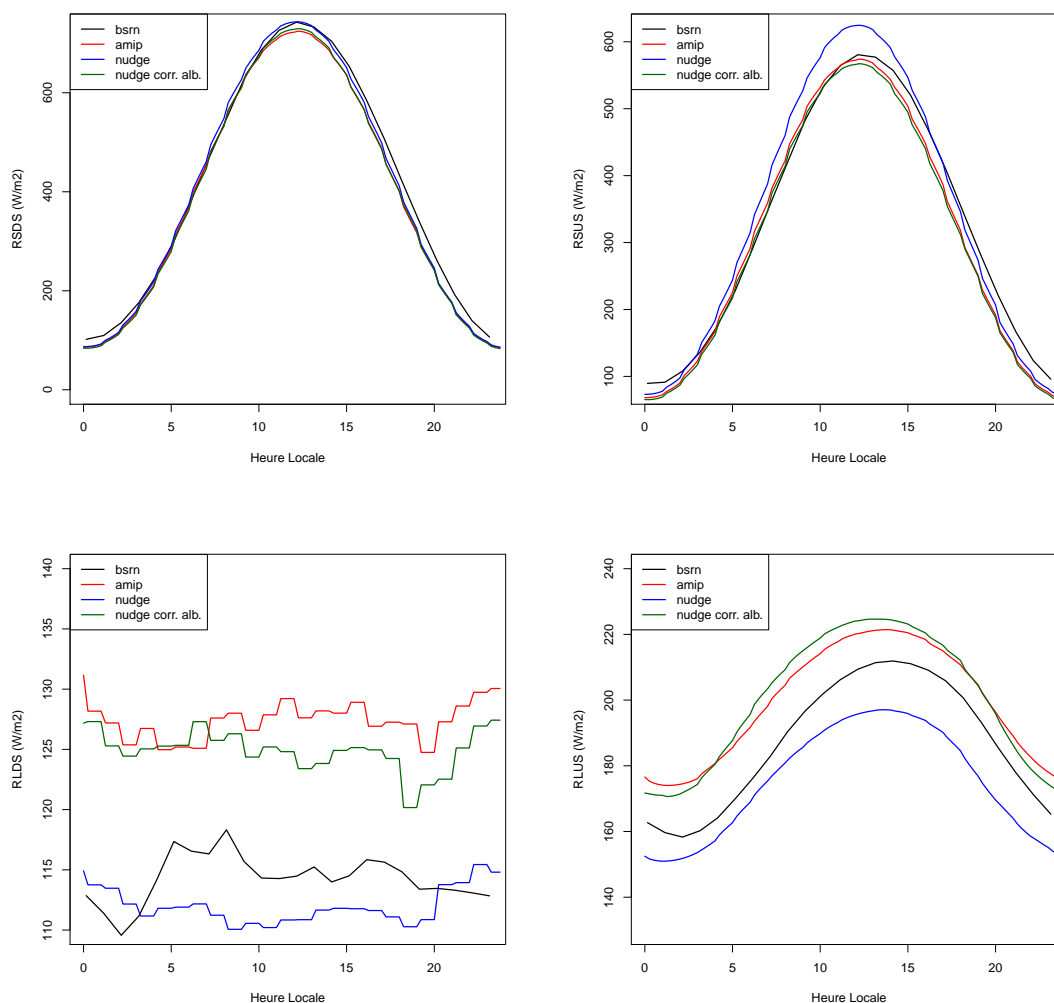


FIGURE 4.16 – Cycle diurne moyen des différents éléments du bilan radiatif de la surface sur le mois de janvier 2015, avec le rayonnement solaire sur la première ligne et infrarouge sur la deuxième ligne. La première colonne correspond au rayonnement descendant et la deuxième au rayonnement montant

Profils verticaux en janvier 2015 La figure 4.17 présente les profils verticaux moyens de température et de vent des simulations guidées en vent avec ou sans correction de l'albédo. Ils sont comparés avec les profils moyens observés sur le mat instrumenté. Les moyennes

sont calculées sur la période de janvier 2015 (pour laquelle un nombre suffisant de données observées est disponible).

On avait montré précédemment que les températures à 2 m des simulations guidées en vent sans correction d'albédo étaient trop froides par rapport aux données du mat instrumenté (Figure 4.14). Les simulations avec l'albédo corrigé présentaient au contraire un biais chaud sur les températures à 2 m. On retrouve ce même constat sur les profils verticaux moyens que ce soit pour les profils diurnes ou nocturnes.

Le profil de vent nocturne (à 2300 heure locale) est peu impacté par la correction de l'albédo et s'écarte sensiblement des observations pour lesquelles une structure de jet de basse couche se dessine nettement malgré la faible altitude des mesures réalisées (avec un maximum à une altitude vers 20 à 30 m), ce qui est caractéristique des couches limites stables. Le profil diurne (à 1100 heure locale) des simulations guidées en vent (avec ou sans correction d'albédo) est plus proche de celui observé. La sous-estimation du jet de basse-couche dans des situations stables est caractéristique de simulations produisant un mélange turbulent surestimé (e.g. Sandu et al., 2013). Cette surestimation du mélange sera mise en évidence sur le cas d'étude GABLS4 dans le chapitre 6.

4.5 Conclusion

Le comportement de ARPEGE-Climat 6.3 est caractérisé par un biais froid sur l'ensemble du continent antarctique (à l'exception des plates-formes de glace). La structure de ce biais est similaire pour la simulation AMIP et la simulation guidée en vent. Le biais froid en été austral peut être relié à une surestimation de l'albédo de surface. Une correction empirique du calcul de ce dernier permet d'obtenir des valeurs d'albédo plus cohérentes avec celles mesurées, en particulier sur la région de Dôme C, que ce soit par rapport à des mesures spatiales ou *in situ*. La simulation guidée en vent avec albédo corrigé présente alors un biais chaud sur la région de Dôme C, plus prononcé pour les températures minimales. L'analyse de ces simulations sur le mois de janvier 2015 et leur comparaison avec les données de la tour instrumentée et de la station BSRN de Dôme C donne des résultats similaires.

Plusieurs indices laissent penser que le biais chaud du modèle avec albédo corrigé peut être relié à une sur-estimation du mélange turbulent pendant la nuit (en particulier une sous-estimation de l'amplitude des cycles diurnes). Le cas GABLS4 basé sur des observations *in situ* à Dôme C en décembre 2009 sera utilisé pour tester ces hypothèses et effectuer une étude plus fine du modèle et de sa sensibilité à sa représentation de la turbulence.

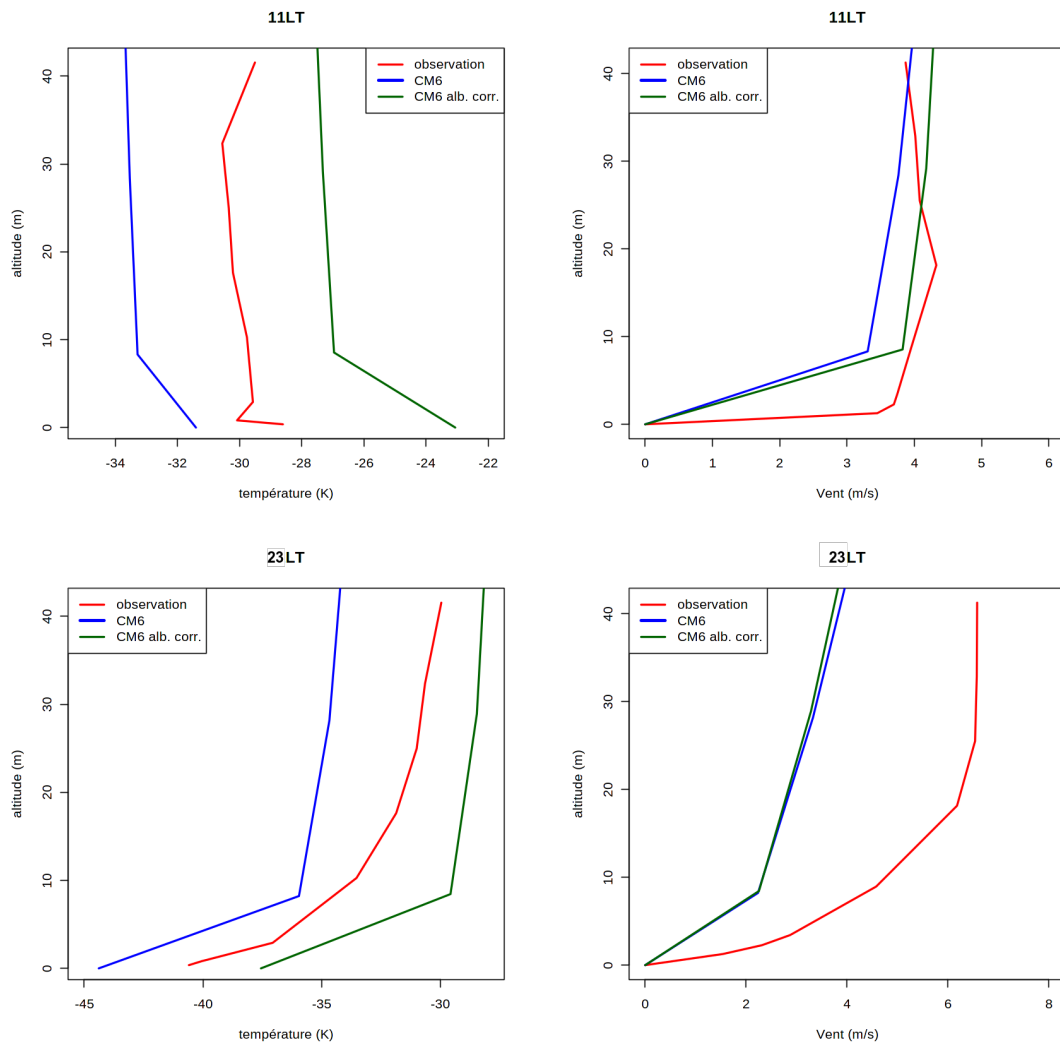


FIGURE 4.17 – Profils verticaux moyens pour janvier 2015 de vent (à droite) et de température (à gauche) à 1100LT (en haut) et à 2300LT (en bas).

Calibration d'un modèle de climat - Calibration statistique

Sommaire

5.1	Calibration de la physique d'un modèle	72
5.1.1	Généralités	72
5.1.2	Problèmes et stratégies	74
5.1.3	La calibration de ARPEGE-Climat 6.3	77
5.2	Vers une nouvelle stratégie. Outils statistiques pour une calibration « orientée processus »	78
5.2.1	La comparaison SCM/LES	80
5.2.2	<i>History Matching</i> et calibration d'un SCM	81
5.3	Conclusions	86

La modélisation numérique est un outil puissant pour étudier des systèmes complexes. Une telle modélisation repose généralement sur un certain nombre d'approximations ou de simplifications. Ces dernières nécessitent l'introduction de paramètres dits libres dans le modèle. Par définition, ces paramètres n'ont pas de valeur définie absolument. Il est pourtant nécessaire de leur en donner une pour réaliser des simulations. La calibration d'un modèle numérique consiste à choisir une valeur pour chacun des paramètres libres de façon à ce que le modèle soit capable de reproduire correctement les phénomènes que l'on cherche à modéliser. On évalue cette capacité du modèle à travers un ensemble de métriques choisies de façon à représenter les principales caractéristiques du phénomène étudié. Elles peuvent être vues comme des cibles à atteindre ou comme des notes caractérisant les performances du modèle. La calibration d'un modèle numérique est un problème scientifique en soi que l'on rencontre dès que l'on veut modéliser des phénomènes complexes. Dans le cas des modèles de climat, cette étape est longtemps restée peu documentée et était vue comme purement technique et relativement laborieuse, relevant plus de l'ingénierie que de la science. Ce travail de calibration peut se rencontrer dans le développement de la partie dynamique du modèle (par exemple dans le réglage d'un terme de diffusion, souvent nécessaire pour des raisons de stabilité numérique du code), mais il concerne essentiellement la partie physique du modèle, *i.e.* l'ensemble des différentes paramétrisations permettant de prendre en compte les phénomènes d'échelles plus petites que celle résolue par le noyau dynamique, ou relevant de phénomènes non représentés par la mécanique des fluides (e.g., rayonnement, microphysique, chimie ...). Durant la dernière décennie, la nécessité de documenter cette étape cruciale du développement des modèles a été soulevée (e.g. Mauritsen et al., 2012, Hourdin et al., 2017, Schmidt

et al., 2017). En particulier, des études ont mis en évidence que les valeurs de ces paramètres sont une source importante d'incertitude pour les projections climatiques. Mauritsen et al. (2012), Golaz et al. (2013) montrent par exemple que pour un modèle donné, différents jeux de paramètres permettant de reproduire correctement le climat actuel peuvent conduire à des projections climatiques sensiblement différentes ou à des climats passés différents.

On présente dans ce chapitre la problématique générale de la calibration d'un modèle de climat ainsi que les stratégies couramment utilisées (section 5.1). En particulier, l'approche suivie pour calibrer ARPEGE-Climat qui a permis d'aboutir à la version 6.3 est décrite (voir section 5.1.3). Une nouvelle approche, développée dans le cadre du projet ANR HighTune, inspirée des techniques d'*History Matching* et reposant sur la comparaison d'une version 1D du modèle à calibrer avec des simulations LES est proposée dans la section 5.2. Cette approche vise non pas à exhiber un jeu de paramètres qui serait le meilleur, contrairement à une approche de type optimisation, mais à définir des régions de l'espace des paramètres dans lesquelles le comportement du modèle est compatible avec les données observées. Cette démarche nécessite d'évaluer le modèle pour un nombre important de jeux de paramètres, ce qui est en pratique complètement irréaliste avec un modèle de climat du fait du coût numérique d'une simulation. Pour palier cette difficulté, la méthode utilisée ici s'appuie largement sur l'utilisation d'émulateurs ou métamodèles, *i.e.* des modèles utilisant des méthodes d'apprentissage statistique, qui « émulent » la réponse du modèle pour une métrique donnée sans avoir besoin de réaliser de nouvelles simulations. Parmi toute une zoologie de modèles statistiques disponibles, la méthodologie présentée ici utilise des émulateurs basés sur des processus gaussiens et est similaire aux techniques de krigeage utilisées en géostatistique.

5.1 Calibration de la physique d'un modèle

5.1.1 Généralités

Aujourd'hui, les modèles numériques jouent un rôle clé dans le domaine des sciences du climat. On peut les voir comme des laboratoires numériques permettant d'améliorer notre compréhension du système climatique et de son fonctionnement. Les modèles de climat se sont peu à peu complexifiés pour tenir compte des interactions entre ses différentes composantes (atmosphère, océan, glaces de mer, surfaces continentales). Le contexte du changement climatique et de son étude a fait sortir le monde de la modélisation du climat du domaine purement académique. Les modèles de climat sont devenus des outils pour comprendre l'évolution récente et future du climat, notamment en lien avec l'activité anthropique, et ainsi d'éclairer les choix des décideurs sur des politiques de réduction des émissions de gaz à effet de serre ou d'adaptation au changement climatique. Pour cela, des projections à différentes échelles de temps sont réalisées par les différents centres de modélisation du climat afin de comprendre et de quantifier les effets de l'activité humaine sur le système climatique, de comprendre et évaluer les mécanismes à l'œuvre, en tenant compte de la complexité du système climatique, et des multiples interactions entre les différentes échelles spatiales et temporelles (Dahan Dalmedico, 2001, Held, 2005). Cependant, les projections climatiques réalisées à l'aide de ces modèles sont entachées d'incertitudes, au premier lieu desquelles les incertitudes sur les futures émissions de gaz à effet de serre. Ces incertitudes sont prises en compte en réalisant des

simulations pour différents scénarios allant d'une réduction drastique de ces émissions à une augmentation continue tout au long du XXI^{ème} siècle. Des incertitudes plus fondamentales, intrinsèques aux modèles numériques, sont liées à leur caractère nécessairement imparfait du fait des hypothèses simplificatrices qui sont faites lors de leur construction ou des lacunes sur la connaissance et la compréhension des phénomènes à l'œuvre dans le système climatique. Elles proviennent en partie des paramétrisations physiques. Une tâche importante dans le développement d'un modèle numérique de climat est le développement de ces paramétrisations qui permettent de prendre en compte l'effet de processus sous-maille qui ne sont pas résolus par le noyau dynamique du modèle ou de processus ne relevant pas de la mécanique des fluides (voir chapitre 3, section 3.2). On peut voir ces dernières comme des approximations qui tentent de résumer des processus complexes et qui doivent être valables pour l'ensemble des climats présents sur Terre, depuis les climats tropicaux aux régions polaires, et au-dessus de tout type de surfaces, qu'elles soient continentales ou océaniques.

Des propositions ont été faites pour remplacer les paramétrisations par des méthodes de type deep learning et des travaux sont en cours dans de nombreux laboratoires de modélisation du climat. Si cette nouvelle approche semble prometteuse (Brenowitz and Bretherton, 2018, Gentine et al., 2018), elle présente l'inconvénient de masquer au sein de la « boîte noire » du machine learning les processus physiques à l'œuvre. Dans le cas de ARPEGE-Climat, les paramétrisations physiques sont construites sur un ensemble de lois ou de principes décrivant les processus mis en jeu. Parmi les paramétrisations les plus critiques pour les modèles de climat et qui continuent de poser des problèmes, on trouve celles concernant la représentation de la convection, de la microphysique des nuages, des interactions nuages-rayonnement et de la turbulence.

Pour formaliser le problème, on rappelle l'équation 3.1 résumant un modèle de climat et introduite au chapitre 3 :

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{D}(\mathbf{x}) + \sum_p \mathcal{P}_p(\mathbf{x}, \boldsymbol{\lambda}_p) \quad (5.1)$$

où $\boldsymbol{\lambda}_p$ est l'ensemble des paramètres libres qui interviennent dans la paramétrisation physique p . Pour chacun de ces paramètres, une valeur est proposée initialement lors du développement des différentes paramétrisations. Ces valeurs peuvent provenir de la théorie, d'une estimation plus ou moins bien documentée, d'expériences numériques (cas d'étude), ou d'observations. On notera que de nombreux paramètres libres ne sont pas directement mesurables. Lors de la calibration de la physique d'un modèle, le travail du modélisateur consiste alors dans un premier temps à identifier parmi l'ensemble des paramètres libres $\boldsymbol{\lambda}_p$ ceux pour lesquels il est nécessaire de discuter la valeur et à identifier pour ces derniers une plage de valeurs potentielles. Il doit ensuite définir des métriques (en général une grandeur scalaire mais on peut également considérer des grandeurs vectorielles) décrivant les caractéristiques du système climatique qu'il souhaite voir correctement reproduites par le modèle. Cette étape est très dépendante des applications visées par le modèle. Par exemple, un modèle de climat devra avoir un bilan d'énergie à l'équilibre ou très proche de l'équilibre alors qu'un modèle de prévision de temps n'aura pas cette contrainte, la durée des simulations réalisées n'étant pas suffisante pour qu'un modèle qui ne serait pas à l'équilibre dérive. Chacune de ces métriques

devra être associée à une incertitude sur la valeur à obtenir afin d'éviter des problèmes de surajustement. Enfin le modélisateur devra trouver des valeurs satisfaisantes des paramètres libres sélectionnés de façon à ce que le modèle reproduise correctement le système climatique, *i.e.* de façon à ce que les métriques correspondent aux observations, modulo leurs incertitudes respectives.

Cette étape est essentielle mais est longtemps restée peu documentée dans la littérature scientifique documentant les différents modèles de climat. De fait, la calibration a longtemps été considérée comme une « cuisine interne » qui relevait plus de l'ingénierie que de la science. Mauritsen et al. (2012) relève que le terme même de calibration d'un modèle n'est pas très bien défini. Elle est souvent considérée comme l'étape finale du développement du modèle, afin que celui-ci reproduise correctement un certain nombre de métriques caractérisant le système climatique (Randall and Wielicki, 1997). Hourdin et al. (2017) propose une définition de la calibration :

"The process of estimating uncertain parameters in order to reduce the mismatch between specific observations and model results is usually referred to as tuning in the climate modelling community."

Si on s'en tient à cette dernière définition, la calibration d'un modèle de climat peut être définie comme un processus objectif d'estimation des paramètres libres de ce modèle pour correspondre à un ensemble prédéfini d'observations, tout en tenant compte de l'incertitude sur ces dernières. Cependant, en raison de la complexité du système climatique, des choix et approximations effectués dans chaque paramétrisation, et en raison des priorités définies dans chaque centre climatique, il émerge une subjectivité dans le réglage des modèles climatiques (Tebaldi and Knutti, 2007). Cette subjectivité, à travers les choix effectués durant cette étape de calibration, a un impact important sur les résultats et les performances d'un modèle (Mauritsen et al., 2012, Golaz et al., 2013). On décrit succinctement dans la section suivante les problèmes principaux et les enjeux de la calibration d'un modèle ainsi que les stratégies généralement utilisées.

5.1.2 Problèmes et stratégies

Pourquoi la calibration d'un modèle de climat est-elle une tâche difficile ? La difficulté de l'étape de calibration d'un modèle de climat est liée d'une part au coût numérique de la réalisation d'une simulation climatique, et d'autre part au nombre important de paramètres à régler. De plus on doit considérer l'existence d'interactions possibles entre les différents processus au sein d'une même paramétrisation, ou entre les différentes paramétrisations. Ainsi, si on note N le nombre de ces paramètres, et si on considère que pour explorer de façon satisfaisante l'espace de tous les réglages possibles du modèle on doit tester une dizaine de valeurs pour chaque paramètre, on obtient un nombre de 10^N simulations à réaliser. L'ordre de grandeur de N pour un modèle de circulation générale étant autour de la centaine, on comprend qu'une telle approche est irréalisable. Une autre difficulté, que l'on a déjà évoquée dans la section précédente, est la diversité des applications des simulations climatiques. Le nombre de ces applications est en constante augmentation et chacune va potentiellement ajouter de nouvelles cibles ou objectifs à l'exercice de calibration du modèle.

L'ajout de ces contraintes va ainsi potentiellement modifier les réglages du modèle. De plus le modèle peut ne pas être capable de satisfaire l'ensemble de ces contraintes en même temps. Enfin, le fait de ne pas pouvoir explorer le comportement du modèle de façon satisfaisante sur l'espace des paramètres rend délicat l'attribution d'un mauvais comportement du modèle. Est-il relié à un défaut ou un choix de calibration ou bien relève-t-il d'une limite intrinsèque d'une paramétrisation donnée ?

Stratégies de calibration d'un modèle de climat Hourdin et al. (2017) présentent les résultats d'un sondage réalisé auprès de centres développant des modèles de climat sur les stratégies de calibration utilisées pour régler leur modèle couplé océan/atmosphère. Il ressort de ce sondage que les objectifs de ces différents centres présentent une certaine diversité qui reflète leurs préoccupations respectives. Cependant l'objectif de garantir un quasi-équilibre énergétique du modèle apparaît nettement comme critique pour l'ensemble des centres. Ainsi des métriques telles que le bilan radiatif net au sommet de l'atmosphère, et dans une moindre mesure, la température moyenne à la surface, apparaissent comme les plus importantes. Ces cibles étant principalement sensibles aux propriétés des nuages et à l'albédo de la surface, les paramètres clés à régler sont ceux qui ont un impact sur la représentation des nuages et sur les propriétés de la surface. Hourdin et al. (2017) citent par exemple des paramètres jouant sur les propriétés optiques des nuages, la microphysique des nuages, l'albédo de la neige ou de la glace. Ces métriques qui apparaissent comme consensuelles sont cependant des métriques globales qui sont le résultat de l'interaction entre de nombreux processus. Régler un modèle de climat sur la base de ces métriques n'est pas raisonnable et est une source potentielle de compensation d'erreurs. Un exemple célèbre de compensation d'erreurs dans la représentation des nuages émergeant de métriques globale est le problème *too few too bright* (Nam et al., 2012) : une sous-estimation de la couverture nuageuse peut être compensée par une sur-estimation de l'effet radiatif des nuages dans le spectre solaire.

De fait la calibration n'intervient pas qu'en toute dernière étape avec une version couplée océan/atmosphère, elle est présente dès le début du développement des paramétrisations physiques. Hourdin et al. (2017) propose d'ailleurs pour la calibration des modèles de climat une approche en trois étapes.

- Lors du développement d'une nouvelle paramétrisation, ou lors de son intégration dans un modèle, des versions SCM du modèle sont généralement utilisées afin de pouvoir tester la paramétrisation dans différents cas. Idéalement, lors de cette étape, on aimerait pouvoir calibrer chaque paramétrisation de façon indépendante des autres paramétrisations. En pratique les interactions entre les différents processus peuvent rendre cet objectif délicat voire impossible à atteindre. Les effets de la grande échelle sont représentés par les conditions de forçage aux limites de la colonne et la paramétrisation ne rétroagit alors pas sur la grande échelle. À l'issue de cette étape, on souhaite avoir identifié pour chacun des paramètres mis en jeu dans ces expériences, une valeur de référence ainsi qu'une plage de valeurs dans laquelle on considère que la paramétrisation se comporte de façon satisfaisante. Les valeurs de référence proposées serviront de point de départ pour l'étape suivante, et les plages de valeurs éventuellement définies serviront à définir l'espace des possibles dans lequel les paramètres peuvent évoluer.
- L'étape suivante est la calibration de chacune des grandes composantes d'un modèle

climatique (atmosphère, surfaces continentales, océan) à travers des simulations globales. Les interactions avec la grande échelle sont alors prises en compte. Chacune de ces composantes est forcée par des conditions aux limites réalistes qui remplacent les interactions avec les autres composantes du système climatique. Pour la composante atmosphérique, on souhaite avoir un modèle satisfaisant à l'ordre 0 un (quasi) équilibre radiatif du modèle, et à l'ordre 1 qu'il garantisse des résultats raisonnables sur un certain nombre de métriques résumant le climat moyen (distribution moyenne méridienne des précipitations par exemple) généralement utilisées pour évaluer les modèles.

- Enfin une dernière étape de calibration est effectuée avec un couplage complet des différentes composantes afin de vérifier que la calibration proposée à l'étape précédente reste pertinente, et éventuellement procéder à des ajustements.

Cette décomposition en trois étapes n'exclut pas des allers et retours entre celles-ci (par exemple du fait de compensations d'erreurs entre paramétrisations et de l'ajout de métriques au fil des étapes). La calibration de ARPEGE-Climat 6.3, qui a accompagné d'importants changements de la physique du modèle, a ainsi nécessité de nombreuses itérations entre ces différentes étapes et est décrite dans la section 5.1.3.

Méthodes de calibration Historiquement, les méthodes utilisées pour calibrer les modèles de climat sont des méthodes itératives de type essai-erreur. Pour un ensemble de métriques donné, on modifie un ou deux paramètres pertinents pour ces métriques. La modification est retenue si elle permet d'améliorer le modèle sur la base de métriques sensibles à ces paramètres. Une nouvelle modification est alors testée et ainsi de suite, jusqu'à obtenir un modèle satisfaisant pour les modélisateurs. La méthode peut alors être répétée pour un autre ensemble de métriques, etc. Cette méthode longue et fastidieuse, qui correspond à un plan d'expérience de type « *one factor at a time* » (OFAT) est connue pour ne pas permettre de révéler efficacement les interactions entre les différents paramètres (Williamson et al., 2015). De plus cette approche est hautement subjective et très peu reproductible. Différentes méthodes ont été proposées pour réaliser un processus de calibration plus objectif et plus reproductible en appréhendant la calibration d'un modèle de climat comme un problème d'optimisation. Une première méthode est alors d'introduire une fonction de coût et d'utiliser un algorithme de minimisation (e.g. Severijns and Hazeleger, 2005, Zhang et al., 2015). Parmi ces méthodes d'optimisation, une des plus populaires est la calibration bayésienne qui est issue de la communauté de Quantification de l'Incertitude et qui a l'avantage de prendre en compte les différentes sources d'incertitude (observations, modèle) et qui vise à déterminer une distribution probabiliste de l'optimum. Si cette méthode peut être intéressante pour calibrer une version 3D du modèle (étape 2 et 3 de la méthodologie présentée ci-dessus), elle pose quelques problèmes si on l'applique à une calibration orientée processus de versions 1D du modèle : en effet cette première étape vise non pas à exhiber un jeu de paramètres optimal mais plutôt à définir une région de l'espace des paramètres dans laquelle la calibration de la version 3D du modèle pourra s'effectuer, réduisant le risque de surajustement du modèle. Une autre méthode issue de la communauté de Quantification de l'Incertitude et appelée *History Matching* permet précisément de répondre à cette question. Cette méthode sera appliquée plus loin pour calibrer une version 1D du schéma de turbulence de ARPEGE-Climat (voir chapitre 6) et est présentée dans la section 5.2.

5.1.3 La calibration de ARPEGE-Climat 6.3

Le développement d'une nouvelle version de ARPEGE-Climat pour les exercices de modélisation CMIP6 a été l'occasion d'une refonte importante de la physique du modèle avec en particulier l'introduction de nouveaux schémas de turbulence, de convection ou de microphysique, ainsi que le couplage avec un nouveau schéma de surface. L'intégration de ces différentes composantes a nécessité un important travail de calibration pour aboutir à la version 6.3 de ARPEGE-Climat. Ce travail est décrit en détail dans Roehrig et al. (2020) et nous en retraçons ici les grandes lignes. Notons tout d'abord que seule la composante atmosphérique du modèle a été calibrée. La calibration du schéma de surface avait été effectuée précédemment (Decharme et al., 2019) et il n'y a pas eu de revisite de ces résultats pour la calibration de ARPEGE-Climat 6.3.

Ce travail se place dans le schéma général de calibration proposé par Hourdin et al. (2017) et reprend en particulier les deux premières étapes, à savoir une étape de calibration orientée processus via des versions SCM du modèle sur des cas d'études, puis via des simulations AMIP de durées variables (allant de quelques années à plusieurs dizaines d'années selon les métriques ciblées). Roehrig et al. (2020) relève que cette calibration de ARPEGE-Climat a nécessité de nombreux allers et retours entre ces deux étapes du fait de l'importance des modifications apportées à la physique du modèle. Pour chacune de ces étapes, la démarche suivie est itérative et empirique : un ou deux paramètres sont modifiés et le comportement du modèle est évalué via différentes métriques décrites dans la suite.

Première phase : Calibration orientée processus Pour chacune de ces itérations, la première étape se basait sur des cas d'études (principalement de convection, profonde ou peu profonde). La calibration se faisait en comparant les résultats des SCM soit avec des observations, soit avec des simulations LES (voir 3.5.2). Cette étape permet de vérifier que le modèle contient la physique nécessaire pour représenter les différents processus en jeu dans ces cas d'étude.

Deuxième phase : Calibration des simulations AMIP Le premier objectif de la calibration des simulations AMIP (pour lesquelles la composante atmosphérique des modèle est forcée par la température des océans) est de garantir un quasi-équilibre énergétique du modèle. C'est une contrainte forte mais qui est indispensable afin d'éviter une dérive de la température moyenne de la Terre quand le modèle sera couplé à un modèle océanique. Le réglage du bilan énergétique du modèle est un travail délicat et demande de régler les différentes composantes du bilan radiatif sans que l'on ait d'observations suffisamment précises pour l'ensemble de celles-ci. Ce travail est pourtant primordial et une différence de $\pm 1 \text{ W.m}^{-2}$ induit une différence de température moyenne de $\pm 1 \text{ K}$ dans le modèle couplé. Ce premier objectif est réalisé en utilisant des métriques globales, telles que les bilans radiatifs solaires et infrarouges à la surface et au sommet de l'atmosphère.

De nouvelles métriques sont alors ajoutées de façon que le modèle reproduise bien les grandes caractéristiques du climat actuel. Parmi celles-ci, on distingue un premier groupe de métriques utilisées systématiquement à chaque itération et qui consiste en des moyennes zonales ou en des cartes de caractéristiques du climat (moyennes annuelles ou saisonnières de température,

de précipitations, de pression, vent, flux radiatifs, fraction nuageuse, effet radiatif des nuages, ...). Un deuxième groupe de métriques plus spécifiques a été utilisé mais de façon moins systématique (distribution quotidienne des précipitations sur des régions données, variabilité intrasaisonnière tropicale, oscillation quasi biennale).

On notera que les paramètres utilisés pour calibrer le modèle proviennent, pour la plupart, de la paramétrisation de la convection et des propriétés microphysiques et radiatives des nuages. Les paramètres contrôlant les propriétés radiatives des nuages n'interviennent que lors de la calibration des simulations AMIP car leur importance n'est visible que sur des contraintes de grande échelle, les cas utilisés lors de la phase « SCM » ne permettant pas actuellement d'étudier les aspects radiatifs des nuages.

Si ce travail a permis d'obtenir un modèle atmosphérique (ARPEGE-Climat 6.3), un modèle couplé (CNRM-CM6-1 – Voltaire et al., 2019) et un modèle système-Terre (CNRM-ESM2-1 – Sférian et al., 2019) à l'état de l'art, la démarche suivie, classique dans la communauté de la modélisation du climat, par laquelle on ne fait varier que un ou deux paramètres à la fois, est insatisfaisante :

- elle représente un travail important et qui peut être fastidieux sans pour autant garantir d'avoir exploré l'ensemble de l'espace des paramètres.
- elle est délicate à formaliser et par suite, difficilement reproductible.

La section suivante présente une approche différente, appelée *History Matching* et relevant des méthodes de quantification de l'incertitude. À différence des approches utilisant des algorithmes d'optimisation (voir section 5.1.2, l'approche *History Matching* vise non pas à exhiber un optimum, mais une région de l'ensemble des paramètres dans laquelle le modèle est compatible avec les observations. Un premier travail un peu prospectif avait été effectué pour calibrer une version AMIP avec des techniques de type *History Matching* pour un certain nombre des paramètres utilisés pour la calibration de ARPEGE-Climat 6.3 (Audouin, 2017). En utilisant des métriques proches de celles utilisées pour la calibration de la version AMIP de ARPEGE-Climat 6.3 la méthode avaient permis d'obtenir des résultats assez proches de ceux obtenus dans Roehrig et al. (2020) pour les paramètres concernés.

5.2 Vers une nouvelle stratégie. Outils statistiques pour une calibration « orientée processus »

La philosophie générale de la méthode de calibration utilisée pour ces travaux de thèse repose sur une approche de type *History Matching*. Le terme même d'*History Matching* provient de l'industrie pétrolière pour laquelle ces méthodes ont été initialement développées. Cette industrie a eu besoin de développer des modèles numériques, appelés modèles de réservoir, afin de prévoir la production future de champs pétrolifères. Ces modèles comportent un certain nombre de paramètres d'entrée pour lesquels on manque souvent d'information. Pour estimer les valeurs de ces paramètres, l'*History Matching* compare les sorties de ces modèles à des historiques de production (*History*) afin de ne retenir que les jeux de paramètres pour lesquels les sorties du modèle correspondent aux historiques de production (*Matching*). Vernon et al. (2010) s'inspire de cette approche pour calibrer un modèle numérique de formation des galaxies qui, à l'instar d'un modèle de climat, comporte un certain nombre de paramètres

dont les valeurs sont peu voire pas documentées. Le coût numérique d'un tel modèle interdit de réaliser un grand nombre de simulations pour régler ces paramètres incertains. Pour pallier cette difficulté, Vernon et al. (2010) propose l'utilisation d'émulateurs statistiques simulant pour un coût négligeable différentes métriques permettant d'évaluer l'adéquation du modèle avec la réalité observée. Vernon et al. (2010) propose une méthode itérative visant à réduire à chaque itération la taille de l'espace des paramètres libres du modèle, en ne conservant que les régions de cet espace compatibles avec les données observées et en tenant compte des différentes sources d'incertitude. Le parallélisme avec les problèmes rencontrés pour la calibration des modèles climatiques inspire un travail équivalent à Daniel Williamson qui applique des méthodes d'*History Matching* à un GCM (Williamson et al., 2013, 2015) et à un modèle de circulation générale océanique (Williamson et al., 2017). Cette approche propose un changement de paradigme pour la calibration des modèles de climat, en ne cherchant pas à exhiber un hypothétique « réglage optimum », mais à déterminer une région de l'espace des paramètres dans laquelle le modèle se rapproche de façon suffisante d'observations ou de simulations de référence, tout en tenant compte des différentes sources d'incertitude du problème (en particulier, les incertitudes sur les références utilisées pour calibrer le modèle).

Une collaboration entre la communauté française de la modélisation du climat (via le projet ANR HIGH-TUNE) et des chercheurs de l'Université d'Exeter (Royaume-Uni) a permis de développer une méthodologie adaptant l'approche développée par Williamson et al. (2013, 2017) au cadre SCM/LES. Un article a été publié pour présenter en détail cette méthodologie (Couvreur et al., 2020b). Cette section s'en inspire largement.

Motivations Les résultats des modèles peuvent être biaisés en raison d'incertitudes dans leurs équations physiques ou paramétrisations, de la spécification des conditions initiales, de la connaissance des facteurs de forçages externes, et des processus et rétroactions non pris en compte. On s'intéresse ici aux incertitudes liées à la paramétrisation des processus non explicitement résolus et en particulier, liées au réglage des paramètres libres mis en jeu dans ces paramétrisations. En particulier, l'approche que l'on se propose de suivre doit permettre de décrire de façon plus objective l'incertitude sur les paramètres libres, tout en tenant compte des incertitudes sur les différentes références qui servent à évaluer le modèle ainsi que les erreurs intrinsèques du modèle, interprétées comme une tolérance à l'erreur, suivant (Williamson et al., 2017, Couvreur et al., 2020b). Cette approche permet d'éviter les problèmes de surajustement, mais aussi de comprendre dans quelle mesure les déficiences d'un modèle relèvent d'une mauvaise calibration des paramétrisations ou de limites intrinsèques liées à la formulation même de ces paramétrisations.

L'approche utilisée dans ce travail est une approche « orientée processus ». Elle se situe donc dans la première étape du cadre général proposé par Hourdin et al. (2017) (voir 5.1.2). Elle utilise une comparaison SCM/LES et vise non pas à proposer un réglage optimal des paramètres libres mais à définir un espace dans lequel le modèle se comporte raisonnablement. C'est dans cet espace « raisonnable » que le réglage d'une version GCM du modèle pourra se faire. On réduit ainsi en partie le risque de compensation d'effets. Ce phénomène de compensation d'effets est très dépendant des métriques utilisées et est plus susceptible d'apparaître sur des métriques globales calculées à partir d'un GCM. Il n'est pas nul pour autant avec des métriques calculées sur des SCM ou des LES. La méthode de calibration classique, qui est basée sur une approche empirique dans laquelle on fait varier un ou deux paramètres à

la fois sur un cas d'étude donné, est insuffisante pour estimer cet espace « raisonnable ». En effet elle ne garantit pas d'explorer toutes les possibilités de réglage du modèle, entre autre les possibilités d'interaction des différentes paramétrisations. L'approche utilisée dans ce travail de thèse résout cette difficulté faisant un large appel à des modèles statistiques appelés émulateurs. Combinés à la méthode d'*History Matching*, l'utilisation de ces émulateurs permet d'explorer de façon quasi exhaustive l'espace des paramètres et de définir une gamme de réglages du modèle qui prend en compte les différentes sources d'incertitude qui interviennent dans le problème de calibration (incertitude sur les références, limites intrinsèques des paramétrisations et incertitudes liées à l'utilisation des émulateurs).

5.2.1 La comparaison SCM/LES

Pourquoi utiliser des SCMs ? La résolution horizontale des GCMs (et même de modèles à aire limitée comme le modèle AROME qui a une résolution horizontale de 1.3 km) fait que les paramétrisations physiques fonctionnent sur chaque colonne de l'atmosphère de façon indépendante des colonnes voisines. L'influence d'une colonne sur les colonnes voisines est traitée par le noyau dynamique du modèle via, en autres, les différents termes d'advection. C'est particulièrement vrai pour les modèles de climat, pour lesquels l'ordre de grandeur de la dimension horizontale d'une maille est de l'ordre de 100 km. Cette caractéristique, moyennant l'introduction d'un forçage des colonnes pour caractériser la grande échelle permet de tester les paramétrisations sur des simulations SCMs (voir 3.4.1) en s'affranchissant du couplage avec la grande échelle. Notons que par conséquent, une approche purement SCM ne permettra pas de traiter des potentielles rétroactions de la physique avec la grande échelle. Le comportement des modèles dans leur version SCM est toutefois souvent représentatif de leur comportement en version 3D (Hourdin et al., 2013, Neggers, 2015, Vignon et al., 2018, Gettelman et al., 2019).

Pourquoi utiliser des LES comme référence ? Calibrer un SCM nécessite de pouvoir l'évaluer et pose donc la question de la référence. Les observations ne sont pas toujours adaptées pour répondre à cette question. Même en passant outre les problèmes de données manquantes ou de mauvaise qualité, on dispose rarement de mesures de l'ensemble des variables d'état d'un modèle (dynamiques et thermodynamiques) à une fréquence spatiale et temporelle suffisante pour répondre aux exigences d'une démarche de calibration d'un modèle. Une difficulté supplémentaire pour comparer un SCM avec des observations est qu'il est très difficile de caractériser les conditions de forçage de grande échelle, ainsi que les conditions initiales.

Une façon de contourner ces difficultés est d'utiliser comme référence non pas des observations, mais des simulations LES (Ayotte et al., 1996, Randall et al., 1996). Les simulations LES fournissent une image tridimensionnelle (3D) complète des mouvements turbulents alors que les observations directes sont rares (Lenderink et al., 2004). Les simulations LES ne résolvent pas explicitement toute la physique et comportent donc des paramétrisations physiques (pour la turbulence et la microphysique par exemple). Il est donc nécessaire de les évaluer par des observations afin de caractériser leurs limites. Les LES peuvent alors être vues comme un intermédiaire qui permet de comparer de façon indirecte les observations et les SCMs. Les

LES sont devenues des outils essentiels (Hourdin et al., 2013, 2020a) pour le développement de paramétrisations physiques représentant les processus de couche limite (convection peu profonde, turbulence).

Le formalisme SCM/LES On rappelle qu'un modèle SCM peut être représenté par l'équation suivante (une version de l'équation 3.21 tenant compte des paramètres libres) :

$$\frac{\partial \mathbf{x}_c}{\partial t} = \sum_p \mathcal{P}_p(\mathbf{x}_c, \lambda_p) + \mathcal{F}_c(\mathbf{x}_c) \quad (5.2)$$

Pour pouvoir être comparées à une simulation d'un SCM, les LES sont réalisées sur un domaine équivalent à celui de la colonne extraite avec le même terme de forçage au bord que le SCM et avec exactement les mêmes conditions initiales. Ces conditions se traduisent pas les deux équations suivantes :

$$\frac{\partial \mathbf{y}}{\partial t} = \mathcal{L}(\mathbf{y}) + \mathcal{F}_c^*(\bar{\mathbf{y}}) \quad (5.3)$$

$$\mathbf{x}_c(t = 0) = \bar{\mathbf{y}}(t = 0) \quad (5.4)$$

où \mathbf{y} est le vecteur d'état de la LES, $\mathcal{L}(\mathbf{y})$ représente le modèle LES complet (dynamique et paramétrisations sous-maille), \mathcal{F}_c^* correspond a un champ 3D qui reproduit le même forçage que celui du modèle uni-colonne et enfin $\bar{\mathbf{y}}(t = 0)$ représente la moyenne sur l'horizontale du vecteur d'état de la LES à l'initialisation de la simulation. Ce formalisme permet de réaliser une comparaison directe d'une simulation SCM avec des LES, et garantit des conditions initiales et un forçage par la grande échelle identique.

La comparaison SCM/LES nécessite l'existence de cas pour lesquels les simulations LES ont été évaluées et couvrant un vaste ensemble de situations différentes. Sous l'égide du Global Atmospheric System Studies (GASS), de nombreux cas couvrant une variété de couches limites convectives ont été construits et sont désormais largement utilisés. Les exercices d'intercomparaison GABLS (Holtlag, 2006, Cuxart et al., 2006, Holtlag et al., 2012, Bosveld et al., 2014b) qui s'intéressent aux couches limites stables rentrent également dans ce cadre et les cas GABLS1 et surtout GABLS4 seront utilisés dans ces travaux pour calibrer et évaluer le modèle ARPEGE-Climat.

Nous présentons dans la suite du chapitre la méthodologie suivie pour calibrer le modèle ARPEGE-Climat en exploitant les possibilités offertes par les comparaisons SCM/LES.

5.2.2 *History Matching* et calibration d'un SCM

La méthode utilisée dans ce travail se base sur les travaux de Daniel Williamson et est le fruit d'une collaboration étroite entre des mathématiciens appliqués de l'université d'Exeter

et la communauté française de la simulation numérique du climat. Cette collaboration s'est placée dans le cadre d'un projet ANR (HighTune) et se poursuit encore à l'heure actuelle. Les outils développés ont pour objectif de permettre une comparaison entre des versions SCM de modèles de climat et des simulations LES qui servent de référence. La comparaison se fait sur la base de métriques représentant la performance du modèle. Cette méthodologie est présentée en détail dans Couvreur et al. (2020b), auquel les travaux réalisés pendant cette thèse ont contribué.

Cette section s'inspire fortement de cet article. On décrit ici les principales étapes de la méthode :

1. Sélection des métriques La première étape est d'identifier des cas d'étude qui permettent d'évaluer le comportement du modèle. Le nombre de cas disponibles pour lesquels on dispose de simulations LES peut être assez limité suivant les conditions météorologiques (couches limites convectives, couches limites stables) et les paramétrisations (turbulence, convection, microphysique, ...) que l'on souhaite évaluer et calibrer. Comme on se place dans le cadre d'une calibration orientée processus, le choix se porte préférentiellement sur des cas pour lesquels le nombre de paramétrisations actives est limité, du moins dans un premier temps. Pour chacun des cas retenus, il faut définir des métriques que l'on calculera à la fois sur la base des simulations SCM, et sur la base des simulations LES. De nombreux types de métriques peuvent être utilisés : les plus simples peuvent être des variables scalaires directement issues du modèle (par exemple la température potentielle à un instant donné et à une altitude donnée), ou bien des statistiques sur ces mêmes variables scalaires (minimum, maximum, moyenne, quantiles, ...). Des métriques de plus grandes dimensions, comme des grandeurs vectorielles (profils verticaux ou séries temporelles) sont également envisageables, via des analyses en composantes principales afin de réduire les dimensions du problème. Des développements sont en cours pour adapter les outils à ce type de métriques (Salter et al., 2019). Le choix des métriques relève de l'expertise du modélisateur et le résultat de la méthode est sensible à ce choix. En particulier, la rapidité de la méthode et sa capacité de discrimination va dépendre des métriques utilisées. Cet aspect est discuté dans Audouin et al. (2020).

Chaque métrique est évaluée sur un ensemble de simulations LES. La moyenne de chaque métrique sur l'ensemble des LES constituera la valeur de référence. L'incertitude sur cette référence est alors simplement la variance de la métrique sur ce même ensemble de LES. Couvreur et al. (2020b) montrent que l'on peut estimer l'incertitude sur les résultats soit avec un ensemble de simulations LES issues de modèles indépendants, soit avec des simulations LES issues d'un même modèle avec des tests de sensibilité à sa formulation (e.g., schéma numérique, paramétrisations).

2. Sélection des paramètres Cette étape peut se faire en amont du choix des métriques. Le nombre de paramètres libres dans un modèle de climat est relativement important (de l'ordre d'une centaine). Pour une paramétrisation donnée, le nombre de paramètres libres est de l'ordre d'une dizaine. Le choix des paramètres qui vont être utilisés pour un exercice de calibration donné est largement basé sur l'expertise des modélisateurs. Une analyse de

sensibilité (avec la méthode de criblage de Morris par exemple) peut être utilisée pour aider à cette sélection.

La sélection des paramètres doit s'accompagner de la définition pour chacun de ces paramètres, d'une plage dans laquelle on pourra faire varier ces paramètres. La définition de la plage initiale est en fait un compromis. Plus celle-ci sera réduite, plus rapide sera la convergence de la méthode. Mais d'un autre côté on conseille généralement de choisir une plage relativement étendue afin d'explorer de façon exhaustive le comportement du modèle.

3. Plan d'expérience Le plan d'expérience consiste en la construction de l'ensemble des points de l'espace des paramètres libres pour lesquels une simulation utilisant le modèle de climat sera effectivement réalisée. Pour des raisons de coût de calcul, la taille du plan d'expérience est relativement réduite. En pratique, on choisit un plan d'une taille au moins égale à dix fois le nombre de paramètres. De nombreuses méthodes existent pour construire des plans d'expériences. On pourra se référer à Franco (2008) pour une revue détaillée de ces méthodes. On utilise ici un plan en hypercube latin pour ses bonnes propriétés de projection : si un des paramètres sélectionnés se révèle peu influent sur le résultat, la projection de l'hypercube latin sur l'espace privé de ce paramètre conserve de bonnes propriétés de remplissage de l'espace. Un plan de n points pour d paramètres est construit en subdivisant pour chaque dimension l'axe en n segments de même longueur. On obtient ainsi un maillage de n^d cellules de même taille. On choisit alors n cellules parmi les n^d possibles de telle manière que les cellules des marges de dimension 1 soient représentées une fois chacune et on tire un point au hasard dans chaque cellule sélectionnée. La figure 5.1 donne un exemple d'hypercube latin. On utilise ici de plus une contrainte de maximisation de la distance minimale entre deux points de l'espace afin d'optimiser encore un peu plus la qualité de remplissage de l'hypercube.

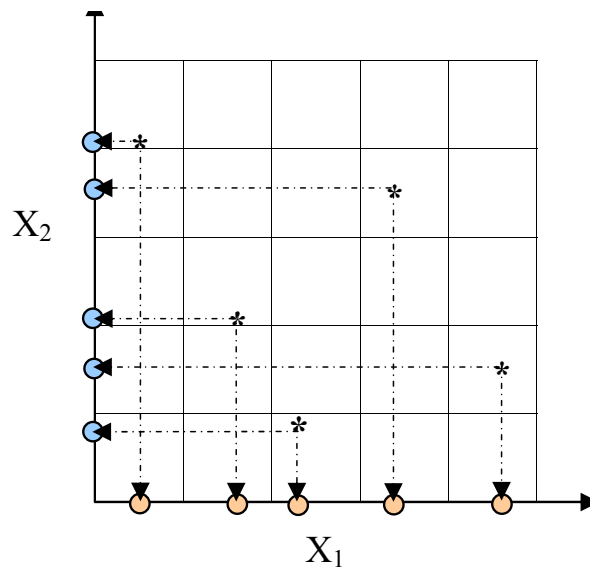


FIGURE 5.1 – Exemple d'hypercube latin de $n = 5$ points en dimension $d = 2$, extrait de (Franco, 2008)

Pour chacun des points du plan d'expérience, une simulation est effectuée avec le SCM et

l'ensemble des métriques sélectionnées est calculé. Ces données serviront d'échantillon d'apprentissage pour construire un modèle statistique qui émule le comportement du modèle quand on fait varier les différents paramètres libres.

4. Construction des émulateurs Explorer un espace de grande dimension nécessite potentiellement un grand nombre de points (souvent de l'ordre du million). Il n'est pas envisageable en général de réaliser un nombre aussi grand de simulations, même avec des versions SCM. On contourne alors cette difficulté en utilisant des méthodes d'apprentissage statistique : pour chaque métrique, on va construire un émulateur qui va simuler la réponse du modèle de climat. Un émulateur est un modèle statistique, construit sur la base d'un échantillon d'apprentissage (dans notre cas, le résultat de l'étape précédente), qui va faire une prévision de la réponse du SCM en un point de l'espace des paramètres pour lequel celle-ci n'est pas connue. De nombreuses méthodes sont désormais disponibles pour construire un tel émulateur (modèles de régression polynomiale, polynôme du chaos, par exemple). Le choix s'est ici porté sur une méthode d'interpolation spatiale basée sur l'utilisation des processus gaussiens (Rasmussen and Williams, 2005), classiquement utilisés pour la calibration des grands codes numériques. Cette méthode, qui s'apparente au krigeage, a pour avantage de fournir une réponse probabiliste et accompagne la prévision qu'elle réalise de l'incertitude associée à cette prévision. Cette incertitude est d'autant plus grande que la prévision se fait sur un point éloigné des points de l'échantillon d'apprentissage. Pour une métrique donnée f , et un jeu de paramètres donné λ , on notera $E[f(\lambda)]$ l'espérance du processus gaussien au point λ , et par $\text{Var}[f(\lambda)]$ sa variance. C'est cette variance qui quantifie l'incertitude sur la prévision effectuée.

Les émulateurs sont systématiquement testés par une méthode de validation croisée de type *Leave One Out* à partir de l'échantillon d'apprentissage, afin de vérifier la qualité de l'émulateur. Quand on parle de qualité de l'émulateur, on n'entend pas ici uniquement de la précision de l'émulateur, mais également de sa capacité à estimer correctement l'incertitude sur la prévision qu'il réalise.

5. History Matching L'approche de type *History Matching* consiste à retenir comme jeux de paramètres acceptables ceux qui conduisent à un modèle compatible avec une référence. Pour cela on explore de façon systématique l'espace des paramètres en comparant pour chaque métrique f la réponse du SCM, simulée par l'émulateur correspondant, avec la référence, calculée comme la moyenne des LES tout en tenant compte des différentes sources d'incertitude. Celles-ci comprennent, outre celle sur la réponse de l'émulateur, l'incertitude sur la référence (soit l'incertitude sur le résultat des LES) ainsi qu'un terme appelé erreur structurelle du modèle de climat et qui traduit le fait que ce modèle n'est pas parfait. En pratique, l'incertitude sur la référence est modélisée comme une grandeur aléatoire de moyenne nulle et de variance $\sigma_{r,f}^2$, cette dernière étant calculée comme la variance de l'ensemble des LES pour la métrique f . La question de l'erreur structurelle est plus délicate. (Williamson et al., 2017) propose de la considérer comme une tolérance à l'erreur, *i.e.* comme une différence entre les résultats du modèle SCM et la référence qui serait jugée acceptable par le modélisateur. On la modélise alors comme une grandeur aléatoire de moyenne nulle (sauf si un biais est mis en évidence) et de variance $\sigma_{d,f}^2$, avec $\sigma_{d,f}$ pris comme étant égal à la tolérance à l'erreur.

Suivant Williamson et al. (2013), on introduit alors la notion d'*implausibility*, que l'on pourrait traduire en bon français par invraisemblance. Pour éviter les confusions avec le terme de vraisemblance qui a un sens précis en modélisation statistique, on gardera le terme anglais dans la suite du texte. Formellement, pour une métrique donnée f , et un jeu de paramètres donné λ , l'*implausibility* $I_f(\lambda)$ est une distance entre la valeur de la métrique fournie par l'émulateur (*i.e* $E[f(\lambda)]$) et la valeur de référence de la métrique r_f , normalisée par les différentes sources d'incertitude, soit :

$$I_f(\lambda) = \frac{|r_f - E[f(\lambda)]|}{\sqrt{\sigma_{r,f}^2 + \sigma_{d,f}^2 + \text{Var}[f(\lambda)]}} \quad (5.5)$$

On considère alors qu'un jeu de paramètres λ conduit à un modèle compatible avec la référence (relativement à la métrique f) si l'*implausibility* $I_f(\lambda)$ est inférieure à un seuil donné T . On fixe ce seuil à 3, suivant la règle de Pukelsheim qui stipule que 95% de la densité de probabilité pour toute distribution unimodale se situe à moins de 3 écarts-types de la moyenne (Pukelsheim, 1994). La valeur de T est volontairement très conservatrice et rend très improbable d'éliminer des points par erreur. L'ensemble des point λ pour lesquels $I_f(\lambda) < T$ constitue la région de l'espace des paramètres dans laquelle le modèle est compatible avec la référence, compte-tenu des différentes sources d'incertitude et qu'on appellera dans la suite l'espace NROY (pour Not-Ruled-Out-Yet Space) selon la formulation proposée dans Williamson et al. (2013). En pratique, l'espace NROY est construit à partir d'un hypercube latin (d'un million de points dans les applications réalisées au cours de ce travail) échantillonnant l'espace des paramètres complets et en calculant l'*implausibility* de chacun de ses points. L'espace restant à l'issue de cette première itération, noté NROY^1 , est donc un sous-ensemble de l'hypercube latin initial.

Dans le cas où plusieurs métriques f_1, f_2, \dots, f_n sont utilisées, la démarche est la même que pour une seule métrique, mais le calcul de l'*implausibility* est différent : on évalue dans un premier temps l'*implausibility* indépendamment pour chacune des métrique, et on prend alors la plus grande des $I_{f_i}(\lambda)$, soit

$$I(\lambda) = \max_{i=1..n} [I_{f_i}(\lambda)] \quad (5.6)$$

Autrement dit l'espace NROY est l'intersection des différents espace NROY évalués pour chaque métrique :

$$\text{NROY} = \bigcap_{i=1..n} \text{NROY}_{f_i}^1 \quad (5.7)$$

6. Iterative Refocusing La formule 5.5 montre que l'*implausibility* est d'autant plus grande que l'erreur absolue (au numérateur) est grande et que les termes d'incertitude (au dénominateur) sont petits. Cela signifie qu'un jeu de paramètres λ pourra être conservé dans le NROY space si $E[f(\lambda)]$ est proche de la valeur de référence ou bien si la somme des incerti-

tudes est grande. Parmi ces dernières, si l'incertitude de l'émulateur $\text{Var}[f(\boldsymbol{\lambda})]$ domine l'erreur structurelle $\sigma_{d,f}^2$ et l'incertitude sur la référence $\sigma_{r,f}^2$, alors des points qui devraient être exclus sont conservés. Afin de réduire l'incertitude des émulateurs, on utilise une méthode appelée *Iterative Refocusing* : on réalise de nouvelles itérations de la méthode (étapes 3 à 5), mais en utilisant comme plan d'expérience initial non pas un hypercube latin échantillonnant tout l'espace des paramètres, mais un plan qui échantillonne seulement l'espace NROY précédent. Ce sous-ensemble, de même taille que le plan d'expérience utilisé lors de la première itération, échantillonne un espace plus petit. La densité plus importante des points permet d'améliorer la précision des émulateurs. Lorsque l'incertitude des émulateurs est suffisamment faible relativement aux autres termes d'incertitude, il ne devient plus possible de réduire l'espace des paramètres et le dernier espace NROY construit constitue la région de l'espace des paramètres compatible avec les références, compte tenu des incertitudes sur les références elle-mêmes et de l'erreur structurelle du modèle. La question de la convergence de la méthode est délicate et est discutée dans Audouin et al. (2020, voir chapitre 6).

Théoriquement, on pourrait appliquer la méthode d'*History Matching* décrite ci-dessus sans utiliser les émulateurs. Le coût numérique d'une simulation SCM permet de réaliser plusieurs dizaines voire centaines de milliers de simulations en un temps raisonnable, ce qui peut être suffisant pour construire l'espace NROY. On n'a alors plus besoin de réaliser plusieurs itérations puisque celles-ci servent à améliorer la précision des émulateurs et de réduire l'incertitude sur leurs résultats. Un tel exercice a d'ailleurs été réalisé afin de vérifier que les résultats obtenus en utilisant les émulateurs étaient similaires à ceux obtenus en utilisant directement des simulations SCM (voir Audouin et al., 2020).

5.3 Conclusions

La calibration de la physique des modèles de climat, une étape importante de leur développement, pose un certain nombre de problèmes auquel les stratégies couramment utilisées peinent à répondre : Comment explorer de façon suffisante l'espace des paramètres libres du modèle ? Comment éviter des phénomènes de compensation d'erreur ? Comment éviter l'écueil du sur-ajustement des paramètres libres ? Comment tenir compte des incertitudes dans les références ? Comment rendre l'étape de calibration reproductible et lever au moins en partie la subjectivité du modélisateur ? Les outils présentés dans ce chapitre se proposent de répondre au moins en partie à ces questions. L'utilisation des émulateurs permet une exploration quasi-exhaustive de l'espace des paramètres. L'introduction des notions d'*implausibility* et d'espace NROY permet de tenir compte des incertitudes et d'éviter le sur-ajustement. Enfin, une fois les métriques et les incertitudes qui y sont associées définies, la méthode proposée est parfaitement reproductible et une partie de la subjectivité du modélisateur est levée. Cette subjectivité n'est toutefois pas entièrement éliminée. Il reste en effet une certaine liberté au modélisateur à travers le choix des métriques d'une part, lui permettant ainsi de mettre l'accent sur tel ou tel aspect des phénomènes simulés et à travers la valeur de l'erreur intrinsèque du modèle d'autre part, qui peut être vue comme une tolérance à l'erreur acceptable (voir Williamson et al., 2017, sur ce sujet). On peut également mentionner le seuil d'*implausibility* au-delà duquel un réglage est éliminé. Ce seuil, initialement fixé à 3, peut ainsi être revu à la baisse par le modélisateur lorsque les vagues successives ne permettent plus de réduire la

taille de l'espace NROY.

La section 5.2 a présenté la méthode dans le cadre de la calibration d'un SCM. Le chapitre 6, à travers un article (Audouin et al., 2020) accepté pour publication dans *Journal of Advances in Modeling Earth Systems* en présentera une application pour la paramétrisation de la turbulence sur un cas de couche limite stable très stratifiée (le cas GABLS4). Hourdin et al. (2020b) montre comment cette approche permet d'articuler la calibration d'une version SCM d'un modèle de climat avec celle d'une version GCM du même modèle et comment cette approche peut permettre d'accélérer le processus de développement d'un modèle de climat. La méthode peut également s'appliquer à la calibration offline d'une paramétrisation donnée (voir Villefranque et al., 2021, pour une application à la calibration d'un schéma de transfert radiatif).

Modéliser la couche limite très stable de GABLS4 avec la physique de ARPEGE-Climat : sensibilité paramétrique ou limites intrinsèques du modèle ?

Sommaire

6.1	Le problème de la représentation des CLS dans les modèles numériques	90
6.2	Résumé	91
6.3	L'article	93
6.4	Compléments	131
6.4.1	Améliorer la visualisation de l'espace NROY	131
6.4.2	Impact des nouveaux réglages sur le cycle diurne de GABLS4	131
6.5	Conclusions	137

L'évaluation du modèle CNRM-CM6 sur l'Antarctique met en évidence un biais froid sur l'ensemble du continent, que ce soit sur la période estivale ou hivernale. On montre dans le chapitre 4 qu'une partie de ce biais froid (au moins pendant l'été austral) peut être relié à un albédo de surface surestimé sur la calotte glaciaire antarctique. Une correction ad-hoc de la formulation du calcul de l'albédo de la neige dans le modèle permet de se rapprocher de l'albédo de référence (CERES). Si l'albédo CERES reste à discuter, la correction permet de réduire considérablement le biais froid sur le plateau pendant l'été. Si la température moyenne est satisfaisante avec cette correction de l'albédo, le cycle diurne est dégradé (voir le focus sur Dôme C, section 4.4), en particulier son amplitude est sous-estimée. Le biais sur la température maximale à 2 m est ainsi réduit jusqu'à le rendre négligeable sur les mois de décembre et de janvier (passant de -5 à 0 K) mais le signe du biais sur la température minimale est inversé : on passe ainsi d'un biais froid de -5 K à un biais chaud $+2$ K). Le refroidissement radiatif contrôlant la formation des vents catabatiques, ce biais chaud des températures nocturnes dégrade la performance du modèle en ce qui concerne la représentation des vents de surface sur le plateau, en particulier près des côtes, là où l'écoulement de gravitation est le plus fort.

Les biais mis en évidence dans ARPEGE-Climat avec correction de l'albédo pendant l'été sont courants dans les modèles numériques sur les couches limites très stratifiées. Ils sont

souvent reliés à la façon dont sont représentés les flux turbulents et en particulier à des formulations de la turbulence qui surestiment de façon artificielle le mélange turbulent (voir la section 6.1). Le cas GABLS4 qui présente un cycle diurne complet d'une couche limite sur le plateau antarctique permet d'étudier en détail le comportement du modèle sur un cas typique de CLS nocturne très stratifié et d'évaluer dans quelle mesure ses mauvaises performances peuvent être reliées au réglage de la paramétrisation de la turbulence. Ce travail est l'occasion de mettre en œuvre les outils de calibration statistique développés dans le cadre du projet ANR HighTune et présentés dans le chapitre 5. Une étude préliminaire à ce travail (non présentée ici) a montré que CNRM-CM6 représentait correctement la phase diurne de GABLS4 avec le développement d'une couche limite convective de 300 m d'épaisseur. D'autre part, Couvreur et al. (2020a) montrent que les simulations LES réalisées dans le cadre du projet GABLS4 avec une résolution initialement prévue de 2 m ne convergent pas sur la phase nocturne (stable) du cycle diurne. Des simulations à plus haute résolution (1 m) se focalisant sur la phase nocturne ont été réalisées et ont donné des résultats satisfaisants (avec notamment une bien meilleure convergence). Ces dernières servent de référence pour le travail présenté dans ce chapitre. Ces deux points ont conduit à réaliser l'analyse de sensibilité et la calibration de ARPEGE-Climat uniquement sur la deuxième partie du cycle diurne de GABLS4 (*i.e.* la phase très stratifiée). Un article présentant ces travaux a été accepté dans *Journal of Advances in Modeling Earth Systems*. Les principaux objectifs de cette étude sont, outre évaluer la capacité du modèle à représenter correctement un cas de couche limite très stratifiée, de documenter sa sensibilité à différents paramètres du schéma de turbulence et de se construire un guide d'utilisation de l'outil statistique discuté dans le chapitre 5 pour la calibration d'un modèle de climat. Ce chapitre reprend cet article (section 6.3) ainsi que quelques résultats complémentaires (6.4), avec en particulier une évaluation de la calibration proposée dans l'article sur le cycle diurne complet de GABLS4.

6.1 Le problème de la représentation des CLS dans les modèles numériques

Les couches limites stables restent délicates à représenter dans les modèles numériques. Fréville et al. (2014) montre par exemple que la réanalyse ERAI présente un biais important sur le plateau antarctique. Ce biais a pu être relié à la profondeur de neige thermiquement active dans le schéma de surface ainsi qu'à une surestimation des flux turbulents en conditions stables (Dutra et al., 2015).

Pourquoi une turbulence *augmentée* ? Les schémas de turbulence utilisés maintiennent un mélange artificiellement trop élevé qui n'est justifié ni par des observations (McCabe and Brown, 2007, Mauritsen and Svensson, 2007, Brown et al., 2008), ni par des résultats de LES (Svensson and Holtslag, 2009). Les raisons d'être de cette turbulence augmentée sont diverses :

- Tenir compte par ce moyen de phénomènes qui ne sont pas explicitement représentés dans les modèles (hétérogénéité du relief, ondes de gravité, variabilité de méso-échelle ...) mais qui vont contribuer au mélange vertical.

- Tenir compte de la non-linéarité de nombreux processus ainsi que des hétérogénéités de terrain Beljaars and Holtslag (1991)
- Éviter le découplage entre la surface et l'atmosphère qui peuvent générer des runaway cooling (Louis, 1979, Steeneveld et al., 2006)
- Prendre en compte le caractère non linéaire du mélange turbulent dans des conditions de couches limites stables

Le risque du runaway cooling Le refroidissement de la surface accroît la stabilisation des premiers niveaux de la CL. Cette stabilisation, si elle s'accompagne d'une forte diminution du mélange turbulent, va entraîner une diminution du flux de chaleur descendant et donc un refroidissement de la surface, formant ainsi une rétroaction positive (e.g. Mauritsen, 2012, Van de Wiel et al., 2012).

Conséquences d'une turbulence augmentée dans les modèles numériques Cette augmentation de la diffusion turbulente tend à détruire les couches limites stables (épaisseur de la CLA trop importante, jet trop faible et situé trop haut, rotation du vent sous-estimée,...) et pose également des problèmes en troposphère libre (en générale assez stratifiée) où une diffusion turbulente trop forte va entraîner un jet peu marqué et des inversions trop faibles. Enfin l'inversion au sommet de la couche limite peut elle aussi être trop faible et participer de la sous-estimation de la couverture en strato-cumulus dans les modèles globaux. Dans beaucoup de modèles opérationnels (ECMWF, UKMO), cette augmentation artificielle de la turbulence continue cependant d'être utilisée, plus par pragmatisme que par une justification rigoureuse (McCabe and Brown, 2007). Elle permet entre autre d'éviter des biais froids fréquemment rencontrés dans les couches limites stables et améliore la représentation des dépressions synoptiques aux moyennes latitudes. Même si ces biais froids en configuration stable peuvent avoir de multiples causes, une diffusion turbulente forte réduira le biais froid en "transportant" le refroidissement radiatif sur une plus grande épaisseur. La représentation de la turbulence des conditions stables et stratifiées dans les modèles numérique (de prévision ou de climat) reste un problème récurrent.

6.2 Résumé

De nombreux modèles (dont ARPEGE-Climat) utilisent des paramétrisations de la turbulence qui surestiment de façon artificielle les flux turbulents (voir 6.1). Cette surestimation qui permet de palier la mauvaise représentation d'un certain nombre de phénomènes (hétérogénéités de terrain, ondes de gravité, ...) a tendance à dégrader les couches limites très stratifiées. Dans le cadre du projet GEWEX Atmospheric Boundary Layer Study (GABLS), quatre exercices d'intercomparaison (voir section 3.5.2.1) de modèles numériques couvrant un ensemble de régimes de couches limites stables ont été menés, allant de couches limites faiblement stratifiées (GABLS1) à des couches limites très stratifiées (GABLS4). Ce dernier exercice se base sur une campagne de mesures réalisée durant le mois de décembre 2009 sur la base franco-italienne de Concordia à Dôme C et couvre un cycle diurne complet. Pendant la journée, une couche limite convective de 400 m se développe. Quand le soleil est suffisamment

bas sur l'horizon, le bilan radiatif de la surface devient négatif et une stratification de l'atmosphère se met en place, avec un gradient de température potentielle très marqué (+15 K sur 40 m). Représenter de tels gradients de température est un défi pour un GCM du fait de sa résolution verticale. Ce cas d'étude sert ici à évaluer la capacité d'ARPEGE-Climat à représenter correctement une couche limite très stratifiée et à utiliser des méthodes statistiques issues de la communauté statistique de quantification des incertitudes pour estimer dans quelle mesure les biais mis en évidence sur le cas GABLS4 relèvent d'un choix de réglage (ou calibration) du modèle ou d'une limite intrinsèque de celui-ci. L'étape de calibration d'un modèle de climat est cruciale dans le développement d'un modèle de climat (Hourdin et al., 2017). Le nombre de paramètres à régler et les nombreuses interactions possibles entre ces paramètres rendent difficile une exploration exhaustive du comportement d'un modèle sur l'ensemble des combinaisons possibles. L'approche utilisée dans ce travail (*History Matching with Iterative Refocusing*) suit celle proposée par Williamson et al. (2013) et s'appuie sur l'utilisation d'émulateurs statistiques qui simulent le comportement du modèle de climat quand on fait varier les valeurs de paramètres, sans pour autant avoir besoin de réaliser de nouvelles simulations. Cette approche ne cherche pas à optimiser la calibration du modèle sur la base de quelques métriques mais à déterminer des régions de l'espace des paramètres dans lequel le modèle se comporte de façon satisfaisante, en tenant compte des différentes sources d'incertitude.

Après une présentation du modèle et des configurations utilisées, la méthodologie et les outils statistiques sont brièvement décrits. Le cas GABLS4 ainsi que les simulations LES qui ont servi de référence pour les exercices de calibration sont également présentés. Le modèle est alors évalué dans différentes configurations de forçage en surface (température ou flux de chaleur sensible) ainsi que pour différentes résolutions verticales : la version standard et une version à 2 m de résolution (respectivement LR et HR dans la suite). Ces différentes configurations mettent en évidence les mêmes défauts du modèle : une couche limite nocturne trop épaisse, ce qui se traduit par un refroidissement nocturne trop faible et un jet de basses couches trop peu marqué et trop haut. La méthode de calibration statistique est appliquée aux différentes configurations de résolution verticale et de forçage décrites plus haut. Un jeu de quatre métriques simples contrôlant le profil vertical de température potentielle au plus fort du refroidissement nocturne ainsi que le profil vertical du vent lorsque le jet est le plus marqué est utilisé. L'exercice montre qu'il est possible de trouver un réglage satisfaisant du modèle dans les différentes configurations. Les combinaisons de paramètres qui sont retenues sont celles qui tendent à diminuer le mélange turbulent. On note en particulier que la borne inférieure à la longueur de mélange qui a été introduite dans la paramétrisation de la turbulence est un paramètre clé. Un autre paramètre, la borne inférieure pour le mélange turbulent, est également crucial pour le réglage du modèle dans sa version LR mais n'a cependant aucun rôle dans la version HR, son action sur le mélange dépendant de la résolution verticale.

Une dernière partie apporte un retour d'expérience sur les aspects les plus sensibles de la méthode. Est ainsi d'abord discuté le choix des métriques et de leur poids relatif dans les résultats. Dans chacune des configurations testées, la méthode de calibration statistique converge après une dizaine d'itérations. Les questions de convergence et d'arrêt des itérations sont discutées, notamment en lien avec les différentes sources d'incertitude qui interviennent (incertitude sur les références et sur les émulateurs). Enfin l'article conclut par une validation de la méthode réalisée en effectuant un exercice similaire mais utilisant uniquement des simulations unicolonne de CNRM-CM6, c'est-à-dire en se passant d'émulateurs. Les résultats obtenus dans les deux expériences sont qualitativement très proches.

6.3 L'article

Erratum Dans la Table 1 de l'article, les valeurs du paramètre **CE** pour ARPEGE-Climat 6.3 et Cuxart et al. (2000) sont erronées : il faut lire 1.18 et 1.43 respectivement.



Modeling the GABLS4 strongly-stable boundary layer with a GCM turbulence parameterization: parametric sensitivity or intrinsic limits?

O. Audouin¹, R. Roehrig¹, F. Couvreux¹, and D. Williamson^{2,3}

¹CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

²Exeter University, Exeter, United Kingdom

³The Alan Turing Institute, 96 Euston Road, London, United Kingdom

Key Points:

- The ARPEGE-Climat 6.3 performance in modeling a stable boundary layer is assessed in a single-column model/large-eddy simulation comparison
- The use of history matching with iterative refocusing indicates turbulence parameterization deficiencies due to poor calibration
- Validation of the statistical framework and guidelines for its use in parameterization development and calibration are discussed

Corresponding author: Olivier Audouin, olivier.audouin@meteo.fr

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1029/2020MS002269](https://doi.org/10.1029/2020MS002269).

This article is protected by copyright. All rights reserved.

Abstract

The representation of stable boundary layers (SBLs) still challenges turbulence parameterizations implemented in current weather or climate models. The present work assesses whether these model deficiencies reflect calibration choices or intrinsic limits in currently-used turbulence parameterization formulations and implementations. This question is addressed for the CNRM atmospheric model ARPEGE-Climat 6.3 in a single-column model/large-eddy simulation (SCM/LES) comparison framework, using the history matching with iterative refocusing statistical approach. The GABLS4 case, which samples a nocturnal strong SBL observed at Dome C, Antarctic Plateau, is used. The standard calibration of the ARPEGE-Climat 6.3 turbulence parameterization leads to a too deep SBL, a too high low-level jet and misses the nocturnal wind rotation. This behavior is found for low and high vertical resolution model configurations. The statistical tool then proves that these model deficiencies reflect a poor parameterization calibration rather than intrinsic limits of the parameterization formulation itself. In particular, the role of two lower bounds that were heuristically introduced during the parameterization implementation to increase mixing in the free troposphere and to avoid runaway cooling in snow- or ice-covered region is emphasized. The statistical tool identifies the space of the parameterization free parameters compatible with the LES reference, accounting for the various sources of uncertainty. This space is non-empty, thus proving that the ARPEGE-Climat 6.3 turbulence parameterization contains the required physics to capture the GABLS4 SBL. The SCM framework is also used to validate the statistical framework and a few guidelines for its use in parameterization development and calibration are discussed.

Plain Language Summary

During the night or in snow- or ice-covered region, a stable atmospheric boundary layer (SBL) often develops. Their representation still challenges turbulence parameterizations implemented in numerical weather or climate models. The present work assesses whether the ARPEGE-Climat atmospheric model deficiencies reflect calibration choices or intrinsic limits in its turbulence parameterization using a statistical approach from the Uncertainty Quantification community. A single-column version of the model is evaluated on the GABLS4 case, a nocturnal strong SBL observed at Dome C, Antarctic plateau, and compared to high-resolution simulations. The standard calibration of the ARPEGE-Climat 6.3 turbulence parameterization leads to a too deep SBL and an incorrect wind pattern and so for different vertical resolutions. The statistical tool proves that these model deficiencies are rectified with proper calibration of the turbulence parameterization. In particular, it is shown that two lower bounds, introduced to increase turbulent mixing, are key to capture the GABLS4 SBL. Finally, the potential and relevance of the Uncertainty Quantification approach applied to the single-column model/large-eddy simulation comparison framework for the calibration of climate models are highlighted and few guidelines for its use are proposed.

1 Introduction

In the atmospheric boundary layer, stably stratified conditions generally develop above a surface colder than the overlying air. These Stable Boundary Layers (SBLs) often result from the advection of warm air over a cold surface or from the cooling of the surface. They are frequently observed over ice or snow surfaces (e.g., polar regions, high-latitude continental regions in wintertime) and over land during nighttime. Their development and intensity (e.g., the vertical stratification) are also strongly modulated by the atmospheric synoptic conditions. Nights with weak or no cloud cover at all also favor SBL occurrence, as such conditions enhance the radiative cooling of the surface (Mahrt, 1998). In contrast, the occurrence of strong near-surface wind reduces the SBL strat-

64 ification, or even inhibits their development, through the maintenance of significant me-
65 chanical mixing (e.g., Van de Wiel et al., 2012).

66 SBL can be classified according to the intensity of their stratification (at first or-
67 der the vertical gradient of potential temperature), ranging from weak SBLs in which
68 the turbulence remains significant, to strong SBLs, in which the turbulence become in-
69 termittent or even disappears (e.g., Mahrt, 1998; Acevedo et al., 2016). In the latter con-
70 ditions, a mechanical decoupling between the atmosphere and the surface can occur (Derbyshire,
71 1999): the temperature inversion close to the surface becomes driven by radiation and
72 soil diffusion and the surface turbulent heat flux cannot sustain the surface energy de-
73 mand enhanced by a strong net surface radiative cooling (e.g., Van de Wiel et al., 2012).
74 Such strong SBLs mostly occur under clear-sky and weak wind conditions, with a strong
75 increase of the near-surface temperature inversion below a critical wind speed (e.g., Vi-
76 gnon, van de Wiel, et al., 2017) .

77 The representation of SBLs in operational climate and weather models is a chal-
78 lenge (e.g., Holtslag et al., 2013): the turbulence is particularly weak and sometimes in-
79 termittent (e.g., Mauritsen & Svensson, 2007), and interacts with other small-scale pro-
80 cesses (e.g., gravity waves – Steeneveld et al., 2008; Tsiringakis et al., 2017). Under the
81 umbrella of the Global Energy and Water Cycle Exchanges (GEWEX) project, the GEWEX
82 Atmospheric Boundary Layer Study (GABLS) has initiated four model intercompari-
83 son projects (Cuxart et al., 2006; Svensson et al., 2011; Bosveld et al., 2014; Bazile
84 et al., 2015) to evaluate and improve the SBL representation in weather and climate mod-
85 els. So far, the GABLS intercomparison exercises revealed:

- 86 1. Large-eddy simulations (LES) are able to consistently capture the main proper-
87 ties of stable boundary layers, at least when their resolution is below a few me-
88 ters for weak to moderate SBL (e.g., Beare et al., 2006) or 1 m for strong SBL (Couvreur,
89 Bazile, et al., 2020). As a result, such LES provide relevant process-level informa-
90 tion to evaluate turbulence parameterizations in an LES/Single-Column Model
91 (SCM) comparison framework (e.g., Randall et al., 1996, 2003)
- 92 2. State-of-the-art turbulence parameterizations, such as those with a 1.5-order tur-
93 bulance closure, are able to reasonably capture the physics of SBLs for a wide range
94 of forcing (e.g., Cuxart et al., 2006; Baas et al., 2018; Vignon, Hourdin, et al., 2017)
- 95 3. Nevertheless, current weather and climate models still simulate SBLs that are too
96 deep, with surface drag that is too strong, low-level jets that are too weak and too
97 high and wind veering with height that is too weak (e.g., Cuxart et al., 2006; Holt-
98 slag et al., 2013).

99 The apparent contradiction between the two last conclusions results from the cal-
100 ibration of weather or climate models which, so far, has required an increased turbulent
101 mixing to reduce the activity of synoptic systems, and thereby improve operational scores
102 (e.g., Sandu et al., 2013), or to prevent runaway surface cooling through long-term me-
103 chanical decoupling with the atmosphere (Derbyshire, 1999). Such a calibration prob-
104 ably reflects the lack of mixing due to processes that are currently not accounted for in
105 weather and climate models (e.g., surface heterogeneities, internal gravity waves, meso-
106 scale variability, impact of subgrid orography).

107 Recently Vignon, Hourdin, et al. (2017), Vignon et al. (2018) and Hourdin, Rio,
108 et al. (2020) showed that it is possible to achieve a reasonable representation of SBLs
109 in a climate model (LMDZ), while maintaining reasonable large-scale performance. Start-
110 ing with an SCM framework built on the very stable boundary layer of GABLS4 (Bazile
111 et al., 2015), Vignon, Hourdin, et al. (2017) underline the importance of (i) the coupling
112 with the surface (snow albedo and thermal inertia) and (ii) the turbulent mixing thresh-
113 olds usually used in current operational turbulence parameterizations (e.g., for the mix-
114 ing length or in stability functions). More specifically, the appropriate calibration of sur-

115 face (snow) properties and the removal of those thresholds in the turbulence parameter-
116 ization allows the LMDZ SCM to capture well the strong temperature gradient close to
117 the surface (in the first 20 meters), observed at Dome C, Antarctica, during an austral
118 summer night. Vignon et al. (2018) further compare 3D LMDZ simulations with local
119 observations collected at Dome C. On the one hand, the SCM improved results are con-
120 sistentlly reported in this 3D configuration, stressing the relevance of the SCM framework
121 for developing and calibrating parameterizations. On the other hand, the new version
122 of the LMDZ model adequately reproduces the annual cycle of Dome C, the two main
123 SBL regimes discussed above, and preserves satisfying large-scale skills. Finally, the ver-
124 tical resolution in the lower part of the boundary layer is also shown to be critical for
125 capturing SBLs that cover only a few tens of meters (see also Steeneveld et al., 2006).

126 Following Vignon, Hourdin, et al. (2017) and Vignon et al. (2018), the general ob-
127 jective of the present work is to document the performance of the CNRM climate atmo-
128 spheric model, namely ARPEGE-Climat 6.3 (Roehrig et al., 2020), to represent SBLs.
129 The focus is here on its turbulence parameterization, which is based on the work of Cuxart
130 et al. (2000). The parameterization of a given process seeks to represent its effects on
131 the large-scale (or resolved) state of the model. It is based on a set of physical theories
132 or empirical relationships to numerically describe the subgrid-scale processes and their
133 effects. Parameterizations introduce a number of constants, called free parameters in the
134 following, which are often difficult to constrain with observations or other references. A
135 parameterization can thus be seen as a function of the model state variables and of these
136 free parameters. Their calibration, or “tuning”, is a critical step in model development
137 for weather or climate applications (e.g., Hourdin et al., 2017). In the present paper, we
138 therefore propose to address the following specific question: Is it possible to calibrate the
139 ARPEGE-Climat turbulence parameterization to achieve a satisfying representation of
140 SBLs, especially those with a strong thermal stratification? In other words, does the ARPEGE-
141 Climat turbulence parameterization contain the required physics to represent appropri-
142 ately strong SBLs?

143 Cuxart et al. (2006) shows that the current turbulence parameterization of ARPEGE-
144 Climat 6.3 is able to capture the main properties of the moderate SBL studied during
145 the first GABLS exercise. We seek to extend this result to the strongly-stratified SBL
146 of the GABLS4 nocturnal phase. We rely on SCM simulations, which have been shown
147 relevant for 3D model configuration (Hourdin et al., 2013; Neggers, 2015; Vignon et al.,
148 2018; Gettelman et al., 2019). We also make use of GABLS4 LES as references, as they
149 have been shown to capture well the properties of the GABLS4 nocturnal phase (Couvreur,
150 Bazile, et al., 2020). Following Couvreur, Hourdin, et al. (2020), we use statistical tools
151 developed in the Uncertainty Quantification community, in particular the history match-
152 ing with iterative refocusing proposed by Williamson et al. (2013) and applied to the SCM/LES
153 comparison. This tool provides the sensitivity analysis of our turbulence parameteriza-
154 tion to its free parameters and identifies which part of the full free parameter space pro-
155 vides SCM simulations consistent with the chosen reference, accounting for the various
156 sources of uncertainty (Williamson et al., 2013, 2017; Couvreur, Hourdin, et al., 2020).
157 Thus, we do not seek to optimize the ARPEGE-Climat turbulence parameterization over
158 the GABLS4 SBL in the traditional meaning of searching for the best set of free param-
159 eters to reproduce this SBL. In contrast, our objective is to document which sets of free
160 parameters, if such sets exist, provide an appropriate representation of the GABLS4 SBL.
161 This information can then be further used in the whole process of the ARPEGE-Climat
162 calibration and combined with similar information based on other 1D cases or on the full
163 3D model configuration. Such an approach should ensure that the full model calibration
164 will keep a reasonable process-level behavior for the GABLS4 SBL (see Couvreur, Hour-
165 din, et al., 2020; Hourdin, Williamson, et al., 2020, for a more detailed discussion).

166 Section 2 introduces the ARPEGE-Climat 6.3 atmospheric model and its turbu-
167 lence parameterization. The relevant free parameters of the parameterization to be used

for calibration are emphasized. Section 3 presents the case study used for the SCM/LES intercomparison and the LES results that serve as a reference. Section 4 describes the statistical framework. Section 5 details the results obtained for two different configurations of the ARPEGE-Climat 6.3 SCM. A third one is rapidly discussed, with more detail provided in Appendix A. Section 6 discusses several aspects of the methodology and section 7 finally concludes the present study.

2 ARPEGE-Climat 6.3

ARPEGE-Climat is a global atmospheric model developed at CNRM for climate studies. Its latest version (6.3, Roehrig et al., 2020) is the atmospheric component of the CNRM ocean-atmosphere climate model CNRM-CM6-1 (Voldoire et al., 2019), and Earth System model CNRM-ESM2-1 (S  f  rian et al., 2019). The following work uses the single column model (SCM) version of ARPEGE-Climat (e.g., Abdel-Lathif et al., 2018), in the context of the GABLS4 framework (see section 3.1). The model physical package is fully described in Roehrig et al. (2020) and therefore we only insist hereafter on the model features relevant for the present study. ARPEGE-Climat 6.3 standard vertical grid consists of 91 vertical levels, following the progressive hybrid σ -pressure discretization of Simmons and Burridge (1981). The altitude of the first 5 model levels is approximately 8.5, 29, 55, 91 and 132 m. The model timestep is 15 minutes. A version of ARPEGE-Climat 6.3 with higher vertical resolution (2 m up to 400 m) is also used in section 5.1. To prevent instabilities, the timestep of this version is reduced to 60 seconds. Note that the use of this 60-s timestep in the 91-level version of ARPEGE-Climat does not impact much the results of the present work. As described in section 3, the SCM configuration is run on a idealized case (stable boundary layer, no moisture, no radiation), in which only the turbulence and surface flux parameterizations are activated. These parameterizations are described hereafter, in a dry context.

2.1 Turbulence parameterization

The turbulence scheme used in ARPEGE-Climat 6.3 follows the work of Redelsperger and Sommeria (1982), Redelsperger and Sommeria (1986), and Cuxart et al. (2000). It relies on the eddy diffusivity approach, coupled to a prognostic equation for the grid-scale-averaged turbulence kinetic energy (TKE) \bar{e} . Given the standard horizontal resolution of ARPEGE-Climat ($\mathcal{O}(100\text{km})$), only the vertical component of turbulent mixing is parameterized. For any variable ψ impacted by turbulent mixing (e.g., wind component u and v , potential temperature θ), the associated second-order turbulent flux $\overline{w'\psi'}$ reads (primes denote fluctuations with respect to the grid-scale average, noted $\bar{\psi}$):

$$\overline{w'\psi'} = -K_\psi \frac{\partial \bar{\psi}}{\partial z} ; \quad K_\psi = \alpha_\psi \mathbf{CM} L_m \sqrt{\bar{e}} \phi_\psi \quad (1)$$

where α_ψ and \mathbf{CM} are free parameters of the parameterization, L_m is the mixing length, and ϕ_ψ is a stability function. ϕ_ψ is taken to 1 for momentum and turbulence kinetic energy ($\psi \in \{u, v, e\}$). For the potential temperature θ , the following formulation is used:

$$\phi_\theta = \frac{1}{1 + C \frac{g}{\theta} \frac{L_m^2}{\bar{e}} \frac{\partial \bar{\theta}}{\partial z}} \quad \text{where } C \text{ is a free parameter.} \quad (2)$$

In Equation 1, \mathbf{CM} modulates all turbulent fluxes in the same way. α_u and α_v are taken to 1, and α_θ is the inverse Prandtl number in neutral condition (i.e. when $\phi_\theta = 1$). In the following, α_e and α_θ will be referred to as **AE** and **AT**, respectively. The turbulent fluxes defined by Equation 1 are computed only above the surface layer, namely from the first model half level. The surface flux boundary condition is provided by the surface flux parameterization (see section 2.2).

211 Eddy diffusivity coefficients K_ψ depend on the intensity of $\bar{\epsilon}$. The time evolution
212 of $\bar{\epsilon}$ is given by:

$$\frac{\partial \bar{\epsilon}}{\partial t} = -\frac{1}{\bar{\rho}} \frac{\partial}{\partial z} (\bar{\rho} \overline{e'w'}) - \left(\overline{u'w'} \frac{\partial \bar{u}}{\partial z} + \overline{v'w'} \frac{\partial \bar{v}}{\partial z} \right) + \frac{g}{\bar{\theta}} \overline{w'\theta'} - \frac{\bar{\epsilon} \sqrt{\bar{\epsilon}}}{L_\epsilon} \quad (3)$$

213 where ρ is the air density, g is the gravity acceleration, and L_ϵ the dissipation length.
214 L_ϵ is assumed to be proportional to the mixing length: $L_\epsilon = \mathbf{CE} L_m$, with \mathbf{CE} a free
215 parameter.

216 The mixing length follows the non-local formulation of Bougeault and Lacarrère
217 (1989) and reads

$$L_m^{\text{BL89}} = \left[\frac{1}{2} \left((L_{\text{up}})^{-2/3} + (L_{\text{down}})^{-2/3} \right) \right]^{-3/2} \quad (4)$$

218 where L_{up} and L_{down} are respectively the maximum upward and downward displacements
219 a parcel can travel within the ambient thermal stratification, given its turbulence kinetic
220 energy, and accounting only for the work of its buoyancy. Close to the surface, the mix-
221 ing length is supposed to be larger than κz where $\kappa = 0.4$ is the Von Kármán constant.
222 A minimum mixing length, \mathbf{LMIN} , is introduced to maintain a minimum vertical mix-
223 ing in stable boundary layers, such that the mixing length L_m finally reads:

$$L_m = \max [L_m^{\text{BL89}}, \min(\mathbf{LMIN}, \kappa z)] \quad (5)$$

224 In case of shallow stable boundary layer, another lower bound, which applies mainly
225 close to the surface, is introduced directly on the turbulent fluxes, to avoid runaway cool-
226 ing of the surface (especially in snow- or ice-covered regions):

$$\overline{w'\psi'} = \max \left[-K_\psi \frac{\Delta \bar{\psi}}{\Delta z}, \alpha_\psi \left(\mathbf{KOZMIN} \left(1 - \frac{z}{\mathbf{ZMAX}} \right) \right) \Delta \bar{\psi} \right] \quad (6)$$

227 where \mathbf{KOZMIN} and \mathbf{ZMAX} are two free parameters, and $\Delta \bar{\psi}$ is the vertical differ-
228 ence of $\bar{\psi}$ between two consecutive model layers (distant of Δz). Above \mathbf{ZMAX} , no lower
229 bound is used. This formulation, through $\Delta \bar{\psi}$ and Δz depends on the vertical discretiza-
230 tion of the model (see also section 5). This empirical formulation is highly questionable,
231 especially in terms of vertical resolution dependence. Nevertheless, we considered it as
232 it is currently implemented in ARPEGE-Climat 6.3. The present work provides the op-
233 portunity to discuss its behavior and possible detrimental impacts with respects to sta-
234 ble boundary layers.

235 The turbulence parameterization thus includes several free parameters that have
236 to be calibrated. Eight parameters have been identified here. The calibration of the pa-
237 rameterization consists in choosing a value for each of them, accounting for both param-
238 eterization performance and physical constraints. In the standard configuration of ARPEGE-
239 Climat 6.3 (see Roehrig et al., 2020), the parameter values follow the work of Cheng et
240 al. (2002) except the parameters \mathbf{LMIN} , \mathbf{KOZMIN} and \mathbf{ZMAX} that were introduced
241 in the course of the parameterization implementation in ARPEGE-Climat and set in a
242 more empirical way. Table 1 provides the values of these parameters as currently used
243 in ARPEGE-Climat 6.3 as well as those initially proposed in Cuxart et al. (2000). Note
244 that the parameter C in Equation 2 is set to 0.143. As the model is not much sensitive
245 to it, the following work does not consider this parameter.

Table 1. Free parameters of the turbulence parameterization. The values in the standard version of ARPEGE-Climat are those from Cheng et al. (2002). Those from the work of Cuxart et al. (2000) are also included. The bottom two lines provide the range of values that we explore for each parameter.

	CM	AE	AT	CE	LMIN	KOZMIN	ZMAX
ARPEGE-Climat 6.3	0.126	2.70	1.13	0.85	10.0	5.10^{-3}	200
Cuxart et al. (2000a)	0.0667	6.0	2.5	0.70	10.0		
Lower bound	0.05	0.50	0.20	0.33	0.0	0.0	30
Upper bound	0.30	6.00	3.00	5.00	10.0	5.10^{-3}	400

2.2 Surface flux parameterization

The SCM configuration of ARPEGE-Climat will be used in two different configurations with respects to the surface boundary conditions, one with prescribed surface sensible heat flux and one with prescribed surface temperature. In both, the roughness lengths for momentum (z_0) and heat (z_{0h}) are prescribed.

2.2.1 Configuration with prescribed surface sensible heat flux

The friction velocity u_* is computed following Paulson (1970):

$$u_* = \frac{\kappa \bar{U}_1}{\ln\left(\frac{z_1}{z_0}\right) - \varphi\left(\frac{z_1}{L_{MO}}\right) + \varphi\left(\frac{z_0}{L_{MO}}\right)} \quad (7)$$

where \bar{U}_1 is the wind intensity ($\bar{U}_1 = \sqrt{\bar{u}_1^2 + \bar{v}_1^2}$) at the first model level of altitude z_1 , and L_{MO} is the Monin-Obukhov length. The similarity function φ is given by Paulson (1970). As L_{MO} depends on u_* , the computation is done iteratively, initialized from a neutrally-stable state (i.e. $L_{MO} = \infty$, knowing $\varphi(0) = 0$). The surface momentum flux is finally given by:

$$F_u = \bar{\rho} u_*^2 \quad (8)$$

2.2.2 Configuration with prescribed surface temperature

In this configuration, the standard version of the ARPEGE-Climat surface scheme is used. The surface momentum and heat fluxes are computed based on the formulations of Mascart et al. (1995) involving the bulk Richardson number Ri_b^0 :

$$Ri_b^0 = \frac{gz_1(\bar{\theta}_1 - \theta_s)}{\frac{1}{2}(\bar{\theta}_1 + \theta_s)\bar{U}_1^2} \quad (9)$$

where $\bar{\theta}_1$ and $\bar{\theta}_s$ are the potential temperature at the first model level and at the surface, respectively. A critical Richardson number $Ri_c = 0.1$ is used as an upper bound of the bulk Richardson number: $Ri_b = \min(Ri_b^0, Ri_c)$. Note this critical Richardson number value is most likely not appropriate in the context of strongly stable boundary layers. It may be considered as a free parameter in future work. The exchange coefficients

267 for momentum and heat in the case of stable states ($Ri_b > 0$) following Mascart et al.
 268 (1995) and Noilhan and Mahfouf (1996) read:

$$C_d = \frac{\kappa^2}{\left(\ln\left(\frac{z}{z_0}\right)\right)^2} \frac{1}{1 + \frac{B_1 Ri_b}{\sqrt{1+B_2 Ri_b}}} \quad (10)$$

$$C_h = \frac{\kappa^2}{\left(\ln\left(\frac{z}{z_0}\right)\right)^2} \left(\frac{\ln(z/z_0)}{\ln(z/z_{0h})}\right) \left(\frac{1}{1 + B_3 Ri_b \sqrt{1 + B_2 Ri_b}}\right) \quad (11)$$

269 and are used to compute the surface fluxes:

$$F_u = \bar{\rho} C_d \bar{U}_1^2 \quad \text{and} \quad F_s = \bar{\rho} C_p C_h \bar{U}_1 (\theta_s - \bar{\theta}_1) \quad (12)$$

270 where C_p is the heat capacity of air at constant pressure.

271 Note that the present surface flux parameterizations include internal free param-
 272 eters ($B_1 = 10$, $B_2 = 5$, $B_3 = 15$), which are not considered in the following analysis.
 273 They are possibly critical for SBLs (e.g. Vignon, Hourdin, et al., 2017), and will be an-
 274 alyzed in a future work.

275 3 Experimental setup and reference simulations

276 3.1 The GABLS4 framework

277 The present study is based on the GABLS4 model intercomparison case (Bazile et
 278 al., 2014, 2015; Couvreux, Bazile, et al., 2020). It focuses on an austral summer diur-
 279 nal cycle of the boundary layer at Dome C, Antarctic Plateau (123.3E, 75.1S, 3223 m
 280 above sea level, local time (LT) = UTC+8 hours) as observed from 11 December 0800
 281 LT to 12 December 0800 LT. During that day, the boundary layer evolved from a 400-
 282 m deep convective regime during daytime to a nighttime very stable regime covering a
 283 depth shallower than 30 m (Vignon, Hourdin, et al., 2017).

284 The GABLS4 model intercomparison encompasses three different stages. The first
 285 one is dedicated to the intercomparison of SCMs with an interactive snow surface scheme.
 286 The second stage prescribes observed surface temperature, thus suppressing several feed-
 287 backs between the atmosphere and the surface. The third stage consists in an idealiza-
 288 tion of GABLS4 stage 2, in which no moisture, no radiation, no large-scale subsidence
 289 and no large-scale advection of temperature are considered.

290 Couvreux, Bazile, et al. (2020) emphasize that the representation of the full GABLS4
 291 diurnal cycle is a challenge for LES as it requires a large domain for the 400-m deep day-
 292 time convective boundary layer and a very high-resolution for the 30-m deep nocturnal
 293 very stable boundary layer. Therefore, Couvreux, Bazile, et al. (2020) proposed a com-
 294plementary setup focused on the GABLS4 nocturnal stable phase, starting at the end
 295 of the convective period (1800 LT, i.e. 10 hours after the start of the original version of
 296 GABLS4 stages) and covering 11 hours (until 0500 LT). This new setup, referred to as
 297 GABLS4-Stage3-10hr, corresponds to the setup used in the present work.

298 The initial conditions are obtained from the ensemble mean of three LES that took
 299 part to the GABLS4 Stage 3 LES intercomparison (Couvreur, Bazile, et al., 2020). GABLS4-
 300 Stage3-10hr uses the same large-scale forcing as in GABLS4 Stage 3, which thus only
 301 includes a large-scale horizontal pressure gradient through a prescribed geostrophic wind.
 302 This geostrophic wind is constant in time ($u_g = 1.25 \text{ m s}^{-1}$ and $v_g = 4.5 \text{ m s}^{-1}$) and

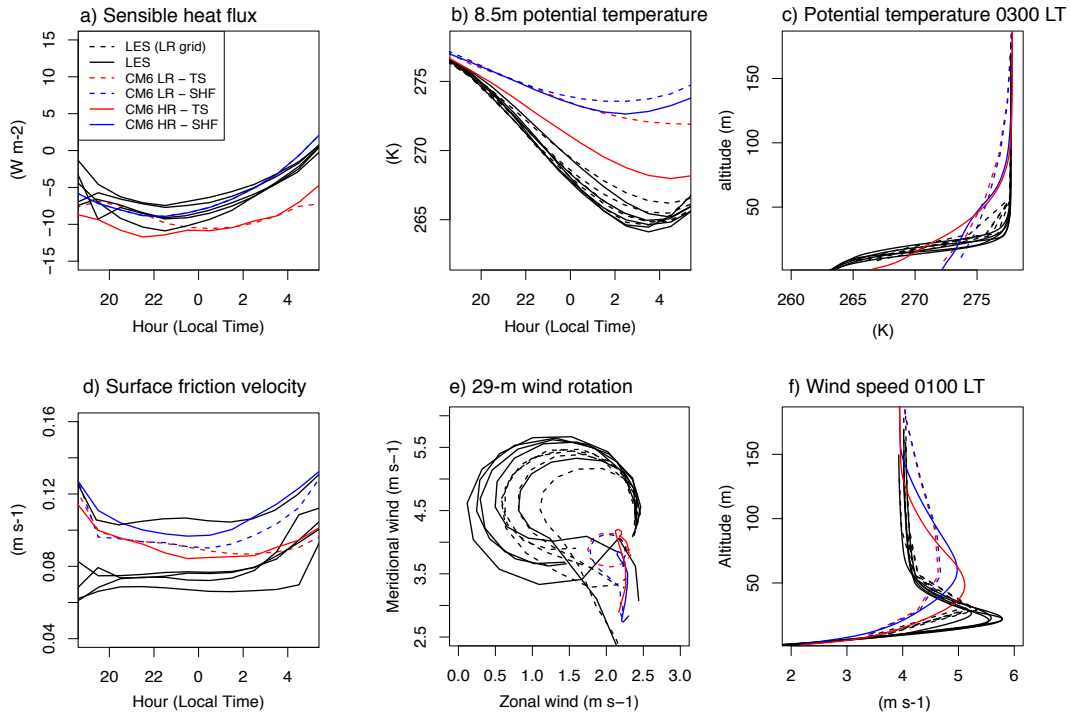


Figure 1. Time evolution of (a) sensible heat flux (W m^{-2}), (b) 8.5-m potential temperature (K), (d) surface friction velocity (m s^{-1}) and (e) 29-m wind rotation, and vertical profile of (c) potential temperature at 0300 LT (K) and (f) wind speed at 0100LT, for the LES (solid black lines), the LES interpolated on CM6-LR vertical levels (dashed black lines), and CM6-LR (dashed lines) and CM6-HR (solid lines) simulations. CM6 simulations are either forced by the surface temperature (red lines) or by the surface sensible heat flux (blue lines).

303 along height. GABLS4 Stage 3 (and thus GABLS4-Stage3-10hr) assumes a dry atmo-
 304 sphere, with no radiation. The surface pressure is held constant to 651 hPa (Dome C
 305 is at 3223 m above sea level), and the surface temperature is prescribed and evolves with
 306 time, following the observations made at Dome C. For the computation of the surface
 307 wind stress, the surface roughness length is set to $z_0 = 10^{-3}$ m for momentum and $z_{0h} =$
 308 10^{-4} m for heat, following Vignon, van de Wiel, et al. (2017) and Couvreur, Bazile, et
 309 al. (2020).

3.2 GABLS4 Large-Eddy Simulations

311 Couvreur, Bazile, et al. (2020) compare 7 LES models over GABLS4-Stage3-10hr.
 312 Two of them are not considered here because of a slightly different setup compared to
 313 the other five (slightly coarser resolution or different roughness lengths). The five remain-
 314 ing LES use an isotropic resolution of 1 m over a $500 \times 500 \times 150 \text{ m}^3$ domain. The LES
 315 compute their surface fluxes (momentum and heat) from the prescribed surface tempera-
 316 ture and roughness lengths using their own parameterization. In such a setup, especially
 317 thanks to the high resolution, the spread among the LES ensemble is rather small, ex-
 318 cept very close to the surface. In particular, there are substantial differences for the sur-
 319 face sensible heat flux (e.g., -6 to -13 W m^{-2} at 2300 LT) or for the friction velocity
 320 (0.07 to 0.11 m s^{-1}) – see Figures 1a,d. This spread is however much smaller than the
 321 spread among the full Stage 3 LES ensemble (Couvreur, Bazile, et al., 2020).

322 Although the setup is idealized compared to the observed situation, observations
 323 have been used qualitatively to evaluate the simulation behavior. The heights of the stable
 324 boundary layer and of the low-level jet are overestimated compared to observations
 325 but this might be due to the neglect of subsidence in the LES simulations. Otherwise,
 326 the jet intensity and the stratification are consistent with observations. The wind turning
 327 at 41 m is also much more realistic in the high-resolution LES than in the coarse-
 328 resolution LES (Couvreur, Bazile, et al., 2020).

329 3.3 SCM configurations

330 The ARPEGE-Climat single column model is run for two different setups. The first
 331 one follows exactly the GABLS4-Stage3-10hr setup detailed in section 3.1. The second
 332 one, derived from GABLS4-Stage3-10hr, and referred to as GABLS4-Stage3-10hr-shf,
 333 prescribes the surface sensible heat flux, instead of the surface temperature, and thus
 334 further removes the coupling between the surface and the atmosphere. In this latter setup,
 335 the prescribed surface sensible heat flux corresponds to the ensemble mean of the 5 LES.

336 In both setups, all model parameterizations are deactivated, except those for the
 337 atmosphere turbulence and the surface fluxes. Note that, in the GABLS4-Stage3-10hr
 338 setup, the surface flux parameterization follows the work of Mascart et al. (1995, , see
 339 section 2.2), while in the GABLS4-Stage3-10hr-shf, it is replaced by the simplified version
 340 described in Paulson (1970, see also section 2.2).

341 We also explore the turbulence parameterization behavior for two different verti-
 342 cal resolutions, namely the standard vertical resolution of ARPEGE-Climat 6.3 (91 verti-
 343 cal levels, with about 15 levels below 1500 m, indicated by LR when relevant), and a
 344 constant high-vertical resolution of 2 m up to an altitude of about 400 m and then de-
 345 creasing as the standard vertical grid up to the model top (indicated by HR when rel-
 346 evant).

347 To conclude, the present work addresses three SCM configurations of ARPEGE-
 348 Climat 6.3: SCM-HR-SHF in which the SCM has a high vertical resolution and is forced
 349 by surface fluxes, SCM-LR-SHF similar to SCM-HR-SHF but with the model standard
 350 vertical grid, and SCM-LR-TS similar to SCM-LR-SHF but forced by the surface tem-
 351 perature. The simulations corresponding to the standard calibration of the turbulence
 352 parameterization (Table 1) are consistently referred to as CM6-HR-SHF, CM6-LR-SHF
 353 and CM-LR-TS.

354 4 Statistical framework : History matching with iterative refocusing

355 The calibration of the free parameters of a parameterization is a difficult task due
 356 to many degrees of freedom (i.e. the number of free parameters) and the computational
 357 cost of operational configuration simulations. For example, the turbulence parameter-
 358 ization used here has eight different free parameters and seven are kept for the calibra-
 359 tion experiments. If one wanted to systematically explore this 7-dimension space, let say
 360 with 10 values in each parameter range, this would require 10^7 simulations. This is clearly
 361 prohibitive for most model configurations, even in an SCM framework. Therefore sur-
 362rogate models, or emulators, are required to appropriately explore the full model param-
 363eter space, with a virtually negligible computational cost. In essence, these surrogate mod-
 364els are mathematical functions that approximate, or emulate, SCM outputs , and which
 365are trained on a few tens to hundreds of well-chosen simulations.

366 In model calibration, one may want to seek for the best or optimal set of model param-
 367eters. As discussed in Hourdin et al. (2017) and in Williamson et al. (2015), this may
 368lead to overtuning issues. Therefore, as advocated in Couvreur, Hourdin, et al. (2020),
 369we adopt a different point of view for model calibration. Namely, we use history match-

370 ing with iterative refocusing as proposed in Williamson et al. (2013). The main objec-
371 tive of the statistical approach is thus not to find the optimal set of parameters, but rather
372 to remove regions of the parameter space in which the model behavior is inappropriate
373 for a given set of performance metrics. The approach thus documents the part of the pa-
374 rameter space which is compatible with the relevant references, accounting for the sev-
375 eral sources of uncertainty. The algorithm is iterative, in the sense that several consec-
376 utive waves are considered. Each wave starts with a small number of new simulations
377 performed with the full model. These new simulations aim at improving the surrogate
378 model accuracy, but only where needed (i.e within the acceptable range defined from the
379 previous wave). The search for acceptable parameter values is thus refocused in a reduced
380 space during the following wave. The iterations (waves) stop when convergence is achieved
381 for the space of acceptable parameters.

382 In the following, history matching with iterative refocusing (or history matching
383 hereafter for the sake of brevity) is applied in a SCM/LES comparison framework as de-
384 tailed in Couvreur, Hourdin, et al. (2020). Its different steps are briefly presented be-
385 low.

- 386 1. **Selection of the targeted metrics.** These metrics aim at summarizing the model
387 behavior. Reference metric values are computed, here using LES results. In the
388 present framework, a metric is a diagnostic that is used to quantify how a sim-
389 ulation (or its emulated version) is close to the reference. It is simply a scalar func-
390 tion of the model free parameters, such as the temperature or the wind speed at
391 a given altitude and at a given time. It may be more sophisticated and include
392 time or vertical integration of a given model variable, or be a function of several
393 model variables. They generally do not include reference information (e.g., such
394 as scores) as that is included in the implausibility (see hereafter).
- 395 2. **Identification of the free parameters.** Free parameters of the model param-
396 eterizations are selected and their possible ranges (generally determined from the
397 modeler expertise) defines the input parameter space. In our case, only param-
398 eters from the turbulence scheme are selected (see section 2.1). Their ranges are
399 given in Table 1.
- 400 3. **Experimental design.** The first wave experimental design (namely the sets of
401 parameters for which a simulation is actually performed) is built by sampling the
402 parameter space previously defined. An optimal sampling of the input param-
403 eter space is ensured using a latin hypercube method (Williamson, 2015).
- 404 4. **Building of the surrogate models.** Metrics defined in step 1 are computed from
405 the first wave simulations (performed in step 3) and used as a training sample to
406 build a surrogate model for each metric. Each surrogate model is a Gaussian-process-
407 based statistical model, a tool well-known in the Uncertainty Quantification com-
408 munity, and which has the advantage to predict both the metric and its uncertainty
409 (e.g., Salter & Williamson, 2016). A prior mean function is imposed as a linear
410 combination of simple functions of the input parameters (monomials, Fourier func-
411 tions and interaction terms chosen according to a predictor selection method – see
412 Williamson et al., 2013). The Gaussian process (GP) assumed a stationary squared
413 exponential kernel, fitted using a Hamiltonian Monte Carlo method (Carpenter
414 et al., 2017). Further details can be found in Couvreur, Hourdin, et al. (2020) and
415 references therein.
- 416 5. **History matching.** The parameter space is then systematically explored using
417 the surrogate models built during the previous step, and the emulated metrics are
418 compared to their respective reference. The space of acceptable parameter values
419 is finally determined by ruling out the parts of the input parameter space which
420 lead to metric values too far from the reference value, accounting for the uncer-
421 tainty on the references, the SCM and the surrogate models. For a given metric

422 f , the following measure $I_f(\boldsymbol{\lambda})$, referred to as implausibility, is thus introduced
 423 to identify which part of the parameter space has to be ruled out:

$$I_f(\boldsymbol{\lambda}) = \frac{|r_f - \mathbb{E}[e(\boldsymbol{\lambda})]|}{\sqrt{\sigma_{r,f}^2 + \sigma_{d,f}^2 + \text{Var}[e(\boldsymbol{\lambda})]}} \quad (13)$$

424 where $\boldsymbol{\lambda}$ is a point of the parameter space and $e(\boldsymbol{\lambda})$ the metric value predicted by
 425 the surrogate model. More precisely, $\mathbb{E}[e(\boldsymbol{\lambda})]$ is the expectation value of the met-
 426 ric Gaussian process, and $\text{Var}[e(\boldsymbol{\lambda})]$ is its variance. The latter is a measure of the sur-
 427rogate model uncertainty for the set of parameters $\boldsymbol{\lambda}$. r_f is the reference met-
 428ric value, and, in the present work, the LES ensemble mean of the metric f . $\sigma_{r,f}^2$
 429is the reference uncertainty and is computed as the LES ensemble variance of the
 430metric f . $\sigma_{d,f}^2$ is the SCM discrepancy or structural error for this metric. This last
 431value is a priori unknown and its estimation could be challenging. Williamson et
 432al. (2017) suggests a “tolerance-to-error” approach for model discrepancy. The im-
 433plausibility $I_f(\boldsymbol{\lambda})$ thus measures the distance between the reference and the pre-
 434dicted metric value, normalized by the sum of uncertainties (supposed to be in-
 435dependent). The implausibility can thus be small either because the predicted value
 436is close to the reference, or because the uncertainties are large. From the implau-
 437sibility, and thus after this first iteration through steps 1 to 5, the Not-Ruled-Out-
 438Yet (NROY) space for the metric f is defined as:

$$\text{NROY}_f^1 = \{\boldsymbol{\lambda} \mid I_f(\boldsymbol{\lambda}) < T\} \quad (14)$$

439 where T is a given threshold taken here as 3, following the 3- σ rule of Pukelsheim
 440 (1994), which states that at least 95% of any unimodal distribution lies within its
 441 3- σ range centered around its mean. This threshold can be reduced once the sur-
 442rogate model for the metric f is sufficiently accurate. In the case of multiple met-
 443rics, the full NROY¹ space is simply defined as the intersection of the NROY spaces
 444 computed for each metric:

$$\text{NROY}^1 = \bigcap_f \text{NROY}_f^1 = \{\boldsymbol{\lambda} \mid I_f(\boldsymbol{\lambda}) < T \text{ for all } f\} \quad (15)$$

445 Note that such an option in combining metrics may lead to multiple testing prob-
 446lems (e.g., Vernon et al., 2010; Couvreur, Hourdin, et al., 2020), but the small num-
 447ber of metrics used in this work reduces this risk.

448 As mentioned above, the implausibility $I_f(\boldsymbol{\lambda})$ for a set of parameters $\boldsymbol{\lambda}$ can be small
 449 either because the SCM emulated metric is close to its reference, or because the emu-
 450lator uncertainty at $\boldsymbol{\lambda}$ is high. Note that $\sigma_{r,f}$ and $\sigma_{d,f}$ are input data to the whole his-
 451tory matching strategy, independent of the surrogate models. The **iterative refocus-**
 452**ing** thus seeks to reduce this emulator uncertainty and thus provide a better estimate
 453of the true space of acceptable sets of parameters. As a result, several iterations (or waves)
 454of the steps 3 to 5 are performed, with an adapted experimental design in step 3. Namely,
 455we only need to improve the emulator accuracy within the current NROY space, thus
 456we only resample NROY¹ after wave 1, or NROY^N after wave N . Once the surrogate
 457model uncertainty is sufficiently reduced so that the other uncertainties dominate the
 458implausibility, the NROY space is not further shrunk by new waves. The iterative pro-
 459cess has thus converged and it is not necessary to perform additional waves. The con-
 460vergence question is discussed in section 6.2.

461 5 Ability of ARPEGE-Climat to simulate the GABLS4 stable bound- 462ary layer

463 The tool introduced in the previous section is used to assess whether the ARPEGE-
 464Climat turbulence scheme is able to capture the main properties of the GABLS4 stable

465 boundary layer. In particular, we determine which part of the space of model free pa-
 466 rameter values is compatible with the LES references. The vertical resolution is poten-
 467 tially a critical aspect for the scheme and therefore we start with the SCM-HR high-resolution
 468 configuration of ARPEGE-Climat. We also remove a degree of freedom in the surface-
 469 atmosphere coupling by prescribing the surface sensible heat flux instead of the surface
 470 temperature, to focus only on the turbulence scheme (SCM-HR-SHF configuration). Then,
 471 section 5.2 discusses results with the standard vertical resolution of ARPEGE-Climat
 472 and with two different surface boundary conditions (SCM-LR-SHF and SCM-LR-TS con-
 473 figurations).

474 5.1 High-resolution SCM configuration forced by surface sensible heat 475 flux (SCM-HR-SHF)

476 5.1.1 ARPEGE-Climat standard calibration

477 As a starting point of the present work, we evaluate the standard calibration of the
 478 ARPEGE-Climat turbulence scheme (parameter values indicated in Table 1), for a high
 479 vertical resolution, and when forced by the LES ensemble mean surface sensible heat flux
 480 (CM6-HR-SHF simulation). Figure 1b presents the time evolution of the 8.5-m poten-
 481 tial temperature from 1800 to 0700 LT. 8.5 m is approximately the altitude of the first
 482 level in the standard vertical resolution model (between the third and fourth level in the
 483 present high resolution configuration). Linear interpolation is used in the high-resolution
 484 model configuration and in LES to compute the 8.5-m potential temperature. The LES
 485 models (solid grey lines) simulate a significant cooling until 0200 LT, from about 277 K
 486 to about 264 K. The minimum potential temperature is reached around 0300 LT. At that
 487 time, the potential temperature vertical gradient between 15 and 29 m varies between
 488 0.4 and 0.8 K m⁻¹ among the LES (Figure 1c). At 0100 LT, a low-level jet is well formed
 489 in all the LES (Figure 1f). The altitude of its peak is similar in all LES, around 22 m,
 490 and its intensity ranges between 5 and 6 m s⁻¹. The inertial rotation of the wind at 29
 491 m is further emphasized in Figure 1e. It is consistent with the theory (e.g., Blackadar,
 492 1957) and representative of the observations collected at Dome C (Gallée et al., 2015).

493 CM6-HR-SHF severely underestimates the 8.5-m potential temperature cooling dur-
 494 ing the first half of the night (Figure 1b, solid blue line). The minimum potential tem-
 495 perature reaches about 273 K at 0200 LT, about 8 K warmer than the LES correspond-
 496 ing value. The potential temperature vertical profile at 0300 LT (Figure 1c) emphasizes
 497 that CM6-HR-SHF simulates a boundary layer, which is too thick and which stability
 498 is underestimated: the potential temperature vertical gradient is at least six times weaker
 499 (0.06 K m⁻¹) than in the LES. Consistently, the CM6-HR-SHF low-level jet is too high,
 500 located near 55 m (Figure 1f). Its intensity is significantly weaker than in all LES but
 501 one. The wind rotation is also strongly underestimated at 29 m (Figure 1e).

502 In order to assess the model sensitivity to its internal turbulent parameters, we choose
 503 to synthesize the model behavior with four scalar metrics (step 1 in section 4). The sen-
 504 sitivity to the choice and number of metrics is discussed in section 6.1. The nocturnal
 505 cooling and boundary layer stability are quantified using the potential temperatures at
 506 2 m and 8 m (referred to as θ_{2m} and θ_{8m} , respectively); these two CM6-HR vertical lev-
 507 els allow to constrain the θ vertical gradient. These two metrics are computed at 0300
 508 LT, when θ is minimum in the LES. The low-level jet structure is measured using the
 509 maximum of the supergeostrophic wind speed and the wind speed at 55 m (referred to
 510 as w_{max} and w_{55m} respectively). The latter altitude corresponds to the level where the
 511 wind returns to its geostrophic value in the LES (it is also the altitude of the CM6-LR
 512 third level). These two last metrics are taken at 0100 LT when the low-level jet is well
 513 established.

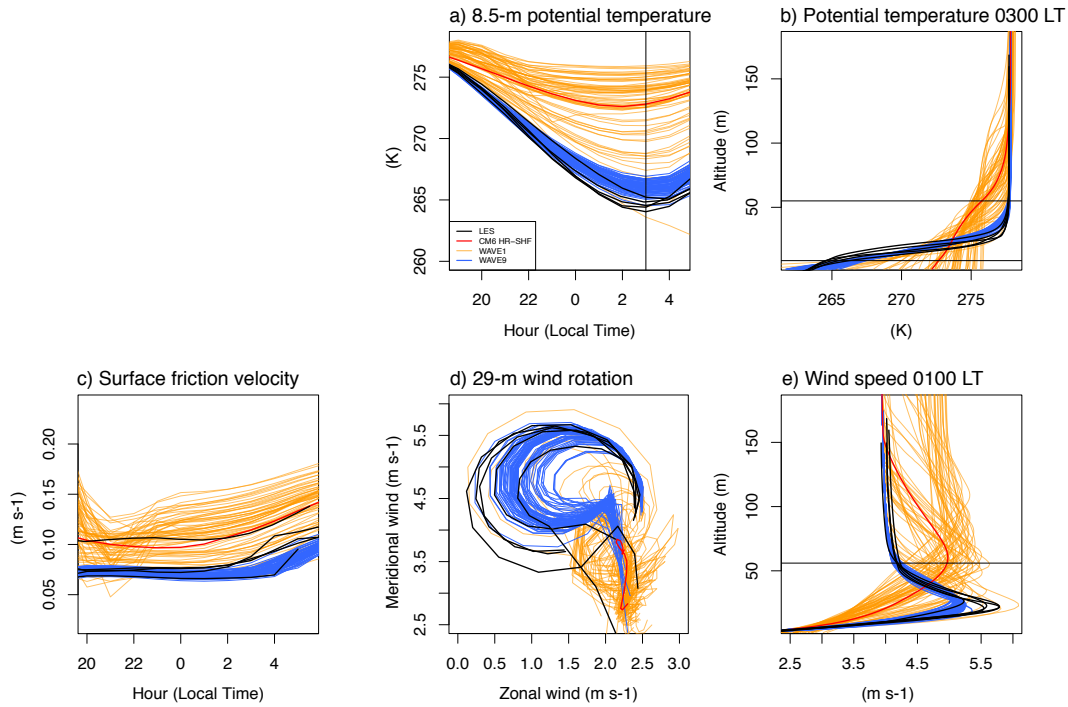


Figure 2. Same as Figure 1 but for the history matching applied to the SCM-HR-SHF configuration. The standard calibration simulation CM6-HR-SHF is indicated with the red line, the Wave 1 70 SCM simulations with the orange lines and the Wave 9 70 SCM simulations with the blue lines. Note that since it is prescribed, the sensible heat flux is not shown.

514

5.1.2 Defining the acceptable range of the turbulence free parameters

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

Based on the seven parameters identified in section 2.1 and their proposed ranges (step 2 of the statistical framework described in section 4), the history matching Wave 1 consists of 70 simulations performed with the SCM-HR-SHF configuration. The associated 70 sets of parameters sample the input parameter space following the experimental design proposed in step 3, namely a Latin hypercube. The results for these 70 simulations are shown by the orange lines in Figure 2. Although the majority of these simulations exhibits a too weak cooling, some of them capture the LES behavior, with both a correct θ vertical profile at 0300 LT and a correct overnight evolution of θ at 8.5 m. Concerning the wind, in most simulations, the low-level jet at 0100 LT is too high, so that the return to the geostrophic wind occurs above 100 m. The wind rotation at 29 m is poorly represented, with a too weak meridional component and a too strong zonal component. Reflecting these first conclusions, the metrics computed for each simulation are most of the time fairly far from those computed with the LES (not shown). Nevertheless, there exist a few simulations, thus a few sets of parameters, for which the chosen metrics have values close to those computed with the LES. This suggests that, at least for each individual metric, there exists an appropriate calibration of the ARPEGE-Climat SCM-HR-SHF configuration. Based on this result, and for the sake of simplicity, in the following, we choose to explore the model performance considering no structural (or tolerance to) error (i.e. $\sigma_{d,f} = 0$ in Equation 13). Such a choice may lead to overtuning (Williamson et al., 2017) and will need to be reconsidered in the context of the full model calibration. It will also be discussed in section 6.2.

536

537

The metric values, computed for each simulation of this Wave 1, are used as a training sample to build a surrogate model for each of the four metrics (step 4 – see section

4 and references therein for their building). These surrogate models then allow us to explore the parameter space more exhaustively by estimating the value of each metric at as many new points as desired. In practice, the complete parameter space is resampled using a new latin hypercube of $\mathcal{O}(10^6)$ points and the surrogate models provide estimate (and uncertainty) of the four metrics for all these new points. As explained in section 4, the following step consists of history matching (step 5): for each point sampled in the parameter space, the implausibility with respect to each metric is calculated, and the maximum implausibility over the four metrics (i.e. the most discriminating metric) is used to characterize NROY¹. Sampled points with an implausibility greater than a threshold set to 3 are ruled out. After this first wave, the remaining space is 3.0% of the initial space, so a large part of the input parameter space is rejected. The NROY space obtained at this stage (not shown) shows that **AE**, **KOZMIN** and **ZMAX** have a negligible influence on the results after this first iteration. The model behavior as a function of the other four parameters indicates a preference for values that significantly reduce the turbulent mixing. Table 2 details for each metric the performance of the 70 simulations performed for this wave 1 and for each of the following waves. For the first wave, all the metrics are very discriminating (more than two thirds of the simulations are incompatible with the LES reference values) except the one representing the jet intensity.

We then iterate several times over steps 3 to 5 described in section 4. Seventy points are randomly sampled in NROY¹ and the corresponding simulations are carried out. These new simulations feed the wave 2. Steps 4 and 5 then provide updated surrogate models in NROY¹ and a new estimate of the NROY space, namely NROY². A third wave can then start. The process is repeated until it is no longer possible to reduce the NROY space (i.e. $\text{NROY}^N \sim \text{NROY}^{N+1}$ if N is the number of a wave). Table 2 indicates no clear further reduction of the NROY space from Wave 6, i.e. the 70 simulations performed at the beginning of Wave 6 are for most of them still in the NROY space estimated at the end of Wave 6 (NROY⁶). This thus suggests convergence of the results and that we can stop the iterations. From Wave 4, w_{\max} , $w_{55\text{m}}$ and $\theta_{8\text{m}}$ metrics are no longer discriminating, i.e. for these metrics, almost no more part of the parameter space is ruled out. For $\theta_{2\text{m}}$, the process evolves more slowly. At Wave 5, 10 simulations are still ruled out because of this metric. From Wave 6 onwards, the number of rejected simulations varies between 2 and 4. The difference between two consecutive waves is then only due to sampling effects of the remaining space. Note that at a given wave N , the current tool version has the limitation to not use the simulations performed during the previous waves $k < N$ to build the wave N emulators. We stopped after 9 waves.

Figure 2 compares the vertical profiles and time series of temperature and wind between the simulations of Wave 1 (orange lines), those of Wave 9 (blue lines), and the reference LES simulations (black lines). The 70 simulations performed in Wave 9 belongs to NROY⁸. Given the chosen metrics and the LES uncertainty, it is not possible to discriminate them (except for 2 of them as indicated in Table 2). Figure 2b shows that the θ vertical profile at 0300LT (time used to compute θ -related metrics) is very close to that of the LES. However, it seems that there is still room for improvement close to the surface but performing an additional iteration does not reduce the near-surface spread of the SCM. Pushing the calibration exercise further would require more simulations but the results obtained are already qualitatively satisfying. Looking at the time series of θ at 8.5 m (Figure 2a), there is also a very clear improvement in the results all along the simulation although we only use the metric at one given time (0300 LT). The results for the wind are good also. The Wave 9 simulations capture well the wind vertical profile at 0100 LT and the 29-m wind veering, falling within the LES spread (Figures 2d,e). Note that the four instantaneous constraints put on the SCM simulations are sufficient to achieve model calibrations that perform well over the entire GABLS4 stable boundary layer, thus illustrating the consistency of the ARPEGE-Climat turbulence scheme physics.

SCM-HR-SHF NROY - Wave 9

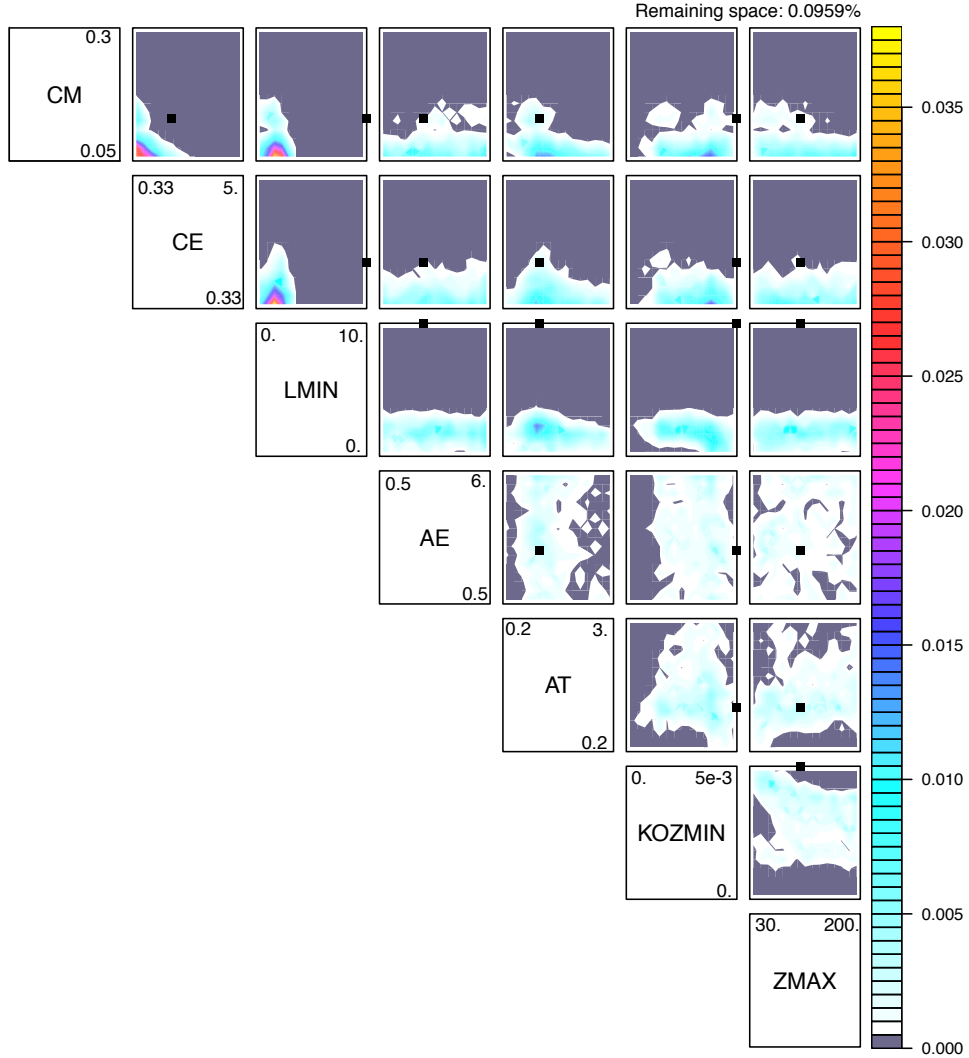


Figure 3. NROY half matrix for Wave 9 of the SCM-HR-SHF history matching (NROY⁹). Each panel displays the fraction of the input space with implausibility lower than the chosen threshold (i.e. NROY⁹, see section 4) as “projected” in a 2-dimension parameter space, the name of the 2 parameters being given in the diagonal of the half matrix. In the diagonal, each parameter box provides the minimum and maximum values to be used for the x -axis of the panels above the box (numbers in the upper part of the box) and the y -axis of the panels to the right of the box (numbers to the right in the box). The black square on each panel locates the parameter values of the ARPEGE-Climat CM6-HR standard calibration. See text for further details.

Figure 3 provides a visualization of the 7-dimension NROY space, obtained after 9 waves, namely NROY⁹. This half matrix tentatively synthesizes the space containing the optimal sets of parameters. Each panel of the half-matrix is a “projection” of NROY⁹ in the 2-dimension space defined by solely two parameters, the names of which are given in the diagonal. For instance, the x -axis of the upper-right panel corresponds to **ZMAX**, varying between its minimum and maximum values provided in Table 1 and indicated in the upper part of its diagonal box (30 and 200 m, respectively) and its y -axis corresponds to **CM**, varying between 0.05 to 0.3 (see the right part of its diagonal box). To numerically characterize NROY⁹ in this 2-dimension parameter space, each panel consists of 15x15 pixels, which values quantify the fraction of the full input space in the other 5 directions that belongs to NROY⁹. This value is estimated thanks to the implausibilities computed over 10⁶ sets of parameters which sample the input parameter space (Latin hypercube). The NROY⁹ fraction of this input space (also equal to the sum of the 15x15 pixel values of each panel) is indicated as text in the upper-right part of the figure (0.0959% in the present case). Grey color of a pixel indicates that the fraction is zero given the input space sampling, meaning that there is mostly no way to fit the reference by varying the other 5 parameters when the 2 panel parameters are fixed at the pixel corresponding values. Dark blue to yellow colors indicate that the combination of the 2 panel parameters are the most likely to provide satisfying results.

As shown in Figure 3, the history matching method selects parameter values that reduce turbulent mixing. The exchange coefficients are significantly reduced (mainly through the high probability of low **CM** and **LMIN** values, see the panels on the first and third rows and on the third column of the half matrix). The calibration also leads to small values of **CE** (panels on the second row and the second column of the half matrix), which controls the dissipation length scale: the TKE dissipation is increased and thus the TKE and the turbulent mixing are reduced. These results are consistent with previous works that generally indicate a too strong turbulent mixing in numerical weather and climate models (e.g., Sandu et al., 2013; Beljaars & Viterbo, 1998). The case of **LMIN** is also consistent with the work of Vignon, Hourdin, et al. (2017) which showed the importance of removing most of the bounds in the turbulence parameterization used in global models to better represent stable boundary layers. The whole framework highlights the high sensitivity of the model results to this parameter (Figure 3), and illustrates the difficulty to calibrate the model for stable boundary layers with **LMIN** values greater than about 4-5 m. The influence of **AT**, which controls the turbulent flux of θ , is significant but not as decisive as that of the three previous parameters. The value currently used in ARPEGE-Climat (black squares on Figure 3) is appropriate. The results are not sensitive to **AE**, so that there is a priori no need to change its current value. This parameter influences the vertical diffusion of TKE. It therefore appears that this term is negligible compared to the other terms in the TKE equation (Equation 3) in the case of the GABLS4 very stable boundary layer (not shown). Finally, **KOZMIN** and **ZMAX** which directly limit the turbulent fluxes, are not relevant to adjust this high-resolution SCM configuration. This is expected given the bound formulation (Equation 6) which tends to zero as the vertical grid spacing goes to zero. These parameters are more critical for the model standard resolution (see section 5.2).

As a preliminary conclusion, the history matching approach demonstrates that the ARPEGE-Climat turbulence scheme contains the required physics to represent the GABLS4 strongly-stable boundary layer, at least for vertical resolution of $\mathcal{O}(1\text{m})$. Besides, if the ARPEGE-Climat standard calibration is not appropriate (i.e. the free parameter standard values are not retained in the NROY space), the standard values of **CM** and **CE** are reasonable (i.e. very close or within the NROY space). It is rather the lower bound **LMIN** on the mixing length that is the most discriminating, a conclusion similar to the one obtained by Vignon, Hourdin, et al. (2017).

Table 2. Evolution of the different metrics (columns) over the successive waves (rows) for the history matching applied to the SCM-HR-SHF configuration. For each metric, the first column indicates the spread (minimum – maximum) of the metric value computed for the 70 simulations performed at the beginning of Wave $N = 1, \dots, 9$. For the LES in the last row, the LES ensemble mean $\pm 3\sigma_{r,f}$ is indicated. The second column (Implausibility $I > 3$) gives the number of these simulations that do not belong to the NROY space estimated at the end of Wave N (namely NROY^N)

	θ_{2m}		θ_{8m}		W_{\max}		W_{55m}	
Wave	Spread	$I > 3$	Spread	$I > 3$	Spread	$I > 3$	Spread	$I > 3$
1	260.3 – 277.9	58	266.3 – 276.5	64	4.8 – 5.7	2	4.1 – 5.3	46
2	252 – 269.4	10	263.2 – 270.7	3	4.7 – 5.8	2	4.1 – 5.6	31
3	255.2 – 265.2	10	262.3 – 268.7	5	5.0 – 5.9	0	4.1 – 4.5	6
4	259.1 – 266.1	7	264.4 – 267.7	4	5.0 – 5.5	0	4.1 – 4.3	0
5	258.9 – 266.5	10	264.4 – 267.3	0	5.0 – 5.4	0	4.1 – 4.5	1
6	261 – 265.1	4	264.7 – 267.4	0	5.0 – 5.4	0	4.1 – 4.3	2
7	261.5 – 264.9	3	264.9 – 267.4	2	5.0 – 5.5	0	4.1 – 4.3	2
8	261.7 – 264.8	4	265.4 – 266.9	0	5.0 – 5.4	0	4.1 – 4.3	0
9	261.7 – 265	2	265.7 – 267	0	5.1 – 5.3	0	4.1 – 4.3	0
LES	263.4 ± 0.43	-	265.0 ± 1.91	-	5.6 ± 0.68	-	4.2 ± 0.06	-

5.2 Standard-resolution SCM configuration forced by surface sensible heat flux (SCM-LR-SHF)

The behavior of the standard-resolution standard-calibration SCM configuration, forced by the LES ensemble-mean surface sensible heat flux (CM6-LR-SHF), is summarized in Figure 1 (dashed blue line). In order to compare the low-resolution SCM results with those of the LES, the latter are regridded onto the vertical grid of CM6-LR. The CM6-LR-SHF overnight cooling is weak, and slightly weaker than in CM6-HR-SHF, thereby indicating sensitivity of the model results to vertical resolution. The minimum potential temperature is reached earlier than in the LES (as for CM6-HR-SHF) and is 8 K warmer. The low-level jet is weakly marked. The altitude of the maximum wind speed is too high (55 m) and the wind speed remains too strong above 55 m, where it should be geostrophic. The wind rotation is slightly better represented than in CM6-HR-SHF but still too weak. The CM6-LR-SHF behavior is thus broadly similar to its high-resolution counterpart, and consistent with an overestimated turbulent mixing. We now investigate whether this behavior is intrinsic to the parameterization (for this standard resolution), or results from a poor calibration for stable boundary layers. Similar scalar metrics to those used in section 5.1 for SCM-HR-SHF are chosen to constrain the θ vertical gradient and the low-level jet structure. The associated altitudes are slightly adapted to be consistent with the SCM standard vertical resolution. The θ gradient is characterized by θ at the first and third model levels at 0300 LT, namely 8.5 m ($\theta_{8.5m}$) and 55 m (θ_{55m}). Because of the rather large uncertainty of θ at the second SCM level (29 m) in the LES (about 5 K), choosing this metric would have been less efficient in reducing the free parameter space. As in the high-resolution experiment, the jet structure is summarized by the wind speed

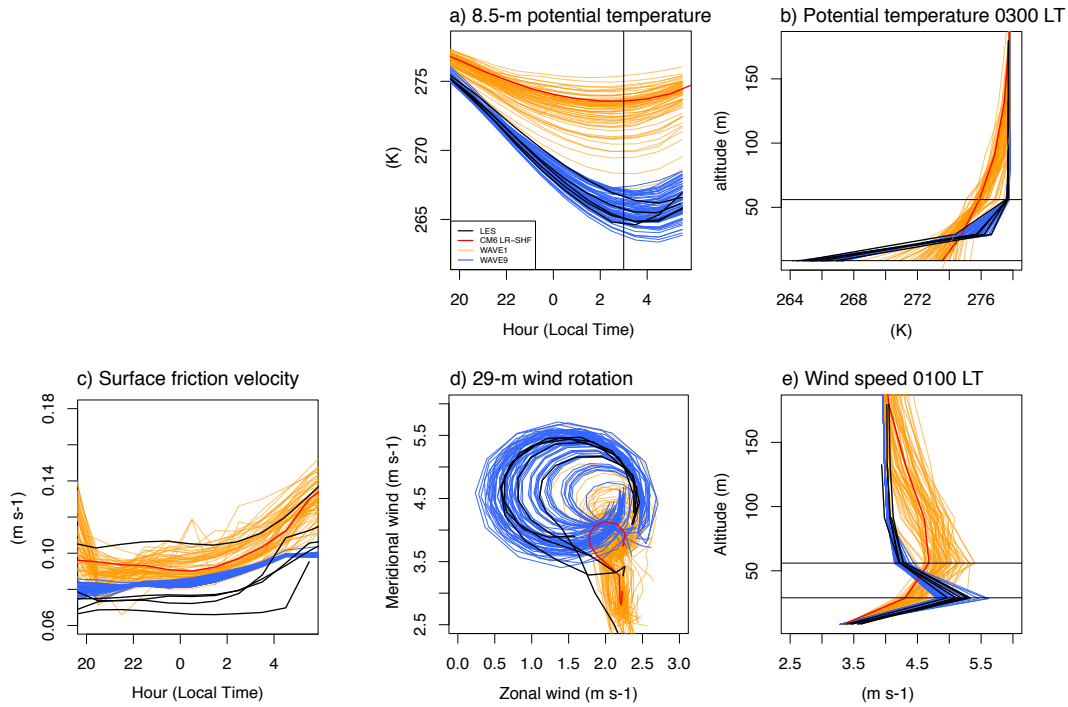


Figure 4. Same as Figure 2 but for the history matching applied to the SCM-LR-SHF configuration. The standard calibration simulation is CM6-LR-SHF (red line).

at 29 m (second SCM-LR level and altitude of the wind maximum – w_{29m}) and at 55 m (third SCM-LR level – w_{55m}).

The Wave 1 simulations present a large spread, with only a few simulations getting close to the LES (Figure 4, orange lines). Note here that before introducing the parameters **KOZMIN** and **ZMAX** in the history matching process, this variety of behavior was not observed and none of the simulations performed well (not shown). Indeed, with the default setting, this minimum bound for the mixing coefficients was systematically reached and any modifications of the other parameters had very little effect as hidden by the minimum bound. The consideration of **KOZMIN** and **ZMAX** in the history matching was crucial in the standard resolution configuration. Table 3 shows the evolution of the metrics for nine waves. From Wave 3, the results are already satisfying for most of the metrics. Only θ_{55m} needs a few more waves to achieve reasonable values. The Wave 9 NROY space, namely $NROY^9$, is finally slightly less than 0.1% of the input parameter space (Figure 5).

The θ vertical profiles at 0300 LT of all the Wave 9 simulations are very close to the LES (Figure 4b, blue lines). The time evolution of the 8.5-m potential temperature still presents some dispersion, which is consistent with the LES uncertainty. Wave 9 simulations also capture well the sharpened wind speed vertical structure at 0100 LT, and much better than CM6-LR (red line). There is still some significant spread in the jet intensity, but again, it reflects the LES discrepancies (Figure 4e). All the simulations reproduce well the 29-m wind rotation (Figure 4d).

The half-matrix of Figure 5 presents the remaining space after nine waves, i.e. $NROY^9$. In contrast with the SCM-HR version, the **KOZMIN** parameter appears critical. The history matching thus shows that the SCM can only behave well over the GABLS4 case for very low values of **KOZMIN**. Given that this parameter is used to maintain signif-

SCM-LR-SHF NROY - Wave 9

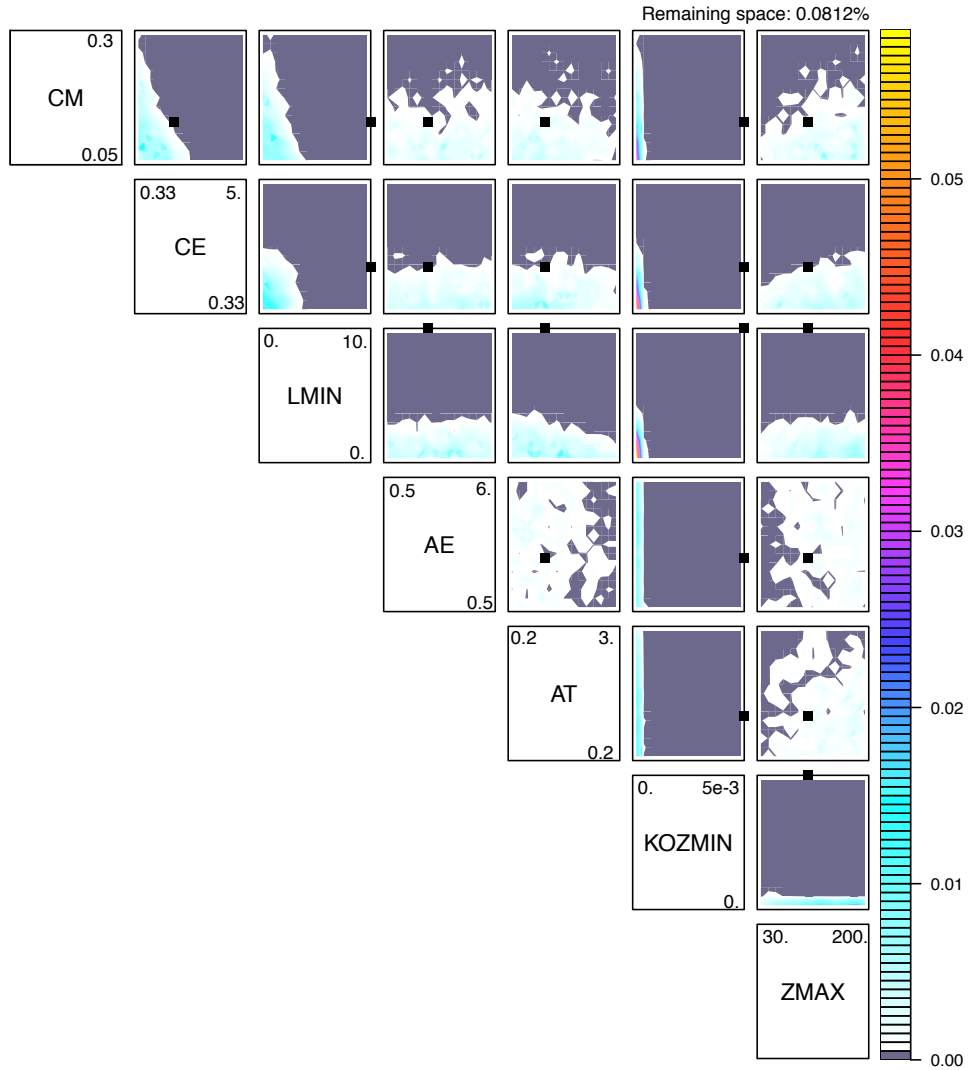


Figure 5. Same as Figure 3, but for the history matching applied to the SCM-LR-SHF configuration (NROY space obtained at Wave 9).

690 icant turbulent fluxes within the first levels of the model, it was rather expected. For the
 691 other parameters, the results are similar to those obtained with the SCM-HR-SHF con-
 692 figuration: parameter values that allow weak turbulent mixing are found the most suit-
 693 able for adjusting the turbulence parameterization. Note that the CM6 values (black square)
 694 of **KOZMIN** and **LMIN** are well beyond their acceptable range (already eliminated
 695 at Wave 1). As for the high-resolution configuration, the SCM behavior is also signif-
 696 icantly sensitive to **CE** and **CM** with acceptable values leading to decrease of the tur-
 697 bulent mixing but the CM6 default values are not ruled out. **AE** and **AT** have no sig-
 698 nificant impact. To summarize, the ARPEGE-Climat turbulence parameterization is able
 699 to capture the GABLS4 stable boundary layer as described by the reference LES, even
 700 with the standard-resolution SCM-LR configuration. Such an acceptable behavior re-
 701 quires to almost remove the two bounds on the turbulent mixing induced by **LMIN** and
 702 **KOZMIN**, while the other free parameters of the parameterization can be kept close
 703 to their original values, similar to those proposed in Cuxart et al. (2000) and Cheng et
 704 al. (2002). The existence of the **LMIN** and **KOZMIN** in the model is thus strongly ques-
 705 tioned by the present results. At least only small values of these parameters are com-
 706 patible with the representation of strongly-stable boundary layers, which suggests that
 707 the associated formulations need to be revised.

Table 3. Same as Table 2, but for the history matching applied to the SCM-LR-SHF configu-
 ration. Note that the LES metric values might be slightly different from those in Table 2, as they
 are computed here using LES data interpolated onto the SCM-LR vertical grid.

	$\theta_{8.5m}$		θ_{55m}		W29m		W55m	
WAVE	Spread	$I > 3$	Spread	$I > 3$	Spread	$I > 3$	Spread	$I > 3$
1	267.9 – 275.7	67	275.3 – 277.2	65	4.6 – 5.5	4	4.4 – 5.4	59
2	263.4 – 271.5	8	276.5 – 277.9	24	4.7 – 6.0	1	4.1 – 4.7	9
3	262.9 – 269.0	1	277.4 – 277.8	2	4.8 – 5.3	0	4.1 – 4.6	3
4	263.2 – 267.8	0	277.4 – 277.8	4	4.8 – 5.4	0	4.1 – 4.6	1
5	263.2 – 268.8	1	277.4 – 277.8	7	4.8 – 5.5	0	4.1 – 4.3	0
6	263.9 – 267.5	0	277.5 – 277.8	2	4.7 – 5.8	1	4.1 – 4.3	0
7	264.4 – 268.4	0	277.5 – 277.8	0	4.8 – 5.7	1	4.1 – 4.4	1
8	263.6 – 268.0	0	277.5 – 277.8	4	4.8 – 5.5	0	4.1 – 4.3	0
9	264.2 – 267.3	0	277.6 – 277.7	0	4.8 – 5.6	0	4.1 – 4.4	3
LES	265.6 ± 2.3	-	277.6 ± 0.32	-	5.2 ± 0.39	-	4.3 ± 0.19	-

708 The history matching approach was also applied to a configuration in which the
 709 surface boundary condition is provided by the surface temperature (SCM-LR-TS con-
 710 figuration), similarly to the LES setup, but also similarly to a more operational use of
 711 the model. In this case, the surface fluxes for momentum and heat follow the formula-
 712 tion described in section 2.2.2. The results are mostly similar to those obtained with the
 713 SCM-LR-SHF configuration used hereabove and therefore are only reported in Appendix
 714 A.

6 Discussion

In this section, we discuss in more detail two features of the calibration statistical approach, that are rather subjective. The selection of metrics is first emphasized as a key step that deserves some caution. We discuss the selection we made in section 2 and analyze the respective role of each selected metrics. Second, we investigate more in depth the convergence criteria, and how we have tackled it. Finally, the SCM computationally-cheap approach provides the opportunity to perform a large simulation ensemble, which can serve as a basis for the evaluation of the full statistical framework.

6.1 Choice of metrics

The choice of metrics in this type of framework is crucial, yet it is difficult to draw up a definitive list in advance. In this section, we illustrate the metric selection procedure in the case of the SCM-LR-SHF history matching and give some guidelines. It might be tempting to choose a large number of metrics to precisely control the model. But for efficiency, especially for the analysis of the results of each wave, the number of metrics should be limited. In particular, redundant metrics should be avoided. Note also that if, in the course of the framework, biases that were not accounted for emerge, new metrics can be added “on the fly” at the beginning of a wave.

The first step consists in conducting an evaluation of the model behavior in order to identify the main model biases (and especially those we care about) and then to design the appropriate metrics to quantify them. We have already seen in section 5.2 that the two main identified biases are a weak nocturnal cooling and an insufficiently marked low-level jet structure. Analysis of the potential temperature profile at 0300 LT (Figure 4b), the time at which the minimum of θ at 8.5 m, the first CM6-LR level, is reached in the LES simulations (Figure 4a), shows that the θ vertical gradient in the nocturnal boundary layer is too weak, with a warm bias (+8 K) at the first model level and a cold bias (about -2 K) at the third model level. The values of θ at these two levels (respectively $\theta_{8.5m}$ and θ_{55m}) are therefore used as metrics to control the model thermodynamics. Concerning the low-level jet structure at 0100 LT (the time at which it is the sharpest in the LES, Figure 4e), the main features to be considered are the altitude and the intensity of the wind maximum (which are respectively too high and too weak in CM6-LR), and the thickness of the jet, which can be represented by the altitude of the wind return to its geostrophic value. The first two features can be summarized by a single metric, namely the wind intensity at the second model level (w_{29m}), which corresponds to the wind maximum in the regridded LES. The third feature can be captured by the wind intensity at the third model level (w_{55m}). Wind intensity at the fourth level could also be chosen and this choice leads to similar results (not shown).

As explained before, the framework used here takes into account the different sources of uncertainty through the notion of implausibility. Note that the lower the uncertainty on the reference metrics the more constraining the metric: particular attention was devoted to the associated reference uncertainty, and metrics with weak reference uncertainty were preferred. When it is significant, at least compared to the Wave 1 simulation errors, it may rapidly dominate the implausibility and the framework will only weakly constrain the model behavior. For the four metrics defined on the basis of model biases, the uncertainty of the LES regarding these metrics is low in front of the biases of Wave 1 simulations (for $\theta_{8.5m}$ and w_{29m}) or almost negligible (for θ_{55m} and w_{55m}).

To understand how each of the selected metrics constrains the model, four history matching experiments are carried out, considering only one of the four metrics at a time. For each metric, the evolution of the remaining space throughout the different iterations is presented in Table 4. The w_{29m} metric is much less discriminating than the other three. 18% of the input parameter space is compatible with the references after 9 waves, as opposed to around 0.5 to 1% for the other metrics. The rather large LES uncertainty for

766 this metric explains part of this result. The importance of the reference uncertainty relative to the bias may be quantified as the ratio between the two quantities (Table 4),
 767 but consideration of this ratio is not sufficient to presuppose the importance of each metric.
 768 The results of these experiments are presented in Figure 6. The θ -related and 55-
 769 m wind speed metrics rule out around 99% of the input parameter space. History match-
 770 ing with only one θ -related metrics (either $\theta_{8.5m}$ or θ_{55m}) lead to simulations with a proper
 771 nocturnal cooling (a little more pronounced when considering only θ_{55m} – compare the
 772 red shading with the grey shading on Figures 6d,h). They also show a low-level jet sig-
 773 nificantly better than CM6-LR with a wind maximum at the right altitude (Figures 6a,e).
 774 However, the return to geostrophic wind is weakly constrained and for many simulations
 775 it occurs at a too high altitude, especially for those resulting from the history match-
 776 ing based on $\theta_{8.5m}$ only (Figure 6a). The history matching based on w_{55m} only leads to
 777 a jet with a structure at 0100 LT very close to the LES (Figure 6m), but with a too weak
 778 nocturnal cooling for many Wave 9 simulations (Figure 6p), even if the improvement com-
 779 pared to CM6-LR is still significant. These different metrics reduce the space of the pa-
 780 rameters differently (Figure 7). If high values for **L**MIN and **K**OZMIN are systemat-
 781 ically eliminated for each of these three metrics (Figures 7a,b,d), they differ for the his-
 782 tory matching results for **C**M, **C**E and **A**T. w_{55m} (Figure 7d) strongly constrains the
 783 sum of **C**M and **C**E but has no influence on **A**T, whereas thermodynamic metrics (Fig-
 784 ures 7a,b) have a less marked sensitivity to **C**M and **C**E but significantly constrain **A**T
 785 by eliminating the highest values. The weaker sensitivity of $\theta_{8.5m}$ and θ_{55m} to **C**M is
 786 explained as the turbulent heat flux depends on the product of **C**M by **A**T (see section
 787 2.1). The w_{29m} metric poorly constrains the nocturnal cooling and the altitude of the
 788 return to geostrophic wind (Figures 6i,l). The usefulness of this particular metric may
 789 be questioned. An experiment conducted without w_{29m} however slightly degrades the
 790 low-level jet structure (not shown) and leads to a slower reduction of the parameter space.
 791

792 In this study, we use simple scalar metrics at given times. The time at which the
 793 metrics are calculated is chosen according to the maximum intensity of the phenomena
 794 they capture. Thus the wind-related metrics are computed at 0100 LT which is the time
 795 when the jet is the sharpest. The θ -related metrics are computed at the time of the strongest
 796 cooling (0300 LT). More complex (vector) metrics could be used, such as vertical pro-
 797 files or time series (e.g., Williamson et al., 2017; Salter et al., 2019), but the results pre-
 798 sented in section 5 emphasize that these very simple metrics are sufficient to constrain
 799 significantly the model behavior over the entire duration of the simulation.

800 We would like to finally stress here that the choices made in section 5 result from
 801 a trial-and-error empirical approach, which we advocate for. To help to analyze the re-
 802 sults, a small number of metrics were selected from a much larger list of potential met-
 803 rics. It is certainly possible to use a greater number of metrics, in practice all those that
 804 the modeler finds relevant. Nevertheless, we emphasize hereafter a few guidelines on this
 805 process of metric selection:

- 806 1. The analysis of the default configuration but also of the first wave simulations is
 807 an important step, which helps to identify model biases or appropriately-simulated
 808 features and build metrics to quantify them.
- 809 2. As the method takes into account the different sources of uncertainty, particular
 810 attention should be paid to the uncertainty in the LES and to ensure that it does
 811 not dominate the biases identified above. Besides, the first wave spread might also
 812 give indications about the discriminatory ability of a given metric.
- 813 3. Metrics do not need to be taken all at once starting at Wave 1 and sometimes it
 814 is preferable to start with few discriminant metrics (with priority metrics being
 815 defined by the modeler expertise) in order to ease the construction of the emu-
 816 lators during the following waves as over a smaller parameter space. This is en-
 817 sured by the tool flexibility, as the modeler can add new metrics from a given wave.
- 818 4. A preference is given to the use of simple metrics, easier to compute and interpret.

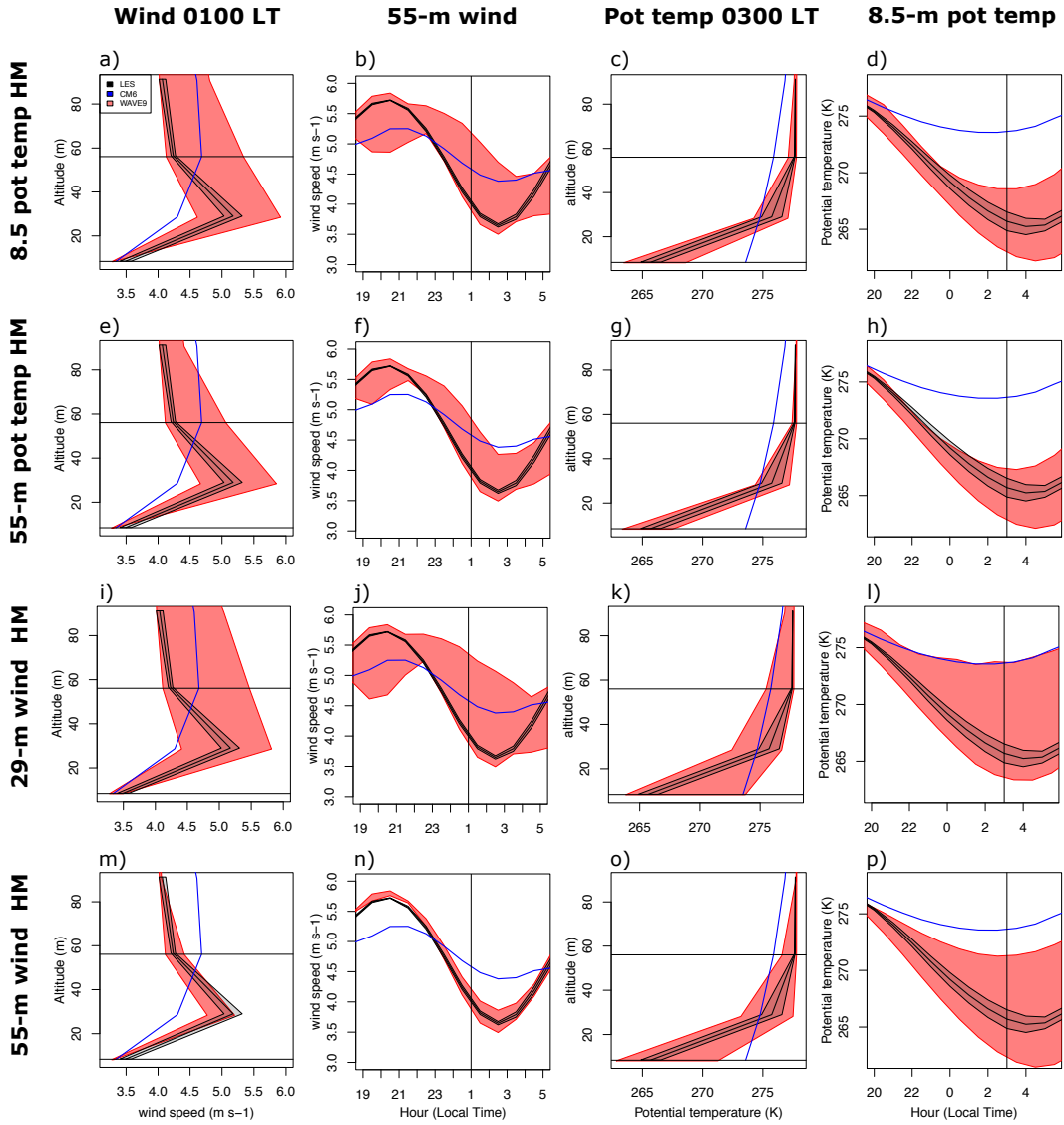


Figure 6. Results of the SCM-LR-SHF history matching experiment, but considering only (a to d) the $\theta_{8.5m}$ metric, (e to h) the θ_{55m} metric, (i to l) the w_{29m} metric and (m to p) the w_{55m} metric. The first column corresponds to the wind speed profile at 0100 LT ($m s^{-1}$), the second column to the wind speed time evolution at 55 m ($m s^{-1}$), the third column to the potential temperature profile at 0300 LT (K) and the fourth column to the 8.5-m potential temperature time evolution (K). On each panel, the blue line indicates the CM6-LR-SHF simulation, the black line the LES ensemble mean with plus or minus one standard deviation as the grey shading and the red shading the Wave 9 simulation envelope.

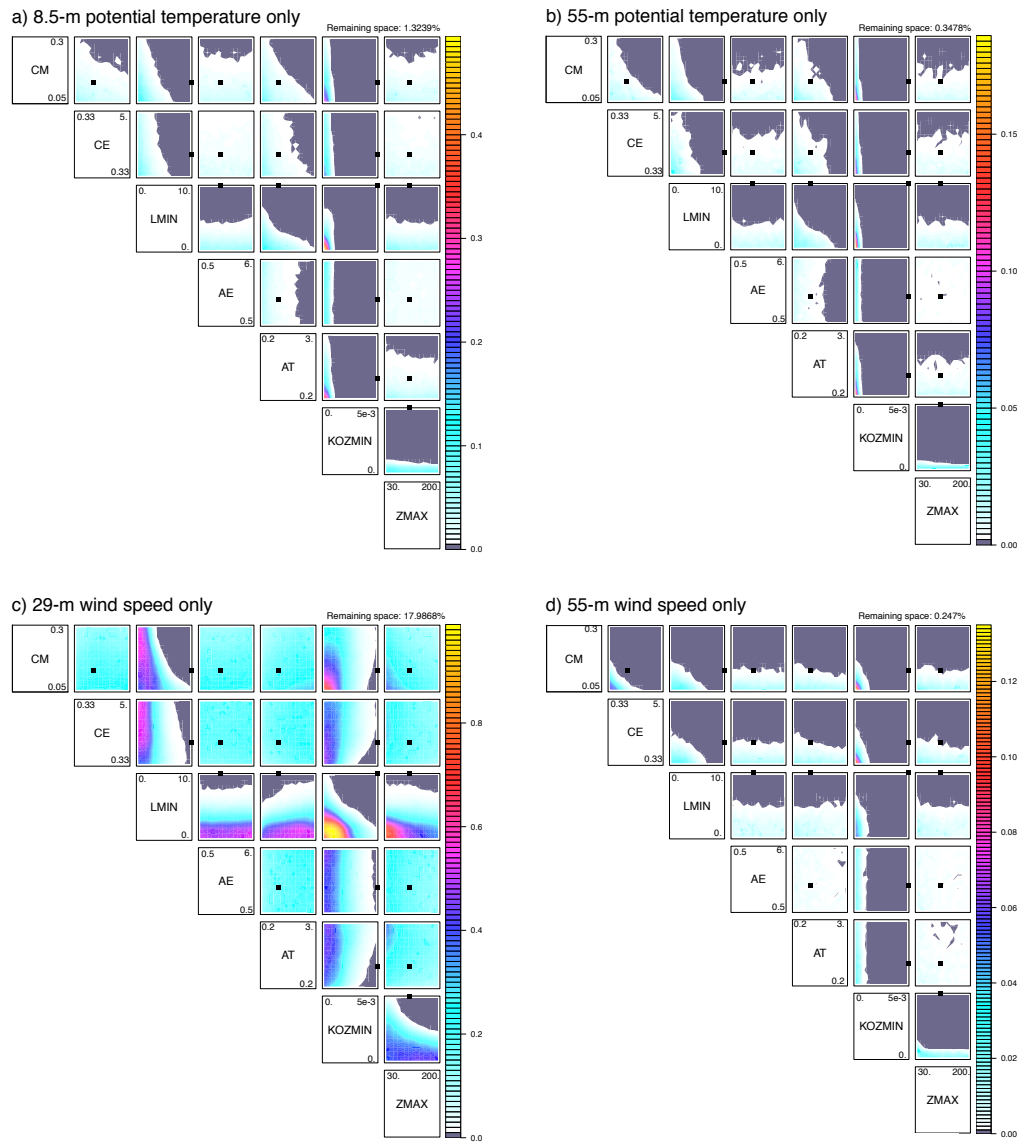


Figure 7. Same as Figure 5, but for the history matching applied to the SCM-LR-SHF configuration considering only (a) the $\theta_{8.5m}$ metric, (b) the θ_{55m} metric, (c) the w_{29m} metric and (d) the w_{55m} metric. On each panel, the NROY space is displayed after 9 waves.

Table 4. Results of the history matching applied to the SCM-LR-SHF configuration, considering only one metric at a time. The first line gives the value of each metric computed with CM6-LR-SHF, the second line gives the value of each reference metric computed as the LES ensemble mean and the third line gives the LES uncertainty computed as the LES ensemble standard deviation. The following lines give the evolution of the NROY space fraction (with respects to the input parameter space) over 9 waves.

Experiment	$\theta_{8.5m}$	θ_{55m}	W _{29m}	W _{55m}
CM6-LR-SHF	273.5	275.5	4.7	4.7
LES	265.7	277.7	5.2	4.2
LES std	0.83	0.027	0.14	0.035
NROY ¹	2.6 %	2.8 %	25.4 %	6.0 %
NROY ²	1.9 %	0.7 %	22.4 %	0.5 %
NROY ³	1.8 %	0.6 %	21.2 %	0.4 %
NROY ⁴	1.6 %	0.5 %	20.8 %	0.3 %
NROY ⁵	1.4 %	0.5 %	19.2 %	0.3 %
NROY ⁶	1.4 %	0.4 %	19.1 %	0.3 %
NROY ⁷	1.4 %	0.4 %	18.7 %	0.3 %
NROY ⁸	1.4 %	0.4 %	18.2 %	0.3 %
NROY ⁹	1.3 %	0.4 %	18.0 %	0.3 %

6.2 Iterative refocusing convergence and its link with the sources of uncertainty

Deciding when to stop the iterations is not trivial. For the different experiments carried out in this study, we decided to stop the iterations when the size of the NROY space no longer decreases as we perform more waves (e.g., Table 5). Using the SCM-LR-SHF experiment, we illustrate here the NROY convergence and analyze the evolution of the different uncertainty sources that we need to take into account to monitor this convergence.

In the SCM-LR-SHF experiment, the NROY space size has converged from about the fourth wave (Table 5). The NROY space estimated at the end of Wave N is defined using the implausibility which depends on the metric value estimated using the surrogate model, the different sources of uncertainty and the chosen threshold. Figure 8 presents the evolution through the different waves of the implausibility distribution (third row), computed for each of the four metrics considered here, as well as the different quantities involved in its computation. For Wave N , the implausibility is computed using the points in the Wave $N-1$ NROY space. From Wave 3, and for each metric, almost all points in the NROY space have an implausibility lower than 3 (the chosen cutoff, cf. section 4) despite the reduced emulator uncertainty within the NROY space compared to the previous wave.

As the NROY space fraction reduces during the successive waves, the 70 SCM runs sample a much smaller space, thus leading to improved surrogate models within the NROY space, with reduced uncertainty. This is particularly the case for the θ_{55m} and w_{55m} met-

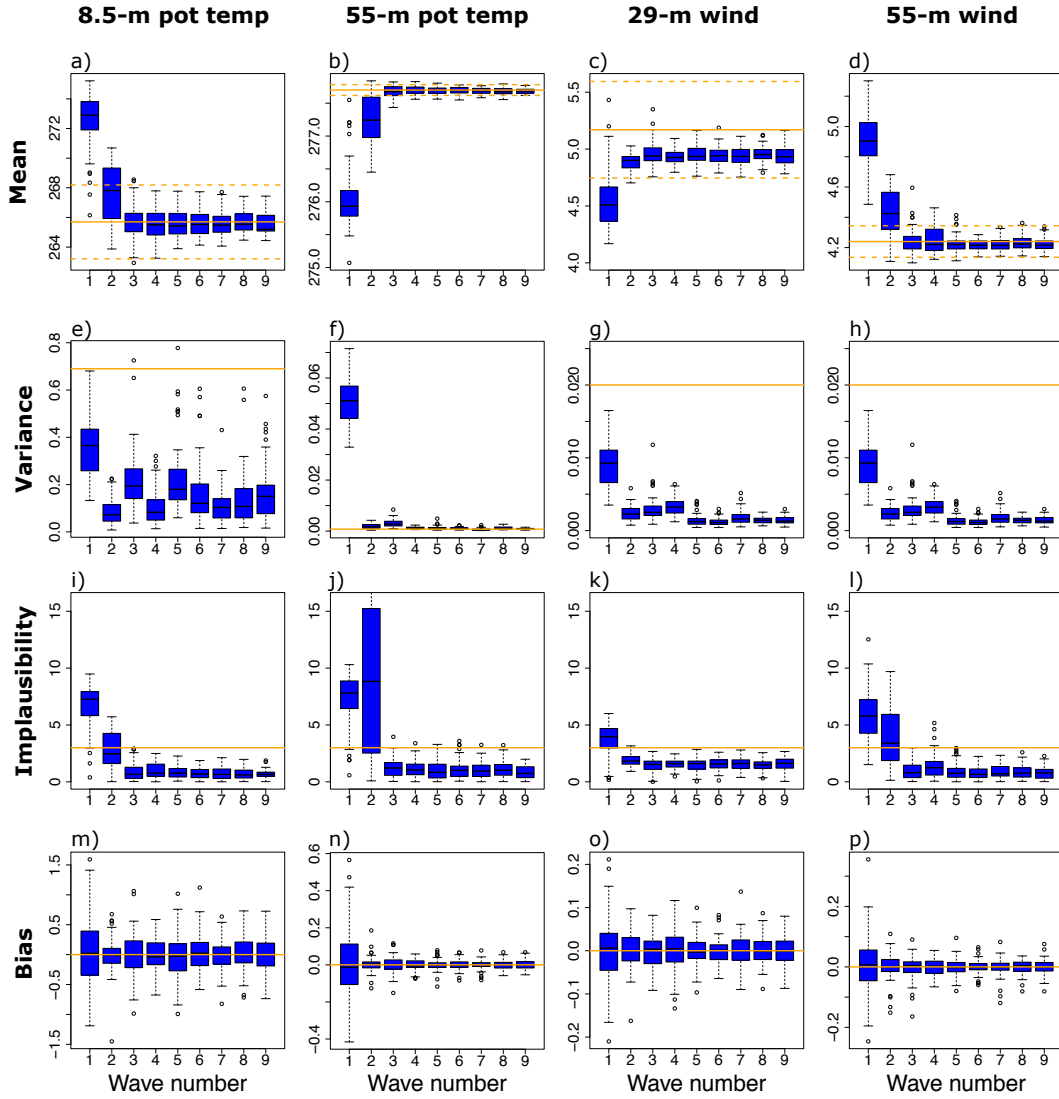


Figure 8. Evolution across the SCM-LR-SHF 9 waves of the distribution of various variables involved in the implausibility computation: (a to d) metrics values with the LES ensemble mean (orange solid line) and plus or minus three standard deviation (dashed orange lines), (e to h) surrogate model variance with the LES ensemble variance (orange solid line), (i to l) implausibility with the cutoff of 3 (orange solid line), and (m to p) the GP-based emulator bias with respect to the SCM ensemble mean of the corresponding wave. The first column correspond to the $\theta_{8.5m}$ metric, the second column to the θ_{55m} metric, the third column to the w_{29m} metric and the fourth column to the w_{55m} metric. The distributions are computed using points sampling the NROY space obtained at the end of the previous wave (input space for Wave 1).

841 rics, for which from Waves 2-3, the surrogate model uncertainty (or variance) falls mostly
 842 below or is of the same order of magnitude as the reference uncertainty (Figures 8f,h).
 843 For $\theta_{8.5m}$, the surrogate model uncertainty is also reduced after Wave 1 (Figure 8e), but
 844 there is not a strong decrease of it. It is already significantly lower than the reference
 845 uncertainty and thus does not play much in the implausibility for this metrics. Besides,
 846 as the NROY space does not anymore reduce dramatically from Wave 2 onwards, the
 847 new SCM runs mostly sample the same space and thus no surrogate model improvement
 848 is expected.

849 The different metrics estimated by the emulators also converge rapidly and from
 850 Wave 3 (Wave 5 for w_{55m}) onwards (Figure 8a-d), only a few outliers fall outside the range
 851 of 3 standard deviations around the reference. In other words, from Wave 5 onwards,
 852 the simulations are consistent with the LES given our four metrics, and taking into ac-
 853 count the LES uncertainty. To go further in the calibration process, the threshold ini-
 854 tially taken to 3 could be changed to a lower value, especially because the emulator un-
 855 certainty is not anymore the dominating uncertainty.

856 The analysis of the metric convergence also emphasizes the question of discrepancy.
 857 If for $\theta_{8.5m}$, θ_{55m} and w_{55m} , the SCM simulations converge to the mean reference, this
 858 is not the case for w_{29m} . For this metric, the LES ensemble mean is about 5.2 m s^{-1} ,
 859 while the SCMs struggle to exceed 5.0 m s^{-1} . If a non-null discrepancy is not necessary
 860 to avoid an empty space, the latter point advocates for taking a structural error of the
 861 model at least equal to 0.2 m s^{-1} . Our choice for no tolerance to error seems reasonable
 862 for exploring a single case study and only a few metrics. It is however not recommended
 863 in a more comprehensive calibration of the full model as it is likely to end up with over-
 864 tuning (Williamson et al., 2017). When no information can help in suggesting or quan-
 865 tifying this discrepancy, tests during the first wave (when computing the implausibili-
 866 ties and identifying the NROY space) are likely to provide some upper bound on it. The
 867 following waves will possibly help to reduce it as the modeler gains a better quantifica-
 868 tion of his model behavior and can further confront what he wants and what the model
 869 can really do.

870 Finally, we want to stress here the importance of analyzing the comparative im-
 871 portance of the different sources of uncertainty when the procedure seems to converge
 872 (i.e. when the space of the parameters no longer reduces significantly from one wave to
 873 the next) before reducing the threshold for example.

874 6.3 Evaluation of the statistical framework: comparison with a 100% 875 SCM approach

876 As SCM low-resolution simulations are computationally cheap, it is possible to per-
 877 form an independent large ensemble of simulations to assess the overall framework used
 878 in section 4, in particular the quality of the GP-based surrogate models. In the present
 879 section, we focus on the SCM-SHF-LR setup for which 10^4 simulations are performed.
 880 The parameters are sampled according to a conventional Latin hypercube of the input
 881 space. The implausibility of each of these 10^4 simulations is evaluated. As there is no
 882 emulator used in this experience, Equation 13 reduces to

$$883 I_f(\boldsymbol{\lambda}) = \frac{|r_f - f(\boldsymbol{\lambda})|}{\sqrt{\sigma_{r,f}^2 + \sigma_{d,f}^2}} \quad (16)$$

884 where $f(\boldsymbol{\lambda})$ is the value of the metric f for the set of parameters $\boldsymbol{\lambda}$, directly computed
 885 from the SCM simulations. As in the previous history matching experiment and in or-
 886 der to compare results, the structural error $\sigma_{d,f}^2$ is taken to zero. This formulation is used
 887 to estimate the NROY space, directly from the 10^4 SCM ensemble. Note that to appro-
 888 priately characterize the NROY space, it is necessary that it contains enough points, at

Table 5. Evolution of the NROY space fraction (with respects to the input parameter space) for the history matching experiments applied to the SCM-HR-SHF and SCM-LR-SHF configurations.

Experiment	SCM-HR-SHF	SCM-LR-SHF
NROY ¹	3.0 %	0.59 %
NROY ²	0.97 %	0.17 %
NROY ³	0.52 %	0.14 %
NROY ⁴	0.39 %	0.12 %
NROY ⁵	0.27 %	0.11 %
NROY ⁶	0.21 %	0.10 %
NROY ⁷	0.16 %	0.09 %
NROY ⁸	0.14 %	0.08 %
NROY ⁹	0.12 %	0.08 %

888 least a few hundreds. The NROY space obtained using the four metrics in section 5.2
 889 is about 0.1% of the input space. It has been numerically evaluated using the implau-
 890 sibility computed over 10^6 points, so that it contains in the end around 10^3 points. With
 891 only 10^4 points in the input space, only an order of 10 points is expected to remain, which
 892 is clearly not sufficient to fully characterize the NROY space and makes its estimate much
 893 sensitive to the input space sampling. This further emphasizes the relevance of using sur-
 894rogate models, even in this SCM framework. Therefore, the statistical framework is eval-
 895 uated using only the w_{29m} metric, as it a priori keeps a sufficient fraction of the initial
 896 space, namely about 18% using emulators (thus about 1 800 points). In parallel, a NROY
 897 space is estimated using the emulators built in section 5.2 on the same 10^4 points in the
 898 parameter space. The use of the same points for both estimates avoids sampling effects
 899 in the comparison.

900 In experiments using emulators, the threshold for implausibility is set at 3. This
 901 choice is deliberately conservative and reduces the risk of ruling out a set of parameters
 902 that would in fact leads to results consistent with the reference. The threshold of 3 fol-
 903 lows the $3\text{-}\sigma$ rule of Pukelsheim (1994) which states that 95% of any unimodal distri-
 904bution lies in the range of $\pm 3\sigma$, σ being the standard deviation of this distribution. To
 905 compare the two methods, keeping this threshold of 3 for the direct approach is not rel-
 906 evant, as in our full SCM approach, there is only one source of uncertainty. Assuming
 907 a Gaussian distribution for the reference, 95% of the distribution is in the range of $\pm 1.96\sigma$
 908 around the mean. We thus reduce the threshold to 1.96 in the full SCM approach when
 909 ruling out parameter values as this provides a more consistent comparison with the GP-
 910 based framework.

911 Figure 9 compares the NROY spaces for the w_{29m} metric, obtained either directly
 912 from the 10^4 SCM simulations (NROY_{SCM}) or with Wave 9 emulator ($\text{NROY}_{\text{GP}}^9$). NROY_{SCM}
 913 is significantly smaller than $\text{NROY}_{\text{GP}}^9$ (9.8% against 18.0%), but the two NROY spaces
 914 are highly consistent. This comparison clearly validates the statistical framework used
 915 in section 5 and the quality of the surrogate models in predicting the metric and its un-
 916 certainty.

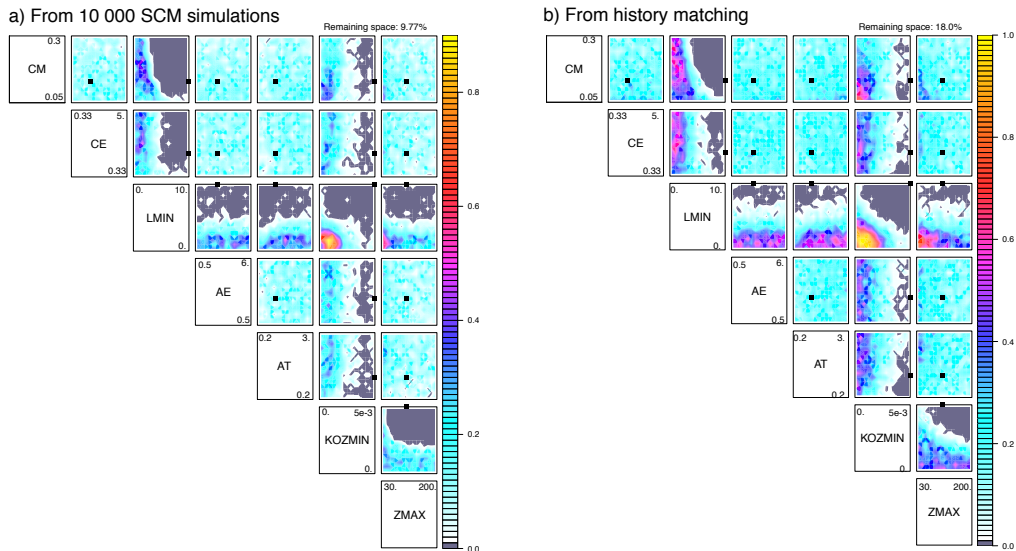


Figure 9. NROY spaces (a) computed from the 10^4 SCM-LR-SHF simulations and (b) computed from the Wave 9 history matching framework applied to SCM-LR-SHF, considering only the w_{29m} metric. See text for details.

7 Conclusion

Stable boundary layers are still critical features for weather and climate models. In the present work, we seek to assess whether these model deficiencies reflect calibration choices, or whether they are more deeply rooted in the formulations and implementations of the turbulence parameterization themselves. In the latter case, this would clearly point to intrinsic limits of current parameterizations and possibly to missing processes key for stable boundary layers. To address this question, we took the example of the CNRM atmospheric model, namely ARPEGE-Climat 6.3, which implements the Cuxart et al. (2000) 1.5-order turbulence parameterization, with a few bounds that were historically and empirically added to prevent undesirable model behavior under certain circumstances (e.g., runaway cooling over Antarctica). At this stage, our example solely makes use of a single-column model framework, based on the very stable boundary layer regime of GABLS4 (Bazile et al., 2015). The Large-Eddy Simulation (LES) ensemble analyzed by Couvreur, Bazile, et al. (2020) serves as reference for evaluating and constraining the model behavior. A statistical approach, based on history matching and Gaussian-Process-based surrogate models, is then used to identify whether there exists or not some calibration of the free parameters of the turbulence parameterization, which can provide satisfying results on this GABLS4 case. More precisely, our framework follows the process-based tuning advocated by Couvreur, Hourdin, et al. (2020) to characterize the part of the free parameter space, which leads to SCM simulations compatible with the LES references, given the various sources of uncertainty.

We have addressed this experience using two vertical resolutions, namely the standard ARPEGE-Climat 6.3 vertical resolution and a LES-type vertical resolution (2 m), and using two configurations for the interaction with the surface (prescribed surface fluxes and prescribed surface temperatures). Using four metrics (two characterizing the temperature profile, and two characterizing the wind profile), sampled at a given time of the GABLS4 nighttime stable boundary-layer regime, we proved that for each SCM configurations, there exist calibrations of the Cuxart et al. (2000) turbulence parameterization, as implemented in ARPEGE-Climat 6.3, which provide results consistent with the

946 LES reference. This indicates in an unambiguous manner that this turbulence param-
 947 eterization contains sufficient physics to capture strongly-stable boundary layers. As ex-
 948 pected, such acceptable model behavior requires calibration that allows to weaken tur-
 949 bulent mixing. This is mostly achieved when strongly reducing the impact of lower bounds
 950 historically introduced for maintaining a minimum turbulent mixing (mixing length and
 951 minimum flux close to the surface). In contrast, and even though it can be revisited, the
 952 calibration of other turbulent parameters can broadly remain consistent with the pre-
 953 vious proposals of Cuxart et al. (2000) and Cheng et al. (2002). This importance of lower
 954 bounds in the turbulence parameterization clearly echoes similar results obtained by Vignon,
 955 Hourdin, et al. (2017) for the climate model LMDZ.

956 The present work is also the opportunity to gather and formalize our experience
 957 with the statistical tools used here and borrowed from the Uncertainty Quantification
 958 community (history matching, GP-based surrogate models). As such, we attempt to pro-
 959 vide guidance for their use in the context of parameterization and atmospheric model
 960 calibration. The importance of the different sources of uncertainty is emphasized,. The
 961 choice of metrics is an important step that is case-dependent: if the analysis of the standard-
 962 calibration model simulation is important by highlighting its main biases, the analysis
 963 of the model behavior over the first wave simulations is equally important. It allows to
 964 explore, to some extent, what the model can do and where the uncertainties lie. We also
 965 illustrate how to understand and tackle the question of the framework convergence by
 966 comparing the emulator uncertainties to the other sources of uncertainty.

967 The present work is a first step in the full calibration of the ARPEGE-Climat at-
 968 mospheric model. It will be first complemented with the addition of new 1D cases that
 969 explore a wide variety of atmospheric regimes. For instance, preliminary new waves that
 970 follow the 9 waves performed with the SCM-LR-TS configuration, and which include ad-
 971 ditional metrics from the first GABLS case (GABLS1, Cuxart et al., 2006), seem to in-
 972 dicate that the moderately-stable boundary layer of GABLS1 requires a minimum mix-
 973 ing. This clearly contrasts with what we obtained here with GABLS4, thus emphasiz-
 974 ing the benefit of using a large diversity of SCM cases (see also Couvreur, Hourdin, et
 975 al., 2020). Then, when the calibration performed in the SCM/LES framework has shrunk
 976 the space of acceptable parameters, it can continue within this Not-Ruled-Out-Yet space
 977 with 3D model configurations. Working in this NROY space constrained with 1D cases
 978 ensures that the model behavior at the process-level remains acceptable, while it dra-
 979 matically reduces the parameter space to be explored with the more expensive 3D con-
 980 figurations. Hourdin, Williamson, et al. (2020) recently provided a first proof of concept
 981 of this whole framework, in the context of boundary-layer clouds.

982 **Appendix A Standard-resolution SCM configuration forced by sur-** 983 **face temperature (SCM-LR-TS)**

984 Similarly to section 5.2, this appendix briefly documents the history matching re-
 985 sults in the case where the ARPEGE-Climat SCM is forced by the surface temperature
 986 (SCM-LR-TS configuration) rather than by the surface sensible heat flux.

987 The standard calibration SCM version (CM6-LR-TS) behaves similarly to CM6-
 988 LR-SHF with an underestimated cooling of the atmospheric low levels and a too weak
 989 low-level jet (Figures 1 and A1). A notable difference is that the model continues to cool
 990 for a longer period (until around 0600 LT). The wind rotation at 29 m remains weakly
 991 captured, as in CM6-LR-SHF.

992 History matching is applied as for SCM-LR-SHF, using the same metrics as in sec-
 993 tion 5.2. 9 waves are performed with a similar convergence as for the SCM-LR-SHF con-
 994 figuration (not shown). The Wave 9 SCM simulations are shown with blue lines on Fig-
 995 ure A1. Overall, the associated sets of free parameters provide improved and satisfying

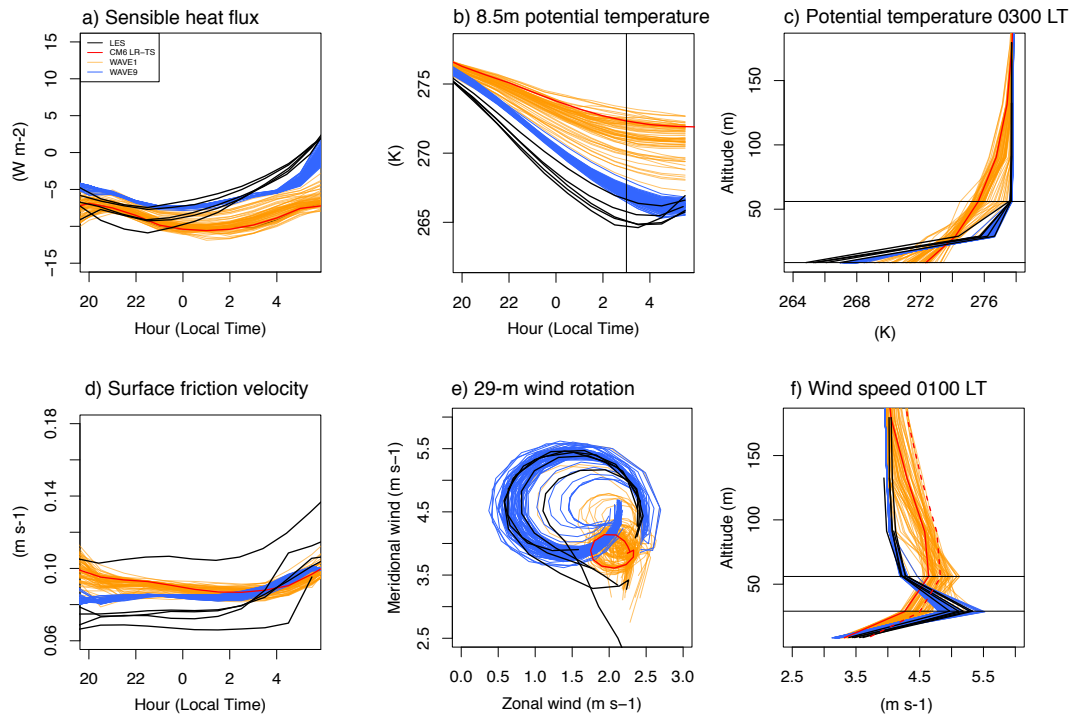


Figure A1. Same as Figure 2 but for history matching applied to the SCM-LR-TS configuration. The standard-calibration simulation CM6-LR-TS is shown with the red line.

996 results over the GABLS4 stable boundary layer, with only a systematic warm bias at 8.5
 997 m at 0300 LT (1.1 to 1.3 K), mainly because of a 2-hour delay in the SBL cooling. Re-
 998 garding the wind, the Wave 9 simulations capture well its vertical structure and rota-
 999 tion, within the LES uncertainty (Figures A1e,f).

1000 Finally, Figure A2 presents the NROY space after the 9 waves for this SCM-LR-
 1001 TS configuration. The conclusions are similar to those made with the SCM-LR-SHF con-
 1002 figuration, with a critical role of **KOZMIN**, **LMIN**, **CM** and **CE**. For these param-
 1003 eters, values leading to low turbulent mixing are required.

1004 Acknowledgments

1005 We are grateful for helpful comments and suggestions from two anonymous reviewers and
 1006 from the associate editor T. Mauritsen. This work received funding from grant HIGH-
 1007 TUNE ANR-16-CE01-0010. It was supported by the DEPHY2 project, funded by the
 1008 French national program LEFE/INSU and the GDR-DEPHY. Daniel Williamson was
 1009 funded by NERC grant: NE/N018486/1 and by the Alan Turing Institute project “Un-
 1010 certainty Quantification of multi-scale and multiphysics computer models: applica-
 1011 tions to hazard and climate models” as part of the grant EP/N510129/1 made to the Alan
 1012 Turing Institute by EPSRC. The LES simulations of the GABLS4-stage3-10hr have been
 1013 kindly provided by G. Matheou, B. Maronga, C. Van Heerwaarden and J. Edwards. The
 1014 High-Tune Explorer (htexplo) tool used here for history matching with iterative refocus-
 1015 ing is available through the open source version control system “subversion” (svn) at [https://](https://svn.lmd.jussieu.fr/HighTune)
 1016 svn.lmd.jussieu.fr/HighTune. A snapshot version of the tool used here is also avail-
 1017 able at <http://doi.org/10.14768/70efa07b-afe3-43a4-8334-050354f9deac>. LES
 1018 and SCM data archiving is underway at <https://doi.org/10.5281/zenodo.4350522>.

SCM-LR-TS NROY - Wave 9

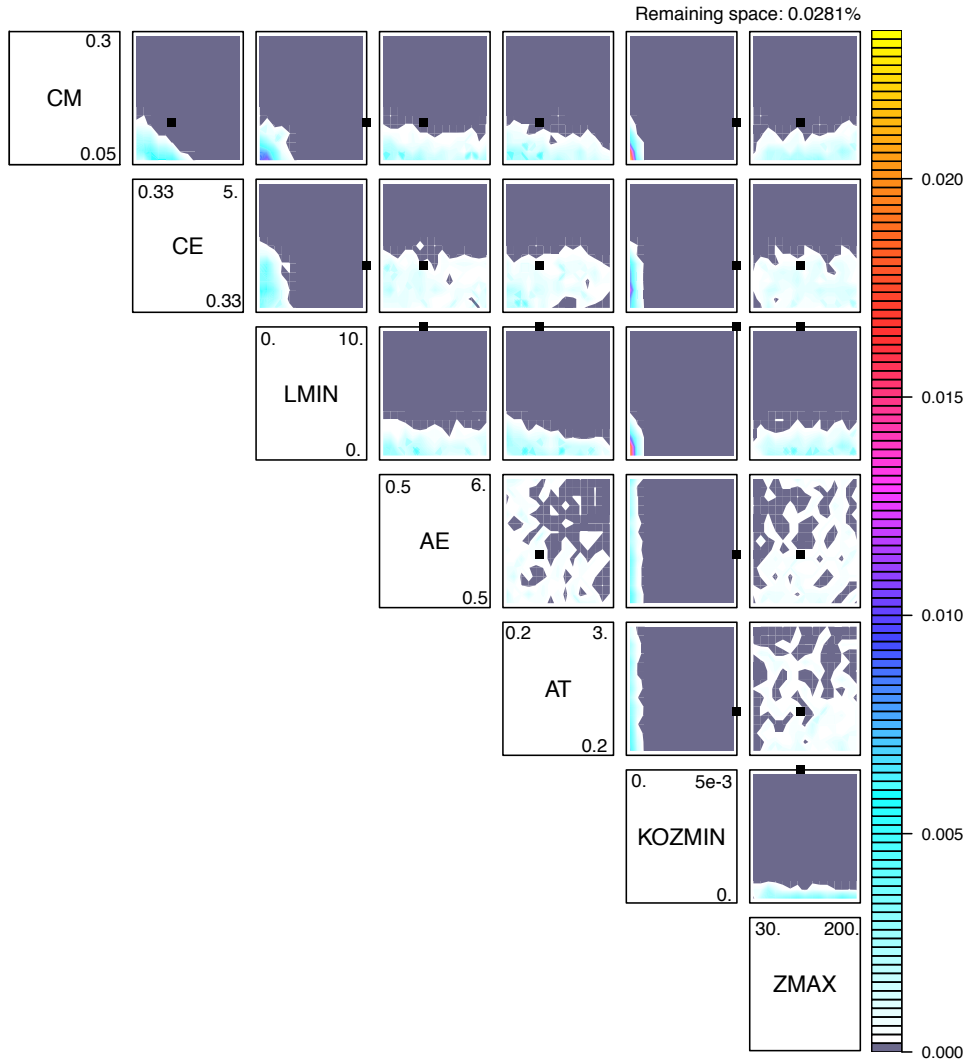


Figure A2. Same as Figure 5, but for the history matching applied to the SCM-LR-TS configuration (NROY space obtained at Wave 9).

References

- 1019
- 1020 Abdel-Lathif, A. Y., Roehrig, R., Beau, I., & Douville, H. (2018). Single-Column
1021 Modeling of Convection During the CINDY2011/DYNAMO Field Campaign
1022 With the CNRM Climate Model Version 6. *Journal of Advances in Modeling*
1023 *Earth Systems*, *10*(3), 578–602. doi: 10.1002/2017MS001077
- 1024 Acevedo, O. C., Mahrt, L., Puhales, F. S., Costa, F. D., Medeiros, L. E., & De-
1025 grazia, G. A. (2016). Contrasting structures between the decoupled and
1026 coupled states of the stable boundary layer. *Quarterly Journal of the Royal*
1027 *Meteorological Society*, *142*(695), 693–702. doi: 10.1002/qj.2693
- 1028 Baas, P., van De Wiel, B., van der Linden, S., & Bosveld, F. (2018). From Near-
1029 Neutral to Strongly Stratified: Adequately Modelling the Clear-Sky Nocturnal
1030 Boundary Layer at Cabauw. *Boundary-Layer Meteorology*, *166*, 217–238. doi:
1031 10.1007/s10546-017-0304-8
- 1032 Bazile, E., Couvreux, F., Le Moigne, P., & Genthon, C. (2015). First Workshop on
1033 the GABLS-4 Intercomparison. *GEWEX News*, *25*(3).
- 1034 Bazile, E., Couvreux, F., Le Moigne, P., Genthon, C., Holtslag, A. A. M., & Svens-
1035 son. (2014). GABLS4: An intercomparison case to study the stable boundary
1036 layer over the Antarctic Plateau. *GEWEX News*, *24*(4).
- 1037 Beare, R. J., Macvean, M. K., Holtslag, A. A. M., Cuxart, J., Esau, I., Golaz, J.-C.,
1038 ... Sullivan, P. (2006). An Intercomparison of Large-Eddy Simulations of the
1039 Stable Boundary Layer. *Boundary-Layer Meteorology*, *118*(2), 247–272. doi:
1040 10.1007/s10546-004-2820-6
- 1041 Beljaars, A., & Viterbo, P. (1998). Role of the boundary layer in a numerical
1042 weather prediction model. *Clear and cloudy boundary layers*, 287–304.
- 1043 Blackadar, A. K. (1957). Boundary layer wind maxima and their significance for the
1044 growth of nocturnal inversions. *Bull. Amer. Meteor. Soc.*, *38*(5), 283–290.
- 1045 Bosveld, F. C., Baas, P., Steeneveld, G.-J., Holtslag, A. A. M., Angevine, W. M.,
1046 Bazile, E., ... Svensson, G. (2014). The Third GABLS Intercomparison
1047 Case for Evaluation Studies of Boundary-Layer Models. Part B: Results and
1048 Process Understanding. *Boundary-Layer Meteorology*, *152*(2), 157–187. doi:
1049 10.1007/s10546-014-9919-1
- 1050 Bougeault, P., & Lacarrère, P. (1989). Parameterization of Orography-Induced Tur-
1051 bulence in a Mesobeta-Scale Model. *Monthly Weather Review*, *117*(8), 1872–
1052 1890.
- 1053 Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M.,
1054 ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal*
1055 *of Statistical Software, Articles*, *76*(1), 1–32. doi: 10.18637/jss.v076.i01
- 1056 Cheng, Y., Canuto, V., & Howard, A. (2002). An improved model for the turbulent
1057 PBL. *Journal of the Atmospheric sciences*, *59*(9), 1550–1565.
- 1058 Couvreux, F., Bazile, E., Rodier, Q., Maronga, B., Matheou, G., Chinita, M. J.,
1059 ... Vignon, E. (2020). Intercomparison of Large-Eddy Simulations of the
1060 Antarctic Boundary Layer for Very Stable Stratification. *Boundary-Layer*
1061 *Meteorology*, *176*(3), 369–400. doi: 10.1007/s10546-020-00539-4
- 1062 Couvreux, F., Hourdin, F., Williamson, D., Roehrig, R., Volodina, V., Villefranque,
1063 N., ... Xu, W. (2020). Process-based climate model development harness-
1064 ing machine learning: I. a calibration tool for parameterization improve-
1065 ment. *Journal of Advances in Modeling Earth Systems*, e2020MS002217.
1066 doi: 10.1029/2020MS002217
- 1067 Cuxart, J., Bougeault, P., & Redelsperger, J.-L. (2000). A turbulence scheme allow-
1068 ing for mesoscale and large-eddy simulations. *Quarterly Journal of the Royal*
1069 *Meteorological Society*, *126*(562), 1–30.
- 1070 Cuxart, J., Holtslag, A. A. M., Beare, R. J., Bazile, E., Beljaars, A., Cheng, A.,
1071 ... Xu, K.-M. (2006). Single-Column Model Intercomparison for a Stably
1072 Stratified Atmospheric Boundary Layer. *Boundary-Layer Meteorology*, *118*(2),
1073 273–303. doi: 10.1007/s10546-005-3780-1

- 1074 Derbyshire, S. H. (1999). Boundary-layer decoupling over cold surfaces as a physical
1075 boundary-instability. *Boundary-Layer Meteorology*, *90*(2), 297–325.
- 1076 Gallée, H., Barral, H., Vignon, E., & Genthon, C. (2015). A case study of a low-level
1077 jet during OPALE. *Atmospheric Chemistry and Physics*, *15*(11), 6237–6246.
1078 doi: 10.5194/acp-15-6237-2015
- 1079 Gettelman, A., Truesdale, J. E., Bacmeister, J. T., Caldwell, P. M., Neale, R. B.,
1080 Bogenschutz, P. A., & Simpson, I. R. (2019). The Single Column Atmosphere
1081 Model Version 6 (SCAM6): Not a Scam but a Tool for Model Evaluation and
1082 Development. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1381–
1083 1401. doi: 10.1029/2018MS001578
- 1084 Holtslag, A. A. M., Svensson, G., Baas, P., Basu, S., Beare, B., Beljaars, A. C. M.,
1085 ... Van De Wiel, B. J. H. (2013). Stable Atmospheric Boundary Layers
1086 and Diurnal Cycles: Challenges for Weather and Climate Models. *Bul-*
1087 *letin of the American Meteorological Society*, *94*(11), 1691–1706. doi:
1088 10.1175/BAMS-D-11-00187.1
- 1089 Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., ... Roehrig,
1090 R. (2013). LMDZ5B: the atmospheric component of the IPSL climate model
1091 with revisited parameterizations for clouds and convection. *Climate Dynamics*,
1092 *40*(9), 2193–2222. doi: 10.1007/s00382-012-1343-y
- 1093 Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q.,
1094 ... Williamson, D. (2017). The Art and Science of Climate Model Tun-
1095 ing. *Bulletin of the American Meteorological Society*, *98*(3), 589–602. doi:
1096 10.1175/BAMS-D-15-00135.1
- 1097 Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin,
1098 N., ... Ghattas, J. (2020). LMDZ6A: The Atmospheric Component of
1099 the IPSL Climate Model With Improved and Better Tuned Physics. *Jour-*
1100 *nal of Advances in Modeling Earth Systems*, *12*(7), e2019MS001892. doi:
1101 10.1029/2019MS001892
- 1102 Hourdin, F., Williamson, D., Rio, C., Couvreux, F., Roehrig, R., Villefranche,
1103 N., ... Volodina, V. (2020). Process-based climate model development
1104 harnessing machine learning: II. model calibration from single column to
1105 global. *Journal of Advances in Modeling Earth Systems*, e2020MS002225. doi:
1106 10.1029/2020MS002225
- 1107 Mahrt, L. (1998). Stratified atmospheric boundary layers and breakdown of models.
1108 *Theoretical and computational fluid dynamics*, *11*(3), 263–279.
- 1109 Mascart, P., Noilhan, J., & Giordani, H. (1995). A modified parameterization of
1110 flux-profile relationships in the surface layer using different roughness length
1111 values for heat and momentum. *Boundary-Layer Meteorology*, *72*(4), 331–344.
1112 doi: 10.1007/BF00708998
- 1113 Mauritsen, T., & Svensson, G. (2007). Observations of stably stratified shear-driven
1114 atmospheric turbulence at low and high Richardson numbers. *Journal of the*
1115 *atmospheric sciences*, *64*(2), 645–655.
- 1116 Neggers, R. A. J. (2015). Exploring bin-macrophysics models for moist convective
1117 transport and clouds. *Journal of Advances in Modeling Earth Systems*, *7*(4),
1118 2079–2104. Retrieved from [https://agupubs.onlinelibrary.wiley.com/](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015MS000502)
1119 [doi/abs/10.1002/2015MS000502](https://doi.org/10.1002/2015MS000502) doi: 10.1002/2015MS000502
- 1120 Noilhan, J., & Mahfouf, J.-F. (1996). The ISBA land surface parameterisation
1121 scheme. *Global and Planetary Change*, *13*(1), 145 - 159. doi: [https://doi.org/](https://doi.org/10.1016/0921-8181(95)00043-7)
1122 [10.1016/0921-8181\(95\)00043-7](https://doi.org/10.1016/0921-8181(95)00043-7)
- 1123 Paulson, C. A. (1970). The Mathematical Representation of Wind Speed and Tem-
1124 perature Profiles in the Unstable Atmospheric Surface Layer. *Journal of Ap-*
1125 *plied Meteorology*, *9*(6), 857–861.
- 1126 Pukelsheim, F. (1994). The Three Sigma Rule. *The American Statistician*, *48*(2),
1127 88–91.
- 1128 Randall, D., Krueger, S., Bretherton, C., Curry, J., Duynkerke, P., Moncrieff, M., ...

- 1129 Wielicki, B. (2003). Confronting Models with Data: The GEWEX Cloud Sys-
 1130 tems Study. *Bulletin of the American Meteorological Society*, 84(4), 455–470.
 1131 doi: 10.1175/BAMS-84-4-455
- 1132 Randall, D., Xu, K.-M., Somerville, R. J., & Iacobellis, S. (1996). Single-column
 1133 models and cloud ensemble models as links between observations and climate
 1134 models. *Journal of Climate*, 9(8), 1683–1697.
- 1135 Redelsperger, J.-L., & Sommeria, G. (1982). Method of representing the turbulence
 1136 associated with precipitations in a three-dimensional model of cloud convec-
 1137 tion. *Boundary-Layer Meteorology*, 24(2), 231–252.
- 1138 Redelsperger, J.-L., & Sommeria, G. (1986). Three-dimensional simulation of a
 1139 convective storm: Sensitivity studies on subgrid parameterization and spatial
 1140 resolution. *Journal of the atmospheric sciences*, 43(22), 2619–2635.
- 1141 Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Gu er emy, J.-F., ...
 1142 S en esi, S. (2020). The CNRM Global Atmosphere Model ARPEGE-Climat 6.3:
 1143 Description and Evaluation. *Journal of Advances in Modeling Earth Systems*,
 1144 12(7), e2020MS002075. doi: <https://doi.org/10.1029/2020MS002075>
- 1145 Salter, J. M., & Williamson, D. (2016). A comparison of statistical emulation
 1146 methodologies for multi-wave calibration of environmental models. *Environ-*
 1147 *metrics*, 27(8), 507–523. doi: 10.1002/env.2405
- 1148 Salter, J. M., Williamson, D., Scinocca, J., & Kharin, V. (2019). Uncertainty Quan-
 1149 tification for Computer Models with spatial output using calibration-optimal
 1150 bases. *Journal of the American statistical association*, 114(528), 1800–1814.
 1151 doi: 10.1080/01621459.2018.1514306
- 1152 Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T., & Balsamo, G. (2013). Why
 1153 is it so difficult to represent stably stratified conditions in numerical weather
 1154 prediction (NWP) models? *Journal of Advances in Modeling Earth Systems*,
 1155 5(2), 117–133.
- 1156 S ef erian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., ...
 1157 Madec, G. (2019). Evaluation of CNRM Earth System Model, CNRM-
 1158 ESM2-1: Role of Earth System Processes in Present-Day and Future Climate.
 1159 *Journal of Advances in Modeling Earth Systems*. doi: 10.1029/2019MS001791
- 1160 Simmons, A. J., & Burridge, D. M. (1981). An Energy and Angular-Momentum
 1161 Conserving Vertical Finite-Difference Scheme and Hybrid Vertical Coordinates.
 1162 *Monthly Weather Review*, 109(4), 758–766.
- 1163 Steeneveld, G.-J., Holtslag, A., Nappo, C., Van de Wiel, B., & Mahrt, L. (2008). Ex-
 1164 ploring the possible role of small-scale terrain drag on stable boundary layers
 1165 over land. *Journal of applied meteorology and climatology*, 47(10), 2518–2530.
- 1166 Steeneveld, G.-J., Van de Wiel, B., & Holtslag, A. (2006). Modeling the evolu-
 1167 tion of the atmospheric boundary layer coupled to the land surface for three
 1168 contrasting nights in CASES-99. *Journal of the atmospheric sciences*, 63(3),
 1169 920–935.
- 1170 Svensson, G., Holtslag, A. A. M., Kumar, V., Mauritsen, T., Steeneveld, G. J.,
 1171 Angevine, W. M., ... Zampieri, M. (2011). Evaluation of the Diurnal Cycle
 1172 in the Atmospheric Boundary Layer Over Land as Represented by a Variety
 1173 of Single-Column Models: The Second GABLS Experiment. *Boundary-Layer*
 1174 *Meteorology*, 140(2), 177–206. doi: 10.1007/s10546-011-9611-7
- 1175 Tsiringakis, A., Steeneveld, G. J., & Holtslag, A. A. M. (2017). Small-scale oro-
 1176 graphic gravity wave drag in stable boundary layers and its impact on synoptic
 1177 systems and near-surface meteorology. *Quarterly Journal of the Royal Meteoro-*
 1178 *logical Society*, 143(704), 1504–1516.
- 1179 Van de Wiel, B. J., Moene, A., & Jonker, H. (2012). The cessation of continuous
 1180 turbulence as precursor of the very stable nocturnal boundary layer. *Journal of*
 1181 *the Atmospheric Sciences*, 69(11), 3097–3115.
- 1182 Vernon, I., Goldstein, M., & Bower, R. G. (2010, 12). Galaxy formation: a Bayesian
 1183 uncertainty analysis. *Bayesian Analysis*, 5(4), 619–669.

- 1184 Vignon, E., Hourdin, F., Genthon, C., Gallee, H., Bazile, E., Lefebvre, M.-P., ...
1185 Van de Wiel, B. J. (2017). Antarctic Boundary-Layer Parametrization in
1186 a General Circulation Model: 1D simulations facing summer observations at
1187 Dome C. *Journal of Geophysical Research: Atmospheres*.
- 1188 Vignon, E., Hourdin, F., Genthon, C., Van de Wiel, B. J. H., Gallée, H., Madeleine,
1189 J.-B., & Beaumet, J. (2018). Modeling the Dynamics of the Atmospheric
1190 Boundary Layer Over the Antarctic Plateau With a General Circulation
1191 Model. *Journal of Advances in Modeling Earth Systems*, 10(1), 98-125. doi:
1192 10.1002/2017MS001184
- 1193 Vignon, E., van de Wiel, B. J. H., van Hooijdonk, I. G. S., Genthon, C., van der
1194 Linden, S. J. A., van Hooft, J. A., ... Casasanta, G. (2017). Stable boundary-
1195 layer regimes at Dome C, Antarctica: observation and analysis. *Quarterly*
1196 *Journal of the Royal Meteorological Society*, 143(704), 1241-1253.
- 1197 Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M.,
1198 ... Waldman, R. (2019). Evaluation of CMIP6 DECK Experiments With
1199 CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, 11(7), 2177-
1200 2213. doi: 10.1029/2019MS001683
- 1201 Williamson, D. (2015). Exploratory ensemble designs for environmental models us-
1202 ing k-extended Latin Hypercubes. *Environmetrics*, 26(4), 268-283. doi: 10
1203 .1002/env.2335
- 1204 Williamson, D., Blaker, A., Hampton, C., & Salter, J. (2015, 9). Identifying and re-
1205 moving structural biases in climate models with history matching. *Climate Dy-*
1206 *namics*, 45, 1299-1324.
- 1207 Williamson, D., Blaker, A. T., & Sinha, B. (2017). Tuning without over-tuning:
1208 parametric uncertainty quantification for the NEMO ocean model. *Geoscienti-*
1209 *fic Model Development*, 10(4), 1789-1816. doi: 10.5194/gmd-10-1789-2017
- 1210 Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L., &
1211 Yamazaki, K. (2013, 10). History matching for exploring and reducing cli-
1212 mate model parameter space using observations and a large perturbed physics
1213 ensemble. *Climate Dynamics*, 41, 1703-1729. doi: 10.1007/s00382-013-1896-4

6.4 Compléments

6.4.1 Améliorer la visualisation de l'espace NROY

La calibration de ARPEGE-Climat dans différentes configurations sur le cas GABLS4 conduit à des espaces des paramètres très réduits (moins d'un millièmme de l'espace initial), ce qui rend la visualisation de l'espace NROY potentiellement difficile, en tout cas critique pour son interprétation. En particulier, dès les premières itérations, les grandes valeurs de **LMIN** > 4 m et de **KOZMIN** $> 5 \times 10^{-4}$ sont systématiquement éliminées. La visualisation de l'espace NROY à l'aide des projections en deux dimensions devient alors délicate du fait d'un effet d'écrasement, en particulier pour le paramètre **KOZMIN** pour lequel l'espace NROY est réduit à une fine bande représentant un dixième de la plage initiale de ce paramètre. L'outil dans sa version actuelle ne permet pas en effet de zoomer facilement sur une région d'intérêt et la structure de l'espace finalement conservé peut être difficile à visualiser si cet espace est de taille très réduite.

Pour palier à cette limitation technique actuelle de l'outil et mettre en avant l'intérêt d'une possibilité de zoom dans la visualisation de l'espace NROY, une nouvelle expérience de calibration a été réalisée en utilisant la configuration SCM-LR-TS sur GABLS4 stage-3-10hrs mais avec une plage initiale réduite pour les deux paramètres **LMIN** et **KOZMIN**. Elle est proche de la plage finale à l'issue de la calibration du modèle présentée dans l'article, soit $[0, 4]$ et $[0, 5 \times 10^{-4}]$ respectivement. La méthode est arrêtée à la neuvième itération et la figure 6.1 représente l'espace NROY ainsi obtenu. Alors que l'expérience initiale semblait montrer que les valeurs les plus faibles de ces deux paramètres étaient systématiquement conservées, la visualisation de l'espace NROY issu de cette nouvelle expérience met en évidence davantage de détails. En particulier pour **KOZMIN** les valeurs trop faibles de ce paramètre sont également systématiquement rejetées. Pour ce paramètre, des valeurs autour de 2.5×10^{-4} semblent plus adaptées. On observe un résultat similaire, bien que moins marqué pour le paramètre **LMIN**. Cette expérience montre que même sur un cas aussi stable que GABLS4, il est nécessaire de conserver une borne inférieure (certes nettement plus faible que dans le réglage initial) pour le mélange turbulent (et dans une moindre mesure pour la longueur de mélange). Rappelons également que l'action du paramètre **KOZMIN**, par construction, dépend de la résolution verticale du modèle et que, dans la version HR, la sensibilité du modèle à ce paramètre est négligeable.

Cette expérience montre donc une piste d'amélioration de l'outil pour lui donner plus de souplesse et pouvoir effectuer des zooms de l'espace NROY quand celui-ci est très réduit sans avoir à relancer une série complète d'itérations .

6.4.2 Impact des nouveaux réglages sur le cycle diurne de GABLS4

La calibration du modèle a été réalisée à partir de la configuration stage3-10hrs de GABLS4, qui correspond à la partie stable du cycle diurne uniquement. Les réglages proposés alors par les outils d'*History Matching* réduisent de façon importante les paramètres **LMIN** et **KOZMIN** qui représentent des limites inférieures pour la longueur de mélange

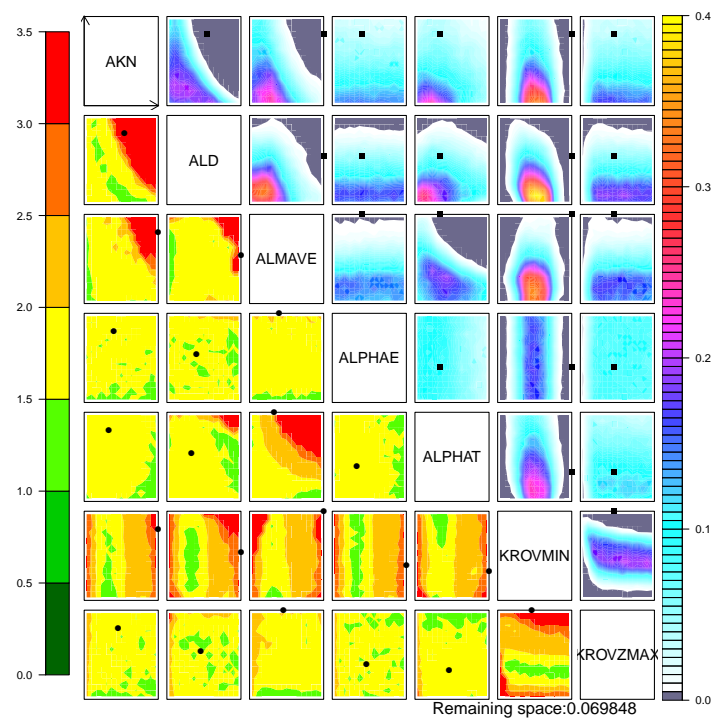


FIGURE 6.1 – Représentation de l’espace des paramètres à l’issue de la neuvième itération de la méthode de tuning statistique pour une gamme initiale des paramètres **L**MIN et **K**OZMIN réduite. Les noms des paramètres apparaissant sur la figure sont les noms des variables dans le code : AKN = CM, ALD = CE, ALMAVE = LMIN, ALPHAE = AE, ALPHAT = AT, KROVMIN = KOZMIN, KROVZMAX = ZMAX

et pour le mélange turbulent (via les coefficients d'échange). Ces seuils ne sont atteints que dans des situations stables voire très stables et n'ont donc d'impact que sur la représentation de la partie nocturne du cycle diurne de GABLS4. Le chapitre suivant (chapitre 7) évaluera l'impact de ces réglages sur un cas de couche limite de stabilité plus modérée. Les résultats proposent également une diminution des paramètres **CM** et **CE** et dans une moindre mesure **AT**. A la différence de **LMIN** et **KOZMIN**, une modification de ces paramètres a un impact potentiel en dehors des cas stables. Si les valeurs initiales de ces trois paramètres ne sont pas systématiquement rejetées, les combinaisons retenues conduisent à des mélanges turbulents plus faibles et les modifications proposées peuvent donc avoir également un impact sur la partie convective du cycle diurne. Alors que l'évaluation initiale de CNRM-CM6 sur le cycle diurne complet de GABLS4 montrait que la CLA convective qui se développe sur les dix premières heures du cas était correctement représentée, nous voulons vérifier ici que ce bon comportement du modèle n'a pas été dégradé par les nouveaux réglages proposés. Cent nouvelles simulations 1D sur le cas GABLS4-Stage3 complet ont donc été réalisées à partir des résultats présentés dans l'article pour la configuration SCM-LR-TS.

La figure 6.2 présente l'évolution de la température potentielle et du module du vent aux trois premiers niveaux du modèle. En ce qui concerne la température, les nouvelles simulations présentent une dispersion négligeable sur les dix premières heures et un comportement quasi identique à CNRM-CM6. Les profils de température et de vent à 1000 LT des cent simulations sur GABLS4-Stage-3 sont très proches du profil initial du cas GABLS4-Stage-3-10hr et la représentation de la partie stable de GABLS4-Stage-3 n'est donc pas modifiée. Concernant le flux de chaleur sensible à la surface et pour la vitesse de friction, on note une des valeurs légèrement supérieures pendant la phase diurne (voir figure 6.3), mais le comportement reste très proche de CNRM-CM6.

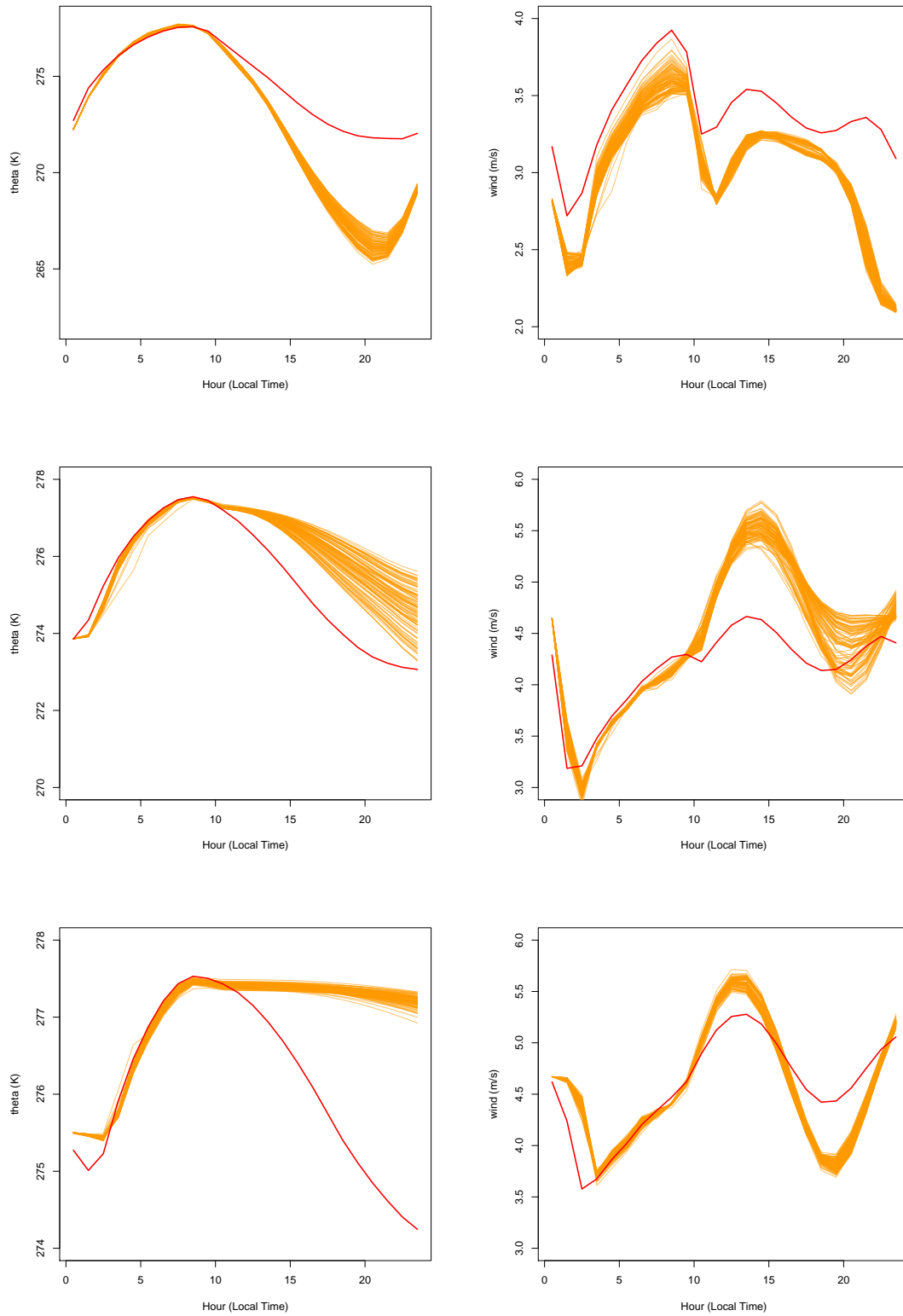


FIGURE 6.2 – Évolution de la température potentielle (colonne de gauche et du module du vent (colonne de droite) sur de haut en bas les trois premiers niveaux du modèle (respectivement ~ 8 m, ~ 29 m et ~ 55 m). Les lignes oranges correspondent aux simulations issues de la calibration et les lignes rouges à la simulation CNRM-CM6.

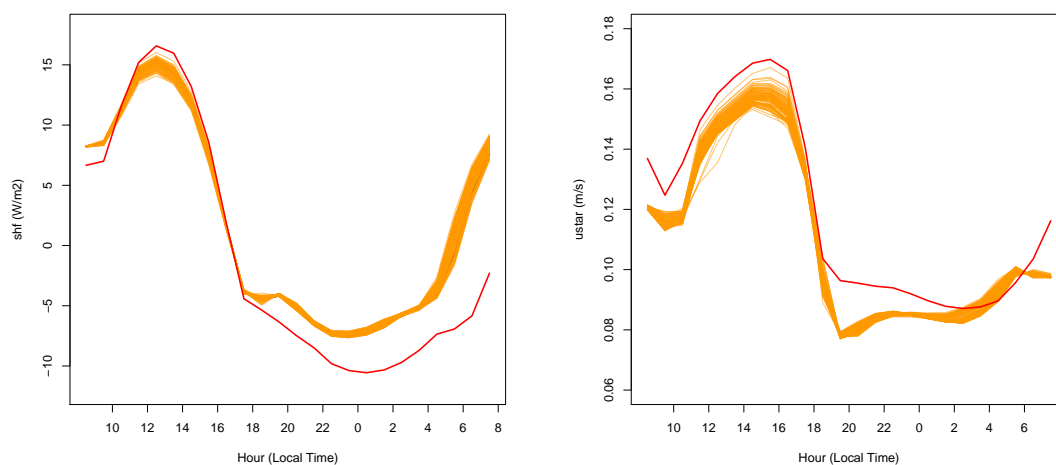


FIGURE 6.3 – Évolution du flux de chaleur sensible à la surface (à gauche) et de la vitesse de friction (à droite). Les lignes oranges correspondent aux simulations issues de la calibration et les lignes rouges à la simulation CNRM-CM6.

Le comportement du modèle semble donc quasiment inchangé pendant la période convective du cycle diurne. Pourtant la calibration statistique a produit des valeurs sensiblement différentes pour les paramètres **CM**, **CE** et **AT** qui jouent un rôle à la fois dans l'évolution de la TKE et dans le calcul des flux turbulents de température et de quantité de mouvement. Dans CNRM-CM6 les effets dus à la paramétrisation de la turbulence dominent nettement ceux dus à la paramétrisation de la convection (figure 6.4), que ce soit pour la température ou pour les composantes zonale et méridienne du vent. Ce résultat reste inchangé au premier niveau du modèle pour les cent simulations : sur la partie convective, les tendances dues à la turbulence sont peu affectées par les modifications et il n'y a pas de rétroaction sur l'activité du schéma de convection. Sur le deuxième niveau, les effets de la turbulence sur la température et le vent sont plus importants par rapport à ceux observés sur CNRM-CM6 mais sont compensés par une activation plus importante du schéma de convection avec des tendances opposées. Sur le troisième niveau du modèle, la modification des tendances dues au schéma de turbulence induit également une augmentation des tendances dues au schéma de convection mais l'équilibre qui en résulte est plus subtil et ne permet pas d'interprétation simple. On se contentera de constater que les deux paramétrisations jouent de concert sur la couche limite convective de GABLS4 et conduisent de façon remarquable toujours au même équilibre.

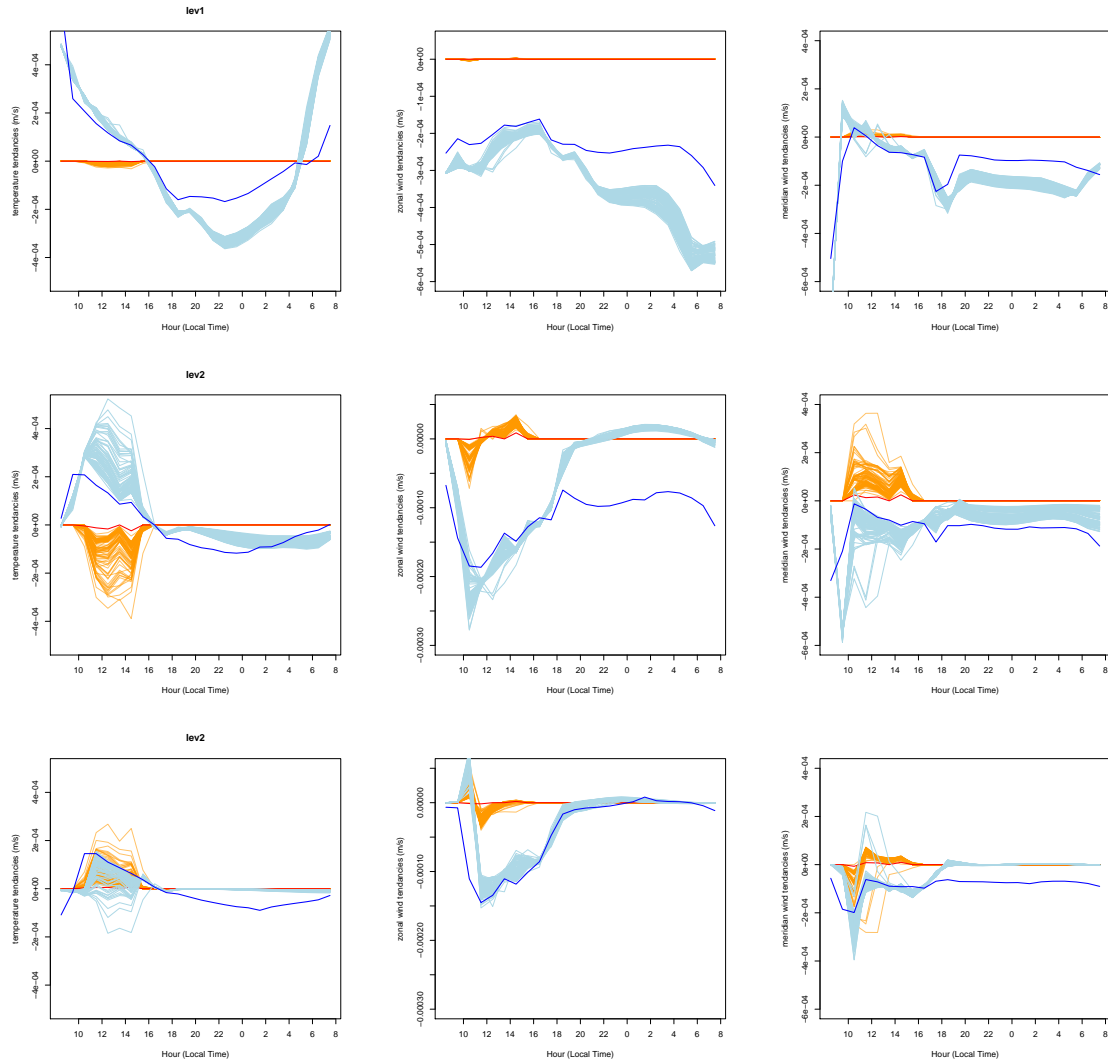


FIGURE 6.4 – Comparaison des tendances de température (première colonne), du vent zonal (au centre) et du vent méridien (à droite) sur GABLS4-Stage-3. La première ligne présente les tendances au premier niveau du modèle (~ 8 m), la deuxième ligne les tendances au deuxième niveau du modèle (~ 29 m) et la troisième ligne les tendances au troisième niveau du modèle (~ 55 m). Les tendances dues à la turbulence sont en bleu clair et celles dues à la convection sont en orange pour les cents simulations issues de la calibration du modèle. Les lignes bleues (resp. rouges) représentent les tendances dues à la turbulence (resp. à la convection) de CNRM-CM6.

6.5 Conclusions

L'application des outils d'*History Matching* pour estimer la sensibilité du modèle aux paramètres de la turbulence et les calibrer a permis de montrer qu'ARPEGE-Climat contient la physique nécessaire pour représenter des couches limites très stratifiées, même à la résolution verticale un peu grossière de la configuration CMIP6 (comparée à celles des LES utilisées pour l'exercice GABLS4). Une des conditions pour que le modèle se comporte de façon satisfaisante a été de réduire de façon significative les seuils introduits dans la paramétrisation pour maintenir un mélange turbulent minimal dans les CLS. Ces seuils ont été introduits entre autres pour palier la mauvaise représentation d'un certain nombre de processus dans les modèles numériques. Sandu et al. (2013) montre que pour le modèle IFS du CEPMMT, ces seuils restent nécessaires pour maintenir les performances du modèle à l'échelle synoptique. Pour la version PNT de ARPEGE, une analyse de sensibilité du modèle au paramètre **LMIN** a montré que les scores du modèle pour la circulation de grande échelle à 500 hPa étaient nettement améliorés pour une valeur supérieure à 50 m (communication personnelle de Laurent Descamps). Une telle valeur conduit toutefois à une très mauvaise représentation des couches limites stables. Le choix a été fait pour ARPEGE-PNT de donner la priorité à la représentation de la couche limite en fixant **LMIN** à zéro. Une autre raison pour le maintien de ces seuils est d'éviter un découplage complet de l'atmosphère avec la surface dans les cas de forte stratification, ce qui pourrait engendrer des phénomènes de "runaway cooling" (fort refroidissement en surface) dans les hautes latitudes, principalement pendant la nuit polaire. Pour ces différentes raisons, les réglages proposés à l'issue de la calibration du modèle par des méthodes d'*History Matching* demandent donc à être évalués sur d'autres couches limites que celle de GABLS4 et sur une plus grande échelle spatiale et temporelle avec des simulations 3D plus longues. Le chapitre suivant propose justement d'évaluer le modèle ARPEGE-Climat et les réglages proposés à l'issue du travail présenté ici sur un cas de couche limite plus modéré (GABLS1).

Extension de la calibration au cas GABLS1

Sommaire

7.1	Analyse de la calibration d'ARPEGE-Climat sur le cas GABLS1	139
7.1.1	Présentation du cas GABLS1	140
7.1.2	Évaluation sur le cas GABLS1	140
7.2	Nouvelle calibration du modèle combinant les cas GABLS4 et GABLS1142	
7.3	Analyse de la nouvelle calibration sur le cas GABLS1	143
7.4	Conclusion	146

Le chapitre précédent proposait une application de la méthode de calibration statistique de type *History Matching* (présentée dans le chapitre 5) au cas de couche limite très stable GABLS4. Cette calibration propose des réglages du modèle conduisant à un mélange turbulent plus faible, via notamment une diminution très sensible des valeurs de différents seuils visant à maintenir un mélange turbulent minimal (voir 3.2.2). Le présent chapitre se propose d'évaluer les réglages ainsi obtenus sur le cas GABLS1 qui correspond à une couche limite de stabilité plus modérée. Cette évaluation montre que la calibration sur le seul cas de GABLS4 n'est pas suffisante. L'expérience de calibration est alors poursuivie afin d'intégrer des métriques calculées sur GABLS1. De nouveaux réglages du modèle sont obtenus qui permettent ainsi de représenter une gamme plus large de CLS.

7.1 Analyse de la calibration d'ARPEGE-Climat sur le cas GABLS1

À l'issue de la calibration d'ARPEGE-Climat sur le cas GABLS4, des nouveaux réglages du schéma de turbulence ont été proposés. Le résultat majeur de cette calibration est la réduction des bornes **L**MIN et **K**OZMIN. Ces bornes, introduites pour éviter les phénomènes de découplages de l'atmosphère avec la surface (et de runaway cooling), ne permettent pas dans leur réglage initial, de représenter la couche très stable de GABLS4. Une réduction drastique des valeurs de ces bornes est donc nécessaire et permet d'obtenir un réglage satisfaisant du modèle.

Ces limites n'étant activées que lorsque le mélange turbulent est faible, et donc pour des couches limites très stables, leur modification ne devrait pas impacter des cas de couches

limites de stabilité modérée. Le réglage des autres paramètres, en particulier **CM** et **CE** qui gouvernent respectivement les termes de production dynamique et de dissipation du bilan de TKE, conduisent à un mélange turbulent plus faible pour tout type de couche limite. On a déjà montré (section 6.4.2) que la représentation de la couche limite convective qui se développe lors de la première partie du cycle diurne de GABLS4, n'est pas détériorée par les nouveaux réglages, le schéma de convection réagissant de façon à compenser la différence de comportement du schéma de turbulence. La question reste ouverte pour une couche limite stable pour laquelle aucune compensation par le schéma de convection n'est possible (ce dernier n'étant activé que sur des cas instables) et c'est le sujet de cette section.

7.1.1 Présentation du cas GABLS1

Le cas GABLS1 (Beare et al., 2006) est le premier cas d'étude de couches limites stables proposé par le projet GABLS (voir section 3.5.2.1 du chapitre 3). Comme le cas GABLS4 stage 3 qui a été utilisé pour calibrer le schéma de turbulence de ARPEGE-Climat, GABLS1 est un cas idéalisé, sans humidité, sans advection ni rayonnement. Il a été construit à partir d'observations réalisées en Arctique et présente une couche limite stable nocturne faiblement stratifiée. C'est un cas de référence qui a été utilisé pour le développement de nombreuses paramétrisations. Le tableau 7.1 présente les principales caractéristiques de GABLS1 et les compare avec celle de GABLS4 stage 3-10hr. Le profil initial de température potentielle est constant ($\theta = 265$ K) jusqu'à 100 m, puis croissant avec un gradient constant de 0.01 Km^{-1} entre 100 et 400 m. Au bout des 9h de simulation, la couche limite présente une hauteur de 150 à 200 m selon les différentes simulations LES, avec un jet nocturne bien établi d'une intensité $\simeq 9.5 \text{ ms}^{-1}$.

	GABLS1	GABLS4 stage 3-10hr
Latitude	73°N	75°06'S
Paramètre de Coriolis	$f = 1.39.10^{-4} \text{ s}^{-1}$	$f = 1.4.10^{-4} \text{ s}^{-1}$
Humidité	Non	Non
Advection	Non	Non
Rayonnement	Non	Non
Vent géostrophique	$u_g = 8.0 \text{ m s}^{-1}$	$u_g = 1.25 \text{ m s}^{-1}$ et $v_g = 4.5 \text{ m s}^{-1}$
Refroidissement	constant (-0.25 K h^{-1})	évolutif (maximum -2.4 K h^{-1})
longueurs de rugosité	$z_0 = z_{0h} = 0.1 \text{ m}$	$z_0 = 10^{-3} \text{ m}$, $z_{0h} = 10^{-4} \text{ m}$
durée	9 h	10 h
domaine des LES	$400 \times 400 \times 400 \text{ m}^{-3}$	$500 \times 500 \times 150 \text{ m}^{-3}$

TABLE 7.1 – Principales caractéristiques des cas GABLS1 et GABLS4 stage 3-10hr.

7.1.2 Évaluation sur le cas GABLS1

Les réglages issus de la calibration sur le cas GABLS4 sont testés ici sur le cas GABLS1. Le cas GABLS1 étant un cas pour lequel le modèle est forcé par la température de surface, nous avons utilisé les résultats de la calibration d'ARPEGE-Climat sur GABLS4 dans la

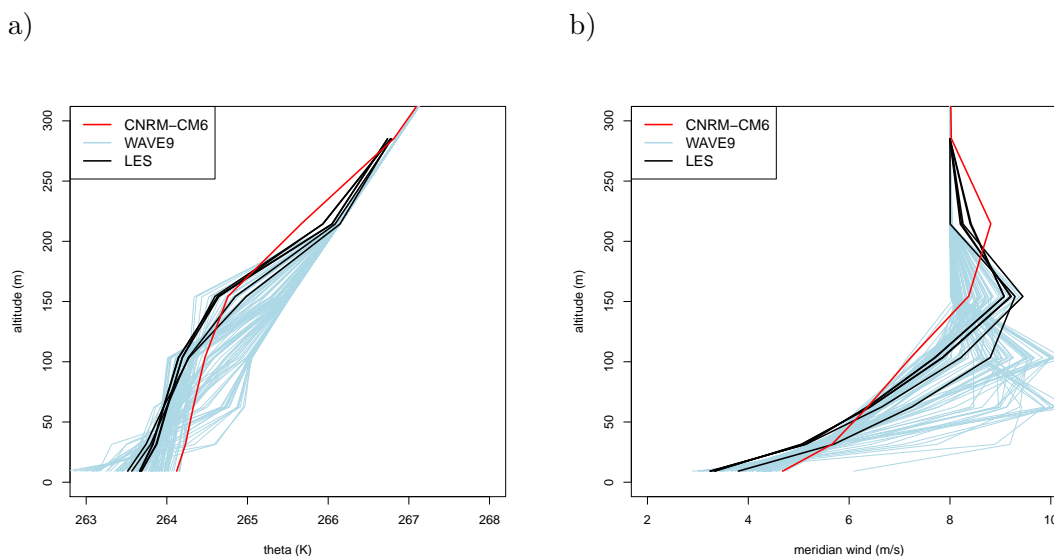


FIGURE 7.1 – Profils verticaux de température potentielle et de vent après 8 h de simulation sur le cas GABLS1. Les simulations sont issues de l'espace NROY de la vague 8 de la calibration de GABLS4. Les LES ayant une résolution verticale de 2 m, leurs profils verticaux ont été projetés (par une interpolation linéaire) sur la grille verticale d'ARPEGE-Climat pour être comparés aux résultats des SCMs.

configuration SCM-LR-TS (voir section 6.4). Plus précisément, cent points de l'espace NROY obtenus à l'issue de la vague 8 de la calibration sur GABLS4 ont été tirés de façon aléatoire et ont été utilisés pour réaliser une vague 9 incluant GABLS4 et GABLS1. La figure 7.1 présente les profils verticaux de température potentielle et du module du vent au bout de huit heures de simulation sur le cas GABLS1. Alors que sur GABLS4, les différentes simulations à l'issue de la calibration étaient très homogènes, les résultats sur GABLS1 sont beaucoup plus dispersés. Si sur certains aspects on peut constater une certaine amélioration (un refroidissement au premier niveau du modèle plus cohérent avec les résultats des LES et une correction du biais de la hauteur de la couche limite même si pour certaines simulations cette correction est trop forte), les nouveaux réglages, en diminuant le mélange turbulent, conduisent cependant souvent à des épaisseurs de couches limites trop faibles. Ce résultat est particulièrement visible sur le profil de vent : un certain nombre de simulations présentent un maximum de vent à des altitudes de 63 et 104 m (soit les troisième et quatrième niveaux du modèle) alors que dans les LES, ce maximum est atteint à 154 m (le cinquième niveau du modèle).

De cette évaluation rapide on peut retenir essentiellement deux conclusions. D'une part, la dispersion des résultats des simulations sur le cas GABLS1 montre que la calibration réalisée sur le seul cas GABLS4 ne permet pas de contraindre suffisamment le modèle. D'autre part, un certain nombre des réglages proposés conduisent à une réduction trop importante du mélange turbulent et donc à des couches limites trop minces quand ils sont appliqués sur le cas GABLS1. On note cependant que parmi les cent simulations qui ont été effectuées, certaines proposent des comportements satisfaisants.

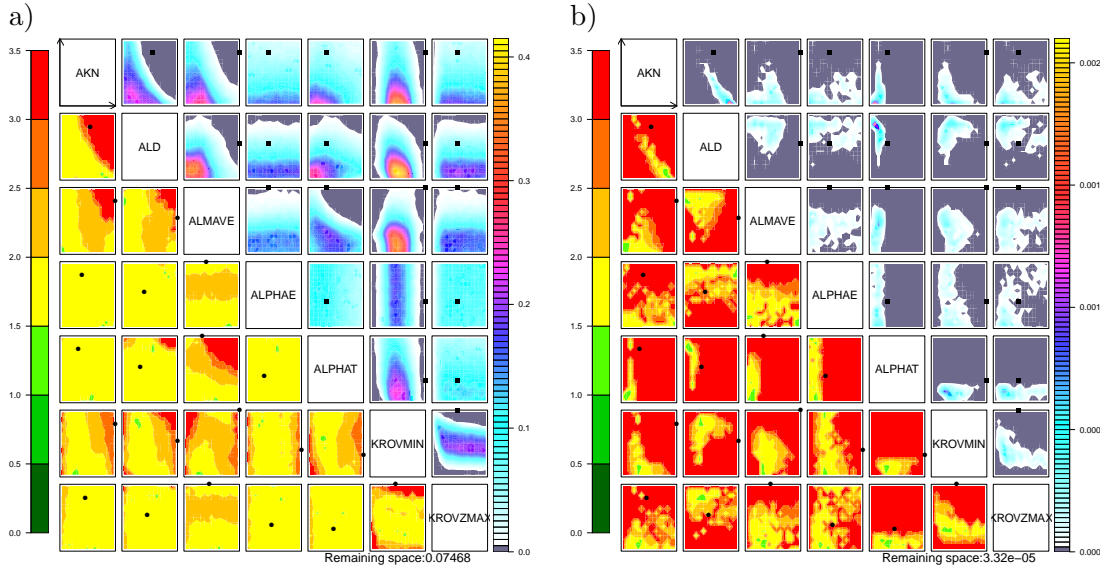


FIGURE 7.2 – À gauche, espace NROY à l’issue de la vague 8 sur le cas GABLS4, forcé en température de surface. À droite, espace NROY à l’issue de la vague 10, avec ajout des métriques calculées sur le cas GABLS1 sur les deux dernières itérations. Les noms des paramètres apparaissant sur la figure sont les noms des variables dans le code : AKN = CM, ALD = CE, ALMAVE = LMIN, ALPHAE = AE, ALPHAT = AT, KROVMIN = KOZMIN, KROVZMAX = ZMAX

7.2 Nouvelle calibration du modèle combinant les cas GABLS4 et GABLS1

Partant de la vague 8 de la calibration d’ARPEGE-Climat en configuration SCM-LR-TS (voir figure 7.2 a), deux nouvelles itérations de la méthode de calibration par *History Matching* sont réalisées. Aux quatre métriques calculées sur GABLS4 sont ajoutées quatre nouvelles métriques calculées sur les profils verticaux de vent et de température potentielle pris à 8 h sur le cas GABLS1. Par analogie avec ce qui a été fait sur GABLS4, deux métriques ont été choisies pour contraindre l’aspect thermique (θ aux premier et cinquième niveaux du modèle), et deux métriques pour contraindre l’aspect dynamique (module du vent aux troisième et cinquième niveaux du modèle). Deux LES qui avaient été réalisées avec une résolution verticale de 2 m (Beare et al., 2006) servent de référence pour la calibration sur GABLS1.

À l’issue de ces deux vagues, l’espace des paramètres a été très significativement réduit (figure 7.2 b) passant de 7.5×10^{-2} à 3.3×10^{-5} de l’espace initial. On doit noter ici qu’il a été nécessaire d’introduire une erreur structurale du modèle $\sigma_d^2 = 1 \text{ K}$ sur les métriques sur GABLS1 concernant la température potentielle afin d’avoir un espace NROY non vide. La question de la valeur à donner à cette erreur reste délicate et on voit ici que le choix qui avait été fait de la considérer comme nulle pour la calibration sur le seul cas GABLS4 était discutable. À défaut de méthode objective pour l’évaluer, on peut conseiller pour de futures expériences de calibration, de la prendre au moins égale à l’incertitude sur la référence utilisée

(ici les LES), et éventuellement de la réduire au fil des vagues (communication personnelle de Daniel Williamson).

Les différences avec le résultat de la calibration sur GABLS4 ne sont pas uniquement quantitatives. Les métriques calculées sur GABLS1 ajoutent en particulier une contrainte sur le couple de paramètres **CM** – **CE**. On rappelle que le paramètre **CM** contrôle le flux de quantité de mouvement et donc la production dynamique de TKE. Ainsi les valeurs fortes de ce paramètre entraînent une production dynamique importante. Le paramètre **CE** contrôle le terme de dissipation dans le bilan de TKE et les valeurs fortes de ce paramètre correspondent à une dissipation faible. Alors que la calibration sur les métriques GABLS4 sélectionnait en première approximation les couples de ces paramètres situés sous la première diagonale, l'ajout des métriques GABLS1 conduit à sélectionner les couples de ces paramètres situés sur la première diagonale et donc à éliminer les combinaisons qui réduisent trop fortement le mélange turbulent.

L'autre différence importante concerne le paramètre **AT**. Le cas GABLS4 seul ne contraint que modérément ce paramètre, même si les faibles valeurs sont privilégiées (figure 7.2a). L'ajout des métriques sur GABLS1 élimine de façon systématique les valeurs les plus fortes de ce paramètre. Ce paramètre intervient dans le terme de production thermique et une faible valeur de ce paramètre se traduit par une plus faible valeur du terme de production thermique en valeur absolue. Or, dans une situation de couche limite stable comme celle de GABLS1, ce terme est négatif et correspond à un puits de TKE. La sélection des faibles valeurs de **AT** va donc lui aussi dans le sens d'une augmentation du mélange turbulent par rapport aux réglages issus de GABLS4.

Ainsi les réglages proposés en ajoutant le cas GABLS1 dans l'expérience de calibration vont dans le sens d'une augmentation du mélange turbulent par rapport aux résultats obtenus lors de la calibration sur le seul cas GABLS4. On met ainsi en évidence l'importance d'avoir une diversité de cas d'étude à disposition pour calibrer un modèle de climat.

7.3 Analyse de la nouvelle calibration sur le cas GABLS1

Nous analysons dans cette partie le comportement du modèle sur la couche limite de GABLS1 après la nouvelle calibration présentée en Section 7.2.

La figure 7.3 présente l'évolution de la température pour les deux niveaux modèle sur lesquels les métriques ont été calculées (soit les premier et cinquième niveaux, respectivement situés à environ 9.2 m et 154.4 m). Le premier constat est la réduction importante de la dispersion du comportement du modèle. Près de la surface (figure 7.3a), les simulations de la vague 11 ont tendance à refroidir trop rapidement sur les premières heures de la simulation. Au contraire sur le niveau 5, le modèle est trop chaud (figure 7.3b). Ces deux constats montrent que le modèle présente encore un mélange trop faible dans la couche limite malgré l'introduction de nouvelles métriques permettant de sélectionner des paramètres qui tendent à l'augmenter. On retrouve ce résultat sur le profil de θ après 8 heures de simulation (heure à laquelle les métriques ont été calculées). La métrique basée sur la température potentielle au premier niveau du modèle a permis de contraindre correctement le modèle, contrairement à

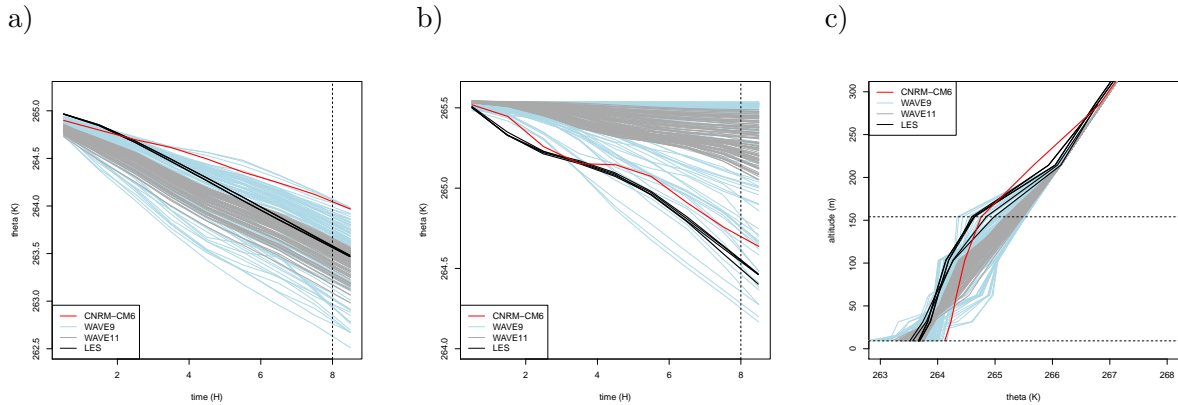


FIGURE 7.3 – Évolution de la température potentielle θ sur les niveaux 1 et 5 du modèle (respectivement 9.2 m et 154.4 m), et profil vertical après 8 heures de simulation. Les lignes pointillées verticales et horizontales représentent respectivement l’heure et les altitudes auxquelles les métriques ont été calculées.

celle sur le cinquième niveau pour lequel l’incertitude sur les LES est plus grande (voir figure 7.3c).

Le résultat de la calibration incluant les métriques calculées sur GABLS1 conduisent à une très bonne représentation du profil de vent, et en particulier de la structure du jet. La figure 7.4 présente les profils après 8 heures de simulations (heure de calcul des métriques) du module du vent (a), et des composantes zonale (b) et méridienne (c). Les métriques choisies (module du vent à 8 h) ont été calculées au troisième et cinquième niveau du modèle, soit les altitudes du maximum des composantes méridienne et zonale respectivement. Non seulement la dispersion des résultats est très réduite, mais la représentation du jet est désormais très satisfaisante. Le maximum du module du vent (et de la composante zonale) est à la bonne altitude (154.4 m, soit le cinquième niveau du modèle). Le profil de la composante méridienne présente également un maximum au même niveau que les simulations LES. Les valeurs du module du vent et de ses composantes sont cohérentes elles aussi avec les résultats des LES à tous les niveaux de la couche limite.

L’évolution du module du vent ainsi que sa rotation à différents niveaux de la couche limite tout au long de la simulation sont présentées sur la figure 7.5. Si à 63 et 104 m (*i.e.* au troisième niveau et quatrième niveau), l’évolution du module du vent dans le modèle est satisfaisante et présente une nette amélioration par rapport aux simulations issues de la calibration sur GABLS4, des différences apparaissent à 154 m (le cinquième niveau), avec en particulier un minimum de vent qui intervient trop tard dans ARPEGE-Climat par rapport à ce que font les LES. Si on s’intéresse à la rotation, les différences apparaissent dès le quatrième niveau avec une rotation trop rapide du vent dans les simulations issues de la vague 11. Cette rotation trop rapide est encore plus évidente au cinquième niveau. Cette différence de comportement ne peut pas s’expliquer avec le modèle du jet de Blackadar (Blackadar, 1957) qui suppose une extinction complète de la turbulence et prévoit une rotation entièrement déterminée par la paramètre de Coriolis. Dans CNRM-CM6, au contraire, cette rotation est trop lente.

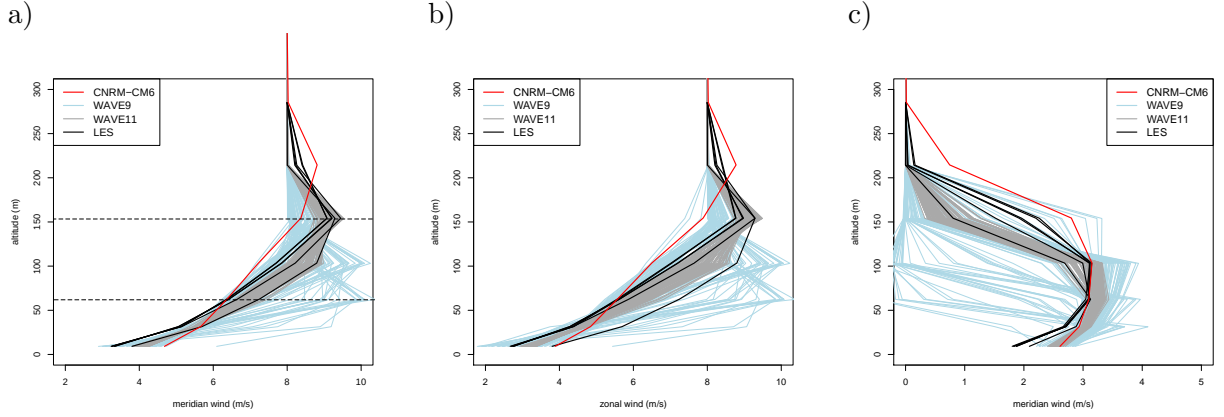


FIGURE 7.4 – Profils verticaux à 8 h du module du vent et de ses composantes zonale et méridienne. Les lignes pointillées indiquent les niveaux du modèle sur lesquels les métriques ont été calculées

Cette différence peut être reliée à la hauteur de la couche limite, que l'on peut estimer en regardant l'altitude à laquelle la TKE s'annule. La figure 7.6 montre l'évolution du profil de TKE tout au long de la simulation. Après 4 heures de simulation, des différences apparaissent entre les simulations de la vague 11 et CNRM-CM6 que l'on retrouve jusqu'à la fin de la simulation : CNRM-CM6 conserve un mélange turbulent jusqu'à une altitude de 200 m alors que dans les simulations de la vague 11, la TKE s'annule dès 150 m. Cette différence impacte la représentation du vent ou de la température. CNRM-CM6 produit encore trop de mélange à 200 m alors que les simulations de la vague 11 ne mélangent plus du tout à cette altitude. On retrouve le constat précédemment posé : la calibration de la turbulence sur le cas GABLS4 conduit à une sous-estimation du mélange turbulent sur le cas GABLS1, et l'ajout de métriques calculées sur ce dernier cas ne corrige que partiellement ce défaut. Il faut rappeler que la calibration du modèle sur le cas GABSL4 a été effectuée en prenant une tolérance à l'erreur nulle. Ajouter une tolérance à l'erreur plus grande pourrait permettre d'aboutir à des résultats plus satisfaisants sur GABLS1.

La différence observée pour la représentation du vent et de la température potentielle mériterait une investigation plus poussée. Une hypothèse pour expliquer cette différence peut être avancée en notant que les paramètres **AT** et **CM** sont reliés entre eux par le nombre de Prandtl turbulent Pr , via la formule suivante (obtenue à partir de l'équation 3.4) :

$$Pr = \frac{\mathbf{CM}}{\mathbf{AT}} \left(1 + C \frac{g}{\theta_{vl}} \frac{L_m^2}{\bar{e}} \frac{\partial \bar{\theta}_{vl}}{\partial z} \right) \quad (7.1)$$

Ce nombre sans dimension quantifie la vitesse de diffusion de la température par rapport à la diffusion dynamique. La valeur de ce nombre est approximativement de 0.8 en condition neutre et augmente avec la stabilité. Une mauvaise valeur de Pr pourrait expliquer la différence observée entre les aspects thermiques et dynamiques. Il pourrait ainsi être intéressant de refaire une expérience de calibration dans laquelle on chercherait à régler le paramètre C intervenant dans l'équation 7.1 et le \mathbf{CM}/\mathbf{AT} plutôt que de chercher à régler de façon

indépendante ces deux paramètres.

7.4 Conclusion

Cette évaluation des résultats de la calibration basée sur un cas de couche limite extrême (GABLS4) sur une couche limite plus modérée montre clairement que la calibration de la turbulence sur un cas unique n'est pas suffisant et milite pour la construction d'une batterie de cas permettant de représenter toute une diversité de couches limites stables. Des variantes du cas GABLS1 faisant varier le refroidissement imposé à la surface (de 0.35 à -1.35 Kh^{-1}) et l'intensité du vent géostrophique (de 6 à 20 ms^{-1}) ont été récemment construites au cNRM (Rodier et al., 2017) et pourraient venir peupler cette batterie de cas de couches limites stables, ainsi que les différents *stages* de GABLS4 avec la possibilité d'avoir un rayonnement et une surface interactifs et éventuellement les cas GABLS2 et GABLS3. Un tel banc d'essai servirait à de futurs exercices de calibration du modèle.

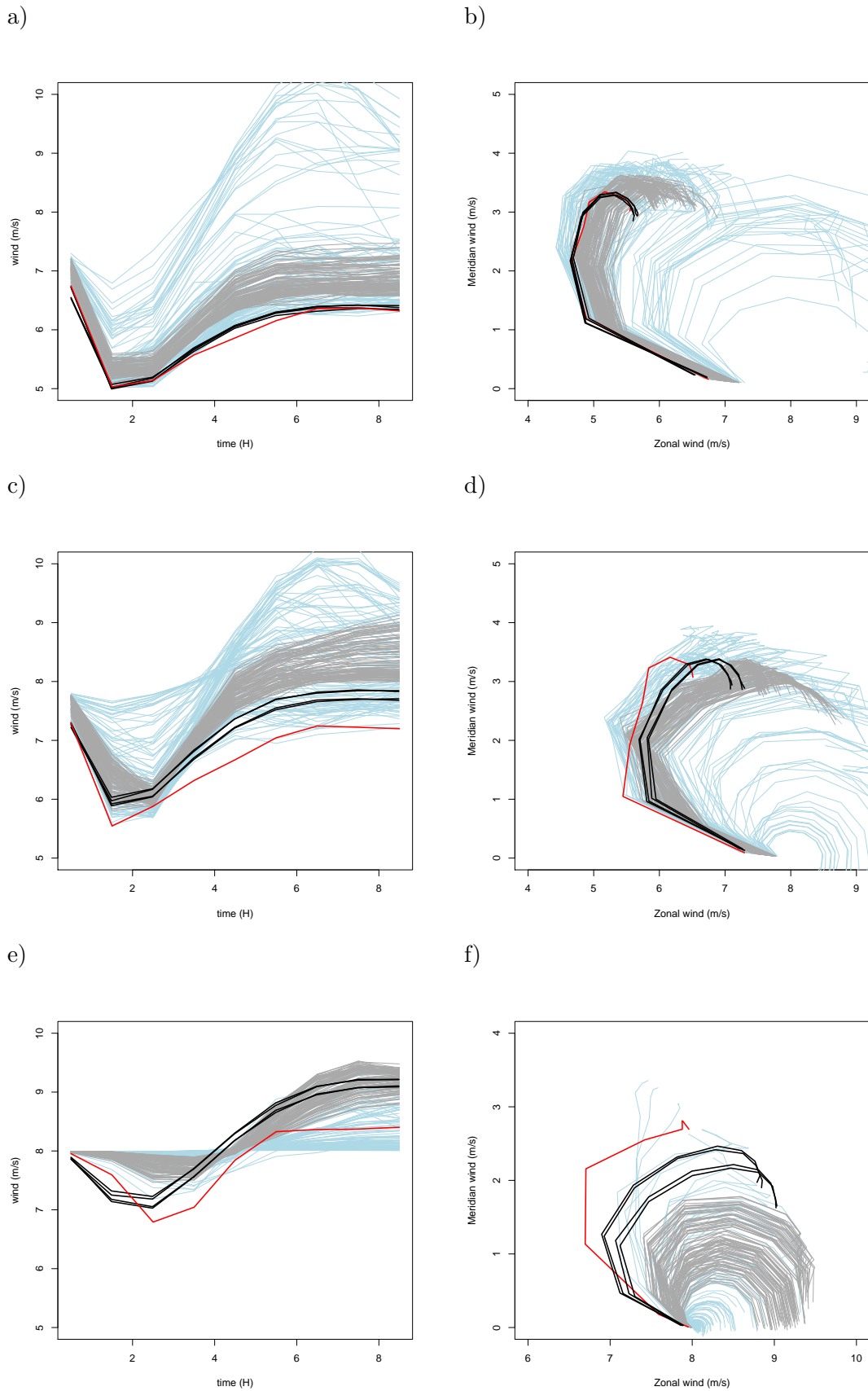


FIGURE 7.5 – a), c) et e) représentent le module du vent sur les niveaux 3, 4 et 5 du modèle (respectivement 60, 110 et 154.4 m). b), d) et f) représentent la rotation du vent sur les niveaux 3, 4 et 5 du modèle

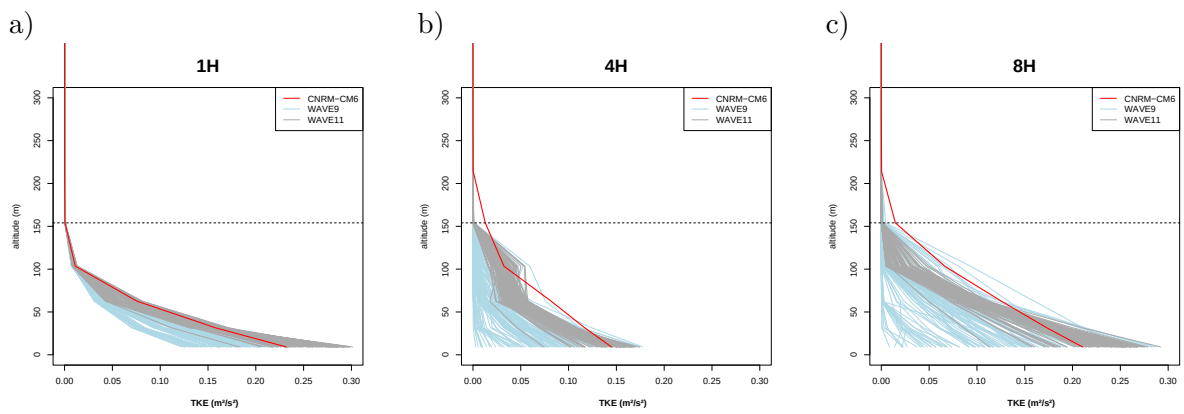


FIGURE 7.6 – Évolution du profil d'énergie cinétique turbulente sur le cas GABLS1 après 1h (a), 4h (b) et 8h(c) de simulations.

Conclusions et perspectives

Au cours de ces travaux de thèse, on s'est intéressé à la représentation des couches limites stables dans un modèle de climat. Celles-ci sont connues pour être difficiles à représenter correctement dans les modèles numériques, que ce soit les modèles de climat ou de prévision numérique du temps. Les couches limites très stratifiées, en particulier, poussent les modèles dans leurs retranchements. Les conditions météorologiques extrêmes en Antarctique en font un lieu privilégié pour leur étude. Alors que le modèle ARPEGE-Climat dans sa version GCM présente un biais froid quasi généralisé sur le plateau antarctique, il ne parvient pas à simuler le cycle diurne du cas d'étude GABLS4. En particulier, la version unicolonne de ARPEGE-Climat ne refroidit pas suffisamment l'atmosphère près de la surface pendant la phase nocturne, quelque que soit le forçage appliqué à la limite inférieure (température ou flux de chaleur sensible). L'objectif de ces travaux a été d'une part de comprendre ce paradoxe, et d'autre part d'évaluer dans quelle mesure la physique de ARPEGE-Climat était suffisante pour représenter des couches limites très stables. Ce dernier point soulève le problème de la calibration des modèles de climat et des stratégies déployées pour la réaliser. Cette étape du développement des modèles longtemps restée peu documentée est depuis quelques années l'objet de nombreuses discussions. Des efforts de recherche ont été entrepris pour proposer des méthodes de calibrations objectives, reproductibles et plus rapides que l'approche essai-erreur qui a longtemps prévalu. L'étude du cas GABLS4 a été l'opportunité de tester une méthode appelée *History Matching*, qui ambitionne de répondre aux points précédents.

Nous résumons dans ce chapitre les principaux résultats des travaux présentés dans ce document. L'évaluation du modèle sur l'Antarctique est synthétisée dans la section 8.1. La section 8.2 rappelle le principe de la méthode d'*History Matching* et présente les résultats de son application aux cas GABLS4 puis GABLS1. La section 8.3 rappelle quelques contributions apportées par ce travail de thèse à l'application des techniques d'*History Matching* à la calibration d'un modèle de climat et revient sur les avantages de ces techniques. Pour chacune de ces sections des perspectives sont proposées. La section 8.4 replace les résultats présentés ici dans le cadre plus large de la calibration d'un modèle de climat et présente des pistes pour poursuivre le travail entamé.

8.1 L'Antarctique

Le chapitre 4 présente une évaluation du modèle ARPEGE-Climat 6.3 sur le continent Antarctique. Le modèle ARPEGE-Climat dans sa version CNRM-CM6 présente un biais froid sur l'ensemble du continent Antarctique, aussi bien en été qu'en hiver, à l'exception notable

des côtes et des plates-formes de glace. On a pu montrer qu'en été, ce biais était vraisemblablement lié à un albédo trop élevé. Une correction ad-hoc du calcul de celui-ci, de façon à le faire correspondre en moyenne avec des données de référence (CERES) a permis de réduire très significativement le biais froid estival, voire même de l'inverser dans certaines régions du plateau antarctique. Une conséquence de cette correction est un cycle diurne dégradé avec une amplitude moindre, sur l'Antarctique Est (voir la section 4.4). Il faut noter que la correction de l'albédo proposée ici est purement empirique et n'est pas destinée à être intégrée dans le schéma de surface. Les résultats montrent cependant qu'une revisite de la paramétrisation de l'albédo devra être effectuée pour représenter correctement le climat antarctique dans ARPEGE-Climat.

Pendant l'hiver, malgré la correction de l'albédo, le biais froid demeure inchangé. Ce dernier ne provient donc pas du biais estival qui perdurerait tout au long de l'hiver austral, à moins de supposer une inertie importante du manteau neigeux avec un temps de réponse supérieur à la durée de la simulation (*i.e.* 3 ans). D'autres pistes sont à explorer pour expliquer ce biais froid persistant :

- L'atmosphère qui surplombe le plateau antarctique est trop sèche dans CNRM-CM6-1 (voir section 4.2). Ce biais conduit à un flux infrarouge descendant trop faible, et donc un déficit d'énergie reçue par la surface pendant l'hiver austral. Cependant la température de l'atmosphère dans les basses couches est aussi trop froide, ce qui limite le contenu en vapeur d'eau de la colonne. Il est finalement difficile d'identifier si le manque de vapeur d'eau est un acteur de ce biais ou simplement une conséquence (avec potentiellement une rétroaction positive).
- Les propriétés physiques du manteau neigeux (émissivité, épaisseur de la neige thermiquement active) pourraient également expliquer une partie des résultats (Vignon et al., 2017a).

Une analyse plus approfondie du comportement du modèle pendant l'hiver austral devra être menée pour apporter une réponse à la question du biais froid hivernal de ARPEGE-Climat sur l'Antarctique.

8.2 Calibration de ARPEGE-Climat sur un cas de couche limite stable

L'étape importante du développement des modèles de climat que constitue la calibration soulève un certain nombre de problèmes (objectivité, reproductibilité, efficacité, prise en compte des incertitudes). Une approche possible pour traiter ces problèmes est de considérer la calibration comme un problème d'optimisation et d'appliquer des méthodes de minimisation à une fonction de coût. Appliquée à la calibration de paramétrisations sur des cas d'étude 1D, cette approche présente un risque important de sur-ajustement. La méthode d'*History Matching* apporte une réponse à ces problèmes tout en évitant le risque de sur-ajustement. L'idée fondamentale est non pas de chercher un « meilleur » jeu de paramètres, mais de mettre en évidence une région de l'espace des paramètres dans laquelle le comportement du modèle est satisfaisant. Pour réduire le coût numérique d'une exploration exhaustive de l'espace des paramètres, l'*History Matching* fait appel à des techniques de *machine learning* afin de construire des émulateurs qui simulent le comportement du modèle quand on perturbe ses

paramètres libres.

Au cours des travaux présentés ici, la calibration par *History Matching* a été appliquée au modèle ARPEGE-Climat sur le cas GABLS4 dans différentes conditions de forçages et pour différentes résolutions verticales. Cette calibration a été effectuée dans un cadre 1D/LES, *i.e.* dans lequel des versions uni-colonne du modèle sont comparées à des simulations haute résolution, dites *Large Eddy Simulations* (LES) qui résolvent de manière explicite les principales structures turbulentes. Ce cadre permet d'être au plus proche des processus physiques de l'atmosphère, et ainsi de limiter les compensations d'erreurs. Une version à haute résolution verticale a permis de montrer que le modèle était capable de représenter des couches limites très stratifiées et de faible épaisseur. Dans la configuration standard de ARPEGE-Climat 6.3, l'épaisseur de la couche limite très stable de GABLS4 ne s'étend que sur deux à trois niveaux du modèle. La physique du modèle, et plus précisément son schéma de turbulence, a pourtant montré, sous réserve d'une bonne calibration de celui-ci, sa capacité à reproduire de façon satisfaisante les résultats des LES. On a montré que les résultats étaient particulièrement sensibles à deux paramètres **LMIN** et **KOZMIN** introduits historiquement dans le modèle pour maintenir un minimum de mélange turbulent dans les situations très stables afin d'éviter un découplage trop important de l'atmosphère avec la surface et les phénomènes de *runaway cooling*. Une représentation satisfaisante de la couche limite stable de GABLS4 passe par une réduction de ces paramètres, ce qui est conforme à ce que l'on trouve dans la littérature (Sandu et al., 2013, Holtslag et al., 2013, Vignon et al., 2017b). On rappelle ici que ces paramètres n'existaient pas dans le schéma de turbulence proposé par Cuxart et al. (2000). Contrairement au modèle LMDZ, pour lequel Vignon et al. (2017a) propose de supprimer complètement la borne minimale pour la longueur de mélange (**LMIN** dans ARPEGE-Climat), les résultats montrent qu'il est nécessaire de la maintenir mais à une valeur nettement plus faible que dans CNRM-CM6. Cette réduction des paramètres **LMIN** et **KOZMIN** n'est cependant pas suffisante pour représenter la couche limite de GABLS4. La calibration sur GABLS4 conduit à des valeurs pour les autres paramètres qui réduise le mélange turbulent, jusqu'à des valeurs pouvant être très faibles.

L'application de ces réglages au cas GABLS1 (chapitre 7) qui présente une couche limite faiblement stratifiée montre que le mélange turbulent est la plupart du temps trop faible, conduisant à une couche limite trop peu profonde. L'introduction de métriques supplémentaires calculées sur GABLS1 conduit à des réglages garantissant un mélange turbulent minimal, via les paramètres **CM**, **CE** et **AT** (modulant les coefficients d'échange ou de la longueur de dissipation). Il faut noter que les paramètres **LMIN** et **KOZMIN** ne sont toutefois pas impactés par ces nouvelles métriques et que les réglages de GABLS4 sont conservés pour ces deux paramètres. Ce résultat, attendu, s'explique par le fait que les bornes minimales que représentent ces paramètres ne sont pas atteintes sur le cas GABLS1. Ces résultats mettent en évidence la limite de l'exercice lorsqu'un seul cas est utilisé pour la calibration d'un schéma de turbulence pour les couches limites stables et militent pour une utilisation conjointe de plusieurs régimes de couches limites stables pour la calibration du modèle. Il sera intéressant de créer un ensemble de cas de CLS pouvant servir de banc d'essai aux les paramétrisations de la turbulence en leur appliquant les méthodes exposées dans ce travail de thèse. Cet ensemble pourrait comprendre entre autres les différentes configurations proposées dans le cadre de GABLS4 (avec la possibilité d'avoir un rayonnement et une surface interactifs), le cas GABLS1, avec des variantes proposant différentes intensités du refroidissement

(de -0.25K/h à -1.35K/h), pour lesquels des simulations LES sont déjà disponibles (Rodier et al., 2017) ou différentes intensités de vent géostrophique, ainsi que, éventuellement, les cas GABLS2 et GABLS3.

L'application des réglages réalisés sur la phase convective de GABLS4 a également montré que si les résultats des simulations restaient inchangés, les interactions entre les paramétrisations de la convection et de la turbulence présentaient des différences avec une certaine dispersion. Il serait souhaitable par la suite d'associer à l'exercice de calibration sur les couches limites stables des situations de couches limites convectives sèches (e.g. le cas AYOTTE, déjà traité par la même méthodologie dans Couvreur et al. (2020b)) afin de mieux contraindre ces interactions.

Pour traiter de la représentation de la couche limite hivernale sur le plateau antarctique et du biais froid de ARPEGE-Climat, la construction d'un cas 1D hivernal à partir des observations de Dôme C et avec une surface interactive pourrait permettre d'évaluer la sensibilité de ce biais froid aux schémas de neige et de turbulence.

8.3 Évaluation de la méthode de calibration par *History Matching*

L'évaluation et la calibration de ARPEGE-Climat sur le cas de la couche limite stable de GABLS4 a été une opportunité de mettre en application les outils d'*History Matching* présentés dans Williamson et al. (2013) pour la calibration des modèles de climat. Ces approches orientées quantification des incertitudes ont mis du temps à être appliquées dans la communauté de la modélisation du climat, alors qu'elles étaient déjà couramment utilisées dans d'autres domaines de modélisation numérique (industrie nucléaire, et astrophysique par exemple). Un des objectifs du projet ANR HighTune était de construire un outil commun pour la modélisation du climat, appliquant les méthodes d'*History Matching* au cadre LES/SCM, couramment utilisé pour des travaux de calibrations orientées processus, première étape de la calibration dans (Hourdin et al., 2017). Le travail réalisé au cours de cette thèse a contribué au développement de cet outil. Un article a été publié pour présenter la méthode d'*History Matching with Iterative Refocusing* appliquée à des modèles 1D (Couvreur et al., 2020b). Un autre article (Audouin et al., 2020) présentant l'application au cas GABLS4, propose de plus quelques réflexions sur l'utilisation de cet outil, en insistant sur l'analyse au cours des différentes itérations, des différentes sources d'incertitude et de leur importance pour la convergence de la méthode.

Ce travail de thèse a montré la pertinence de la méthode pour la calibration des modèles de climat, et plus largement des modèles numériques de l'atmosphère. Outre la possibilité d'explorer de façon quasi-exhaustive l'espace des paramètres libres du modèle, via l'utilisation des émulateurs, la méthode utilisée dans ce travail remet en cause l'existence d'un « jeu de paramètres optimal ». La méthode d'*History Matching* se donne au contraire comme objectif de définir une région de l'espace des paramètres qui soit compatible avec les références dont on dispose (observations, LES), en tenant compte des différentes sources d'incertitude qui interviennent dans le processus de calibration.

Par la possibilité d'explorer de façon exhaustive l'espace des paramètres, l'outil permet en outre de d'estimer de façon objective les limites des paramétrisations physiques en montrant dans quelle mesure elles relèvent d'un choix de calibration ou de limites intrinsèques.

Enfin quelques pistes d'amélioration de l'outil sont proposées pour permettre une meilleure visualisation des résultats (voir 6.4.1) ainsi que pour suivre de façon plus simple l'évolution des différentes sources d'incertitude. Plus généralement il conviendrait de continuer de développer des diagnostics et améliorer la visualisation des différentes sorties de l'outil afin d'exploiter toute l'information qu'il peut fournir. Enfin l'utilisation pour la construction des émulateurs d'une vague donnée des simulations des vagues précédentes qui étaient conservées dans l'espace NROY permettrait d'améliorer la convergence et l'efficacité de la méthode.

8.4 Stratégie de calibration d'ARPEGE-Climat

Le travail réalisé au cours de cette thèse et présenté dans ce document prend sa place dans une stratégie plus générale de réglage du modèle. On fournit ici quelques pistes pour poursuivre le travail entamé.

1. La représentation des couches limites stables dans le modèle nécessite de poursuivre le travail commencé ici, en l'étendant à d'autres cas 1D (voir section 8.2) mais également en 3D (en développant des métriques appropriées aux couches limites stables). L'*History Matching* tel que présenté ici doit permettre non seulement de répondre à des questions de calibration et d'évaluer plus finement les limites des paramétrisations actuelles, mais également d'aider au développement de nouvelles formulations (par exemple une nouvelle longueur de mélange comme proposé dans Rodier et al. (2017)). La méthode fournit des outils qui peuvent être utilisés pour évaluer de façon objective différentes formulations d'une paramétrisation et choisir celle qui serait la plus adaptée.
2. Le travail réalisé ne concernait que la paramétrisation de la turbulence. Or les interactions entre les différentes paramétrisations et donc entre les différents processus (rayonnement, surface, voir convection) sont importantes et doivent faire l'objet de futurs travaux. Pour le cas de couches limites très stables, les différentes configurations de GABLS4 (avec rayonnement, avec interaction avec la surface, sur 24h...) peuvent fournir un cadre pour ce travail.
3. La calibration d'un modèle de climat doit prendre en compte toute une variété de situations afin que le modèle soit capable de représenter aussi bien la couche limite convective du Sahara, les couches limites océaniques ou les couches très stables des régions polaires. L'*History Matching* fournit un cadre pour réaliser une calibration 1D multi-cas. L'erreur intrinsèque du modèle, interprétée et utilisée comme une tolérance à l'erreur, pour le moment encore peu exploitée, sera alors une notion clé, pour éviter des espaces NROY vides et en introduisant l'expertise et les attentes du modélisateur (voir 7.2).
4. La calibration de la turbulence sur des cas de couches limites très stables montre le besoin de réévaluer les limites introduites de façon un peu artificielle pour le mélange turbulent. Or une des raisons pour lesquelles ces bornes ont été introduites était la

représentation de la turbulence dans la troposphère libre et de son impact sur la circulation générale. Une évaluation de la paramétrisation de la turbulence dans ce contexte reste à faire, avec probablement des développements spécifiques à entreprendre.

5. Enfin la calibration sur des cas 1D n'est que la première étape de la calibration d'un modèle de climat. Ce travail fournit une base (via la construction des espaces NROY) pour une calibration du modèle 3D, dans un premier temps via des simulations guidées en vent, qui permettent de mettre en évidence les biais relevant de la physique seule, puis via des simulations AMIP afin de prendre en compte les interactions avec la dynamique. Hourdin et al. (2020b) montre un exemple de l'application d'une telle combinaison d'un ensemble de cas 1D et de métriques 3D (sur des simulations AMIP) pour la calibration d'un modèle 3D en s'intéressant plus particulièrement à la représentation des nuages bas.

Bibliographie

- E. A. Anderson. A point energy and mass balance model of a snow cover, 1976. URL <https://repository.library.noaa.gov/view/noaa/6392>. Technical Report.
- J. M. Arblaster and G. A. Meehl. Contributions of External Forcings to Southern Annular Mode Trends. *Journal of Climate*, 19(12) :2896–2905, 06 2006. ISSN 0894-8755. doi : 10.1175/JCLI3774.1. URL <https://doi.org/10.1175/JCLI3774.1>.
- S. Argentini, I. Pietroni, G. Mastrantonio, A. P. Viola, G. Dargaud, and I. Petenko. Observations of near surface wind speed, temperature and radiative budget at dome c, antarctic plateau during 2005. *Antarctic Science*, 26(1) :104–112, 2014.
- Aristidi, E., Agabi, K., Azouit, M., Fossat, E., Vernin, J., Travouillon, T., Lawrence, J. S., Meyer, C., Storey, J. W. V., Halter, B., Roth, W. L., and Walden, V. An analysis of temperatures and wind speeds above dome c, antarctica. *A&A*, 430(2) :739–746, 2005. doi : 10.1051/0004-6361:20041876. URL <https://doi.org/10.1051/0004-6361:20041876>.
- O. Audouin. Calibration statistique de arpege-climat. Master’s thesis, 2017.
- O. Audouin, R. Roehrig, F. Couvreur, and D. Williamson. Modeling the gabls4 strongly-stable boundary layer with a gcm turbulence parameterization : parametric sensitivity or intrinsic limits? *Journal of Advances in Modeling Earth Systems*, n/a(n/a) :e2020MS002269, 2020. doi : <https://doi.org/10.1029/2020MS002269>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002269>.
- K. W. Ayotte, P. P. Sullivan, A. Andr n, S. C. Doney, A. A. Holtslag, W. G. Large, J. C. McWilliams, C.-H. Moeng, M. J. Otte, J. J. Tribbia, et al. An evaluation of neutral and convective planetary boundary-layer parameterizations relative to large eddy simulations. *Boundary-Layer Meteorology*, 79(1-2) :131–175, 1996.
- R. J. Beare, M. K. Macvean, A. A. Holtslag, J. Cuxart, I. Esau, J.-C. Golaz, M. A. Jimenez, M. Khairoutdinov, B. Kosovic, D. Lewellen, et al. An intercomparison of large-eddy simulations of the stable boundary layer. *Boundary-Layer Meteorology*, 118(2) :247–272, 2006.
- J. Beaumet. *Changement climatique en Antarctique :  tudes   l’aide d’un mod le atmosph rique de circulation g n rale   haute r solution r gionale*. PhD thesis, 2018.
- P. Bechtold, J. W. M. Cuijpers, P. Mascart, and P. Trouilhet. Modeling of Trade Wind Cumuli with a Low-Order Turbulence Model : Toward a Unified Description of Cu and Se Clouds in Meteorological Models. *Journal of Atmospheric Sciences*, 52(4) :455 – 463, 1995. doi : 10.1175/1520-0469(1995)052<0455:MOTWCW>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/atsc/52/4/1520-0469_1995_052_0455_motwcw_2_0_co_2.xml.
- A. Beljaars and A. Holtslag. Flux parameterization over land surfaces for atmospheric models. *Journal of Applied Meteorology*, 30(3) :327–341, 1991.

- A. Betts. Non-precipitating cumulus convection and its parameterization. *Quarterly Journal of the Royal Meteorological Society*, 99(419) :178–196, 1973.
- A. K. Blackadar. Boundary layer wind maxima and their significance for the growth of nocturnal inversions. *Bull. Amer. Meteor. Soc*, 38(5) :283–290, 1957.
- A. Bodas-Salcedo, M. J. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. A. Klein, Y. Zhang, R. Marchand, J. M. Haynes, R. Pincus, and V. O. John. Cosp : Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, 92(8) :1023 – 1043, 2011. doi : 10.1175/2011BAMS2856.1. URL https://journals.ametsoc.org/view/journals/bams/92/8/2011bams2856_1.xml.
- F. C. Bosveld, P. Baas, G.-J. Steeneveld, A. A. Holtslag, W. M. Angevine, E. Bazile, E. I. de Bruijn, D. Deacu, J. M. Edwards, M. Ek, et al. The third GABLS intercomparison case for evaluation studies of boundary-layer models. part B : results and process understanding. *Boundary-Layer Meteorology*, 152(2) :157–187, 2014a.
- F. C. Bosveld, P. Baas, E. van Meijgaard, E. I. de Bruijn, G.-J. Steeneveld, and A. A. Holtslag. 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, 2014b.
- P. Bougeault and P. Lacarrère. Parameterization of orography-induced turbulence in a mesobeta-scale model. *Monthly Weather Review*, 117(8) :1872–1890, 1989.
- T. J. Bracegirdle and G. J. Marshall. The reliability of antarctic tropospheric pressure and temperature in the latest global reanalyses. *Journal of Climate*, 25(20) :7138 – 7146, 2012. doi : 10.1175/JCLI-D-11-00685.1. URL <https://journals.ametsoc.org/view/journals/clim/25/20/jcli-d-11-00685.1.xml>.
- T. J. Bracegirdle, E. Shuckburgh, J.-B. Sallee, Z. Wang, A. J. S. Meijers, N. Bruneau, T. Phillips, and L. J. Wilcox. Assessment of surface winds over the Atlantic, Indian, and Pacific Ocean sectors of the Southern Ocean in CMIP5 models : historical bias, forcing response, and state dependence. *Journal of Geophysical Research : Atmospheres*, 118(2) :547–562, 2013. doi : <https://doi.org/10.1002/jgrd.50153>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50153>.
- T. J. Bracegirdle, C. R. Holmes, J. S. Hosking, G. J. Marshall, M. Osman, M. Patterson, and T. Rackow. Improvements in Circumpolar Southern Hemisphere Extratropical Atmospheric Circulation in CMIP6 Compared to CMIP5. *Earth and Space Science*, 7(6) :e2019EA001065, 2020. doi : <https://doi.org/10.1029/2019EA001065>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019EA001065>. e2019EA001065 2019EA001065.
- N. D. Brenowitz and C. S. Bretherton. Prognostic validation of a neural network unified physics parameterization. *Geophysical Research Letters*, 45(12) :6289–6298, 2018. doi : 10.1029/2018GL078510. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078510>.
- W. S. Broecker, S. L. Peacock, S. Walker, R. Weiss, E. Fahrbach, M. Schroeder, U. Mikolajewicz, C. Heinze, R. Key, T.-H. Peng, and S. Rubin. How much deep water is formed

- in the southern ocean? *Journal of Geophysical Research : Oceans*, 103(C8) :15833–15843, 1998. doi : <https://doi.org/10.1029/98JC00248>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JC00248>.
- D. Bromwich, Z. Guo, L. Bai, and Q. shi Chen. Modeled antarctic precipitation. part i : Spatial and temporal variability*. *Journal of Climate*, 17 :427–447, 2004.
- D. H. Bromwich, J. P. Nicolas, K. M. Hines, J. E. Kay, E. L. Key, M. A. Lazzara, D. Lubin, G. M. McFarquhar, I. V. Gorodetskaya, D. P. G., T. Lachlan-Cope, and N. P. M. van Lipzig. Tropospheric clouds in antarctica. *Reviews of Geophysics*, 50(1), 2012. doi : <https://doi.org/10.1029/2011RG000363>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011RG000363>.
- D. H. Bromwich, J. P. Nicolas, A. J. Monaghan, M. A. Lazzara, L. M. Keller, G. A. Weidner, and A. B. Wilson. Central west antarctica among the most rapidly warming regions on earth. *Nature Geoscience*, 6(2) :139–145, Feb 2013. ISSN 1752-0908. doi : 10.1038/ngeo1671. URL <https://doi.org/10.1038/ngeo1671>.
- A. R. Brown, R. J. Beare, J. M. Edwards, A. P. Lock, S. J. Keogh, S. F. Milton, and D. N. Walters. Upgrades to the boundary-layer scheme in the met office numerical weather prediction model. *Boundary-Layer Meteorology*, 128(1) :117–132, Jul 2008. ISSN 1573-1472. doi : 10.1007/s10546-008-9275-0. URL <https://doi.org/10.1007/s10546-008-9275-0>.
- E. Brun, E. Martin, V. Simon, C. Gendre, and C. Coleou. An Energy and Mass Model of Snow Cover Suitable for Operational Avalanche Forecasting. *Journal of Glaciology*, 35 (121) :333–342, 1989. doi : 10.3189/S0022143000009254.
- E. Brun, P. David, M. Sudul, and G. Brunot. A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. *Journal of Glaciology*, 38(128) :13–22, 1992. doi : 10.3189/S0022143000009552.
- A. Budretsky. The new air temperature absolute minimum, 1984. URL http://www.aari.aq/publication/abs_min/abs_min_en.html.
- G. Casasanta, I. Pietroni, I. Petenko, and S. Argentini. Observed and modelled convective mixing-layer height at dome c, antarctica. *Boundary-Layer Meteorology*, 151(3) :597–608, Jun 2014. ISSN 1573-1472. doi : 10.1007/s10546-014-9907-5. URL <https://doi.org/10.1007/s10546-014-9907-5>.
- J. J. Cassano, T. R. Parish, and J. C. King. Evaluation of Turbulent Surface Flux Parameterizations for the Stable Surface Layer over Halley, Antarctica*. *Monthly Weather Review*, 129 (1) :26–46, 01 2001. ISSN 0027-0644. doi : 10.1175/1520-0493(2001)129<0026:EOTSFP>2.0.CO;2. URL [https://doi.org/10.1175/1520-0493\(2001\)129<0026:EOTSFP>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0026:EOTSFP>2.0.CO;2).
- M. G. P. Cavitte, F. Parrenin, C. Ritz, D. A. Young, B. Van Liefferinge, D. D. Blankenship, M. Frezzotti, and J. L. Roberts. Accumulation patterns around dome c, east antarctica, in the last 73 kyr. *The Cryosphere*, 12(4) :1401–1414, 2018. doi : 10.5194/tc-12-1401-2018. URL <https://tc.copernicus.org/articles/12/1401/2018/>.
- H. Chepfer, S. Bony, D. Winker, G. Cesana, J. L. Dufresne, P. Minnis, C. J. Stubenrauch, and S. Zeng. The gcm-oriented calipso cloud product (calipso-goccp). *Journal of Geophysical*

- Research : Atmospheres*, 115(D4), 2010. doi : <https://doi.org/10.1029/2009JD012251>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012251>.
- M. Collins, R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A. J. Weaver, and M. Wehner. *Long-term climate change : Projections, commitments and irreversibility*, pages 1029–1136. Cambridge University Press, 2013. doi : 10.1017/CBO9781107415324.024.
- W. M. Connolley. The antarctic temperature inversion. *International Journal of Climatology*, 16(12) :1333–1342, 1996. doi : 10.1002/(SICI)1097-0088(199612)16:12<1333::AID-JOC96>3.0.CO;2-6.
- F. Couvreux, E. Bazile, Q. Rodier, B. Maronga, G. Matheou, M. J. Chinita, J. Edwards, B. J. H. van Stratum, C. C. van Heerwaarden, J. Huang, , A. F. Moene, A. Cheng, V. Fuka, S. Basu, E. Bou-Zeid, G. Canut, and E. Vignon. Intercomparison of large-eddy simulations of the antarctic boundary layer for very stable stratification. *Boundary-Layer Meteorology*, 176(3) :369–400, Sep 2020a. ISSN 1573-1472. doi : 10.1007/s10546-020-00539-4. URL <https://doi.org/10.1007/s10546-020-00539-4>.
- F. Couvreux, F. Hourdin, D. Williamson, R. Roehrig, V. Volodina, N. Villefranque, O. Audouin, C. Rio, J. Salter, E. Bazile, F. Brient, B. Diallo, F. Favot, R. Honnert, M. Lefebvre1, J. Madeleine, I. Musat, and Q. Rodier. Process-based climate model development harnessing uncertainty quantification. part i : a calibration tool for parameterization improvement. *Journal of Advances in Modeling Earth Systems*, 2020b.
- J. Cuxart, C. Yagüe, G. Morales, E. Terradellas, J. Orbe, J. Calvo, A. Fernández, M. Soler, C. Infante, P. Buenestado, et al. Stable atmospheric boundary-layer experiment in spain (SABLES 98) : a report. *Boundary-layer meteorology*, 96(3) :337–370, 2000.
- J. Cuxart, A. Holtslag, R. J. Beare, E. Bazile, A. Beljaars, A. Cheng, L. Conangla, M. Ek, F. Freedman, R. Hamdi, et al. Single-column model intercomparison for a stably stratified atmospheric boundary layer. *Boundary-Layer Meteorology*, 118(2) :273–303, 2006.
- A. Dahan Dalmedico. History and epistemology of models : Meteorology (1946–1963) as a case study. *Arch. Hist. Exact Sci.*, 55 :395–422, 2001. doi : 10.1007/s004070000032.
- I. Das, R. E. Bell, T. A. Scambos, M. Wolovick, T. T. Creyts, M. Studinger, N. Frearson, J. P. Nicolas, J. T. M. Lenaerts, and M. R. van den Broeke. Influence of persistent wind scour on the surface mass balance of antarctica. *Nature Geoscience*, 6(5) :367–371, May 2013. doi : 10.1038/ngeo1766. URL <https://doi.org/10.1038/ngeo1766>.
- B. Decharme, E. Brun, A. Boone, C. Delire, P. Le Moigne, and S. Morin. Impacts of snow and organic soils parameterization on northern Eurasian soil temperature profiles simulated by the ISBA land surface model. *The Cryosphere*, 10(2) :853–877, 2016. doi : 10.5194/tc-10-853-2016. URL <https://tc.copernicus.org/articles/10/853/2016/>.
- B. Decharme, C. Delire, M. Minvielle, J. Colin, J.-P. Vergnes, A. Alias, D. Saint-Martin, R. Séférian, S. Sénési, and A. Voldoire. Recent changes in the isba-ctrip land surface system for use in the cnrm-cm6 climate model and in global off-line hydrological applications. *Journal of Advances in Modeling Earth Systems*, 11(5) :1207–1252, 2019. doi :

- <https://doi.org/10.1029/2018MS001545>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001545>.
- D. P. Dee, S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kállberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart. The ERA-Interim reanalysis : configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656) :553–597, 2011. doi : <https://doi.org/10.1002/qj.828>. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.828>.
- S. Derbyshire. Stable boundary layers : Observations, models and variability part i : Modelling and measurements. *Boundary-Layer Meteorology*, 74(1) :19–54, 1995.
- Q. Ding, E. J. Steig, D. S. Battisti, and M. Küttel. Winter warming in west antarctica caused by central tropical pacific warming. *Nature Geoscience*, 4(6) :398–403, Jun 2011. ISSN 1752-0908. doi : [10.1038/ngeo1129](https://doi.org/10.1038/ngeo1129). URL <https://doi.org/10.1038/ngeo1129>.
- H. Douville, J.-F. Royer, and J.-F. Mahfouf. A new snow parameterization for the météo-france climate model. part ii : Validation in a 3-d gcm experiments. *Climate Dynamics*, 12(1) :37–52, Nov 1995a. ISSN 1432-0894. doi : [10.1007/BF00208761](https://doi.org/10.1007/BF00208761). URL <https://doi.org/10.1007/BF00208761>.
- H. Douville, J.-F. Royer, and J.-F. Mahfouf. A new snow parameterization for the météo-france climate model. part i : Validation in stand-alone experiments. *Climate Dynamics*, 12(1) :21–35, Nov 1995b. ISSN 1432-0894. doi : [10.1007/BF00208760](https://doi.org/10.1007/BF00208760). URL <https://doi.org/10.1007/BF00208760>.
- P. J. Durack and K. E. Taylor. Pcmdi amip sst and sea-ice boundary conditions version 1.1.3, 2017. URL <https://doi.org/10.22033/ESGF/input4MIPs.1735>.
- E. Dutra, I. Sandu, G. Balsamo, A. Beljaars, E. Vignon, E. Brun, and H. Freville. Understanding the ECMWF winter surface temperature biases over Antarctica. (762), 2015. doi : [10.21957/77ohp5te](https://doi.org/10.21957/77ohp5te). URL <https://www.ecmwf.int/node/15262>.
- M. Déqué, c. Drevet, A. Braun, and D. Cariolle. The arpege/ifs atmosphere model : a contribution to the french community climate modelling. *Climate Dynamics*, 10 :249–266, 1994. doi : <https://doi.org/10.1007/BF00208992>.
- T. D. Foster. Abyssal water mass formation off the eastern wilkes land coast of antarctica. *Deep Sea Research Part I : Oceanographic Research Papers*, 42(4) :501–522, 1995. ISSN 0967-0637. doi : [https://doi.org/10.1016/0967-0637\(95\)00002-N](https://doi.org/10.1016/0967-0637(95)00002-N). URL <https://www.sciencedirect.com/science/article/pii/096706379500002N>.
- Y. Fouquart and B. Bonnel. Computations of solar heating of the earth’s atmosphere : a new parameterization. *Beitr. Phys. Atmosph.*, 53(1) :35–62, 1980.
- J. Franco. *Planification d’expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*. PhD thesis, 2008.

- P. Fretwell, H. D. Pritchard, D. G. Vaughan, J. L. Bamber, N. E. Barrand, R. Bell, C. Bianchi, R. G. Bingham, D. D. Blankenship, G. Casassa, G. Catania, D. Callens, H. Conway, A. J. Cook, H. F. J. Corr, D. Damaske, V. Damm, F. Ferraccioli, R. Forsberg, S. Fujita, Y. Gim, P. Gogineni, J. A. Griggs, R. C. A. Hindmarsh, P. Holmlund, J. W. Holt, R. W. Jacobel, A. Jenkins, W. Jokat, T. Jordan, E. C. King, J. Kohler, W. Krabill, M. Riger-Kusk, K. A. Langley, G. Leitchenkov, C. Leuschen, B. P. Luyendyk, K. Matsuoka, J. Mouginot, F. O. Nitsche, Y. Nogi, O. A. Nost, S. V. Popov, E. Rignot, D. M. Rippin, A. Rivera, J. Roberts, N. Ross, M. J. Siegert, A. M. Smith, D. Steinhage, M. Studinger, B. Sun, B. K. Tinto, B. C. Welch, D. Wilson, D. A. Young, C. Xiangbin, and A. Zirizzotti. Bedmap2 : improved ice bed, surface and thickness datasets for antarctica. *The Cryosphere*, 7(1) :375–393, 2013. doi : 10.5194/tc-7-375-2013. URL <https://tc.copernicus.org/articles/7/375/2013/>.
- H. Fréville, E. Brun, G. Picard, N. Tatarinova, L. Arnaud, C. Lanconelli, C. Reijmer, and M. Van Den Broeke. Using modis land surface temperatures and the crocus snow model to understand the warm bias of era-interim reanalyses at the surface in antarctica. *The Cryosphere*, 8(4) :1361–1373, 2014.
- J.-F. Geleyn. Interpolation of wind, temperature and humidity values from model levels to the height of measurement. *Tellus A*, 40A(4) :347–351, 1988. doi : <https://doi.org/10.1111/j.1600-0870.1988.tb00352.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0870.1988.tb00352.x>.
- C. Genthon, G. Krinner, and M. Sacchetti. Interannual antarctic tropospheric circulation and precipitation variability. *Climate Dynamics*, 21(3) :289–307, Sep 2003. ISSN 1432-0894. doi : 10.1007/s00382-003-0329-1. URL <https://doi.org/10.1007/s00382-003-0329-1>.
- C. Genthon, G. Krinner, and H. Castebrunet. Antarctic precipitation and climate-change predictions : horizontal resolution and margin vs plateau issues. *Annals of Glaciology*, 50(50) :55–60, 2009. doi : 10.3189/172756409787769681.
- C. Genthon, M. S. Town, D. Six, V. Favier, S. Argentini, and A. Pellegrini. Meteorological atmospheric boundary layer measurements and ecmwf analyses during summer at dome c, antarctica. *Journal of Geophysical Research : Atmospheres*, 115(D5), 2010. doi : <https://doi.org/10.1029/2009JD012741>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012741>.
- C. Genthon, D. Six, V. Favier, M. Lazzara, and L. Keller. Atmospheric Temperature Measurement Biases on the Antarctic Plateau. *Journal of Atmospheric and Oceanic Technology*, 28(12) :1598–1605, Dec. 2011. doi : 10.1175/JTECH-D-11-00095.1. URL <https://ui.adsabs.harvard.edu/abs/2011JAOT...28.1598G>.
- C. Genthon, D. Six, H. Gallée, P. Grigioni, and A. Pellegrini. Two years of atmospheric boundary layer observations on a 45-m tower at dome c on the antarctic plateau. *Journal of Geophysical Research : Atmospheres*, 118(8) :3218–3232, 2013.
- C. Genthon, D. Six, C. Scarchilli, V. Ciardini, and M. Frezzotti. Meteorological and snow accumulation gradients across dome c, east antarctic plateau. *International Journal of Climatology*, 36(1) :455–466, 2016. doi : 10.1002/joc.4362. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.4362>.

- C. Genthon, L. Piard, E. Vignon, J.-B. Madeleine, M. Casado, and H. Gallée. Atmospheric moisture supersaturation in the near-surface atmosphere at dome c, antarctic plateau. *Atmospheric Chemistry and Physics*, 17(1) :691–704, 2017. doi : 10.5194/acp-17-691-2017. URL <https://acp.copernicus.org/articles/17/691/2017/>.
- P. Gentine, M. Pritchard, S. Rasp, G. Reinaudi, and G. Yacalis. Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, 45(11) : 5742–5751, 2018. doi : 10.1029/2018GL078202. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078202>.
- A. Gettelman, J. E. Truesdale, J. T. Bacmeister, P. M. Caldwell, R. B. Neale, P. A. Bogenschutz, and I. R. Simpson. The single column atmosphere model version 6 (scam6) : Not a scam but a tool for model evaluation and development. *Journal of Advances in Modeling Earth Systems*, 11(5) :1381–1401, 2019. doi : <https://doi.org/10.1029/2018MS001578>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001578>.
- J.-C. Golaz, L. W. Horowitz, and H. Levy II. Cloud tuning in a coupled climate model : Impact on 20th century warming. *Geophysical Research Letters*, 40(10) :2246–2251, 2013. doi : <https://doi.org/10.1002/grl.50232>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50232>.
- D. Gong and S. Wang. Definition of antarctic oscillation index. *Geophysical Research Letters*, 26(4) :459–462, 1999. doi : 10.1029/1999GL900003. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999GL900003>.
- J. Griggs and J. Bamber. Antarctic ice-shelf thickness from satellite radar altimetry. *Journal of Glaciology*, 57(203) :485–498, 2011. doi : 10.3189/002214311796905659.
- J. F. Guérémy. A continuous buoyancy based convection scheme : one- and three-dimensional validation. *Tellus A*, 63(4) :687–706, 2011. doi : 10.1111/j.1600-0870.2011.00521.x. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0870.2011.00521.x>.
- K.-J. Ha, Y.-K. Hyun, H.-M. Oh, K.-E. Kim, and L. Mahrt. Evaluation of Boundary Layer Similarity Theory for Stable Conditions in CASES-99. *Monthly Weather Review*, 135(10) : 3474–3483, 10 2007. ISSN 0027-0644. doi : 10.1175/MWR3488.1. URL <https://doi.org/10.1175/MWR3488.1>.
- I. M. Held. The Gap between Simulation and Understanding in Climate Modeling. *Bulletin of the American Meteorological Society*, 86(11) :1609–1614, 11 2005. ISSN 0003-0007. doi : 10.1175/BAMS-86-11-1609. URL <https://doi.org/10.1175/BAMS-86-11-1609>.
- H. Hersbach, W. Bell, P. Berrisford, A. Horányi, M.-S. J., J. Nicolas, R. Radu, D. Schepers, A. Simmons, C. Soci, and D. Dee. Global reanalysis : goodbye ERA-Interim, hello ERA5. *ECMWF Newsletter*, pages 17–24, 04 2019. doi : 10.21957/vf291hehd7. URL <https://www.ecmwf.int/node/19027>.
- S. Hoch, P. Calanca, R. Philipona, and A. Ohmura. Year-round observation of longwave radiative flux divergence in greenland. *Journal of Applied Meteorology and Climatology*, 46 (9) :1469–1479, 2007.

- A. Holtslag, G. Svensson, S. Basu, B. Beare, F. Bosveld, and J. Cuxart. Overview of the gewex atmospheric boundary layer study (gabls). In *Workshop on Workshop on Diurnal cycles and the stable boundary layer, 7-10 November 2011*, pages 11–23. ECMWF, 2012.
- A. Holtslag, G. Svensson, P. Baas, S. Basu, B. Beare, A. Beljaars, F. Bosveld, J. Cuxart, J. Lindvall, G. Steeneveld, et al. Stable atmospheric boundary layers and diurnal cycles : challenges for weather and climate models. *Bulletin of the American Meteorological Society*, 94(11) :1691–1706, 2013.
- A. A. Holtslag. Preface : GEWEX Atmospheric Boundary-layer Study (GABLS) on stable boundary layers. *Boundary-Layer Meteorology*, 118(2) :243–246, 2006.
- F. Hourdin, J.-. Grandpeix, C. Rio, S. Bony, A. Jam, F. Cheruy, N. Rochetin, L. Fairhead, A. Idelkadi, I. Musat, J.-L. Dufresne, A. Lahellec, M.-P. Lefebvre, and R. Roehrig. LMDZ5B : the atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, 40(9) :2193–2222, May 2013. ISSN 1432-0894. doi : 10.1007/s00382-012-1343-y. URL <https://doi.org/10.1007/s00382-012-1343-y>.
- F. Hourdin, T. Mauritsen, A. Gettelman, J.-C. Golaz, V. Balaji, Q. Duan, D. Folini, D. Ji, D. Klocke, Y. Qian, F. Rauser, C. Rio, L. Tomassini, M. Watanabe, and D. Williamson. The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, 98(3) :589–602, 2017. doi : 10.1175/BAMS-D-15-00135.1. URL <https://doi.org/10.1175/BAMS-D-15-00135.1>.
- F. Hourdin, C. Rio, J.-Y. Grandpeix, J.-B. Madeleine, F. Cheruy, N. Rochetin, A. Jam, I. Musat, A. Idelkadi, L. Fairhead, M.-A. Foujols, L. Mellul, A.-K. Traore, J.-L. Dufresne, O. Boucher, M.-P. Lefebvre, E. Millour, E. Vignon, J. Jouhaud, F. B. Diallo, F. Lott, G. Gastineau, A. Caubel, Y. Meurdesoif, and J. Ghattas. 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, 2020a. doi : <https://doi.org/10.1029/2019MS001892>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001892>. e2019MS001892 10.1029/2019MS001892.
- F. Hourdin, D. Williamson, C. Rio, F. Couvreur, R. Roehrig, N. Villefranque, I. Musat, L. Fairhead, F. B. Diallo, and V. Volodina. Process-based climate model development harnessing machine learning : Ii. model calibration from single column to global. *Journal of Advances in Modeling Earth Systems*, n/a(n/a) :e2020MS002225, 2020b. doi : <https://doi.org/10.1029/2020MS002225>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002225>. e2020MS002225 2020MS002225.
- S. R. Hudson and R. E. Brandt. A look at the surface-based temperature inversion on the antarctic plateau. 18 :1673–1696, 06 2005. doi : 10.1175/JCLI3360.1.
- IPCC. Climate change 2013 : The physical science basis. Technical report, 2013. URL <https://www.ipcc.ch/report/ar5/wg1/>.
- IPCC. Special report on the ocean and cryosphere in a changing climate. Technical report, 2019. URL <https://www.ipcc.ch/srocc/>.

- S. S. Jacobs, A. F. Amos, and P. M. Bruchhausen. Ross sea oceanography and antarctic bottom water formation. *Deep Sea Research and Oceanographic Abstracts*, 17(6) :935–962, 1970. doi : [https://doi.org/10.1016/0011-7471\(70\)90046-X](https://doi.org/10.1016/0011-7471(70)90046-X).
- S. Kato, F. G. Rose, D. A. Rutan, T. J. Thorsen, N. G. Loeb, D. R. Doelling, X. Huang, W. L. Smith, W. Su, and S.-H. Ham. Surface Irradiances of Edition 4.0 Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Data Product. *Journal of Climate*, 31(11) :4501–4527, 05 2018. ISSN 0894-8755. doi : [10.1175/JCLI-D-17-0523.1](https://doi.org/10.1175/JCLI-D-17-0523.1). URL <https://doi.org/10.1175/JCLI-D-17-0523.1>.
- J. E. Kay, L. Bourdages, N. B. Miller, A. Morrison, V. Yettella, H. Chepfer, and B. Eaton. Evaluating and improving cloud phase in the community atmosphere model version 5 using spaceborne lidar observations. *Journal of Geophysical Research : Atmospheres*, 121(8) : 4162–4176, 2016. doi : <https://doi.org/10.1002/2015JD024699>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD024699>.
- J.-C. King and J. Turner. *Antarctic Meteorology and Climatology*. Cambridge University Press, 1997.
- J. C. King, W. M. Connolley, and S. H. Derbyshire. Sensitivity of modelled antarctic climate to surface and boundary-layer flux parametrizations. *Quarterly Journal of the Royal Meteorological Society*, 127(573) :779–794, 2001. doi : [10.1002/qj.49712757304](https://doi.org/10.1002/qj.49712757304). URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712757304>.
- C. Lanconelli, M. Busetto, E. G. Dutton, G. König-Langlo, M. Maturilli, R. Sieger, V. Vitale, and T. Yamanouchi. Polar baseline surface radiation measurements during the international polar year 2007–2009. *Earth System Science Data*, 3(1) :1–8, 2011. doi : [10.5194/essd-3-1-2011](https://doi.org/10.5194/essd-3-1-2011). URL <https://essd.copernicus.org/articles/3/1/2011/>.
- P. Le Moigne, J. Colin, and B. Decharme. Impact of lake surface temperatures simulated by the FLake scheme in the CNRM-CM5 climate model. *Tellus A : Dynamic Meteorology and Oceanography*, 68(1), 2016. doi : [10.3402/tellusa.v68.31274](https://doi.org/10.3402/tellusa.v68.31274). URL <https://doi.org/10.3402/tellusa.v68.31274>.
- P. Lemoigne, C. Albergel, A. Boone, S. Belamari, B. Decharme, M. Dumont, V. Masson, E. Brun, J.-C. Calvet, H. Giordani, S. Lafont, S. Morin, V. Vionnet, S. Faroux, A.-L. Gibelin, C. Lebeaupin, J.-F. Mahfouf, E. Martin, D. Mironov, J. Noilhan, P. Tulet, and B. Van Den Hurk. SURFEX Scientific documentation, 2018.
- G. Lenderink, A. P. Siebesma, S. Cheinet, S. Irons, C. G. Jones, P. Marquet, F. M. üLLER, D. Olmeda, J. Calvo, E. Sánchez, and P. M. M. Soares. 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, 2004. doi : [10.1256/qj.03.122](https://doi.org/10.1256/qj.03.122). URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/qj.03.122>.
- Q. Libois, G. Picard, L. Arnaud, S. Morin, and E. Brun. Modeling the impact of snow drift on the decameter-scale variability of snow properties on the antarctic plateau. *Journal of Geophysical Research : Atmospheres*, 119(20) :11,662–11,681, 2014. doi : [10.1002/2014JD022361](https://doi.org/10.1002/2014JD022361). URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD022361>.

- D. K. Lilly. Models of cloud-topped mixed layers under a strong inversion. *Quarterly Journal of the Royal Meteorological Society*, 94(401) :292–309, 1968. doi : 10.1002/qj.49709440106. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49709440106>.
- J. Lindvall and G. Svensson. Wind turning in the atmospheric boundary layer over land. *Quarterly Journal of the Royal Meteorological Society*, 145(724) :3074–3088, 2019. doi : <https://doi.org/10.1002/qj.3605>. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3605>.
- N. G. Loeb, D. R. Doelling, H. Wang, W. Su, C. Nguyen, J. G. Corbett, L. Liang, C. Mitrescu, F. G. Rose, and S. Kato. Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *Journal of Climate*, 31(2) :895–918, 01 2018. ISSN 0894-8755. doi : 10.1175/JCLI-D-17-0208.1. URL <https://doi.org/10.1175/JCLI-D-17-0208.1>.
- J.-F. Louis. A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorology*, 17(2) :187–202, 1979.
- I. Mahlstein, P. R. Gent, and S. Solomon. Historical Antarctic mean sea ice area, sea ice trends, and winds in CMIP5 simulations. *Journal of Geophysical Research : Atmospheres*, 118 (11) :5105–5110, 2013. doi : <https://doi.org/10.1002/jgrd.50443>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50443>.
- L. Mahrt. Nocturnal boundary-layer regimes. *Boundary-Layer Meteorology*, 88(2) :255–278, Aug 1998. ISSN 1573-1472. doi : 10.1023/A:1001171313493. URL <https://doi.org/10.1023/A:1001171313493>.
- L. Mahrt. Stratified atmospheric boundary layers. *Boundary-Layer Meteorology*, 90(3) : 375–396, 1999.
- L. Mahrt. Stably stratified atmospheric boundary layers. *Annual Review of Fluid Mechanics*, 46 :23–45, 2014.
- S. Malardel. *Fondamentaux de météorologie*. Cépadués éditions, 2005.
- G. J. Marshall. Half-century seasonal relationships between the southern annular mode and antarctic temperatures. *International Journal of Climatology*, 27(3) :373–383, 2007. doi : 10.1002/joc.1407. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.1407>.
- P. Mascart, J. Noilhan, and H. Giordani. A modified parameterization of flux-profile relationships in the surface layer using different roughness length values for heat and momentum. *Boundary-layer meteorology*, 72 :331–344, 1995.
- V. Masson-Delmotte, M. Kageyama, P. Braconnot, S. Charbit, G. Krinner, C. Ritz, É. Guilyardi, J. Jouzel, A. Abe-Ouchi, M. Crucifix, R. M. Gladstone, C. D. Hewitt, A. Kitoh, A. N. Legrande, O. Marti, U. Merkel, T. Motoi, R. Ohgaito, B. Otto-Bliesner, W. R. Peltier, I. Ross, P. J. Valdes, G. Vettoretti, S. L. Weber, F. Wolk, and Y. Yu. Past and future polar amplification of climate change : climate model intercomparisons and ice-core constraints. *Climate Dynamics*, 26 :513–529, 2006. doi : 10.1007/s00382-005-0081-9. URL <https://hal.archives-ouvertes.fr/hal-00124765>.

- T. Mauritsen. Advancing closures for stably stratified turbulence in global atmospheric models. *ECMWF, Shinfield Park, UK*, 2012.
- T. Mauritsen and G. Svensson. Observations of stably stratified shear-driven atmospheric turbulence at low and high Richardson numbers. *Journal of the atmospheric sciences*, 64(2) :645–655, 2007.
- T. Mauritsen, B. Stevens, E. Roeckner, T. Crueger, M. Esch, M. Giorgetta, H. Haak, J. Jungclaus, D. Klocke, D. Matei, U. Mikolajewicz, D. Notz, R. Pincus, H. Schmidt, and L. Tomasini. Tuning the climate of a global model. *Journal of Advances in Modeling Earth Systems*, 4(3), 2012. doi : 10.1029/2012MS000154. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012MS000154>.
- G. A. Maykut. Energy exchange over young sea ice in the central arctic. *Journal of Geophysical Research : Oceans*, 83(C7) :3646–3658, 1978. doi : 10.1029/JC083iC07p03646. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JC083iC07p03646>.
- A. McCabe and A. R. Brown. The role of surface heterogeneity in modelling the stable boundary layer. *Boundary-Layer Meteorology*, 122(3) :517–534, Mar 2007. ISSN 1573-1472. doi : 10.1007/s10546-006-9119-8. URL <https://doi.org/10.1007/s10546-006-9119-8>.
- E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough. Radiative transfer for inhomogeneous atmospheres : RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research : Atmospheres*, 102(D14) :16663–16682, 1997. doi : 10.1029/97JD00237. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97JD00237>.
- M. Morlighem, E. Rignot, T. Binder, D. Blankenship, R. Drews, G. Eagles, O. Eisen, F. Ferraccioli, R. Forsberg, P. Fretwell, V. Goel, J. S. Greenbaum, H. Gudmundsson, J. Guo, V. Helm, C. Hofstede, I. Howat, A. Humbert, W. Jokat, N. B. Karlsson, W. S. Lee, K. Matsuoka, R. Millan, J. Mouginot, J. Paden, F. Pattyn, J. Roberts, S. Rosier, A. Ruppel, H. Seroussi, E. C. Smith, D. Steinhage, B. Sun, M. R. v. d. Broeke, T. D. v. Ommen, M. v. Wessem, and D. A. Young. Deep glacial troughs and stabilizing ridges unveiled beneath the margins of the antarctic ice sheet. *Nature Geoscience*, 13(2) :132–137, Feb 2020. ISSN 1752-0908. doi : 10.1038/s41561-019-0510-8. URL <https://doi.org/10.1038/s41561-019-0510-8>.
- L. Mortarini, D. Cava, U. Giostra, O. Acevedo, L. Nogueira Martins, P. E. Soares de Oliveira, and D. Anfossi. Observations of submeso motions and intermittent turbulent mixing across a low level jet with a 132-m tower. *Quarterly Journal of the Royal Meteorological Society*, 144(710) :172–183, 2018. doi : <https://doi.org/10.1002/qj.3192>. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3192>.
- Météo-France. Stratégie scientifique de météo-france 2013-2020, 2014. URL https://www.umr-cnrm.fr/IMG/pdf/strategie_scientifique_web.pdf.
- C. Nam, S. Bony, J.-L. Dufresne, and H. Chepfer. The ‘too few, too bright’ tropical low-cloud problem in cmip5 models. *Geophysical Research Letters*, 39(21), 2012. doi : <https://doi.org/10.1029/2012GL053421>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL053421>.

- R. A. J. Neggers. Exploring bin-macrophysics models for moist convective transport and clouds. *Journal of Advances in Modeling Earth Systems*, 7(4) :2079–2104, 2015. doi : <https://doi.org/10.1002/2015MS000502>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015MS000502>.
- J. Noilhan and J.-F. Mahfouf. The isba land surface parameterisation scheme. *Global and Planetary Change*, 13(1) :145 – 159, 1996. doi : [https://doi.org/10.1016/0921-8181\(95\)00043-7](https://doi.org/10.1016/0921-8181(95)00043-7). URL <http://www.sciencedirect.com/science/article/pii/0921818195000437>.
- J. Noilhan and S. Planton. A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review*, 117(3) :536–549, 1989.
- C. Palerme, J. E. Kay, C. Genthon, T. L’Ecuyer, N. B. Wood, and C. Claud. How much snow falls on the antarctic ice sheet? *The Cryosphere*, 8(4) :1577–1587, 2014. doi : 10.5194/tc-8-1577-2014. URL <https://tc.copernicus.org/articles/8/1577/2014/>.
- C. Palerme, C. Genthon, C. Claud, J. E. Kay, N. B. Wood, and T. L’Ecuyer. Evaluation of current and projected antarctic precipitation in cmip5 models. *Climate Dynamics*, 48(1) :225–239, Jan 2017. ISSN 1432-0894. doi : 10.1007/s00382-016-3071-1. URL <https://doi.org/10.1007/s00382-016-3071-1>.
- T. R. Parish and D. H. Bromwich. The surface windfield over the antarctic ice sheets. *Nature*, 328 :41–54, 1987. doi : <https://doi.org/10.1038/328051a0>.
- T. R. Parish and D. H. Bromwich. Reexamination of the Near-Surface Airflow over the Antarctic Continent and Implications on Atmospheric Circulations at High Southern Latitudes*. *Monthly Weather Review*, 135(5) :1961–1973, 05 2007. ISSN 0027-0644. doi : 10.1175/MWR3374.1. URL <https://doi.org/10.1175/MWR3374.1>.
- C. A. Paulson. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *Journal of Applied Meteorology*, 9(6) :857–861, 1970.
- I. Petenko, S. Argentini, G. Casasanta, M. Kallistratova, R. Sozzi, and A. Viola. Wavelike structures in the turbulent layer during the morning development of convection at dome c, antarctica. *Boundary-Layer Meteorology*, 161(2) :289–307, Nov 2016. ISSN 1573-1472. doi : 10.1007/s10546-016-0173-6. URL <https://doi.org/10.1007/s10546-016-0173-6>.
- I. Pietroni, S. Argentini, I. Petenko, and R. Sozzi. Measurements and parametrizations of the atmospheric boundary-layer height at dome c, antarctica. *Boundary-Layer Meteorology*, 143(1) :189–206, Apr 2012. ISSN 1573-1472. doi : 10.1007/s10546-011-9675-4. URL <https://doi.org/10.1007/s10546-011-9675-4>.
- J. Piriou, J. Gueremy, and Y. Bouteloup. A subgrid convection scheme for representing dry, moist and precipitating convection in large-scale models, PCMT. Part 1 : equations. Technical Report 32, CNRM, 2018. URL https://www.umr-cnrm.fr/IMG/pdf/gueremy_part1_note-cnrm_f.pdf.
- J.-M. Piriou, J.-L. Redelsperger, J.-F. Geleyn, J.-P. Lafore, and F. Guichard. An Approach for Convective Parameterization with Memory : Separating Microphysics and Transport in Grid-Scale Equations. *Journal of the Atmospheric Sciences*, 64(11) :4127–4139, 11

2007. ISSN 0022-4928. doi : 10.1175/2007JAS2144.1. URL <https://doi.org/10.1175/2007JAS2144.1>.
- L. M. Polvani, D. W. Waugh, G. J. P. Correa, and S.-W. Son. Stratospheric Ozone Depletion : The Main Driver of Twentieth-Century Atmospheric Circulation Changes in the Southern Hemisphere. *Journal of Climate*, 24(3) :795–812, 02 2011. ISSN 0894-8755. doi : 10.1175/2010JCLI3772.1. URL <https://doi.org/10.1175/2010JCLI3772.1>.
- M. Previdi, K. L. Smith, and L. M. Polvani. The antarctic atmospheric energy budget. part i : Climatology and intraseasonal-to-interannual variability. *Journal of Climate*, 26(17) : 6406–6418, 08 2013. doi : 10.1175/JCLI-D-12-00640.1. URL <https://doi.org/10.1175/JCLI-D-12-00640.1>.
- F. Pukelsheim. The three sigma rule. *The American Statistician*, 48(2) :88–91, 1994. URL <http://www.jstor.org/stable/2684253>.
- D. A. Randall and B. A. Wielicki. Measurements, Models, and Hypotheses in the Atmospheric Sciences. *Bulletin of the American Meteorological Society*, 78(3) :399–406, 03 1997. ISSN 0003-0007. doi : 10.1175/1520-0477(1997)078<0399:MMOHIT>2.0.CO;2. URL [https://doi.org/10.1175/1520-0477\(1997\)078<0399:MMOHIT>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0399:MMOHIT>2.0.CO;2).
- D. A. Randall, K.-M. Xu, R. J. Somerville, and S. Iacobellis. Single-column models and cloud ensemble models as links between observations and climate models. *Journal of Climate*, 9 (8) :1683–1697, 1996.
- C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005. ISBN 026218253X.
- J.-L. Redelsperger and G. Sommeria. Method of representing the turbulence associated with precipitations in a three-dimensional model of cloud convection. *Boundary-Layer Meteorology*, 24(2) :231–252, 1982.
- J.-L. Redelsperger and G. Sommeria. Three-dimensional simulation of a convective storm : Sensitivity studies on subgrid parameterization and spatial resolution. *Journal of the atmospheric sciences*, 43(22) :2619–2635, 1986.
- L. A. Roach, C. M. Bitz, C. Horvat, and S. M. Dean. Advances in modeling interactions between sea ice and ocean surface waves. *Journal of Advances in Modeling Earth Systems*, 11(12) :4167–4181, 2019. doi : <https://doi.org/10.1029/2019MS001836>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001836>.
- Q. Rodier. *Paramétrisation de la turbulence atmosphérique dans la couche limite stable*. PhD thesis, 2017.
- Q. Rodier, V. Masson, F. Couvreux, and A. Paci. Evaluation of a buoyancy and shear based mixing length for a turbulence scheme. *Frontiers in Earth Science*, 5 :65, 2017.
- R. Roehrig, I. Beau, D. Saint-Martin, A. Alias, B. Decharme, J.-F. Guérémy, A. Voldoire, A.-L. Ahmat Younous, E. Bazile, S. Belamari, S. Blein, D. Bouniol, Y. Bouteloup, J. Cattiaux, F. Chauvin, M. Chevallier, J. Colin, H. Douville, P. Marquet, M. Michou, P. Nabat, T. Oudar, P. Peyrillé, J.-M. Piriou, D. Salas y Melia, R. Sférian, and S. Sénési. The cnrm global

- atmosphere model arpege-climat 6.3 : description and evaluation. *Journal of Advances in Modeling Earth Systems*, n/a(n/a) :e2020MS002075, 2020. doi : 10.1029/2020MS002075. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002075>.
- R. Rohde, R. Muller, R. Jacobsen, E. Muller, and S. Perlmutter. A new estimate of the average earth surface land temperature spanning 1753 to 2011. *Geoinformatics and Geostatistics : An Overview*, 1(1), 2013a. doi : 10.4172/2327-4581.1000101.
- R. Rohde, R. Muller, S. Perlmutter, and A. Rosenfeld. Berkeley earth temperature averaging process. *Geoinformatics and Geostatistics : An Overview*, 1(2), 2013b. doi : 10.4172/2327-4581.1000103.
- M.-L. Roussel, F. Lemonnier, C. Genthon, and G. Krinner. Brief communication : Evaluating Antarctic precipitation in ERA5 and CMIP6 against CloudSat observations. *The Cryosphere*, 14(8), 2020. doi : 10.5194/tc-14-2715-2020. URL <https://doi.org/10.5194/tc-14-2715-2020>.
- J.-F. Rysman, A. Lahellec, E. Vignon, C. Genthon, and S. Verrier. Characterization of atmospheric ekman spirals at dome c, antarctica. *Boundary-Layer Meteorology*, 160(2) : 363–373, Aug 2016. ISSN 1573-1472. doi : 10.1007/s10546-016-0144-y. URL <https://doi.org/10.1007/s10546-016-0144-y>.
- J. M. Salter, D. B. Williamson, J. Scinocca, and V. Kharin. Uncertainty quantification for computer models with spatial output using calibration-optimal bases. *Journal of the American Statistical Association*, 114(528) :1800–1814, 2019. doi : 10.1080/01621459.2018.1514306. URL <https://doi.org/10.1080/01621459.2018.1514306>.
- I. Sandu, A. Beljaars, P. Bechtold, T. Mauritsen, and G. Balsamo. Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models? *Journal of Advances in Modeling Earth Systems*, 5(2) :117–133, 2013.
- T. A. Scambos, G. G. Campbell, A. Pope, T. Haran, A. Muto, M. Lazzara, C. H. Reijmer, and M. R. van den Broeke. Ultralow surface temperatures in east antarctica from satellite thermal infrared mapping : The coldest places on earth. *Geophysical Research Letters*, 45(12) :6124–6133, 2018. doi : 10.1029/2018GL078133.
- E. Schlosser, B. Stenni, M. Valt, A. Cagnati, J. G. Powers, K. W. Manning, M. Raphael, and M. G. Duda. Precipitation and synoptic regime in two extreme years 2009 and 2010 at dome c, antarctica – implications for ice core interpretation. *Atmospheric Chemistry and Physics*, 16(8) :4757–4770, 2016. doi : 10.5194/acp-16-4757-2016. URL <https://acp.copernicus.org/articles/16/4757/2016/>.
- G. A. Schmidt, D. Bader, L. J. Donner, G. S. Elsaesser, J.-C. Golaz, C. Hannay, A. Molod, R. B. Neale, and S. Saha. Practice and philosophy of climate model tuning across six us modeling centers. *Geoscientific Model Development*, 10(9) :3207–3223, 2017. doi : 10.5194/gmd-10-3207-2017. URL <https://gmd.copernicus.org/articles/10/3207/2017/>.
- C. A. Severijns and W. Hazeleger. Optimizing parameters in an atmospheric general circulation model. *Journal of Climate*, 18(17) :3527 – 3535, 2005. doi : 10.1175/JCLI3430.1. URL <https://journals.ametsoc.org/view/journals/clim/18/17/jcli3430.1.xml>.

- A. Shapiro and E. Fedorovich. Analytical description of a nocturnal low-level jet. *Quarterly Journal of the Royal Meteorological Society*, 136(650) :1255–1262, 2010.
- A. Shapiro, E. Fedorovich, and S. Rahimi. A unified theory for the great plains nocturnal low-level jet. *Journal of the Atmospheric Sciences*, 73(8) :3037–3057, 2016.
- K. L. Smith, M. Previdi, and L. M. Polvani. The Antarctic Atmospheric Energy Budget. Part II : The Effect of Ozone Depletion and its Projected Recovery. *Journal of Climate*, 26(24) :9729–9744, 12 2013. ISSN 0894-8755. doi : 10.1175/JCLI-D-13-00173.1. URL <https://doi.org/10.1175/JCLI-D-13-00173.1>.
- G.-J. Steeneveld, B. Van de Wiel, and A. Holtslag. Modeling the evolution of the atmospheric boundary layer coupled to the land surface for three contrasting nights in CASES-99. *Journal of the atmospheric sciences*, 63(3) :920–935, 2006.
- G.-J. Steeneveld, A. Holtslag, C. Nappo, B. Van de Wiel, and L. Mahrt. Exploring the possible role of small-scale terrain drag on stable boundary layers over land. *Journal of applied meteorology and climatology*, 47(10) :2518–2530, 2008.
- R. B. Stull. *Meteorology for scientists and engineers : a technical companion book with Ahrens' Meteorology Today*. Brooks/Cole, 1988.
- J. Sun, L. Mahrt, R. M. Banta, and Y. L. Pichugina. Turbulence regimes and turbulence intermittency in the stable boundary layer during CASES-99. *Journal of the Atmospheric Sciences*, 69(1) :338–351, 2012.
- J. Sun, C. J. Nappo, L. Mahrt, D. Belušić, B. Grisogono, D. R. Stauffer, M. Pulido, C. Staquet, Q. Jiang, A. Pouquet, et al. Review of wave-turbulence interactions in the stable atmospheric boundary layer. *Reviews of geophysics*, 53(3) :956–993, 2015.
- G. Svensson and A. A. Holtslag. Analysis of model results for the turning of the wind and related momentum fluxes in the stable boundary layer. *Boundary-layer meteorology*, 132(2) :261–277, 2009.
- G. Svensson and J. Lindvall. Evaluation of near-surface variables and the vertical structure of the boundary layer in cmip5 models. *Journal of Climate*, 28(13) :5233–5253, 2015.
- G. Svensson, A. A. Holtslag, V. Kumar, T. Mauritsen, G. Steeneveld, W. Angevine, E. Bazile, A. Beljaars, E. I. de Bruijn, A. Cheng, et al. Evaluation of the diurnal cycle in the atmospheric boundary layer over land as represented by a variety of single-column models : the second GABLS experiment. *Boundary-layer meteorology*, 140(2) :177–206, 2011.
- R. Sférian, P. Nabat, M. Michou, D. Saint-Martin, A. Voldoire, J. Colin, B. Decharme, C. Delire, S. Berthet, M. Chevallier, S. Sénési, L. Franchisteguy, J. Vial, M. Mallet, E. Joetzjer, O. Geoffroy, J.-F. Guérémy, M.-P. Moine, R. Msadek, A. Ribes, M. Rocher, R. Roehrig, D. Salas-y Mélia, E. Sanchez, L. Terray, S. Valcke, R. Waldman, O. Aumont, L. Bopp, J. Deshayes, C. Éthé, and G. Madec. Evaluation of cnrm earth system model, cnrm-esm2-1 : Role of earth system processes in present-day and future climate. *Journal of Advances in Modeling Earth Systems*, 11(12) :4182–4227, 2019. doi : 10.1029/2019MS001791. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001791>.

- C. Tebaldi and R. Knutti. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 365(1857) :2053–2075, 2007. doi : 10.1098/rsta.2007.2076.
- J. Turner, T. Bracegirdle, T. Phillips, G. J. Marshall, and J. Scott Hosking. An Initial Assessment of Antarctic Sea Ice Extent in the CMIP5 Models. *Journal Of Climate*, 26(5) : 1473–1484, 2013. doi : <https://doi.org/10.1175/JCLI-D-12-00068.1>.
- B. Van de Wiel, A. Moene, O. Hartogensis, H. De Bruin, and A. Holtslag. Intermittent turbulence in the stable boundary layer over land. part III : A classification for observations during CASES-99. *Journal of the atmospheric sciences*, 60(20) :2509–2522, 2003.
- B. J. Van de Wiel, A. Moene, G. Steeneveld, P. Baas, F. Bosveld, and A. Holtslag. A conceptual view on inertial oscillations and nocturnal low-level jets. *Journal of the atmospheric sciences*, 67(8) :2679–2689, 2010.
- B. J. H. Van de Wiel, A. F. Moene, H. J. J. Jonker, P. Baas, S. Basu, J. M. M. Donda, J. Sun, and A. A. M. Holtslag. The Minimum Wind Speed for Sustainable Turbulence in the Nocturnal Boundary Layer. *Journal of the Atmospheric Sciences*, 69(11) :3116–3127, 11 2012. ISSN 0022-4928. doi : 10.1175/JAS-D-12-0107.1. URL <https://doi.org/10.1175/JAS-D-12-0107.1>.
- M. R. van den Broeke and N. P. M. van Lipzig. Factors Controlling the Near-Surface Wind Field in Antarctica*. *Monthly Weather Review*, 131(4) :733–743, 04 2003. ISSN 0027-0644. doi : 10.1175/1520-0493(2003)131<0733:FCTNSW>2.0.CO;2. URL [https://doi.org/10.1175/1520-0493\(2003\)131<0733:FCTNSW>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0733:FCTNSW>2.0.CO;2).
- M. R. van den Broeke, N. P. M. van Lipzig, and E. van Meijgaard. Momentum Budget of the East Antarctic Atmospheric Boundary Layer : Results of a Regional Climate Model. *Journal of the Atmospheric Sciences*, 59(21) :3117–3129, 11 2002. ISSN 0022-4928. doi : 10.1175/1520-0469(2002)059<3117:MBOTEA>2.0.CO;2. URL [https://doi.org/10.1175/1520-0469\(2002\)059<3117:MBOTEA>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<3117:MBOTEA>2.0.CO;2).
- J. M. van Wessem, C. H. Reijmer, J. T. M. Lenaerts, W. J. van de Berg, M. R. van den Broeke, and E. van Meijgaard. Updated cloud physics in a regional atmospheric climate model improves the modelled surface energy balance of antarctica. *The Cryosphere*, 8(1) :125–135, 2014. doi : 10.5194/tc-8-125-2014. URL <https://tc.copernicus.org/articles/8/125/2014/>.
- N. Vercauteren, V. Boyko, D. Faranda, and I. Stiperski. Scale interactions and anisotropy in stable boundary layers. *Quarterly Journal of the Royal Meteorological Society*, 145(722) :1799–1813, 2019. doi : <https://doi.org/10.1002/qj.3524>. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3524>.
- I. Vernon, M. Goldstein, and R. G. Bower. Galaxy formation : a bayesian uncertainty analysis. *Bayesian Anal.*, 5(4) :619–669, 12 2010. doi : 10.1214/10-BA524. URL <https://doi.org/10.1214/10-BA524>.
- E. Vignon. *La couche limite extrême du Plateau Antarctique et sa représentation dans les modèles de climat*. PhD thesis, 2017.

- E. Vignon, F. Hourdin, C. Genthon, H. Gallee, E. Bazile, M.-P. Lefebvre, J.-B. Madeleine, and B. J. Van de Wiel. Antarctic Boundary-Layer Parametrization in a General Circulation Model : 1D simulations facing summer observations at Dome C. *Journal of Geophysical Research : Atmospheres*, 2017a.
- E. Vignon, B. J. H. van de Wiel, I. G. S. van Hooijdonk, C. Genthon, S. J. A. van der Linden, J. A. van Hooft, P. Baas, W. Maurel, O. Traullé, and G. Casasanta. Stable boundary-layer regimes at dome c, antarctica : observation and analysis. *Quarterly Journal of the Royal Meteorological Society*, 143(704) :1241–1253, 2017b.
- E. Vignon, F. Hourdin, C. Genthon, B. J. H. Van de Wiel, H. Gallée, J.-B. Madeleine, and J. Beaumet. Modeling the dynamics of the atmospheric boundary layer over the antarctic plateau with a general circulation model. *Journal of Advances in Modeling Earth Systems*, 10(1) :98–125, 2018. doi : 10.1002/2017MS001184.
- N. Villefranque, S. Blanco, F. Couvreur, R. Fournier, J. Gautrais, R. J. Hogan, F. Hourdin, V. Volodina, and D. Williamson. 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*, 2021.
- V. Vionnet, E. Brun, S. Morin, A. Boone, S. Faroux, P. Le Moigne, E. Martin, and J.-M. Willemet. The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2. *Geoscientific Model Development*, 5(3) :773–791, 2012. doi : 10.5194/gmd-5-773-2012. URL <https://gmd.copernicus.org/articles/5/773/2012/>.
- A. Voldoire, D. Saint-Martin, S. Sénési, B. Decharme, A. Alias, M. Chevallier, J. Colin, J.-F. Guérémy, M. Michou, M.-P. Moine, P. Nabat, R. Roehrig, D. Salas y Méliá, R. Sférian, S. Valcke, I. Beau, S. Belamari, S. Berthet, C. Cassou, J. Cattiaux, J. Deshayes, H. Douville, C. Ethé, L. Franchistéguy, O. Geoffroy, C. Lévy, G. Madec, Y. Meurdesoif, R. Msadek, A. Ribes, E. Sanchez-Gomez, L. Terray, and R. Waldman. Evaluation of cmip6 deck experiments with cnrm-cm6-1. *Journal of Advances in Modeling Earth Systems*, 11(7) :2177–2213, 2019. doi : 10.1029/2019MS001683. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001683>.
- G. Walker. World weather. *Quarterly Journal of the Royal Meteorological Society*, 54(226) :79–87, 1928. doi : 10.1002/qj.49705422601. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49705422601>.
- Z. Wan, S. Hook, and G. Hulley. MOD11C3 MODIS/Terra Land Surface Temperature/Emissivity Monthly L3 Global 0.05Deg CMG V006, 2015. URL <https://doi.org/10.5067/MODIS/MOD11C3.006>.
- S. G. Warren. Optical properties of snow. *Reviews of Geophysics*, 20(1) :67–89, 1982. doi : 10.1029/RG020i001p00067. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG020i001p00067>.
- D. Williamson, M. Goldstein, L. Allison, A. Blaker, P. Challenor, L. Jackson, and K. Yamazaki. History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble. *Climate Dynamics*, 41 :1703–1729, 10 2013. doi : 10.1007/s00382-013-1896-4.

- D. Williamson, A. T. Blaker, C. Hampton, and J. Salter. Identifying and removing structural biases in climate models with history matching. *Climate Dynamics*, 45(5) :1299–1324, Sep 2015. ISSN 1432-0894. doi : 10.1007/s00382-014-2378-z. URL <https://doi.org/10.1007/s00382-014-2378-z>.
- D. B. Williamson, A. T. Blaker, and B. Sinha. Tuning without over-tuning : parametric uncertainty quantification for the nemo ocean model. *Geoscientific Model Development*, 10(4) :1789–1816, 2017. doi : 10.5194/gmd-10-1789-2017. URL <https://www.geosci-model-dev.net/10/1789/2017/>.
- D. M. Winker, W. H. Hunt, and M. J. McGill. Initial performance assessment of CALIOP. *Geophysical Research Letters*, 34(19), 2007. doi : 10.1029/2007GL030135. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007GL030135>.
- N. B. Wood, T. S. L’Ecuyer, F. L. Bliven, and G. L. Stephens. Characterization of video disdrometer uncertainties and impacts on estimates of snowfall rate and radar reflectivity. *Atmospheric Measurement Techniques*, 6(12) :3635–3648, 2013. doi : 10.5194/amt-6-3635-2013. URL <https://amt.copernicus.org/articles/6/3635/2013/>.
- N. B. Wood, T. S. L’Ecuyer, A. J. Heymsfield, G. L. Stephens, D. R. Hudak, and P. Rodriguez. Estimating snow microphysical properties using collocated multisensor observations. *Journal of Geophysical Research : Atmospheres*, 119(14) :8941–8961, 2014. doi : <https://doi.org/10.1002/2013JD021303>. URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013JD021303>.
- T. Zhang, L. Li, Y. Lin, W. Xue, F. Xie, H. Xu, and X. Huang. An automatic and effective parameter optimization method for model tuning. *Geoscientific Model Development*, 8(11) : 3579–3591, 2015. doi : 10.5194/gmd-8-3579-2015. URL <https://gmd.copernicus.org/articles/8/3579/2015/>.