



**HAL**  
open science

# Représentation des erreurs de modélisation dans le système de prévision d'ensemble régional PEARO

Meryl Wimmer

► **To cite this version:**

Meryl Wimmer. Représentation des erreurs de modélisation dans le système de prévision d'ensemble régional PEARO. Météorologie. Université Paul Sabatier - Toulouse III, 2021. Français. NNT : 2021TOU30148 . tel-03616084

**HAL Id: tel-03616084**

**<https://theses.hal.science/tel-03616084v1>**

Submitted on 22 Mar 2022

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

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



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

---

---

Présentée et soutenue le *07/12/2021* par :

**Meryl WIMMER**

---

### Représentation des erreurs de modélisation dans le système de prévision d'ensemble régional PEARO

---

---

#### JURY

Jean-Pierre CHABOUREAU	Physicien, Université Toulouse 3 Paul Sabatier	Président du Jury
Eric BLAYO	Professeur d'Université, Université Grenoble-Alpes	Rapporteur
Clémentine PRIEUR	Professeure d'Université, Université Grenoble-Alpes	Rapporteur
Stéphane VANNITSEM	Directeur de Recherche, Institut Royal Météorologique de Belgique	Examinateur
Agnès LAGNOUX	Maîtresse de Conférence, Université Toulouse Jean Jaurès	Invitée
Loïc BERRE	Directeur de recherche des Ponts, des Eaux et des Forêts, Météo-France	Directeur de thèse
Laurent DESCAMPS	Ingénieur de recherche des Travaux de la Météorologie, Météo-France	Co-directeur de thèse

---

#### École doctorale et spécialité :

*SDU2E : Océan, Atmosphère, Climat*

#### Unité de Recherche :

*Centre National de Recherches Météorologiques*

#### Directeur(s) de Thèse :

*Loïc BERRE et Laurent DESCAMPS*

#### Rapporteurs :

*Eric BLAYO et Clémentine PRIEUR*







# Remerciements

Je tiens à remercier avant tout ceux qui m'ont accompagné pendant ces trois années, mes encadrants. Un grand merci à mes directeurs de thèse, Loik BERRE et Laurent DESCAMPS, que ce soit pour leurs conseils, leurs disponibilités ou leurs questions toujours pertinentes. A Laurent aussi pour s'être retrouvé avec moi, prisonniers de ces bus dont on ne voit jamais la fin du voyage à Reading, et de manière générale, toute cette petite escapade à l'ECMWF pour y apprendre les paramétrisations physiques. Je remercie également mes autres encadrants, tapis dans l'ombre, mais bien présents. Je remercie ainsi chaleureusement Laure RAYNAUD pour sa disponibilité, sa rigueur. Je remercie aussi Yann SEITY, toujours présent pour quelques questions sur le modèle AROME. Je tiens, bien évidemment, à souligner toutes ces petites conversations qu'on a pu avoir tous ensemble et qui ont rendu ce travail de thèse possible.

Pour avoir lu cette thèse de quelques pages jusqu'au bout, je tiens à remercier Eric BLAYO et Clémentine PRIEUR pour m'avoir fait l'honneur de rapporter ma thèse. Je remercie également les examinateurs, Stéphane VANNITSEM et Jean-Pierre CHABOUREAU.

Mes pensées vont aussi pour les membres de mon comité de suivi de thèse. Je remercie, de ce fait, François BOUTTIER pour ces remarques pertinentes, Agnès LAGNOUX, à la fois pour sa participation à mon jury de thèse mais également pour sa curiosité et son intérêt pour ce monde de la météorologie bien loin de sa spécialité. Je l'en remercie pour avoir pris le temps de comprendre. I would like to also thank Inger-Lise FROGNER for her participation to my thesis follow up committee, but also for inviting me to present my work at the HIRLAM EPS team during the HIRLAM Working Weeks. In the same way, I would like to thank Pirkka, Ulf, Alan, Ake, Mauri, Alfons, Pau, and all the HIRLAM-EPS community for their advices and warm welcome.

D'un point de vue plus scientifique, je tiens également à remercier tous ceux, qui ont participé de près ou de loin, à l'aboutissement de cette thèse. Je cite donc, parmi tant d'autres, Bruno SUDRET, Robert FAIVRE, Rachel HONNERT, Fleur COUVREUX, Yves BOUTELOUP, Pascal MARQUET, Ghislain FAURE, ...

Une petite pensée va évidemment à tous ceux qui ont contribué à faire de ces trois années, de beaux moments passés avec eux. En particulier, je pense à RECYF pour ses pauses cafés toujours aussi mouvementées, mais également mes co-bureaux qui ont su me supporter. Mes remerciements vont aussi pour l'équipe SIRES, l'administration du CNRM, ou encore mes élèves et l'équipe pédagogique de l'ENM. Je remercie aussi les doctorants du CNRM, Arnaud, Emilie, Matteo, Ivana, Mayeul, Yamina, Axelle, Léna et tous les autres. Une énorme merci à Thomas pour avoir répondu à toutes mes questions, même un vendredi soir à 17h.

## REMERCIEMENTS

---

Merci également à Fab' pour ses anecdotes toujours croustillantes.

Ces trois années de thèse vont aussi de paire avec le stage que j'ai effectué au CNRM. Pour cette raison, je remercie également toute la communauté NAWDEX, en particulier Jean-Marcel, Marie ou encore Didier.

Pour finir, toute ma gratitude va pour Philippe ARBOGAST et Gwendal RIVIERE, sans qui cette thèse n'aurait jamais eu lieu. Merci à Philippe d'avoir d'abord tout simplement créé ce sujet de thèse. Mais principalement, un grand merci à ces deux maîtres du PV pour leur soutien, leur écoute, leur transmission de cette passion du PV. Merci enfin à Gwendal pour sa disponibilité, la confiance qu'il m'a accordé mais aussi pour m'avoir offert cette possibilité de continuer dans le monde de la recherche.

# Table des matières

Remerciements	i
Liste des figures	xi
Liste des tableaux	xiii
Liste des acronymes	xviii
<b>Introduction</b>	<b>1</b>
<b>I Prévision d'Ensemble</b>	<b>5</b>
<b>1 Incertitudes des modèles de prévision du temps</b>	<b>7</b>
1.1 Prévision Numérique du temps . . . . .	7
1.1.1 Modèles globaux . . . . .	7
1.1.2 Modèles régionaux . . . . .	10
1.2 A quoi ressemble un modèle de prévision numérique du temps? . .	12
1.2.1 Coeur dynamique . . . . .	12
1.2.2 Paramétrisation physique . . . . .	15
1.3 Aspect chaotique de l'atmosphère . . . . .	20
1.4 Illustration des incertitudes des modèles dans le cas d'une dépres- sion des moyennes latitudes . . . . .	22
1.4.1 Dépression Stalactite du 30/09/2016 au 02/10/2016 et bandes transporteuses d'air chaud . . . . .	22
1.4.2 Incertitudes liées à une paramétrisation : sensibilité aux schémas de convection profonde . . . . .	24
1.4.3 Incertitudes liées à la paramétrisation-même : sensibilité aux paramètres de la microphysique . . . . .	29
<b>2 Représenter la distribution de probabilité de l'état de l'atmo- sphère via la Prévision d'Ensemble</b>	<b>39</b>
2.1 Perturbation des conditions initiales . . . . .	40
2.1.1 Perturbation aléatoire . . . . .	41
2.1.2 Perturbation des modes instables de la dynamique . . . . .	41
2.1.3 Assimilation d'Ensemble . . . . .	42
2.1.4 Descente d'échelle des modèles ensemblistes globaux . . . . .	43
2.2 Perturbation des conditions de surface . . . . .	44

2.2.1	Perturbation des schémas de surface . . . . .	44
2.2.2	Perturbation des analyses de surface . . . . .	45
2.2.3	Perturbation de la topographie . . . . .	45
2.3	Perturbation des conditions de bords pour les modèles à aire limitée	45
2.3.1	A partir de prévisions d'ensemble globales . . . . .	46
2.3.2	A partir de prévisions globales déterministes . . . . .	47
2.4	Erreurs de modélisation dues aux paramétrisations . . . . .	49
2.4.1	Approche multi-modèles . . . . .	49
2.4.2	Approche multi-physiques . . . . .	50
2.4.3	SKEB . . . . .	50
2.4.4	SPPT . . . . .	51
2.4.5	Perturbation de paramètres . . . . .	52
2.4.6	Paramétrisations physiques stochastiques . . . . .	55
2.5	Perturbation de la dynamique . . . . .	55
2.6	Évaluation des systèmes de prévision d'ensemble par des scores proba- bilistes . . . . .	56
2.6.1	Scores déterministes . . . . .	57
2.6.2	Diagramme de Talagrand - Histogramme de rang . . . . .	58
2.6.3	Diagramme de fiabilité . . . . .	58
2.6.4	Courbe ROC . . . . .	59
2.6.5	Brier et CRPS . . . . .	61
2.6.6	Rapport Dispersion/Erreur . . . . .	62
2.6.7	Calcul des scores probabilistes . . . . .	62
<b>3</b>	<b>Prévision d'Ensemble AROME : Outils et objectifs</b>	<b>71</b>
3.1	AROME . . . . .	71
3.1.1	Coeur dynamique . . . . .	71
3.1.2	Paramétrisations physiques et dynamiques . . . . .	72
3.2	Prévision d'Ensemble régionale PEARO . . . . .	75
3.2.1	EDA : l'Assimilation d'Ensemble AROME . . . . .	75
3.2.2	Perturbation de la surface . . . . .	76
3.2.3	Couplage avec la PEARP . . . . .	76
3.2.4	Erreurs de modélisation . . . . .	77
3.3	Objectif de la thèse : perturbation des paramètres d'entrée des pa- ramétrisations . . . . .	78
3.3.1	Limite de la méthode SPPT . . . . .	78
3.3.2	Test des méthodes de perturbations de paramètres dans la PEARO . . . . .	78
 <b>II Analyse de sensibilité du modèle AROME aux para- mètres incertains des paramétrisations physiques et dy- namique</b>		<b>85</b>
<b>4</b>	<b>Des méthodes d'analyse de sensibilité locales aux méthodes glo- bales</b>	<b>87</b>
4.1	Principe des analyses de sensibilité . . . . .	87
4.2	Etape 2 : plan d'expérience . . . . .	88
4.3	Etape 4 : Mesure de l'influence et utilisation des indices de Sobol' .	90

---

4.4	Méthode de Morris . . . . .	93
<b>5</b>	<b>Application des méthodes d'analyse de sensibilité globale au modèle AROME</b>	<b>99</b>
5.1	Identification d'une liste de paramètres incertains . . . . .	99
5.2	Choix de la sortie d'AROME pour analyser l'impact des paramètres	103
5.3	Configuration des méthodes de Morris et de Sobol' . . . . .	104
5.3.1	Choix des paramètres de Morris . . . . .	104
5.3.2	Estimation des indices de Sobol' . . . . .	105
5.4	Réduction du coût de calcul . . . . .	106
<b>6</b>	<b>Sensibilité du modèle AROME à divers paramètres d'entrée des paramétrisations</b>	<b>111</b>
6.1	Résumé de l'article . . . . .	111
6.2	Article . . . . .	112
<b>III</b>	<b>Représentation des erreurs de modélisation dans la PEARO</b>	<b>141</b>
<b>7</b>	<b>Optimisation des méthodes de perturbations de paramètres</b>	<b>143</b>
7.1	Échantillon de 1000 PP à partir du jeu de données de l'analyse de sensibilité . . . . .	143
7.2	Optimisation de la méthode PP selon le CRPS . . . . .	144
7.3	Optimisation de la méthode RPP . . . . .	146
7.3.1	A partir de distributions gaussiennes . . . . .	147
7.3.2	A partir des distributions du B-CRPS-PP . . . . .	148
7.3.3	A partir des distributions des 50 meilleurs B-CRPS-PP et du plan d'expérience de Morris . . . . .	151
7.3.4	A partir des distributions obtenues avec un émulateur de CRPS . . . . .	154
7.4	Réduction à la perturbation de 8 paramètres . . . . .	160
<b>8</b>	<b>Évaluation des nouvelles méthodes de représentation des erreurs de modélisation</b>	<b>165</b>
8.1	Evaluation globale des méthodes PP et RPP . . . . .	165
8.1.1	Impact de la perturbation de paramètres . . . . .	166
8.1.2	Impact de l'optimisation selon le CRPS . . . . .	166
8.1.3	Optimisation selon une ou 3 saisons . . . . .	167
8.1.4	Réduction à 8 paramètres . . . . .	168
8.2	Optimisation de la méthode PP . . . . .	168
8.3	Optimisation de la méthode RPP . . . . .	172
8.4	Configuration opérationnelle . . . . .	174
8.5	Caractéristique des différentes représentations de l'erreur de modélisation . . . . .	178
8.5.1	Biais des membres du B-CRPS-PP . . . . .	178
8.5.2	Différence de dispersion entre l'été 2018 et 2020 . . . . .	182
8.6	Evaluation des méthodes de représentation des erreurs de modélisation sur différents cas d'étude . . . . .	184

---

## TABLE DES MATIÈRES

---

8.6.1	Cas de convection sur la France . . . . .	184
8.6.2	Cas de la tempête Gabriel . . . . .	190
8.6.3	Bilan des études de cas . . . . .	194
8.7	Synthèse des résultats . . . . .	194
<b>Conclusion</b>		<b>197</b>
<b>Annexes</b>		<b>201</b>
A	Impact des schémas de convection profonde sur la dynamique de la dépression Stalactite	i
B	ANASYG - Représentation schématique des situations météorologiques	lxiii

# Liste des figures

1.1	Erreur quadratique moyenne comparée aux radiosondages, de la prévision à 48h du géopotential à 500hPa sur le domaine Europe pour différents modèles globaux de mai 2012 et mai 2021 (D'après controle.meteo.fr) . . . . .	10
1.2	Schémas représentant les mailles des principaux modèles globaux : (a) maille étirée du modèle français ARPEGE (Météo-France), (b) maille icosaédrique triangulaire du modèle allemand ICON (DWD), (c) maille « cubed-sphere » du modèle américain GFS (NCEP), (d) maille octaédrique gaussienne réduite du modèle européen IFS (CEPMMT), (e) grille Yin, (f) grille Yang et (g) maille « Ying-Yang » du modèle canadien GEM (CMC) (Adapté de Météo-France et Ullrich et al. (2017)) . . . . .	15
1.3	A gauche, géopotential à 500hPa en couleur et pression au niveau de la mer en isoligne, le 2 octobre 2016 à 12hUTC. Le vol F7 du Falcon20 de SAFIRE réalisé pendant la période d'observation intense n°6 de NAWDEX apparaît en trait noir épais, la section verticale en gris. A droite, image dans le canal visible du capteur VIIRS du satellite Suomi NPP (NASA Worldview) avec la trajectoire du vol F7 du Falcon 20 de SAFIRE en rouge . . . . .	23
1.4	A gauche, évolution temporelle durant la campagne NAWDEX de la corrélation entre le champ de géopotential à 500 hPa des analyses ARPEGE et les 4 régimes de temps. A droite, anomalie de géopotential pour ces 4 régimes de temps hivernaux de l'Océan Atlantique Nord. (Avec l'aimable autorisation de Gwendal Rivière) . . . . .	24
1.5	En haut : Carte montrant la position et la pression (en couleur) le long des trajectoires lagrangiennes de 48h répondant au critère de la bande transporteuse d'air chaud et croisant le vol F7 (représenté en trait noir) pour a) la simulation avec B85 et b) avec PCMT. En bas : Vue méridionale du chauffage le long des trajectoires de la bande transporteuse d'air chaud pour c) la simulation avec B85 et d) avec PCMT. Les points violets représentent l'intersection des trajectoires avec l'iso-0°C. . . . .	26
1.6	Coupes verticales, le 2 octobre 2016 à 15hUTC, à 58°N (latitude du deuxième segment du vol F7) de vent zonal (en couleur) et vorticit� potentielle (en iso-ligne noir ; les zones de vorticit� potentielle sup�rieure � 2PVU sur marqu�e par des pointill�s) pour a) l'analyse IFS, b) B85 et c) PCMT. . . . .	28

1.7	Contenu en glace le long du vol F7 : a) mesuré par la plateforme RALI, b) prévu par ARPEGE avec le schéma de convection profonde PCMT et une vitesse de chute de la neige à 1.5m/s dans le schéma de microphysique de Lopez (2002), c) prévu par ARPEGE avec le schéma de convection profonde PCMT et une vitesse de chute de la neige à 0.6m/s dans le schéma de microphysique de Lopez (2002), (Avec l'aimable autorisation de Gwendal Rivière) . . .	31
2.1	Schéma d'échantillonnage de la distribution de probabilité de l'état futur de l'atmosphère par les différents membres de la prévision d'ensemble. Une prévision d'ensemble fournissant plusieurs scénarios possibles, l'un d'eux peut être plus proche de la réalité qu'un modèle déterministe. . . . .	40
2.2	Schéma de la méthode de « breeding » (traduit et adapté de Callado et al. (2013)) . . . . .	43
2.3	Schéma de la méthode SLAF. Les prévisions déterministes globales sont représentées en vert et les membres de la prévision d'ensemble à aire limitée en rouge et bleu. . . . .	48
2.4	Diagramme de fiabilité pour les ensembles bien calibré, sur-dispersif, sous-dispersif, avec biais positif et négatif (traduit et adapté de Wilks (2011)) . . . . .	59
2.5	Exemple de courbes ROC pour des ensembles $A$ et $B$ (adapté de Wilks (2011)) . . . . .	61
2.6	Définition du CRPS : distance (orange) entre la distribution de probabilité cumulée prévue (noir) et observée (bleu). Exemple pour une température observée à 17°C . . . . .	62
3.1	Domaine de prévision d'AROME-France . . . . .	72
3.2	Processus microphysiques entre la vapeur d'eau ( $r_v$ ), la pluie ( $r_r$ ), la neige ( $r_s$ ), le graupel ( $r_g$ ), la glace primaire ( $r_i$ ) et les gouttelettes nuageuses ( $r_c$ ) : eva = évaporation, dep = déposition de vapeur, cnd = condensation, hen = nucléation hétérogène, ber = effet Bergeron-Findeisen, hon = nucléation homogène, mlt = fonte, dry, wet = accroissement sec et humide, acc = accréation, aut = autoconversion, shd = perte d'eau, cfr = solidification par contact, agg = agrégation, rim = givrage, cvm = fonte de conversion. . . . .	74
3.3	Déroulement de l'activation des paramétrisations physiques pendant un pas de temps du modèle AROME. . . . .	75
4.1	Exemple de plan d'expérience comprenant 10 échantillons des paramètres $X_1$ et $X_2$ par hypercube latin (noir) et hypercube latin centré (blanc) (Adapté de Faivre et al. (2013)) . . . . .	89
4.2	Effets compris dans les différents indices de Sobol'. Exemple pour 3 paramètres, $a$ , $b$ et $c$ . Les indices de premier ordre sont symbolisés en bleu, ceux de second ordre en orange et de troisième ordre en vert. L'indice de Sobol' total du paramètre $a$ comprend l'ensemble des effets inclus dans le triangle violet. . . . .	91

4.3	Exemple de plan d'expérience de Morris pour $k = 2$ paramètres ( $X_1$ et $X_2$ ) avec les caractéristiques suivantes : $r = 5$ , $p = 5$ , $\Delta = \frac{1}{p-1}$ . $A$ , $B$ , $C$ , $D$ et $E$ sont les points initiaux des trajectoires de Morris . . . . .	94
5.1	Erreur relative de prédiction dite « Leave-One-Out », moyennée sur toutes les saisons, jours, échéances des différentes variables météorologiques et sortie du modèle. En noir, l'erreur de prédiction du biais moyen, en gris de la MAE et blanc de la RMSE . . . . .	106
5.2	Scores avec et sans hypothèse hydrostatique et test de différents pas de temps : Non-hydrostatique/60s (bleu) ; Hydrostatique/60s (rose) ; Hydrostatique/75s (vert) ; Hydrostatique/90s (jaune) . . . . .	108
7.1	Schéma synthétisant l'échantillonnage des 1000 PP à partir du plan d'expérience de Morris ainsi que l'optimisation de la méthode PP selon le CRPS de pluie et de vent . . . . .	145
7.2	Schéma explicatif du oRPP gaussien. . . . .	148
7.3	Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le membre 1 (a) et le membre 2 (b) du oRPP gaussien. Le point noir représente la valeur par défaut des paramètres. . . . .	149
7.4	Schéma comparant les tirages aléatoires des 21 paramètres perturbés, effectués pour chaque membre d'une prévision d'ensemble avec les méthodes RPP et oRPP (ex pour 50 B-eCRPS-PP) . . . . .	150
7.5	Histogramme normalisé (intégrale unitaire) de la distribution de probabilité sur les 8 valeurs des paramètres perturbés pour le B-CRPS-PP. La verticale rouge représente la valeur par défaut des paramètres. . . . .	151
7.6	Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le B-CRPS-PP. Ces distributions sont obtenues à partir de celles de la figure 7.3 et d'une estimation par noyau. Le point noir représente la valeur par défaut des paramètres. . . . .	152
7.7	Schéma expliquant la génération de nombre aléatoire suivant une loi de probabilité continue issue d'histogramme et validation de la méthode (exemple à partir de l'histogramme de VSIGQSAT du plan d'expérience de Morris) . . . . .	152
7.8	Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé à partir des 50B-CRPS-PP. Le point noir représente la valeur par défaut des paramètres. . . . .	153
7.9	Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le plan d'expérience de Morris. Le point noir représente la valeur par défaut des paramètres. . . . .	153
7.10	Erreurs relatives moyennes de validation (en bleu) et d'entraînement (en rouge) sur des échantillons correspondants respectivement à 30% et 70% de la base d'apprentissage. Les barres d'erreur indiquent les quantiles 25% et 75% alors que les minima et maxima apparaissent sous la forme d'astérisques. Notons deux axes des ordonnées, à gauche pour l'erreur relative moyenne de validation et à droite pour celle de l'entraînement. . . . .	158

7.11	Schéma d'explication du rééchantillonnage et augmentation de la base de données d'apprentissage . . . . .	159
7.12	Erreur de validation pour différentes tailles d'augmentation de la base d'apprentissage . . . . .	160
7.13	Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le oRPPMat. Le point noir représente la valeur par défaut des paramètres. . . . .	161
7.14	Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le oRPPDot. Le point noir représente la valeur par défaut des paramètres. . . . .	161
8.1	Taux d'amélioration par rapport au SPPT du (a) CRPS et du (b) rapport Dispersion/Erreur pendant la période estivale du 01 Mai 2018 au 30 Juillet 2018 . . . . .	168
8.2	Taux d'amélioration par rapport au SPPT du (a) CRPS et du (b) rapport Dispersion/Erreur pendant la période hivernale du 01 Décembre 2018 au 31 Janvier 2019 . . . . .	169
8.3	Impact de l'optimisation, pendant la période estivale, de la méthode PP sur les scores de la force de vent à 10m : (a) Aire sous la courbe ROC, (b) Diagramme de fiabilité (seuil : 20km/h) ; et de pluie accumulée pendant 3h : (c) Aire sous la courbe de ROC, (d) Diagramme de fiabilité (seuil : 1mm). Les scores du SPPT (bleu), Mean PP (Rouge), des scores maximaux et minimaux des 1000 PP (enveloppe rose), du Be-CRPS-PP (noir) et du B-CRPS-PP (vert) sont tracés en fonction des échéances. . . . .	171
8.4	Impact de l'optimisation, pendant la période hivernale, de la méthode PP sur (a) l'aire sous la courbe ROC pour la force de vent à 10m (seuil : 20km/h) et (b) le diagramme de fiabilité de pluie accumulée pendant 3h (seuil : 1mm) . Les scores du SPPT (bleu), Mean PP (Rouge), des scores maximaux et minimaux des 1000 PP (enveloppe rose), du Be-CRPS-PP (noir) et du B-CRPS-PP (vert) sont tracés le long des échéances. . . . .	172
8.5	Impact de l'optimisation, pendant la période estivale, de la méthode RPP sur le CRPS de la force de vent à 10m (a) CRPS, (b) l'aire sous la courbe ROC des précipitations accumulées pendant 3h (seuil : 1mm) et (c) le diagramme de fiabilité des précipitations sur 3h (seuil : 1mm). Les scores du SPPT (bleu), B-CRPS-PP (vert), RPP (magenta), oRPP-B-CRPS-PP (gris) et oRPP (jaune) sont tracés le long des échéances. . . . .	175
8.6	Impact de l'optimisation, pendant la période estivale, de la méthode RPP sur (a) le CRPS de la force de vent à 10m et (b) sur le diagramme de fiabilité de pluie accumulée pendant 3h (seuil : 1mm) . Les scores du SPPT (bleu) RPP (magenta), oRPP-Morris (bleu clair), oRPP-B-CRPS-PP (gris) et oRPP-50B-CRPS-PP (jaune) sont tracés le long des échéances. . . . .	176

8.7	Impact de l'optimisation, pendant la période estivale, de la méthode RPP sur le CRPS de la force de vent à 10m. Les scores du SPPT (bleu), RPP (magenta), oRPP-B-CRPS-PP (gris), oRPPMat (bleu clair) et oRPPDot (vert) sont tracés le long des échéances. . . . .	176
8.8	Dispersion de (a) la force de vent à 10m et (b) des précipitations accumulées pendant 3h ainsi que (c) CRPS de vent à 10m, en fonction des échéances pendant la période estivale 2018 pour le B-CRPS-PP (vert), le B-CRPS-PP dans la configuration opérationnelle (vert foncé), SPPT (bleu), SPPT dans la configuration opérationnelle (bleu clair) ainsi que le B-CRPS-PP et SPPT combinés dans la configuration opérationnelle (rouge) . . . . .	179
8.9	Biais selon les échéances, pendant la période estivale, du (a) vent à 10m (b), rafales, et (c) pluie accumulée pendant 3h, pour les différents membres du B-CRPS-PP. La ligne noire tiretée correspond au biais pour un membre utilisant les valeurs par défaut des paramètres (correspondant au modèle AROME) . . . . .	181
8.10	Série temporelle du biais de vent à 10m (km/h), à 21hUTC pendant la période estivale pour les différents membres du B-CRPS-PP. La ligne noire tiretée correspond au biais pour un membre utilisant les valeurs par défaut des paramètres (autrement dit correspondant au modèle AROME). . . . .	182
8.11	Étude de la différence d'amélioration de la dispersion entre les étés 2018 et 2020 . . . . .	185
8.12	Cas d'étude du 01 Juillet 2018 . . . . .	187
8.13	Observations de précipitations cumulées sur 3h, le 01 Juillet 2018 à (a) 15h UTC, (b) 18h UTC et (c) 24h UTC ainsi qu'observations de précipitations cumulées sur 24h (d) pendant la journée du 01 Juillet 2018 (Produit ANTILOPE combinant données radar et pluviomètres)	188
8.14	Prévisions de la PEARO du 30 juin 2018 à 21hUTC du cumul de précipitation sur 24h prévu pour la journée du 01 Juillet 2018 (de 3 à 27h d'échéance) : Colonne de gauche Q90, Colonne de droite Qmax; en haut : SPPT, au milieu : B-CRPS-PP, en bas : oRPP utilisant des distributions gaussiennes. . . . .	191
8.15	Cas d'étude des 29 et 30 Janvier 2019 . . . . .	192
8.16	Cumul de précipitation sur 48h entre le 28 Janvier 2019, 21h UTC et le 30 Janvier 2019, 21h UTC : (a) observation ANTILOPE, (b) Q75 pour le SPPT, (c) Q90 pour le SPPT, (d) Q75 pour le B-CRPS-PP et (e) Q90 pour le B-CRPS-PP . . . . .	195
B.1	Légende des symboles utilisées dans les ANASYG . . . . .	lxiv



# Liste des tableaux

1.1	Caractéristiques de différents modèles globaux . . . . .	13
1.2	Caractéristiques de différents modèles régionaux . . . . .	13
2.1	Perturbation des paramètres dépendant des membres de la Prév- ision d'Ensemble, de la date initiale, des échéances et de l'espace (points de grille) pour les quatre méthodes de perturbation de pa- ramètres . . . . .	53
2.2	Tableau de Contingence . . . . .	60
5.1	Liste des 21 paramètres incertains identifiés par les experts phy- siciens. La première colonne indique le schéma de paramétrisation dans lequel le paramètre est impliqué. Son nom est indiqué dans la seconde colonne. La troisième décrit la fonction du paramètre alors que la quatrième donne la valeur par défaut du paramètre, utilisé dans AROME. Les valeurs extrêmes acceptées pour chacun des paramètres sont renseignés dans les deux dernières colonnes. . .	102
5.2	Période pour l'analyse de sensibilité et calcul des scores . . . . .	103
5.3	Définition des indices de sensibilité, avec $f(x, s, d, h)$ , la sortie du modèle dépendant de l'espace $x$ , du jour de début de prévision $d$ et de l'échéance $h$ . $I_i$ est l'indice de sensibilité du paramètre $X_i$ . $S$ est le nombre de saison de 31 jours, $D$ est de nombre de jours pour chaque saison, $H$ est le nombre d'échéances et $X$ est le nombre de données spatiales (nombre de stations au sol pour les scores déterministes (environ 1500), ou nombre de points de grille pour les champs moyennés (environ 200.000)). . . . .	105
5.4	Temps de calcul pour les différentes configurations d'AROME à 2.5km	107
7.1	p-value d'un test de Kolmogorov-Smirnov réalisé entre les distribu- tions des 50B-CRPS-PP et de Morris pour chaque paramètre . . . .	154
7.2	Liste des différents systèmes de prévisions d'ensemble testés avec différentes optimisations. Les versions perturbant 8 paramètres sont en gris. . . . .	163



# Liste des acronymes

<b>ACCESS-G</b>	Australian Community Climate and Earth-System Simulator Global.
<b>AEARO</b>	Assimilation d'Ensemble AROME.
<b>ALADIN</b>	Aire Limitée Adaptation dynamique Développement International.
<b>ALARO</b>	ALADIN-AROME.
<b>ANASYG</b>	ANALyses SYnoptiques Graphiques.
<b>ANTILOPE</b>	ANalyse par spaTiaLisation hOraire des PrEcipitations.
<b>AROME</b>	Application of Research to Operations at MEscale.
<b>ARPEGE</b>	Action de Recherche Petite Echelle Grande Echelle.
<b>ARW</b>	Advanced Research WRF.
<b>ASMO</b>	Adaptive Surrogate Modelling-based Optimization.
<b>B85</b>	Bougeault (1985).
<b>BESK</b>	Binary Electronic Sequence Calculator.
<b>BoM</b>	Bureau of Meteorology / service météorologique australien.
<b>CAPE</b>	Convective Available Potential Energy.
<b>CDF</b>	Fonction de Distribution Cumulée.
<b>CEPMMT</b>	Centre Européen pour les Prévisions Météorologiques à Moyen Terme.
<b>CFL</b>	Courant–Friedrichs–Lew.
<b>CIFK</b>	Concitional Instability of the First Kind.
<b>CISK</b>	Conditional Instability of the Second Kind.
<b>CLINK</b>	Complete Link.
<b>CMA</b>	China Meteorological Administration / service météorologique chinois.
<b>CMC</b>	Centre Météorologique Canadien.
<b>CNRM</b>	Centre National de Recherches Météorologiques.
<b>COSMO-D2</b>	COnsortium for Small scale MOdelling-Deutschland.

<b>CRPS</b>	Continuous Ranked Probability Score.
<b>DRPS</b>	Discrete Ranked Probability Score.
<b>DWD</b>	Deutscher Wetterdienst / service météorologique allemand.
<b>ECUME</b>	Exchange Coefficients from Unified Multicampaigns Estimates.
<b>ENIAC</b>	Electronic Numerical Integrator And Computer.
<b>FMI</b>	Finnish Meteorological Institute / Finnish Meteorological Institute.
<b>GDAPS-KIM</b>	Global Data Assimilation and Prediction System based on the Korea Integrated Model.
<b>GDPS</b>	Global Deterministic Prediction System.
<b>GEM</b>	Global Environmental Multiscale Model.
<b>GFS</b>	Global Forecast System.
<b>GRAPES</b>	Global and Regional Assimilation and Prediction System.
<b>GSM</b>	Global Spectral Model.
<b>HarmonEPS</b>	HARMONIE-EPS.
<b>HARMONIE</b>	Hirlam Aladin Regional/Meso-scale Operational NWP In Europe.
<b>HIRLAM</b>	High Resolution Limited Area Model.
<b>ICON</b>	ICOsahedral Nonhydrostatic.
<b>IFS</b>	Integrated Forecasting System.
<b>IFS-HRES</b>	Integrated Forecasting System High RESolution.
<b>ipSPPT</b>	Independent physical parametrization-based SPPT.
<b>ISBA</b>	Interactions between Soil, Biosphere and Atmosphere.
<b>iSPPT</b>	Independant SPPT.
<b>JMA</b>	Japan Meteorological Agency / service météorologique japonais.
<b>KMA</b>	Korean Meteorological Administration / service météorologique coréen.
<b>LAI</b>	Indice de surface foliaire / Leaf Area Index.
<b>LFM</b>	Local Forecast Model.
<b>LHS</b>	Latin Hypercube Sampling.

<b>MAE</b>	Erreur Absolue Moyenne / Mean Absolute Error.
<b>McICA</b>	Monte Carlo Independent Column Approximation.
<b>MOGREPS</b>	Met Office Global and Regional Ensemble Prediction System.
<b>MSM</b>	Meso-Scale Model.
<b>NAO</b>	North Atlantic Oscillation.
<b>NAWDEX</b>	North Atlantic Waveguide Downstream impact EXperiment.
<b>NCAR</b>	National Center for Atmospheric Research.
<b>NCEP</b>	National Centers for Environmental Prediction / groupe de centres météorologiques des Etats-Unis.
<b>NMMB</b>	Nonhydrostatic Multiscale Model on B-grid.
<b>NWP</b>	Numerical Weather Prediction.
<b>OARA</b>	Observations Automatiques au-delà de RADOME.
<b>OAT</b>	One at A Time.
<b>OMM</b>	Organisation météorologique mondiale.
<b>PCE</b>	Polynomial Chaos Expansion.
<b>PCMT</b>	Prognostic Condensates Microphysics and Transport.
<b>PEARO</b>	Prévision d'Ensemble AROME.
<b>PEARP</b>	Prevision d'Ensemble ARPEGE.
<b>PG</b>	Processus Gaussien.
<b>PMMC</b>	Pergaud et al. (2009).
<b>PP</b>	Perturbed Parameter.
<b>pSPPT</b>	Physical parametrization-based SPPT.
<b>PV</b>	Vorticité Potentielle.
<b>Qmax</b>	quantile maximum.
<b>RADOME</b>	Réseau d'Acquisition de Données et d'Observations Météorologiques Etendues.
<b>RALI</b>	RAdar-LIdar.
<b>RMSE</b>	Racine de l'Erreur Quadratique Moyenne / Root Mean Square Error.
<b>ROC</b>	Receiver Operating Characteristic.
<b>RP</b>	Random Parameter.
<b>RPP</b>	Random Perturbed Parameter.
<b>RRTM</b>	Rapid Radiative Transfer Model.

<b>SAFIRE</b>	Service des Avions Français Instrumentés pour la Recherche en Environnement.
<b>SDT</b>	Stochastic Diabatic Tendency.
<b>SHUM</b>	Stochastic boundary-layer humidity.
<b>SKEB</b>	Stochastic Kinetic Energy Backscatter.
<b>SLAF</b>	Scaled Lagged Average Forecast.
<b>SLHD</b>	Diffusion Horizontale Semi-Lagrangienne.
<b>SLT</b>	Scaled Time-Lagged Perturbations.
<b>SMHI</b>	Sveriges Meteorologiska och Hydrologiska Institut/Institut suédois de météorologie et d'hydrologie.
<b>SPP</b>	Stochastic Parameter Perturbation.
<b>SPPT</b>	Stochastically Perturbed Parameterization Tendencies.
<b>SREPS</b>	Short Range Ensemble Prediction System.
<b>SST</b>	Température de la surface de la mer / Sea Surface Temperature.
<b>STC</b>	Stochastic Trigger of Convection.
<b>SURFEX</b>	SURFace EXternalisée.
<b>TEB</b>	Town Energy Budget.
<b>TIGGE</b>	THORPEX Interactive Grand Global Ensemble.
<b>UGAMP</b>	Universities Global Atmospheric Modelling Project.
<b>UM</b>	Unified Model.
<b>UM-UKV</b>	Unified Model United Kingdom Variable resolution.
<b>WISHE</b>	Wind Induced Surface Heat Exchange.
<b>WRF</b>	Weather Research and Forecasting.
<b>ZAMG</b>	Zentralanstalt für Meteorologie und Geodynamik / service météorologique autrichien.

# Introduction

---

... un battement d'aile d'une mouette peut changer  
à tout jamais l'évolution future de l'atmosphère

---

Edward Lorenz - The predictability of a flow  
which possesses many scales of motion (1969)

Depuis toujours, l'humanité s'intéresse au ciel et à ses humeurs. Que ce soit dans le cadre de la météorologie de tous les jours affectant le quotidien de chacun ou pour des phénomènes violents tels que tempêtes, ouragans ou orages mettant en danger la vie de certains, il est nécessaire de prévoir avec précision l'état futur de l'atmosphère. Chaque jour, les prévisionnistes étudient les simulations produites par les modèles de prévision numérique du temps, en cherchant à élaborer des bulletins de prévision ou de vigilance les plus précis et les plus justes possible. Cependant, malgré des modèles météorologiques de plus en plus réalistes et complexes, la précision et la qualité de leurs simulations restent limitées par le caractère chaotique de l'évolution de l'atmosphère.

En effet, de par son aspect chaotique, l'atmosphère est difficilement prévisible. Une petite erreur dans l'estimation de l'état initial de l'atmosphère peut vite conduire à une mauvaise prévision de son état futur. C'est ce qu'on appelle « l'effet papillon » ou la sensibilité aux conditions initiales. Une première solution est alors d'améliorer autant que possible les conditions initiales mais aussi les modèles météorologiques afin de réduire les erreurs des prévisions. Cependant, un phénomène donné n'est généralement prévisible qu'en dessous d'un certain temps dit horizon de prévisibilité<sup>1</sup>. Au delà, la prévision ne peut être que probabiliste. On ne cherche plus alors à savoir, par exemple, si la température sera de 25°C demain à Toulouse, mais plutôt à connaître la probabilité de dépasser les 25°C demain à Toulouse.

Historiquement, les prévisionnistes ont longtemps utilisé les modèles de prévision dits déterministes pour prévoir l'occurrence d'un évènement. Cependant, depuis quelques années, des systèmes dits de prévisions d'ensemble voient le jour dans la plupart des centres météorologiques. Ils cherchent à prendre en compte les

---

1. Joly, A. et Descamps, L. *De l'incertitude de la prévision ... à la prévision de l'incertitude*, D'après l'exposé au colloque : Comment manier l'incertitude des prévisions météorologiques et des projections climatiques? Organisé le 3 avril 2008 à Paris par le Conseil Supérieur de la Météorologie, 2009.

sources d'incertitude en proposant plusieurs prévisions d'un même modèle atmosphérique visant ainsi à représenter l'ensemble des possibilités de l'état futur de l'atmosphère. Ces systèmes dits ensemblistes doivent alors représenter toutes les sources d'incertitudes intervenant dans la prévision numérique du temps : celles dues à la connaissance imparfaite de l'état initial de l'atmosphère et celles dues aux erreurs de représentation numérique de l'évolution de l'état de l'atmosphère. Chacune d'entre elles étant responsables de l'incertitude de la prévision, elles doivent toutes deux être prises en compte dans la prévision d'ensemble. Cependant, dans ce travail de thèse, nous nous intéresserons seulement aux secondes : les erreurs de modélisation.

Que ce soit en raison de la discrétisation de l'atmosphère ou en raison des approximations réalisées afin d'obtenir des méthodes de résolution des équations de la mécanique des fluides suffisamment rapides pour fournir les prévisions à temps, les modèles de prévision du temps sont entachés d'erreurs. De manière générale, il est souvent considéré que les principales erreurs de modélisation proviennent de la représentation des phénomènes physiques.

A ce jour, dans le système de prévision d'ensemble à l'échelle régionale Prévision d'Ensemble AROME (PEARO), opérationnel à Météo-France, les erreurs de modélisation sont représentées par l'ajout de perturbations aléatoires aux tendances de l'état de l'atmosphère dues aux phénomènes physiques. Cette méthode a cependant le désavantage de rendre difficile l'interprétation physique de l'incertitude. C'est pourquoi une volonté de mettre en place une nouvelle représentation de l'erreur de modélisation dans la PEARO a émergé. Celle-ci doit permettre une meilleure interprétabilité physique de l'incertitude, et être basée sur la perturbation de paramètres incertains des schémas physiques du modèle de prévision numérique du temps AROME. Ainsi, l'objectif de cette thèse est de mettre en place et d'évaluer une telle représentation de l'erreur de modélisation dans la PEARO afin de vérifier son potentiel de complétion voire de remplacement de la méthode actuellement employée.

Dans cette perspective, des analyses de sensibilité ont donc été mises en place afin d'évaluer l'incertitude d'un certain nombre de paramètres physiques. Celles-ci sont essentielles à l'identification et à la meilleure compréhension de l'incertitude du modèle. Elles donnent, par ailleurs, un bon aperçu du potentiel d'utilisation des perturbations de paramètres pour représenter une partie de l'erreur de modélisation. En parallèle, l'implémentation de telles méthodes de représentation de l'erreur de modélisation a été réalisée et évaluée.

Le premier chapitre de cette thèse s'intéressera donc tout d'abord à l'incertitude des modèles de prévision numérique du temps. Après avoir présenté dans les grandes lignes ces modèles atmosphériques, nous illustrerons, sur un cas de tempête, les conséquences de l'incertitude de modélisation de l'atmosphère sur la prévision des vents d'altitude. Cet exemple permettra de mettre en évidence la nécessité d'employer la prévision d'ensemble pour prendre en compte ces incertitudes. Les différentes techniques mises en oeuvre jusqu'à ce jour pour prendre en compte les sources d'incertitude au sein des systèmes ensemblistes de prévision numérique du temps seront ensuite présentées au chapitre 2. Un point sera fait sur les méthodes d'évaluation des ensembles. Un chapitre présentant le système de prévision d'ensemble à l'échelle régionale, PEARO, opérationnel à Météo-France, ainsi que les objectifs de la thèse conclura cette première partie.

La seconde partie de ce mémoire sera dédiée à l'analyse de sensibilité. Le chapitre 4 présentera différentes méthodes alors que les chapitres 5 et 6 se focaliseront respectivement sur leur application au modèle AROME et sur les résultats obtenus. On s'intéressera en particulier à l'impact des paramètres incertains sur les prévisions de temps sensible, c'est-à-dire de champs météorologiques tels que le vent ou le cumul de précipitations. Une compréhension de cette incertitude selon l'espace et le temps sera présentée, tandis que les paramètres les plus à même de représenter l'erreur de modélisation au sein de la PEARO seront identifiés.

La troisième partie de ce mémoire sera consacrée à la représentation de l'erreur de modélisation dans la PEARO. Le chapitre 7 présentera les méthodes innovantes mises en place pour représenter l'erreur de modélisation de manière optimale à partir de la perturbation de paramètres. Dans le dernier chapitre, nous évaluerons ces différentes méthodes de manière globale, mais aussi à travers des études de cas. Une attention sera aussi portée aux différentes caractéristiques des diverses méthodes testées.

Enfin, une conclusion viendra clore ce document.



**Première partie**

**Prévision d'Ensemble**



# Chapitre 1

## Incertitudes des modèles de prévision du temps

### 1.1 Prévision Numérique du temps

#### 1.1.1 Modèles globaux

Dès l'Antiquité, l'Homme s'est très vite intéressé à la compréhension de l'atmosphère et de ses états. A titre d'exemple, Aristote, dans *Les Météorologiques* cherchait déjà à expliquer, avec plus ou moins de succès, des phénomènes météorologiques tels que le vent ou la foudre. Cependant, il faudra attendre le milieu du XIX<sup>e</sup> siècle pour que l'on commence à chercher à prévoir les états futurs de l'atmosphère de façon scientifique.

Lors de la guerre de Crimée (1853-1856) qui opposa l'empire Ottoman allié de la France et du Royaume-Uni, à l'Empire de Russie, la flotte alliée, située dans le port de Balaklava en Mer Noire, fut ravagée par une tempête survenue le 14 novembre 1854 (Landsberg, 1954). En particulier, 41 bâtiments coulèrent dont le navire *Henri IV* (Liais, 1865) et le steamer *Prince*, ce dernier transportant les uniformes hivernaux de l'armée anglaise (Landsberg, 1954). Cette tempête fut suivie d'un hiver rigoureux qui plongea les armées alliées dans des conditions difficiles où près de 12000 français moururent de froid, de maladie ou de faim (Edgerton, 1999). Ce phénomène météorologique fut étudié en détails par Emmanuel Liais qui réussit à localiser la trajectoire de la dépression sur l'Europe entière grâce aux relevés barométriques (Liais, 1865). Il montra en particulier que cette dépression arrivée sur les côtes occidentales françaises le 12 novembre 1854, aurait pu être prévue en Crimée deux jours plus tard et que les armées auraient pu être prévenues à l'avance si les données barométriques avaient été transmises par télégraphe (Liais, 1865). Urbain Le Verrier, dans un compte rendu à l'Académie Française (Le Verrier, 1855), suggéra alors de réunir les observations envoyées par télégraphe pour produire quotidiennement des cartes d'observations à la surface, sur toute l'Europe. Ainsi, à la demande de Napoléon III, Urbain Le Verrier mit en place le premier réseau d'observations français ayant pour but de prévenir les marins des tempêtes (Locher, 2009). Ce réseau, comprenant à l'origine 24 stations dont 13 reliées par télégraphe, passa à 59 observations sur toute l'Europe en 1865

(AstroFiles.net). Et c'est le 17 septembre 1863 qu'Hypolyte Marié-Davy réalisa la première prévision météorologique en France à 24h pour le port d'Hambourg : « Menace à l'ouest sur l'océan » (Locher, 2009).

Remarquons que l'utilisation du télégraphe à des fins météorologiques avait déjà été réalisé à l'observatoire de Greenwich en 1849 par James Glaisher (Locher, 2009). Celui-ci réunissait les observations d'une trentaine de gares anglaises et les transmettaient au Daily News qui publiait le soir-même, la météo de la journée (Rlarriott, 1903). Cependant, il ne s'agit pas ici de prévisions. En revanche, de l'autre côté de l'Atlantique, dès 1843, James Espy organisa un réseau d'observations utilisant le télégraphe dès 1849 (Lagrange, 1930). Très vite, des suivis de déplacement de tornades ont été réalisés (Locher, 2009). L'idée fût aussi émise par Henri Piddington à Calcutta lors de ses études sur les cyclones tropicaux (Liais, 1865).

Il faudra ensuite attendre la fin du XIX<sup>e</sup> et le début du XX<sup>e</sup> siècle, pour que les prémices de la prévision numérique du temps voient le jour. Cleveland Abbe, un météorologue américain, après avoir remarqué vers 1890 que « la météorologie est essentiellement l'application à l'atmosphère de l'hydrodynamique et de la thermodynamique » (Willis and Hooke, 2006; Lynch and Moor, 2008), proposa d'utiliser une approche mathématique pour prévoir les états futurs de l'atmosphère (Abbe, 1901; Lynch and Moor, 2008). En 1904, Vilhelm Bjerknes eût l'idée d'appliquer les lois de la dynamique des fluides et de la thermodynamique à des situations réelles de l'atmosphère pour établir des prévisions (OMM). Plus particulièrement, il identifia 7 variables définissant l'état de l'atmosphère (pression, température, masse volumique, humidité, vent zonal, vent méridien et vitesse verticale) ainsi que 7 équations primitives (équations tridimensionnelles de Navier-Stokes, équation de continuité, d'état ainsi que les première et seconde lois de la thermodynamique) (Lynch and Moor, 2008; Bjerknes, 2009). Il proposa ensuite une méthode de prévision du temps basée sur deux étapes : la première, de type diagnostique, établissant l'état initial de l'atmosphère à partir d'observations et la seconde, pronostique, utilisant les 7 équations pour calculer l'évolution des 7 variables définissant cet état de l'atmosphère (Bjerknes, 2009; Lynch and Moor, 2008).

Cependant, ces équations n'ont pas de solutions analytiques, donc seules des méthodes numériques peuvent permettre d'approximer la solution exacte. En 1922, Lewis Richardson a ainsi formulé le premier modèle numérique hydrostatique de prévision du temps en discrétisant les équations primitives simplifiées de l'atmosphère établies par Bjerknes. Pour cela, son modèle considérait une atmosphère terrestre découpée selon une maille rectangulaire de 200km (latitude) par 230km (longitude), comprenant 3 couches atmosphériques à 4, 7 et 12km d'altitude et calculait les états futurs de l'atmosphère avec un pas temporel de 3h (Richardson, 1922). Pendant 2 ans, il tenta de réaliser une prévision, à 6h d'échéance, de la variation de pression à partir de conditions initiales du 20 Mai 1910 établies par Bjerknes (Lynch and Moor, 2008). Il trouva une variation de 145hPa qui s'avéra complètement fautive, voire irréaliste physiquement : la véritable variation de pression observée était quasi nulle, de moins de 1hPa (Richardson, 1922; Lynch and Moor, 2008). Richardson attribua cette erreur aux incertitudes des conditions initiales du vent (Richardson, 1922; Lynch and Moor, 2008). Plus tard, dans les années 1920, Courant, Friedrichs et Lewy montrèrent que, dans le cadre de la résolution numérique des équations aux dérivées partielles, le pas de temps

est conditionné par la résolution spatiale (Condition CFL ; Friedrichs and Lewy, 1928; Lynch and Moor, 2008). Ainsi, il fut prouvé, qu'une partie de l'erreur de calcul de Richardson résidait dans le choix d'un pas de temps trop long pour son modèle (3h), celui-ci ne devant pas excéder 30 minutes.

La quantité énorme de calculs inhérents à la résolution du modèle de Richardson a amené son auteur à proposer des formulaires de calculs pour prévoir la météo à temps. Il a estimé ainsi que 64000 personnes étaient nécessaires pour mettre en oeuvre cette technique nommée « usine à prévoir le temps » (Richardson, 1922).

Grâce à l'arrivée des premiers ordinateurs dans les années 1940, le calcul d'une solution numérique approchée des équations de la dynamique de l'atmosphère et le projet de Richardson sont devenus de plus en plus réalisables. Malgré la puissance des machines de calculs de l'époque, Jule Charney reprend et simplifie le système d'équations de la dynamique sous la forme d'équations quasi-géostrophiques en 1948 (Lynch and Moor, 2008). En 1950, avec Agnar Fjørtoft et John von Neumann, ils utilisent l'Electronic Numerical Integrator And Computer (ENIAC) pour résoudre les équations de ce modèle barotrope (Persson, 2005; Lynch and Moor, 2008). La machine a produit en 33 jours et 33 nuits, trois prévisions à 24h de l'écoulement atmosphérique à 500 hPa sur la Terre entière (Persson, 2005; Lynch and Moor, 2008). Ces résultats encourageants ouvrirent très vite la voie à la prévision numérique du temps à l'échelle globale.

Dès décembre 1954, grâce à un Binary Electronic Sequence Calculator (BESK), l'institut suédois de Météorologie et d'Hydrologie (SMHI) a produit des prévisions opérationnelles à 72h en temps réel à partir d'un modèle barotrope à une couche mis en place par Carl-Gustaf Rossby (Persson, 2005; Harper et al., 2007; Lynch and Moor, 2008). Par la suite, les Etats-Unis ont suivi le pas avec une production opérationnelle dès 1955. Puis, les modèles se sont complexifiés en se basant directement sur les équations primitives, devenant opérationnels aux Etats-Unis et Allemagne de l'Ouest en 1966, puis au Royaume-Uni en 1972 et Australie en 1977 (Leslie and Dietachmeyer, 1992; Lynch and Moor, 2008).

En parallèle, en 1955, Norman Phillips proposa le premier modèle de circulation générale de l'atmosphère hémisphérique, quasi-géotrophique, à 2 couches, capable de représenter l'intensification des tempêtes via l'instabilité barocline (Phillips, 1956; Hamilton, 2008). Ce modèle peut être considéré comme le premier modèle de climat. En 1959, Syukuro Manabe sous la direction de Joseph Smagorinsky réalisa un modèle global à 500km de résolution et 9 niveaux verticaux allant du sol à 10hPa (Manabe et al., 1965). Ce modèle de Manabe créé en 1965 restera longtemps une référence en la matière (Hamilton, 2008).

Avec la montée en puissance des machines de calculs suivant la loi de Moore (2006), les modèles d'atmosphère deviennent toujours plus performants (Bauer et al., 2015). A titre d'exemple, la figure 1.1 montre l'erreur quadratique moyenne de prévision du géopotential à 500hPa à 48h d'échéance sur l'Europe comparée aux données de radiosondage pour différents modèles globaux entre mai 2012 et 2021. Les différents modèles présentent une amélioration de plus de 2m du géopotential à 500hPa. En particulier, on constate une forte amélioration de près de 5m du modèle allemand lorsque celui-ci a de nouveau été l'objet d'intérêt à partir de novembre 2016 (courbe noire). A ce jour, les modèles du Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT ; vert) et de Météo-France (bleu foncé) font partie des meilleurs mondiaux pour la prévision de ce champ aux

Erreur Quadratique Moyenne de prévision du géopotential (en m) par rapport aux radiosondages

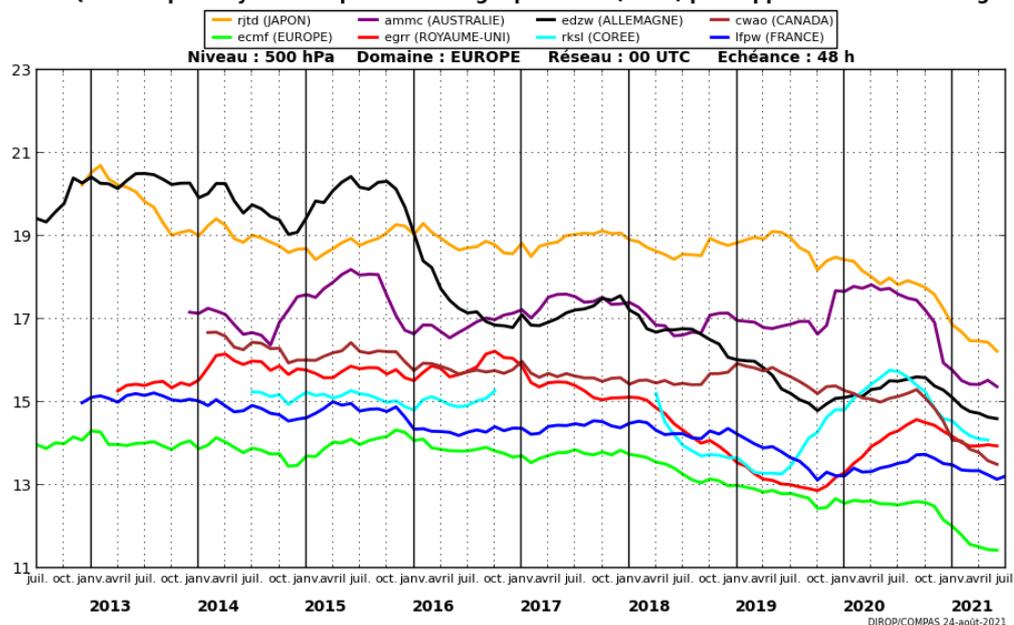


FIGURE 1.1 – Erreur quadratique moyenne comparée aux radiosondages, de la prévision à 48h du géopotential à 500hPa sur le domaine Europe pour différents modèles globaux de mai 2012 et mai 2021 (D'après *controle.meteo.fr*)

courtes échéances. Par ailleurs, les prévisions de ces modèles globaux atteignent des échéances toujours plus lointaines, avec un gain estimé d'un jour de prévision par décade (Bauer et al., 2015). Ces prévisions se font aussi à de plus en plus fine échelle. On estime que la résolution horizontale a été améliorée d'un facteur 10 et la résolution verticale d'un facteur 5 entre 1990 et 2010 (Hamilton, 2008). Ainsi, aujourd'hui, les modèles globaux opérationnels sont capables de prévoir l'état de l'atmosphère sur une dizaine de jours avec une résolution spatiale horizontale d'environ 10km et plus de 100 niveaux verticaux. Le tableau 1.1 présente les caractéristiques actuelles des principaux modèles globaux opérationnels.

### 1.1.2 Modèles régionaux

Malgré la performance des modèles globaux opérationnels, ceux-ci gardent, encore aujourd'hui, une résolution spatiale assez grossière. En effet, avec une résolution horizontale d'environ 10km, les phénomènes météorologiques de taille inférieure à la maille du modèle, dits phénomènes sous-maille, tels que les orages, ne peuvent être explicitement représentés. Dès les années 1970, il a été très vite nécessaire de créer des modèles régionaux imbriqués dans les modèles globaux afin de réaliser des prévisions sur une petite zone du globe, à de plus fines échelles. Ces modèles imbriqués dits aussi modèles à aire limitée, ont la particularité d'avoir les conditions de bords latérales fournies par un modèle global ce qui, comme le constatait dès 1962, Jule Charney pose des problèmes dans le cas des équations en eau peu profonde, par exemple (Mesinger, 2001).

Le premier modèle à aire limitée vit le jour au SMHI. Les premiers tests de modèles imbriqués y ont débuté dès 1967 pour aboutir à une version opérationnelle en 1969 (Bengtsson and Moen, 1971; Mesinger, 2001). Ce modèle imbriqué

quasi-géostrophique à 3 niveaux, avait une résolution de 150km et était couplé au même modèle global à 300km de résolution (Bengtsson and Moen, 1971; Mesinger, 2001). Le service météorologique anglais a ensuite utilisé le même type d'imbrication, mais pour un modèle basé sur les équations primitives. Le modèle régional rectangulaire centré sur les îles britanniques comportait 10 niveaux verticaux et 64 x 48 points de grille séparés de 100km (Mesinger, 2001). L'Agence Météorologique Japonaise (JMA) réalisa en 1973, le premier couplage entre deux modèles différents. Le modèle régional basé sur les équations primitives comportait 6 niveaux et avait une résolution horizontale de 152km alors que le modèle coupleur de 304km de résolution, était cette fois-ci quasi-géostrophique, sur l'hémisphère-Nord (Okamura, 1975; Mesinger, 2001).

En raison d'un manque de moyens informatiques, de nombreux centres météorologiques ne pouvant produire des prévisions à l'échelle globale, se sont focalisés sur ces modèles à aire limitée. La fin des années 1980 fût l'essor de ces nouveaux modèles. En particulier, le centre météorologique yougoslave de Belgrade fut le premier, en Janvier 1978 à produire des prévisions opérationnelles avec un modèle régional couplé aux prévisions d'un modèle global d'un autre centre météorologique (centre allemand, actuel DWD ; Mesinger, 2001).

Avec une résolution horizontale à l'origine de l'ordre de la centaine de kilomètres, ces modèles régionaux atteignent désormais une résolution de 1 à 5 km et peuvent ainsi modéliser explicitement les phénomènes physiques de moyenne échelle tels que la convection profonde amenant à la formation d'orage (cumulonimbus). A de telles résolutions, les mouvements verticaux doivent être pris en compte plus précisément : l'hypothèse hydrostatique n'est plus valide, la force de pression verticale n'est plus en équilibre avec la gravitation, les modèles deviennent non hydrostatiques. Ces modèles, dits à convection profonde résolue, ont fait l'objet de nombreuses recherches jusqu'à la fin du XX<sup>e</sup> siècle. Leur première application en temps réel se fit dans le cadre de la prévision de tornades en 1989 (Clark et al., 2016). Aujourd'hui, la plupart des centres nationaux météorologiques utilisent en opérationnel des modèles régionaux à de telles résolutions (voir tableau 1.2).

Cependant, en raison de la finesse de leur résolution, ces modèles ne peuvent être utilisés à des fins opérationnelles à l'échelle globale, même si des premiers tests ainsi que des projets commencent à voir le jour (par exemple Stevens et al., 2019). Par ailleurs, malgré leur haute résolution spatiale, ces modèles ne peuvent résoudre explicitement l'intégralité de la convection, il reste toujours des phénomènes physiques plus petits que la maille du modèle. Cette zone « grise » où les phénomènes atmosphériques sont partiellement résolus et doivent être aussi représentés, reste aujourd'hui un défi pour les modélisateurs de l'atmosphère. A l'heure actuelle, la recherche s'intéresse à des modèles régionaux pouvant représenter l'atmosphère à l'échelle horizontale de 500m.

Enfin, les modèles à échelle convective étant très réalistes et proposant des champs de précipitations semblables aux données radar, des modèles de prévisions immédiates à haute résolution virent le jour. Basés sur une mise à jour très fréquente des observations, ils fournissent des prévisions de quelques heures.

## 1.2 A quoi ressemble un modèle de prévision numérique du temps ?

### 1.2.1 Coeur dynamique

Prévoir les états futurs d'un système dynamique, c'est résoudre les équations aux dérivées partielles faisant intervenir la variable temporelle et caractérisant l'évolution de ce système. Lors de l'étude de systèmes mécaniques simples, les équations différentielles en présence sont solvables explicitement dès lors que les conditions initiales sont connues (ex : oscillateur harmonique). Par opposition, l'atmosphère est régie par les équations de la dynamique des fluides de Navier-Stokes qui, de par leurs caractéristiques mathématiques, ne peuvent être résolues analytiquement. Ainsi, pour calculer les états futurs de l'atmosphère décrits par un vecteur d'état comprenant 7 variables, la seule possibilité trouvée à ce jour repose sur l'utilisation de techniques d'analyse numérique. En effet, ces méthodes basées sur une discrétisation de l'espace et du temps, permettent d'approximer les états futurs des différentes variables caractérisant l'atmosphère, en calculant, pas de temps après pas de temps, la progression de l'atmosphère d'un état initial vers un état prédit à un instant  $t$ .

Ainsi, les modèles de prévision numérique du temps considèrent une atmosphère découpée selon une grille tridimensionnelle. Concernant la coordonnée verticale, différents types de coordonnées peuvent être utilisés : altitude, masse, pression, pression hydrostatique, ou hybride (voir section 3.1). Différentes grilles de discrétisation horizontale ont été proposées. La figure 1.3 présente en particulier la maille étirée du modèle global français ARPEGE (coefficient d'étirement  $c = 2.2$ ), la maille icosaédrique triangulaire du modèle allemand ICON, la maille « cubed-sphere » du modèle américain GFS, la maille octaédrique gaussienne réduite du modèle européen IFS et celle appelée « Ying-Yang » du modèle canadien GEM. L'évolution temporelle de l'atmosphère est quant à elle discrétisée selon un pas de temps  $\Delta t$ . En chacun des points de la grille et à chaque pas de temps, les modèles calculent les futures valeurs du vecteur d'état en estimant les tendances temporelles de ces variables pronostiques grâce aux équations de la dynamique. Généralement, la méthode numérique utilisée dans la plupart des modèles météorologiques est le schéma semi-implicite qui moyenne les tendances sur un ou plusieurs pas de temps. Il est dit semi-implicite car les termes linéaires sont résolus de manière implicite alors que les termes non-linéaires utilisent des formules explicites (Benacchio and Wood, 2016).

Le schéma semi-implicite n'étant pas inconditionnellement stable numériquement, celui-ci peut amener à des prévisions irréalistes. Ainsi, il est souvent couplé à des schémas centrés implicites itératifs, aussi nommés prédicteur-correcteur dans le cas d'un schéma à 2 itérations, afin de converger vers le schéma implicite de Crank-Nicolson plus stable (Seity, 2020).

Centre Météorologique (Pays)	Modèle	Résolution horizontale	Niveaux	Echéance maximale	Condition initiale
CEPMMT (Europe)	IFS-HRES	9km	137	10 jours	4D-Var
Météo-France (France)	ARPEGE	5 à 24 km	105	4 jours	4D-Var
Met Office (Royaume-Uni)	UM	10km	70	6 jours	4D-Var
DWD (Allemagne)	ICON	13km	90	7 jours	Hybrid 3D-Var
CMC (Canada)	GDPS/GEM	0.14°	84	10 jours	Hybrid 4D-En-Var
CMA (Chine)	GRAPES	10km (0.15°)	90	10 jours	4D-Var
BoM (Australie)	ACCESS-G	12km	70	10 jours	4D-Var
KMA (Corée du Sud)	GDAPS-KIM	12km	91	12 jours	Hybrid 4DEnVAR
NCEP (Etats-Unis)	GFS	13km (0.25°)	127	16 jours	4D-En-Var
JMA (Japon)	GSM	20km (0.1875°)	128	132-264h	Hybrid 4D-Var

TABLE 1.1 – Caractéristiques de différents modèles globaux

Centre Météorologique (Pays)	Modèles	Résolution horizontale	Niveaux	Echéance maximale	Condition initiale
DWD (Allemagne)	COSMO-D2	2.2km	65	27h	KENDA
Météo-France (France)	AROME	1.3km	90	48h	3D-Var
JMA (Japon)	MSM	5km	76	39-51h	4D-Var
JMA (Japon)	LFM	2km	76	10h	3D-Var
Met-Norway (Norvège)	HARMONIE-AROME	2.5km	65	66h	3D-Var
FMI (Finlande)	HIRLAM	7.5km	65	54h	4D-Var
ZAMG (Autriche)	ALARO	4.8km	60	72h	Analyse de surface
Met Office (Royaume-Uni)	UM-UKV	1.5km	70	54h	4D-Var

TABLE 1.2 – Caractéristiques de différents modèles régionaux

Différentes techniques d'estimation des dérivées spatiales impliquées dans les équations ont été mises au point. Avant les années 1970, la plupart des modèles étaient basés sur des méthodes en point de grille où les dérivées impliquées étaient calculées par différences finies explicites (Williamson, 2007). Cependant, pour assurer qualité et stabilité des prévisions, ces méthodes sont soumises à la condition CFL : pendant un pas de temps, le transport des phénomènes météorologiques par le vent (advection) et la propagation d'une onde ne doivent pas dépasser un carreau de la grille du modèle. Pour remédier en partie à ce problème, des méthodes de transformée spectrale proposées par Eliassen et al. (1970) et Orszag peuvent être utilisées : le calcul des dérivées du vecteur d'état se fait dans l'espace spectral des harmoniques de ce vecteur. L'efficacité de la méthode est améliorée avec la Fast Fourier Transform (Williamson, 2007).

Enfin, pour caractériser le phénomène d'advection, les méthodes semi-lagrangiennes ont permis un gain significatif des modèles de prévision du temps : Hortal (1998) constate une amélioration d'un facteur 72 grâce à ces changements algorithmiques. Celles-ci, basées sur la méthode des caractéristiques, modélisent l'advection par le vent de différentes quantités via le calcul de trajectoires lagrangiennes (aussi appelées caractéristiques). Ainsi, à partir des champs météorologiques connus à l'instant  $t$ , on cherche à estimer les champs météorologiques en chaque point de grille à l'instant  $t + \Delta t$  (pour les schémas semi-lagrangien à 2 pas de temps). Pour cela, des rétrotrajectoires (trajectoires dans le passé) ensemencées sur les points de grille du modèle à l'instant  $t + \Delta t$  sont calculées. Celles-ci transportent, par le vent à l'instant  $t$ , les parcelles d'air issues des points de grilles du modèle vers leur point de départ (pied de la caractéristique ou point d'origine de la trajectoire) à l'instant  $t$ . Une fois le point de départ estimé, des méthodes d'interpolation permettent de retrouver, à partir des champs météorologiques déjà connus à l'instant  $t$ , la valeur des variables pronostiques à ces points de départ (différents des points de grille) à l'instant  $t$  et donc, implicitement, aux points de grille à l'instant  $t + \Delta t$ . Le champ de vent utilisé pour l'advection étant pris à l'instant  $t$  est donc supposé constant sur l'intervalle  $[t, t + \Delta t]$ . Cependant, celui-ci n'est pas optimal pour modéliser le transport à l'instant  $t + \Delta t$ . Des schémas permettent alors de mieux estimer, par itérations successives, le champ de vent advectif (Seity, 2020).

Avec les méthodes semi-lagrangiennes, la condition CFL devient moins contraignante sur le choix du pas de temps. Les méthodes autorisent alors de plus grands pas de temps, améliorant significativement la rapidité des calculs. Le seul critère en vigueur est le critère de Lipschitz interdisant le croisement de trajectoires.

En parallèle de ces méthodes semi-lagrangiennes et spectrales, un schéma de diffusion horizontale est ajouté afin de reproduire la cascade d'énergie de Kolmogorov et donc de dissiper l'énergie cinétique à l'échelle de la maille par les processus sous mailles. Ainsi, ce schéma apparaît comme un filtre évitant une accumulation de bruit à l'échelle de la maille par blocage spectral.

Ainsi, la résolution des équations de la dynamique des fluides (équations de Navier-Stokes) permet de modéliser l'écoulement de l'atmosphère autour de la Terre. Cette partie, nommée le **coeur dynamique**, assure principalement la propagation horizontale des mouvements de l'atmosphère. Les ondes de gravité ainsi que les ondes acoustiques y sont aussi développées et des techniques numériques sont parfois utilisées pour filtrer les ondes non météorologiques.

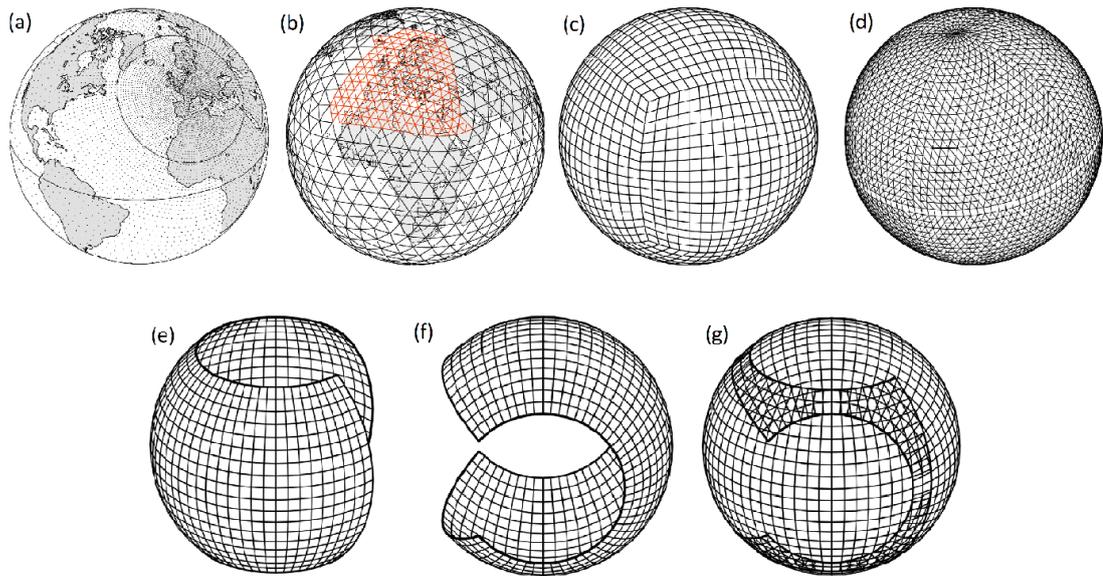


FIGURE 1.2 – Schémas représentant les mailles des principaux modèles globaux : (a) maille étirée du modèle français ARPEGE (Météo-France), (b) maille icosaédrique triangulaire du modèle allemand ICON (DWD), (c) maille « cubed-sphere » du modèle américain GFS (NCEP), (d) maille octaédrique gaussienne réduite du modèle européen IFS (CEPMMT), (e) grille Yin, (f) grille Yang et (g) maille « Ying-Yang » du modèle canadien GEM (CMC) (Adapté de Météo-France et Ulrich et al. (2017))

## 1.2.2 Paramétrisation physique

Les termes de forçage des équations définissant les tendances temporelles des variables pronostiques, ne doivent pas non plus être négligés. Contrairement au cœur dynamique, ces forçages concernent des phénomènes physiques ayant lieu à des échelles plus petites que celle du modèle et ne pouvant être résolus explicitement par celui-ci. Par exemple, aucun modèle météorologique n'a une résolution assez fine pour reproduire la formation de gouttes de pluie à partir de gouttelettes nuageuses ou la réflexion du rayonnement solaire sur les molécules composant l'air. Ces phénomènes physiques dits sous-maille comprennent essentiellement le rayonnement, la microphysique nuageuse (changement d'état des hydrométéores), la convection (interaction vitesse verticale et changements de phase, formation de nuages cumuliformes tels que les cumulus et cumulonimbus), la représentation de la couche limite atmosphérique et sa turbulence, les ondes de gravités orographiques et les caractéristiques de la surface (végétation, ville, ...). Pour rendre compte de ces processus physiques, des schémas dits de **paramétrisations**, basés sur des lois physiques statistiques, vont reproduire l'influence moyenne de ces phénomènes de petite échelle à l'échelle de la maille du modèle. Ces lois sont propres à chaque paramétrisation et estiment, par exemple, le chauffage, flux turbulents et source secondaire.

Plus précisément, ces paramétrisations intègrent des variables internes et les relient aux variables d'état à l'instant  $t$  en modélisant les processus à l'origine d'une altération des variables d'état. Concernant la température, ces schémas modélisent les processus diabatiques produisant un chauffage dû aux flux radiatifs, aux changements d'état de l'eau ou à la friction. D'autres forçages doivent être pris

en compte, tels que les flux turbulents et la production de quantités secondaires tels que les aérosols ou les composants chimiques de l'atmosphère.

Cependant, en raison d'un fort coût numérique mais aussi par manque de connaissance, ces schémas ne peuvent représenter l'ensemble des phénomènes physiques et leur complexité. Seuls les phénomènes physiques prépondérants sont modélisés et même simplifiés (Stensrud, 2011). Par exemple, ces processus sont considérés comme homogènes (homogénéité statistique) horizontalement ; les équations les régissant sont donc unidimensionnelles en  $z$  (hauteur). Autrement dit, ces paramétrisations représentent la physique uniquement sur une colonne atmosphérique indépendante (Hourdin, 2011).

Puisque chaque modèle utilise des paramétrisations qui lui sont propres, nous présentons par la suite les paramétrisations mises en place à Météo-France, et plus particulièrement dans les modèles AROME et ARPEGE (convection profonde).

**Rayonnement :** Le chauffage (tendance de température) dû au rayonnement étant proportionnel à la divergence des flux radiatifs, la paramétrisation du rayonnement doit estimer le bilan de flux radiatifs solaire et thermique au sommet et à la base de chaque couche atmosphérique supposée suffisamment fine pour que les propriétés optiques y soient constantes (Hogan, 2019; Seity, 2020). Pour ce faire, il faut prendre en compte différents phénomènes tels que :

- l'absorption du rayonnement par les différentes espèces chimiques présentes dans l'atmosphère ;
- la diffusion du rayonnement de l'échelle moléculaire à l'échelle d'un nuage ;
- la réflexion du rayonnement solaire à la surface terrestre due à son albédo ;
- l'émission infra-rouge par la Terre et l'atmosphère.

Ainsi, à partir du rayonnement solaire au sommet de l'atmosphère et son angle zénithal (dépendant de l'heure), mais aussi à partir des caractéristiques de la surface (albédo, température de surface, émissivité,...) et des caractéristiques de l'atmosphère (présence de nuages, gaz, aérosols,...), ce schéma peut, en théorie, calculer quasiment parfaitement la luminance sur l'ensemble de l'atmosphère. Cependant, réaliser ces calculs pour tout angle solaire d'incidence, longueur d'onde, épaisseur optique et espèce chimique donnés est très coûteux et ne peut être utilisé tel quel dans les modèles de prévision du temps. Pour remédier à ce problème, à chaque niveau vertical, la couche atmosphérique est supposée isotrope (étude du rayonnement uniquement sur la verticale) et avec des propriétés optiques homogènes. Ainsi, seuls les flux verticaux sont estimés. Par ailleurs, les nuages y sont représentés de manière approchée, c'est-à-dire en considérant un pourcentage de zone nuageuse au sein d'une maille de la grille du modèle et en utilisant des hypothèses sur le recouvrement nuageux sur la verticale.

Pour ce qui est du rayonnement infra-rouge, le schéma de paramétrisation est relativement similaire à celui du rayonnement solaire, seuls les phénomènes inhérents au rayonnement solaire tel que l'angle zénithal ne sont pas pris en compte.

**Microphysique nuageuse :** Les schémas de microphysique reproduisent l'évolution des différentes classes d'hydrométéores contenus dans l'atmosphère tels que l'eau nuageuse, les gouttes de pluie, les cristaux de glace primaire (aussi dit « pristine ice »), la neige, les graupels, la grêle, le grésil, ... En particulier, ils s'attachent à décrire les processus microphysiques exo ou endothermiques entre les différents hydrométéores. Ainsi, ils permettent d'estimer les flux de chaleur latente et donc de calculer le chauffage dû à ces processus. Historiquement, pour décrire ces hydrométéores, deux approches ont été mises en place : les modèles à moments dits « bulk » et les modèles détaillés (Planche, 2011; Khain et al., 2015).

Dans les schémas à moments, chaque classe d'hydrométéores est décrite par un spectre en nombre ou en masse totale des hydrométéores. La forme de ce spectre est fixée et suit en générale une forme exponentielle de type Marshall-Palmer, log-normale ou gamma. Celle-ci est décrite par les moments d'ordre  $i$  ayant une signification physique, c'est-à-dire, le moment d'ordre 0 correspond au nombre total de particules, celui d'ordre 3 à la masse totale et celui d'ordre 6 à la réflectivité radar (Milbrandt and Yau, 2005). La complexité de ces schémas de microphysique dépend ainsi du nombre de moments considérés pour décrire le spectre de chaque hydrométéore mais aussi du nombre de classes d'hydrométéores considérés. Les modèles les plus simples ne décrivent que deux classes (Kessler, 1969) alors que les modèles les plus aboutis peuvent aller jusqu'à cinq voire six classes (Hong and Lim, 2006) .

Les schémas de microphysique détaillés, également appelés spectraux ou « bin », considèrent, quant à eux, une distribution en taille des hydrométéores. Cette fois-ci, le spectre des hydrométéores ne suit pas une distribution fixée. Il est discrétisé en taille où chaque point de grille correspond à une boîte nommée « bin » dans la littérature anglaise ; ces schémas demandant ainsi un nombre de variables proportionnel au nombre de boîte (« bin »), soit environ 200-300 variables. Ainsi, les processus microphysiques déplacent les hydrométéores dans la distribution en taille : une gouttelette nuageuse grossissant pour donner une goutte de pluie se déplacera dans la boîte suivante dans la distribution en taille. Deux approches existent. Les modèles « bin » considèrent chaque boîte avec des intervalles de taille fixe où les hydrométéores changent de boîte en grossissant. Les méthodes lagrangiennes quant à elles, gardent un nombre constant d'hydrométéores par boîte mais modifient l'intervalle de chaque boîte. Malgré leur réalisme, ces schémas, de par le très grand nombre de variables et équations nécessaires, demandent un coût en mémoire et en temps de calcul important et ne sont donc pas utilisés en prévision numérique du temps opérationnel. Cependant, des alternatives mixant schémas détaillés et à moments ont été mises en place. Par exemple, le schéma d'Onishi and Takahashi (2012), utilise une représentation détaillée pour les espèces intervenant dans les processus chauds (vapeur, eau liquide) et une représentation à moments pour les processus froids (glace, eau surfondue, ...).

Ainsi, chaque hydrométéore peut être vu comme un réservoir dont le contenu varie en fonction des processus microphysiques. Pour représenter cette évolution, il faut prendre en compte les phénomènes de nucléation (condensation/sublimation sur des noyaux de condensation dès qu'il y a sursaturation), collision-collection (modification du moment d'ordre 0), autoconversion, accrétion (gouttelettes en pluie), aggrégation de cristaux de glace, givrage, fonte, évaporation, sublimation, éclatement de gouttelettes, sédimentation, etc. Ceux-ci, calculés uniquement si le

réservoir source n'est pas vide, suivent en général des équations empiriques complexes. Cependant, en raison de la définition des classes d'hydrométéores, certains processus deviennent des phénomènes à seuil dans les modèles. Par exemple, à partir d'une certaine taille (seuil d'autoconversion de la pluie), les gouttes nuageuses s'agrègent par collection pour former des petites gouttes de pluie soumises à la sédimentation. L'ordre de calcul des différents processus, propre à chaque schéma de microphysique, est important et définit la qualité du schéma. Les processus rapides interviennent en premier alors que la sédimentation agit toujours en dernier.

Parfois couplé à un schéma de diffusion horizontale, le schéma semi-lagrangien se charge ensuite d'advecter, par le vent moyen, les hydrométéores estimées par le schéma de microphysique. Enfin, un schéma de condensation sous maille vient souvent compléter le schéma microphysique, en permettant, via l'utilisation de distributions statistiques, de ne créer un nuage que sur une fraction de la maille du modèle.

**Convection profonde :** Pour représenter la formation de nuages cumuliformes au sein de la maille d'un modèle global, il est nécessaire de modéliser les mouvements d'ascendance dit « updrafts » à l'origine de la convection précipitante (Bechtold, 2019a,b,c).

Ne pouvant cependant représenter toutes les ascendances réelles sous-mailles, il faut alors modéliser une ascendance moyenne représentant le mouvement vertical total dans la maille. Il s'agit donc de calculer un flux de masse en prenant en compte les phénomènes d'entraînement (apport d'air sec de l'environnement dans l'ascendance) et de déentraînement (perte de l'ascendance dans l'environnement), tout en prenant en compte les phénomènes de microphysique générant le nuage cumuliforme (Bechtold, 2009). L'activation de ces schémas est soumise à une fonction de déclenchement qui peut dépendre de la vitesse verticale (Bretherton et al., 2004) et/ou de la flottabilité (Berg and Stull, 2005). Ainsi, différentes hypothèses de fermeture ont été proposées. On compte trois principaux schémas conceptuels (Piriou, 2020). Dans un premier temps, les schémas dits Conditional Instability of the First Kind (CIFK) s'appuient sur l'hypothèse que la convection est pilotée par la Convective Available Potential Energy (CAPE), c'est à dire l'énergie disponible pour qu'une particule d'air puisse s'élever et créer un cumulus. Dans ce cas, seul le mouvement ascendant est pris en compte. Dans les schémas Conditional Instability of the Second Kind (CISK), la convection est supposée pilotée par une convergence d'humidité en surface qui va générer l'ascendance. Ces modèles font donc intervenir un mouvement horizontal. Enfin, les modèles Wind Induced Surface Heat Exchange (WISHE) supposent une convection liée à une humidification dans les basses couches de l'atmosphère due à une forte évaporation au sol.

Tous les modèles de convection n'utilisent pas la notion de flux de masse (Bechtold, 2009). Par exemple, le schéma de Kuo (1974) se base sur un bilan d'humidité alors que le schéma de Betts and Miller (1986) calcule les tendances de température et humidité dues à la convection à partir d'un ajustement du profil de référence de température et d'humidité. Cependant, tous ces schémas modélisent les précipitations dues à la convection profonde.

**Turbulence / Convection peu profonde :** Pour représenter la couche limite atmosphérique et ses tourbillons, les schémas de turbulence considèrent que ces tourbillons peuvent être représentés comme les fluctuations de température et de vent autour d'un état moyen de l'atmosphère, selon la décomposition de Reynolds (Stull, 1988). Le système d'équations obtenu à partir des équations de Navier-Stokes comprend des termes dits de Reynolds dont l'estimation est l'enjeu principal des schémas de turbulence. L'une des principales méthodes consiste à relier ces flux turbulents aux gradients verticaux de vent et température par un coefficient de diffusion (K-gradient) (Sandu, 2019a,b,c). Différentes approches ont été mises en place pour estimer ces coefficients (théorie de la similitude de Monin-Obukhov, définition des profils de K de Troen et Mahrt,...). Pour les modèles de grande échelle, ces transports sont calculés en fonction des variables du modèle (vent, humidité, température). Cependant, pour les modèles de moyenne échelle, tels que les modèles résolvant la convection profonde, il est nécessaire de prendre en compte l'intensité de la turbulence en représentant l'évolution diurne de l'énergie cinétique turbulente.

Au delà des modèles de turbulence locale, des modèles de convection peu profonde s'attachent à modéliser les ascendances et la turbulence pour la convection peu précipitante. Ces modèles utilisant la méthode Eddy-Diffusivity/Mass-Flux peuvent à la fois modéliser la turbulence de petites échelles en faisant l'analogie avec la diffusion mais aussi les thermiques (ascendance) à l'origine de la convection peu profonde avec la notion de flux de masse (Siebesma and Teixeira, 2000). Ainsi, les flux turbulents humides  $y$  sont décrits à la fois par la méthode des K-gradients et par un flux de masse.

**Surface :** Les schémas de surface permettent de modéliser l'évolution de la surface. Ils estiment en particulier l'évolution de la température et l'humidité du sol connaissant les caractéristiques de ce sol (relief, albédo, longueur de rugosité) et ses flux radiatifs (Balsamo, 2019a,b,c). Ainsi, chaque maille du modèle est caractérisée par différents types de surface en proportions variables (ville, végétation, mer, lac). Pour chacun d'eux, l'estimation des flux de chaleur latente et sensible définissant les tendances d'humidité et température est différente. En particulier, pour la végétation, cette estimation dépend de l'état de la végétation représentée par, entre autres, le LAI (Leaf Area Index; Bréda, 2008). L'évolution de l'albédo est aussi prise en compte en modélisant la fonte de la neige. Notons que la température du sol et l'albédo calculés par le schéma de surface sont utilisés dans le schéma de rayonnement pour estimer le rayonnement solaire réfléchi et le rayonnement infra-rouge émis par la surface.

Ces schémas s'intéressent aussi aux effets du relief sur l'atmosphère (onde de gravité, traînée orographique, etc). Ainsi, ils calculent aussi les tendances du vent proportionnelles à la divergence du flux de quantité de mouvement tout comme le schéma de turbulence (cf ci-dessous). Comme les modèles atmosphériques représentent une atmosphère commençant au premier niveau du modèle situé à quelques mètres du sol, les quantités calculées par les schémas de surface sont reportées à ce premier niveau vertical du modèle et non à la surface-même.

### 1.3 Aspect chaotique de l'atmosphère

---

« Pourquoi les météorologistes ont-ils tant de peine à prédire le temps avec quelque certitude ? Pourquoi les chutes de pluie, les tempêtes elles-mêmes nous semblent-elles arriver au hasard ... un dixième de degré en plus ou en moins en un point quelconque, le cyclone éclate ici et non pas là ... Si on avait connu ce dixième de degré, on aurait pu le savoir d'avance, mais les observations n'étaient ni assez serrées, ni assez précises, et c'est pour cela que tout semble dû à l'intervention du hasard. Ici encore nous retrouvons le même contraste entre une cause minime, inappréciable pour l'observateur, et des effets considérables, qui sont quelquefois d'épouvantables désastres. »

---

Henri Poincaré - Science et Méthode (1908)

En parallèle de toutes les avancées sur la modélisation de l'atmosphère, une découverte va révolutionner la vision de la prévisibilité des états futurs de l'atmosphère, passant d'une vision déterministe à une vision ensembliste.

Dès la fin du XIX<sup>e</sup> siècle, Henri Poincaré s'est intéressé au problème à trois corps, c'est-à-dire aux interactions gravitationnelles entre 3 corps. Ces travaux menant à la découverte d'une sensibilité de la solution des équations aux conditions initiales, posent les axes de la théorie du chaos. En particulier, Henri Poincaré y propose de nombreux concepts spécifiques à cette théorie (espace de phases, notion de bifurcation, ... ; Letellier, 2020a,b).

Par la suite, en 1963, Edward Lorenz s'est intéressé à la convection naturelle de Rayleigh-Bénard (Lorenz, 1963). Ce phénomène se produit, lorsqu'un fluide est situé dans un gradient thermique vertical important : une source chaude en dessous et une source froide au dessus. La surface chaude située en bas va réchauffer le fluide à son contact par conduction. Or, à partir d'une certaine valeur du nombre de Rayleigh (nombre caractérisant la proportion entre conduction et convection), le fluide va perdre en densité, être plus léger que son environnement et va donc s'élever par la poussée d'Archimède. Une fois arrivée au contact de la source froide, cette parcelle de fluide va se refroidir, s'alourdir en gagnant en densité et va donc retomber. Ces rouleaux de convection sont bien connus des atmosphériciens puisqu'ils ont lieu au sein de la couche limite et expliquent les cellules de Bénard.

Dans son étude de la convection naturelle, Edward Lorenz a simplifié les équations jusqu'à obtenir un système dynamique de trois équations différentielles à 3 degrés de libertés :

$$\frac{dx}{dt} = Pr (y(t) - x(t)), \quad (1.1)$$

$$\frac{dy}{dt} = Ra_r x(t) - y(t) - x(t) z(t) \quad (1.2)$$

$$\frac{dz}{dt} = x(t) y(t) - \beta z(t), \quad (1.3)$$

où  $x$  est proportionnel à l'intensité du mouvement de convection,  $y$  à la différence de température entre les ascendances et subsidences, et  $z$  à l'écart du profil vertical de température au profil linéaire. Les degrés de liberté sont  $Pr$  le Nombre de Prandtl,  $Ra_r$  le nombre de Rayleigh réduit ( $Ra_r = Ra/Ra_c$ , rapport entre le nombre de Rayleigh et le nombre de Rayleigh critique correspondant à la situation où la convection commence à l'emporter sur la conduction) et  $\beta$  la longueur caractéristique de l'expérience  $\beta = 8/3 > 0$ .

Grâce à l'arrivée des ordinateurs, Lorenz calcule les 6000 premières itérations de la solution de ce système pour  $Pr = 10$ ,  $Ra_r = 28$  et  $\beta = 8/3$ . Il montre en particulier l'instabilité de la solution à long terme. En représentant la solution du système dans l'espace des phases  $(x, y, z)$  à l'instar d'Henri Poincaré, Lorenz découvre que la trajectoire du système se ramène toujours à la même structure topologique quelles que soient les conditions initiales (excepté pour les conditions initiales situées aux points fixes). Cette structure en forme de papillon est aujourd'hui connue sous le nom d'attracteur étrange de Lorenz, apportant étonnement de l'ordre dans le chaos atmosphérique. Cependant, ces trajectoires restent non périodiques, rendant le système chaotique.

Ainsi, toutes les solutions du système étant instables à long terme, le système devient sensible aux conditions initiales. Une différence minime dans ces conditions initiales amènent à long terme à des différences notoires de la position du système sur l'attracteur étrange et donc des différences notoires de l'état de l'atmosphère.

La notion de prévisibilité apparaît alors. On définit ainsi l'horizon de prévisibilité, l'instant au delà duquel le système initialisé par des conditions proches prend une trajectoire radicalement différente. En dessous de cet horizon de prévisibilité, l'erreur quadratique moyenne d'une prévision est inférieure à la différence quadratique moyenne entre deux états de l'atmosphère choisis aléatoirement (Joly and Descamps, 2009). Au delà de cet horizon de prévisibilité, il est en revanche, impossible de prévoir les états futurs de l'atmosphère.

La prévisibilité des phénomènes météorologiques est par ailleurs liée à leur taille caractéristique. L'horizon de prévisibilité est beaucoup plus faible pour un phénomène de petite taille que pour un phénomène d'échelle synoptique. Ainsi, il est donc illusoire de chercher à prévoir une structure au delà de sa durée de vie. Par exemple, cet horizon se situe autour de huit à quinze jours pour les régimes de temps, deux à trois jours pour les dépressions, ou quatre à huit heures pour la convection profonde (cumulonimbus) (Joly and Descamps, 2009).

Enfin, d'un point de vue plus pratique, les réseaux d'observations n'étant pas parfaits que ce soit au niveau de l'incertitude de mesures ou de la répartition spatiale des stations de mesures, les conditions initiales des modèles météorologiques opérationnels ne peuvent être connues parfaitement. Ainsi, en raison de la sensibilité aux conditions initiales et à l'aspect chaotique de l'atmosphère vue pré-

cédemment, cette incertitude rend les prévisions météorologiques opérationnelles incertaines. Aux incertitudes des conditions initiales s'ajoutent les incertitudes des modèles eux-mêmes, pouvant amener à des prévisions différentes.

## 1.4 Illustration des incertitudes des modèles dans le cas d'une dépression des moyennes latitudes

L'objectif de cette section est d'illustrer l'incertitude de la prévision d'une dépression à l'aide du modèle global français ARPEGE. Plus particulièrement, on se focalisera sur les incertitudes liées à la représentation de la convection profonde mais aussi à la valeur même d'un paramètre incertain dans la représentation de la microphysique nuageuse. Ces illustrations sont tirées d'études préliminaires effectuées en partenariat avec le Laboratoire de Météorologie Dynamique. Elles ont été présentées à de nombreux workshop et font l'objet actuellement d'articles soumis (Rivière et al., 2021; Wimmer et al., 2021) au journal *Weather and Climate Dynamics* dont un publié (Rivière et al., 2021).

### 1.4.1 Dépression Stalactite du 30/09/2016 au 02/10/2016 et bandes transporteuses d'air chaud

Nous allons nous intéresser à la dépression nommée « Stalactite », particulièrement bien étudiée pendant la période d'observation intense n°6 de la campagne de mesure internationale NAWDEX (North Atlantic Waveguide Downstream impact Experiment ; Schäfler et al., 2018). Cette dépression, née au large de Terre-Neuve le 30 septembre 2016, s'est intensifiée lors de sa traversée de l'Océan Atlantique Nord. Elle a atteint sa maturité près de l'Islande avec un minimum dépressionnaire avoisinant les 960hPa le 02 octobre 2016 à 21hUTC. En particulier, quelques heures avant ce point culminant, l'avion instrumenté Falcon 20 de l'équipe Service des Avions Français Instrumentés pour la Recherche en Environnement (SAFIRE) a réalisé, entre autres, le vol F7, échantillonnant une zone soumise à de nombreux processus diabatiques dont l'estimation est incertaine, la bande transporteuse d'air chaud de la dépression. Ce vol se situant exactement dans la zone nuageuse d'ascendance de la bande transporteuse d'air chaud (voir figure 1.3), il servira de base pour les deux illustrations de l'incertitude de la prévision du modèle ARPEGE.

Cette dépression et son évolution sont particulièrement intéressantes pour plusieurs raisons. D'une part, cette tempête a connu un creusement explosif avoisinant les 26hPa en 24h, faisant d'elle une bombe météorologique. D'autre part, son intensification est étroitement liée à la prévisibilité d'un anticyclone sur la Scandinavie et d'un changement de régime de temps sur l'Océan Atlantique Nord. En effet, la figure 1.4 montre la corrélation entre le champ de géopotential à 500hPa de l'analyse (prévision corrigée par les observations) ARPEGE et différents régimes de temps, pendant le cycle de vie de la dépression Stalactite et le blocage (anticyclone scandinave) succédant. Plus la corrélation avec un régime de temps est élevée, plus la situation météorologique s'apparente au régime de temps correspondant.

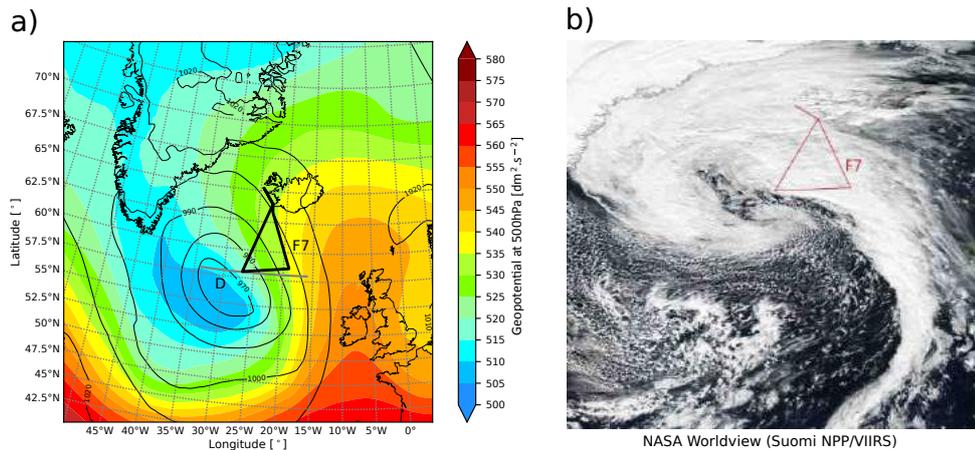


FIGURE 1.3 – *A gauche, géopotential à 500hPa en couleur et pression au niveau de la mer en isoligne, le 2 octobre 2016 à 12hUTC. Le vol F7 du Falcon20 de SAFIRE réalisé pendant la période d’observation intense n°6 de NAWDEX apparaît en trait noir épais, la section verticale en gris. A droite, image dans le canal visible du capteur VIIRS du satellite Suomi NPP (NASA Worldview) avec la trajectoire du vol F7 du Falcon 20 de SAFIRE en rouge*

Au niveau de l’Océan Atlantique Nord, quatre régimes de temps sont possibles. Le premier concerne les situations dites de blocage scandinave, assimilées à un anticyclone d’altitude prépondérant sur la péninsule scandinave comme décrit sur la figure 1.4b. Deux autres types de circulation avec un rail dépressionnaire zonal sont possibles. Ceux-ci se basent sur l’indice d’Oscillation Nord Atlantique (North Atlantic Oscillation (NAO)), défini comme l’écart par rapport à la moyenne climatique de la différence de pression entre Lisbonne et Reykjavik. Lorsque cet indice est positif, donc régime de temps NAO positif, le rail des dépressions est situé plutôt vers le Nord de l’Europe, tandis que pour un indice NAO négatif, les dépressions ont tendances à aller vers le Sud de l’Europe. Le dernier régime de temps est la dorsale atlantique où un anticyclone s’étend sur l’océan Atlantique Nord, forçant les dépressions à arriver par le Nord de l’Europe. Sur la figure 1.4 a, la dépression Stalactite naît dans un régime NAO positif assez marqué. Mais au cours de maturation de la dépression, un régime de blocage va se mettre en place, suggérant une intensification d’un anticyclone sur la Scandinavie. Ce changement de situation, coïncidant avec la phase de maturation du cyclone extra-tropical, est ainsi étroitement lié à la dynamique de la dépression. Il est donc fondamental de bien prévoir l’évolution de la tempête et de son impact sur l’ondulation du Jet Stream et l’anticyclone d’altitude en aval du jet. Plus particulièrement, la zone d’intérêt se situe dans la bande transporteuse d’air chaud de la dépression Stalactite.

En effet, ces bandes transporteuses d’air chaud sont des courants de masse d’air intervenant dans les dépressions et amenant l’air chaud et humide des tropiques vers les pôles. Lors de leur déplacement vers les pôles, ces masses d’air s’élèvent en altitude et forment des nuages par condensation due à l’ascendance. Ce transport lagrangien comporte de nombreux phénomènes diabatiques donc sources de chauffage, tels que la condensation, le refroidissement radiatif en haute altitude

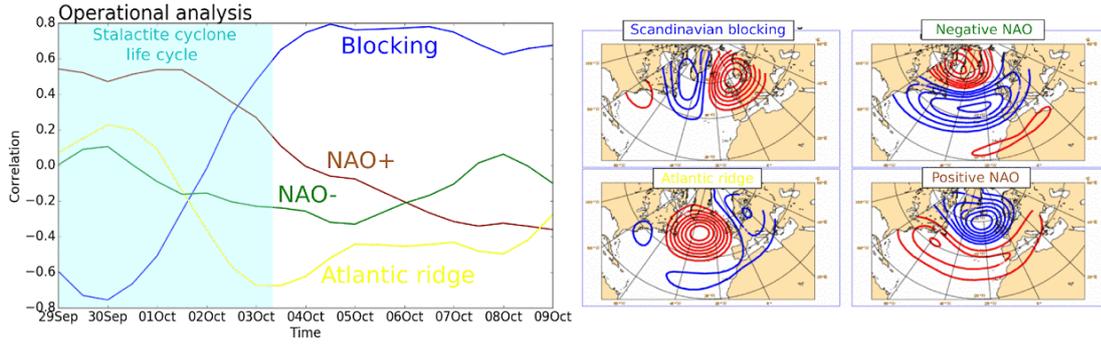


FIGURE 1.4 – A gauche, évolution temporelle durant la campagne NAWDEX de la corrélation entre le champ de géopotential à 500 hPa des analyses ARPEGE et les 4 régimes de temps. A droite, anomalie de géopotential pour ces 4 régimes de temps hivernaux de l’Océan Atlantique Nord. (Avec l’aimable autorisation de Gwendal Rivière)

ou encore la friction due à la turbulence. Ceux-ci ont un effet direct sur l’onde de Rossby du Jet Stream (Schäfler et al., 2018) de par leur altération de la Vorticité Potentielle (PV). Cette quantité conservative dans une atmosphère adiabatique, caractérise la rotation d’une parcelle d’air par rapport à une parcelle d’air située à l’équateur, selon l’équation suivante (Ertel, 1942) :

$$q = \frac{1}{\rho} ((2\boldsymbol{\Omega} + \boldsymbol{\nabla} \wedge \mathbf{v}) \cdot \boldsymbol{\nabla} \theta) \quad (1.4)$$

où  $q$  est la vorticité potentielle,  $\rho$  la masse volumique de l’air,  $\boldsymbol{\Omega}$  la vitesse angulaire de la rotation de la Terre,  $\mathbf{v}$  le vent,  $\theta$  la température potentielle. La tendance temporelle du PV est donc liée au gradient de chauffage  $\boldsymbol{\nabla} \dot{\theta}$  par la relation suivante (Stoelinga, 1996) :

$$\frac{dq}{dt} = \frac{1}{\rho} \left( (2\boldsymbol{\Omega} + \boldsymbol{\nabla} \wedge \mathbf{v}) \cdot \boldsymbol{\nabla} \dot{\theta} + \boldsymbol{\nabla} \wedge \mathbf{F} \cdot \boldsymbol{\nabla} \theta \right) \quad (1.5)$$

où  $\mathbf{F}$  est la force de frottement. Ainsi, lorsqu’il y a un maximum de chauffage, il y a production de PV en dessous et à l’Ouest de ce chauffage et destruction au dessus et à l’Est (Joos and Wernli, 2012). Ainsi, une altération du PV due à la présence d’un chauffage dans la bande transporteuse d’air chaud entraîne nécessairement une modification du rotationnel de vent, donc de l’ondulation du Jet Stream d’altitude. Comme suggéré par Dirren et al. (2003) et Schäfler et al. (2018), il est donc nécessaire de modéliser au mieux les phénomènes diabatiques au sein des bandes transporteuses d’air chaud.

### 1.4.2 Incertitudes liées à une paramétrisation : sensibilité aux schémas de convection profonde

Il s’agit ici d’étudier l’impact de différentes représentations numériques de la convection profonde sur la prévision de la dynamique de la dépression Stalactite. Deux simulations de cette dépression avec le modèle global ARPEGE sont donc

analysées. La première utilise le schéma de convection profonde de Bougeault (1985), nommé B85 et utilisé dans la version opérationnelle d'ARPEGE. La seconde comprend le schéma Prognostic Condensates Microphysics and Transport (PCMT), décrit dans Piriou et al. (2007) et intervenant dans la version climatique du modèle ARPEGE (ARPEGE-Climat). Ces deux schémas utilisent le concept de flux de masse pour décrire la convection. Cependant, le schéma de B85 a une fermeture en humidité alors que PCMT a une fermeture en CAPE. Par ailleurs, couplé aux schémas de microphysique et de transports, PCMT présente un plus fort entraînement.

Ces simulations sont toutes deux initialisées au 01 octobre 2016 à 12h, soit 24h avant la phase de maturité de la dépression et entre 26 et 28h avant la réalisation du vol F7. Le modèle ARPEGE ayant une maille étirée, la version utilisée a une résolution horizontale d'environ 10km sur la France et 15 sur l'Islande. En revanche, les sorties du modèle sont à 0.5° de longitude et latitude. Le pas de temps du modèle a été modifié à 450s.

Pour représenter la bande transporteuse d'air chaud, des trajectoires et des rétrotrajectoires lagrangiennes de 24h initialisées le long du vol F7 ont été calculées grâce à un modèle de trajectographie inspiré d'un schéma de type semi-lagrangien (Rivière et al., 2021). Mis bout à bout, ces trajectoires et rétrotrajectoires forment des trajectoires lagrangiennes de 48h. Seules celles répondant au critère de la bande transporteuse d'air chaud, c'est à dire avec une ascendance de 300hPa en 24h glissantes, ont été sélectionnées. Remarquons que le critère choisi ici est un peu moins strict que celui proposé par Joos and Wernli (2012) (ascendance de 600hPa en 48h) mais reste représentatif des bandes transporteuses d'air chaud.

La figure 1.5 représente les trajectoires de la bande transporteuse d'air chaud croisant le vol F7. Sur les figures 1.5a et b, les trajectoires issues des simulations B85 et PCMT respectivement, se déplacent du Sud vers le Nord en passant de 900hPa à 350hPa. Celles-ci s'élèvent bien au fur et à mesure qu'elles vont vers les pôles, il s'agit bien de la bande transporteuse d'air chaud. Après avoir croisé le vol F7 représenté en trait noir, la bande transporteuse d'air chaud se sépare en deux catégories : une première branche s'enroule autour du coeur cyclonique de la dépression et une seconde est emportée par les flux anticycloniques d'altitude.

Les figures 1.5c et d montrent une vue méridionale de ces trajectoires pour B85 et PCMT respectivement. Cette fois-ci, il s'agit du chauffage représenté en couleur le long de ces trajectoires. On constate un chauffage net d'environ 2K/h pour les deux simulations dans la partie ascendante de la bande transporteuse d'air chaud. Ce chauffage est principalement dû à la condensation de l'eau lorsque l'air s'élève et se refroidit. En dessous de l'iso-0°C marqué par les points violets, un refroidissement (bleu) est visible pour les deux simulations mais est beaucoup plus marqué avec B85. Ceci est le résultat de la fonte endothermique de la glace lorsqu'elle passe sous l'iso-0°C. On note en particulier une nette différence de chauffage entre les deux simulations. PCMT présente un plus fort chauffage en dessous de l'iso-0°C comparé à B85. Ceci signifie un chauffage plus bas avec le schéma de convection profonde PCMT ayant une conséquence directe sur le PV et le vent.

Les figures 1.6b et c représentent respectivement pour B85 et PCMT, une coupe verticale, le 2 octobre 2016 à 15h (soit 27h d'échéance) et à 58°N (temps et latitude du deuxième segment zonal du vol F7). Le vent horizontal est représenté en

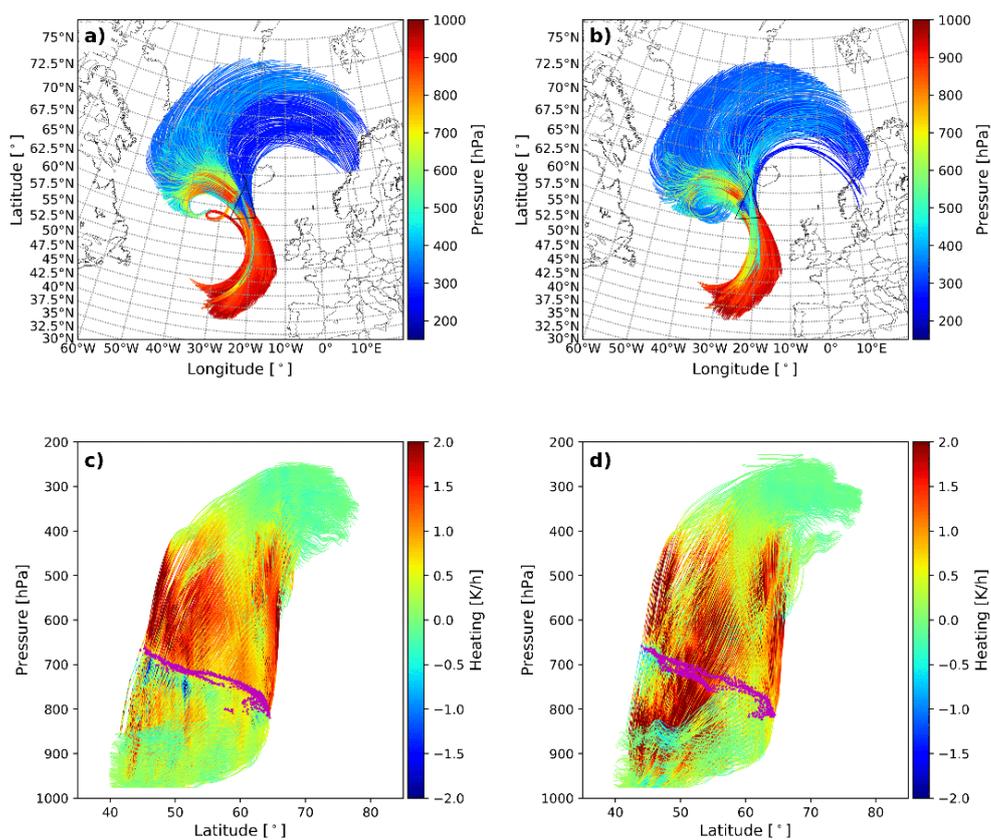


FIGURE 1.5 – *En haut* : Carte montrant la position et la pression (en couleur) le long des trajectoires lagrangiennes de 48h répondant au critère de la bande transporteuse d'air chaud et croisant le vol F7 (représenté en trait noir) pour a) la simulation avec B85 et b) avec PCMT. *En bas* : Vue méridionale du chauffage le long des trajectoires de la bande transporteuse d'air chaud pour c) la simulation avec B85 et d) avec PCMT. Les points violets représentent l'intersection des trajectoires avec l'iso-0°C.

couleur tandis que le PV apparaît en isoligne. Pour mieux visualiser la séparation de l'air stratosphérique de l'air troposphérique par la tropopause située à 2PVU, la stratosphère est représentée par les zones en pointillés. Afin d'estimer le réalisme de ces deux simulations, nous comparons ces coupes verticales à celle obtenue avec l'analyse du modèle du CEPMMT, IFS (Figure 1.6a). Nous ne considérerons pas ici l'analyse du modèle ARPEGE, car celle-ci comprend le schéma de convection profonde de Bougeault (1985), et ne peut donner des éléments de comparaison acceptables. Par opposition, le modèle IFS utilise un autre schéma de convection profonde, celui de Tiedtke (1989).

Concernant le vent horizontal, pour les trois simulations, un fort vent de près de 70m/s (250km/h) est visible en haute altitude. Celui-ci correspond au Jet Stream. Sa position horizontale varie très peu pour les trois simulations, précisément autour de 27°W. Sur la verticale, celui-ci présente un maximum autour de 375hPa pour les trois simulations en revanche, son extension verticale est différente. Des vents de plus de 40m/s (150km/h) s'étendent jusqu'à 550hPa dans l'analyse IFS et 575hPa avec B85. Cependant, ils descendent jusqu'à 650hPa avec PCMT. Ainsi, le Jet Stream est plus bas en altitude avec le schéma PCMT qu'avec B85. B85 étant plus proche de l'analyse IFS, celui-ci semble être le plus réaliste.

Cette différence de vent sur la verticale est liée à une différence de PV. En effet, on remarque que des valeurs stratosphériques de PV situées à l'Ouest du Jet Stream, atteignent des altitudes beaucoup plus basses dans la simulation avec PCMT, jusqu'à 650hPa. Par contre, cette anomalie de tropopause est moins creusée avec B85 et atteint seulement 450hPa. En parallèle, une anomalie négative de PV située à l'Est du Jet Stream est beaucoup plus marquée dans les deux simulations ARPEGE que l'analyse IFS. Celle-ci est située autour de 400 hPa avec B85 et 450hPa dans PCMT. Cette anomalie négative va générer une circulation anticyclonique qui va renforcer le Jet Stream. Ces anomalies étant plus basses en altitude avec PCMT, le Jet Stream est donc un peu plus bas dans cette simulation.

L'origine de cette anomalie de PV négative plus basse dans la simulation avec PCMT est en lien direct avec les résultats montrés à la figure 1.5. En effet, le chauffage de la bande transporteuse d'air chaud (située à l'Est du Jet Stream) étant situé dans des couches de l'atmosphère plus basses avec PCMT, il va générer une destruction de PV à des altitudes plus basses, conformément à l'équation 1.5. C'est pourquoi, l'anomalie négative de PV ainsi que le Jet Stream sont plus profonds avec PCMT qu'avec B85.

Ainsi, modifier la représentation de la convection profonde dans un modèle global ne se limite pas à une altération de la modélisation des orages. On montre ici que ces paramétrisations peuvent avoir un impact sur la dynamique même d'une dépression de par l'altération différente du PV qu'elles engendrent. En particulier, la structure verticale du PV est différente entre les deux schémas de convection profonde. Le schéma PCMT semble générer un chauffage en trop basse altitude, provoquant une descente erronée du Jet Stream. Or, une mauvaise prévision du Jet Stream peut avoir des répercussions sur la prévision des anticyclones en aval, comme la situation de blocage suivant la dépression Stalactite.

Remarquons que dans cette étude, le schéma de convection profonde B85 semble le plus réaliste comparé aux analyses. Cependant, il ne faut pas conclure trop hâtivement que ce schéma est le meilleur. En effet, dans un des articles soumis concernant cette étude (Rivière et al., 2021), il a été montré que pour deux

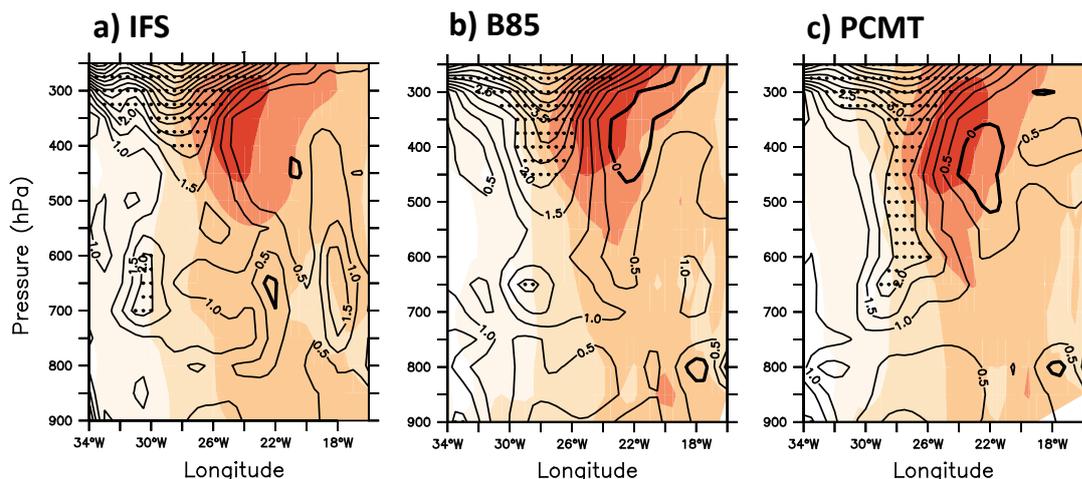


FIGURE 1.6 – Coupes verticales, le 2 octobre 2016 à 15hUTC, à 58°N (latitude du deuxième segment du vol F7) de vent zonal (en couleur) et vorticité potentielle (en iso-ligne noir ; les zones de vorticité potentielle supérieure à 2PVU sur marquée par des pointillés) pour a) l'analyse IFS, b) B85 et c) PCMT.

échéances différentes, les conclusions diffèrent. Aux courtes échéances, le schéma B85 représente mieux une structure double du Jet Stream à 24h d'échéance, comparé à l'analyse IFS mais aussi aux données du vol F6. Mais cette structure, trop bien marquée, reste encore présente, à tort, aux plus longues échéances (30h d'échéance). Alors que le schéma PCMT, même s'il peine à reproduire la structure double aux courtes échéances (24h), modélise assez justement le Jet Stream aux longues échéances (30h).

Notons par ailleurs que la même analyse opérationnel ARPEGE utilisant B85 a été utilisée comme conditions initiales. Ceci n'est pas sans conséquence sur les résultats puisque les conditions initiales ont un fort impact sur les prévisions à courtes échéances, expliquant ainsi en partie la meilleure performance de B85 à 24h d'échéance.

De part leurs impacts sur le chauffage et par conséquent sur la vorticité potentielle et le vent, les schémas de convection profonde peuvent engendrer des prévisions proposant des positions et évolutions différentes du Jet Stream, illustrant ainsi l'incertitude sur la dynamique de la dépression Stalactite. Une seule modélisation de la convection profonde ne pouvant prévoir la dynamique de la dépression Stalactite parfaitement, il est donc nécessaire de prendre en compte les deux simulations pour les prévisions météorologiques.

**Note :** Les résultats issus de cette étude sur l'impact des schémas de convection profonde sur la bande transporteuse d'air chaud de la dépression Stalactite font actuellement l'objet de deux articles soumis au journal *Weather and Climate Dynamics*. Ceux-ci sont disponibles en annexes avec de plus amples informations. Pour le premier article (Rivière et al., 2021), déjà publié, j'ai produit les trajectoires lagrangiennes et analysé les données ainsi que réalisé l'ensemble des études préliminaires. L'article a principalement été écrit par Gwendal Rivière, premier auteur. Cependant, j'ai participé à la rédaction de plusieurs paragraphes tels que le 2.1.1 ainsi qu'à la correction et relecture. J'ai rédigé la quasi totalité du deuxième

article (à l'exception de l'introduction), produit encore une fois les trajectoires lagrangiennes, analysé et réalisé la plupart des graphiques montrés.

### 1.4.3 Incertitudes liées à la paramétrisation-même : sensibilité aux paramètres de la microphysique

L'incertitude des modèles météorologiques ne se limite pas uniquement aux schémas utilisés pour représenter les phénomènes physiques. Au sein-même de ces paramétrisations, la présence de paramètres dont la valeur est incertaine est aussi responsable des erreurs de prévisions. Pour illustrer cet effet, nous proposons ici d'étudier l'impact de la vitesse de chute des flocons de neige modélisés dans le schéma de microphysique de Lopez (2002) sur le contenu en glace dans les nuages avec le modèle global ARPEGE.

Tout comme précédemment, l'étude se concentre sur le vol F7 réalisé pendant la période d'observation intense n°6 de NAWDEX. Il s'agit de comparer deux simulations ARPEGE utilisant le schéma de convection profonde PCMT (utilisé dans ARPEGE-Climat), mais différant par la valeur de vitesse verticale des flocons de neige modélisée dans le schéma de microphysique nuageuse de Lopez (2002). Plus particulièrement, une première expérience utilise la valeur actuelle par défaut d'ARPEGE, c'est-à-dire une vitesse de 1.5m/s, tandis que la deuxième utilise la valeur de 0.6m/s, valeur utilisée à l'origine dans les anciennes versions d'ARPEGE. Ce changement de valeur avait été réalisé pour une meilleure représentation des cirrus (nuages glacés d'altitude) dans ARPEGE. En effet, Mitchell et al. (2008) montre qu'une modification de la vitesse de chute de la neige entraîne une couverture nuageuse des cirrus suffisamment importante pour que le forçage radiatif net dû aux nuages atteigne  $-5W.m^{-2}$  aux tropiques. Les mêmes conditions initiales que celles considérées dans l'étude de la sous-section précédente sont utilisées.

Pour visualiser l'impact de ce paramètre sur les prévisions ARPEGE, nous allons nous focaliser sur le contenu en glace des nuages le long du vol F7. A partir des mesures de la plateforme RALI embarquée à bord du Falcon 20 et comprenant un radar et un lidar, l'algorithme VarCloud (mis en place à l'origine pour ces mêmes appareils embarqués sur les satellites CloudSat et CALIPSO), permet de déduire avec précision le contenu en glace le long de la trajectoire de l'avion instrumenté (Delanoë et al., 2010). La figure 1.7a représente ainsi ce contenu en glace déduit des mesures de la plateforme RAdar-LIdar (RALI), le long du vol F7. Pour faciliter la comparaison avec les simulations, les données télédéteectées sont interpolées sur la grille horizontale de sortie du modèle ARPEGE ( $0.5^\circ$  de longitude et latitude). Les zones blanches, correspondent aux zones sèches et aux zones nuageuses sans glace. Il ne s'agit donc pas d'une visualisation de la présence de nuage, car le contenu en eau liquide nuageuse, entre autres, n'est pas représenté.

Les figures 1.7b et c correspondent aux contenus en glace prévus par les deux simulations d'ARPEGE pour des valeurs différentes de vitesse de chute de la neige (respectivement 1.5m/s et 0.6m/s). Seules les valeurs prévues à l'emplacement des données mesurées sont représentées (application du même masque que les données RALI). Tout d'abord, nous constatons que pour les deux simulations, de par la résolution horizontale d'environ 15km du modèle ARPEGE sur cette zone, les fines structures de méso-échelles apparaissant dans le contenu en glace mesuré,

sont inexistantes dans les prévisions. Malgré cet aspect lissé des données prévues, les deux simulations sous-estiment grandement les contenus en glace le long du vol F7. Avec une vitesse de chute de 1.5m/s, ce contenu atteint jusqu'à  $500g/m^3$  (fig. 1.7b) alors que la simulation avec 0.6m/s peut modéliser des contenus plus élevés jusqu'à  $900g/m^3$  (fig. 1.7c), ce qui est très éloigné des  $7000g/m^3$  mesurés (fig. 1.7b). Cependant, nous constatons bien une plus forte densité de glace nuageuse dans les basses couches atmosphériques comme dans les observations télédéteectées.

Notons que quelle que soit la valeur de vitesse de chute de la neige utilisée dans le modèle ARPEGE, celui-ci prévoit mal le contenu en glace. Cette sous-estimation des modèles atmosphériques a déjà été observée pour des modèles climatiques entre autres (par exemple, Heymsfield et al., 2016). Cependant, l'ancienne valeur de vitesse de chute de 0.6m/s semble mieux appropriée pour modéliser le contenu en glace dans les nuages au sein de la bande transporteuse d'air chaud de la dépression Stalactite. En effet, une plus forte vitesse de chute des flocons de neige va engendrer un plus fort taux de précipitation de la neige et donc réduire le contenu en glace dans les nuages. Ceci explique le plus faible contenu en glace observé pour une vitesse de chute de 1.5m/s.

Ainsi, le paramètre réglant la vitesse de chute de la neige dans le modèle ARPEGE est un bon exemple d'incertitude du modèle due aux paramètres. Des valeurs différentes peuvent amener à des contenus en glace presque doublés, passant de  $500g/m^3$  avec 1.5m/s à  $900g/m^3$  avec 0.6m/s. Heymsfield et al. (2016) ayant montré que la vitesse de chute est proportionnelle au contenu en glace, fixer la vitesse de chute de la neige à une valeur unique pour des raisons de simplification des schémas de microphysique est ainsi une source d'erreur non négligeable du contenu en glace dans les modèles globaux. Pour les modèles régionaux, cette vitesse de chute n'est pas fixée et est mieux estimée. Cependant, d'autres paramètres sont tout autant incertains.

**Conclusion :** Pour améliorer la prévision des phénomènes météorologiques et repousser l'horizon de prévisibilité, il s'avère nécessaire de réduire les sources d'incertitudes des modèles météorologiques déterministes telles que celles dues aux paramétrisations physiques ou à leurs paramètres. Cependant, en raison de l'aspect chaotique de l'atmosphère, il est illusoire de chercher à mettre au point le modèle météorologique parfait. Il faut donc plutôt prendre en compte ces sources d'incertitudes dans les prévisions. Plutôt que de prévoir un unique état futur de l'atmosphère, on peut chercher à prévoir l'ensemble des états futurs possibles, soit chercher à estimer la distribution de probabilité des états futurs de l'atmosphère. Une possibilité est de calculer explicitement cette distribution de probabilités par la loi de Liouville et Fokker-Plank. Cependant, ces lois sont impossibles à appliquer pour des systèmes de grandes tailles tels que l'atmosphère. Une solution apparue dans les années 1990 est d'évaluer la probabilité d'occurrence d'un évènement et de quantifier l'incertitude qui pèse sur sa prévision via la technique de **prévision d'ensemble**.

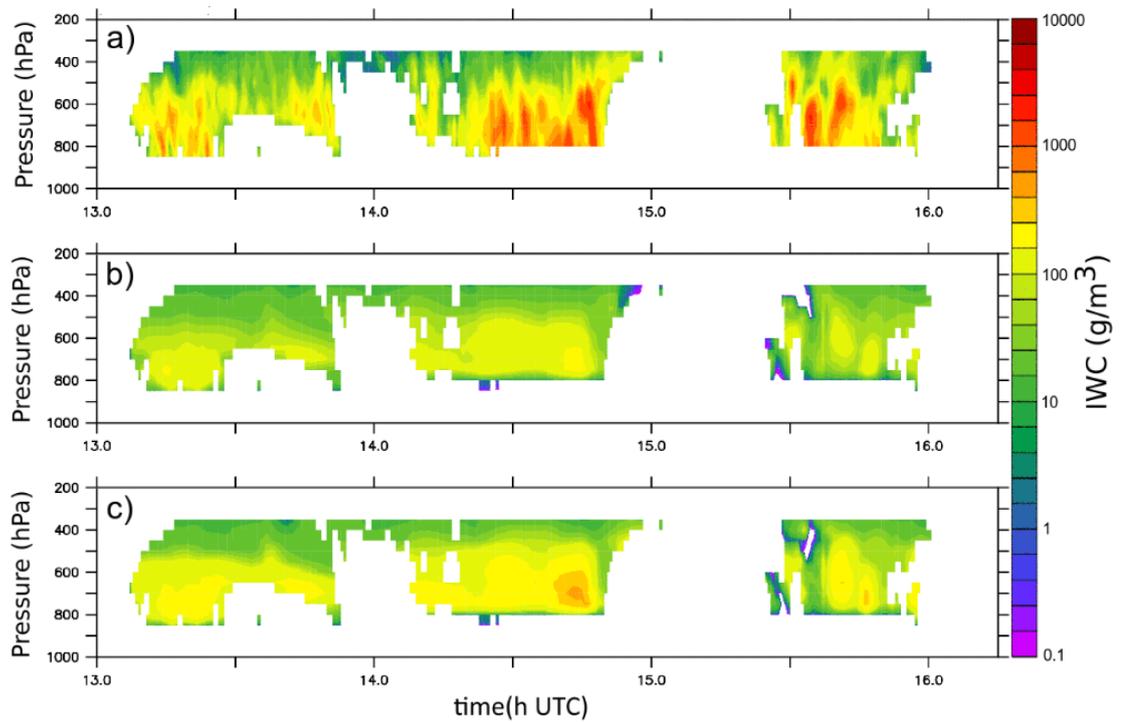


FIGURE 1.7 – Contenu en glace le long du vol F7 : a) mesuré par la plateforme RALI, b) prévu par ARPEGE avec le schéma de convection profonde PCMT et une vitesse de chute de la neige à 1.5m/s dans le schéma de microphysique de Lopez (2002), c) prévu par ARPEGE avec le schéma de convection profonde PCMT et une vitesse de chute de la neige à 0.6m/s dans le schéma de microphysique de Lopez (2002), (Avec l'aimable autorisation de Gwendal Rivière)

## Bibliographie

- C. Abbe. The physical basis of long-range weather forecasts. *Monthly Weather Review*, 29(12) :551 – 561, 1901. doi : 10.1175/1520-0493(1901)29[551c:TPBOLW]2.0.CO;2.
- AstroFiles.net. Urbain le verrier (1811-1877). <https://www.astrofiles.net/astronomie-urbain-le-verrier>. [Accessed : 16/06/2021].
- G. Balsamo. Parameterization of continental surfaces in coupled earth system modelling. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019a.
- G. Balsamo. Land surface : Part II - introduction to warm processes(vegetation and carbon). In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019b.
- G. Balsamo. Land surface : Part III - introduction to cold processes. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019c.
- P. Bauer, A. Thorpe, and G. Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525 :47–55, 2015.
- P. Bechtold. Atmospheric moist convection, 2009. URL [https://www.ecmwf.int/sites/default/files/elibrary/2017/Atmospheric\\_moist\\_convection.pdf](https://www.ecmwf.int/sites/default/files/elibrary/2017/Atmospheric_moist_convection.pdf).
- P. Bechtold. Numerical weather prediction parametrization of diabatic - convection I : General circulation and concepts. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019a.
- P. Bechtold. Numerical weather prediction parametrization of diabatic - convection II : The mass flux approach and the ifs scheme. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019b.
- P. Bechtold. Numerical weather prediction parametrization of diabatic - convection III : Forecasting and diagnostics. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019c.
- T. Benacchio and N. Wood. Semi-implicit semi-lagrangian modelling of the atmosphere : A met office perspective. *Communications in Applied and Industrial Mathematics*, 7 :4–25, 2016. doi : 10.1515/caim-2016-0020.
- L. Bengtsson and L. Moen. An operational system for numerical weather prediction. In : *Satellite and computer applications to synoptic meteorology : Lectures presented during the scientific discussions at the fifth session of the Commission for Synoptic Meteorology*, 283 :63–88, 1971.
- L. Berg and R. Stull. A simple parameterization coupling the convective daytime boundary layer and fair-weather cumuli. *Journal of the Atmospheric Sciences*, 62 :1976–1988, 2005.

- A. K. Betts and M. J. Miller. A new convective adjustment scheme. part II : Single column tests using gate wave, bomex, atex and arctic air-mass data sets. *Quarterly Journal of the Royal Meteorological Society*, 112(473) :693–709, 1986. doi : 10.1002/qj.49711247308.
- V. Bjercknes. The problem of weather prediction, considered from the viewpoints of mechanics and physics. *Meteorologische Zeitschrift*, 18(6) :663–667, 2009. doi : 10.1127/0941-2948/2009/416.
- P. Bougeault. A simple parameterization of the large-scale effects of cumulus convection. *Monthly Weather Review*, 113 :2105–2121, 1985.
- C. Bretherton, J. McCa, and H. Grenier. A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. part i : Description and 1d results. *Monthly Weather Review*, 132, 2004. doi : 10.1175/1520-0493(2004)132<0864:ANPFSC>2.0.CO;2.
- N. Bréda. Leaf area index. In S. E. Jørgensen and B. D. Fath, editors, *Encyclopedia of Ecology*, pages 2148–2154. Academic Press, Oxford, 2008. doi : 10.1016/B978-008045405-4.00849-1.
- P. Clark, N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez. Convection-permitting models : a step-change in rainfall forecasting. *Meteorological Applications*, 23(2) :165–181, 2016. doi : 10.1002/met.1538.
- J. Delanoë et al. Combined cloudsat-calipso-modis retrievals of the properties of ice clouds. *Journal of Geophysical Research : Atmosphere*, 115(D4), 2010. doi : 10.1029/2009JD012346.
- S. Dirren, M. Didone, and H. Davies. Diagnoses of "forecast-analysis" differences of a weather prediction system. *Geophysical Research Letters*, 30 (20), 2003. doi : 10.1029/2003GL017986.
- R. B. Edgerton. *Death or glory : the legacy of the Crimean War / Robert B. Edgerton*. Westview Press, Boulder, CO, 1999. ISBN 0813335701.
- E. Eliassen, B. Machenhauer, and E. Rasmussen. On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. *Institut for Teoretisk Meteorologi, University of Copenhagen*, 2, 1970.
- H. Ertel. Ein neuer hydrodynamischer wirbelsatz. *Meteorologische Zeitschrift (Braunschweig)*, 59 :277–281, 1942. doi : 10.1127/0941-2948/2004/0013-0527.
- K. Friedrichs and H. Lewy. Über die eindeutigkeit und das abhängigkeitsgebiet der lösungen beim anfangswertproblem linearer hyperbolischer differentialgleichungen. *Mathematische Annalen*, 98 :192–204, 1928.
- K. Hamilton. *Numerical Resolution and Modeling of the Global Atmospheric Circulation : A Review of Our Current Understanding and Outstanding Issues*, pages 7–27. Springer New York, New York, NY, 2008. ISBN 978-0-387-49791-4. doi : 10.1007/978-0-387-49791-4\_1.

- K. Harper, L. W. Uccellini, E. Kalnay, K. Carey, and L. Morone. 50th anniversary of operational numerical weather prediction. *Bulletin of the American Meteorological Society*, 88(5) :639 – 650, 2007. doi : 10.1175/BAMS-88-5-639.
- A. Heymsfield, M. Krämer, N. Wood, A. Gettelman, P. Field, and G. Liu. Dependence of the ice water content and snowfall rate on temperature, globally : Comparison of in-situ observations, satellite active remote sensing retrievals and global climate model simulations. *Journal of Applied Meteorology and Climatology*, 56, 2016. doi : 10.1175/JAMC-D-16-0230.1.
- R. Hogan. Radiative transfer in numerical models of the atmosphere. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019.
- S.-Y. Hong and J. Lim. The wrf single-moment 6-class microphysics scheme (wsm6). *Journal of the Korean Meteorological Society*, 42(2) :129–151, 2006.
- M. Hortal. Aspects of the numerics of the ECMWF model. In *Seminar on Recent Developments in Numerical Methods for Atmospheric Modelling, 7-11 September 1998*, pages 127–143, Shinfield Park, Reading, 1998. ECMWF, ECMWF. URL <https://www.ecmwf.int/node/10015>.
- F. Hourdin. Les paramétrisations physiques. In C. Edition, editor, *Le climat à découvert*, chapter 7, pages 162–165. 2011.
- A. Joly and L. Descamps. De l’incertitude de la prévision ... à la prévision de l’incertitude. 2009. D’après l’exposé au colloque : Comment manier l’incertitude des prévisions météorologiques et des projections climatiques ? Organisé le 3 avril 2008 à Paris par le Conseil Supérieur de la Météorologie.
- H. Joos and H. Wernli. Influence of microphysical processes on the potential vorticity development in a warm conveyor belt : a case-study with the limited-area model cosmo. *Quarterly Journal of the Royal Meteorological Society*, 138 (663) :407–418, 2012. doi : 10.1002/qj.934.
- E. Kessler. *On the Distribution and Continuity of Water Substance in Atmospheric Circulations*, pages 1–84. American Meteorological Society, Boston, MA, 1969. doi : 10.1007/978-1-935704-36-2\_1.
- A. P. Khain, K. D. Beheng, A. Heymsfield, A. Korolev, S. O. Krichak, Z. Levin, M. Pinsky, V. Phillips, T. Prabhakaran, A. Teller, S. C. van den Heever, and J.-I. Yano. Representation of microphysical processes in cloud-resolving models : Spectral (bin) microphysics versus bulk parameterization. *Reviews of Geophysics*, 53(2) :247–322, 2015. doi : 10.1002/2014RG000468.
- H. Kuo. Further studies of the parameterization of the influence of cumulus convection on large-scale flow. *Journal of the Atmospheric Sciences*, 31 :1232–1240, 1974. doi : 10.1175/1520-0469(1974)031<1232:FSOTPO>2.0.CO;2.
- E. Lagrange. Les cartes synoptiques et la prévision du temps. *Ciel et Terre*, 46 : 171, 1930.
- H. Landsberg. Storm of Balaklava and the daily weather forecast. *The Scientific Monthly*, 79(6) :347–352, 1954.

- U. Le Verrier. Rapport sur l'organisation de l'observatoire impérial de paris et projet d'organisation. 1855.
- L. Leslie and G. Dietachmeyer. Real-time limited area numerical weather prediction in australia : a historical perspective. *Australian Meteorological Magazine*, 41 :61–77, 1992.
- C. Letellier. Lorenz et la théorie du chaos. 2020a. URL <https://webtv.univ-rouen.fr/permalink/v125f76041e25ajixk9e/>.
- C. Letellier. Systèmes dynamiques chaotiques : des concepts aux applications. In *Grands Séminaires de l'Observatoire Midi-Pyrénées*, 2020b. URL <https://www.youtube.com/watch?v=pth2gX-yZTY>.
- E. Liais. *L'espace céleste et la nature tropicale : description physique de l'univers, d'après des observations personnelles faites dans les deux hémisphères*, volume (XIII-606 p.). Garnier frères (Paris), 1865. Contributeurs : Babinet, Jacques (1794-1872). Préfacier et Illustrateur : Dargent, Yan (1824-1899).
- F. Locher. *Les météores de la modernité : la dépression, le télégraphe et la prévision savante du temps (1850-1914)*, volume 4. Belin, Revue d'histoire moderne et contemporaine, 2009. doi : 10.3917/rhmc.564.0077.
- P. Lopez. Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Quarterly Journal of the Royal Meteorological Society*, 128(579) :229–257, 2002. doi : 10.1256/00359000260498879.
- E. N. Lorenz. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2) :130–148, 1963. doi : 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.
- P. Lynch and G. Moor. Les origines de la prévision numérique du temps et de la modélisation climatique. *La Météorologie*, 8(63) :14–24, 2008. doi : 10.4267/2042/21887.
- S. Manabe, J. Smagorinsky, and R. Strickler. Simulated climatology of general circulation with a hydrologic cycle. *Monthly Weather Review*, 93 :769–798, 1965.
- F. Mesinger. *Limited Area Modeling : Beginnings, state of the art, outlook*, pages 91–118. 2001. ISBN 3-928903-22-5.
- J. Milbrandt and M. Yau. A multimoment bulk microphysics parameterization. part II : A proposed three-moment closure and scheme description. *Journal of the Atmospheric Sciences*, 62 :3065–3081, 2005. doi : 10.1175/JAS3535.1.
- D. L. Mitchell, P. Rasch, D. Ivanova, G. McFarquhar, and T. Nousiainen. Impact of small ice crystal assumptions on ice sedimentation rates in cirrus clouds and gcm simulations. *Geophysical Research Letters*, 35(9), 2008. doi : 10.1029/2008GL033552.

- G. E. Moore. Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Society Newsletter*, 11(3) :33–35, 2006. doi : 10.1109/N-SSC.2006.4785860.
- Y. Okamura. Computational design of a limited-area prediction model. *Journal of the Meteorological Society of Japan*, 53 :175–188, 1975.
- OMM. Techniques d'interprétation des produits de prévision numérique du temps pour la météorologie aéronautique : établies par le groupe de travail de l'utilisation de techniques modernes en météorologie aéronautique relevant de la commission de météorologie aéronautique.
- R. Onishi and K. Takahashi. A warm-bin-cold-bulk hybrid cloud microphysical model. *Journal of the Atmospheric Sciences*, 69 :1474–1497, 2012. doi : 10.1175/JAS-D-11-0166.1.
- S. Orszag. Transform method for the calculation of vector-coupled sums : Application to the spectral form of the vorticity equation. *Journal of the Atmospheric Sciences*.
- A. Persson. Early operational numerical weather prediction outside the USA : an historical introduction. part 1 : Internationalism and engineering nwp in sweden, 1952–69. *Meteorological Applications*, 12(2) :135–159, 2005. doi : 10.1017/S1350482705001593.
- N. Phillips. The general circulation of the atmosphere : A numerical experiment. *Quarterly Journal of the Royal Meteorological Society*, 82(352) :123, 1956.
- J.-M. Piriou. Convection : causalité, sensibilité, transitions, interaction d'échelles, limites des modèles. In *Cours de formation permanente*, 2020.
- 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 :4127–4139, 2007. doi : 10.1175/2007JAS2144.1.
- C. Planche. *Développement et évaluation d'un modèle tridimensionnel de nuage mixte à microphysique détaillée : Application aux précipitations orographiques*. Theses, Université Blaise Pascal - Clermont-Ferrand II, 2011.
- L. Richardson. *Weather Prediction by Numerical Process*. Cambridge University Press : Cambridge, MA, 1922. Reprinted 2006 by Cambridge University Press with a new introduction by Peter Lynch.
- G. Rivière, M. Wimmer, P. Arbogast, J.-M. Piriou, J. Delanoë, C. Labadie, Q. Cazenave, and J. Pelon. The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during nawdex iop6. *Weather and Climate Dynamics*, 2(4) :1011–1031, 2021. doi : 10.5194/wcd-2-1011-2021. URL <https://wcd.copernicus.org/articles/2/1011/2021/>.

- W. Rlarriott. The earliest telegraphic daily meteorological reports and weather maps. *Quarterly Journal of the Royal Meteorological Society*, 29(126) :123–131, 1903. doi : 10.1002/qj.49702912609.
- I. Sandu. Turbulence parametrization (with a focus on the boundary layer). In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019a.
- I. Sandu. Parameterization of surface fluxes : Outline. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019b.
- I. Sandu. Parametrization of turbulent fluxes in the outer layer. In *ECMWF Training Course : Parametrization of subgrid physical processes*, 2019c.
- A. Schäfler, G. Craig, H. Wernli, P. Arbogast, J. D. Doyle, R. McTaggart-Cowan, J. Methven, G. Rivière, F. Ament, M. Boettcher, M. Bramberger, Q. Cazenave, R. Cotton, S. Crewell, J. Delanoë, A. Dörnbrack, A. Ehrlich, F. Ewald, A. Fix, C. M. Grams, S. L. Gray, H. Grob, S. Groß, M. Hagen, B. Harvey, L. Hirsch, M. Jacob, T. Kölling, H. Konow, C. Lemmerz, O. Lux, L. Magnusson, B. Mayer, M. Mech, R. Moore, J. Pelon, J. Quinting, S. Rahm, M. Rapp, M. Rautenhaus, O. Reitebuch, C. A. Reynolds, H. Sodemann, T. Spengler, G. Vaughan, M. Wendisch, M. Wirth, B. Witschas, K. Wolf, and T. Zinner. The North Atlantic Waveguide and Downstream Impact experiment. *Bulletin of the American Meteorological Society*, 99 :1607–1637, 2018.
- Y. Seity. Prévision numérique à mésoéchelle. In *Cours de formation initiale IENM2*, 2020.
- P. Siebesma and J. Teixeira. An advection-diffusion scheme for the convective boundary layer, description and 1d results. In *Proceedings of 14th symposium on boundary layers and turbulence*. Aspen, USA, 2000.
- D. Stensrud. *Parameterization schemes : Keys to understanding numerical weather prediction models*, volume 9780521865401. Cambridge University Press, United Kingdom, 2011. ISBN 0521865409. doi : 10.1017/CBO9780511812590. Publisher Copyright : © D. Stensrud 2007. Copyright : Copyright 2017 Elsevier B.V., All rights reserved.
- B. Stevens, M. Satoh, L. Auger, J. Biercamp, C. Bretherton, X. Chen, P. Düben, F. Judt, M. Khairoutdinov, D. Klocke, C. Kodama, L. Kornblueh, S.-J. Lin, P. Neumann, W. Putman, N. Röber, R. Shibuya, B. Vannière, P. Vidale, and L. Zhou. Dyamond : the dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, 6 : 61, 2019. doi : 10.1186/s40645-019-0304-z.
- M. Stoelinga. A potential vorticity-based study of the role of diabatic heating and friction in a numerically simulated baroclinic cyclone. *Monthly Weather Review*, 124, 1996.
- R. B. Stull. *An Introduction to Boundary Layer Meteorology*. Springer, Dordrecht, 1988. doi : 10.1007/978-94-009-3027-8.

- M. Tiedtke. A comprehensive mass flux scheme for cumulus parametrization in large-scale models. *Monthly Weather Review*, 117 :1779–1800, 1989.
- P. A. Ullrich, C. Jablonowski, J. Kent, P. H. Lauritzen, R. Nair, K. A. Reed, C. M. Zarzycki, D. M. Hall, D. Dazlich, R. Heikes, C. Konor, D. Randall, T. Dubos, Y. Meurdesoif, X. Chen, L. Harris, C. Kühnlein, V. Lee, A. Qaddouri, C. Girard, M. Giorgetta, D. Reinert, J. Klemp, S.-H. Park, W. Skamarock, H. Miura, T. Ohno, R. Yoshida, R. Walko, A. Reinecke, and K. Viner. DCMIP2016 : a review of non-hydrostatic dynamical core design and intercomparison of participating models. *Geoscientific Model Development*, 10(12) :4477–4509, 2017. doi : 10.5194/gmd-10-4477-2017.
- D. L. Williamson. The evolution of dynamical cores for global atmospheric models. *Journal of the Meteorological Society of Japan*, 85B :241–269, 2007. doi : 10.2151/jmsj.85B.241.
- E. P. Willis and W. H. Hooke. Cleveland Abbe and american meteorology, 1871–1901. *Bulletin of the American Meteorological Society*, 87(3) :315 – 326, 2006. doi : 10.1175/BAMS-87-3-315.
- M. Wimmer, G. Rivière, P. Arbogast, J.-M. Piriou, J. Delanoë, C. Labadie, Q. Caizenave, and J. Pelon. Diabatic processes modulating the vertical structure of the jet stream above the cold front of an extratropical cyclone : sensitivity to deep convection schemes. *Weather and Climate Dynamics Discussions*, 2021 : 1–30, 2021. doi : 10.5194/wcd-2021-76. URL <https://wcd.copernicus.org/preprints/wcd-2021-76/>.

## Chapitre 2

# Représenter la distribution de probabilité de l'état de l'atmosphère via la Prédiction d'Ensemble

Afin de représenter la distribution de probabilité de l'état futur de l'atmosphère et son évolution temporelle (Figure 2.1), une prédiction d'ensemble est composée de plusieurs prévisions appelées membres dont chacune est un échantillon de la distribution de probabilité de l'état de l'atmosphère. Leur évolution temporelle permet de prévoir l'évolution de cette distribution, dont l'estimation est d'autant meilleure que le nombre d'échantillons, donc de membres, est élevé. Les différents membres de la prédiction d'ensemble doivent alors représenter l'ensemble des sources d'incertitudes inhérentes à la prédiction.

Une première source d'incertitude se trouve dans l'assimilation de données aboutissant aux conditions initiales, appelées analyses. Celles-ci sont calculées à partir d'une ébauche, c'est-à-dire une prédiction récente issue d'une précédente analyse, qui est corrigée par les observations pendant le cycle d'assimilation. Différentes sources d'incertitudes interviennent alors. Tout d'abord, des erreurs sont présentes dans les observations même, notamment les erreurs instrumentales et de représentativité, mais aussi celles dues à la répartition spatio-temporelle hétérogène des observations. L'ébauche étant une prédiction récente, elle est également entachée d'erreurs comme toute prédiction. Les erreurs d'observation et d'ébauche, ainsi que les approximations et simplifications employées dans la mise au point des algorithmes d'assimilation, contribuent ainsi aux erreurs affectant les conditions initiales.

Une autre source d'incertitudes non négligeable et très documentée réside au sein même des modèles météorologiques. Des erreurs dues aux méthodes de calcul ou aux hypothèses et simplifications utilisées, entre autres dans les paramétrisations physiques, rendent les prévisions incertaines.

La discrétisation de l'atmosphère avec le choix de sa résolution ainsi que la représentation des limites de l'atmosphère sont aussi source d'incertitude. Que ce soient les incertitudes dues aux conditions de bord au niveau de la surface ou au

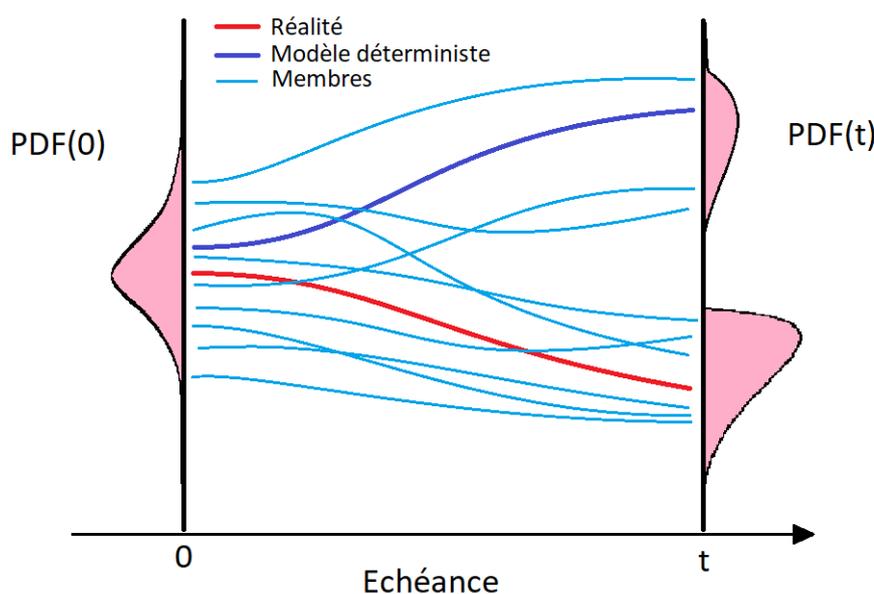


FIGURE 2.1 – Schéma d'échantillonnage de la distribution de probabilité de l'état futur de l'atmosphère par les différents membres de la prévision d'ensemble. Une prévision d'ensemble fournissant plusieurs scénarios possibles, l'un d'eux peut être plus proche de la réalité qu'un modèle déterministe.

sommet de l'atmosphère, celles-ci doivent apparaître dans la prévision d'ensemble. A ce jour, seules les erreurs de la modélisation de la surface ont été étudiées. Dans le cas des modèles à aire limitée, différentes techniques ont été mises en place pour rendre compte des incertitudes pesant sur les méthodes de couplage avec les modèles globaux.

Ainsi quatre principales sources d'incertitudes sont connues, elles peuvent-être liées :

- aux conditions initiales,
- aux conditions de surface,
- aux conditions latérales,
- au modèle-même.

Ce chapitre s'attache alors à décrire les méthodes utilisées à ce jour pour représenter ces sources d'incertitudes dans les prévisions d'ensemble.

## 2.1 Perturbation des conditions initiales

Dans les modèles météorologiques déterministes, les conditions initiales sont fournies par des analyses valides à l'instant d'initialisation du modèle,  $t_0$ . Résultats de la technique d'assimilation de données, celles-ci sont obtenues à partir de la correction par des observations, de prévisions, appelées ébauches.

Cependant, en raison de la sensibilité de l'atmosphère aux conditions initiales, il est nécessaire de prendre en compte les incertitudes pesant sur les analyses

servant de conditions initiales aux prévisions déterministes et ensemblistes. Ainsi, différentes méthodes ensemblistes échantillonnant la distribution des états initiaux de l'atmosphère compte tenu des erreurs d'observations et d'ébauches ont été mises en place. Il s'agit ici de présenter, dans les grandes lignes, les méthodes les plus courantes.

### 2.1.1 Perturbation aléatoire

La méthode dite de « perturbation aléatoire » des conditions initiales repose sur l'ajout d'une perturbation à une analyse de contrôle valide à  $t_0$  (souvent la dernière analyse disponible du modèle déterministe). Cette perturbation se base sur des statistiques d'erreur passée d'analyse calculées sur une longue période en utilisant par exemple, les écarts entre analyses de différents centres météorologiques (Du et al., 2018). Les différences entre ces analyses sont particulièrement prononcées dans les régions avec peu d'observations ou encore dans les régions avec une forte variabilité météorologique. Cependant, cette méthode n'utilise pas une erreur représentative de l'état d'écoulement atmosphérique du jour et peut souffrir d'un manque de cohérence spatiale due à l'utilisation d'analyses issues de modèles différents. Une autre version consiste à perturber aléatoirement les éléments de la matrice de covariance des erreurs d'ébauche dite matrice B.

### 2.1.2 Perturbation des modes instables de la dynamique

Deux grandes approches, longtemps étudiées et encore utilisées en opérationnel pour certaines, se basent sur la croissance des modes instables de la dynamique. L'une se base sur la croissance future des perturbations alors que la seconde s'appuie sur les perturbations passées.

#### Vecteurs singuliers

La croissance des erreurs de prévision dues à la perturbation des conditions initiales suivant des directions privilégiées dans l'espace des phases, appliquer des perturbations selon ces directions dans les analyses va donc permettre de focaliser la représentation des erreurs dues aux conditions initiales dans les directions qui sont sources d'incertitudes.

En suivant l'hypothèse que la croissance d'une perturbation initiale représentative de l'erreur d'analyse suit une évolution linéaire en fonction du temps sur une période approximative de deux jours (Lacarra and Talagrand, 2010), un modèle météorologique peut donc être approximé par son tangent linéaire  $M$  (c'est-à-dire le modèle linéarisé autour de la trajectoire prise par le modèle lors de l'assimilation).  $M$  représente alors cette évolution tangente linéaire sur ces deux jours. Par définition, les éléments propres de  $M^*M$  (où  $M^*$  est la matrice adjointe de  $M$ ), sont les valeurs et vecteurs singuliers, soit les directions représentant les erreurs d'analyses.

Plus concrètement, la production des perturbations à partir des vecteurs singuliers est basée sur des combinaisons linéaires de ces vecteurs, ceci afin de prendre en compte l'ensemble des vecteurs singuliers et leurs caractéristiques spatiales et énergétiques (Buizza and Palmer, 1995). Ces perturbations vont le plus souvent

se localiser sur les zones d'instabilités baroclines sous forme de paquets d'ondes orientées vers l'Ouest. Au cours de leur évolution, ces perturbations souvent situées au départ en milieu de troposphère vont se déplacer vers les basses couches et au niveau de la tropopause.

Différents paramètres sont à prendre en compte pour le calcul de ces vecteurs. En particulier, la période dite d'optimisation (i.e. pendant laquelle des perturbations croissent) ne doit pas dépasser la période de validité de l'hypothèse de linéarité. Par ailleurs, au cours du calcul des vecteurs singuliers, le choix d'une norme est nécessaire. La norme couramment choisie est la norme énergie totale. Il est à noter qu'une généralisation de la méthode des vecteurs singuliers pour un modèle non-linéaire a été mise au point et est nommée Conditional Nonlinear Optimal Perturbation (Mu et al., 2003).

## Breeding

Une autre approche possible pour perturber l'analyse est d'utiliser comme perturbation, la différence entre deux prévisions valides à  $t_0$  mais initialisées par des analyses différentes (Figure 2.2). A l'origine, cette première perturbation est aléatoire, mais à chaque cycle d'assimilation, la perturbation est définie comme la différence entre la prévision non perturbée (issue d'une analyse de contrôle) et celle perturbée issue du cycle précédent (Toth and Kalnay, 1993; Callado et al., 2013).

Les perturbations sont remises à l'échelle et peuvent être soit ajoutées, soit soustraites à la dernière analyse :

$$x_k^a = x^a \pm \alpha(x_{H(k),perturb} - x_{H(k)}) \quad (2.1)$$

où  $k$  est le numéro du membre,  $x^a$  est la dernière analyse déterministe à  $t_0$ ,  $\alpha$  est un facteur d'amplitude et  $x_{H(k)}$  est une précédente prévision initialisée au temps  $H$  et valide à  $t_0$ .

Ces perturbations, nommées « bred vectors » vont croître au fur et à mesure du cycle d'assimilation et représenter l'incertitude qui croît pendant ce cycle.

Notons que même si les vecteurs singuliers et les « bred vectors » ne s'intéressent qu'à la dynamique (du futur et du passé), ils ne tiennent pas compte de l'impact des observations sur la réduction de l'incertitude.

### 2.1.3 Assimilation d'Ensemble

L'assimilation de données repose sur une succession temporelle (appelée cyclage) d'étapes d'analyse et de prévision : dans le cas d'AROME, une nouvelle analyse est calculée à chaque heure de la journée, puis une prévision d'une heure est lancée à partir de cette analyse afin de fournir une ébauche à l'analyse suivante, et ainsi de suite. Ce cyclage conduit à une évolution temporelle spécifique des erreurs associées : de nouvelles erreurs d'observation et de modèle sont introduites à chaque étape d'analyse et de prévision, et elles se propagent au cours des étapes d'analyse/prévision, tout en étant progressivement atténuées.

Afin de simuler ce processus, les méthodes d'Assimilation d'Ensemble reposent sur la mise en oeuvre d'un ensemble de plusieurs analyses et prévisions, qui sont

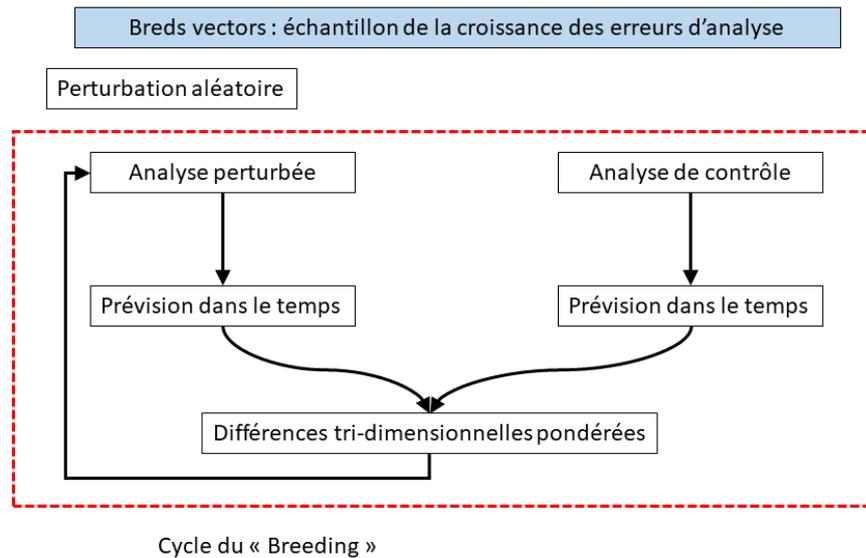


FIGURE 2.2 – Schéma de la méthode de « breeding » (traduit et adapté de Callado et al. (2013))

perturbées et cyclées temporellement. Les perturbations appliquées sont construites de façon à être représentatives des erreurs en jeu, d'une part au niveau des observations (à l'aide de tirages aléatoires des covariances d'erreur d'observation) et d'autre part au niveau du modèle (à l'aide de méthodes d'inflation multiplicative par exemple, ou d'autres méthodes décrites en section 2.4).

Les analyses sont souvent calculées de façon semblable au système déterministe, à l'aide de méthodes variationnelles notamment. Elles peuvent également être calculées sous la forme d'un filtre de Kalman d'ensemble (Evensen, 1994). Une variante de cette approche (appelée Ensemble Transform Kalman Filter (Bishop et al., 2001)) consiste à ne pas perturber les observations, mais à modifier le calcul de l'analyse, de façon à ce que la dispersion des analyses soit cohérente avec l'expression théorique des covariances d'erreur d'analyse.

#### 2.1.4 Descente d'échelle des modèles ensemblistes globaux

Dans le cas des modèles ensemblistes à aire limitée, une possibilité de génération de conditions initiales est de réutiliser les analyses ou prévisions à très courte échéance générées par un modèle ensembliste global. Ainsi, l'analyse servant de conditions initiales du  $i^{\text{ème}}$  membre de la prévision d'ensemble régionale consiste en l'analyse du modèle déterministe régional  $x_{regional}^a$  à laquelle on ajoute ou retranche une perturbation. Cette dernière est définie comme la différence entre l'analyse du  $j^{\text{ème}}$  membre du modèle ensembliste global ( $x_{j,PEglobale}^a$ ) et la moyenne des analyses des différents membres du modèle ensembliste global ( $\frac{1}{N} \sum_j x_{j,PEglobale}^a$ ) (Raynaud and Bouttier, 2016; Tennant, 2015), tous deux interpolés sur la grille du modèle régional. Peralta et al. (2012) propose d'utiliser à la place de la moyenne des analyses des différents membres du modèle ensembliste global, l'analyse du modèle déterministe global. Cette perturbation peut aussi être modulée selon la verticale par un facteur  $\alpha$  définissant un profil de modulation verticale interdisant une perturbation en basse troposphère afin de garder une cohérence au niveau des analyses

de surface. La formulation des conditions initiales est donc la suivante, pour le cas utilisant la moyenne des analyses du modèle ensembliste global :

$$x_{i,PEreg}^a = x_{regional}^a + \alpha \left( x_{j,PEglobale}^a - \frac{1}{N} \sum_j^N x_{j,PEglobale}^a \right) \quad (2.2)$$

où  $i = 1, 2, \dots, M$  le  $i^{\text{ème}}$  membre de la prévision d'ensemble régionale comprenant  $M$  membres et  $j = 1, 2, \dots, N$  le  $j^{\text{ème}}$  membre de la prévision d'ensemble globale comprenant  $N$  membres.

## 2.2 Perturbation des conditions de surface

Malgré la représentation des incertitudes dans les conditions initiales et de modélisation, les ensembles restent sous dispersifs dû, d'après Sutton et al. (2006) à un manque de perturbation de la surface. Plusieurs solutions sont alors possibles : perturbation de paramètres des schémas de surface, des analyses de surface et de la topographie.

### 2.2.1 Perturbation des schémas de surface

Pour représenter l'incertitude des schémas de surface, la méthode la plus utilisée est la perturbation de paramètres. Ceux-ci sont modifiés de manière différente pour chaque membre de la prévision d'ensemble mais restent constants tout au long de l'intégration du modèle. Par exemple, Lavaysse et al. (2013) modifient les valeurs de paramètres tels que la fraction de végétation et de neige dans chaque maille du modèle, le LAI (Leaf Area Index), l'albédo, la longueur de rugosité ou encore l'épaisseur de la neige. Gehne et al. (2019) proposent d'y adjoindre la conductivité hydraulique du sol alors que Bouttier et al. (2016) perturbe l'indice de végétation, le coefficient de chaleur de la végétation ou les flux à la surface océanique.

Ces paramètres peuvent être constants sur l'ensemble de la surface, mais être aussi variables dans l'espace. Dans ce cas, un générateur de perturbation est utilisé, comme dans Bouttier et al. (2016), produisant ainsi des champs de perturbations gaussiennes d'amplitudes et de structures spatiales fixées avec des motifs de 400km de large. L'amplitude de ces champs de perturbations est optimisée pour chaque paramètre. Selon le type de paramètres, ces perturbations peuvent être multiplicatives ou additives.

Les tendances d'humidité et de température en sortie du schéma de surface peuvent aussi être perturbées (Gehne et al., 2019) tout comme les champs pronostiques d'humidité, température du sol ou température de surface de la mer (SST) (Tennant and Beare, 2014; Deng et al., 2016). Ces champs étant pronostiques, ils ne sont perturbés qu'en début de prévisions. Le masque terre-mer est aussi perturbé dans Lavaysse et al. (2013).

Enfin, Gao et al. (2021) a testé aussi différentes valeurs d'un paramètre de couplage surface-atmosphère.

Toujours pour représenter l'incertitude des schémas de surface, une autre solution réside dans l'emploi de différents schémas de surface. Gao et al. (2021) a alors utilisé quatre paramétrisations différentes de surface (Noah, Rapid Update Cycle, Community Land Model et la version multi paramétrisation du modèle Noah).

### 2.2.2 Perturbation des analyses de surface

Concernant la perturbation des analyses de surface, Sutton et al. (2006) proposent d'utiliser deux analyses différentes de l'humidité de surface. Dans le même ordre d'idée, Gao et al. (2021) utilisent pour WRF, quatre analyses de l'humidité surfacique issues de quatre centres différents : analyse du modèle GFS (CNEP), ré-analyse ERA5 (CEPMMT), analyse du système chinois d'assimilation des données surfaciques du CMA (Land Data Assimilation system) et du système global d'assimilation des données surfaciques (Global Land Data assimilation) de la NASA. Une alternative proposée par Deng et al. (2016) est d'ajouter, à une analyse de surface, une perturbation proportionnelle à la différence entre deux analyses de surface déterministes. Wang et al. (2010) perturbent en revanche, les conditions initiales de surface à partir d'une descente d'échelle d'un modèle global ensembliste.

### 2.2.3 Perturbation de la topographie

Différentes études ont également montré l'impact de la perturbation de la topographie dans les modèles sur les prévisions, en particulier, sur la localisation spatiale des précipitations orographiques. Par exemple, un aplanissement du relief corse peut annihiler un épisode de précipitations intense sur cette île, alors qu'une meilleure résolution des Alpes atteignant 3000 m d'altitude améliore la prévision d'un tel phénomène (Scheffknecht et al., 2016).

Concernant les prévisions d'ensembles, Li et al. (2017) ont montré une amélioration de celles-ci en perturbant la topographie. En particulier, ils ont utilisé trois interpolations et deux lissages différents pour représenter l'incertitude sur le relief dans le modèle WRF. Les motifs de perturbations obtenus sont similaires à ceux obtenus avec la perturbation des conditions initiales et de la modélisation. Cependant, un ensemble ne peut se limiter à la perturbation du terrain car cette modification est beaucoup plus limitée que celle des conditions initiales et de modélisation.

## 2.3 Perturbation des conditions de bords pour les modèles à aire limitée

Concernant les modèles à aire limitée, une autre source d'incertitude est à prendre en compte. Puisque ces modèles nécessitent des conditions de bords provenant de modèles coupleurs (soit un modèle avec une aire limitée plus étendue, soit un modèle global), une incertitude pèse quant aux choix de ce dernier.

De nombreuses études ont montré l'impact des modèles coupleurs sur les prévisions de modèles régionaux. A titre d'exemple, dans leur étude préliminaire de

l'évènement de précipitation intense ayant eu lieu sur la Corse le 14 septembre 2006, Lambert and Argence (2008) montrent qu'à grande échelle, le choix du coupleur est peu significatif. En revanche, à haute résolution, les deux modèles coupleurs utilisés donnent des résultats différents. Couplé au modèle IFS, le modèle de recherche Méso-NH localise mieux les précipitations sur la Corse. En revanche, couplées à ARPEGE, la simulation représente mieux l'intensité du phénomène. Ce cas n'est pas isolé : les prévisionnistes rencontrent souvent ce genre de configurations rendant difficile le choix de coupleur entre ARPEGE et IFS, par exemple. Afin de rendre compte de l'incertitude sur les conditions de bord dans les modèles ensemblistes à aire limitée, différentes techniques ont été mises en place.

### 2.3.1 A partir de prévisions d'ensemble globales

De manière analogue aux modèles régionaux déterministes qui sont une descente d'échelle de leur équivalent global, les prévisions d'ensemble régionales sont en général couplées aux membres des modèles ensemblistes globaux. Cependant, les prévisions d'ensemble globales présentent dans leur grande majorité, plus de membres que leur équivalent régional. Il faut donc sélectionner  $N_R$  membres parmi les  $N_G$  membres de la prévision d'ensemble globale pour alimenter les conditions de bord des  $N_R$  membres de la prévision d'ensemble régionale.

Une première méthode consiste à choisir, tout simplement les  $N_R$  premiers membres de la prévision d'ensemble globale. Il est aussi possible de tirer aléatoirement  $N_R$  membres parmi les  $N_G$  membres de l'ensemble global. Cependant, des méthodes de clustering optimisant la variabilité des membres globaux choisis sont bien plus performantes. Plus particulièrement, ces méthodes utilisent une métrique synthétisant la différence de plusieurs champs météorologiques entre deux membres de la prévision d'ensemble globale, regroupent ensuite les  $N_G$  membres globaux en  $N_R$  groupes dits « cluster » et cherchent enfin le membre représentatif de chacun des  $N_R$  clusters.

La méthode la plus utilisée est la méthode de Molteni et al. (2001). La métrique proposée et utilisée pour mesurer la distance entre les membres de la prévision d'ensemble globale est la somme sur chaque point de grille, de l'erreur quadratique moyenne du géopotential à 700hPa et de précipitation à 120h d'échéance entre deux membres globaux. Cependant, l'étude de Molteni et al. (2001) se focalisant sur les Alpes, d'autres champs sont proposés pour des cas plus généraux tels que le vent horizontal, le géopotential et l'humidité relative à 500, 700 et 850hPa et à 30h d'échéance (Bouttier and Raynaud, 2018). D'autres variantes utilisent la pression au niveau de la mer, le vent et la température à 850 et 950hPa et à 24 et 36h comme pour HarmonEPS (Frogner et al., 2019) ou encore, dans le cas d'épisode Méditerranéen, uniquement le géopotential à 500hPa et les flux d'humidité à 925hPa (Nuissier et al., 2012). Notons que plusieurs échéances peuvent être utilisées pour calculer la métrique, en particulier pour les versions opérationnelles de HarmonEPS ou encore PEARO.

Même des membres issus de précédents réseaux peuvent être utilisés. Dans son étude sur le couplage des membres de la prévision d'ensemble AROME à celle d'ARPEGE, Nuissier et al. (2012) ont aussi testé la méthode de Molteni et al. (2001) utilisant à la fois des membres de la PEARP la plus récente mais aussi des membres du réseau précédent de la PEARP. Cette méthode peut amener à un

ensemble très varié lorsqu'une situation météorologique est peu prévisible.

Pour regrouper les membres globaux en  $N_R$  clusters, la méthode de clustering Complete Link (CLINK) calcule, à partir des métriques précédemment définies, la dissimilarité entre deux clusters comme la distance maximale entre les membres des clusters. Le membre représentatif du cluster est alors celui qui minimise l'indice de représentativité défini comme le ratio entre l'erreur quadratique moyenne de la distance moyenne entre membres d'un même cluster et l'erreur quadratique moyenne de la distance moyenne avec les membres des autres clusters.

Cependant, cette méthode a l'inconvénient de trop favoriser les membres extrêmes de la prévision d'ensemble globale. Une première solution consiste à considérer le membre le plus proche du centre du cluster (méthode « Molteni central » ; Bouttier and Raynaud, 2018). Une autre solution, mise en place par Bouttier and Raynaud (2018), est d'utiliser la méthode de Ward pénalisée dans laquelle la dissimilarité entre deux clusters est définie comme la variance de l'union de ces deux clusters. Cette méthode présente l'avantage de mieux répartir les  $N_G$  membres des modèles ensemblistes globaux dans les  $N_R$  clusters et de générer des clusters plus peuplés avec des tailles plus homogènes. Plus précisément, ceci évite d'avoir des clusters comportant un seul et unique membre alors que d'autres contiennent la quasi totalité des  $N_G$  membres de l'ensemble global.

### 2.3.2 A partir de prévisions globales déterministes

Dans le cas où aucun modèle global ensembliste n'est disponible, une solution réside dans l'utilisation de modèles globaux déterministes. Les systèmes de prévision d'ensemble régionaux établis par Gebhardt et al. (2011) et García-Moya et al. (2011) couplent chaque membre de la prévision d'ensemble régionale à différents modèles globaux déterministes. Par exemple, dans la prévision d'ensemble régionale SREPS utilisé au centre météorologique espagnol (García-Moya et al., 2011), les 25 membres sont couplés à IFS (CEPMMT), GFS (NCEP), ARPEGE (Météo-France), GSM (JMA) et GDPS (GEM, CMC). Cette méthode multimodèles, très coûteuse, a l'avantage d'utiliser des modèles coupleurs ayant une meilleure résolution : de par leur coût de calcul, les modèles ensemblistes utilisent souvent des résolutions souvent moins fines que leur équivalents déterministes.

Une autre alternative est proposée pour le modèle HarmonEPS (Frogner et al., 2019; Callado et al., 2018; García-Moya et al.; Clancy et al., 2018). Cette fois-ci la méthode Scaled Lagged Average Forecast (SLAF) n'utilise qu'un seul modèle déterministe coupleur, mais elle se base sur l'initialisation de ce modèle à différents instants dits « réseaux ». Cette méthode est uniquement possible grâce au fait que ces modèles globaux ont une longueur d'intégration plus élevée que les modèles régionaux et atteignent ainsi de plus longues échéances. Dans le cas du modèle HarmonEPS, le modèle déterministe global coupleur est IFS. Cette fois-ci, l'incertitude est représentée par une « erreur du jour », différence de prévisions du modèle coupleur à un instant donné, initialisées à des instants séparés de  $H$  heures. Chaque membre utilisant une erreur du jour différente, un facteur est appliqué pour que l'amplitude de cette erreur soit identique entre chaque membre. Ce coefficient peut être proportionnel à la norme de l'énergie totale ou être fixé après optimisation. Comme celui-ci peut être appliqué positivement ou négativement, une paire de membre est associée à chaque erreur du jour. Ainsi, pour

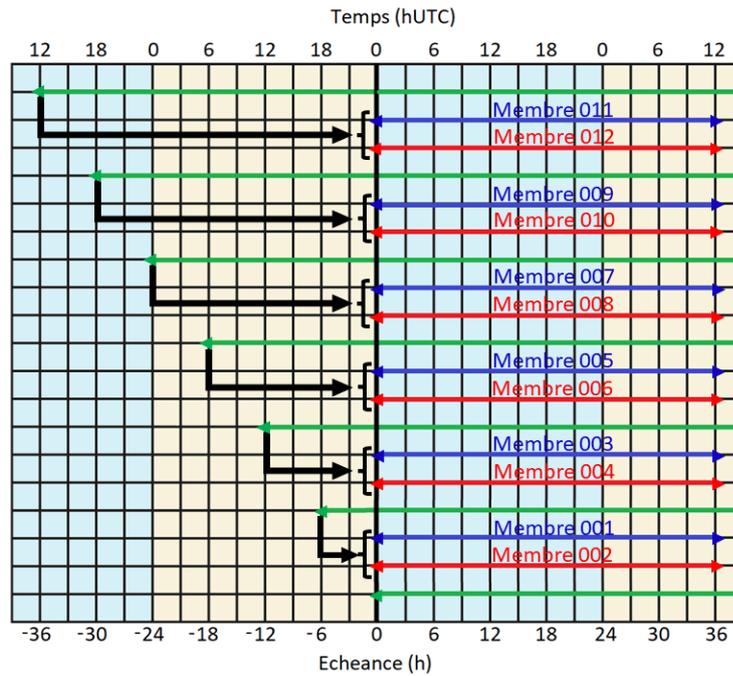


FIGURE 2.3 – Schéma de la méthode SLAF. Les prévisions déterministes globales sont représentées en vert et les membres de la prévision d'ensemble à aire limitée en rouge et bleu.

un membre  $m$ , les conditions de bord utilisées pour une prévision commençant à l'instant  $t = 0$ , suivront la formulation suivante :

$$BC_m(0) = IFS(0) \pm K_m(IFS(h) - IFS(h - H)) \quad (2.3)$$

où  $h$  est l'heure d'initialisation du modèle IFS. La figure 2.3 schématise cette méthode pour une prévision d'ensemble régionale de 36h initialisée à 0hUTC et couplée à un modèle valide au même instant (0hUTC). En particulier, ce schéma représente en fonction du temps (ou échéance), la disponibilité des données du modèle couplé et la durée d'intégration du modèle régional ensembliste. Les prévisions déterministes du modèle couplé (soit IFS pour HarmonEPS) sont représentées en vert et couvrent 90h d'échéance. Les différents membres de 36h de l'ensemble régional sont représentés en rouge (application du facteur d'amplitude positivement) et bleu (application du facteur d'amplitude négativement). Dans cet exemple, le décalage entre 2 réseaux est de 6h. Ainsi, pour les membres 1 et 2, l'erreur du jour est définie comme la différence entre la prévision déterministe du réseau le plus récent (0hUTC) et celui de 18hUTC la veille (6h avant le réseau de 0hUTC). Pour les membres 3 et 4, cette différence est calculée entre les réseaux déterministes de 12hUTC et 18hUTC la veille, etc.

Tout comme la méthode multimodèles, le SLAF présente l'avantage d'utiliser un couplage avec des modèles déterministes à haute résolution. Il a aussi l'avantage d'être moins coûteux puisque les prévisions déterministes des précédents réseaux sont déjà disponibles. Cependant, en raison de la longueur limitée des prévisions du modèle couplé, le nombre de membres est très vite limité. Sur la figure 2.3, pour des prévisions d'ensemble régionale de 36h, on ne peut donc avoir plus de 12

membres.

Comparées à des méthodes aléatoires utilisant une « erreur du jour » issue de prévisions réalisées pour une date antérieure choisie aléatoirement (Magnusson et al., 2008), les méthodes SLAF et multimodèles sont mieux représentatives de l'incertitude de couplage et offrent un plus fort étalement de la probabilité des états de l'atmosphère (soit une plus grande diversité des membres de la prévision d'ensemble). Par exemple, dans son étude, Frogner et al. (2019) ont montré le réel gain obtenu avec la méthode SLAF en la comparant à celle de perturbations aléatoires de (Magnusson et al., 2008). Cette dernière utilise une erreur du jour non pas calculée à partir de réseaux valides à l'instant de la prévision, mais à partir de prévisions valides à des dates situées dans une période de plus ou moins 20 jours, mais d'une année différente. Cependant, en raison de son nombre limité de membres, Frogner et al. (2019) recommandent plutôt d'utiliser les modèles ensemblistes globaux, surtout si une méthode de clustering est appliquée. La méthode SLAF peut être utilisée, même si elle est un peu moins performante qu'une descente d'échelle d'un modèle ensembliste global, parce qu'elle dégrade moins les ensembles qu'une méthode de tirage aléatoire des membres globaux. Il faut noter par ailleurs, que toutes ces méthodes ont très peu d'impact en début de prévision, comparé à la représentation des erreurs de modélisation mais surtout des erreurs initiales, qui ont le plus fort impact jusqu'à environ 12h. Elles sont, en revanche, plus effectives aux longues échéances.

## 2.4 Erreurs de modélisation dues aux paramétrisations

Concernant les erreurs de modélisation, une des principales sources d'erreur dans les modèles météorologiques se situe dans la modélisation des processus diabatiques par les paramétrisations. En effet, les paramétrisations représentant, par définition, les effets moyens de la physique sur une colonne atmosphérique, celles-ci ne peuvent reproduire parfaitement la réalité dans ses moindres détails. En outre, par manque de connaissance, certains phénomènes, comme certains processus microphysiques, sont absents ou mal représentés dans les schémas de physique. Enfin, pour des raisons de coûts de calcul, certains processus tels que le rayonnement doivent être simplifiés malgré le fait, qu'en théorie, il est tout à fait possible de modéliser ces phénomènes avec précision. Ainsi, différentes études montrent qu'une source non négligeable d'erreur dans les prévisions provient des paramétrisations-mêmes (Bauer et al., 2015) et doit être représentée dans les prévisions d'ensemble. Depuis la fin des années 1990, date à laquelle l'exploration de l'erreur de modèle a débuté, différentes méthodes ont été mises en place pour échantillonner cette incertitude.

### 2.4.1 Approche multi-modèles

Une des premières techniques pour représenter l'erreur de modélisation dans les prévisions d'ensemble a été d'utiliser l'approche multi-modèles (Gebhardt et al., 2011; Iversen et al., 2011). En effet, chaque modèle étant différent, avec une dis-

crétisation de l'atmosphère, un coeur dynamique et des paramétrisations physiques qui lui sont propres, considérer différents modèles permet à la fois d'échantillonner l'ensemble des erreurs de modélisation, à la fois dues à la physique mais aussi à la dynamique. Par exemple, le système de prévision d'ensemble régional à courte échéance de l'agence météorologique espagnole, SREPS produit un ensemble de prévisions régionales de 72h à partir de cinq modèles déterministes régionaux : HARMONIE-AROME (du consortium HIRLAM), ALARO (du consortium ALARO), WRF-ARW (du centre météorologique NCAR) et NMMB (du centre météorologique NCEP). Ceux-ci sont couplés à différents modèles déterministes globaux, à savoir GEM (du centre météorologique CMC), IFS (CEPMMT), GSM (du centre météorologique JMA), GFS (du centre météorologique NCEP) et ARPEGE (Météo-France) (Gonzalez et al., 2020).

Cette méthode a l'avantage d'offrir une large diversité de membres de la prévision d'ensemble et prend en compte les différents biais de chaque modèle. Cependant, la qualité des modèles n'étant pas la même (par ex : résolution différente), les membres de la prévision d'ensemble ne sont pas équiprobables. Enfin, cette méthode demandant un coût de maintenance et de développement considérable, reste peu utilisée.

### 2.4.2 Approche multi-physiques

De manière analogue à l'approche multi-modèles, les ensembles dits multi-physiques utilisent plusieurs représentations d'un même processus physique. Ainsi, chaque membre d'un tel ensemble empruntera un jeu de paramétrisations physiques différent.

L'un des premiers ensembles utilisant cette technique a été élaboré au centre météorologique canadien et utilisait différentes paramétrisations de la convection profonde, rayonnement, traînée orographique et surface (Houtekamer et al., 1996). Encore aujourd'hui, la prévision d'ensemble ARPEGE (PEARP) opérationnelle à Météo-France repose sur des paramétrisations différentes pour la convection profonde, convection peu profonde, turbulence, rayonnement, ondes orographiques et dans une version précédente, aussi la paramétrisation de la surface océanique (Descamps et al., 2015).

Tout comme l'approche multi-modèles, cette méthode produit des membres diversifiés mais peu équiprobables du fait de la qualité différente des membres selon les paramétrisations. Cette diversité des membres semble principalement liée aux schémas de convection profonde. Ponzano et al. (2020) ont montré une forte ressemblance entre les membres utilisant le même schéma de convection profonde pour les épisodes Méditerranéens alors que Rivière et al. (2021) ont montré, dans un contexte de dépression extra-tropicale, une ressemblance selon la fermeture de ce schéma.

### 2.4.3 SKEB

En parallèle des approches multi-modèles et multi-physiques, des approches stochastiques ont vu le jour dans les années 1990.

La cascade d'énergie turbulente, transférant l'énergie des grands tourbillons atmosphériques aux plus petits qui la dissipent, est mal prise en compte dans les

modèles. En effet, la rétroaction énergétique des phénomènes de petites échelles, c'est-à-dire des phénomènes paramétrisés, sur les phénomènes de grande échelle, c'est-à-dire à l'échelle résolue du modèle, n'est pas modélisée. La méthode Stochastic Kinetic Energy Backscatter (SKEB), mise en place par Shutts (2004, 2005), cherche à simuler cette source d'erreur, considérée comme un forçage, en introduisant une perturbation aléatoire bidimensionnelle de la fonction de courant, celle-ci étant représentative de la rotation des vents horizontaux donc des tourbillons de la cascade d'énergie turbulente. Dans certains cas, les tendances de température potentielle sont aussi perturbées (Berner et al., 2011). Ces perturbations sont pondérées par le taux de dissipation instantanée et totale lié à la dissipation numérique et énergétique due à la convection profonde et aux ondes orographiques (Shutts, 2004; Berner et al., 2009). Une simplification de ce schéma proposée pour WRF considère une dissipation constante spatialement et temporellement de l'énergie (Berner et al., 2011). Celle-ci est fixée et est non dépendante de l'état de l'atmosphère. Cette méthode propose une forte diversité des membres en particulier dans l'atmosphère libre et la couche limite atmosphérique, mais celle-ci reste limitée dans les tropiques.

#### 2.4.4 SPPT

Une autre technique de perturbation stochastique des paramétrisations a vu le jour au CEPMMT. Initialement nommée schéma stochastique des tendances diabatiques (Stochastic Diabatic Tendency (SDT); Buizza et al., 1999), la méthode Stochastically Perturbed Parameterization Tendencies (SPPT; Palmer et al., 2009) suppose que l'erreur des paramétrisations est proportionnelle aux tendances produites par ces mêmes paramétrisations. Ainsi, l'idée est ici de perturber les tendances totales diabatiques des variables d'état (vent, température et humidité spécifique) issues des paramétrisations. A chaque pas de temps du modèle, la tendance d'une variable  $x$  est calculée de la manière suivante :

$$\frac{dx}{dt} = A + D \quad (2.4)$$

où  $A$  est la tendance adiabatique calculée par la dynamique et  $D$  la somme des tendances diabatiques issues de chaque paramétrisation. La méthode SPPT ajoute une perturbation multiplicative à  $D$  telle que :

$$D(x, y, z) \rightarrow D(x, y, z) \times (1 + \mu(z) \times r(x, y)) \quad (2.5)$$

où  $\mu(z)$  est un profil vertical de poids allant de 0 à la surface, à 1 au milieu de l'atmosphère pour diminuer jusqu'à 0 au sommet de l'atmosphère, et  $r(x, y)$  est la perturbation bi-dimensionnelle qui suit une loi normale, d'écart-type  $\sigma$  et de corrélation horizontale homogène et isotrope. Cette dernière provient d'un générateur de motifs stochastiques gaussiens. Ces motifs sont corrélés spatialement selon une longueur caractéristique définie, et évoluent dans l'espace spectral avec un temps caractéristique fixé. Ces temps et longueurs caractéristiques peuvent être optimisés pour chaque application du SPPT et diffèrent pour chacune des tendances des variables d'état (température, vent, humidité, ...; Buizza et al., 1999). Cependant,

généralement, la même perturbation est appliquée à toutes les variables.

La méthode SPPT apporte une diversité des membres plus limitée que la méthode SKEB en particulier dans l'atmosphère libre. Tout comme le SKEB, l'étalement de la distribution de probabilité échantillonnée par le SPPT n'est pas suffisamment large dans les tropiques. En revanche, la combinaison de ces deux méthodes permet d'améliorer la diversité des membres.

Cependant, puisque la perturbation des tendances est proportionnelle à la valeur-même de la tendance diabatique, si deux paramétrisations produisent des tendances dont la somme est nulle, celles-ci ne seront donc pas perturbées avec la méthode SPPT. C'est pourquoi, Christensen et al. (2017) proposent une modification du schéma nommé Independent SPPT (iSPPT) où une perturbation de même type que celle du SPPT, est appliquée indépendamment à chaque tendance issue de chaque paramétrisation et non à la somme des tendances contrairement à la méthode SPPT originale. En appliquant la perturbation avant le calcul de la somme des tendances, ceci permet d'éviter les phénomènes compensatoires. Notons que contrairement à la première version du SPPT de Buizza et al. (1999), les motifs de perturbations sont maintenant identiques entre les variables d'état (température, vent et humidité).

Dans le SPPT, tout comme dans le iSPPT, les perturbations sont appliquées à la fin de chaque pas de temps du modèle, une fois les tendances de chaque paramétrisation calculées. Or, au sein-même d'un pas de temps, les paramétrisations sont appelées dans un ordre bien précis. Par exemple, dans le modèle AROME, le rayonnement est d'abord calculé et fournit ses résultats aux schémas de convection peu profonde qui lui-même fournit des données d'entrées au schéma de turbulence dont les sorties sont reprises dans la microphysique.

Ainsi, Wastl et al. (2019a) proposent cette fois-ci un Physical parametrization-based SPPT (pSPPT), perturbant les tendances aux cours du pas de temps. Là où le iSPPT perturbe les tendances du rayonnement à la fin du pas de temps et ne les injecte pas dans le schéma de convection peu profonde, Wastl et al. (2019a) perturbent la tendance de rayonnement avant que celle-ci ne soit injectée dans le schéma de convection peu profonde, et ainsi de suite pour les différentes paramétrisations jusqu'à la fin du pas de temps. Une autre version du pSPPT, le Independent physical parametrization-based SPPT (ipSPPT), utilise par contre des motifs différents de perturbations entre les différentes tendances physiques (température, vent et humidité; Wastl et al., 2019a), contrairement au pSPPT.

## 2.4.5 Perturbation de paramètres

Une approche différente de la méthode SPPT (s'intéressant aux sorties des paramétrisations) perturbe, cette fois-ci, les paramètres en entrée de ces paramétrisations. Celle-ci se base sur le fait que les paramétrisations physiques considèrent un grand nombre de paramètres dont les valeurs sont incertaines. Cette incertitude peut venir du fait que la valeur de certains d'entre eux a été déterminée empiriquement, grâce aux données relevées pendant des campagnes de mesures. L'incertitude repose donc indirectement sur celle des données observées. D'autres paramètres intervenant dans des équations de fermeture ou des équations trop simplifiées ont été déterminés à force d'intercomparisons modèle-observation.

Pour représenter l'incertitude provenant de ces paramètres, différentes mé-

thodes de perturbations ont été mises en place. Celles-ci sont résumés dans le tableau 2.1. Contrairement aux méthodes multi-physiques, ces méthodes nécessitent peu de maintien des modèles. Néanmoins, l'une des difficultés majeures rencontrées réside dans le choix, par les experts physiciens, des paramètres et de leur intervalles de perturbations n'engendrant pas de simulations atmosphériques irréalistes.

Méthode	Membres	Date initiale	Echéance	Espace
PP	✓			
RPP	✓	✓		
RP	✓	✓	✓	
SPP	✓	✓	✓	✓

TABLE 2.1 – *Perturbation des paramètres dépendant des membres de la Prévision d'Ensemble, de la date initiale, des échéances et de l'espace (points de grille) pour les quatre méthodes de perturbation de paramètres*

**PP :** La méthode appelée ici Perturbed Parameter (PP) utilise des valeurs différentes de paramètres incertains des paramétrisations, pour chaque membre (Murphy et al., 2004; Gebhardt et al., 2011; Stainforth et al., 2005; Reynolds et al., 2011). Ces valeurs restent constantes à chaque utilisation du modèle (réseau, date initiale), au cours de l'intégration du modèle (échéance) et à chaque point de grille (espace).

Cette méthode a été appliquée au modèle COSMO-DE-EPS par Gebhardt et al. (2011) sous le nom d'approche « Multi-paramètres ». Les 20 membres de cet ensemble diffèrent par la valeur de 5 paramètres issus de la convection peu profonde, de la microphysique et de la turbulence. Ceux-ci utilisent aléatoirement soit la valeur par défaut de la version déterministe COSMO-DE, soit une (pour 3 paramètres) ou deux (pour 2 paramètres) valeurs alternatives proposées par les experts des paramétrisations physiques du modèle COSMO.

Dans le même ordre d'idée, Murphy et al. (2004) proposent de perturber 29 paramètres issus des schémas de microphysique, convection, rayonnement, diffusion, traînée orographique, surface et turbulence d'un modèle climatique pour générer un ensemble de 53 membres. Cette méthode qu'ils ont nommé « Perturbed Physics Ensemble », modifie les paramètres les uns après les autres, selon 3 valeurs différentes proposées par les experts physiciens, à savoir : minimum, maximum et une valeur intermédiaire n'étant pas forcément au centre de l'intervalle de perturbation défini par les minima et maxima. Par ailleurs, notons que les valeurs par défaut du modèle peuvent être les minima, maxima ou valeurs intermédiaires.

Au lieu d'utiliser des valeurs bien définies pour chaque paramètre, il est aussi possible de tirer aléatoirement les valeurs des paramètres dans un intervalle et selon une distribution déterminée par les experts physiciens. En particulier, les valeurs peuvent être choisies par un tirage de Monte-Carlo ou mieux encore par un échantillonnage par hypercube latin comme proposé par Hacker et al. (2011) ou Christensen et al. (2015).

**RPP :** La méthode PP ne prend pas en compte la variabilité de l'incertitude selon la situation météorologique. Pour palier cet inconvénient, une version que l'on

nommera ici "Random Perturbed Parameter (RPP)", modifie donc la valeur des paramètres de chaque membre à chaque prévision initialisée à une date différente (dite réseau). Il est à noter que ces deux méthodes se rapprochent de la méthode multi-physiques puisqu'il s'agit d'utiliser des paramétrisations différentes de par le réglage des valeurs des paramètres (Berner et al., 2017).

**RP :** De manière plus stochastique, le modèle ensembliste MOGREPS du MetOffice utilise une technique nommé Random Parameter (RP) perturbant les paramètres issus des paramétrisations pendant l'intégration du modèle. Ceux-ci changent de valeurs toutes les 30 ou 60 minutes selon les paramètres (Bowler et al., 2008; Baker et al., 2014). Cette perturbation suit l'équation suivante :

$$P(t) = \bar{P} + r \times (P(t-1) - \bar{P}) + \epsilon \quad (2.6)$$

où  $P(t)$  est la valeur du paramètre au temps  $t$ ,  $P(t-1)$  est la valeur précédente du paramètre,  $\bar{P}$  est la valeur moyenne (souvent valeur par défaut) du paramètre,  $r = 0.95$  l'auto-corrélation du paramètre et  $\epsilon$  un terme de choc stochastique, tiré aléatoirement, selon une distribution uniforme dans l'intervalle  $[-\frac{(P_{max}-P_{min})}{3}, +\frac{(P_{max}-P_{min})}{3}]$  où  $P_{min}$  et  $P_{max}$  sont les valeurs maximales et minimales de  $P$  définies par les experts physiiciens. Des tests sont réalisés afin de vérifier que le paramètre garde une valeur dans l'intervalle  $[P_{min}; P_{max}]$ . Une optimisation de  $r$  et  $\epsilon$  a été effectuée sur des prévisions de 72h, en maximisant l'étalement de la distributions des membres. Afin d'éviter une optimisation coûteuse de ces deux quantités, McCabe et al. (2016) proposent une dépendance de  $r$  et  $\epsilon$  dépendant du temps caractéristique fixé pour la perturbation des paramètres.

**SPP :** Enfin, il existe des méthodes de perturbation stochastique nommées Stochastic Parameter Perturbation (SPP) (e.g. Ollinaho et al., 2017; Leutbecher et al., 2017; Christensen et al., 2015; Ollinaho et al., 2013; Jankov et al., 2019, ...) qui appliquent une perturbation spatio-temporelle des paramètres de type :

$$P = \exp(r)\bar{P}, r \sim \mathcal{N}(\mu, \sigma) \quad (2.7)$$

où  $P$  est la valeur du paramètre,  $\bar{P}$  la valeur par défaut de  $P$  dans le modèle déterministe,  $r$  une perturbation suivant une distribution, généralement gaussienne, de moyenne  $\mu$  (soit nulle, soit choisie pour générer une médiane correspondant à la valeur par défaut) et d'écart-type  $\sigma$  déterminés, pour tout  $P$ , par les experts ou suite à plusieurs expériences de type test-correction. Cette perturbation  $\exp(r)$  varie dans le temps et l'espace suivant un motif, indépendant pour chaque paramètre et chaque membre, et identique à chaque niveau de l'atmosphère. Comme pour la perturbation temporelle du RP (McCabe et al., 2016), une auto-régression de premier ordre permet d'obtenir des corrélations spatio-temporelles de type gaussienne dans l'espace spectral. Entre deux pas de temps séparés de  $\Delta t$ , nous avons :

$$\hat{r}(t + \Delta t) = \phi \hat{r}(t) + s\epsilon(t); \quad \phi = \exp(-\frac{\Delta t}{\tau}); \quad s^2 = A \exp(-l^2 R_{\oplus}^{-2} n(n+1)) \quad (2.8)$$

où  $\hat{r}(t)$  sont les coefficients spectraux de  $r$  à l'instant  $t$ ,  $\phi$  la corrélation de  $\hat{r}$  pendant l'intervalle de temps  $\Delta t$ ,  $\tau$  l'échelle de temps de corrélation ;  $s$  l'écart-type dépendant de la longueur d'onde totale  $n$ ,  $A$  un facteur permettant de régler la variance de la perturbation,  $l$  l'échelle de longueur de corrélation,  $R_{\oplus}$  le rayon de la Terre et  $\epsilon$  est un nombre aléatoire suivant une loi normale centrée réduite. Une comparaison entre les méthodes PP et SPP appliquées à la paramétrisation de la convection montre que les schémas de perturbation de paramètres (PP) semblent améliorer les performances des prévisions d'ensemble (Christensen et al., 2015). Cependant, a contrario de la méthode SPP, la méthode PP induit des biais non désirés d'humidité dans les tropiques pour les différents membres. Enfin, des méthodes hybrides combinant SKEB, SPPT ou encore SPP ont été testées (Leutbecher et al., 2017; Jankov et al., 2017, 2019). Par exemple, le schéma « Hybrid stochastically perturbed parametrization » propose des schémas pSPPT pour le rayonnement, la convection peu profonde et la microphysique mais un schéma SPP pour la turbulence (Wastl et al., 2019b). La raison de l'utilisation d'un schéma SPP uniquement pour la turbulence est simplement due à l'instabilité numérique générée par la méthode pSPPT sur les tendances de la turbulence.

#### 2.4.6 Paramétrisations physiques stochastiques

Depuis quelques années, des recherches sont menées pour mettre au point des paramétrisations physiques stochastiques. Celles-ci, en plus de fournir les tendances moyennes habituelles de vent, température et humidité sur une maille du modèle, calculent aussi la variance de ces tendances au sein de la maille à partir des variables internes de la paramétrisation.

Par exemple, dans un schéma de convection peu profond, Sakradzija and Klocke (2018) considèrent que le flux de masse moyen dans la maille est constitué de plusieurs flux de masses internes à la maille dont le nombre et la taille sont variables. Ainsi, un tirage aléatoire du nombre de nuages convectifs et, pour chacun de ces nuages, une distribution de leur flux de masse permettent d'obtenir une version stochastique des schémas de convection peu profonde. D'autres paramètres tels que la fraction de flux de masse humide dans la maille sont perturbés (Sakradzija et al., 2015).

Dans le cas de la turbulence, Kober and Craig (2016) proposent d'ajouter à la tendance d'une variable  $x$  calculée par le schéma de turbulence, une perturbation proportionnelle à la variance de  $x$  correspondant à l'incertitude des flux de chaleur à la surface. Cette dernière est calculée en implémentant les équations définissant la variance de  $x$  élaborées par Mellor and Yamada (1982).

### 2.5 Perturbation de la dynamique

En complément de la perturbation de la physique, des méthodes s'intéressant à l'incertitude du coeur dynamique voient le jour. En parallèle de l'approche multi-physiques, la prévision d'ensemble globale proposée par Houtekamer et al. (1996) et mise en place au centre météorologique canadien, perturbe un paramètre de la dynamique : le temps caractéristique de la diffusion horizontale (e-folding time). Plus récemment, un schéma SPPT a été appliqué aux tendances issues de

la diffusion par Koo and Hong (2014) et Qiao et al. (2017). Celui-ci applique un nombre aléatoire aux tendances dynamiques, dépendant du temps et de l'espace. Un travail a été réalisé afin de maintenir la stabilité du modèle. Enfin, concernant l'incertitude des conditions initiales, une version hybride du SPPT, perturbant à la fois les tendances physiques et dynamiques, a été testé dans un filtre de Kalman, améliorant ainsi grandement les covariances d'erreur d'ébauche (Lim et al., 2020). Actuellement, des études sur la perturbation des points d'ensemencement dans le schéma semi-lagrangien sont en cours d'étude au CEPMMT (Leutbecher et al., 2017).

Différentes méthodes de production d'ensemble ont été présentées ici. Notons que d'autres techniques ont été proposées, telles qu'un ensemble utilisant des prévisions déterministes décalées dans le temps ou issues de différents centres météorologiques (ensemble du pauvre).

## 2.6 Évaluation des systèmes de prévision d'ensemble par des scores probabilistes

Au delà de simplement représenter l'ensemble des erreurs et sources d'incertitudes des modèles météorologiques, une prévision d'ensemble se doit de satisfaire certains critères.

Une prévision d'ensemble doit être fiable et donc être en accord avec ce qui est observé dans la réalité. Tout comme la confiance s'inscrit dans la durée, la fiabilité d'une prévision d'ensemble se juge, non pas sur un unique évènement, mais sur une statistique d'évènements. Il s'agit donc de vérifier pour un grand nombre de prévisions, que la probabilité prévue d'occurrence d'un phénomène est bien conforme aux fréquences d'occurrence observées. Par exemple, si sur une année, la prévision d'ensemble propose  $n = 100$  occurrences de l'évènement  $E$  avec une probabilité de 20% (20% des membres proposent une occurrence), alors celle-ci est fiable si, dans l'année, on observe effectivement ce phénomène dans 20% des  $n = 100$  cas, soit exactement 20 fois. Le modèle le plus fiable est alors celui proposant des probabilités correspondantes à celles observées en moyenne dans l'atmosphère, c'est à dire, la climatologie. Cependant, même si celui-ci est cohérent avec les observations en moyenne, il reste peu utile pour les prévisionnistes. C'est pourquoi une prévision d'ensemble se doit aussi d'aider à discriminer l'occurrence ou non d'un évènement.

Le second critère pour évaluer une prévision d'ensemble est donc sa résolution. Elle doit alors être capable de distinguer les prévisions amenant à des probabilités d'occurrence différentes.

Enfin, une prévision d'ensemble se doit de représenter l'incertitude de l'atmosphère sans sous-estimation ou sur-estimation de celle-ci. Si la prévision d'ensemble propose une distribution de probabilité des états de l'atmosphère trop peu étalée, elle va proposer des prévisions très peu variées et va donc sous-estimer les incertitudes. Il y aura donc sur-confiance en le modèle. Si par malchance, le modèle n'est pas fiable, alors la prévision d'ensemble proposera à tort avec confiance des évènements non existants. Cette configuration peut donc amener à un grand nombre de fausses alarmes pour les prévisionnistes. A l'opposé, avec une prévision d'ensemble

surestimant les incertitudes et proposant une distribution de probabilité des états de l'atmosphère trop dispersée, il y aura un manque de confiance en ce modèle et une difficulté pour les prévisionnistes à identifier l'évolution réelle de l'atmosphère.

Afin d'évaluer la capacité des prévisions d'ensemble à bien représenter l'incertitude des prévisions météorologiques et donc la distribution de probabilité des états prévus de l'atmosphère, différents scores probabilistes ont été proposés dont certains sont présentés ici. Pour plus de détails, le lecteur pourra se référer à Talagrand et al. (1999), Candille and Talagrand (2005) ou encore Candille et al. (2007).

### 2.6.1 Scores déterministes

Avant toute chose, intéressons-nous aux scores déterministes mesurant la qualité des prévisions d'une variable d'état  $x$ , plus précisément leur distance à l'observation  $o$ . Il s'agit, entre autres, du biais moyen, de l'erreur moyenne absolue (MAE) et de la racine de l'erreur quadratique moyenne (RMSE) définis par les équations suivantes :

$$Biais = \frac{1}{R} \sum_{i=1}^R (x_i - o_i) \quad (2.9)$$

$$MAE = \frac{1}{R} \sum_{i=1}^R |x_i - o_i| \quad (2.10)$$

$$RMSE = \sqrt{\frac{1}{R} \sum_{i=1}^R (x_i - o_i)^2} \quad (2.11)$$

où  $R$  est le nombre d'évènements,  $x_i$  et  $o_i$  les valeurs prévues et observées de l'évènement  $i$ . De par leur définition, ces scores déterministes donnent une interprétation différente de l'erreur de prévision. Le biais est plutôt utilisé afin d'identifier des erreurs systématiques. Par ailleurs, la RMSE met davantage en valeur les erreurs importantes que la MAE.

D'un point de vue ensembliste, différentes versions de ces scores peuvent être définies. Par exemple, pour le biais moyen, il peut s'agir :

- du biais de la moyenne de l'ensemble composés de  $N$  membres soit :

$$Biais = \frac{1}{R} \sum_{i=1}^R \frac{1}{N} \left( \sum_{j=1}^N (x_{i,j} - o_i) \right) \quad (2.12)$$

- ou du biais moyen de  $N$  membres de l'ensemble soit :

$$Biais = \frac{1}{N} \sum_{j=1}^N \left( \frac{1}{R} \sum_{i=1}^R (x_{i,j} - o_i) \right) \quad (2.13)$$

## 2.6.2 Diagramme de Talagrand - Histogramme de rang

Parmi les scores probabilistes, on compte le diagramme de Talagrand également appelé histogramme de rang qui est un score évaluant la fiabilité de la distribution de probabilité de la variable  $x$  prévue par la prévision d'ensemble. Plus particulièrement, il s'agit de comparer la valeur observée de  $x$  dit  $o$  à la distribution de probabilité échantillonnée par les  $N$  membres de la prévision d'ensemble. Rangées dans l'ordre croissant, les  $N$  valeurs de  $x$ , issues des  $N$  membres de la prévision d'ensemble, définissent des quantiles de  $x$  et par la même, définissent aussi  $N + 1$  intervalles, dont  $N - 1$  valeurs intermédiaires et 2 aux extrémités. On regarde alors dans quel intervalle (entre quels quantiles) se situe la valeur de  $o$ . Le processus étant réitéré pour un grand nombre de prévisions et observations, on compte le nombre de fois où l'observation est située dans un intervalle, permettant ainsi d'obtenir un histogramme du nombre d'observations en fonction des quantiles de la prévision d'ensemble.

Si l'observation est une réalisation tout aussi indépendante que les  $N$  réalisations effectuées par la prévision d'ensemble, alors cette observation suit la même distribution de probabilité que les  $N$  prévisions et l'observation peut être vue comme une  $N + 1$  ième réalisations. Celle-ci doit donc avoir autant de chance de se retrouver dans un intervalle que dans un autre.

Ainsi, pour une prévision d'ensemble fiable, l'histogramme de rang doit être plat. Un histogramme asymétrique en forme de J suggère un biais négatif de la prévision d'ensemble, c'est-à-dire que puisque les observations sont trop souvent au delà des valeurs maximales de la prévision d'ensemble, le modèle a tendance à sous-estimer les valeurs de  $x$ . A l'inverse, un histogramme en forme de L dénote un biais positif, une surestimation de  $x$ . Un histogramme symétrique en forme de U révèle une surreprésentation des observations dans les quantiles extrêmes : les observations étant trop souvent éloignées des valeurs extrêmes de la prévision d'ensemble, celle-ci est sous-dispersive. A l'inverse, un histogramme en forme de  $\Lambda$  indique une sur-dispersion de la prévision d'ensemble.

## 2.6.3 Diagramme de fiabilité

Dans le même ordre d'idée que le diagramme de Talagrand, le diagramme de fiabilité permet également d'évaluer la fiabilité de la distribution de probabilité prévue de la variable  $x$ . Celui-ci trace la fréquence d'occurrence d'un évènement en fonction de sa probabilité prévue par la prévision d'ensemble (définie comme le ratio entre le nombre de membres prévoyant l'évènement et le nombre total de membres). Ainsi, si un ensemble est parfaitement fiable, la fréquence d'observation doit correspondre aux probabilités d'occurrence et le diagramme de fiabilité doit suivre la bissectrice.

Si la courbe est située au dessus de la bissectrice, alors la fréquence d'observation est donc trop importante comparée aux probabilités prévues, la prévision d'ensemble a tendance à produire des valeurs trop faibles de  $x$  et présente donc un biais négatif (Figure 2.4, en bas). A l'inverse, une courbe située en-dessous de la bissectrice dénotera un biais positif (Figure 2.4, à haut).

Un diagramme de fiabilité surestimant la fréquence d'observation pour les faibles probabilités et la sous-estimant pour les fortes probabilités, suggérera une

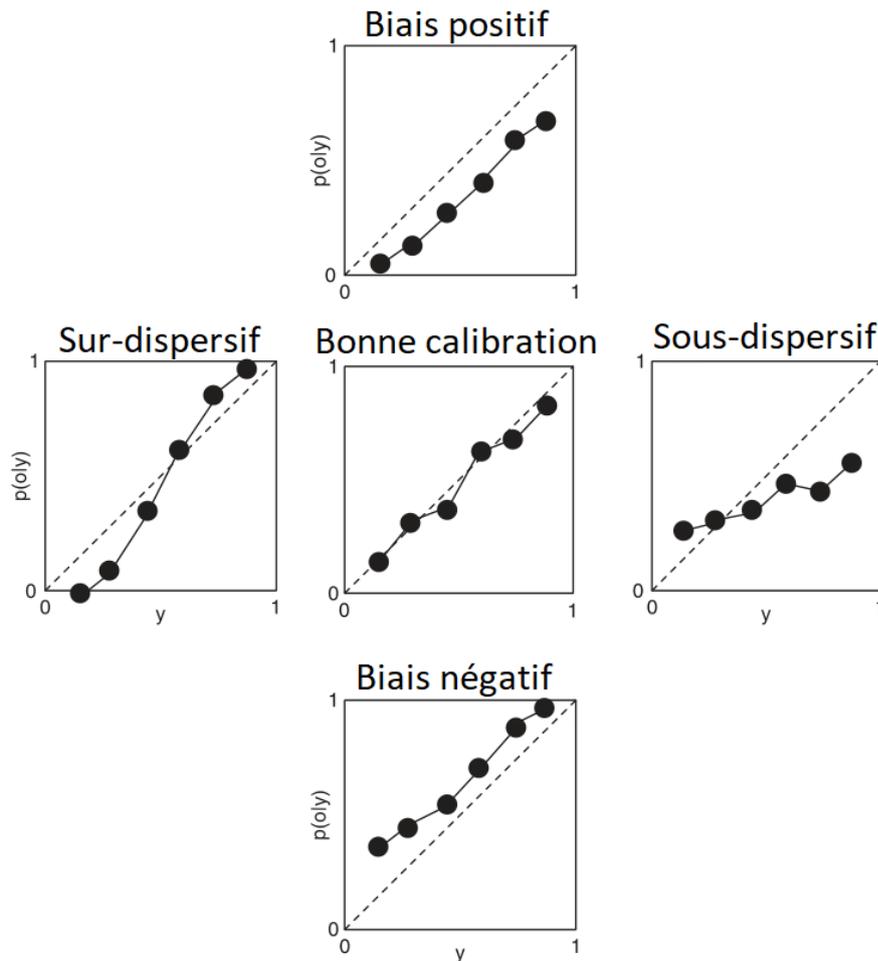


FIGURE 2.4 – Diagramme de fiabilité pour les ensembles bien calibré, sur-dispersif, sous-dispersif, avec biais positif et négatif (traduit et adapté de Wilks (2011))

surdispersion de la prévision d'ensemble (Figure 2.4, à gauche). En effet, si la fréquence d'observation est sur-estimée pour les faibles probabilités, la prévision d'ensemble propose alors de trop faibles valeurs de  $x$ . En outre, si la fréquence d'observations est sous-estimée pour les fortes probabilités, autrement dit, il y a trop d'observations pour des valeurs de  $x$  en dessous des valeurs prévues alors la prévision d'ensemble produit des valeurs de  $x$  trop importantes. Il y a bien surdispersion. A l'opposé, un diagramme de fiabilité sous-estimant les fréquences d'observations pour les faibles quantiles et les sur-estimant pour les forts quantiles, dénotera une sous-dispersion de la prévision d'ensemble (Figure 2.4, à droite).

## 2.6.4 Courbe ROC

A l'inverse des diagrammes de rang et de fiabilité, la courbe Receiver Operating Characteristic (ROC) s'intéresse à la discrimination de la prévision d'ensemble, plus particulièrement à sa capacité à choisir entre deux situations possibles. Plus concrètement, elle évalue la prévision d'occurrence d'évènements binaires de type « est-ce que la température sera supérieure à 25 °C ? ».

Ce score ROC se base sur la table de contingence, décrite dans le tableau 2.2

et regroupant les  $N$  prévisions de l'ensemble en 4 catégories :  $a$ , celles ayant prévu l'occurrence d'un évènement effectivement observé,  $b$  celles ayant prévu à tort un évènement non observé,  $c$  celles ayant manqué la prévision d'un phénomène et  $d$  celles ayant bien prévu la non-occurrence du phénomène. A partir de cette table, on définit le taux de bonnes prévisions  $H$  (hite rate en anglais) comme la proportion d'évènements bien prévus lorsque ceux-ci sont effectivement observés, et le taux de fausses alarmes  $F$  (False alarm rate en anglais) comme la proportion d'évènements prévus lorsqu'ils ne sont pas observés :

$$H = \frac{a}{a + c} \quad (2.14)$$

$$F = \frac{b}{b + d} \quad (2.15)$$

Évènement	observé	non observé
prévu	a	b
non prévu	c	d

TABLE 2.2 – *Tableau de Contingence*

Dans le cas d'une prévision probabiliste, l'occurrence d'un évènement est prévue si la probabilité affectée à celui-ci est supérieure à une probabilité dite de coupure  $p_c$ . Pour différentes probabilités de coupure, il est possible alors de calculer les taux de fausses alarmes et de bonnes prévisions.

La courbe ROC représente ainsi le taux de bonnes prévisions en fonction du taux de fausses alarmes pour les différentes probabilités de coupure (Figure 2.5) Sachant que, pour une probabilité de coupure nulle, l'évènement est systématiquement prévu et, a contrario, pour une probabilité de coupure de 1, l'évènement n'est jamais prévu, la courbe ROC passe toujours par les points  $(0, 0)$  et  $(1, 1)$ . Un système de prévision est parfait si son taux de fausses alarmes est nul et son taux de bonnes prévisions est égale à 1 ; il doit donc tendre vers le point  $(0, 1)$ . Ainsi, la courbe ROC doit se rapprocher au mieux des segments  $(0, 0) - (0, 1)$  et  $(0, 1) - (1, 1)$ . Par exemple, sur la figure 2.5, la courbe ROC notée  $A$  est plus proche de ces segments et est donc meilleure que celle notée  $B$ . A l'inverse, un système est non discriminant s'il prévoit autant de fois un phénomène qu'il ne le prévoit pas :  $a = c$  et  $b = d$ . Dans ce cas, sa courbe de ROC correspond à la bissectrice. Pire encore, un système peut être discriminant, mais à tort, en prévoyant un évènement qui n'aura pas lieu : la courbe ROC a des taux de fausses alarmes et de bonnes prévisions se rapprochant du point  $(1, 0)$ .

Une façon de synthétiser le diagramme de ROC est de s'intéresser à l'aire sous la courbe ROC comprise entre 0 et 1. Plus la courbe se rapproche du point  $(0, 1)$ , plus l'aire sous la courbe de ROC approchera la valeur de 1, plus la prévision d'ensemble est discriminante. Une aire inférieure à 0.5 dénotera un système faussement déterminant.

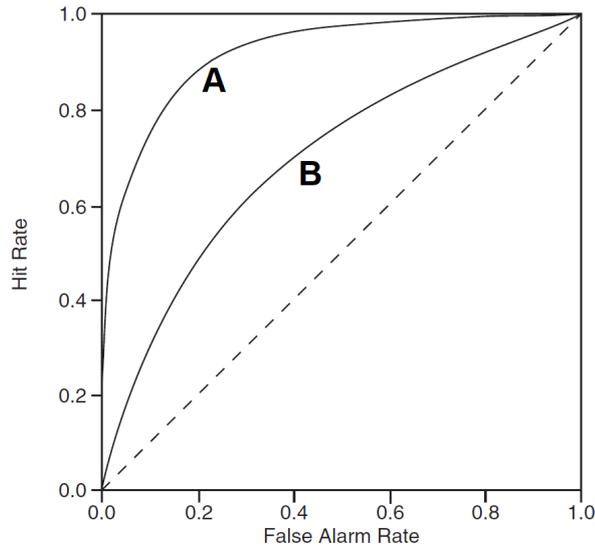


FIGURE 2.5 – Exemple de courbes ROC pour des ensembles A et B (adapté de Wilks (2011))

### 2.6.5 Brier et CRPS

Le score de Brier (1950) s'intéresse uniquement aux prévisions d'évènements binaires et est défini comme la distance quadratique, moyennée sur  $R$  évènements, entre la probabilité d'occurrence d'un  $i^{\text{ème}}$  évènement,  $p_i$  et l'observation  $o_i$  de cet évènement :

$$BS = \frac{1}{R} \sum_{i=1}^R (p_i - o_i)^2 \quad (2.16)$$

où  $R$  est le nombre d'évènements étudiés. Puisque ce score concerne des évènements binaires,  $o_i$  vaut 1 si l'évènement a lieu et 0 si ce n'est pas le cas. En revanche, la probabilité d'occurrence d'un  $i^{\text{ème}}$  évènement,  $p_i$ , définie comme la proportion de membres ayant prévu l'évènement, varie, quant à elle, entre 0 et 1. Ce score mesurant la distance moyenne entre la prévision et l'observation dans l'espace des probabilités, il est d'autant meilleur qu'il est faible. Il peut par ailleurs être décomposé en trois termes (Murphy, 1973). L'un représentant la fiabilité doit être le plus faible possible alors que le terme de résolution doit être élevé. Le dernier terme est un terme d'incertitude sur l'occurrence de l'évènement.

On définit par ailleurs le Discrete Ranked Probability Score (DRPS) comme la moyenne des scores de Brier obtenus pour différents seuils de dépassement. Dans le même ordre d'idée, le Continuous Ranked Probability Score (CRPS) calcule cette moyenne pour une variable  $x$  continue :

$$CRPS = \int_{-\infty}^{\infty} BS(x) dx = \int_{-\infty}^{\infty} F(x) - H(x - o) dx \quad (2.17)$$

où  $F(x)$  est la distribution cumulative de probabilités de l'évènement prévu et  $H$  la distribution d'Heaviside associée à l'observation. Le CRPS mesure alors la distance entre la distribution de probabilité prévue cumulée et celle observée (Figure 2.6). Cette dernière est un échelon d'Heaviside valant 0 si la variable  $x$  est inférieure à l'observation et égale à 1 dans le cas contraire. Plus le CRPS est faible, meilleure

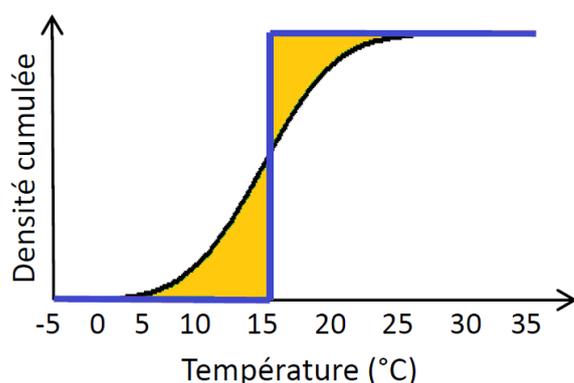


FIGURE 2.6 – Définition du CRPS : distance (orange) entre la distribution de probabilité cumulée prévue (noir) et observée (bleu). Exemple pour une température observée à 17°C

sera la prévision d'ensemble.

## 2.6.6 Rapport Dispersion/Erreur

L'étalement de la distribution de probabilités échantillonnée par les membres de la prévision d'ensemble doit être représentatif de l'incertitude de la prévision. Cette dispersion doit donc être cohérente avec l'erreur même de la prévision pour que la prévision d'ensemble soit fiable. Pour ce faire, on définit le rapport Dispersion/Erreur (Spread-Skill ratio), une mesure de la fiabilité, comme étant le ratio entre la dispersion de l'ensemble (soit l'écart-type de la distribution de probabilités) et la racine de l'erreur quadratique moyenne de la moyenne de l'ensemble :

$$SpSkRat = \frac{\sqrt{\frac{1}{R} \sum_{i=1}^R \left( \frac{1}{N} \sum_{j=1}^N (x_{i,j} - \bar{x})^2 \right)}}{\sqrt{\frac{1}{R} \sum_{i=1}^R \left( \frac{1}{N} \sum_{j=1}^N (x_{i,j} - o_i) \right)^2}} \quad (2.18)$$

où  $R$  est le nombre d'évènements et  $N$  le nombre de membres de la prévision d'ensemble.

Une prévision d'ensemble représente bien l'incertitude de la prévision si l'écart-type de la prévision d'ensemble est équivalent à l'erreur même, soit un rapport Dispersion/Erreur de 1. Une valeur de rapport Dispersion/Erreur supérieure à 1 dénotera une surdispersion de l'ensemble alors qu'une valeur inférieure à 1 suggérera une sous-dispersion.

## 2.6.7 Calcul des scores probabilistes

Dans ces travaux, nous avons utilisé un logiciel de scores existant et développé au CNRM pour calculer les différents scores probabilistes et déterministes précédemment présentés.

Les observations utilisées pour les calcul des scores proviennent de stations du réseau d'observations français Réseau d'Acquisition de Données et d'Observations Météorologiques Etendues (RADOME) et des données SYNOP du système mon-

dial de télécommunication de l'OMM. RADOME comprend environ 600 stations de mesures synoptiques issues du Réseau synoptique de base régional (RSBR), du Réseau synoptique national complémentaire (RSNC) situé dans les Centres départementaux météorologiques, sur les aéroports, dans les sémaphores ainsi que le Réseau régional automatisé (RRA) (Tardieu and Leroy, 2003). Les données SYNOP fournies par des stations d'observations à la surface, sur la France et ses pays voisins ainsi que les mesures réalisées sur environ 40 bateaux et 25 bouées sont aussi prises en compte. D'autres stations issues des réseaux OARA, institutionnels (Centrale Nucléaire de Production d'Electricité), d'intérêt commun étrangers (Suisse), entre autres, sont aussi considérées.

Ainsi, après élimination des stations placées sous liste noire dans le cycle d'assimilation de données, des mesures de température à 2m, humidité relative à 2m (calculée à partir des températures et températures au point de rosée à 2m et pression), vent à 10m, rafales sur 1h à 10m et précipitations accumulées sur 1, 3, 6, 12 et 24h sont disponibles sur près de 1500 stations. Pour environ 500 stations, des mesures de rayonnement, température au sol, visibilité, couverture nuageuse etc, sont également disponibles.

Une interpolation des données du modèle aux positions des observations avec la méthode des plus proches voisins, est réalisée afin d'estimer les erreurs de modèle vis-à-vis de ces observations. Concernant la température à 2m, l'effet de l'altitude sur la température est corrigé.

Des tests de significativité à 95% sont réalisés afin de vérifier la fiabilité de la différence du score d'un ensemble par rapport à un ensemble de référence. Il s'agit ici de tester si la différence moyenne de scores entre les deux ensembles est statistiquement significative quel que soit l'échantillon de dates choisies. Ainsi, pour vérifier cette fiabilité, une méthode de rééchantillonnage dit « bootstrap » est utilisée. Plus précisément, pour des scores moyennés sur  $n$  jours supposés indépendants, la méthode recalcule  $10 + 10 \times n$  fois, la différence de scores moyennée sur  $n$  dates tirées aléatoirement avec remise parmi les  $n$  dates de départ. Ces  $10 + 10 \times n$  différences de scores sont ensuite comparées à la différence de référence calculée sur les  $n$  dates de départ : sur ces  $10 + 10 \times n$  tirages aléatoires, on compte le nombre de fois où la différence de scores entre les deux ensembles a un signe différent de celle de référence. Autrement dit, on compte le nombre de fois où il y a détérioration du score au lieu d'amélioration. Le test de significativité consiste en la comparaison du ratio entre ce nombre et le nombre de tirages ( $10 + 10 \times n$ ), soit le taux de détérioration du score, avec un seuil de significativité à 5%.

**Conclusion :** Après s'être intéressé aux différentes méthodes de représentation des sources d'incertitude dans les prévisions d'ensemble ainsi qu'aux méthodes d'évaluation de ces ensembles, nous allons porter notre intérêt sur le modèle ensembliste faisant l'objet de notre étude, la PEARO.

## Bibliographie

- L. H. Baker, A. C. Rudd, S. Migliorini, and R. N. Bannister. Representation of model error in a convective-scale ensemble prediction system. *Nonlinear Processes in Geophysics*, 21(1) :19–39, 2014. doi : 10.5194/npg-21-19-2014.

- P. Bauer, A. Thorpe, and G. Brunet. The quiet revolution of numerical weather prediction. *Nature*, 525 :47–55, 2015. doi : 10.1038/nature14956.
- J. Berner, G. J. Shutts, M. Leutbecher, and T. N. Palmer. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences*, 66(3) :603 – 626, 2009. doi : 10.1175/2008JAS2677.1.
- J. Berner, S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder. Model uncertainty in a mesoscale ensemble prediction system : Stochastic versus multiphysics representations. *Monthly Weather Review*, 139(6) :1972 – 1995, 2011. doi : 10.1175/2010MWR3595.1.
- J. Berner, U. Achatz, L. Batté, L. Bengtsson, A. De la Cámara, H. Christensen, M. Colangeli, D. Coleman, D. Crommelin, S. Dolaptchiev, C. Franzke, P. Friedrichs, P. Imkeller, H. Järvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajan, T. Palmer, C. Penland, M. Sakradzija, J.-S. Von Storch, A. Weisheimer, M. Weniger, P. William, and J.-I. Yano. Stochastic parameterization : Towards a new view of weather and climate models. *Bulletin of the American Meteorological Society*, 98(3) :565–588, 2017. doi : 10.1175/BAMS-D-15-00268.1.
- C. H. Bishop, B. J. Etherton, and S. J. Majumdar. Adaptive sampling with the ensemble transform kalman filter. part I : Theoretical aspects. *Monthly Weather Review*, 129(3) :420 – 436, 2001. doi : 10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.
- F. Bouttier and L. Raynaud. Clustering and selection of boundary conditions for limited-area ensemble prediction. *Quarterly Journal of the Royal Meteorological Society*, 144(717) :2381–2391, 2018. doi : 10.1002/qj.3304.
- F. Bouttier, L. Raynaud, O. Nuissier, and B. Ménétrier. Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. *Quarterly Journal of the Royal Meteorological Society*, 142(S1) :390–403, 2016. doi : 10.1002/qj.2622.
- N. E. Bowler, A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare. The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 134(632) :703–722, 2008. doi : 10.1002/qj.234.
- G. W. Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1) :1 – 3, 1950. doi : 10.1175/1520-0493(1950)0780001:VOFEIT2.0.CO;2.
- R. Buizza and T. N. Palmer. The singular-vector structure of the atmospheric global circulation. *Journal of the Atmospheric Sciences*, 52(9) :1434–1456, 1995. doi : 10.1175/1520-0469(1995)052<1434:TSVSOT>2.0.CO;2.
- R. Buizza, M. Milleer, and T. N. Palmer. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125(560) :2887–2908, 1999. doi : 10.1002/qj.49712556006.

- A. Callado, P. Escribà, J. A. García-Moya, J. Montero, C. Santos, D. Santos-Muñoz, and J. Simarro. Ensemble forecasting. In Y. Zhang and P. Ray, editors, *Climate Change and Regional/Local Responses*, chapter 1. IntechOpen, 2013. doi : 10.5772/55699.
- A. Callado, P. Escribà, and J. A. García-Moya. AEMET- SREPS. In C. Santos-Burguete, editor, *Física del caos en la predicción meteorológica*, chapter 22, pages 331–356. Agencia Estatal de Meteorología (Spanish Meteorological Agency), 2018. doi : 10.31978/014-18-009-X.
- G. Candille and O. Talagrand. Evaluation of probabilistic prediction systems for a scalar variable. *Quarterly Journal of the Royal Meteorological Society*, 131 (609) :2131–2150, 2005. doi : 10.1256/qj.04.71.
- G. Candille, C. Côté, P. L. Houtekamer, and G. Pellerin. Verification of an ensemble prediction system against observations. *Monthly Weather Review*, 135 (7) :2688 – 2699, 2007. doi : 10.1175/MWR3414.1.
- H. Christensen, I. Moroz, and T. Palmer. Stochastic and perturbed parameter representations of model uncertainty in convection parameterization. *Journal of the Atmospheric Sciences*, 72, 2015. doi : 10.1175/JAS-D-14-0250.1.
- H. M. Christensen, S.-J. Lock, I. M. Moroz, and T. N. Palmer. Introducing independent patterns into the stochastically perturbed parametrization tendencies (SPPT) scheme. *Quarterly Journal of the Royal Meteorological Society*, 143 (706) :2168–2181, 2017. doi : 10.1002/qj.3075.
- C. Clancy, R. Darcy, S. Gallagher, E. Gleeson, A. Hally, J. Hanley, and E. Whelan. Met Éireann updates. Technical report, 2018. ALADIN-HIRLAM Newsletter 10.
- G. Deng, Y. Zhu, J. Gong, D. Chen, R. Wobus, and Z. Zhang. The effects of land surface process perturbations in a global ensemble forecast system. *Advances in Atmospheric Sciences*, 33(10) :1199–1208, 2016. doi : 10.1007/s00376-016-6036-8.
- L. Descamps, C. Labadie, A. Joly, E. Bazile, P. Arbogast, and P. Cébron. PEARP, the Météo-France short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 141(690) :1671–1685, 2015. doi : 10.1002/qj.2469.
- J. Du, J. Berner, R. Buizza, M. Charron, P. Houtekamer, D. Hou, I. Jankov, M. Mu, X. Wang, M. Wei, and H. Yuan. *Ensemble Methods for Meteorological Predictions*, pages 1–52. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN 978-3-642-40457-3. doi : 10.1007/978-3-642-40457-3\_13-1.
- G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research : Oceans*, 99(C5) :10143–10162, 1994. doi : 10.1029/94JC00572.

- I.-L. Frogner, U. Andrae, J. Bojarova, A. Callado, P. Escribà, H. Feddersen, A. Hally, J. Kauhanen, R. Randriamampianina, A. Singleton, G. Smet, S. van der Veen, and O. Vignes. HarmonEPS the HARMONIE ensemble prediction system. *Weather and Forecasting*, 34(6) :1909 – 1937, 2019. doi : 10.1175/WAF-D-19-0030.1.
- Z. Gao, J. Zhu, Y. Guo, X. Yan, X. Wang, H. Li, and S. Li. Comparison of different land-surface perturbation methods in short-range ensemble forecasts. *Atmospheric and Oceanic Science Letters*, 14(3) :1674–2834, 2021. doi : 10.1016/j.aosl.2021.100045.
- J. A. García-Moya, A. Callado, P. Escribà, and C. Santos. SLAF implementation in armonEPS : First results. ALADIN-HIRLAM The 25th Workshop All Staff Meeting, Elsinore, 13-16 April 2015. URL <http://www.umr-cnrm.fr/aladin/IMG/pdf/slaf.pdf>.
- J.-A. García-Moya, A. Callado, P. Escribà, C. Santos, D. S.-M. nOz, and J. Simarro. Predictability of short-range forecasting : a multimodel approach. *Tellus A : Dynamic Meteorology and Oceanography*, 63(3) :550–563, 2011. doi : 10.1111/j.1600-0870.2010.00506.x.
- C. Gebhardt, S. Theis, M. Paulat, and Z. Ben Bouallègue. Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbation and variations of later boundaries. *Atmospheric Research*, 100 :168–177, 2011. doi : 10.1016/j.atmosres.2010.12.008.
- M. Gehne, T. M. Hamill, G. T. Bates, P. Pegion, and W. Kolczynski. Land surface parameter and state perturbations in the global ensemble forecast system. *Monthly Weather Review*, 147(4) :1319 – 1340, 2019. doi : 10.1175/MWR-D-18-0057.1.
- S. Gonzalez, A. Callado, M. Martínez, and B. Elvira. The AEMET- $\gamma$ SREPS over the antarctic peninsula and the impact of kilometeric-resolution EPS on logistic activities on the continent. *Advances in Science and Research*, 17 :209–217, 2020. doi : 10.5194/asr-17-209-2020.
- J. P. Hacker, C. Snyder, S.-Y. Ha, and M. Pocerlich. Linear and non-linear response to parameter variations in a mesoscale model. *Tellus A : Dynamic Meteorology and Oceanography*, 63(3) :429–444, 2011. doi : 10.1111/j.1600-0870.2010.00505.x.
- P. Houtekamer, L. Lefaivrem, J. Derome, H. Ritchie, and H. Mitchell. A system simulation approach to ensemble prediction. *Monthly Weather Review*, 124 : 1225–1242, 1996. doi : 10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2.
- T. Iversen, A. Deckmyn, C. Santos, K. Sattler, J. B. Bremnes, H. Feddersen, and I.-L. Frogner. Evaluation of ‘GLAMEPS’—a proposed multimodel EPS for short range forecasting. *Tellus A : Dynamic Meteorology and Oceanography*, 63(3) : 513–530, 2011. doi : 10.1111/j.1600-0870.2010.00507.x.

- I. Jankov, J. Berner, J. Beck, H. Jiang, J. B. Olson, G. Grell, T. G. Smirnova, S. G. Benjamin, and J. M. Brown. A performance comparison between multiphysics and stochastic approaches within a north american RAP ensemble. *Monthly Weather Review*, 145(4) :1161 – 1179, 2017. doi : 10.1175/MWR-D-16-0160.1.
- I. Jankov, J. Beck, J. Wolff, M. Harrold, J. B. Olson, T. Smirnova, C. Alexander, and J. Berner. Stochastically perturbed parameterizations in an HRRR-Based ensemble. *Monthly Weather Review*, 147(1) :153 – 173, 2019. doi : 10.1175/MWR-D-18-0092.1.
- K. Kober and G. C. Craig. Physically based stochastic perturbations (PSP) in the boundary layer to represent uncertainty in convective initiation. *Journal of the Atmospheric Sciences*, 73(7) :2893 – 2911, 2016. doi : 10.1175/JAS-D-15-0144.1.
- M.-S. Koo and S.-Y. Hong. Stochastic representation of dynamic model tendency : Formulation and preliminary results. *Asia-Pacific Journal of Atmospheric Sciences*, 50 :497–506, 2014. doi : 10.1007/s13143-014-0039-0.
- J.-F. Lacarra and O. Talagrand. Short range evolution of small perturbation in a barotropic model. *Tellus A*, 40A :81 – 95, 2010. doi : 10.1111/j.1600-0870.1988.tb00408.x.
- D. Lambert and S. Argence. Preliminary study of an intense rainfall episode in Corsica, 14 september 2006. *Advances in Geosciences*, 16 :125–129, 2008. doi : 10.5194/adgeo-16-125-2008.
- C. Lavaysse, M. Carrera, S. Bélair, N. Gagnon, R. Frenette, M. Charron, and M. K. Yau. Impact of surface parameter uncertainties within the canadian regional ensemble prediction system. *Monthly Weather Review*, 141(5) :1506 – 1526, 2013. doi : 10.1175/MWR-D-11-00354.1.
- M. Leutbecher, S.-J. Lock, P. Ollinaho, S. T. K. Lang, G. Balsamo, P. Bechtold, M. Bonavita, H. M. Christensen, M. Diamantakis, E. Dutra, S. English, M. Fisher, R. M. Forbes, J. Goddard, T. Haiden, R. J. Hogan, S. Juricke, H. Lawrence, D. MacLeod, L. Magnusson, S. Malardel, S. Massart, I. Sandu, P. K. Smolarkiewicz, A. Subramanian, F. Vitart, N. Wedi, and A. Weisheimer. Stochastic representations of model uncertainties at ecmwf : state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, 143(707) : 2315–2339, 2017. doi : 10.1002/qj.3094.
- J. Li, J. Du, Y. Liu, and X. Jianyu. Similarities and differences in the evolution of ensemble spread using various ensemble perturbation methods including terrain perturbation. *Acta Meteorologica Sinica*, 75(1) :123, 2017.
- S. Lim, M.-S. Koo, K. In-Hyuk, and S. K. Park. Model error representation using the stochastically perturbed hybrid physical-dynamical tendencies in ensemble data assimilation system. *Applied Sciences*, 10, 2020. doi : 10.3390/app10249010.
- L. Magnusson, E. Källén, and J. Nycander. Initial state perturbations in ensemble forecasting. *Nonlinear Processes in Geophysics*, 15(5) :751–759, 2008. doi : 10.5194/npg-15-751-2008. URL <https://npg.copernicus.org/articles/15/751/2008/>.

- A. McCabe, R. Swinbank, W. Tennant, and A. Lock. Representing model uncertainty in the met office convection-permitting ensemble prediction system and its impact on fog forecasting. *Quarterly Journal of the Royal Meteorological Society*, 142(700) :2897–2910, 2016. doi : 10.1002/qj.2876.
- G. Mellor and T. Yamada. Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20 :851–875, 1982.
- F. Molteni, R. Buizza, C. Marsigli, A. Montani, F. Nerozzi, and T. Paccagnella. A strategy for high-resolution ensemble prediction. I : Definition of representative members and global-model experiments. *Quarterly Journal of the Royal Meteorological Society*, 127(576) :2069–2094, 2001. doi : 10.1002/qj.49712757612.
- M. Mu, W. S. Duan, and B. Wang. Conditional nonlinear optimal perturbation and its applications. *Nonlinear Processes in Geophysics*, 10(6) :493–501, 2003.
- A. H. Murphy. A new vector partition of the probability score. *Journal of Applied Meteorology and Climatology*, 12(4) :595 – 600, 1973. doi : 10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2.
- J. Murphy, D. Sexton, D. Barnett, G. Jones, M. Webb, M. Collins, and D. Stainforth. Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, 430 :768–72, 2004. doi : 10.1038/nature02771.
- O. Nuissier, B. Joly, B. Vié, and V. Ducrocq. Uncertainty of lateral boundary conditions in a convection-permitting ensemble : a strategy of selection for mediterranean heavy precipitation events. *Natural Hazards and Earth System Sciences*, 12(10) :2993–3011, 2012. doi : 10.5194/nhess-12-2993-2012.
- P. Ollinaho, M. Laine, A. Solonen, H. Haario, and H. Järvinen. NWP model forecast skill optimization via closure parameter variations. *Quarterly Journal of the Royal Meteorological Society*, 139(675) :1520–1532, 2013. doi : 10.1002/qj.2044.
- P. Ollinaho, S.-J. Lock, M. Leutbecher, P. Bechtold, A. Beljaars, A. Bozzo, R. M. Forbes, T. Haiden, R. J. Hogan, and I. Sandu. Towards process-level representation of model uncertainties : stochastically perturbed parametrizations in the ECMWF ensemble. *Quarterly Journal of the Royal Meteorological Society*, 143(702) :408–422, 2017. doi : 10.1002/qj.2931.
- T. Palmer, R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer. Stochastic parametrization and model uncertainty. Technical Report 598, ECMWF Technical Memoranda, Shinfield Park, Reading, 2009. URL <https://www.ecmwf.int/node/11577>.
- C. Peralta, Z. Ben Bouallègue, S. E. Theis, C. Gebhardt, and M. Buchhold. Accounting for initial condition uncertainties in COSMO-DE-EPS. *Journal of Geophysical Research : Atmosphere*, 117(D7), 2012. doi : 10.1029/2011JD016581.
- M. Ponzano, B. Joly, L. Descamps, and P. Arbogast. Systematic errors analysis of heavy precipitating events prediction using a 30-year hindcast dataset. *Natural Hazards and Earth System Sciences*, 20 :1369–1389, 2020. doi : 10.5194/nhess-20-1369-2020.

- X. Qiao, S. Wang, and J. Min. A stochastic perturbed parameterization tendency scheme for diffusion (SPPTD) and its application to idealized supercell simulation. *Monthly Weather Review*, 145, 2017. doi : 10.1175/MWR-D-16-0307.1.
- L. Raynaud and F. Bouttier. Comparison of initial perturbation methods for ensemble prediction at convective scale. *Quarterly Journal of the Royal Meteorological Society*, 142(695) :854–866, 2016. doi : 10.1002/qj.2686.
- C. A. Reynolds, J. A. Ridout, and J. G. Mclay. Examination of parameter variations in the U. S. navy global ensemble. *Tellus A : Dynamic Meteorology and Oceanography*, 63(5) :841–857, 2011. doi : 10.1111/j.1600-0870.2011.00532.x.
- G. Rivière, M. Wimmer, P. Arbogast, J.-M. Piriou, J. Delanoë, C. Labadie, Q. Cazenave, and J. Pelon. The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during nawdex iop6. *Weather and Climate Dynamics*, 2021 :1–32, 2021. doi : 10.5194/wcd-2021-38. URL <https://wcd.copernicus.org/preprints/wcd-2021-38/>.
- M. Sakradzija and D. Klocke. Physically constrained stochastic shallow convection in realistic kilometer-scale simulations. *Journal of Advances in Modeling Earth Systems*, 10, 2018. doi : 10.1029/2018MS001358.
- M. Sakradzija, A. Seifert, and T. Heus. Fluctuations in a quasi-stationary shallow cumulus cloud ensemble. *Nonlinear Processes in Geophysics*, 22(1) :65–85, 2015. doi : 10.5194/npg-22-65-2015.
- P. Scheffknecht, E. Richard, and D. Lambert. A highly localized high-precipitation event over Corsica. *Quarterly Journal of the Royal Meteorological Society*, 142 (S1) :206–221, 2016. doi : 10.1002/qj.2795.
- G. Shutts. A stochastic kinetic energy backscatter algorithm for use in ensemble prediction systems. Technical Report 449, ECMWF Technical Memoranda, Shinfield Park, Reading, 2004. URL <https://www.ecmwf.int/node/12214>.
- G. Shutts. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 131(612) :3079–3102, 2005. doi : 10.1256/qj.04.106.
- D. A. Stainforth, T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Piani, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe, and M. R. Allen. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, 433(7024) :403–406, 2005. doi : 10.1038/nature03301.
- C. Sutton, T. M. Hamill, and T. T. Warner. Will perturbing soil moisture improve warm-season ensemble forecasts? A proof of concept. *Monthly Weather Review*, 134(11) :3174 – 3189, 2006. doi : 10.1175/MWR3248.1.
- O. Talagrand, R. Vautard, and B. Strauss. Evaluation of probabilistic prediction systems. *Proc. ECMWF Workshop on Predictability*, 4.2, 1999.
- J. Tardieu and M. Leroy. Radome, le réseau temps réel d'observation au sol de météo-france. *La Météorologie*, 8(40) :40–43, 2003. doi : 10.4267/2042/36262.

- W. Tennant. Improving initial condition perturbations for mogreps-uk. *Quarterly Journal of the Royal Meteorological Society*, 141(691) :2324–2336, 2015. doi : 10.1002/qj.2524.
- W. Tennant and S. Beare. New schemes to perturb sea-surface temperature and soil moisture content in MOGREPS. *Quarterly Journal of the Royal Meteorological Society*, 140(681) :1150–1160, 2014. doi : 10.1002/qj.2202.
- Z. Toth and E. Kalnay. Ensemble forecasting at NMC : The generation of perturbations. *Bulletin of the American Meteorological Society*, 74(12) :2317 – 2330, 1993. doi : 10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.
- Y. Wang, A. Kann, M. Bellus, J. Pailleux, and C. Wittmann. A strategy for perturbing surface initial conditions in LAMEPS. *Atmospheric Science Letters*, 11(2) :108–113, 2010. doi : 10.1002/asl.260.
- C. Wastl, Y. Wang, A. Atencia, and C. Wittmann. Independent perturbations for physics parametrization tendencies in a convection-permitting ensemble (pSPPT). *Geoscientific Model Development*, 12(1) :261–273, 2019a. doi : 10.5194/gmd-12-261-2019.
- C. Wastl, Y. Wang, A. Atencia, and C. Wittmann. A hybrid stochastically perturbed parametrization scheme in a convection-permitting ensemble. 147(6) : 2217 – 2230, 2019b. doi : 10.1175/MWR-D-18-0415.1.
- D. Wilks. Chapter 3 - empirical distributions and exploratory data analysis. In D. S. Wilks, editor, *Statistical Methods in the Atmospheric Sciences*, volume 100 of *International Geophysics*, pages 23–70. Academic Press, 2011. doi : <https://doi.org/10.1016/B978-0-12-385022-5.00003-8>.

# Chapitre 3

## Prévision d'Ensemble AROME : Outils et objectifs

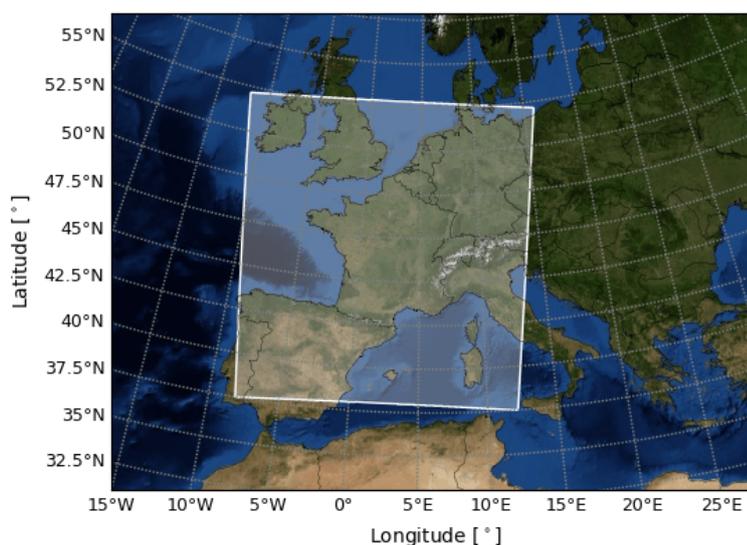
Le travail de thèse réalisé ici s'appuie sur la Prévision d'Ensemble AROME (PEARO ; Bouttier et al., 2012), la prévision d'ensemble opérationnelle à l'échelle régionale à Météo-France, basée sur le modèle AROME-France (Seity et al., 2011). Nous allons tout d'abord présenter le modèle AROME et la configuration de la PEARO avant d'argumenter les objectifs de la thèse.

### 3.1 AROME

Le modèle déterministe Application of Research to Operations at MESoscale (AROME) est un modèle à échelle convective, opérationnel à Météo-France depuis 2008 (Seity et al., 2011). Il produit des prévisions à 48h d'échéance sur un domaine à aire limitée couvrant l'Europe occidentale et centré sur la France (figure 3.1). Ce domaine de 1600 km par 1600 km a une résolution horizontale de 1.3km et 90 niveaux verticaux (dont le plus bas est à 5m). Les conditions de surface et initiales proviennent d'une assimilation de données 3D-Var (Brousseau et al., 2016) et les conditions latérales sont issues du modèle opérationnel global de Météo-France, ARPEGE (Courtier et al., 1991; Déqué et al., 1994). Ce dernier possède une maille étirée gaussienne réduite dont la résolution horizontale avoisine les 5km sur la France et sa résolution verticale est composée de 105 niveaux. La méthode de Davies (Radnóti, 1995) utilisant une zone de relaxation d'une largeur de 8 points de grille permet de coupler les deux modèles.

#### 3.1.1 Coeur dynamique

Le coeur dynamique d'AROME est issu du modèle ALADIN (Aire Limitée Adaptation Développement INternational; Bubnová et al., 1995; Bénard et al., 2010). Il s'agit d'un modèle non-hydrostatique résolvant les équations d'Euler compressibles dont les variables pronostiques sont le vent horizontal, la température, ainsi que deux variables reliées à la pression et divergence verticale. La coordonnée verticale utilisée est hybride en masse; il s'agit d'une pression hydrostatique pondérée afin que celle-ci suive la forme du terrain (relief).

FIGURE 3.1 – *Domaine de prévision d'AROME-France*

Ce modèle spectral utilise un schéma semi-implicite ainsi qu'un schéma semi-lagrangien à deux pas de temps pour la discrétisation temporelle. Le schéma semi-implicite est complété, pour les raisons de stabilité numérique, par un schéma prédicteur-correcteur à deux itérations. Ceci permet alors d'utiliser un pas de temps relativement long de 50s. Une double transformation spectrale de Fourier de certaines variables d'état permet ensuite de calculer les dérivées spatiales ainsi que la diffusion numérique dans l'espace spectral. Cette diffusion numérique est linéaire d'ordre 4 avec la même intensité pour toutes les variables pronostiques. Cependant, puisque les hydrométéores ne sont pas convertis dans l'espace spectral, un schéma de Diffusion Horizontale Semi-Lagrangienne (SLHD) est ajouté (Váňa et al., 2008) afin de diffuser les hydrométéores à partir du champ de déformation dynamique.

### 3.1.2 Paramétrisations physiques et dynamiques

Contrairement à ALADIN, les paramétrisations physiques d'AROME sont issues du modèle Mésos-NH (Lafore et al., 1998).

**Rayonnement :** La modélisation du rayonnement dans AROME utilise le schéma de rayonnement utilisé jusqu'en 2016 dans le modèle IFS. Le schéma de Fouquart and Bonnel (1980) discrétise les calculs du rayonnement solaire en 6 bandes spectrales. La modélisation de Nielsen et al. (2014) fournit les propriétés optiques dans les nuages. Les rayons effectifs des particules sont calculés différemment pour les particules liquides (à partir de l'eau liquide nuageuse dans Martin et al. (1994)) et pour les particules glacées (à partir de la température dans Ou and Liou (1995)). Les noyaux de condensations sont considérés comme constants et leur concentration diffèrent selon que le point de grille considéré se trouve sur terre ou sur mer. Le rayonnement thermique en revanche, utilise le schéma Rapid Radiative Transfer Model (RRTM; Mlawer et al., 1997). Les différentes espèces chimiques altérant ce rayonnement thermique ainsi que les aérosols sont issus de la climatologie de (Tegen et al., 1997). L'ozone est remis à jour tous les mois à

partir des données de l'UGAMP (Li and Shine, 1995). Afin de réduire le temps de calcul, le schéma de rayonnement complet est appelé toutes les 15 minutes, soit tous les 18 pas de temps. Aux pas de temps intermédiaires, les flux solaires sont modifiés uniquement par la variation de l'angle solaire zénithal.

**Microphysique nuageuse :** Le schéma de microphysique, à un seul moment, considère cinq contenus spécifiques d'hydrométéores en plus de la vapeur d'eau ( $r_v$ ) : la pluie ( $r_r$ ), la neige ( $r_s$ ), le graupel ( $r_g$ ) pour les hydrométéores précipitants ainsi que la glace primaire ( $r_i$ ) et les gouttelettes nuageuses ( $r_c$ ) pour la composition des nuages (ICE3; Pinty and Jabouille, 1998). Ceux-ci sont alors des variables prognostiques supplémentaires à celles des équations d'Euler (tout comme l'énergie cinétique turbulente apparaissant dans le schéma de turbulence). La grêle étant considérée comme de gros graupels, seules trois classes d'hydrométéores glacés sont admises, d'où le nom ICE3 du schéma. Les processus chauds sont régis par un schéma de Kessler. Ainsi, près de 25 processus reproduisant l'évolution de ces hydrométéores sont calculés explicitement et sont schématisés dans la figure 3.2. Le spectre des hydrométéores est supposé de la forme d'une distribution gamma généralisée. Une loi en puissance relie la masse et la vitesse de chute des hydrométéores précipitants. Afin d'améliorer ICE3, un schéma de condensation sous-maille, utilisant la variance d'humidité au sein de la maille calculée par le schéma de turbulence, a été introduit (Bougeault, 1982; Bechtold et al., 1995). Par ailleurs, un schéma statistique de sédimentation, utilisant une fonction de distribution de probabilité, réalise un bilan des hydrométéores sur la verticale afin de déterminer les flux descendants d'hydrométéores précipitants (Bouteloup et al., 2011).

**Turbulence :** La turbulence dans la couche limite atmosphérique est modélisée par l'équation prognostique de l'énergie cinétique turbulente avec le schéma de Cuxart et al. (2000) et utilise la longueur de mélange établie par le schéma de Bougeault and Lacarrere (1989). Cette dernière est calculée à partir de la distance parcourue par une parcelle d'air adiabatique lors de mouvements ascendants et descendants dus à la flottabilité, jusqu'à ce qu'elle ait consommé toute son énergie cinétique turbulente.

**Convection peu profonde :** AROME ayant une résolution horizontale assez fine pour résoudre explicitement la convection profonde, seul le schéma de convection peu profonde de (PMMC; Pergaud et al., 2009) est utilisé. Celui-ci représente les thermiques secs ainsi que les cumulus par un schéma en flux de masse dit « Eddy Diffusivity Mass Flux » (Siebesma and Teixeira, 2000). Il représente donc une seule ascendance avec les phénomènes d'entraînement et de déentraînement. Ces derniers dépendent de la vitesse verticale et de la flottabilité dans les thermiques secs, alors que dans les nuages, ils suivent la formulation de Kain and Fritsch (1990).

**Surface :** En chaque point de grille d'AROME, la surface, représentée par le schéma externe SURFEX (Le Moigne et al., 2009), est décrite par quatre types de sol en proportion variable : la terre, la ville, la mer et les lacs. La paramétrisation Interactions between Soil, Biosphere and Atmosphere (ISBA; Noilhan and

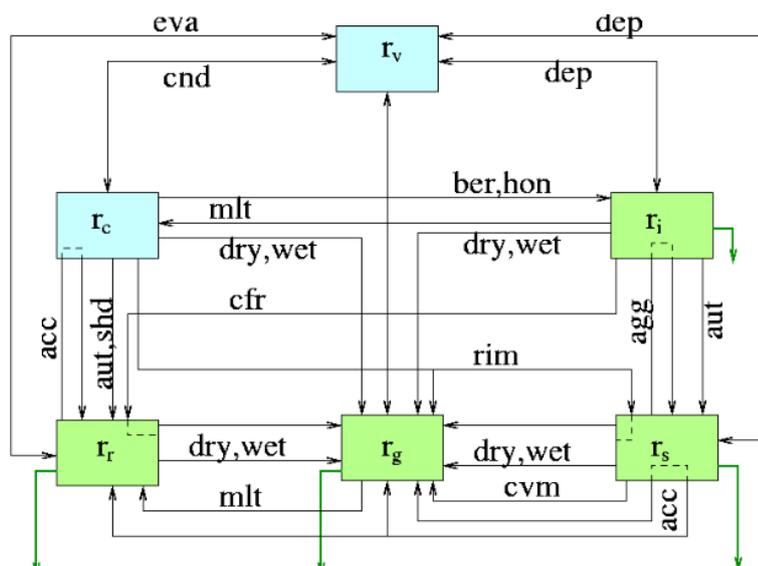


FIGURE 3.2 – Processus microphysiques entre la vapeur d'eau ( $r_v$ ), la pluie ( $r_r$ ), la neige ( $r_s$ ), le graupel ( $r_g$ ), la glace primaire ( $r_i$ ) et les gouttelettes nuageuses ( $r_c$ ) : *eva* = évaporation, *dep* = déposition de vapeur, *cnd* = condensation, *ber* = effet Bergeron-Findeisen, *hon* = nucléation homogène, *mlt* = fonte, *dry, wet* = accroissement sec et humide, *acc* = accréation, *aut* = autoconversion, *shd* = perte d'eau, *cfr* = solidification par contact, *agg* = agrégation, *rim* = givrage, *cvm* = fonte de conversion.

Planton, 1989) modélise la terre et son évolution en deux niveaux verticaux. Le schéma Town Energy Budget (TEB ; Masson, 2000) reproduit, entre autres, les canyons et les îlots de chaleur urbains. La mer est modélisée par le schéma Exchange Coefficients from Unified Multicampaigns Estimates (ECUME ; Belamari, 2005), reproduisant les flux océaniques à partir de campagnes de mesures passées. La formulation de Charnock (1955) est utilisée pour la représentation des lacs. L'orographie provient quant à elle de la base de donnée Global 30 Arc-Second Elevation Data Set alors que la base de données ECOCLIMAP fournit les données physiographiques (Masson et al., 2003).

Ainsi, l'ordre d'apparition des paramétrisations pendant un pas de temps du modèle est le suivant (voir figure 3.3 pour un schéma). Tout d'abord, ICE3 (condensation sous-maille) met à jour les contenus en eau et glace nuageuses et établit un état thermodynamique équilibré, sur lequel vont travailler les paramétrisations physiques. Ensuite, le schéma de rayonnement calcule les flux radiatifs ; dont ceux à la surface sont ensuite réutilisés dans SURFEX. Ce dernier fournit les flux de température potentielle et d'humidité à PMMC, ainsi que le flux de quantité de mouvements horizontal au schéma de turbulence. Notons que le calcul de l'albédo et autres propriétés optiques dans SURFEX sont pris en compte par le schéma de rayonnement au pas de temps suivant. Le schéma de convection peu profonde estime ensuite, les contenus spécifiques d'eau et glace nuageuse. Puis, la turbulence est activée. Elle permet d'estimer la variance de la saturation sous maille pour l'ajustement d'ICE3 au prochain pas de temps. Enfin, le schéma de microphysique est pleinement réalisé et se termine par le calcul des précipitations grâce à la représentation de la sédimentation. Le résultat de ce dernier calcul sera réutilisé au pas de temps suivant par SURFEX. Chacune de ces paramétrisa-

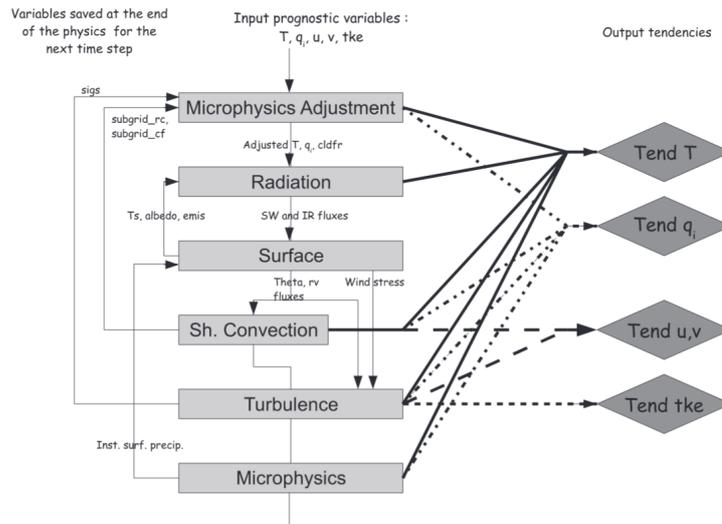


FIGURE 3.3 – Déroulement de l’activation des paramétrisations physiques pendant un pas de temps du modèle AROME.

tions fournit alors un échantillon de tendances de température, humidité, vent et énergie cinétique turbulente qui sont sommées et ajoutées aux tendances dues à l’advection et à la dynamique pour estimer l’évolution de l’état de l’atmosphère.

## 3.2 Prévision d’Ensemble régionale PEARO

Dans sa première version, la PEARO comptait 12 membres, et en compte 16 depuis Juillet 2019. Dans cette étude, nous utiliserons la PEARO avec les 12 membres initiaux. Contrairement à AROME, la PEARO a une résolution horizontale de 2.5 km (1.3km pour AROME) mais les niveaux verticaux restent identiques (90 niveaux). Son pas de temps est de 60s et elle fournit des prévisions de 45 à 51h d’échéance pour 4 réseaux quotidiens (45h pour initialisation à 3UTC et 15UTC puis 51h pour 9UTC et 21UTC).

### 3.2.1 EDA : l’Assimilation d’Ensemble AROME

Les conditions initiales des 12 membres de la PEARO sont issues de 12 membres tirés aléatoirement parmi les 25 membres de l’Assimilation d’Ensemble AROME (AEARO) (Bouttier et al., 2012).

Plus précisément, l’analyse servant de condition initiale au membre  $i$  de la PEARO ( $x_{i,PEARO}^a$ ) utilise l’analyse du  $j^{\text{ème}}$  membre de l’AEARO ( $x_{j,AEARO}^a$ ). La différence de cette dernière avec l’analyse moyenne de l’AEARO ( $\frac{1}{N} \sum_{j=1}^N x_{j,AEARO}^a$ ), permet de fournir les perturbations des conditions initiales appliquées à l’analyse AROME  $x_{AROME}^a$ . La formulation est donc la suivante :

$$x_{i,PEARO}^a = x_{AROME}^a + \left( x_{j,AEARO}^a - \frac{1}{N} \sum_{j=1}^N x_{j,AEARO}^a \right) \quad (3.1)$$

où  $i = 1, 2, \dots, M$  le  $i^{\text{ème}}$  membre de la PEARO comprenant  $M = 12$  membres et  $j = 1, 2, \dots, N$  le  $j^{\text{ème}}$  membre, tiré aléatoirement, de l'AEARO comprenant  $N = 25$  membres.

Les membres de l'assimilation d'ensemble AROME sont couplés à ceux de l'assimilation d'ensemble ARPEGE et correspondent à 25 assimilations différentes de données dites 3D-Var. Ces analyses assimilent pendant des cycles de 3h, les données de radio-sondages, stations au sol, profileurs de vent, bouées, bateaux, avions, GPS, satellites, radar de vent Doppler et réflectivité radar. Dans les 25 assimilations, les données observées sont perturbées selon une distribution gaussienne (Brousseau et al., 2011). La variance de cette perturbation des observations est définie par les statistiques d'erreurs des instruments de mesures associés. Les ébauches sont également perturbées au travers du cyclage temporel de l'EDA, et les perturbations d'analyse résultent ainsi des perturbations d'observation et d'ébauche. Chaque étape de prévision sur 3h conduit à son tour à des perturbations de prévision; celles-ci sont amplifiées de façon multiplicative, pour que la dispersion des prévisions soit cohérente avec la variance d'erreur de prévision, estimée à l'aide des écarts à l'analyse déterministe. Le facteur d'amplification peut être vu comme une mesure de l'amplitude d'erreur modèle accumulée au cours de chaque prévision de 3h.

### 3.2.2 Perturbation de la surface

La représentation des erreurs de surface, décrites dans Bouttier et al. (2016), se focalise sur la perturbation des paramètres suivants : indice de végétation, coefficient de chaleur de la végétation, LAI, albédo terrestre, longueur de rugosité, température de surface terrestre et océanique, humidité des différents types de sol (sauf lac et ville), épaisseur de la neige et flux à la surface océanique. La plupart des paramètres sont perturbés par un facteur  $\alpha$  de façon multiplicative ( $x = x(1 + \alpha)$ ), seule la température du sol est perturbée de manière additive ( $x = x + \alpha$ ).

Les perturbations sont de types gaussiennes dont l'amplitude est optimisée pour chaque paramètre de surface. La longueur de corrélation spatiale est cependant fixée à 400km pour tous les paramètres.

### 3.2.3 Couplage avec la PEARP

Chacun des 12 membres de la PEARO est couplé, à résolution horaire, à l'un des 34 membres de la Prévision d'Ensemble ARPEGE de Météo-France (PEARP; Descamps et al., 2015). Ce modèle ensembliste à l'échelle globale à courte échéance, basée sur le modèle ARPEGE, comprend 35 membres (dont un membre de contrôle non perturbé) allant jusqu'à 90h voire 108h d'échéance. Deux réseaux sont disponibles et proposent des prévisions de 90 et 108h initialisées à 6UTC et 18UTC respectivement. Deux autres runs initialisés à 0UTC et 12UTC fournissent des prévisions de seulement 48h. Le modèle ARPEGE utilisant une maille étirée dont le facteur d'étirement est fixé à 2.2, la résolution TL1198 de la PEARP fournit une résolution de 7.5km sur la France et 37km aux antipodes. Le modèle comprend 90 niveaux verticaux allant de 14m à 1hPa. Le pas de temps est de 514.3s.

La PEARP est initialisée par 35 des 50 membres de l'Assimilation d'Ensemble ARPEGE choisis aléatoirement. Cette méthode, combinée à la méthode des vec-

teurs singuliers calculés sur l'Europe, l'Atlantique Nord, l'hémisphère Sud, la bande intertropicale et sur le reste de l'hémisphère Nord, permet de représenter les erreurs sur les conditions initiales. Ces vecteurs singuliers ont un temps caractéristique de 18h sur le domaine Euro-Atlantique ainsi que sur les tropiques alors qu'il est de 24h sur les autres domaines. Ces vecteurs utilisent une norme en énergie cinétique pour les tropiques et totale sèche pour les moyennes latitudes.

L'erreur de modélisation est représentée par la technique multi-physiques utilisant dix physiques différentes. Celles-ci se basent sur deux schémas de convection profonde différents : B85 (Bougeault, 1985) et PCMT (Piriou et al., 2007). Elles comprennent aussi deux versions du schéma de turbulence de Cuxart et al. (2000) : avec et sans prise en compte de l'advection. Trois schémas de convection peu profonde sont proposés : le schéma Kain and Fritsch (1993) amélioré par Bechtold et al. (2001), un schéma Eddy-Diffusivity/Mass-Flux basé sur l'approche de Kain and Fritsch (1993) et le schéma de Pergaud et al. (2009). Une modification du calcul des ondes orographiques dans Catry et al. (2008) intervient pour certains membres. Enfin, pour modéliser la rayonnement solaire, les schémas de Fouquart and Bonnel (1980) et de Mlawer et al. (1997) avec McICA1 et 2 (Monte Carlo Independent Column Approximation) alors que le rayonnement thermique utilise le schéma de Mlawer et al. (1997). En revanche, les paramétrisations des flux océaniques, de surface, de longueur de mélange et de microphysiques (Belamari, 2005; Le Moigne et al., 2009; Bougeault and Lacarrere, 1989; Lopez, 2002) restent inchangés pour l'ensemble des membres de la PEARP.

Le choix des 12 membres PEARP coupleurs est effectué avec une méthode de clustering (comme décrit en 2.4) utilisant le critère de Ward pénalisé. La métrique permettant de mesurer la distance entre les membres de PEARP est fonction du vent horizontal, géopotentiel et humidité relative à 500, 700 et 850hPa, à 12, 18, 24, 30, 36 et 42h d'échéance (Bouttier and Raynaud, 2018).

### 3.2.4 Erreurs de modélisation

Concernant les erreurs de modélisation, celles-ci sont représentées avec le schéma SPPT. Celui-ci applique des perturbations aléatoires multiplicatives aux tendances totales de vent, température et contenu en vapeur d'eau en sortie des paramétrisations physiques. Tout comme la perturbation des paramètres de surface, les perturbations de la tendance totale sont multiplicatives, de types gaussiennes et suivent la formule suivante :

$$D(x, y, z) \rightarrow D(x, y, z) \times (1 + \mu(z) \times r(x, y)) \quad (3.2)$$

où  $D(x, y, z)$  est la tendance totale,  $\mu(z)$  un profil vertical de poids, dont les valeurs vont de 1 en milieu de l'atmosphère et décroît linéairement jusqu'à 0 à la surface ainsi qu'au sommet de l'atmosphère, et  $r(x, y)$  une perturbation bidimensionnelle gaussienne d'écart-type  $\sigma$ . Cette dernière présente une corrélation spatiale horizontale homogène et isotrope avec une longueur caractéristique de l'ordre de 500km. Cette perturbation évolue dans l'espace spectral avec un temps caractéristique d'environ 6h et est bornée entre  $[-2\sigma, 2\sigma]$ .

### 3.3 Objectif de la thèse : perturbation des paramètres d'entrée des paramétrisations

#### 3.3.1 Limite de la méthode SPPT

Malgré l'efficacité de la méthode SPPT pour représenter l'erreur de modélisation dans la PEARO, celle-ci comporte quelques inconvénients.

En reposant sur l'hypothèse forte que les erreurs de modélisation des tendances des variables d'état par les paramétrisations physiques sont proportionnelles aux tendances même, le SPPT présente le désavantage de n'appliquer aucune perturbation lorsque les tendances totales sont nulles. C'est pourquoi, la méthode ne peut que modifier l'intensité des phénomènes et non leur déclenchement. Ceci est particulièrement problématique pour des phénomènes à seuil tels que la convection. Quelques alternatives ont été proposées pour résoudre ce problème. Sachant que le SPPT perturbe l'intensité de la convection mais pas son déclenchement, le schéma Stochastic Trigger of Convection (STC) propose d'ajouter un terme stochastique dans la fonction de déclenchement de la convection en perturbant le changement de température d'une parcelle d'air (Du et al., 2018). Une autre possibilité est la méthode Stochastic boundary-layer humidity (SHUM) se basant sur le fait que le déclenchement de la convection dépend de l'humidité présente dans les basses couches. Elle perturbe ainsi l'humidité spécifique dans la couche limite atmosphérique à chaque pas de temps, de la même façon que le SPPT perturbe la tendance d'humidité spécifique. Cette technique améliore particulièrement les ensembles dans les tropiques mais reste sans effets pendant les périodes hivernales dans les extra-tropiques (Whitaker et al., 2013).

Par ailleurs, en perturbant les tendances totales de température, vent et humidité issues des différentes paramétrisations, il est impossible d'associer l'incertitude du modèle à un phénomène physique particulier. Les perturbations associées à ce schéma ne sont donc pas interprétables physiquement. En outre, l'application du schéma SPPT engendre des incohérences. Par exemple, les flux de surface sont calculés à partir des tendances non perturbées, ce qui engendre une non conservation de l'énergie. Les méthodes pSPPT et iSPPT permettent de résoudre en partie ce problème.

D'un point de vue plus technique, la calibration des motifs de perturbations, leur amplitudes et corrélations spatio-temporelles, demande de nombreux essais pour être optimisée et est donc très coûteuse numériquement. Par ailleurs, pour des raisons de stabilité numérique, le SPPT n'est pas appliqué en basses couches, ni au sommet de l'atmosphère, et ne représente donc pas l'incertitude de modélisation sur l'intégralité de la colonne atmosphérique.

#### 3.3.2 Test des méthodes de perturbations de paramètres dans la PEARO

L'objectif principal de la thèse est alors de développer, régler et évaluer dans la Prévision d'Ensemble AROME (PEARO) une autre représentation des erreurs de modélisation basée cette fois-ci sur la perturbation de paramètres en entrée des paramétrisations physiques.

Actuellement, les méthodes perturbant de tels paramètres reposent uniquement sur les connaissances des experts en paramétrisations physiques. Ceux-ci fournissent des listes de paramètres dont la valeur est mal connue et qui sont donc susceptibles d'être représentatifs des erreurs de modélisation. Les experts fournissent aussi les gammes de valeurs des paramètres ainsi que les distributions de probabilité correspondantes en se basant sur des données empiriques. Par exemple, différentes campagnes de mesures proposent des valeurs disparates d'un paramètre, fournissant ainsi une gamme de valeurs dans laquelle perturber le dit paramètre.

Cependant, malgré l'incertitude sur leur valeur, tous les paramètres fournis par les experts n'engendrent pas tous une incertitude notable sur les prévisions. Afin d'éviter une perturbation inutile engendrant un coût de maintien des schémas de perturbations de paramètres non nécessaires, il peut être plus judicieux d'effectuer tout d'abord une pré-sélection des paramètres ayant une influence sur les sorties du modèle et pouvant par conséquent représenter la majeure partie de l'erreur de modélisation. Pour ce faire, des méthodes d'analyse de sensibilité permettant d'identifier et de quantifier l'influence des paramètres d'entrée d'un modèle sur ses sorties peuvent être appliquées aux modèles météorologiques. Dans cette optique, une application de ces méthodes au modèle atmosphérique AROME a été réalisée afin d'identifier et d'étudier plus en détail l'influence des paramètres sur des variables de temps sensibles. Cette étude fait l'objet de la partie II.

Dans une deuxième étape, différentes techniques de perturbation des paramètres ont été testées dans la PEARO et comparées au SPPT actuellement implémenté dans la prévision opérationnelle. Une attention particulière est portée à l'optimisation de ces méthodes. Ces travaux font l'objet de la partie III.

## Bibliographie

- 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 Sc clouds in meteorological models. *Journal of the Atmospheric Sciences*, 52(4) :455 – 463, 1995. doi : 10.1175/1520-0469(1995)052<0455:MOTWCW>2.0.CO;2.
- P. Bechtold, E. Bazile, F. Guichard, P. Mascart, and E. Richard. A mass flux convection scheme for regional and global models. *Quarterly Journal of the Royal Meteorological Society*, 127 :869–886, 2001. doi : 10.1002/qj.49712757309.
- S. Belamari. Report on uncertainty estimates of an optimal bulk formulation for surface turbulent fluxes,. Technical report, MERSEA IP Deliverable 412, 2005.
- P. Bénard, J. Vivoda, J. Mašek, P. Smolikova, K. Yessad, C. Smith, R. Brozkova, and J.-F. Geleyn. Dynamical kernel of the Aladin-NH spectral limited-area model : Revised formulation and sensitivity experiments. *Quarterly Journal of the Royal Meteorological Society*, 136 :155 – 169, 2010. doi : 10.1002/qj.522.
- P. Bougeault. Cloud-ensemble relations based on the gamma probability distribution for the higher-order models of the planetary boundary layer. *Journal of the*

- Atmospheric Sciences*, 39(12) :2691 – 2700, 1982. doi : 10.1175/1520-0469(1982)039<2691:CERBOT>2.0.CO;2.
- P. Bougeault. A simple parameterization of the large-scale effects of cumulus convection. *Monthly Weather Review*, 113 :2105–2121, 1985.
- P. Bougeault and P. Lacarrere. Parameterization of orography-induced turbulence in a mesobeta-scale model. *Monthly Weather Review*, 117(8) :1872 – 1890, 1989. doi : 10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.
- Y. Bouteloup, Y. Seity, and E. Bazile. Description of the sedimentation scheme used operationally in all Météo-France NWP models. *Tellus A*, 63 :300 – 311, 2011. doi : 10.1111/j.1600-0870.2010.00484.x.
- F. Bouttier and L. Raynaud. Clustering and selection of boundary conditions for limited-area ensemble prediction. *Quarterly Journal of the Royal Meteorological Society*, 144(717) :2381–2391, 2018. doi : 10.1002/qj.3304.
- F. Bouttier, B. Vié, O. Nuissier, and L. Raynaud. Impact of stochastic physics in a convection-permitting ensemble. *Monthly Weather Review*, 140(11) :3706 – 3721, 2012. doi : 10.1175/MWR-D-12-00031.1.
- F. Bouttier, L. Raynaud, O. Nuissier, and B. Ménétrier. Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. *Quarterly Journal of the Royal Meteorological Society*, 142(S1) :390–403, 2016. doi : 10.1002/qj.2622.
- P. Brousseau, L. Berre, F. Bouttier, and G. Desroziers. Background-error covariances for a convective-scale data-assimilation system : AROME–France 3D-Var. *Quarterly Journal of the Royal Meteorological Society*, 137(655) :409–422, 2011. doi : 10.1002/qj.750.
- P. Brousseau, Y. Seity, D. Ricard, and J. Léger. Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, 142(699) :2231–2243, 2016. doi : 10.1002/qj.2822.
- R. Bubnová, G. Hello, P. Bénard, and J.-F. Geleyn. Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Monthly Weather Review*, 123(2) :515 – 535, 1995. doi : 10.1175/1520-0493(1995)123<0515:IOTFEE>2.0.CO;2.
- B. Catry, J.-F. Geleyn, F. Bouyssel, J. Cedilnik, R. Brozkova, M. Derková, and Richard. A new sub-grid scale lift formulation in a mountain drag parameterisation scheme. *Meteorologische Zeitschrift*, 17 :193–208, 2008. doi : 10.1127/0941-2948/2008/0272.
- H. Charnock. Wind stress on a water surface. *Quarterly Journal of the Royal Meteorological Society*, 81(350) :639–640, 1955. doi : 10.1002/qj.49708135027.

- P. Courtier, C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas. The Arpege project at Météo France. In *Seminar on Numerical Methods in Atmospheric Models, 9-13 September 1991*, volume II, pages 193–232, Shinfield Park, Reading, 1991. ECMWF, ECMWF. URL <https://www.ecmwf.int/node/8798>.
- J. Cuxart, P. Bougeault, and J. L. Redelsperger. A turbulence scheme allowing for mesoscale and large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, 126 :1–30, 2000. doi : 10.1002/qj.49712656202.
- M. Déqué, C. Dreveton, A. Braun, and D. Cariolle. The ARPEGE/IFS atmosphere model : a contribution to the french community climate modelling. *Climate Dynamic*, 10 :249–266, 1994. doi : 10.1007/BF00208992.
- L. Descamps, C. Labadie, A. Joly, E. Bazile, P. Arbogast, and P. Cébron. PEARP, the Météo-France short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 141(690) :1671–1685, 2015. doi : 10.1002/qj.2469.
- J. Du, J. Berner, R. Buizza, M. Charron, P. Houtekamer, D. Hou, I. Jankov, M. Mu, X. Wang, M. Wei, and H. Yuan. *Ensemble Methods for Meteorological Predictions*, pages 1–52. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN 978-3-642-40457-3. doi : 10.1007/978-3-642-40457-3\_13-1.
- Y. Fouquart and B. Bonnel. Computations of solar heating of the earth’s atmosphere—a new parameterization. *Beiträge zur Physik der Atmosphäre*, 53 :35–62, 1980.
- J. Kain and J. Fritsch. *Convective Parameterization for Mesoscale Models : The Kain-Fritsch Scheme*, volume The Representation of Cumulus Convection in Numerical Models of *Meteorological Monographs*, chapter 16, pages 165–170. American Meteorological Society, Boston, MA, 1993. doi : 10.1007/978-1-935704-13-3\\_\_16.
- J. S. Kain and J. M. Fritsch. A one-dimensional entraining/detraining plume model and its application in convective parameterization. *Journal of the Atmospheric Sciences*, 47(23) :2784 – 2802, 1990. doi : 10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2.
- J. P. Lafore, J. Stein, N. Asencio, P. Bougeault, V. Ducrocq, J. Duron, C. Fischer, P. Hérel, P. Mascart, V. Masson, J. P. Pinty, J. L. Redelsperger, E. Richard, and J. Vilà-Guerau de Arellano. The Meso-NH atmospheric simulation system. part I : adiabatic formulation and control simulations. *Advances in Geosciences*, 16(1) :90–109, 1998. doi : 10.1007/s00585-997-0090-6.
- P. Le Moigne, A. Boone, J.-C. Calvet, B. Decharme, S. Faroux, A.-L. Gibelin, C. Lebeaupin, J.-F. Mahfouf, E. Martin, V. Masson, D. Mironov, J. Noilhan, P. Tulet, and B. Van Den Hurk. Surfex scientific documentation. Technical report, Météo-France, 2009. Note de centre du Groupe de Météorologie à Moyenne Echelle, available at : <http://www.cnrm.meteo.fr/surfex/>.
- D. Li and K. Shine. A 4-dimensional ozone climatology for ugamp models. Technical report, 1995. UGAMP Internal Report No. 35.

- P. Lopez. Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Quarterly Journal of the Royal Meteorological Society*, 128(579) :229–257, 2002. doi : 10.1256/00359000260498879.
- G. M. Martin, D. W. Johnson, and A. Spice. The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *Journal of the Atmospheric Sciences*, 51(13) :1823 – 1842, 1994. doi : 10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2.
- V. Masson. A physically-based scheme for the urban energy budget in atmospheric models. *Boundary-Layer Meteorology*, 94(3) :357–397, 2000. doi : 10.1023/a:1002463829265.
- V. Masson, J.-L. Champeaux, F. Chauvin, C. Meriguet, and R. Lacaze. A global database of land surface parameters at 1-km resolution in meteorological and climate models. *Journal of Climate*, 16(9) :1261 – 1282, 2003. doi : 10.1175/1520-0442(2003)16<1261:AGDOLS>2.0.CO;2.
- 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 : Atmosphere*, 102 (D14) :16663–16682, 1997. doi : 10.1029/97JD00237.
- K. P. Nielsen, E. Gleeson, and L. Rontu. Radiation sensitivity tests of the harmonie 37h1 nwp model. *Geoscientific Model Development*, 7(4) :1433–1449, 2014. doi : 10.5194/gmd-7-1433-2014. URL <https://gmd.copernicus.org/articles/7/1433/2014/>.
- J. Noilhan and S. Planton. A simple parameterization of land surface processes for meteorological models. *Monthly Weather Review*, 117(3) :536 – 549, 1989. doi : 10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2.
- S. Ou and K. Liou. Ice microphysics and climatic temperature feedback. *Atmospheric Research*, 35(2) :127–138, 1995. doi : 10.1016/0169-8095(94)00014-5.
- J. Pergaud, V. Masson, S. Malardel, and F. Couvreux. A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction. *Boundary-Layer Meteorology*, 132(1) :83, 2009. doi : 10.1007/s10546-009-9388-0.
- J. Pinty and P. Jabouille. A mixed-phased cloud parameterization for use in a mesoscale non-hydrostatic model : simulations of a squall line and of orographic precipitation. In A. M. Society, editor, *Preprints of Conference On Cloud Physics*, pages 217–220, Everett, WA, 1998.
- 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 :4127–4139, 2007. doi : 10.1175/2007JAS2144.1.

- G. Radnóti. Comments on : A spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations. *Monthly Weather Review*, 123(10) :3122 – 3123, 1995. doi : 10.1175/1520-0493(1995)123<3122:COSLAF>2.0.CO;2.
- Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The AROME-France convective-scale operational model. *Monthly Weather Review*, 139(3) :976–991, 2011. doi : 10.1175/2010MWR3425.1.
- P. Siebesma and J. Teixeira. An advection-diffusion scheme for the convective boundary layer, description and 1d results. In *Proceedings of 14th symposium on boundary layers and turbulence*. Aspen, USA, 2000.
- I. Tegen, P. Hollrig, M. Chin, I. Fung, D. Jacob, and J. Penner. Contribution of different aerosol species to the global aerosol extinction optical thickness : Estimates from model results. *Journal of Geophysical Research : Atmosphere*, 102(D20) :23895–23915, 1997. doi : 10.1029/97JD01864.
- F. Váňa, P. Bénard, J.-F. Geleyn, A. Simon, and Y. Seity. Semi-lagrangian advection scheme with controlled damping : An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society*, 134(631) :523–537, 2008. doi : 10.1002/qj.220.
- J. Whitaker, P. Pegion, and T. Hamill. Representing model uncertainty in data assimilation (using ensembles). *EMC/NCEP/NOAA seminar*, 2013. <http://www.emc.ncep.noaa.gov/seminars/index.html>.



## Deuxième partie

Analyse de sensibilité du modèle  
AROME aux paramètres  
incertains des paramétrisations  
physiques et dynamique



# Chapitre 4

## Des méthodes d'analyse de sensibilité locales aux méthodes globales

Afin d'identifier les paramètres des paramétrisations physiques d'AROME susceptibles de représenter la majeure partie de l'erreur de modélisation, nous avons cherché à qualifier et quantifier l'influence de ces paramètres sur les prévisions AROME grâce à l'utilisation de méthodes d'analyse de sensibilité. Leur principe et les éléments nécessaires à leur mise en place (plan d'expérience et définition de la mesure d'influence) seront présentés dans le présent chapitre. Nous allons en particulier nous attarder sur les méthodes qui seront appliquées au modèle AROME. Cette application fera l'objet du chapitre 5 alors que les résultats seront présentés au chapitre 6.

### 4.1 Principe des analyses de sensibilité

Par définition, les modèles de simulation numérique de systèmes physiques cherchent à reproduire les systèmes réels en se basant sur des équations mathématiques discrétisées les régissant. Ces équations relient alors les données d'entrées aux données de sorties du modèle, faisant ainsi intervenir de nombreux paramètres. Cependant, ces modèles de simulation numérique présentent des incertitudes dont il peut être intéressant de connaître l'origine et les impacts. L'analyse de sensibilité entre alors dans le processus d'élaboration des modèles numériques, en permettant de diagnostiquer les sources d'incertitudes présentes parmi les paramètres d'entrées des modèles (Faivre et al., 2013). Elle permet de hiérarchiser, parfois même de quantifier, ces sources et de mieux appréhender leur influence sur les sorties des modèles numériques (Saltelli et al., 2000). Ces informations sont particulièrement utiles lorsque l'on souhaite savoir quel paramètre doit être connu avec précision pour réduire l'incertitude. En d'autres termes, ces méthodes sont utiles à la fois pour estimer les sources d'incertitude d'un modèle, mais aussi pour optimiser celui-ci.

Deux grands types d'analyse de sensibilité ont été mises au point. Les méthodes

locales, se basant sur le calcul des dérivées de la sortie  $y$  du modèle par rapport à un paramètre d'entrée  $X_i$ , s'intéressent aux influences de petites variations des paramètres autour d'une valeur de référence. Elles sont surtout utilisées pour étudier le comportement d'un modèle autour de valeurs bien précises. A l'inverse, les méthodes globales s'intéressent aux effets des paramètres sur la variabilité de la sortie  $y$  du modèle lorsque ces paramètres varient sur leur domaine d'incertitude entier. De ce fait, seules ces méthodes globales sont considérées ici.

Afin de mettre en place ces analyses de sensibilité globales, plusieurs étapes sont requises. Il faut tout d'abord définir les gammes des valeurs possibles des  $k$  paramètres du modèle, ainsi que leurs distributions de probabilité. La seconde étape doit déterminer les différentes valeurs de paramètres à tester. Pour cela, un *plan d'expérience* doit être choisi selon le type de modèle et de paramètres. Pour chacune des valeurs de paramètres produites par le plan d'expérience, il faut calculer les sorties du modèle correspondantes. Enfin, à partir de l'ensemble des données de sorties, on calcule les *indices de sensibilité*  $I_i$  de chaque paramètre  $X_i$ . Les étapes sont ainsi les suivantes :

1. définition des intervalles de perturbations et distributions de probabilité des paramètres
2. définition du plan d'expérience
3. calcul des sorties du modèle correspondant aux valeurs de paramètre définies par le plan d'expérience
4. calcul des indices de sensibilité

La première étape relevant du domaine des experts des systèmes modélisés, nous nous intéressons ici, uniquement aux trois dernières étapes.

## 4.2 Etape 2 : plan d'expérience

Un plan d'expérience définit l'ensemble des valeurs des paramètres d'entrée du modèle prises pour les différentes simulations nécessaires à l'estimation de l'influence de ces paramètres. Il s'agit donc d'identifier les différents points d'échantillons de l'espace des valeurs de paramètres. Plusieurs plans sont possibles et chacun d'entre eux a une utilisation et un coût différents. Ici, il ne s'agit pas de donner une liste exhaustive de l'ensemble des plans d'expérience possibles. Nous n'en présentons que les plus usuels.

Un des premiers plans d'expérience possible est tout simplement un plan généré par une méthode de Monte-Carlo. Il suffit alors de tirer aléatoirement la valeur de chaque paramètre pour définir une simulation. En répétant ce processus, on définit alors un plan d'expérience permettant d'obtenir différents échantillons de l'espace des paramètres d'entrée du modèle. Cependant, pour estimer l'influence des paramètres sur l'ensemble de leur domaine de perturbation, et ce, de manière suffisamment précise, ce plan d'expérience peut devenir très vite coûteux. Il existe donc des plans maximisant la dispersion des échantillons de l'espace des paramètres d'entrée.

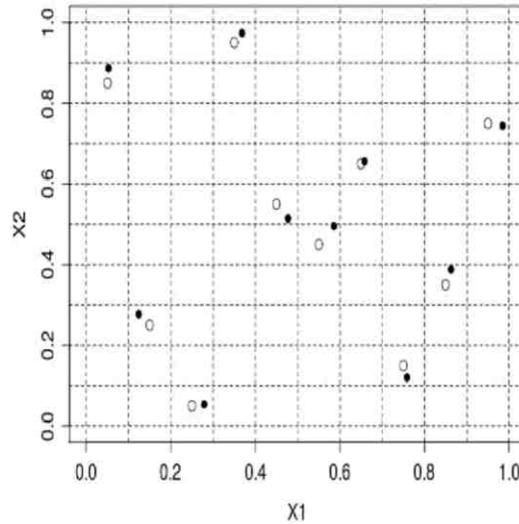


FIGURE 4.1 – Exemple de plan d'expérience comprenant 10 échantillons des paramètres  $X_1$  et  $X_2$  par hypercube latin (noir) et hypercube latin centré (blanc) (Adapté de Faivre et al. (2013))

Par exemple, l'échantillonnage par hypercubes latin (dit Latin Hypercube Sampling, LHS) propose un plan d'expérience permettant de répartir plus uniformément les points d'échantillon sur l'espace de perturbation des paramètres (McKay et al., 1979). Plus précisément, chaque domaine de perturbation de paramètre est découpé en  $n$  segments de probabilité  $1/n$ . Dans chacun de ces segments, on tire aléatoirement une valeur de paramètre. Une fois les  $n$  valeurs de chaque paramètre obtenues pour chacun des segments, on réalise une combinaison de ces valeurs de manière aléatoire. Ainsi, on obtient un échantillon de  $k \times n$  points dans l'espace des paramètres où  $k$  est le nombre de paramètres. Notons que l'échantillonnage par hypercube latin centré utilise la valeur au centre de chacun des  $n$  segments. Sur la figure 4.1, on présente un exemple d'échantillonnage par hypercube latin (noir) et celui centré (blanc). Ces plans permettent d'échantillonner l'ensemble du domaine de perturbation des paramètres sans sur-échantillonner certaines zones contrairement à la méthode de Monte-Carlo. Par contre, il ne prend pas en compte la répartition conjointe des différents paramètres.

Pour prendre en compte les distributions conjointes des paramètres, une solution est de tester toutes les combinaisons de valeurs de paramètres possibles. Ce plan dit factoriel complet demande alors  $n^k$  simulations, pour  $k$  paramètres avec  $n$  valeurs différentes testées et est l'un des plus coûteux (Faivre et al., 2013). Les plans factoriels fractionnaires sont, en revanche, plus raisonnables et ne proposent qu'une partie des combinaisons d'un plan factoriel complet. Cependant, ce type de plan suppose que certaines interactions entre paramètres sont inexistantes. Par exemple, un plan de force 2 vérifie que chaque couple de valeurs de paramètres n'est représenté qu'une fois.

D'autres méthodes d'analyse de sensibilité, dites de criblage, cherchent uniquement à identifier les paramètres les plus influents, sans quantifier leur influence. Elles s'adressent principalement aux modèles dont on cherche à connaître la sensibilité pour un très grand nombre de paramètres dont les interactions sont négligées. Elles supposent alors que peu de paramètres sont influents.

Par exemple, parmi les méthodes de criblage, on compte la bifurcation séquentielle qui définit deux groupes de paramètres et élimine le moins influent. En réitérant ce processus, cette méthode converge vers le groupe de paramètres le plus influent. Cependant, ceci demande d’avoir une idée des paramètres influents lors de la définition des deux premiers groupes de paramètres. De façon similaire, le criblage par groupe perturbe les paramètres réunis par groupe et identifie les paramètres influents comme ceux appartenant au groupe ayant le plus fort impact sur les sorties du modèle. Dans un plan d’expérience séquentiel, le nombre de simulations n’est pas connu à l’avance : de nouvelles expériences sont définies en fonction des résultats précédents. Enfin, un des plans d’expérience les plus simples est le plan dit « One at a Time » (OAT). Celui-ci consiste à perturber un paramètre après l’autre, d’une première valeur à une seconde. Contrairement aux autres méthodes de criblage qui peuvent demander un très grand nombre de simulations, le plan OAT nécessite seulement  $k + 1$  simulations mais a l’inconvénient d’estimer les effets des paramètres avec peu de précision. Le plan d’expérience de Morris (1991), composé de plusieurs plans OAT, permet en revanche de rendre la méthode OAT plus robuste. Cette méthode fera plus particulièrement l’objet de la section 4.4 puisqu’elle définit à la fois un plan d’expérience et des indices de sensibilité qui lui sont propres.

### 4.3 Etape 4 : Mesure de l’influence et utilisation des indices de Sobol’

Après avoir choisi le plan d’expérience utilisé, les sorties correspondants aux valeurs de paramètres du plan d’expérience sont alors simulées dans l’étape 3 et une mesure de l’influence des paramètres peut être effectuée.

Si le modèle est linéaire, différents coefficients de corrélation ou de régression peuvent être utilisés. A titre d’exemple, on définit les coefficients de corrélation linéaire dits de Pearson (Wilks, 2011) comme :

$$I_i = \frac{Cov(X_i, y)}{\sqrt{V(X_i)V(y)}} \quad (4.1)$$

où  $V$  est la variance et  $Cov$  la covariance. Ils permettent de mesurer la relation linéaire entre un paramètre  $X_i$  et la sortie  $y$  du modèle. Si le modèle est bien linéaire, plus le coefficient de Pearson est élevé, plus le paramètre sera influent.

La plupart des modèles n’étant pas linéaires, les méthodes se basant sur la variance, dites ANOVA (« ANalysis Of VAriance »), sont alors utilisées. En se basant sur le théorème de décomposition d’Hoeffding, la variance de la sortie du modèle  $V(y) = V$  peut être décomposée en la somme suivante :

$$V = \sum_{i=1}^k V_i + \sum_{1 \leq i < j \leq k} V_{ij} + \dots + V_{1\dots k} \quad (4.2)$$

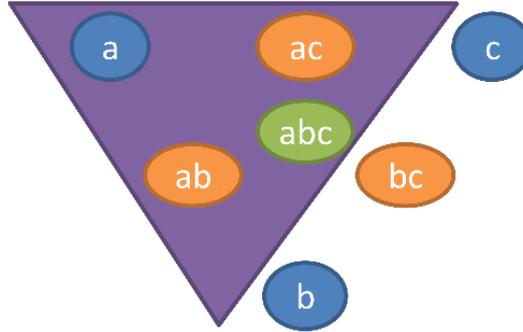


FIGURE 4.2 – Effets compris dans les différents indices de Sobol'. Exemple pour 3 paramètres,  $a$ ,  $b$  et  $c$ . Les indices de premier ordre sont symbolisés en bleu, ceux de second ordre en orange et de troisième ordre en vert. L'indice de Sobol' total du paramètre  $a$  comprend l'ensemble des effets inclus dans le triangle violet.

où

$$V_i = \mathbb{V}(\mathbb{E}[Y|X_i])$$

$$V_{ij} = \mathbb{V}(\mathbb{E}[Y|X_i, X_j]) - V_i - V_j$$

A partir de cette décomposition, Sobol' (1990) définit alors les indices de sensibilité comme la proportion de la variance de la sortie du modèle due à un paramètre ou un groupe de paramètres. Les indices de Sobol' d'ordre 1 sont alors, pour un paramètre  $X_i$  :

$$S_i = \frac{V_i}{V}, \quad (4.3)$$

Là où les indices d'ordre 1 fournissent une estimation de l'influence directe des paramètres, les indices d'ordres supérieurs permettent d'estimer les interactions entre paramètres. En particulier, les indices d'ordre 2 donnent l'influence du couple de paramètres  $(X_i, X_j)$  :

$$S_{i,j} = \frac{V_{ij}}{V}. \quad (4.4)$$

Pour un grand nombre de paramètres, il devient complexe d'interpréter l'ensemble des indices de Sobol'. La somme de l'ensemble des indices de Sobol' valant 1, Homma and Saltelli (1996) préconisent alors d'utiliser l'indice de Sobol' total, prenant en compte toutes les contributions d'un paramètre à la variance de la sortie du modèle :

$$I_i = \sum_{j \# i} S_j = 1 - \frac{\mathbb{V}(\mathbb{E}[Y|X_{\sim i}])}{V}, \quad (4.5)$$

où  $\#i$  représente tous les ensembles d'indices contenant l'indice  $i$ , et  $X_{\sim i}$  correspond à tous les paramètres sauf le  $i^{\text{ème}}$ . Ces indices représentent alors l'effet total d'un paramètre (représenté en violet dans la figure 4.2), c'est-à-dire à la fois son effet principal (indice de Sobol' d'ordre 1, schématisé en bleu) et ses interactions avec d'autres paramètres (indices d'ordre supérieur, représenté en orange pour les ordres 2 et vert pour ordre 3).

Afin d'estimer les indices de Sobol', il est nécessaire de calculer précisément

la variance de la sortie du modèle, connaissant les différents paramètres  $X_i$ . Une première possibilité est d'utiliser un plan d'expérience de type Monte-Carlo. Mais ceci nécessite pas moins de 1000 simulations pour chaque paramètre (Jacques, 2011). Par ailleurs, pour obtenir une précision de l'indice de Sobol' de 10%, près de 10 000 simulations par paramètre sont nécessaires (Iooss, 2011). Des plans d'expériences tels que l'hypercube latin peuvent réduire ce nombre de simulations.

Une solution pour réduire le nombre de simulations nécessaires pour l'utilisation de la méthode de Sobol' est la méthode FAST (Fourier Amplitude Sensitivity Test) qui utilise une transformation de Fourier pour exprimer la sortie du modèle dans l'espace spectral. Les indices de Sobol' sont alors calculés analytiquement à partir des harmoniques de la série de Fourier.

Une alternative pour réduire davantage le coût de calcul de la méthode de Sobol' est l'utilisation de méta-modèles, également appelés émulateurs ou encore surfaces de réponse (Faivre et al., 2013; Saltelli et al., 2000). Ces modèles statistiques sont des approximations mathématiques du modèle originel et permettent de générer des sorties émulées du modèle avec un très faible coût de calcul. Différentes fonctions mathématiques peuvent être utilisés pour approximer un modèle originel, comme les processus gaussiens ou encore un polynôme. Il existe une très grande variété d'émulateurs, nous avons choisi ici de ne présenter que l'expansion du polynôme du chaos puisqu'il sera utilisé dans la suite de ce travail (PCE; Sudret, 2008; Le Gratiet et al., 2016).

Considérons un modèle de simulation numérique  $M(x) = y$ . Celui-ci est projeté sur une base de polynômes orthogonaux et approximé par une combinaison linéaire :

$$M(x) = \sum_{\alpha \in \mathbb{A}} b_{\alpha} \Psi_{\alpha}(x) + \epsilon = \mathbf{B}^T \Psi(\mathbf{X}) + \epsilon(\mathbf{X}) \quad (4.6)$$

où  $\mathbb{A} \subset \mathbb{N}^k$ ,  $\Psi_{\alpha}$  sont des polynômes orthogonaux multivariés,  $b_{\alpha}$  sont les coefficients à déterminer afin que le métamodèle s'adapte aux données réelles du modèle,  $\mathbf{B} = \{b_0, \dots, b_{P-1}\}$ ,  $\Psi = \{\Psi_0, \dots, \Psi_{P-1}(x)\}$ . L'ensemble des coefficients  $b_{\alpha}$  ne pouvant être déterminés, une troncature doit être imposée pour délimiter le nombre de coefficients à estimer ( $P$ ). Ceux-ci sont ensuite déterminés en minimisant l'erreur moyenne quadratique résiduelle entre le méta-modèle et des données originelles du modèle. Ces dernières, appelées données d'apprentissage, sont issues d'un plan d'expérience comprenant un échantillon des sorties du modèle et servent à ajuster le méta-modèle au vrai modèle.

Là où la plupart des méta-modèles sont utilisés pour générer rapidement les milliers de simulations nécessaires aux calculs de la variance et donc des indices de Sobol', l'expansion du chaos polynomial a la particularité de fournir les indices de Sobol' directement à partir des coefficients du chaos polynomial :

$$S_u = \frac{\sum_{\alpha \in A_u} b_{\alpha}^2}{\sum_{\alpha \in A \setminus 0} b_{\alpha}^2} \quad (4.7)$$

pour un groupe de paramètres  $u$  avec  $A_u = \{\alpha \in \mathbb{A} : i \in u \Leftrightarrow \alpha_i \neq 0\}$ .

Les indices de Sobol', très populaires, ont été utilisés à de maintes reprises, dans de nombreux domaines, en particulier en sciences environnementales, tels que l'hydrologie (par ex : Nossent et al., 2011; Sarrazin et al., 2016), la climatologie

(Shi et al., 2019), la chimie atmosphérique (Girard et al., 2016) ou l'observation atmosphérique (Rieutord, 2017). Cette approche présentant l'avantage d'apporter une compréhension riche de l'influence des paramètres, elle peut aussi convenir au modèle de prévision numérique du temps tel qu'AROME si elle est utilisée avec un émulateur en raison du coût du modèle de prévision du temps et d'une telle analyse de sensibilité.

## 4.4 Méthode de Morris

En parallèle de la méthode de Sobol', la méthode de Morris peut être utilisée pour qualifier uniquement (et non quantifier) l'influence des paramètres. Faisant partie des méthodes de criblage, la méthode de Morris (Morris, 1991) utilise un plan d'expérience « One-At-a-Time » (OAT). Ce plan d'expérience permettant d'estimer l'influence des paramètres avec peu de précision, celui-ci est réitéré  $r$  fois afin de rendre la méthode plus robuste et plus précise (réitération entre 10 et 50 par paramètres; Campolongo et al., 2007).

Plus précisément, considérons un modèle avec  $k$  paramètres incertains, notés  $X_i, i = 1 \dots, k$ . Ces paramètres, supposés continus et généralement, de probabilité uniforme sur leur domaine de perturbation, sont discrétisés en  $p$  niveaux. On définit alors un point initial dans l'espace discrétisé des  $k$  paramètres, définissant aléatoirement une première valeur de ces paramètres. Chaque paramètre est ensuite perturbé dans un ordre aléatoire, d'une valeur  $\Delta$  proportionnelle à  $\frac{1}{p-1}$ . Notons que la variation  $\Delta$  peut être positive ou négative. Une fois l'ensemble des paramètres perturbés, on obtient une trajectoire de Morris, c'est à dire un plan OAT, contenant  $k + 1$  simulations. Puisqu'une seule trajectoire de Morris n'échantillonne qu'une partie infime de l'espace des paramètres, la méthode échantillonne  $r$  trajectoires de Morris. Par exemple, la figure 4.3 présente un plan d'expérience de Morris pour  $k = 2$  paramètres ( $X_1$  et  $X_2$ ), chacun perturbé entre 0 et 1, et discrétisé en  $p = 5$  niveaux. Celui-ci comprend  $r = 5$  trajectoires de Morris utilisant un pas  $\Delta = \frac{1}{p-1}$  et initialisées aux points  $A, B, C, D$  et  $E$ , choisis aléatoirement.

Afin d'optimiser la répartition spatiale des trajectoires de Morris, Campolongo et al. (2007) suggèrent de maximiser la distance euclidienne des trajectoires. Pour cela, on calcule la somme de la distance euclidienne entre chaque point de  $M$  trajectoires. Seules les  $r$  trajectoires les plus éloignées sont considérées.

L'influence des paramètres peut ensuite être mesurée le long d'une trajectoire de Morris. On définit alors l'effet élémentaire  $EE_i$  du paramètre  $X_i$ , comme une mesure de la sensibilité du modèle à ce paramètre :

$$EE_i = \frac{f(X_1, X_2, \dots, X_i + \Delta, \dots, X_k) - f(X_1, X_2, \dots, X_i, \dots, X_k)}{\Delta}, \quad (4.8)$$

où  $f(X_1, X_2, \dots, X_k)$  est la sortie scalaire du modèle. Ces effets élémentaires sont alors calculés pour chacune des trajectoires de Morris, permettant ainsi d'obtenir un échantillon de la distribution des effets élémentaires de chacun des paramètres. Ainsi, plus le nombre de trajectoires de Morris est élevé, meilleure sera l'estimation de cette distribution. Campolongo et al. (2007) suggère d'utiliser entre 10 et 50 trajectoires, et Faivre et al. (2013) fixe un minimum à 5 trajectoires.

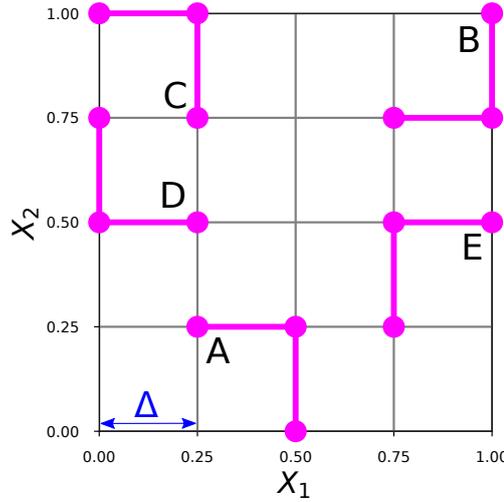


FIGURE 4.3 – Exemple de plan d'expérience de Morris pour  $k = 2$  paramètres ( $X_1$  et  $X_2$ ) avec les caractéristiques suivantes :  $r = 5$ ,  $p = 5$ ,  $\Delta = \frac{1}{p-1}$ . A, B, C, D et E sont les points initiaux des trajectoires de Morris

Pour décrire la distribution des effets élémentaires, Morris (1991) et Campolongo et al. (2007) proposent différents indices de sensibilité pour chaque paramètre  $X_i$  :

$$\mu_i = \mathbb{E}(EE_i), \quad \mu_i^* = \mathbb{E}(|EE_i|), \quad \sigma_i = \text{std}(EE_i),$$

où  $\mathbb{E}$  est l'espérance et  $\text{std}$  l'écart-type.  $\mu_i^*$  caractérise l'impact direct des paramètres sur les sorties du modèle. Cet indice est préféré à  $\mu$  qui a l'inconvénient de ne pas prendre en compte les effets élémentaires compensatoires entre différentes trajectoires. Plus  $\mu_i^*$  est élevé, plus le paramètre  $X_i$  aura une influence directe sur les sorties du modèle. A l'inverse,  $\sigma_i$  représente plutôt l'écart-type des effets élémentaires et donc plutôt l'importance des interactions entre paramètres et la non linéarité des effets élémentaires. Plus celui-ci est élevé, plus les interactions peuvent être fortes ou plus le modèle présente de fortes non linéarités. Ces deux aspects étant indissociables, ils représentent une des limites de la méthode de Morris.

Ainsi, cette méthode est principalement utilisée pour donner un premier aperçu de l'influence des paramètres et éliminer les paramètres les moins influents. C'est pourquoi cette méthode est plutôt classée dans les méthodes de criblage. Notons que des améliorations ont été proposées afin d'obtenir des indices de sensibilité d'ordre 2 (Cropp and Braddock, 2002; Campolongo et al., 2007).

Enfin, on peut définir un indice de sensibilité  $I_i$  du paramètre  $X_i$  représentant l'influence totale du paramètre. Puisque  $\mu_i^*$  et  $\sigma_i$  représentent tous deux des mesures complémentaires de l'influence des paramètres, nous définissons  $I_i$  comme  $\sqrt{\mu_i^{*2} + \sigma_i^2}$ , comme proposé par Ciric (2012).

Cette méthode d'analyse de sensibilité a été appliquée avec succès à divers modèles complexes, tels qu'un modèle de batterie (van Rensburg et al., 2019), de bilan énergétique de bâtiments (Sanchez et al., 2014; Branger et al., 2015; Yang et al., 2016), de localisation de sources sismiques (Franczyk, 2019), d'eaux usées (Ruano et al., 2011; Cosenza et al., 2013; Vanrolleghem et al., 2015), de barrage en remblai (Ren et al., 2019), de combat (Alam et al., 2004), d'agronomie

(Confalonieri et al., 2010a,b) ou encore de médecine (Wu et al., 2013; Link et al., 2018). En sciences de l'atmosphère, ce criblage a été utilisé pour des modèles de climat (Covey et al., 2013) ou de prévision numérique du temps (Di et al., 2015; Quan et al., 2016; Di et al., 2017). Elle présente donc un bon potentiel pour l'analyse de sensibilité du modèle AROME.

**Conclusion :** Ainsi, au vu des différentes méthodes d'analyse de sensibilité possibles, nous nous sommes focalisés sur deux d'entre elles, la méthode de Morris et les indices de Sobol'.

## Bibliographie

- F. M. Alam, K. R. McNaught, and T. J. Ringrose. Using morris' randomized oat design as a factor screening method for developing simulation metamodels. In *Proceedings of the 36th Conference on Winter Simulation*, WSC '04, page 949–957. Winter Simulation Conference, 2004. ISBN 0780387864.
- F. Branger, L.-G. Giraudet, C. Guivarch, and P. Quirion. Global sensitivity analysis of an energy–economy model of the residential building sector. *Environmental Modelling and Software*, 70 :45 – 54, 2015. ISSN 1364-8152. doi : 10.1016/j.envsoft.2015.03.021.
- F. Campolongo, J. Cariboni, and A. Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software*, 22 :1509–1518, January 2007. doi : doi:10.1016/j.envsoft.2006.10.004.
- C. Ciric. Extrapolation des données de biotests pour la protection des communautes par une approche couplant biotests de laboratoire, études sur mésocosme et modélisation écologique. diploma thesis, Universite Claude Bernard Lyon 1, 2012.
- R. Confalonieri, G. Bellocchi, S. Bregaglio, M. Donatelli, and M. Acutis. Comparison of sensitivity analysis techniques : A case study with the rice model warm. *Ecological Modelling*, 221 :1897–1906, 08 2010a. doi : 10.1016/j.ecolmodel.2010.04.021.
- R. Confalonieri, G. Bellocchi, S. Tarantola, M. Acutis, M. Donatelli, and G. Genovese. Sensitivity analysis of the rice model warm in europe : Exploring the effects of different locations, climates and methods of analysis on model sensitivity to crop parameters. *Environmental Modelling and Software*, 25 :479–488, 2010b. doi : 10.1016/j.envsoft.2009.10.005.
- A. Cosenza, G. Mannina, P. A. Vanrolleghem, and M. B. Neumann. Global sensitivity analysis in wastewater applications : A comprehensive comparison of different methods. *Environmental Modelling and Software*, 49 :40–52, 2013. doi : 10.1016/j.envsoft.2013.07.009.
- C. Covey, D. D. Lucas, J. Tannahill, X. Garaizar, and R. Klein. Efficient screening of climate model sensitivity to a large number of perturbed input parameters.

- Journal of Advances in Modeling Earth Systems*, 5(3) :598–610, 2013. doi : 10.1002/jame.20040.
- R. Cropp and R. Braddock. The new morris method : an efficient second-order screening method. *Reliability Engineering and System Safety*, 78 :77–83, 2002. doi : 10.1016/S0951-8320(02)00109-6.
- Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, J. Quan, J. Li, C. Miao, A. Ye, and C. Tong. Assessing wrf model parameter sensitivity : A case study with 5 day summer precipitation forecasting in the greater beijing area. *Geophys. Res. Lett.*, 42(2) :579–587, 2015. doi : 10.1002/2014GL061623.
- Z. Di, Q. Duan, W. Gong, A. Ye, and C. Miao. Parametric sensitivity analysis of precipitation and temperature based on multi-uncertainty quantification methods in the weather research and forecasting model. *Science China Earth Sciences*, 60 :876–898, 03 2017. doi : 10.1007/s11430-016-9021-6.
- R. Faivre, B. Iooss, S. Mahévas, D. Makowski, and H. Monod. *Analyse de sensibilité et exploration de modèles*. Collection Savoir-Faire. Editions Quae, 2013.
- A. Franczyk. Using the morris sensitivity analysis method to assess the importance of input variables on time-reversal imaging of seismic sources. *Acta Geophysica*, 67 :1525–1533, 2019. ISSN 1895-7455. doi : 10.1007/s11600-019-00356-5.
- S. Girard, V. Mallet, I. Korsakissok, and A. Mathieu. Emulation and sobol’ sensitivity analysis of an atmospheric dispersion model applied to the fukushima nuclear accident. *Journal of Geophysical Research : Atmospheres*, 121(7) :3484–3496, 2016. doi : 10.1002/2015JD023993.
- T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52(1) :1 – 17, 1996. ISSN 0951-8320. doi : 10.1016/0951-8320(96)00002-6.
- B. Iooss. Revue sur l’analyse de sensibilité globale de modèles numériques. *Journal de la Societe Française de Statistique*, 152(1) :1–23, 2011.
- J. Jacques. Pratique de l’analyse de sensibilité : comment évaluer l’impact des entrées aléatoires sur la sortie d’un modèle mathématique. *Pub. IRMA*, 71(III), 2011.
- L. Le Gratiet, S. Marelli, and B. Sudret. Metamodel-based sensitivity analysis : polynomial chaos expansions and Gaussian processes. In *Handbook of Uncertainty Quantification - Part III : Sensitivity analysis*. 2016.
- K. G. Link, M. T. Stobb, J. Di Paola, K. B. Neeves, A. L. Fogelson, S. S. Sindi, and K. Leiderman. A local and global sensitivity analysis of a mathematical model of coagulation and platelet deposition under flow. *PLoS ONE*, 13(7) : 1–38, 07 2018. doi : 10.1371/journal.pone.0200917.
- M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2) :239–245, 1979. doi : 10.2307/1268522.

- M. D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33 :161–174, May 1991. doi : 10.2307/1269043.
- J. Nossent, P. Elsen, and W. Bauwens. Sobol' sensitivity analysis of a complex environmental model. *Environmental Modelling and Software - ENVSOFT*, 26, 12 2011. doi : 10.1016/j.envsoft.2011.08.010.
- J. Quan, Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, A. Ye, and C. Miao. An evaluation of parametric sensitivities of different meteorological variables simulated by the wrf model. *Quart. J. Roy. Meteor. Soc.*, 142(700) :2925–2934, 2016. doi : 10.1002/qj.2885.
- J. Ren, W. Zhang, and J. Yang. Morris sensitivity analysis for hydrothermal coupling parameters of embankment dam : A case study. *Mathematical Problems in Engineering*, 2019, 2019. doi : 10.1155/2019/2196578.
- T. A. Rieutord. *Sensitivity analysis of a filtering algorithm for wind lidar measurements*. Theses, Institut National Polytechnique de Toulouse (INP Toulouse), Nov. 2017.
- M. V. Ruano, J. Ribes, J. Ferrer, and G. Sin. Application of the morris method for screening the influential parameters of fuzzy controllers applied to wastewater treatment plants. *Water Science and Technology*, 63(10) :2199–2206, 05 2011. doi : 10.2166/wst.2011.442.
- A. Saltelli, K. Chan, and E. Scott. *Sensitivity Analysis*, volume 134. John Wiley and Sons, 01 2000. ISBN 0471998923.
- D. G. Sanchez, B. Lacarrière, M. Musy, and B. Bourges. Application of sensitivity analysis in building energy simulations : Combining first- and second-order elementary effects methods. *Energy and Buildings*, 68 :741 – 750, 2014. ISSN 0378-7788. doi : 10.1016/j.enbuild.2012.08.048.
- F. Sarrazin, F. Pianosi, and T. Wagener. Global sensitivity analysis of environmental models : Convergence and validation. *Environmental Modelling and Software*, 79 :135 – 152, 2016. ISSN 1364-8152. doi : 10.1016/j.envsoft.2016.02.005.
- Y. Shi, W. Gong, Q. Duan, J. Charles, C. Xiao, and H. Wang. How parameter specification of an earth system model of intermediate complexity influences its climate simulations. *Progress in Earth and Planetary Science*, 6(1) :46–64, 2019. ISSN 2197-4284. doi : 10.1186/s40645-019-0294-x.
- I. M. Sobol'. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2(1) :112–118, 1990. doi : 10.1016/S0378-4754(00)00270-6. (In russian).
- B. Sudret. Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering and System Safety*, 93 :964–979, 07 2008. doi : 10.1016/j.res.2007.04.002.

- A. J. van Rensburg, G. van Schoor, and P. van Vuuren. Stepwise global sensitivity analysis of a physics-based battery model using the morris method and monte carlo experiments. *Journal of Energy Storage*, 25 :100875, 2019. ISSN 2352-152X. doi : 10.1016/j.est.2019.100875.
- P. Vanrolleghem, G. Mannina, A. Cosenza, and M. Neumann. Global sensitivity analysis for urban water quality modelling : Terminology, convergence and comparison of different methods. *Journal of Hydrology*, 522, 03 2015. doi : 10.1016/j.jhydrol.2014.12.056.
- D. Wilks. Chapter 3 - empirical distributions and exploratory data analysis. In D. S. Wilks, editor, *Statistical Methods in the Atmospheric Sciences*, volume 100 of *International Geophysics*, pages 23–70. Academic Press, 2011. doi : <https://doi.org/10.1016/B978-0-12-385022-5.00003-8>.
- J. Wu, R. Dhingra, M. Gambhir, and J. Remais. Sensitivity analysis of infectious disease models : Methods, advances and their application. *Journal of the Royal Society, Interface*, 10 :20121018, 06 2013. doi : 10.1098/rsif.2012.1018.
- S. Yang, W. Tian, E. Cubi, Q. Meng, Y. Liu, and L. Wei. Comparison of sensitivity analysis methods in building energy assessment. *Procedia Engineering*, 146 : 174–181, 12 2016. doi : 10.1016/j.proeng.2016.06.369.

# Chapitre 5

## Application des méthodes d'analyse de sensibilité globale au modèle AROME

Après avoir identifié, au chapitre précédent, les différentes méthodes d'analyse de sensibilité qui feront l'objet de ce travail, nous allons à présent nous attacher à l'application-même des méthodes de Morris et de Sobol' au modèle AROME. Il s'agit d'abord d'identifier les paramètres dont nous souhaitons analyser l'incertitude. Une seconde section portera sur le choix des sorties du modèle AROME que nous allons étudier. Puis, nous expliquerons en détail la configuration des méthodes de Morris et Sobol' pour l'application à AROME. Enfin, une réduction du coût de calcul de ces analyses de sensibilité sera proposée.

### 5.1 Identification d'une liste de paramètres incertains

Parmi les centaines de paramètres intervenant dans les paramétrisations du modèle de prévision numérique du temps AROME, nous définissons comme incertains les paramètres pour lesquels les campagnes de mesure n'ont pas permis de connaître la valeur avec précision, ou encore ceux dont le réglage a été revu à de nombreuses reprises par les physiciens lors des différentes versions du modèle AROME. Par ailleurs, des tests préliminaires, réalisés à Météo-France avec un modèle AROME centré sur les Antilles ont montré la sensibilité des prévisions de pluie du modèle à certains paramètres du schéma de diffusion (Faure et al, communication personnelle). Ceux-ci ont donc été définis comme incertains. Enfin, la mise en place d'une technique objective de calibration des paramètres des schémas de turbulence et convection peu profonde du modèle AROME a permis de montrer l'incertitude pesant sur la valeur de certains paramètres (Couvreur et al., 2020).

Ainsi, avec l'aide des experts physiciens, 21 paramètres provenant des paramétrisations physiques et dynamiques ont été identifiés comme source d'incertitude du modèle AROME. Ceux-ci interviennent dans les schémas de rayonnement (Fou-

quart and Bonnel, 1980; Mlawer et al., 1997), microphysique nuageuse (Pinty and Jabouille, 1998), turbulence (Cuxart et al., 2000; Bougeault and Lacarrere, 1989), surface (Le Moigne et al., 2009), convection peu profonde (Pergaud et al., 2009) et de diffusion (Váňa et al., 2008) et sont listés dans le tableau 5.1.

**Rayonnement :** Concernant le schéma de rayonnement, deux paramètres, RS-WINHF et RLWINHF, dont le réglage a été revu depuis la version originale du modèle AROME ont été sélectionnés. Ceux-ci correspondent, respectivement pour le rayonnement solaire et thermique, à des coefficients permettant de rendre compte de l’inhomogénéité nuageuse au sein de la maille du modèle. Autrement dit, ils permettent de pondérer les propriétés radiatives de l’atmosphère telles que la diffusion, la réfraction ou l’épaisseur optique, selon que la maille est composée d’un seul nuage uniforme (nuage stratiforme) ou de plusieurs petits nuages convectifs. Lorsque ces paramètres sont élevés (à 1), les propriétés optiques sont plus élevées, correspondant ainsi à un nuage uniforme (stratiforme), alors que des faibles valeurs (0.6) se réfèrent plutôt à une forte inhomogénéité nuageuse soit plutôt des nuages convectifs.

**Microphysique nuageuse :** Les seuils d’autoconversion, définissant indirectement la distinction entre deux types d’hydrométéores au sein du schéma de microphysique, ne peuvent être estimés précisément puisque cette distinction entre hydrométéores n’existe pas dans la nature : il y a continuité entre gouttelette nuageuse et goutte de pluie. En effet, l’autoconversion étant le phénomène de collision entre deux hydrométéores formant un nouvel hydrométéore d’une autre catégorie, il faut alors définir un seuil pour déterminer le type du nouvel hydrométéore créé lors d’une collision. Par exemple, si deux cristaux de glace entrent en collision et forment un nouveau cristal de glace plus gros, un seuil doit alors définir si ce nouveau cristal est considéré comme un flocon de neige précipitable ou comme un cristal de glace en suspension dans l’atmosphère. Pour la pluie, le seuil correspond à la différence entre eau nuageuse et pluie. Ainsi, RCRIAUTI et RCRIAUTC caractérisant ces seuils respectivement, pour la neige et la pluie font partie des paramètres identifiés dont la valeur est incertaine. Plus ces paramètres sont élevés, moins il y a, respectivement de neige ou de pluie.

En plus de ces paramètres, le paramètre VSIGQSAT du schéma de condensation sous-maille est aussi considéré. Celui-ci contrôle la variance de la distribution de probabilité sous-maille de l’écart au point de saturation. Il vient s’ajouter à la variance diagnostiquée par le schéma de turbulence et a été ajouté afin de garder une variance non nulle dans les cas pour lesquels la turbulence est faible. Plus ce paramètre est élevé, plus la probabilité de saturation au sein de la maille sera élevée, plus il y a aura d’eau condensée sur une fraction de la grille.

**Turbulence :** Dans le schéma représentant la turbulence, sept paramètres ont été sélectionnés. XLINI définit la longueur de mélange minimale utilisée dans le schéma de Bougeault and Lacarrere (1989). Plus il est élevé, plus le phénomène de dissipation et donc de turbulence est important. XCTD, noté  $C_\theta = C_q = C_{q\theta}$  dans Cuxart et al. (2000) est une constante dans la dissipation de la température et de l’humidité par la turbulence. XCTP ( $C_h = C_s = C_S$  dans Cuxart et al. (2000)) et XCEP ( $C_m$  dans Cuxart et al. (2000)) sont des coefficients de corrélation

température ou pression de vapeur et pression-vent, respectivement. XCED est un coefficient intervenant dans la dissipation de l'énergie cinétique turbulente ( $C_\epsilon$  dans Cuxart et al. (2000)). XPHI\_LIM définit les seuils critiques de l'inverse des nombres de Schmidt et Prandlt, notés  $\Phi_3$  et  $\Psi_3$  dans Cuxart et al. (2000), le nombre de Prandlt définissant le rapport entre convection et conduction alors que le nombre de Schmidt définit le rapport entre transfert de quantité de mouvement et transport de matière par diffusion. Enfin, XCET est un coefficient dans le terme de transport de l'équation de l'énergie cinétique turbulente, noté  $C_\epsilon$  dans Cuxart et al. (2000).

**Diffusion :** Pour ce qui est du schéma de diffusion des hydrométéores, trois paramètres sont retenus. SLHDEPSH est un paramètre sans dimension définissant l'intensité horizontale de la diffusion sous forme de lissage laplacien. SLHDKMIN et SLHDKMAX en revanche imposent les valeurs minimales et maximales du coefficient d'amortissement dans l'interpolation du vecteur d'état au point de trajectoires déterminées par le schéma semi-lagrangien. Par ailleurs, SLHDKMIN contrôle aussi la proportion entre deux interpolations (polynôme de Lagrange cubique et quadratique). Plus ces paramètres sont élevés, plus la diffusion est importante.

**Surface :** Deux paramètres sont retenus dans le schéma de surface SURFEX. XRIMAX est le nombre de Richardson critique, c'est-à-dire le seuil à partir duquel l'atmosphère est turbulente. Il définit alors les flux turbulents produits par le schéma de surface. Ceux-ci seront repris par le schéma de turbulence comme condition à la surface et vont donc modifier indirectement le comportement du schéma de turbulence. Plus ce paramètre est faible, plus les phénomènes d'entraînement et détrainement sont faibles. XFRACZ0 est un coefficient pondérant la longueur de rugosité sur les reliefs. Plus ce paramètre est élevé, plus la longueur de rugosité est faible, plus la traînée orographique est faible.

**Convection peu profonde :** Enfin, dans le schéma de convection peu profonde, XCMF est un coefficient contrôlant l'intensité du flux de masse en basse couche et donc de l'ascendance convective. XABUO et XBDETR sont des coefficients définissant l'intensité des phénomènes de flottabilité et de détrainement dans le calcul de la vitesse verticale de l'ascendance. XENTR\_DRY est un coefficient pondérant l'apport d'air sec environnant par entraînement dans ce flux de masse.

Au delà de cette simple identification, les experts physiciens ont aussi fourni des gammes de perturbations de ces paramètres. Celles-ci ont été estimées à partir de la littérature ou encore à partir des nombreux tests et corrections réalisés par les experts physiciens. Une attention a été donnée quant au réalisme de ces intervalles : les valeurs connues comme non-physiques ont été interdites. Notons que les valeurs par défaut de ces paramètres dans le modèle AROME ne sont pas nécessairement centrées au milieu des intervalles de perturbations. Aucune distribution de probabilité ni valeur de paramètres n'étant privilégiée, une distribution de probabilité uniforme est donc suggérée pour chacun des paramètres.

Une synthèse des paramètres et de leurs intervalles de variation est donnée dans le tableau 5.1.

PARTIE II. ANALYSE DE SENSIBILITÉ DU MODÈLE AROME AUX PARAMÈTRES INCERTAINS DES PARAMÉTRISATIONS PHYSIQUES ET DYNAMIQUE

Paramétrisation	Paramètre	Signification physique	Défaut	Minimum	Maximum
Rayonnement	RSWINHF	Facteur d'inhomogénéité pour le rayonnement solaire	1	0.6	1
	RLWINHF	Facteur d'inhomogénéité pour le rayonnement thermique	1	0.6	1
Microphysique	RCRIAUTI	Seuil d'autoconversion de la neige	0.2e-3	0.2e-4	0.25e-3
	RCRIAUTC	Seuil d'autoconversion de la pluie	1e-3	0.4e-3	1e-3
	VSIGQSAT	Constante pour la condensation sous-maille	0.02	0	0.1
	XLINI	Longueur de mélange minimale	0	0	0.2
Turbulence	XCTD	Constante pour la dissipation des fluctuations de température et humidité	1.2	0.98	1.2
	XCTP	Constante pour la corrélation de Température/Pression de vapeur	4.65	1.035	22.22
	XCEP	Constante pour la corrélation Vent-Pression	2.11	0.225	4.0
	XCED	Constante pour la dissipation de la TKE	0.85	0.4	2
	XPHL_LIM	Seuil pour $Sc^{-1}$ et $Pr^{-1}$	3	1	4.5
	XCET	Constante pour le transport de TKE	0.4	0.072	1.512
Diffusion	SLHDEPSH	Force de la diffusion	0.060	0.01	0.09
	SLHDKMIN	Minimum de la fonction de diffusion	0	-1	1
	SLHDKMAX	Maximum de la fonction de diffusion	6	4	12
Surface	XRIMAX	Nombre de Richardson critique	0.2	0	0.3
	XFRACZ0	Coefficient pour la traînée orographique	5	2	10
Convection peu profonde	XCNMF	Coefficient de fermeture dans les basses couches	0.065	0	0.1
	XABUO	Coefficient de la flottabilité	1	0.7	1.5
	XBDETR	Coefficient du déentraînement	1e-6	0	1
	XENTR_DRY	Coefficient pour l'entraînement sec	0.55	0.1	0.699

TABLE 5.1 – Liste des 21 paramètres incertains identifiés par les experts physiciens. La première colonne indique le schéma de paramétrisation dans lequel le paramètre est impliqué. Son nom est indiqué dans la seconde colonne. La troisième décrit la fonction du paramètre alors que la quatrième donne la valeur par défaut du paramètre, utilisé dans AROME. Les valeurs extrêmes acceptées pour chacun des paramètres sont renseignées dans les deux dernières colonnes.

## 5.2 Choix de la sortie d'AROME pour analyser l'impact des paramètres

Puisque nous souhaitons étudier l'influence des paramètres précédemment évoqués sur les prévisions du modèle AROME, il nous faut alors définir la sortie du modèle sur laquelle étudier l'impact de ces paramètres. Différents champs météorologiques de temps sensible sont alors considérés afin d'évaluer la sensibilité du modèle sur divers phénomènes météorologiques. Nous avons porté notre attention sur :

- la force du vent à 10m (ff10m) ;
- les rafales sur 1h à 10m (ffgust) ;
- les précipitations de pluie cumulées sur 1h, 3h, 6h et 24h (prec01, prec03, prec06, prec24) ;
- la couverture nuageuse totale (cloud) ;
- l'humidité relative à 2m (RH2m) ;
- la température à 2m (T2m) ;
- le rayonnement solaire global descendant cumulé sur 1h au niveau de la surface (Sol01).

L'influence des paramètres pouvant dépendre de la situation météorologique et de la saison, les analyses de sensibilité ont été effectuées sur trois saisons de 31 jours définies dans le tableau 5.2. L'impact des paramètres est étudié sur des prévisions initialisées à 21hUTC pour chacune des dates indiquées dans le tableau. Celles-ci utilisent les conditions de surface et initiales issues de l'analyse opérationnelle d'AROME dite 3D-Var (Brousseau et al., 2016). Les conditions latérales sont fournies par les prévisions globales opérationnelles du membre de contrôle de la PEARP, utilisant les paramétrisations par défaut d'ARPEGE (Courtier et al., 1991; Déqué et al., 1994). Notons que ces prévisions à 48h d'échéance étant séparées de 2 ou 3 jours, celles-ci peuvent être supposées indépendantes puisqu'une faible corrélation temporelle est attendue entre elles.

Saison	Période	Intervalle entre 2 prévisions
Été 2018	du 01/05/2018 au 30/07/2018	tous les 3 jours
Automne 2018	du 01/10/2018 au 30/11/2018	tous les 2 jours
Hiver 2018-2019	du 01/12/2018 au 30/01/2019	tous les 2 jours

TABLE 5.2 – Période pour l'analyse de sensibilité et calcul des scores

Les méthodes de Morris et Sobol' utilisées dans notre étude demandant des sorties scalaires du modèle, deux types de sorties scalaires sont utilisées : les champs météorologiques moyens et des scores déterministes tels que le biais moyen, la racine carrée de l'erreur quadratique moyenne (RMSE) et l'erreur absolue moyenne

(MAE). Ces derniers, définis au paragraphe 2.6.1, sont calculés à partir des observations SYNOP et RADOME comme présenté au paragraphe 2.6.7.

Disposant des sorties du modèle ( $f(x, s, d, h)$ ) pour chacune des  $D = 31$  journées ( $d$ ) des  $S = 3$  saisons ( $s$ ) considérées, sur  $H = 16$  échéances ( $h$ ) allant de 3 à 48h avec un pas de 3h ainsi que sur chaque station au sol  $x$  pour les scores déterministes ou sur chaque point de grille  $x$  pour les champs météorologiques, plusieurs indices de sensibilité  $I_i$  du paramètre  $X_i$  peuvent être obtenus en considérant différentes définitions de l'opérateur de moyenne. Le tableau 5.3 référence l'ensemble des indices de sensibilité considérés. A partir de sorties du modèle  $f(x, s, d, h)$  (champ météorologique, biais, RMSE ou MAE) moyennées spatialement sur chaque station de mesure ou point de grille  $x$  ( $f(d, h) = \frac{1}{X} \sum_x f(x, d, h)$ ), on définit alors, pour chacune des variables de surface précédemment définies, quatre types d'indices de sensibilité :

- Globaux : représentant l'influence moyenne des paramètres sur une année (été, automne, hiver). Ils sont obtenus à partir des indices de sensibilité moyennés sur les 3 saisons, jours, et les différentes échéances (équation 5.1).
- Saisonniers : représentant l'influence moyenne des paramètres sur chaque saison. Ils sont obtenus à partir des indices de sensibilité moyennés sur les 31 jours de la saison et les différentes échéances considérées (équation 5.2).
- Journaliers : représentant l'influence moyenne des paramètres en fonction du jour. Ils sont obtenus, pour chaque saison et jour, à partir des indices de sensibilité moyennés sur les différentes échéances considérées (équation 5.3).
- Horaires : représentant l'influence moyenne des paramètres en fonction des échéances. Ils sont obtenus, pour chaque saison et échéance, à partir des indices de sensibilité moyennés sur la totalité des jours d'une saison (équation 5.4).

Un dernier type d'indices de sensibilité dit spatial, s'intéresse à l'influence sur le domaine AROME-France des paramètres (équation 5.5). Dans ce cas, les indices de sensibilité sont calculés à partir des sorties du modèle moyennées selon les jours ( $f(x, s, h) = \frac{1}{D} \sum_d f(x, s, d, h)$ ).

Notons que des versions des méthodes de Morris et Sobol' pour des sorties vectorielles permettent d'aboutir à des indices de sensibilité uniques (par ex, Gamboa et al., 2013). Cependant, elles n'ont pas été considérées ici car nous souhaitons étudier, de manière simple, l'influence des paramètres selon différents points de vue (saisons, échéances, dates, espace, variables météorologiques).

## 5.3 Configuration des méthodes de Morris et de Sobol'

### 5.3.1 Choix des paramètres de Morris

Concernant l'analyse de sensibilité de Morris, nous avons choisi un plan d'expérience comprenant  $r = 12$  trajectoires de Morris, dont l'espacement a été maximisé

Influence des paramètres	Définition des indices de sensibilité
Globale	$I_{i,global} = \frac{1}{S} \frac{1}{D} \frac{1}{H} \sum_s \sum_d \sum_h I_i \left( \frac{1}{X} \sum_x f(x, s, d, h) \right) \quad (5.1)$
Saisonnnière	$I_{i,saisonnier}(s) = \frac{1}{D} \frac{1}{H} \sum_d \sum_h I_i \left( \frac{1}{X} \sum_x f(x, s, d, h) \right) \quad (5.2)$
Journalière	$I_{i,journalier}(d, s) = \frac{1}{H} \sum_h I_i \left( \frac{1}{X} \sum_x f(x, s, d, h) \right) \quad (5.3)$
Horaire	$I_{i,horaire}(h, s) = \frac{1}{D} \sum_d I_i \left( \frac{1}{X} \sum_x f(x, s, d, h) \right) \quad (5.4)$
Spatiale	$I_{i,spatial}(x, s) = \frac{1}{H} \sum_h I_i \left( \frac{1}{D} \sum_d f(x, s, d, h) \right) \quad (5.5)$

TABLE 5.3 – Définition des indices de sensibilité, avec  $f(x, s, d, h)$ , la sortie du modèle dépendant de l'espace  $x$ , du jour de début de prévision  $d$  et de l'échéance  $h$ .  $I_i$  est l'indice de sensibilité du paramètre  $X_i$ .  $S$  est le nombre de saison de 31 jours,  $D$  est de nombre de jours pour chaque saison,  $H$  est le nombre d'échéances et  $X$  est le nombre de données spatiales (nombre de stations au sol pour les scores déterministes (environ 1500), ou nombre de points de grille pour les champs moyennés (environ 200.000)).

avec la méthode de Campolongo et al. (2007). D'après Campolongo et al. (2007) et Di et al. (2015), ceci est un bon compromis entre le coût de calcul du modèle et la robustesse de la méthode. Cette méthode, appliquée aux  $k = 21$  paramètres précédemment décrits, demande alors  $r(k + 1) = 12 \times (21 + 1) = 264$  simulations du modèle AROME. Par ailleurs, nous avons choisi une discrétisation des paramètres en 8 niveaux avec un pas de perturbation  $\Delta = \frac{4}{p-1}$  suffisamment large pour mesurer l'impact de la perturbation globale et non locale des paramètres. L'indice de sensibilité  $I_i$  utilisé pour cette méthode est alors l'indice de Morris total défini au chapitre précédent, soit  $I_i = \sqrt{\mu_i^{*2} + \sigma_i^2}$

### 5.3.2 Estimation des indices de Sobol'

Comme évoqué au chapitre précédent, l'estimation des indices de Sobol' nécessite des milliers de simulations pour estimer le plus finement possible la variance de la sortie du modèle. Or, de par le coût de calcul du modèle AROME, un tel plan d'expérience ne peut être réalisé. Nous avons décidé alors d'employer un émulateur, l'expansion du Chaos Polynomial, utilisant la méthode Least-Angle Regression, afin d'estimer analytiquement ces indices de Sobol'. Disposant d'ores et déjà de 264 échantillons grâce au plan d'expérience de Morris, un polynôme du chaos d'ordre 5 est alors construit indépendamment pour chaque sortie scalaire de biais, MAE et RMSE moyennées sur les différentes stations au sol, des différentes variables météorologiques, saisons, jours et échéances. Ainsi, ceci représente un total de 44 640 chaos polynomial à construire.

Pour valider chacun d'entre eux, la méthode « Leave-One-Out » est utilisée.

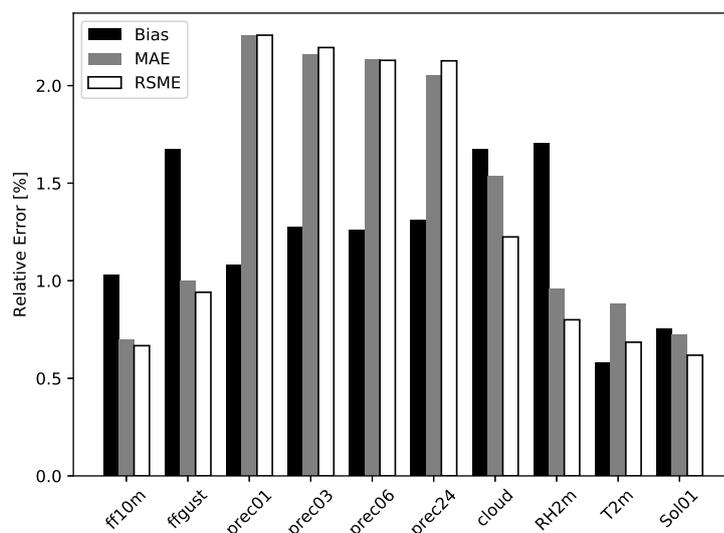


FIGURE 5.1 – Erreur relative de prédiction dite « Leave-One-Out », moyennée sur toutes les saisons, jours, échéances des différentes variables météorologiques et sortie du modèle. En noir, l’erreur de prédiction du biais moyen, en gris de la MAE et blanc de la RMSE

A partir d’un apprentissage sur l’ensemble des données sauf une, le métamodèle prédit la valeur de l’unique donnée non apprise. Son erreur de prédiction, relative à la vraie valeur, est alors estimée. En réitérant sur l’ensemble des données d’apprentissage, on obtient alors une erreur moyenne relative de prédiction. Celles-ci, moyennées sur les prédictions pour les différentes saisons, jours et échéances, sont présentées en figure 5.1. Ces erreurs ne dépassant pas les 2.25% valident l’utilisation de tels émulateurs pour estimer analytiquement les indices de Sobol’. Plus précisément, nous constatons que l’émulation de vent à 10m (ff10m), de température à 2m (T2m) et rayonnement solaire à la surface accumulé pendant une heure (Sol01) présentent les meilleures erreurs (plus petites que 1%), quelque soit le score émulé. Les erreurs d’émulation les plus importantes concernent l’émulation du biais de rafales à 10m (ffgust), couverture nuageuse totale (cloud) et humidité relative à 2m (RH2m). Pour ce qui est de l’émulation de MAE et RMSE, les plus fortes erreurs sont observées pour les précipitations.

Seuls les indices de Sobol’ totaux et de second ordre seront étudiés. Les émulateurs étant valides, des indices de sensibilité  $I_i$  peuvent alors être calculés analytiquement comme présenté en section 4.3.

## 5.4 Réduction du coût de calcul

L’analyse de sensibilité de Morris demandant 264 sorties du modèle AROME pour les 31 dates des trois saisons, un total de 24 552 prévisions est alors nécessaire. Dans le but de réduire le temps de calcul de la production d’un tel échantillon de données et pour se rapprocher de la configuration de la PEARO opérationnelle, nous avons, tout d’abord, utilisé le modèle AROME avec la même résolution spatiale que la PEARO, soit 2.5km.

Pour encore réduire le coût de calcul, différentes solutions ont été envisa-

gées : utilisation d'une version hydrostatique du modèle AROME sans schéma prédicteur-correcteur dans le schéma semi-implicite et augmentation du pas de temps du modèle. En effet, les gains en temps de calcul peuvent être importants selon les différentes configurations (voir tableau 5.4). On constate un gain de près de 40% en temps de calcul en passant de la version non hydrostatique à la version hydrostatique sans schéma prédicteur-correcteur. L'augmentation du pas de temps de 60s à 75s ou 90s permet d'ajouter un gain de 16% et 29% respectivement.

<b>Non Hydrostatique</b>	<b>Schéma Prédicteur-Correcteur</b>	<b>Pas de temps</b>	<b>Temps de calcul</b>
Oui	Oui	60s	72min
Non	Non	60s	44min
Non	Non	75s	37min
Non	Non	90s	31min

TABLE 5.4 – *Temps de calcul pour les différentes configurations d'AROME à 2.5km*

Afin de vérifier que de telles modifications du modèle AROME restent sans conséquence significative sur les prévisions du modèle, nous avons calculé les scores probabilistes des différentes variables météorologiques considérées pour les différentes configurations de la PEARO listées dans le tableau 5.4 (Figure 5.2). Plus particulièrement, ces scores ont été calculés sur 31 jours allant du 15 septembre au 15 octobre 2018 avec le réseau de 21hUTC de la PEARO. Une première configuration de référence (bleu) correspond à une PEARO utilisant la version d'AROME opérationnelle, soit non-hydrostatique, avec schéma prédicteur-correcteur et un pas de temps de 60s. Une deuxième simulation (rose) utilise une version hydrostatique sans schéma prédicteur-correcteur et même pas de temps (60s). Les deux autres ensembles ne diffèrent du précédent que par le pas de temps utilisé, soit 75s et 90s (vert et jaune respectivement). Nous présentons ici uniquement les scores pour lesquels la différence entre ces configurations est la plus importante. Pour chacune des échéances, la significativité à 95% de la différence avec la version non hydrostatique a été calculée avec la méthode décrite au paragraphe 2.6.7 et est marquée par des points apparaissant aux échéances concernées.

C'est pour le CRPS de vent à 10m que l'on constate les plus fortes différences entre les configurations. Pour la quasi totalité des échéances, la configuration hydrostatique avec un pas de temps de 90s (jaune) est significativement différente de la référence. Pour la version à 75s (vert), cette différence est plus modérée. En revanche, pour la version ne différant que par l'utilisation de l'hydrostatisme et schéma prédicteur-correcteur (rose), les différences ne sont significatives qu'en fin d'après midi, lors de la plus forte activité de la convection. Ceci est en partie dû au fait que dans l'hypothèse non hydrostatique, les vitesses verticales de la convection sont mieux modélisées. Tout comme le CRPS d'humidité relative, les aires sous la courbe de ROC pour le vent à 10m dépassant les 20km/h et les précipitations cumulées sur 3h dépassant 2mm montrent que les configurations utilisant des pas de temps plus grands (jaune et vert) s'écartent significativement des scores de la version de référence. En revanche pour un pas de temps identique à la référence, la différence entre les aires sous la courbe ROC n'est significative que pour un nombre très limité d'échéance (1 à 3).

Notons que si les scores sont améliorés avec l'utilisation d'une version hydrosta-

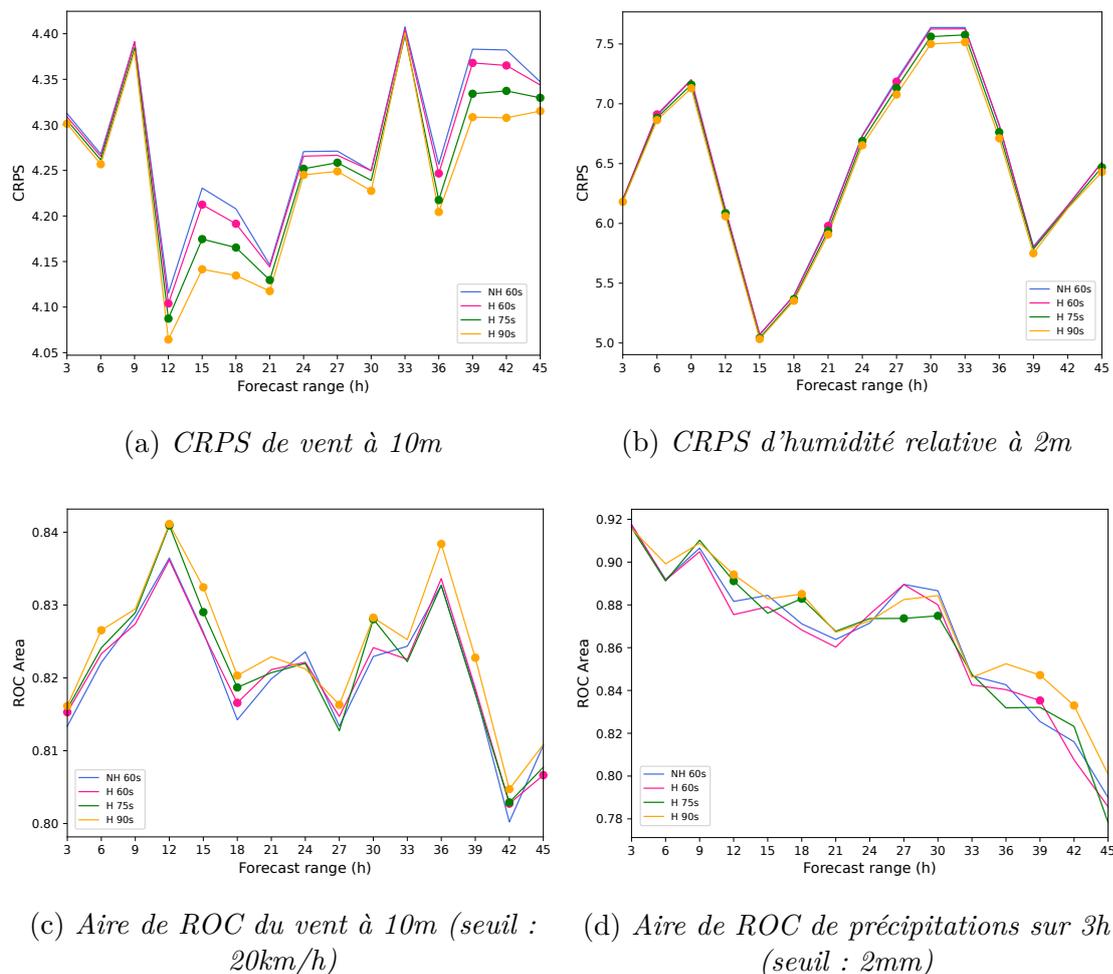


FIGURE 5.2 – Scores avec et sans hypothèse hydrostatique et test de différents pas de temps : Non-hydrostatique/60s (bleu) ; Hydrostatique/60s (rose) ; Hydrostatique/75s (vert) ; Hydrostatique/90s (jaune)

tique du modèle et une augmentation du pas de temps, ceci est dû au fait que les champs météorologiques sont plus lissés et provoque ainsi de plus faibles erreurs due à la double peine (erreur comptée doublement : absence de l'évènement prévu au niveau de l'observation et fausse alarme en un temps ou point différent)

Ainsi, la version hydrostatique du modèle AROME proposant des scores globalement similaires à son équivalent non hydrostatique, l'utilisation de l'hypothèse hydrostatique est alors retenue. Par ailleurs, le schéma prédicteur-correcteur, servant principalement à stabiliser le schéma semi-implicite, devient inutile puisque la version hydrostatique est plus stable que son homologue non hydrostatique. Il apporte même un surplus de calcul non nécessaire. Notons que dans la première version du modèle AROME à 2.5km, résolution utilisée pour cette étude, il n'y avait qu'une seule itération dans le schéma semi-implicite, soit seulement la partie prédicteur et non la partie correcteur. Nous décidons alors de supprimer ce schéma, revenant ainsi à réaliser l'unique étape de prédiction.

Enfin, les scores pour une version d'AROME avec des pas de temps plus grands étant un peu plus éloignés de la version avec un pas de temps de 60s, nous décidons de ne pas changer le pas de temps du modèle et de rester à 60s. Ceci est d'autant plus raisonnable que le schéma de microphysique dépend de ce pas de temps.

Changer cette variable pourrait rendre les résultats des analyses de sensibilité non valides pour la version opérationnelle d'AROME.

## Bibliographie

- P. Bougeault and P. Lacarrere. Parameterization of orography-induced turbulence in a mesobeta-scale model. 117(8) :1872 – 1890, 1989. doi : 10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.
- P. Brousseau, Y. Seity, D. Ricard, and J. Léger. Improvement of the forecast of convective activity from the arome-france system. *Quarterly Journal of the Royal Meteorological Society*, 142(699) :2231–2243, 2016. doi : <https://doi.org/10.1002/qj.2822>. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2822>.
- F. Campolongo, J. Cariboni, and A. Saltelli. An effective screening design for sensitivity analysis of large models. 22 :1509–1518, January 2007. doi : doi:10.1016/j.envsoft.2006.10.004.
- P. Courtier, C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas. The arpege project at meteo france. In *Seminar on Numerical Methods in Atmospheric Models, 9-13 September 1991*, volume II, pages 193–232, Shinfield Park, Reading, 1991. ECMWF, ECMWF. URL <https://www.ecmwf.int/node/8798>.
- F. Couvreux, F. Hourdin, D. Williamson, R. Roehrig, V. Volodina, N. Villefranque, C. Rio, O. Audouin1, J. Salter, E. Bazile, F. Brient, F. Favot, R. Honnert, M.-P. Lefebvre, J.-B. Madeleine, Q. Rodier, and W. Xu. Process-based climate model development harnessing machine learning : I. a calibration tool for parameterization improvement. n/a(n/a) :e2020MS002217, 2020. doi : 10.1029/2020MS002217. e2020MS002217 2020MS002217.
- J. Cuxart, P. Bougeault, and J. L. Redelsperger. A turbulence scheme allowing for mesoscale and large-eddy simulations. 126 :1–30, 2000. doi : 10.1002/qj.49712656202.
- Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, J. Quan, J. Li, C. Miao, A. Ye, and C. Tong. Assessing wrf model parameter sensitivity : A case study with 5 day summer precipitation forecasting in the greater beijing area. 42(2) :579–587, 2015. doi : 10.1002/2014GL061623.
- M. Déqué, C. Drevet, A. Braun, and D. Cariolle. The arpege/ifs atmosphere model : a contribution to the french community climate modelling. 10 : 249–266, 1994. doi : 10.1007/BF00208992. URL <https://doi.org/10.1007/BF00208992>.
- Y. Fouquart and B. Bonnel. Computations of solar heating of the earth's atmosphere—a new parameterization. 53 :35–62, 1980.
- F. Gamboa, A. Janon, T. Klein, and A. Lagnoux. Sensitivity analysis for multi-dimensional and functional outputs. 8, 11 2013. doi : 10.1214/14-EJS895.

- P. Le Moigne, A. Boone, J.-C. Calvet, B. Decharme, S. Faroux, A.-L. Gibelin, C. Lebeaupin, J.-F. Mahfouf, E. Martin, V. Masson, D. Mironov, J. Noilhan, P. Tulet, and B. Van Den Hurk. Surfex scientific documentation. Technical report, Météo-France, 2009. Note de centre du Groupe de Météorologie à Moyenne Echelle, available at : <http://www.cnrm.meteo.fr/surfex/>.
- 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. 102(D14) :16663–16682, 1997. doi : 10.1029/97JD00237.
- J. Pergaud, V. Masson, S. Malardel, and F. Couvreux. A parameterization of dry thermals and shallow cumuli for mesoscale numerical weather prediction. 132 (1) :83, 2009. doi : 10.1007/s10546-009-9388-0.
- J. Pinty and P. Jabouille. A mixed-phased cloud parameterization for use in a mesoscale non-hydrostatic model : simulations of a squall line and of orographic precipitation. In A. M. Society, editor, *Preprints of Conference On Cloud Physics*, pages 217–220, Everett, WA, 1998.
- F. Váňa, P. Bénard, J.-F. Geleyn, A. Simon, and Y. Seity. Semi-lagrangian advection scheme with controlled damping : An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. 134(631) :523–537, 2008. doi : 10.1002/qj.220.

# Chapitre 6

## Sensibilité du modèle AROME à divers paramètres d'entrée des paramétrisations

Dans ce chapitre, nous allons nous intéresser aux résultats des analyses de sensibilités de Morris et Sobol' présentées au chapitre 4 et appliquées aux paramètres incertains des paramétrisations du modèle AROME comme indiqué au chapitre 5. Ces résultats ont fait l'objet d'un article de recherche soumis au *Quarterly Journal of the Royal Meteorological Society*, dont j'ai réalisé l'ensemble des graphiques, analysé les données et rédigé l'intégralité sous les conseils et corrections de mes encadrants. Après une première révision par l'éditeur, l'article a été accepté pour publication (Wimmer et al.). Les méthodes ayant déjà été présentées, nous référons le lecteur à la partie 4 de l'article.

### 6.1 Résumé de l'article

Des analyses de sensibilité aux modifications de certains paramètres du modèle régional AROME ont été réalisées afin de déterminer ceux ayant le plus fort impact sur les prévisions de champs météorologiques, tels que la force du vent à 10m (ff10m), les rafales associées (ffgust), les précipitations cumulées sur 1, 3, 6 et 24h (prec01, prec03, prec06, prec24), la couverture nuageuse totale (cloud), l'humidité relative à 2m (RH2m), la température (T2m) et le rayonnement solaire descendant au niveau de la surface et accumulé sur 1h (Sol01).

Une première analyse de sensibilité utilisant la méthode de Morris a été appliquée aux 21 paramètres identifiés par les experts physiciens, provenant de six paramétrisations physiques et dynamiques du modèle AROME. Parce que la sensibilité des prévisions aux paramètres peut varier selon la saison, l'analyse a été répétée pour les 3 saisons définies dans le tableau 5.2. Sur la totalité des 3 saisons, cette analyse a permis d'identifier les 8 paramètres les plus influents sur les diverses variables météorologiques étudiées et provenant des différentes paramétrisations : RSWINHF, VSIGQSAT, XCTP, XCEP, XCED, SLHDEPSH, XFRACZ0 et XCMF. Notons que la méthode de Sobol' utilisant l'expansion du chaos polyno-

mial comme émulateur pour calculer analytiquement les indices de Sobol' des 21 paramètres, confirme les résultats de l'analyse de Morris mais suggère cependant de réduire la liste des paramètres influents à VSIGQSAT, XCTP, XCEP, XCED et SLHDEPSH.

Au delà de la simple identification des paramètres les plus influents, une étude plus poussée à partir de l'analyse de sensibilité de Morris a permis d'étudier plus en détails la sensibilité du modèle selon différentes échelles spatio-temporelles. Tout d'abord, la comparaison des résultats selon les 3 saisons d'étude a montré l'importance des paramètres de la convection en été. Au cours d'une même saison, la liste des paramètres influents peut varier selon la situation météorologique. Il est donc nécessaire de réaliser des analyses de sensibilité sur de longues périodes afin d'obtenir des résultats plus robustes. Une analyse de l'impact des paramètres au cours des échéances de prévision, a montré la présence d'un cycle diurne particulièrement visible sur la période estivale. Celui-ci est cependant quasiment inexistant pour une analyse de sensibilité menée sur la période hivernale. Des cartes d'influence des paramètres ont montré une forte dépendance à la valeur même des champs météorologiques. Cependant, nous avons noté une forte dépendance au type de surface (mer/terre, relief/plaine) pour certains des paramètres influençant le vent et les rafales. Enfin, les indices de Sobol' de second ordre ont permis d'identifier quelques interactions, principalement entre les paramètres identifiés comme les plus influents.

## 6.2 Article

# Sensitivity analysis of the convective-scale AROME model to physical and dynamical parameters

WIMMER M.<sup>1</sup>, RAYNAUD L.<sup>1</sup>, DESCAMPS L.<sup>1</sup>, BERRE L.<sup>1</sup>, SEITY Y.<sup>1</sup>

<sup>1</sup> Centre National de Recherches Météorologiques, Université Toulouse III - Paul Sabatier, Météo-France, CNRS, Toulouse, France

**Correspondence:** M. Wimmer, Centre National de Recherches Météorologiques, Université Toulouse III - Paul Sabatier, 42 Avenue Gaspard Coriolis, 31057 Toulouse France

Email: meryl.wimmer@umr-cnrm.fr

**Abstract.** A global sensitivity analysis of the convective-scale AROME model is performed in order to determine the most influential parameters on the forecast of different near-surface variables. For that purpose, the Morris method is applied to 21 parameters from six different physical and dynamical parametrization schemes, over different seasons. Results highlight a set of eight parameters with a noticeable influence on most variables, in particular 10-meter wind speed and precipitation forecasts. The sensitivity of parameter uncertainties is also examined on different spatio-temporal scales. A clear diurnal cycle of parameters influence is observed in summer, in close connection with the convective activity. In addition, the spatial distribution of parameters influence is mostly consistent with the underlying distribution of weather forecasts. A Sobol' sensitivity analysis, based on surrogate models, mostly confirms Morris conclusions and highlights some interactions between parameters.

**Keywords.** Sensitivity Analysis, Morris screening, Sobol' indices, AROME-EPS, parametrization, mesoscale, model uncertainty, ensemble prediction

## 1 Introduction

Convection-permitting Numerical Weather Prediction (NWP) models are now widely used in most operational weather prediction centers, with horizontal resolutions usually between 1 and 4km (Clark et al., 2016). These models have the advantage, compared to global models, to provide a realistic representation of small-scale phenomena. However, they are still subject to imperfections. One source of forecast errors is related to the representation of unresolvable subgrid-scale physical processes with parametrization schemes, which often provide a simplified modelling of complex physical processes (Stensrud, 2011). Furthermore, these schemes involve many uncertain parameters, whose accuracy is crucial for forecast performances. Generally, the calibration of parameters values is performed manually by searching for the configurations that optimize the forecast-observation

24 comparison. However, in recent years, more objective techniques using emulators have also been  
25 developed (Couvreur et al., 2020).

26 In addition to deterministic forecasts, convective-scale ensemble prediction systems (EPSs)  
27 have been developed to estimate forecast uncertainties (Gebhardt et al., 2008; Bouttier et al., 2012;  
28 Romine et al., 2014; Hagelin et al., 2017; Frogner et al., 2019). Properly accounting for model errors  
29 in these systems is still challenging and different methods have been proposed (Buizza et al., 1999;  
30 Shutts, 2005; Wang et al., 2011). Perturbing uncertain parameters is an appealing approach that has  
31 lead to promising results using different approaches such as Random Parameter (McCabe et al., 2016)  
32 and Stochastically Perturbed Parametrization (SPP, Ollinaho et al., 2013), which directly perturb  
33 some parametrization parameters. An efficient implementation of such methods requires identifying  
34 the relevant parameters to perturb, along with their uncertainty ranges. These choices often rely on  
35 experts knowledge.

36 A better understanding of the impact of parameters on forecasts could help improve the pa-  
37 rameters tuning in deterministic forecasts or the representation of parameters uncertainty in EPS.  
38 For that purpose, objective methods such as Sensitivity Analysis (SA) can be used to measure the  
39 influence of these parameters. These methods consist in sampling the uncertain input parameters  
40 space, in order to identify and sometimes quantify the impact of parameters perturbations on model  
41 outputs (Saltelli et al., 2000). One can distinguish between local SA methods which look at the  
42 impact of small parameter perturbations around nominal values, and global methods which examine  
43 the influence of simultaneous perturbations of all inputs over their entire variation ranges. Global  
44 methods are particularly suited for non-linear behaviours such as those encountered in NWP models.

45 Global SA have already been applied to complex models from different scientific domains includ-  
46 ing atmospheric sciences. For instance, these analyses were conducted for land-surface models (Liu  
47 et al., 2004; Bastidas et al., 2006; Hou et al., 2012; Li et al., 2013), coupled ocean-atmosphere mod-  
48 els (Marzban et al., 2014), global climate models (Qian et al., 2015; Zhang et al., 2012; Stainforth  
49 et al., 2005; Covey et al., 2013) and atmospheric chemical models (Lucas and Prinn, 2005). Regard-  
50 ing NWP models, Di et al. (2015) and Quan et al. (2016) examined the sensitivity of the Weather  
51 Research and Forecasting (WRF) model to a set of parameters from different physical parametriza-  
52 tions, with a focus on nine rainy events. Di et al. (2017) extended these studies over a one-month  
53 summer period and compared results from two SA methods. However, to the authors' knowledge,  
54 no global SA has been performed in an operational convective-scale model over long periods.

55 In this context, the aim of this study is to apply two global SA methods in the French convective-  
56 scale AROME model, in order to identify the most influential parameters from all physical parametriza-  
57 tions and some dynamical schemes.

58 The outline of the paper is the following. Section 2 describes the global sensitivity analysis  
59 methods used. The experimental setup is introduced in section 3 and results are discussed in section  
60 4. Finally, some concluding remarks are provided in section 5.

## 61 **2 Global Sensitivity Analysis**

62 Amongst global SA, three types of methods can be distinguished : (1) screening methods as,  
63 for instance, the Morris method ; (2) variance decomposition methods such as Fourier Amplitude  
64 Sensitivity Testing (FAST) or Sobol' indices method; (3) regression-based methods like, for example,  
65 Pearson coefficient method. In this section, the two state-of-the-art SA methods used in this study  
66 are presented: Morris screening and Sobol' indices.

## 67 2.1 Morris sensitivity analysis

68 The Morris screening (Morris, 1991), also called Elementary Effect method, is a "One-At-a-  
 69 Time" (OAT) design, meaning that each parameter is perturbed one after the other. Screening  
 70 methods are often time-consuming as they require a large number of simulations, especially for a  
 71 high level of accuracy. However, the Morris approach can give robust results with a small number  
 72 of simulations (about 10 or 50 per parameters, Campolongo et al., 2007). Furthermore, another  
 73 noticeable advantage of the Morris approach is that it does not require any specific assumption on  
 74 output monotony properties or parameters influence.

75 This method has been successfully applied to various complex models such as physics-based  
 76 battery (van Rensburg et al., 2019), building energy budget (Sanchez et al., 2014; Branger et al.,  
 77 2015; Yang et al., 2016), seismic sources location (Franczyk, 2019), wastewater model (Ruano et al.,  
 78 2011; Cosenza et al., 2013; Vanrolleghem et al., 2015), hydrothermal of embankment dam (Ren  
 79 et al., 2019), combat (Alam et al., 2004), agronomics (Confalonieri et al., 2010a,b), medicine (Wu  
 80 et al., 2013; Link et al., 2018). In atmospheric sciences, this screening method has been applied to  
 81 a climate model (Covey et al., 2013) and NWP models (Di et al., 2015; Quan et al., 2016; Di et al.,  
 82 2017).

83 The principle of the method is the following. Let us consider a model with  $k$  uncertain pa-  
 84 rameters, denoted  $X_i, i = 1 \dots, k$ , assumed to be continuous and uniformly distributed over specific  
 85 intervals. After a normalization between 0 and 1, these parameters are discretized in  $p$  levels. The  
 86 method starts with initial values of each parameter, randomly sampled in the  $k$  parameters space.  
 87 All parameters are then perturbed one-by-one, in a random order, by a value  $\Delta$  proportional to  $\frac{1}{p-1}$ .  
 88 When all parameters are perturbed, we obtain  $k + 1$  simulations, forming a Morris trajectory.  
 89 For example, Figure 1 presents a Morris design for  $k = 2$  parameters ( $X_1$  and  $X_2$ ), perturbed between 0  
 90 and 1 and discretized in  $p = 5$  values. There are  $r = 5$  Morris trajectories with a perturbation step  
 91  $\Delta = \frac{1}{p-1}$ , starting at points  $A, B, C, D$  and  $E$ , which are randomly chosen.

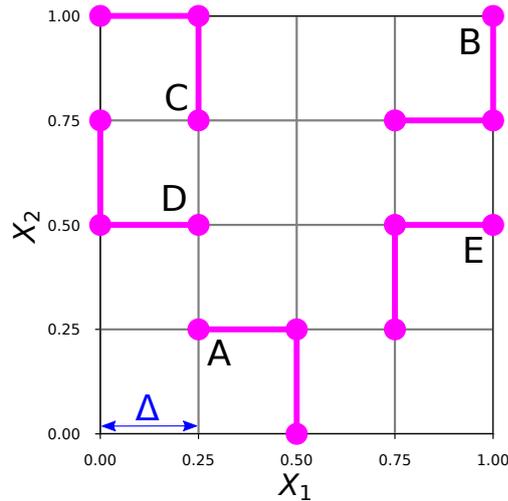


Figure 1: Example of a Morris design for  $k = 2$  parameters ( $X_1$  and  $X_2$ ) with the following settings:  $r = 5$ ,  $p = 5$ ,  $\Delta = \frac{1}{p-1}$ .  $A, B, C, D$  and  $E$  are the starting points of the Morris trajectories.

92 Along a Morris trajectory, the influence of each parameter  $X_i$  is measured with the correspond-

ing elementary effect  $EE_i$  defined as follows :

$$EE_i = \frac{f(X_1, X_2, \dots, X_i + \Delta, \dots, X_k) - f(X_1, X_2, \dots, X_i, \dots, X_k)}{\Delta}, \quad (1)$$

where  $f(X_1, X_2, \dots, X_k)$  is a scalar model output. Note that the perturbation  $\Delta$  can be either positive or negative depending on trajectory orientation (direct or opposite). Therefore, along a trajectory, one elementary effect is computed for each parameter.

As a single Morris trajectory samples a very small part of the  $k$  parameters space,  $r$  trajectories are necessary to sample the whole domain. For example, in Figure 1, the Morris trajectory starting at point  $A$  is not sufficient, 4 other trajectories starting from points  $B$ ,  $C$ ,  $D$  and  $E$  are needed to better sample the whole parameters space  $(X_1, X_2)$ . In order to properly explore the parameters space, we follow Campolongo et al. (2007), who suggest to maximise distances between trajectories. This is achieved by computing the sum of euclidean distances between points from  $M$  different trajectories and by selecting the  $r$  furthest ones. These  $r$  trajectories provide a sample of the whole elementary effects distribution, referred as  $F_i$  for the parameter  $X_i$ . The higher  $r$  is, the more accurate the estimation of  $F_i$  is.

As proposed by Campolongo et al. (2007), the  $F_i$  distribution can be summarised with two sensitivity indices, for each parameter  $X_i$  :

$$\mu_i^* = \mathbb{E}(|EE_i|), \quad \sigma_i = std(EE_i),$$

where  $\mathbb{E}$  is the expected value and  $std$  the standard deviation.  $\mu_i^*$  characterises the direct impact of a parameter on model outputs: the higher  $\mu_i^*$  is, the more influential the parameter is.  $\sigma_i$  represents not only the impact amplitude, but also interactions between parameters and nonlinear effects. High  $\sigma_i$  values can correspond to a strong interaction between considered parameters, or to nonlinear effects. As  $\mu_i^*$  and  $\sigma_i$  represent two complementary measures of the parameters influence, in the remainder of the paper, the total parameter impact will be evaluated with the Total Morris Indices (TMI), given by  $TMI_i = \sqrt{\mu_i^{*2} + \sigma_i^2}$ , as proposed by Ciric (2012).

A major limitation of this method is that parameters interactions and non-linearity properties cannot be distinguished. This is the reason why the Morris approach is mostly used to provide a first insight into the parameters influence. Generally, it is often completed with other methods such as the Sobol' indices described in the next section. Some improvements have also been proposed to determine second-order effect (Cropp and Braddock, 2002; Campolongo et al., 2007) or to account for dependent parameters (Ge and Menendez, 2017), but they are not considered in this study.

## 2.2 Sobol' indices method

The second method, proposed by Sobol' (1990), is a variance-based approach. Whatever the method used for Sobol' indices estimation, they have been very popular in several domains such as hydrology (for instance, Nossent et al. (2011); Sarrazin et al. (2016)), climatology (Shi et al., 2019), atmospheric chemistry (Girard et al., 2016) or atmospheric observation (Rieutord, 2017). This approach has the advantage of providing a deep understanding of parameters influence and interactions without requiring any hypothesis on the studied model. Furthermore, for a more global application, the method has been extended to vectorial outputs (Gamboa et al., 2013) and dependent variables (Chastaing, 2013) in particular.

More precisely, this method uses the Hoeffding decomposition theorem to break down the model

129 output variance  $\mathbb{V}(Y) = V$  (where  $Y = f(X_1, X_2, \dots, X_i, \dots, X_k)$  is the model output) into a sum  
 130 of statistics of increasing dimensions :

$$V = \sum_{i=1}^k V_i + \sum_{1 \leq i < j \leq k} V_{ij} + \dots + V_{1\dots k} \quad (2)$$

where:

$$V_i = \mathbb{V}(\mathbb{E}[Y|X_i])$$

$$V_{ij} = \mathbb{V}(\mathbb{E}[Y|X_i, X_j]) - V_i - V_j$$

...

$$V_{1\dots k} = V - \sum_{i=1}^k V_i - \sum_{1 \leq i < j \leq k} V_{ij} - \dots - \sum_{1 \leq i_1 < \dots < i_{p-1} \leq k} V_{i_1 \dots i_{p-1}}$$

131 According to this variance decomposition, Sobol' (1990) defines sensitivity indices as the pro-  
 132 portion of the model output variance due to a parameter or a group of parameters. These indices  
 133 are commonly referred to as the Sobol' indices. The most frequently used are the first-order Sobol'  
 134 indices, defined, for a  $i^{th}$  parameter  $X_i$ , by:

$$S_i = \frac{V_i}{V}, \quad (3)$$

135 and the second-order indices, defined by:

$$S_{i,j} = \frac{V_{ij}}{V}. \quad (4)$$

136 These second-order indices provide information about interactions between parameters  $X_i$  and  $X_j$ .  
 137 As the sum of all Sobol' indices is equal to one, the higher the Sobol' indice  $S_i$  is, the more influential  
 138 the parameter  $X_i$  is.

139 However, when the number of parameters is too large, the interpretation of all Sobol' indices  
 140 can be complex and the total Sobol' indices (Homma and Saltelli, 1996) can be used to summarise  
 141 the overall contribution of a parameter :

$$S_{Ti} = \sum_{j \# i} S_j = 1 - \frac{\mathbb{V}(\mathbb{E}[Y|X_{\sim i}])}{V}, \quad (5)$$

142 where  $\#i$  represents all indices ensembles containing index  $i$ , and  $X_{\sim i}$  corresponds to all parameters  
 143 except the  $i^{th}$  one.

144 Sobol' indices estimation requires to compute the model output variance  $Y$ , given the parameter  
 145  $X_i$ . According to Jacques (2011), a possible approach is to use a Monte-Carlo sampling with, at  
 146 least, 1.000 simulations for each parameter. To achieve an accuracy on the Sobol indices of about  
 147 10%, a number of 10.000 simulations per parameter is necessary (Iooss, 2011). Designs such as Latin  
 148 Hypercube sampling, stratified sampling or quasi-Monte Carlo can reduce the number of simulations.  
 149 However, compared to the Morris screening this method is much more costly.

150 The use of statistical surrogate models, also called model emulators or metamodels, can be a  
 151 useful alternative to generate thousands of simulations with a very low computational cost. Surrogate  
 152 models are simple mathematical functions that aim at approximating a model output in a computa-  
 153 tionally efficient way. They can be based on polynoms, neural networks, kriging or random forests

154 (Faivre et al., 2016). In this paper, the Polynomial Chaos Expansion (PCE) is used. One advantage  
155 of that surrogate is that the Sobol' indices can be computed analytically from PCE coefficients. For  
156 more details, the reader can refer to Sudret (2008).

### 157 **3 Experimental setup**

158 In this section we present the NWP model used, the selected parameters and the configuration  
159 of the two sensitivity analysis methods.

#### 160 **3.1 AROME model**

161 The French convection-permitting AROME model (Seity et al., 2011) is a spectral and non-  
162 hydrostatic limited-area model. Its dynamical core comes from the ALADIN model (Aire Limitée  
163 Adaptation Développement INternational, Bénard et al., 2010), while the physical parametrizations  
164 are taken from the research model MESO-NH (Lafore et al., 1998).

165 Microphysics is represented with a bulk, one-moment scheme using six water species: water  
166 vapour, cloud liquid water, cloud pristine ice, precipitation rain, snow and graupel (ICE3, Pinty and  
167 Jabouille, 1998). Bouteloup et al. (2011) improved this scheme by adding a PDF-based sedimentation  
168 scheme. The Planetary Boundary Layer turbulence is represented with the Cuxart et al. (2000)'s  
169 scheme, that uses a Turbulent Kinetic Energy (TKE) prognostic equation. Shallow convection is  
170 simulated with an Eddy Diffusivity Kain-Fritsch scheme (Pergaud et al., 2009). Surface is described  
171 with tiles of land, towns, sea and inland waters for each AROME grid box, in the SURFEX scheme  
172 (Le Moigne et al., 2009). Finally, AROME uses a previous version of ECMWF's radiation scheme,  
173 the Rapid Radiative Transfer Model (RRTM) code, to compute longwave (Mlawer et al., 1997) and  
174 shortwave radiation (Fouquart and Bonnel, 1980). A Semi-Lagrangian Horizontal Diffusion (SLHD)  
175 scheme simulates the dissipation of energy at small-scales by controlling the amplitude of numerical  
176 diffusion of hydrometeors (Váňa et al., 2008). This numerical diffusion scheme can be considered as  
177 a parametrization as it represents a subgrid process, despite the fact it is more related to dynamics.

178 AROME has been running operationally since 2008 and produces 2-day forecasts, over a limited  
179 domain covering Western Europe and centred on France. This domain covers an area of size about  
180 1600 km×1600 km with a horizontal resolution of 1.3km. For the present study, the AROME model  
181 has been run with a 2.5km horizontal resolution and 90 vertical levels. Surface and initial conditions  
182 come from the 3D-Var AROME analysis (Brousseau et al., 2016), and lateral conditions are provided  
183 by the global operational ARPEGE model (Courtier et al., 1991; Déqué et al., 1994).

#### 184 **3.2 Setup for parameters SA**

185 Discussions with experts in physical processes lead to the identification of 21 uncertain param-  
186 eters from 6 different parametrizations: radiation, microphysics, turbulence, surface, convection and  
187 SLHD. They are summarised in Table 1.

188 In the radiation scheme, RSWINHF and RLWINHF are inhomogeneity coefficients for shortwave  
189 and longwave radiations. They represent the inhomogeneity of clouds by weighting optical properties  
190 such as diffusion, scattering and optical thickness. The higher these parameters are, the more uniform  
191 the cloud is and the higher the optical properties are. On the contrary, the smaller the parameters  
192 are, the more heterogeneous the cloud is and the smaller the optical properties are. Consequently,  
193 higher values correspond to stratiform clouds and lower values to cumuliform clouds.

194 Regarding microphysics, RCRIAUTI and RCRIAUTC are autoconversion thresholds for snow  
 195 and rain respectively. The higher RCRIAUTI is, the less snow is created. The higher RCRIAUTC, the  
 196 less rain is produced. VSIGQSAT is a parameter from the subgrid condensation scheme that is linked  
 197 to the width of the Probability Density Function (PDF) of the subgrid variance of the saturation  
 198 departure. The higher VSIGQSAT, the larger the PDF and the easiest for the model to condense  
 199 water in a fraction of the grid.

200 In the Cuxart et al. (2000)'s scheme, 7 parameters are selected. XLINI is the minimum of mixing  
 201 length. The higher XLINI is, the higher the dissipation is. XCTD is a constant in the dissipation of  
 202 temperature and relative humidity. This parameter is written  $C_\theta = C_q = C_{q\theta}$  in Cuxart et al. (2000).  
 203 XCTP is a constant for the temperature-vapour pressure correlations, denoted  $C_h = C_s = C_S$  in  
 204 Cuxart et al. (2000). XCEP is a constant for wind-pressure correlations, namely  $C_m$  in Cuxart et al.  
 205 (2000). XCED is a coefficient involved in the dissipation of the turbulent kinetic energy, which  
 206 corresponds to  $C_\epsilon$  in Cuxart et al. (2000). XPHLLIM is a threshold for the reverse Schmidt and  
 207 Prandtl number denoted  $\Phi_3$  and  $\Psi_3$  in Cuxart et al. (2000). At last, XCET is a constant in the  
 208 transport term in the equation of the TKE, referred as  $C_e$  in Cuxart et al. (2000).

209 In the SLHD scheme, SLHDEPSH is a dimensionless parameter representing the intensity of  
 210 horizontal Laplacian smoothing. SLHDKMIN and SLHDKMAX are, respectively, the minimum and  
 211 maximum values for the  $\kappa$  diffusion function described in Váňa et al. (2008).

212 In the surface scheme, XRIMAX is the critical Richardson number, that impacts the planetary  
 213 boundary layer thickness. The smaller XRIMAX is, the lower the entrainment rate is. XFRACZ0 is  
 214 a tuning coefficient of orographic drag. The smaller XFRACZ0 is, the stronger the drag is.

215 Finally, in the shallow convection scheme, XCMF is the closure coefficient at bottom level,  
 216 denoted  $M_0$  in Pergaud et al. (2009). XABUO and XBDETR are coefficients in the calcula-  
 217 tion of ascending vertical velocity, particularly, in the buoyancy and detrainment term, respectively.  
 218 XENTR.DRY is a coefficient in the dry entrainment calculation.

219 For each parameter, experts also suggested a uniform distribution for the perturbation range  
 220 reported in Table 1. This table also presents the default values used in the operational AROME. Note  
 221 that in some cases, the default value is not in the middle of the proposed range to avoid nonphysical  
 222 settings.

## 223 3.3 Experimental design

### 224 3.3.1 Morris screening

225 Regarding the Morris sensitivity analysis, an experimental design of  $r = 12$  trajectories using the  
 226 Campolongo et al. (2007) maximizing distance method is applied to the space of the  $k = 21$  parameters  
 227 previously described. Each parameter perturbation range is discretized in  $p = 8$  levels and the perturbation  
 228 step used is  $\Delta = \frac{4}{p-1}$ . This screening thus requires running  $r(k + 1) = 12 \times (21 + 1) = 264$  AROME  
 229 simulations, which is a good compromise between computational efficiency and the robustness of the results.  
 230 Based on Campolongo et al. (2007) suggestion and Di et al. (2015) findings, this Morris design should be  
 231 sufficient to provide valuable and robust sensitivity results.

232 As the parameters influence is likely to depend on the season and on the day, the analysis is conducted  
 233 for 3 periods of 31 days corresponding to:

- 234 — Summer 2018 : from 01/05/2018 to 30/07/2018, every three days
- 235 — Fall 2018 : from 01/10/2018 to 30/11/2018, every two days
- 236 — Winter 2018-2019 : from 01/12/2018 to 30/01/2019, every two days

Scheme	Parameter	Physical meaning	Default	Range
Radiation	RSWINHF	Shortwave inhomogeneity factor	1	0.6 - 1
	RLWINHF	Longwave inhomogeneity factor	1	0.6 - 1
Microphysics	RCRIAUTI	Snow Autoconversion threshold	0.2e-3	0.2e-4 - 0.25e-3
	RCRIAUTC	Rain Autoconversion threshold	1e-3	0.4e-3 - 1e-3
	VSIGQSAT	Constant for subgrid condensation	0.02	0 - 0.1
Turbulence	XLINI	Minimum mixing length	0	0 - 0.2
	XCTD	Constant for dissipation of temperature and vapor pressure fluctuations	1.2	0.98 - 1.2
	XCTP	Constant for temperature and vapor pressure correlations	4.65	1.035 - 22.22
	XCEP	Constant for wind-pressure correlations	2.11	0.225 - 4.0
	XCED	Constant for dissipation of TKE	0.85	0.4 - 2
	XPHI_LIM	Threshold value for $Sc^{-1}$ and $Pr^{-1}$	3	1 - 4.5
Diffusion	XCET	Constant for transport of TKE	0.4	0.072 - 1.512
	SLHDEPSH	Strength of SLHD	0.060	0.01 - 0.09
	SLHDKMIN	Diffusion function minimum	0	-1 - 1
	SLHDKMAX	Diffusion function maximum	6	4 - 12
Surface	XRIMAX	Critical Richardson Number	0.2	0 - 0.3
	XFRACZ0	Coefficient of orographic drag	5	2 - 10
Convection	XCMF	Closure coefficient at bottom level	0.065	0 - 0.1
	XABUO	Coefficient of the buoyancy	1	0.7 - 1.5
	XBDETR	Coefficient of the detrainment	1e-6	0 - 1
	XENTR_DRY	Coefficient for dry entrainment	0.55	0.1 - 0.699

Table 1: List of AROME parameters selected for the sensitivity analysis. The first column indicates the corresponding parameterization schemes. The third column provides a short description of the parameters. Default values (used in the operational AROME configuration) and minimum/maximum allowed values are given in columns four and five for each parameter.

237 The 264 AROME simulations are run once per day, starting at 21 UTC, and provide 48h forecasts.  
 238 Note that, because of the 3-day or 2-day interval between successive forecasts, temporal correlations are  
 239 expected to be small in the dataset.

240 This configuration thus represents a total of 24.552 AROME forecasts to produce. In order to reduce  
 241 the computational time, an hydrostatic version of the AROME model with an horizontal resolution of 2.5km  
 242 has been used. It has been verified that this hydrostatic option does not degrade the forecast performances.  
 243 It is finally important to mention that no simulation crashed or drifted towards unrealistic forecasts.

### 244 3.3.2 Scalar model outputs used to evaluate the model sensitivity

245 Morris and Sobol' methods measure the parameters influence with respect to a scalar model output.  
 246 In this study, scalar outputs considered are either an averaged forecast field or a performance metric such  
 247 as mean bias, root-mean square error (RMSE) and mean absolute error (MAE). These scores are computed  
 248 using SYNOP (surface synoptic observations) and the French real-time meteorological observations network  
 249 (Tardieu and Leroy, 2003), with the following formulae :

$$245 \quad Bias = \frac{1}{n} \sum_{k=1}^n (y_k - o_k) \quad ; \quad RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k - o_k)^2} \quad ; \quad MAE = \frac{1}{n} \sum_{k=1}^n |y_k - o_k| \quad (6)$$

250 where  $n$  is the number of in-situ measures,  $y_k$  is the model output interpolated to the location of the  $k^{th}$   
 251 observation (with a nearest neighbour approach) and  $o_k$  is the  $k^{th}$  observation.

252 Hence, for all 264 daily simulations, Morris sensitivity indices,  $\mu_i^*(f(d, h))$  and  $\sigma_i(f(d, h))$ , are com-  
 253 puted for each parameter, where  $f(d, h)$  is a spatially-averaged forecast field value or verification metrics,  
 254 corresponding to the forecast issued on day  $d$  and valid at forecast range  $h$ .

255 The parameters influence is examined on different meteorological variables, including 10-meter wind  
 256 speed (ff10m), 10-meter wind gust (ffgust), 1-hourly, 3-hourly, 6-hourly and 24-hourly accumulated precipita-  
 257 tion (prec01, prec03, prec06, prec24), total cloud cover (cloud), 2-meter relative humidity (RH2m), 2-meter  
 258 temperature (T2m) and 1-hourly downward global solar radiation (Sol01).

### 259 3.3.3 Sobol' indices estimation

260 For the Sobol' method, a PCE is used to compute analytically Sobol' indices. Because of computational  
 261 cost, storage capacity and time limitations, the 5-order PCE is built directly from the 264 Morris simulations.  
 262 PCEs are independently computed for each scalar model output (mean bias, MAE, RMSE) of each variable  
 263 (ff10m, ffgust, prec01, ... ), season (summer, fall, winter), day (31 for each season) and forecast range (16  
 264 forecast ranges from 3 to 48h lead times). This represents a total of 44 640 PCEs to build. All of these  
 265 surrogates are validated with the cross-validation Leave-One-Out method (Stone, 1974; Geisser, 1975).

266 Figure 2 represents this Leave-One-Out relative emulation error, averaged over different seasons, days  
 267 and forecast ranges, for the different emulated metrics of all meteorological variables. Emulation errors for the  
 268 three metrics are very small, with values between 0.5% and 2.25%. This thus indicates a good performance  
 269 of the PCE method. In more details, the emulations of wind at 10m (ff10m), temperature at 2m (T2m)  
 270 and solar radiation (Sol01) have the smallest errors (less than 1%), whatever the emulated score (mean  
 271 bias, MAE, RMSE). However, stronger errors are detected for the mean bias emulation of wind gust at 10m  
 272 (ffgust), total cloud cover (cloud) and relative humidity at 2m (RH2m). On the other hand, emulation of  
 273 precipitation gives the highest errors with MAE and RMSE.

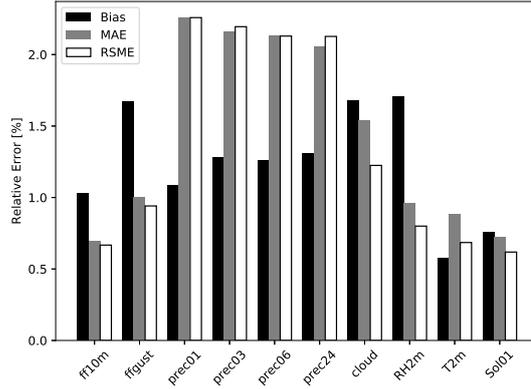


Figure 2: Mean Leave-One-Out relative emulation error averaged over all seasons, days and forecast ranges for different meteorological variables and scalar model outputs. The surrogate errors for mean bias emulation are in black, for MAE emulation in grey and for RMSE emulation in white.

## 4 Results

This section presents sensitivity results from the Morris sensitivity analysis on different variables and model outputs. Different points of view of the parameter sensitivity are considered leading to TMI ( $\sqrt{\mu_i^* + \sigma_i}$ ) calculations with different formulae listed in table 2.

Most sensitivity indices are calculated from the spatially averaged model outputs  $f(d, h) = \frac{1}{X} \sum_x f(x, d, h)$  (formulae 7, 8, 9 and 10). From  $f(d, h)$ , three types of indices are computed :

- Global : represents the mean influence of parameters over a year (summer, fall and winter), it is obtained from an average over seasons, days and forecast ranges
- Seasonal : represents the mean influence of parameters for each season, it is obtained from an average over days and forecast ranges
- Daily : represents the mean influence of parameters for each day of each season, it is obtained from an average over forecast ranges
- Hourly : represents the mean influence of parameters as a function of forecast range, it is averaged over all days of a season

The last type of TMI, hereafter called spatial TMI, allows for a spatial analysis of the influence of parameters. In that case, spatial TMI are computed on spatial model outputs but temporally averaged to reduce computational cost, as defined in 11.

The last paragraph is devoted to results with the Sobol' method, where TSIs' are presented only for the global influence (following formula 7).

### 4.1 Global parameters influence with the Morris design

As a first global insight into the parameters influence, the Total Morris Indices calculated from equation 7 are considered.

Figure 3 shows the sensitivity of each parameter for different meteorological variables. As results for the four scalar outputs (averaged field and the 3 verification scores) are very similar, the sensitivity averaged over these four outputs is presented here and in the remainder of the paper. Furthermore, as TMI values change with the meteorological variable, TMIs are normalised, for each meteorological variable, by their respective maximum value over all parameters. Hence, normalised values vary between 0 and 1, with higher values corresponding to a larger parameter sensitivity.

In the Figure 3, VSIGQSAT, a parameter from the microphysics scheme that controls the cloud formation (especially through subgrid condensation) has a large indirect impact on solar radiation and also on

Parameters influence	$TMI_i$ 's formulae
Global	$TMI_i = \frac{1}{S} \frac{1}{D} \frac{1}{H} \sum_s \sum_d \sum_h \sqrt{\mu^* \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2 + \sigma \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2} \quad (7)$
Seasonal	$TMI_i(s) = \frac{1}{D} \frac{1}{H} \sum_d \sum_h \sqrt{\mu^* \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2 + \sigma \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2} \quad (8)$
Daily	$TMI_i(d, s) = \frac{1}{H} \sum_h \sqrt{\mu^* \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2 + \sigma \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2} \quad (9)$
Hourly	$TMI_i(h, s) = \frac{1}{D} \sum_d \sqrt{\mu^* \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2 + \sigma \left( \frac{1}{X} \sum_x f(x, d, h) \right)^2} \quad (10)$
Spatial	$TMI_i(x, s) = \frac{1}{H} \sum_h \sqrt{\mu^* \left( \frac{1}{D} \sum_d f(x, d, h) \right)^2 + \sigma \left( \frac{1}{D} \sum_d f(x, d, h) \right)^2} \quad (11)$

Table 2: Different TMI definitions with  $f(x, h, d)$ , the model output depending on the space  $x$ , the forecast starting day  $d$  and the forecast range  $h$ .  $\mu^*$  and  $\sigma$  are the Morris parameters calculation,  $S$  is the number of seasons of 31 days,  $D$  is the number of days for each season,  $H$  is the number of forecast ranges and  $X$  is the number of spatial data (number of observations for performance metric (about 1500) or number of grid points for mean forecast (about 200.000)).

304 temperature at 2m through a parasol cooling effect. As a second-order impact, this parameter also indirectly  
305 influences the precipitation (all accumulation ranges) and, to a lesser extent, other meteorological fields.

306 However, as this parameter defines the cloud formation, its influence on total cloud cover is particularly  
307 high and the influence of other parameters, such as RSWINHF, seems negligible on that meteorological field.  
308 Even if RSWINHF is one of the less influential parameter on total cloud cover, it is the second most important  
309 one for the solar radiation influence as it is a parameter involved in the shortwave optical properties. Hence,  
310 the influence of that parameter on clouds is quite subjective and depends on the choice of the meteorological  
311 variable used to study the parameters impact on clouds. As VSIGQSAT has a too strong influence on total  
312 cloud cover, the measure of parameters influence on clouds with total cloud cover is not the most appropriated.  
313 Other meteorological fields, such as the one proposed here (downward solar radiation), must be considered.

314 Regarding turbulence parameters, three of them have a large influence on several meteorological fields :  
315 XCTP, XCEP and XCED. As XCTP occurs in the equations of the correlations between wind and temperature  
316 or vapour fluctuations, it has a strong impact on these meteorological fields, near the surface. XCEP is  
317 involved directly in the calculation of wind fluctuations. Consequently, its impact on wind at 10 meters is  
318 the strongest. Finally, XCED defines the diffusion of the turbulent kinetic energy (TKE). Since the TKE is  
319 proportional to the squared wind fluctuations and so explicitly appears in the wind gusts calculation, XCED  
320 is clearly linked to wind gust calculation. This parameter also impacts the wind at 10 meters. Furthermore,  
321 these three parameters also influence the precipitation field but these impacts are more indirect and not  
322 clearly understood.

323 In the diffusion scheme, SLHDEPSH controls the hydrometeors diffusion. Consequently, this parameter  
324 is known to have a large impact on the precipitation rate. We can notice that this impact becomes stronger  
325 for longer accumulation ranges, compared to other influential parameters on precipitation. This parameter  
326 does not seem to have an influence on other meteorological fields.

327 In the surface scheme, XFRACZ0 controls the orographic drag, and consequently impacts the friction  
328 at the surface and then the near-surface wind.

329 The XCMF parameter from shallow convection scheme seems to have an influence on all variables.  
 330 Since XCMF describes the mass flux at low levels, it impacts the convective clouds formation, their radiative  
 331 impact on the ground and also convective precipitation. As it characterises the convective plumes at the  
 332 bottom level, it defines the wind convergence and the wind force at the surface. This explains why this  
 333 parameter affects winds at 10 meters, accumulated precipitations, solar radiation, temperature and relative  
 334 humidity at 2 meters.

335 We finally notice that all parameters impact the precipitation forecasts. That result shows the impli-  
 336 cation of all parametrizations in the calculation of precipitation.

337 From the above analysis, one can deduce a list of the 8 most influential parameters on all meteorological  
 338 variables : RSWINHF, VSIGQSAT, XCTP, XCEP, XCED, SLHDEPSH, XFRACZ0 and XCMF. These param-  
 339 eters appear in the top 5 ranking for at least one meteorological variable. Furthermore, they are involved in  
 340 all parametrizations.

341 The 13 remaining parameters have only marginal impacts on model outputs. The limited sensitivity of  
 342 AROME forecasts to these parameters can be explained either by their contribution to second-order physical  
 343 phenomena (for example, XCTD is involved in second-order temperature, humidity fluctuations), or possibly  
 344 by a too small variation range.

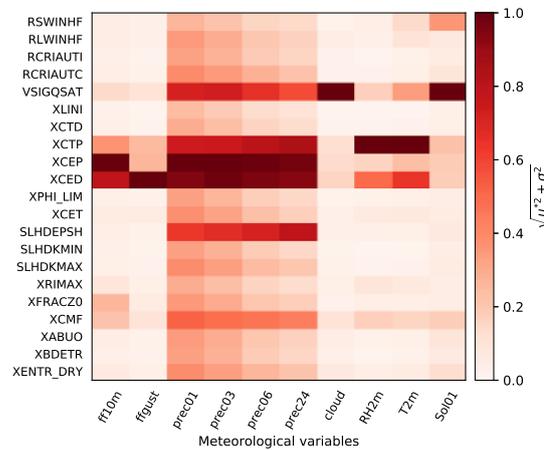


Figure 3: TMI for each parameter as a function of meteorological variable. The higher the sensitivity is, the darker the colour is, the more influential the parameter is.

## 345 4.2 Seasonal parameters influence

346 As a second step, we examine the seasonal dependence of the parameter influence, using seasonal TMI  
 347 calculated with equation 8. Figures 4a and 4b present, respectively for summer and winter, TMIs (averaged  
 348 for all model outputs: Mean Bias, MAE, RMSE, mean field) for each parameter and meteorological field. To  
 349 better compare TMIs for different model outputs and seasons they are normalized, for each meteorological  
 350 field, by the maximum TMI value of the 3 seasons (summer, fall, winter).

351 All the most influential parameters have a strong variability according to season. In particular, we can  
 352 notice that parameters from radiation, microphysics and shallow convection schemes (RSWINHF, VSIGQSAT  
 353 and XCMF) have a larger influence on model outputs during summer (Fig.4a) than in winter (Fig.4b).  
 354 Indeed, during summertime, there are more convective clouds and the radiation intensifies. That is why  
 355 radiating parameters (RSWINHF) are more sensitive to downward solar radiation (Sol01) in summer (Fig.4a).  
 356 Furthermore, as shallow convection and microphysic schemes are more active to represent convective clouds,  
 357 parameters from these parametrization (XCMF and VSIGQSAT) have a stronger influence. Moreover, as  
 358 the convective activity induces higher hydrometeors contents and the kind of associated precipitation is  
 359 mainly convective, hydrometeor contents and consequently precipitations are more smoothed by the diffusion  
 360 scheme in convective situations. This is particularly true as the SLHD scheme takes more action in convective

361 plumes where the wind strain field is more important due to high values of vertical velocity. This explains  
 362 why SLHDEPSH has a stronger influence in summer than in winter.

363 Generally, as there are more physical phenomena involved in summer, all major parameters have a  
 364 larger influence on model outputs during that period.

365 On the other hand, parameters from turbulence schemes (XCTP, XCEP and XCED) and from orographic  
 366 drag representation (XFRACZ0) have a larger influence on wind and wind gust in winter. Hence, the variations  
 367 of wind forecasts performance are mainly related to turbulence parameters. We explain this behaviour by a  
 368 stronger wind shear, in average in the AROME-France domain during wintertime, and also by more frontal  
 369 turbulence and storms on that period.

370 Finally, the 8 most influential parameters identified with the global analysis remain the most influential  
 371 ones for summertime. For wintertime this list can be reduced to only 4 (VSIQOSAT, XCTP, XCEP and  
 372 XCED).

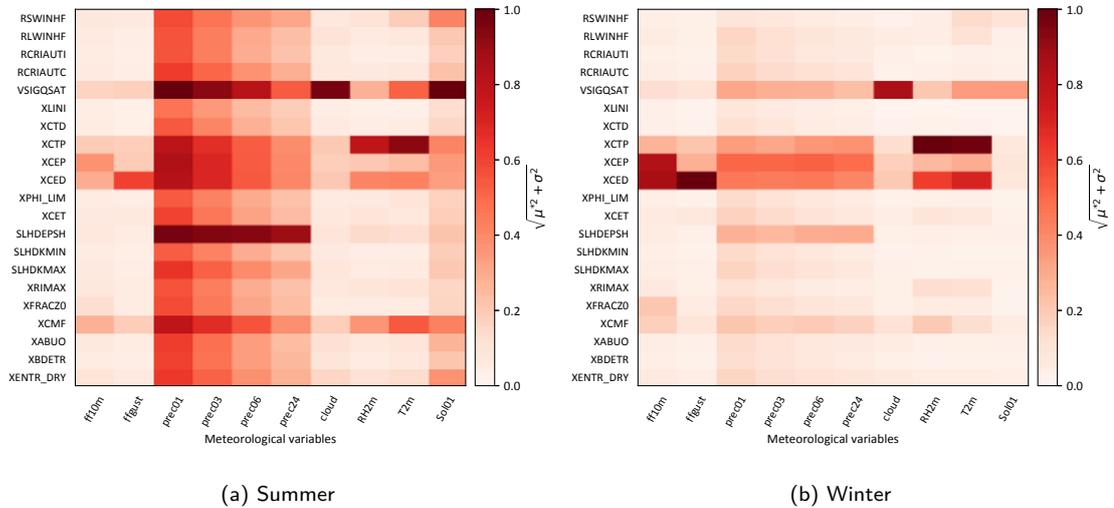


Figure 4: TMI for each parameter as a function of meteorological variable, and for each season: (a) summer and (b) winter. The higher the sensitivity is, the darker the colour is, the more influential the parameter is.

### 373 4.3 Daily parameters influence

374 In this section, we examine the daily variations of the parameters influence over a season. These daily  
 375 results are obtained using equation 10. Because of a larger influence of parameters during summer, we present  
 376 hereafter results only for that period.

377 Figure 5 shows the parameters influence averaged over the different forecast ranges (from 3h to 48h  
 378 lead time), for each day of the summer period. As results are not very sensitive to the choice of the model  
 379 outputs used to compute TMI, results show averaged TMI for all model outputs (Mean Bias, MAE, RMSE,  
 380 mean field). The different values for a given variable are normalized by the corresponding maximum value  
 381 over the considered period.

382 The most influential parameters remain the same over the whole period, but we can notice some  
 383 daily variations in the sensitivity intensity. For example, on Figures 5a and 5b the influence of the most  
 384 influential parameters on wind at 10m and wind gusts (VSIQOSAT, XCTP, XCEP, XCED and XCMF) is  
 385 generally unchanged. However, a larger sensitivity in the forecast of the 27th of July 2018 at 21UTC is  
 386 detected for XCEP and XCED from the boundary layer scheme. On this day, an extratropical cyclone crossed  
 387 the AROME-France domain. Its pressure minimum passed on the North of the domain, near the United  
 388 Kingdom. A cold front crossed France during the 28th and 29th of July 2018. Hence, this stronger influence  
 389 on wind at 10m and wind gust may be due to the gust front associated with the cold front.

390 On Figure 5c, the daily parameters influence on 3-hourly accumulated precipitation is much more  
 391 variable according to the weather situation. In particular, VSIGQSAT and SLHDEPSH have a larger influence  
 392 for forecasts initialised from the 28th of May to the 9th of June and from the 30th of June to the 3rd of July  
 393 2018. These periods correspond to stronger convective situations with many warnings for thunderstorms.  
 394 Thus, the impact of these parameters is clearly linked to the convective activity in the model. On the opposite,  
 395 no significant rainfall was observed between the 12th and the 27th of June, which explains the small influence  
 396 of parameters during that period.

397 The parameters influence clearly depends on the weather situation. The interpretation of the parameters  
 398 sensitivity may change significantly. This shows the necessity to perform sensitivity analysis over large periods  
 399 in order to cover a large number of different meteorological situations and get more robust results.

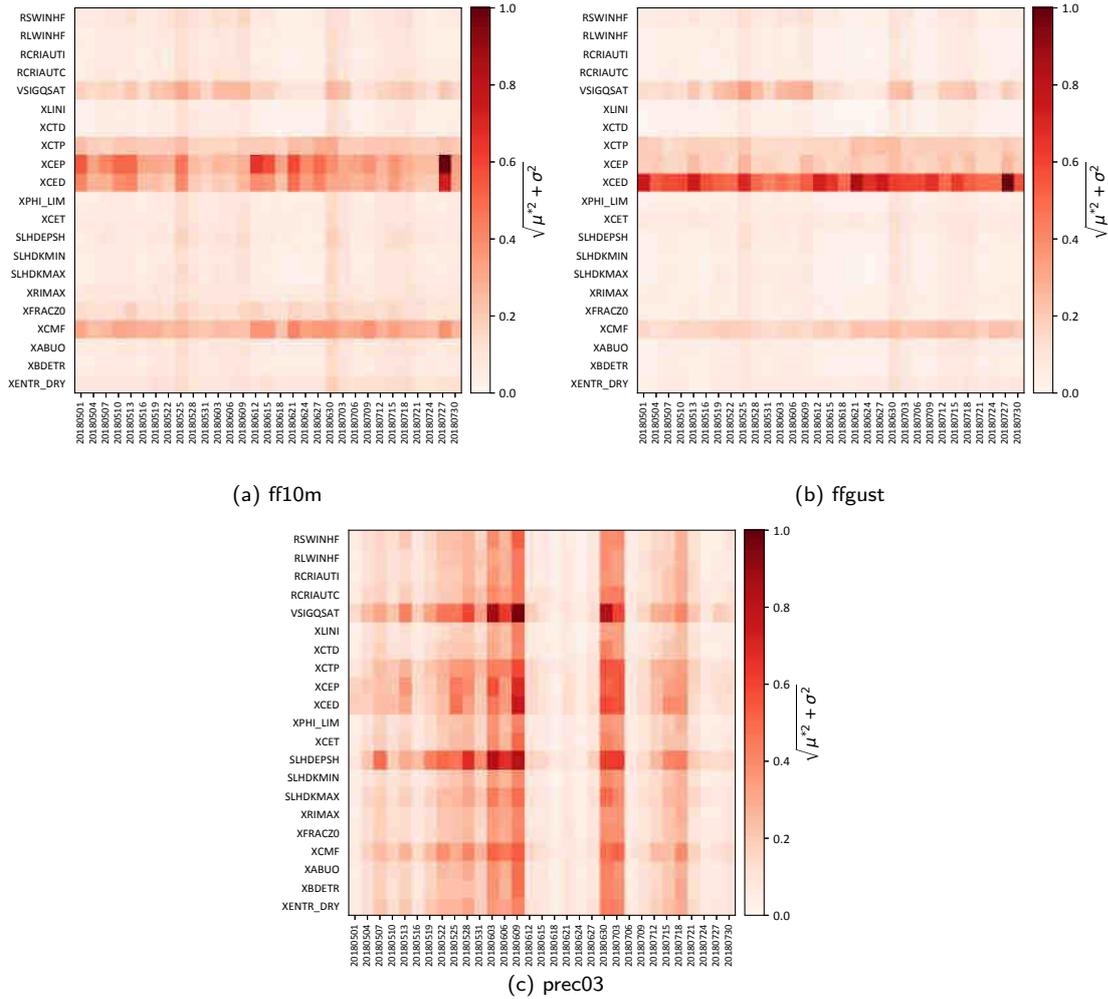


Figure 5: Daily TMI for each day of the summer 2018 period, for each parameter and for (a) wind at 10m, (b) wind gust at 10m and (c) 3h precipitation. The higher the sensitivity is, the darker the colour is, the more influential the parameter is.

#### 4.4 Hourly parameters influence

We present in this section the diurnal cycle of parameters influence on some meteorological variables and for the summer and winter periods. TMI are computed using equation 10. Similarly to previous results, these hourly TMIs are averaged over all model outputs considered, for wind at 10m, 6-hourly accumulated precipitation and temperature at 2 meters in summer (Fig. 6a, b and c) and winter (Fig. 7a, b and c). TMIs are normalized by the maximum value over all forecast ranges. We note that the 0h lead time in Figures 6

406 and 7 corresponds to 21 UTC.

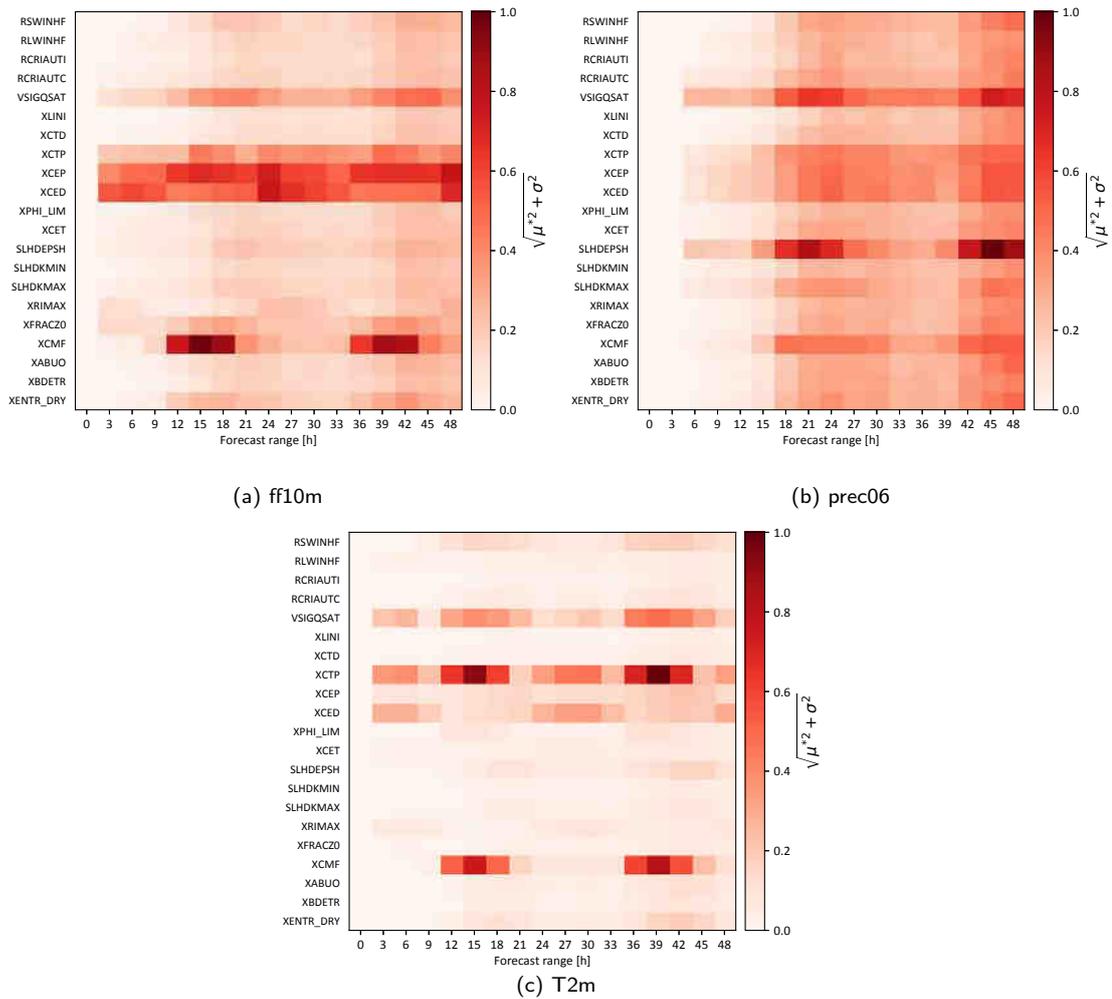


Figure 6: TMI as a function of forecast range, for each parameter during the summer 2018 period. (a) 10-meter wind, (b) 6h accumulated precipitation and (c) temperature at 2m. The higher the sensitivity is, the darker the colour is, the more influential the parameter is.

407 On Figure 6, we can point out that the most influential parameters are consistent with those identified  
 408 in the global sensitivity analysis, namely : RSWINHf, XCTP, XCEP, XCED, XFRACZ0, SLHDEPSH, XCMF  
 409 and VSIGQSAT. However, their effects on model outputs differ according to the lead time. For wind at  
 410 10m (Fig. 6a), VSIGQSAT and XCMF influences exhibit a clear diurnal cycle with an increase between  
 411 12-18h and 36-42h forecast range, which corresponds to the period of convective activity. We notice that  
 412 VSIGQSAT presents a slightly 3-hour temporal shift compared to XCMF. This shift can be explained by the  
 413 time difference between the formation of convective plumes in the boundary layer and the creation of clouds  
 414 at the top of the boundary layer. Turbulence parameters also present a diurnal cycle but it is less noticeable.  
 415 XCEP (wind/pressure correlation) and XCTP (temperature/vapor pressure correlation) affect wind at 10  
 416 meters between 15-18h and 39-42h lead time, during afternoons when the turbulence in the boundary layer  
 417 is more intense and correlations between wind fluctuations are stronger. On the contrary, XCED (diffusion of  
 418 TKE) has a larger sensitivity at night (3-9h and 24-30h lead time) due to the increase of the TKE dissipation  
 419 during that moment of the day.

420 Regarding precipitation (Fig. 7b), all turbulence parameters have the same diurnal cycle with a slightly  
 421 larger influence during the afternoon (18-30h and 42-48h), clearly linked to convective precipitation. We  
 422 can also notice that the sensitivities of SLHDEPSH and VSIGQSAT on precipitation present a strong diurnal  
 423 cycle, particularly correlated with the convection activity (at 18-24h and 42-48h).

424 Concerning temperature at 2 meters (Fig. 6c), the diurnal cycle is clearly more contrasted compared  
 425 to wind at 10 meters and precipitation. XCTP exhibits a smaller but noticeable effect on temperature during  
 426 the night, between 3-6h, 24-30h, in addition to the one observed during the day.

427 We also notice that the influence of parameters linked to the convective activity (VSIGQSAT and  
 428 XCMF) is slightly delayed by 6 hours, according to the meteorological variable: from 12-18h forecast ranges  
 429 for wind at 10m and temperature to 18-24h for 6h-precipitation (for J+1). This is due to the accumulation  
 430 range of precipitation.

431 Regarding the winter period (Fig. 7), consistently with the seasonal analysis, a much smaller influence  
 432 of convection, diffusion and microphysic parameters, such as XCMF, SLHDEPSH and VSIGQSAT, is observed  
 433 for wind at 10 meters, precipitation and temperature at 2 m. The main differences is the absence of diurnal  
 434 cycles particularly for wind at 10m meters (Fig. 7a). A diurnal cycle is still observed for the impact of XCTP  
 435 and XCED on temperature at 2 meters (Fig. 7c), but it is much smaller than in summer (Fig. 6c). Finally,  
 436 no diurnal cycle appears in the influence of parameters on precipitation during winter (Fig. 7b). However,  
 437 for the two seasons, an increase of the parameters impact on precipitation according to forecast lead time is  
 438 noticed (Fig. 6b and 7b), suggesting cumulative effects along forecasts range of the parameters impact.

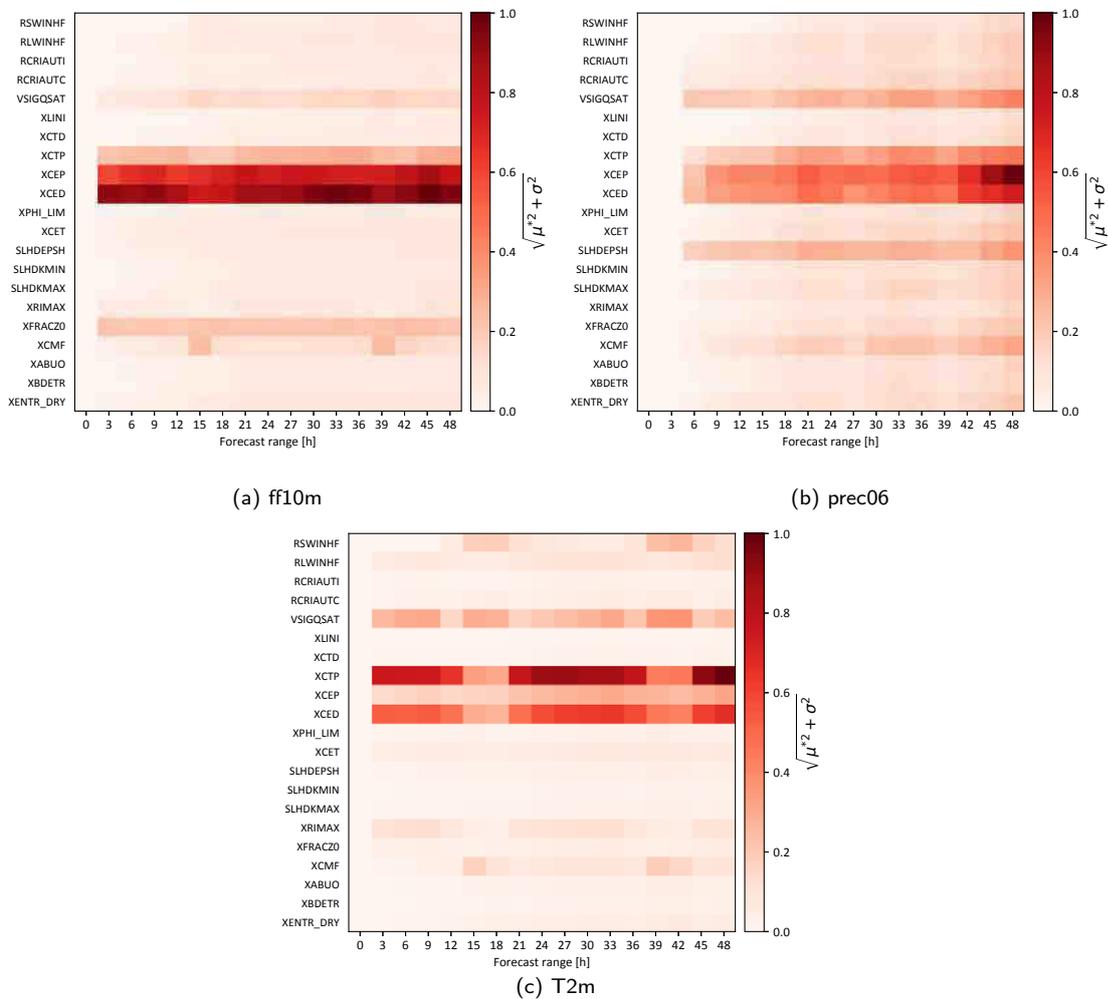


Figure 7: TMI as a function of forecast range, for each parameter during winter 2018-2019. (a) 10-meter wind, (b) 6h accumulated precipitation and (c) temperature at 2m. The higher the sensitivity is, the darker the colour is, the more influential the parameter is.

## 439 4.5 Spatial influence of most influential parameters

440 In order to examine the spatial distribution of parameters sensitivity, figures 8 (a, b), 9(a), 10(a) and  
441 11(a) map the response of seasonal mean forecasts to the parameters variations during summertime. These  
442 TMI are computed following equation 11 and are normalized by the maximum value in the AROME-France  
443 domain.

444 Similar calculations can be performed using forecast verification metrics as scalar outputs. However,  
445 these scores provide a sparse information with only one sensitivity value at each observation location. The  
446 advantage of directly looking at the forecast fields is that it provides an information over the whole AROME-  
447 France domain. Sensitivities are presented for five of the most influential parameters regarding 10-meter  
448 wind, 10-meter wind gust, 6-h precipitation and total cloud cover. As the interpretation of results for the  
449 different seasons is very similar, only the summer period is shown here.

450 In addition to the sensitivity maps, figures 8(c), 9(b), 10(b) and 11(b) also present the AROME forecast  
451 fields, averaged over the summer period and all forecast ranges.

452 Figures 8a and 8b present the sensitivity of the 10-meter wind forecast field to parameters XCEP  
453 (constant for wind-pressure correlations) and XFRACZ0 (coefficient for orographic drag) respectively, during  
454 the summer period. According to previous results, XCEP is certainly one of the most important parameters  
455 for that meteorological variable. Its sensitivity is particularly high in the Channel and in the Gulf of Lion (Fig.  
456 8a). Figure 8c presents the seasonal mean of operational AROME forecast of wind at 10 meters during this  
457 summer, and these two regions correspond to windy zones. Thus, this parameter seems to have a stronger  
458 influence on windy areas. XFRACZ0, used in the SURFEX scheme, is one of the 8 most influential parameters  
459 on wind. Contrary to XCEP, this parameter has a very local impact and its influence does not depend on  
460 seasonal wind. One can see on figure 8b that its impact predominates in orographic areas such as the Alps  
461 or the Pyrenees.

462 XCED was previously identified as the most influential parameter for 10-meter wind gusts. Figure 9a  
463 indicates a clear land/sea contrast of its impact during summer, with higher sensitivity values over land,  
464 where the turbulence is more active than over sea, due to a larger rugosity over land. Thus, no clear link with  
465 the seasonal mean wind gust forecast can be done (Fig. 9b). However, the sensitivity to XCED is locally  
466 enhanced in well-known windy regions such as the Rhône valley.

467 Figure 10a presents the sensitivity of precipitation to the diffusion parameter SLHDEPSH, during  
468 summer. Its sensitivity is clearly correlated with the seasonal field of precipitations (Fig. 10b). For example,  
469 the Western part of France, the north of Spain, the south-west of Germany and the Alps are rainy locations,  
470 where SLHDEPSH has a strong influence. On the opposite, the sensitivity is much more limited in the rainy  
471 region near Ireland, United-Kingdom or Sardinia. This may be due to less convective precipitation but more  
472 stratiform precipitation on these areas.

473 Finally, the sensitivity of total cloud cover to VSIGQSAT is shown in Figure 11a. This parameter has  
474 a stronger influence over seas, which can be explained by a higher relative humidity that favours subgrid  
475 condensation. However, we note a higher influence over the Mediterranean and North seas than over the  
476 Atlantic. This behaviour does not seem to be correlated with the mean total cloud cover field (Fig. 11b). This  
477 behaviour could be explained by a stronger evaporation rate over Mediterranean sea. This is not a surprise  
478 that VSIGQSAT has a high influence over this region: as the relative humidity is higher, the condensation  
479 point is easier to reach. However, for the cold North sea, which is not particularly known to have a high  
480 evaporation rate, the strong influence of VSIGQSAT is not clearly understood.

481 The spatial distribution of parameters sensitivity is overall consistent with the underlying distribution  
482 of weather forecasts. However, this spatial study shows that some parameters can have a spatially wide  
483 influence related to sea/land differences for instance, while others have a more local impact, e.g. in relation  
484 with orographic features.

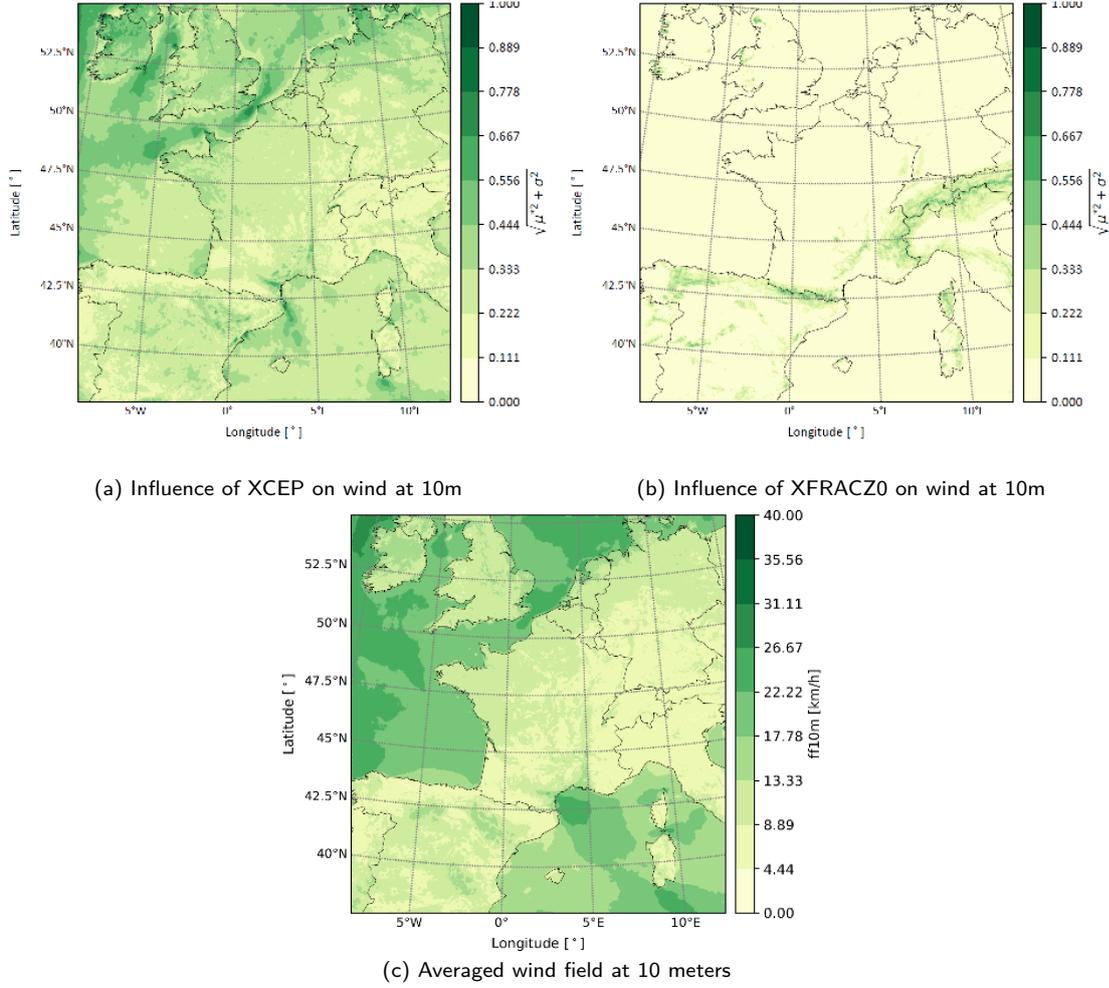


Figure 8: Comparison between influences of XCEP (a) and XCFRACZ0 (b) and the AROME forecasts of wind field at 10 meters, averaged over summer 2018 and forecast ranges (c).

#### 4.6 Interactions between parameters with the Sobol' method

In addition to the Morris SA, a Sobol' analysis has also been conducted. However, because of the high number of meta-models needed, the first-order, second-order and Total Sobol' Indices are analytically calculated from PCE coefficients built only for the spatially-averaged scores as scalar model outputs. Furthermore, the same global averaging (over all days, seasons and forecast ranges) used in section 4.1 is applied to compute the Global Total Sobol' Indices (GTSI):

$$GTSI_i = \frac{1}{S} \frac{1}{D} \frac{1}{H} \sum_s \sum_d \sum_h S_{T,i}(f(d, h)) \quad (12)$$

where  $f(d, h) = \frac{1}{X} \sum_x f(x, d, h)$  is the spatial-averaged scores (Mean Bias, MAE, RMSE), and  $S_{T,i}$  is the Total Sobol Indices analytically calculated from PCE coefficients. Hence, GTSI represents an overall sensitivity measure, that can be compared to the global Total Morris Indices ( $\sqrt{\mu_i^{*2} + \sigma_i^2}$ ) defined in equation 7.

Figure 12 presents the GTSI for each parameter and for different meteorological variables. This figure can directly be compared to Figure 3. Generally, the information provided by GTSI is quite similar to that provided by TMI. However some differences can be noticed. It is interesting to observe that there is a higher contrast between the impact of the most influential parameters and the lesser influential ones. For example, the impact of XCEP on wind at 10 m dominates, in particular in comparison with parameters XCED and XCTP, whose impact was larger according to the Morris sensitivity (Fig. 3). The same remark is true for all meteorological variables and particularly for precipitation: the Morris sensitivity indicated a relatively high

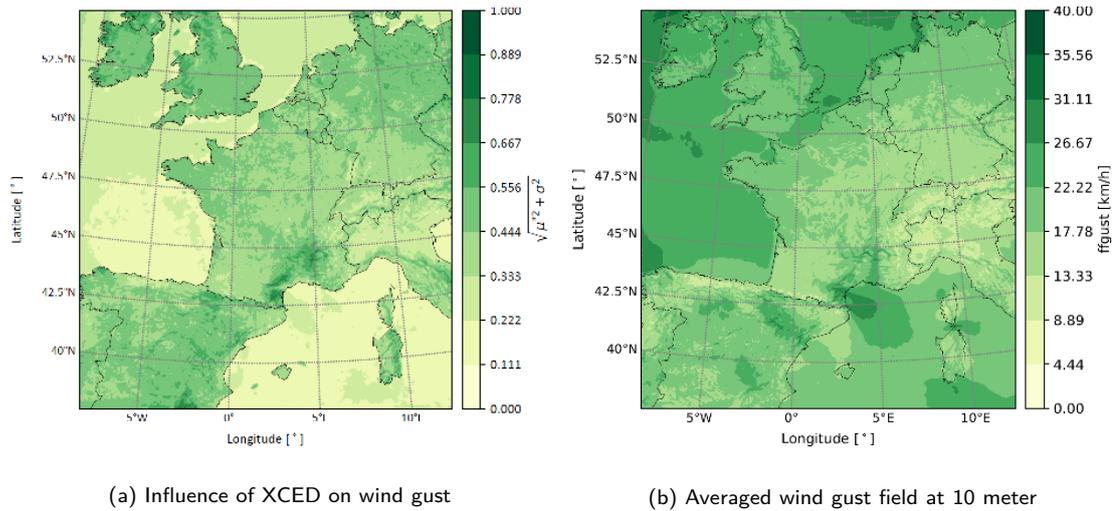


Figure 9: Comparison between the influence of XCED (a) and the AROME forecasts of wind gust field at 10 meters, averaged over summer 2018 and forecast ranges (b).

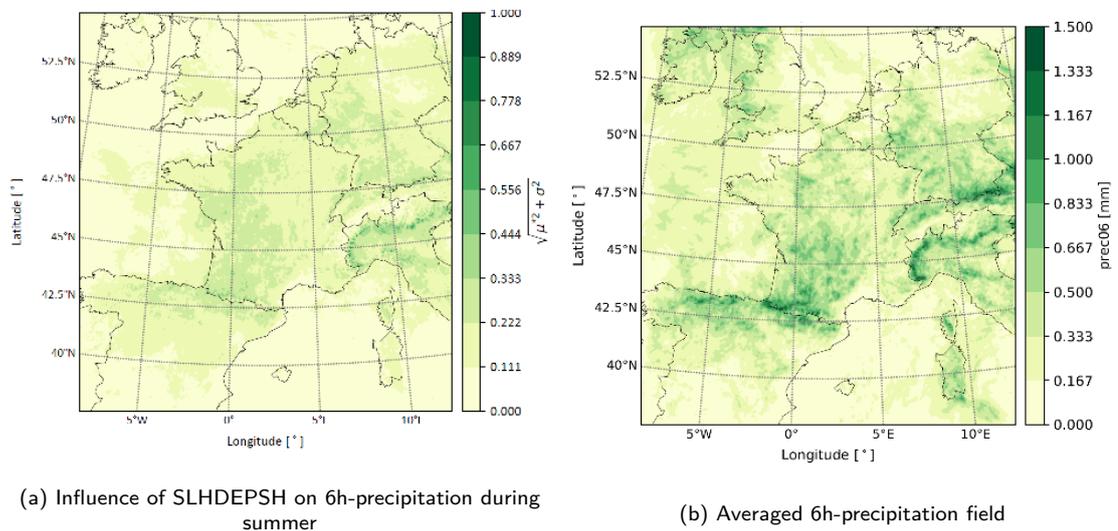


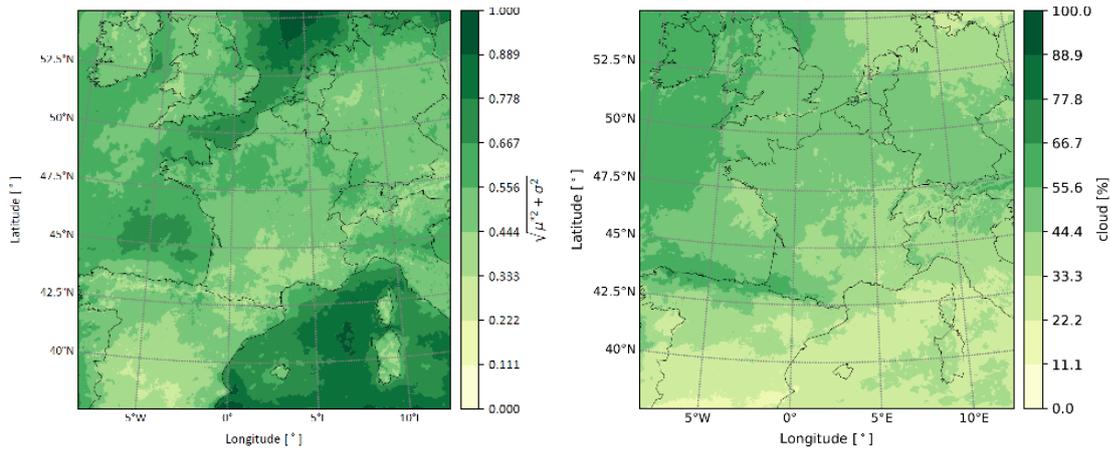
Figure 10: Comparison between the influence of SLHDEPSH (a) and the AROME forecasts of 6h-precipitation field, averaged over summer 2018 and forecast ranges (b).

501 impact of most parameters, while the Sobol indices better highlight the contributions of VSIGQSAT, XCEP  
 502 and SLHDEPSH. Furthermore, the impact of XCEP and XCED on precipitation and solar radiation is smaller  
 503 with the Sobol' method. Finally, the list of the most influential parameters with the Sobol' method may  
 504 change: it suggests to reduce the list to VSIGQSAT, XCTP, XCEP, XCED and SLHDEPSH.

505 The good agreement between Morris and Sobol' results tend to confirm the robustness of the PCEs.  
 506 It is then interesting to look at the additional information about interactions between parameters, provided  
 507 by the second-order Sobol' indices. It is worth mentioning, however, that given the relatively basic sampling  
 508 used for training the PCEs, these interactions may be only partially estimated.

509 Figure 13 presents the second-order Sobol' indices for each couple of parameters and for different  
 510 meteorological variables. Similarly to the Global Total Sobol' Indices, these second-order indices are computed  
 511 from the spatially-averaged scores and are then temporally averaged as in equation 12.

512 First of all, we remark numerous interactions for wind gusts and 6h-precipitation suggesting high  
 513 interactions between parametrizations, as previously noticed in section 4.1 (Fig. 13b and 13c). On the other  
 514 hand, only few strong interactions are noticeable for 10-meter wind, total cloud cover, relative humidity and



(a) Influence of VSIGQSAT on total cloud cover

(b) Averaged total cloud cover field

Figure 11: Comparison between the influence of VSIGQSAT (a) and the AROME forecasts of total cloud cover, averaged over summer 2018 and forecast ranges.

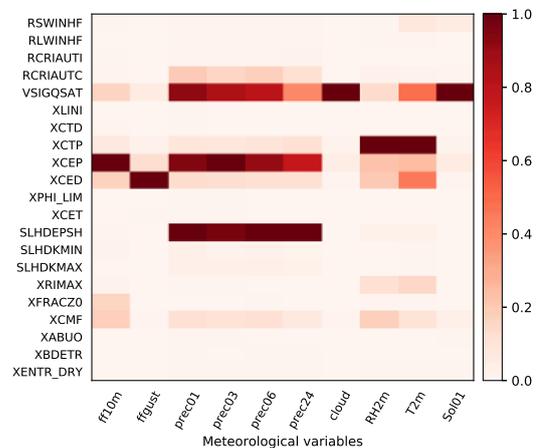


Figure 12: Total Sobol' indices, for each parameter as a function of meteorological variable. The higher the Total Sobol' Indice is, the darker the colour is, the more influential the parameter is.

515 temperature.

516 For wind gusts (Fig. 13b), interactions occur mainly between parameters from the turbulence scheme.  
 517 Two interactions seem particularly important: XCED (diffusion of TKE) with XCEP (wind/pressure correla-  
 518 tion) and with XCTP (temperature/vapour pressure correlation). As XCEP and XCTP are involved indirectly  
 519 in the wind gust calculation, which strongly depends on XCED, these interactions are not surprising. We  
 520 also remark that the interaction between XCED and XCEP also appears for 6h-precipitation (Fig. 13c),  
 521 temperature at 2 meters (Fig. 13f) and particularly for 10 meters wind (Fig. 13a). Indeed, it involves XCED  
 522 which occurs in the wind gust calculation and XCEP in the wind correlation. The interaction between XCED  
 523 and XCTP for wind gusts (Fig. 13b) also appears for relative humidity and temperature (Fig. 13e and 13f).  
 524 For these two meteorological variables, the interaction between XCTP and XCMF is also quite important.  
 525 This shows the link between the turbulence and the shallow convection schemes.

526 The most influential parameter on total cloud cover is VSIGQSAT, and this parameter is in interaction  
 527 with several others (Fig. 13d), in particular with XCEP. This interaction also appears for the impact on  
 528 precipitation (Fig. 13c).

529 Some other interactions can be noticed such as (1) RSWINHF with XCEP (resp. XCED) on the  
 530 precipitation (Fig. 13c) (resp. wind gust (Fig. 13b)) forecasts, (2) XCET with XCMF and XCED on the  
 531 wind gust forecasts (Fig. 13b) and (3) VSIGQSAT with XCTP on the total cloud cover forecasts (Fig. 13d).

532 For downward solar radiation (not shown) there is only one interaction between RSWINHF and VSIGQSAT.  
533 This result is not surprising since this meteorological variable clearly depends on the solar optical properties  
534 altered by RSWINHF and on the cloud parasol effect linked to the cloud formation controlled by VSIGQSAT.

535 Finally, we can notice that XLINI (minimum mixing length), which has one of the smallest direct  
536 influence on AROME forecasts, seems to have a significant indirect impact through its interaction with  
537 XCEP on the wind at 10 meters. This interaction also appears for the precipitation forecasts (Fig. 13c).

538 Thus, regarding Figure 13, five main interactions are detected for the influence on different meteo-  
539 rological variables : XCED-XCEP, XCED-XCTP, XCMF-XCTP, VSIGQSAT-XCEP and XLINI-XCEP. These  
540 interactions mainly concern the 8 most influential parameters except the last one.

## 541 **5 Outlooks and conclusions**

542 In this study, a sensitivity analysis of the French convective-scale AROME model to a set of 21 param-  
543 eters from physical and dynamical schemes has been presented and discussed. The Morris screening, which  
544 perturbs one parameter at a time, has been successfully applied to several meteorological variables and scalar  
545 model outputs, in order to examine the parameters influence on different spatial and temporal scales. In  
546 addition, the Sobol' method has mostly confirmed the Morris results.

547 The global parameters influence, averaged over approximately a year of data, allows us to identify  
548 eight major parameters, namely, RSWINHF, VSIGQSAT, XCTP, XCEP, XCED, SLHDEPSH, XFRACZ0 and  
549 XCMF, that have a noticeable impact on different near-surface meteorological variables, especially wind and  
550 precipitation. Three parameters in this list occur in the turbulence scheme (XCTP, XCEP and XCED) and  
551 affect almost all weather variables, which highlights the importance of properly tuning this parametrization.  
552 VSIGQSAT, a parameter from the subgrid condensation scheme, such as XCMF from shallow convection  
553 scheme, have large direct or indirect impacts on most meteorological variables. These two parameters  
554 seem to be relevant for the cloud representation in the AROME model. Finally, parameters RSWINHF,  
555 SLHDEPSH and XFRACZ0 are in the top 5 of most influential parameters for at least one meteorological  
556 variables. RSWINHF, which controls the short-wave inhomogeneity factor, is obviously one of the most  
557 sensitive parameter on solar radiation at the surface. SLHDEPSH has a strong influence only on precipitation.  
558 XFRACZ0, from SURFEX scheme, particularly affect the wind at 10 meters. The Sobol' method suggests to  
559 reduce the list to five parameters: VSIGQSAT, XCTP, XCEP, XCED and SLHDEPSH.

560 A deeper investigation of sensitivity results also revealed interesting variations on different temporal  
561 scales. In particular, a diurnal cycle is observed for several parameters in the summer period, and is closely  
562 related to convective activity. It is also shown that convection-related parameters have a larger influence  
563 in summer. Spatial sensitivities also indicate that the spatial distribution of parameters influence is closely  
564 linked to the underlying distribution of the forecast field. Finally, the Sobol' sensitivity analysis gives a first  
565 insight into the interactions between parameters.

566 This work provides a helpful insight into the black-box of parametrizations and may serve as a baseline  
567 for different applications. For instance, the tuning of parameters values for deterministic forecasting could  
568 primarily focus on the most influential parameters identified. In addition, this study can be seen as a  
569 preliminary step towards the implementation of perturbed parameter approaches in EPS in order to define  
570 the most important parameters to perturb. This aspect will be further explored in a forthcoming paper.  
571 Moreover, the daily variations of parameters impacts suggest that it is necessary to perturb parameters  
572 during the integration of each forecast. More specifically, a time-scale of 6h is recommended. In the same  
573 way, the spatial study indicates the necessity to perturb parameters in space. Spatial scale must cover all the  
574 scales between the land/sea (synoptic) to plain/relief scale (local).

575 Finally, some influences of parameters are still not clearly understood. A more detailed study on  
576 physical processes based on the impact on physical tendencies could help to better understand impacts and  
577 interactions of these parameters.

578 **Acknowledgements** The authors would like to thank Agnès Lagnoux, Clémentine Prieur, Robert Faivre  
579 and Bruno Sudret for their helps and advices on sensitivity analysis, but also Rachel Honnert for its expertise  
580 on the turbulence scheme. Our sincere gratitude also goes to Inger-Lise Frogner, the HIRLAM-EPS team,  
581 François Bouttier and Philippe Arbogast for their suggestions and remarks.

## 582 References

- 583 F. M. Alam, K. R. McNaught, and T. J. Ringrose. Using morris' randomized oat design as a factor screen-  
584 ing method for developing simulation metamodels. In *Proceedings of the 36th Conference on Winter*  
585 *Simulation*, WSC '04, page 949–957. Winter Simulation Conference, 2004. ISBN 0780387864.
- 586 L. A. Bastidas, T. S. Hogue, S. Sorooshian, H. V. Gupta, and W. J. Shuttleworth. Parameter sensitivity  
587 analysis for different complexity land surface models using multicriteria methods. *Journal of Geophysical*  
588 *Research: Atmospheres*, 111(D20), 2006. ISSN 2156-2202. doi: 10.1029/2005JD006377.
- 589 P. Bénard, J. Vivoda, J. Mašek, P. Smolikova, K. Yessad, C. Smith, R. Brozkova, and J.-F. Geleyn. Dynamical  
590 kernel of the aladin-nh spectral limited-area model: Revised formulation and sensitivity experiments. *Quart.*  
591 *J. Roy.Meteor. Soc.*, 136:155 – 169, 01 2010. doi: 10.1002/qj.522.
- 592 Y. Bouteloup, Y. Seity, and E. Bazile. Description of the sedimentation scheme used operationally in all  
593 météo-france nwp models. *Tellus A*, 63:300 – 311, 03 2011. doi: 10.1111/j.1600-0870.2010.00484.x.
- 594 F. Bouttier, B. Vié, O. Nuisser, and L. Raynaud. Impact of stochastic physics in a convection-permitting  
595 ensemble. *Mon. Wea. Rev.*, 140:3706–3721, 11 2012. doi: 10.1175/MWR-D-12-00031.1.
- 596 F. Branger, L.-G. Giraudet, C. Guivarch, and P. Quirion. Global sensitivity analysis of an energy–economy  
597 model of the residential building sector. *Environmental Modelling and Software*, 70:45 – 54, 2015. ISSN  
598 1364-8152. doi: 10.1016/j.envsoft.2015.03.021.
- 599 P. Brousseau, Y. Seity, D. Ricard, and J. Léger. Improvement of the forecast of con-  
600 vective activity from the arome-france system. *Quarterly Journal of the Royal Meteor-*  
601 *ological Society*, 142(699):2231–2243, 2016. doi: <https://doi.org/10.1002/qj.2822>. URL  
602 <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2822>.
- 603 R. Buizza et al. Stochastic representation of model uncertainties in the ecmwf ensemble prediction system.  
604 *Quart. J. Roy.Meteor. Soc.*, 125(560):2887–2908, 1999. doi: 10.1002/qj.49712556006.
- 605 F. Campolongo, J. Cariboni, and A. Saltelli. An effective screening design for sensitivity analy-  
606 sis of large models. *Environmental Modelling and Software*, 22:1509–1518, January 2007. doi:  
607 doi:10.1016/j.envsoft.2006.10.004.
- 608 G. Chastaing. *Generalized Sobol sensitivity indices for dependent variables*. Theses, Université de Grenoble,  
609 Sept. 2013.
- 610 C. Ciric. Extrapolation des données de biotests pour la protection des communautés par une approche couplant  
611 biotests de laboratoire, études sur mésocosme et modélisation écologique. diploma thesis, Université Claude  
612 Bernard Lyon 1, 2012.
- 613 P. Clark, N. Roberts, H. W. Lean, S. Ballard, and C. Charlton-Perez. Convection-permitting models: a step-  
614 change in rainfall forecasting. *Meteorological Applications*, 23:165–181, 2016. doi: 10.1002/met.1538.
- 615 R. Confalonieri, G. Bellocchi, S. Bregaglio, M. Donatelli, and M. Acutis. Comparison of sensitivity analysis  
616 techniques: A case study with the rice model warm. *Ecological Modelling*, 221:1897–1906, 08 2010a. doi:  
617 10.1016/j.ecolmodel.2010.04.021.
- 618 R. Confalonieri, G. Bellocchi, S. Tarantola, M. Acutis, M. Donatelli, and G. Genovese. Sensitivity analysis  
619 of the rice model warm in europe: Exploring the effects of different locations, climates and methods of  
620 analysis on model sensitivity to crop parameters. *Environmental Modelling and Software*, 25:479–488,  
621 2010b. doi: 10.1016/j.envsoft.2009.10.005.

- 622 A. Cosenza, G. Mannina, P. A. Vanrolleghem, and M. B. Neumann. Global sensitivity analysis in wastewater  
623 applications: A comprehensive comparison of different methods. *Environmental Modelling and Software*,  
624 49:40–52, 2013. doi: 10.1016/j.envsoft.2013.07.009.
- 625 P. Courtier, C. Freyrier, J.-F. Geleyn, F. Rabier, and M. Rochas. The arpege project at meteo france. In  
626 *Seminar on Numerical Methods in Atmospheric Models, 9-13 September 1991*, volume II, pages 193–232,  
627 Shinfield Park, Reading, 1991. ECMWF, ECMWF. URL <https://www.ecmwf.int/node/8798>.
- 628 F. Couvreur, F. Hourdin, D. Williamson, R. Roehrig, V. Volodina, N. Villefranche, C. Rio, O. Audouin1,  
629 J. Salter, E. Bazile, F. Brient, F. Favot, R. Honnert, M.-P. Lefebvre, J.-B. Madeleine, Q. Rodier, and W. Xu.  
630 Process-based climate model development harnessing machine learning: I. a calibration tool for parameter-  
631 ization improvement. *Journal of Advances in Modeling Earth Systems*, n/a(n/a):e2020MS002217, 2020.  
632 doi: 10.1029/2020MS002217. e2020MS002217 2020MS002217.
- 633 C. Covey, D. D. Lucas, J. Tannahill, X. Garaizar, and R. Klein. Efficient screening of climate model sensitivity  
634 to a large number of perturbed input parameters. *Journal of Advances in Modeling Earth Systems*, 5(3):  
635 598–610, 2013. doi: 10.1002/jame.20040.
- 636 R. Cropp and R. Braddock. The new morris method: an efficient second-order screening method. *Reliability*  
637 *Engineering and System Safety*, 78:77–83, 2002. doi: 10.1016/S0951-8320(02)00109-6.
- 638 J. Cuxart, P. Bougeault, and J.-L. Redelsperger. A turbulence scheme allowing for mesoscale and large-eddy  
639 simulations. *Quart. J. Roy. Meteor. Soc.*, 126(562):1–30, 2000. doi: 10.1002/qj.49712656202.
- 640 Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, J. Quan, J. Li, C. Miao, A. Ye, and C. Tong. Assessing  
641 wrf model parameter sensitivity: A case study with 5 day summer precipitation forecasting in the greater  
642 beijing area. *Geophys. Res. Lett.*, 42(2):579–587, 2015. doi: 10.1002/2014GL061623.
- 643 Z. Di, Q. Duan, W. Gong, A. Ye, and C. Miao. Parametric sensitivity analysis of precipitation and temperature  
644 based on multi-uncertainty quantification methods in the weather research and forecasting model. *Science*  
645 *China Earth Sciences*, 60:876–898, 03 2017. doi: 10.1007/s11430-016-9021-6.
- 646 M. Déqué, C. Drevet, A. Braun, and D. Cariolle. The arpege/ifs atmosphere model: a contribution to the  
647 french community climate modelling. *Clim. Dynam.*, 10:249–266, 1994. doi: 10.1007/BF00208992. URL  
648 <https://doi.org/10.1007/BF00208992>.
- 649 R. Faivre, B. looss, S. Mahévas, D. Makowski, and H. Monod. *Analyse de sensibilité et exploration de modèles:*  
650 *Application aux sciences de la nature et de l'environnement*. Collection Savoir-faire. Quae éditions, 2016.  
651 ISBN 9782759219063.
- 652 Y. Fouquart and B. Bonnel. Computations of solar heating of the earth's atmosphere—a new parameteriza-  
653 tion. *Beitrag zur Physik der Atmosphäre*, 53:35–62, 1980.
- 654 A. Franczyk. Using the morris sensitivity analysis method to assess the importance of input variables on  
655 time-reversal imaging of seismic sources. *Acta Geophysica*, 67:1525–1533, 2019. ISSN 1895-7455. doi:  
656 10.1007/s11600-019-00356-5.
- 657 I.-L. Frogner, A. T. Singleton, M. O. Køltzow, and U. Andrae. Convection-permitting ensembles: Challenges  
658 related to their design and use. *Quart. J. Roy. Meteor. Soc.*, 145(S1):90–106, 2019. doi: 10.1002/qj.3525.
- 659 F. Gamboa, A. Janon, T. Klein, and A. Lagnoux. Sensitivity analysis for multidimensional and functional  
660 outputs. *Electronic Journal of Statistics*, 8, 11 2013. doi: 10.1214/14-EJS895.
- 661 Q. Ge and M. Menendez. Extending morris method for qualitative global sensitivity analysis of models with de-  
662 pendent inputs. *Reliability Engineering and System Safety*, 162, 01 2017. doi: 10.1016/j.res.2017.01.010.
- 663 C. Gebhardt, S. Theis, P. Krahe, and V. Renner. Experimental ensemble forecasts of precipitation based on  
664 a convection-resolving model. *Atmospheric Science Letters*, 9(2):67–72, 2008. doi: 10.1002/asl.177.
- 665 S. Geisser. The predictive sample reuse method with applications. *Journal of American Statistical Association*,  
666 70(350), 1975. doi: 10.2307/2285815.
- 667 S. Girard, V. Mallet, I. Korsakissok, and A. Mathieu. Emulation and sobol' sensitivity analysis of an at-

668 mospheric dispersion model applied to the fukushima nuclear accident. *Journal of Geophysical Research:*  
669 *Atmospheres*, 121(7):3484–3496, 2016. doi: 10.1002/2015JD023993.

670 S. Hagelin, J. Son, R. Swinbank, A. McCabe, N. Roberts, and W. Tennant. The met office convective-scale  
671 ensemble, mogreps-uk. *Quart. J. Roy.Meteor. Soc.*, 143(708):2846–2861, 2017. doi: 10.1002/qj.3135.

672 T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability*  
673 *Engineering and System Safety*, 52(1):1 – 17, 1996. ISSN 0951-8320. doi: 10.1016/0951-8320(96)00002-6.

674 Z. Hou, M. Huang, R. L. L. G. Lin, and D. M. Ricciuto. Sensitivity of surface flux simulations to hydrologic pa-  
675 rameters based on an uncertainty quantification framework applied to the Community Land Model. *Journal*  
676 *of Geophysical Research: Atmospheres*, 117(D15), 2012. ISSN 2156-2202. doi: 10.1029/2012JD017521.

677 B. looss. Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la Societe Française*  
678 *de Statistique*, 152(1):1–23, 2011.

679 J. Jacques. Pratique de l'analyse de sensibilité : comment évaluer l'impact des entrées aléatoires sur la sortie  
680 d'un modèle mathématique. *Pub. IRMA*, 71(III), 2011.

681 J. P. Lafore, J. Stein, N. Asencio, P. Bougeault, V. Ducrocq, J. Duron, C. Fischer, P. Hérelil, P. Mascart,  
682 V. Masson, J. P. Pinty, J. L. Redelsperger, E. Richard, and J. Vilà-Guerau de Arellano. The meso-nh  
683 atmospheric simulation system. part i: adiabatic formulation and control simulations. *Annales Geophysicae*,  
684 16(1):90–109, 1998. doi: 10.1007/s00585-997-0090-6.

685 P. Le Moigne, A. Boone, J.-C. Calvet, B. Decharme, S. Faroux, A.-L. Gibelin, C. Lebeaupin, J.-F. Mahfouf,  
686 E. Martin, V. Masson, D. Mironov, J. Noilhan, P. Tulet, and B. Van Den Hurk. Surfex scientific docu-  
687 mentation. Technical report, Météo-France, 2009. Note de centre du Groupe de Météorologie à Moyenne  
688 Echelle, available at: <http://www.cnrm.meteo.fr/surfex/>.

689 J. Li, Q. Y. Duan, W. Gong, A. Ye, Y. Dai, C. Miao, Z. Di, C. Tong, and Y. Sun. Assessing parameter  
690 importance of the common land model based on qualitative and quantitative sensitivity analysis. *Hydrology*  
691 *and Earth System Sciences*, 17(8):3279–3293, 2013. doi: 10.5194/hess-17-3279-2013.

692 K. G. Link, M. T. Stobb, J. Di Paola, K. B. Neeves, A. L. Fogelson, S. S. Sindi, and K. Leiderman. A local  
693 and global sensitivity analysis of a mathematical model of coagulation and platelet deposition under flow.  
694 *PLoS ONE*, 13(7):1–38, 07 2018. doi: 10.1371/journal.pone.0200917.

695 Y. Liu, H. Gupta, S. Sorooshian, L. Bastidas, and W. Shuttleworth. Exploring parameter sensitivities of the  
696 land surface using a locally coupled land-atmosphere model. *Journal of Geophysical Research: Atmospheres*,  
697 109:21101–, 11 2004. doi: 10.1029/2004JD004730.

698 D. D. Lucas and R. G. Prinn. Parametric sensitivity and uncertainty analysis of dimethylsulfide oxidation in  
699 the clear-sky remote marine boundary layer. *Atmospheric Chemistry and Physics*, 5(6):1505–1525, 2005.  
700 doi: 10.5194/acp-5-1505-2005.

701 C. Marzban, S. Sandgathe, J. Doyle, and N. Lederer. Variance-based sensitivity analysis: Preliminary results  
702 in coamps. *Mon. Wea. Rev.*, 142:2028–2042, 05 2014. doi: 10.1175/MWR-D-13-00195.1.

703 A. McCabe, R. Swinbank, W. Tennant, and A. Lock. Representing model uncertainty in the met office  
704 convection-permitting ensemble prediction system and its impact on fog forecasting. *Quart. J. Roy.Meteor.*  
705 *Soc.*, 142(700):2897–2910, 2016. doi: 10.1002/qj.2876.

706 E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough. Radiative transfer for inho-  
707 mogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave. *Journal of Geophysical*  
708 *Research: Atmospheres*, 102(D14):16663–16682, 1997. doi: 10.1029/97JD00237.

709 M. D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33:  
710 161–174, May 1991. doi: 10.2307/1269043.

711 J. Nossent, P. Elsen, and W. Bauwens. Sobol' sensitivity analysis of a complex environmental model. *Envi-*  
712 *ronmental Modelling and Software - ENVSOFT*, 26, 12 2011. doi: 10.1016/j.envsoft.2011.08.010.

713 P. Ollinaho, M. Laine, A. Solonen, H. Haario, and H. Järvinen. Nwp model forecast skill optimization via clo-

714 sure parameter variations. *Quart. J. Roy. Meteor. Soc.*, 139(675):1520–1532, 2013. doi: 10.1002/qj.2044.

715 J. Pergaud, V. Masson, S. Malardel, and F. Couvreur. A parameterization of dry thermals and shallow cumuli  
716 for mesoscale numerical weather prediction. *Bound. Lay. Meteorol.*, 132(1):83, 2009. doi: 10.1007/s10546-  
717 009-9388-0.

718 J. Pinty and P. Jabouille. A mixed-phased cloud parameterization for use in a mesoscale non-hydrostatic  
719 model: simulations of a squall line and of orographic precipitation. In A. M. Society, editor, *Preprints of*  
720 *Conference On Cloud Physics*, pages 217–220, Everett, WA, 1998.

721 Y. Qian, H. Yan, Z. Hou, G. Johannesson, S. Klein, D. Lucas, R. Neale, P. Rasch, L. Swiler, J. Tannahill,  
722 H. Wang, M. Wang, and C. Zhao. Parametric sensitivity analysis of precipitation at global and local  
723 scales in the community atmosphere model cam5. *Journal of Advances in Modeling Earth Systems*, 7(2):  
724 382–411, 2015. doi: 10.1002/2014MS000354.

725 J. Quan, Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, A. Ye, and C. Miao. An evaluation of parametric  
726 sensitivities of different meteorological variables simulated by the wrf model. *Quart. J. Roy. Meteor. Soc.*,  
727 142(700):2925–2934, 2016. doi: 10.1002/qj.2885.

728 J. Ren, W. Zhang, and J. Yang. Morris sensitivity analysis for hydrothermal coupling parameters of embank-  
729 ment dam: A case study. *Mathematical Problems in Engineering*, 2019, 2019. doi: 10.1155/2019/2196578.

730 T. A. Rieutord. *Sensitivity analysis of a filtering algorithm for wind lidar measurements*. Theses, Institut  
731 National Polytechnique de Toulouse (INP Toulouse), Nov. 2017.

732 G. S. Romine, C. S. Schwartz, J. Berner, K. R. Fossell, C. Snyder, J. L. Anderson, and M. L. Weisman.  
733 Representing forecast error in a convection-permitting ensemble system. *Mon. Wea. Rev.*, 142(12):4519 –  
734 4541, 2014. doi: 10.1175/MWR-D-14-00100.1.

735 M. V. Ruano, J. Ribes, J. Ferrer, and G. Sin. Application of the morris method for screening the influential  
736 parameters of fuzzy controllers applied to wastewater treatment plants. *Water Science and Technology*,  
737 63(10):2199–2206, 05 2011. doi: 10.2166/wst.2011.442.

738 A. Saltelli, K. Chan, and E. Scott. *Sensitivity Analysis*, volume 134. John Wiley and Sons, 01 2000. ISBN  
739 0471998923.

740 D. G. Sanchez, B. Lacarrière, M. Musy, and B. Bourges. Application of sensitivity analysis in building energy  
741 simulations: Combining first- and second-order elementary effects methods. *Energy and Buildings*, 68:741  
742 – 750, 2014. ISSN 0378-7788. doi: 10.1016/j.enbuild.2012.08.048.

743 F. Sarrazin, F. Pianosi, and T. Wagener. Global sensitivity analysis of environmental models: Convergence  
744 and validation. *Environmental Modelling and Software*, 79:135 – 152, 2016. ISSN 1364-8152. doi:  
745 10.1016/j.envsoft.2016.02.005.

746 Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The  
747 arome-france convective-scale operational model. *Mon. Wea. Rev.*, 139(3):976–991, 2011. doi:  
748 10.1175/2010MWR3425.1.

749 Y. Shi, W. Gong, Q. Duan, J. Charles, C. Xiao, and H. Wang. How parameter specification of an earth  
750 system model of intermediate complexity influences its climate simulations. *Progress in Earth and Planetary*  
751 *Science*, 6(1):46–64, 2019. ISSN 2197-4284. doi: 10.1186/s40645-019-0294-x.

752 G. Shutts. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J.*  
753 *Roy. Meteor. Soc.*, 131(612):3079–3102, 2005. doi: 10.1256/qj.04.106.

754 I. M. Sobol'. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie*,  
755 2(1):112–118, 1990. doi: 10.1016/S0378-4754(00)00270-6. (In russian).

756 D. Stainforth, T. Aina, C. Christensen, M. Collins, N. Faull, D. Frame, J. Kettleborough, S. Knight, A. Mar-  
757 tin, J. Murphy, C. Piani, D. Sexton, L. Smith, R. Spicer, A. Thorpe, and M. Allen. Uncertainty in  
758 predictions of the climate response to rising levels of greenhouse gases. *Nat.*, 433:403–6, 02 2005. doi:  
759 10.1038/nature03301.

- 760 D. Stensrud. *Parameterization schemes: Keys to understanding numerical weather prediction models*, volume  
761 9780521865401. Cambridge University Press, United Kingdom, Jan. 2011. ISBN 0521865409. doi:  
762 10.1017/CBO9780511812590. Publisher Copyright: © D. Stensrud 2007. Copyright: Copyright 2017  
763 Elsevier B.V., All rights reserved.
- 764 M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical*  
765 *Society: Series B (Methodological)*, 36(2):111–133, 1974. doi: 10.1111/j.2517-6161.1974.tb00994.x.
- 766 B. Sudret. Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering and System*  
767 *Safety*, 93:964–979, 07 2008. doi: 10.1016/j.ress.2007.04.002.
- 768 J. Tardieu and M. Leroy. Radome, le réseau temps réel d'observation au sol de météo-france. *La Météorologie*,  
769 40:40–43, 2003.
- 770 A. J. van Rensburg, G. van Schoor, and P. van Vuuren. Stepwise global sensitivity analysis of a physics-based  
771 battery model using the morris method and monte carlo experiments. *Journal of Energy Storage*, 25:  
772 100875, 2019. ISSN 2352-152X. doi: 10.1016/j.est.2019.100875.
- 773 P. Vanrolleghem, G. Mannina, A. Cosenza, and M. Neumann. Global sensitivity analysis for urban water  
774 quality modelling: Terminology, convergence and comparison of different methods. *Journal of Hydrology*,  
775 522, 03 2015. doi: 10.1016/j.jhydrol.2014.12.056.
- 776 F. Váña, P. Bénard, J.-F. Geleyn, A. Simon, and Y. Seity. Semi-lagrangian advection scheme with controlled  
777 damping: An alternative to nonlinear horizontal diffusion in a numerical weather prediction model. *Quart.*  
778 *J. Roy.Meteor. Soc.*, 134(631):523–537, 2008. doi: 10.1002/qj.220.
- 779 Y. Wang et al. The central european limited-area ensemble forecasting system: Aladin-laef. *Quart. J.*  
780 *Roy.Meteor. Soc.*, 137(655):483–502, 2011. doi: 10.1002/qj.751.
- 781 J. Wu, R. Dhingra, M. Gambhir, and J. Remais. Sensitivity analysis of infectious disease models: Methods,  
782 advances and their application. *Journal of the Royal Society, Interface*, 10:20121018, 06 2013. doi:  
783 10.1098/rsif.2012.1018.
- 784 S. Yang, W. Tian, E. Cubi, Q. Meng, Y. Liu, and L. Wei. Comparison of sensitivity analysis methods in building  
785 energy assessment. *Procedia Engineering*, 146:174–181, 12 2016. doi: 10.1016/j.proeng.2016.06.369.
- 786 Y. Zhang, S. Xie, C. Covey, D. D. Lucas, P. Gleckler, S. A. Klein, J. Tannahill, C. Doutriaux, and R. Klein.  
787 Regional assessment of the parameter-dependent performance of CAM4 in simulating tropical clouds.  
788 *Geophys. Res. Lett.*, 39(14), 2012. ISSN 1944-8007. doi: 10.1029/2012GL052184.

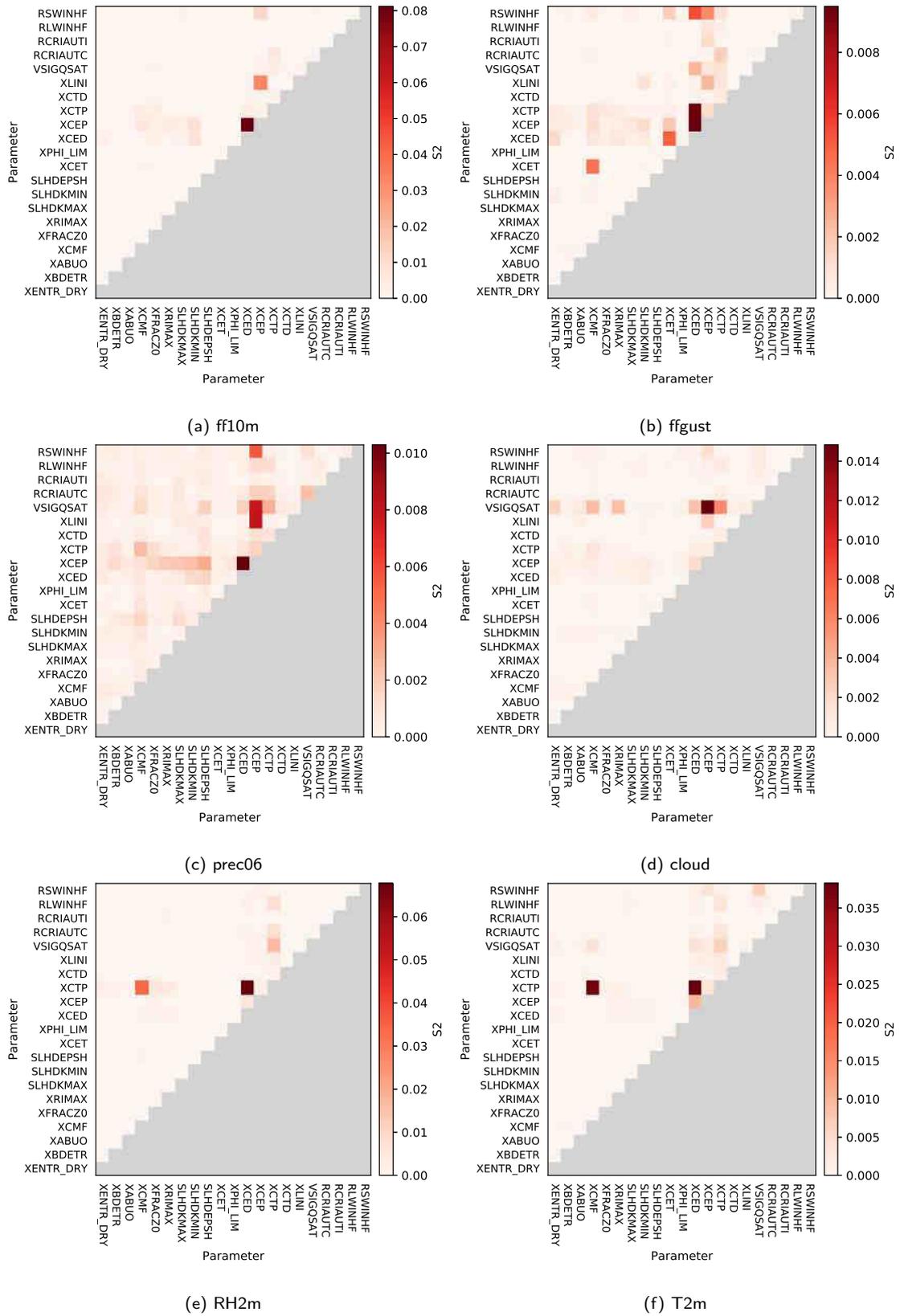


Figure 13: Second-order Sobol' indices calculated for (a) wind at 10 meters, (b) 10m-wind gusts, (c) 6h-precipitation, (d) total cloud cover, (e) relative humidity at 2m and (f) temperature at 2m. The higher the Sobol' indices is, the darker the colour is, the higher the interaction between parameters is.

**Conclusion :** Les résultats des différentes analyses de sensibilité montrant bien une sensibilité des prévisions du modèle AROME aux différents paramètres identifiés par les experts, ceux-ci présentent alors une incertitude qu'il faut représenter dans la PEARO. L'idée de perturber ces paramètres pour chacun des membres de la PEARO est donc justifiée pour représenter une partie de l'erreur de modélisation. Ceci fera alors l'objet de la partie III.

## Bibliographie

M. Wimmer, L. Raynaud, L. Descamps, L. Berre, and Y. Seity. Sensitivity analysis of the convective-scale arome model to physical and dynamical parameters. *Quarterly Journal of the Royal Meteorological Society*, n/a(n/a). doi : <https://doi.org/10.1002/qj.4239>. URL <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.4239>.

## Troisième partie

# Représentation des erreurs de modélisation dans la PEARO



# Chapitre 7

## Optimisation des méthodes de perturbations de paramètres

Après avoir justifié, dans la précédente partie, l’emploi de la perturbation de paramètres pour représenter l’erreur de modélisation, nous allons à présent nous focaliser sur cette représentation-même de l’erreur de modélisation dans la PEARO. Tout d’abord, dans ce chapitre, nous allons nous intéresser aux différentes méthodes de perturbation de paramètres mises en place pour représenter l’erreur de modélisation dans la PEARO. Pour cela, nous nous sommes d’abord attachés à la méthode « Perturbed Parameters », en réutilisant les simulations effectuées lors de l’analyse de sensibilité de Morris. Une optimisation de cette méthode selon le score probabiliste CRPS a été proposée. La méthode « Random Perturbed Parameter », prenant en compte la variabilité journalière, a aussi été implémentée et optimisée. Enfin, une réduction du nombre de paramètres à perturber, à partir des résultats de l’analyse de Morris, a été testée. Les différentes méthodes présentées ici feront l’objet d’une évaluation dans le chapitre 8.

### 7.1 Échantillon de 1000 PP à partir du jeu de données de l’analyse de sensibilité

De par la construction même du plan d’expérience de l’analyse de sensibilité de Morris, la base de données produite pour cette étude comprend des simulations différant uniquement par la valeur des paramètres : pour chaque trajectoire de Morris  $r$ , deux simulations diffèrent uniquement par la valeur d’un paramètre. Avec les  $r = 12$  trajectoires de Morris utilisées et les  $k = 21$  paramètres testés, cette base de données comprend  $N_M = r(k + 1) = 264$  simulations disponibles. Celles-ci peuvent être assimilées à une prévision d’ensemble à 264 membres où seule l’erreur de modélisation est représentée par la perturbation de paramètres, méthode dite « Perturbed Parameters » (PP). Puisque l’objectif de l’analyse de sensibilité est d’étudier l’impact seul des paramètres sur les prévisions, les conditions initiales, de surface et latérales n’ont pas été perturbées. Il n’y a donc pas de représentation de l’incertitude de ces conditions dans cet ensemble dit PP.

En parallèle, la méthode SPPT actuellement implémentée dans la PEARO

comprend  $N_{Mb} = 12$  membres. Afin de comparer proprement cette méthode au PP comprenant 264 membres, le même nombre d'échantillons de la distribution de probabilité de l'état de l'atmosphère, donc le même nombre de membres, est nécessaire. Il faut donc créer un sous-échantillon de 12 simulations parmi les 264 du plan d'expérience de Morris pour ainsi créer un PP avec 12 membres. Afin d'examiner la variabilité des performances obtenues avec différentes valeurs de paramètres, il faut alors réaliser plusieurs prévisions d'ensemble PP différant uniquement par la valeur des paramètres conformément aux 264 simulations de Morris.

Une première solution est de réaliser toutes les combinaisons possibles de  $N_{Mb} = 12$  prévisions parmi les  $N_M = 264$  originales du plan d'expérience de Morris. Or, ceci demande près de  $10^{11}$  milliards de combinaisons :

$$\binom{N_M}{N_{Mb}} = \frac{N_M!}{N_{Mb}!(N_M - N_{Mb})!} = \frac{264!}{12! \times (264 - 12)!} \approx 1.85 \times 10^{20} \quad (7.1)$$

Pour des raisons de coût de calcul, il est impossible de réaliser toutes ces combinaisons. Nous proposons donc d'utiliser une méthode de sous-échantillonnage pour sélectionner uniquement quelques combinaisons.

En répétant 1000 fois un tirage aléatoire avec remise de 12 simulations parmi les 264 du plan d'expérience de Morris, on construit 1000 prévisions d'ensemble de 12 membres où seule l'erreur de modélisation est représentée par la méthode PP (voir figure 7.1).

Pour chacun des membres des 1000 ensembles PP, nous disposons de la valeur des paramètres correspondant mais aussi des prévisions et des écarts aux observations ( $y-o$ ). L'analyse de sensibilité de Morris ayant été effectuée pour trois saisons de 31 jours, définies dans le tableau 5.2, les 1000 ensembles utilisant la méthode PP fournissent ainsi des prévisions et erreurs aux observations correspondantes, sur ces mêmes périodes, soit un total de 93 jours.

A partir des écarts aux observations, les différents scores probabilistes définis en section 2.6 ont été calculés pour chacun des 1000 ensembles et ce, pour les trois saisons précédemment définies. Nous appellerons par la suite « Mean PP » la prévision d'ensemble fictive correspondant aux scores probabilistes moyennés sur les 1000 ensembles.

## 7.2 Optimisation de la méthode PP selon le CRPS

Les 1000 PP ayant des scores probabilistes différents, il peut être intéressant de chercher celui offrant les meilleurs scores. Comme proposé par Díaz-Isaac et al. (2019) et Houtekamer et al. (2021), nous avons optimisé la méthode PP en cherchant la prévision d'ensemble présentant le meilleur score probabiliste CRPS parmi les 1000 PP (Figure 7.1). Ce score calculant la performance d'un ensemble pour tous seuils, il peut être vu comme une mesure de la performance globale d'une prévision d'ensemble. En effet, pouvant être décomposé selon la fiabilité et la résolution, ce score a l'avantage de prendre en compte les deux caractéristiques principales d'une prévision d'ensemble. Il est donc un bon candidat pour l'optimisation d'un tel ensemble. D'autres scores peuvent également être utilisés pour optimiser une prévision d'ensemble. Par exemple, Díaz-Isaac et al. (2019) ont uti-

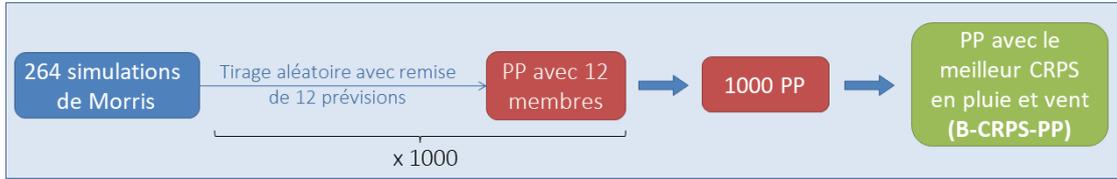


FIGURE 7.1 – Schéma synthétisant l'échantillonnage des 1000 PP à partir du plan d'expérience de Morris ainsi que l'optimisation de la méthode PP selon le CRPS de pluie et de vent

lisé l'écart au diagramme de rang idéalement plat. Cependant, ce score ne rend pas compte de la résolution de l'ensemble. Ekblom et al. (2020) se sont, eux, basés sur le rapport dispersion/erreur. Par ailleurs, notons que d'autres méthodes, telles que l'algorithme génétique (Díaz-Isaac et al., 2019; Houtekamer et al., 2021) ou le « simulated annealing » (Díaz-Isaac et al., 2019), non considérées ici mais déjà appliquées à la prévision d'ensemble, permettent aussi d'estimer la prévision d'ensemble optimale.

Le CRPS étant calculé pour chacune des échéances  $h$ , il faut donc chercher à minimiser ce score orienté négativement pour toutes les échéances, soit minimiser la quantité  $I_{CRPS}$  définie par :

$$I_{CRPS} = \sum_{t=0}^{N_{ech}} CRPS(h) \quad (7.2)$$

où  $N_{ech}$  correspond au nombre d'échéances. Cependant, la force du vent et des pluies étant particulièrement exploitées avec la PEARO, il est suggéré de réaliser une optimisation dite à plusieurs objectifs, c'est-à-dire une optimisation à la fois sur le CRPS de la force de vent à 10m et des précipitations cumulées pendant 3h.

De telles méthodes d'optimisation de paramètres à plusieurs objectifs ont déjà été utilisées pour des modèles déterministes. A titre d'exemple, Gong et al. (2015) et Duan et al. (2017) proposent diverses combinaisons de scores de différentes variables pour prendre en compte les différents objectifs. Gong et al. (2015), dans le cas de l'optimisation de paramètres d'un modèle de surface déterministe, se sont appuyés sur la RMSE de  $N_{var} = 6$  variables de surface, alors que Duan et al. (2017) ont optimisé la MAE de pression et température ( $N_{var} = 2$ ), à tous les niveaux du modèle de prévision numérique du temps WRF. Malgré leurs applications différentes, Gong et al. (2015) et Duan et al. (2017) montrent tous deux que la combinaison  $F$  de scores normalisés par un score par défaut prend mieux en compte l'optimisation sur les différents objectifs (ex : pour le cas de la RMSE) :

$$F = \sum_{v=1}^{N_{var}} \frac{RMSE_v}{RMSE_{v,defaut}} \quad (7.3)$$

où  $v$  itère sur les  $N_{var}$  objectifs,  $RMSE_{v,defaut}$  étant la RMSE pour les valeurs par défaut des paramètres.

Dans le même ordre d'idée, nous proposons ici d'optimiser la somme pondérée

par le maximum des CRPS pour  $N_{var} = 2$  objectifs (vent à 10m et pluie sur 3h) :

$$F = \sum_{v=1}^{N_{var}} \frac{I_{CRPS_v}}{\max(CRPS_v(t))} \quad (7.4)$$

où  $v$  correspond à la pluie pour  $v = 1$  et à la force du vent à 10m pour  $v = 2$ . Contrairement à Gong et al. (2015) et Duan et al. (2017), nous normalisons ici par le maximum de CRPS, sur tous les PP et sur les différentes échéances (afin que les valeurs de CRPS soient comprises entre 0 et 1 pour le vent et la pluie). Les valeurs de CRPS de vent à 10m étant presque d'un ordre de grandeur plus grand que ceux de la pluie, cette normalisation permet de ramener les différents CRPS dans la même gamme de valeur et d'ainsi attribuer un poids similaire au CRPS de vent et de pluie dans l'optimisation.

Ainsi pour chaque PP, la quantité  $F$  a été calculée grâce aux CRPS disponibles pour chaque saison. Le PP présentant le minimum de  $F$  est sélectionné et dit par la suite « Best CRPS PP » (B-CRPS-PP). Pour l'été, il s'agit du PP n°966, que l'on nommera Be-CRPS-PP. Pour l'automne et l'hiver, celui-ci est identique et correspond au PP n°426 que l'on différenciera et notera Ba-CRPS-PP pour l'automne et Bh-CRPS-PP pour l'hiver.

Un PP optimal sur les trois saisons (considérées comme représentant une année) peut aussi être identifié. Celui-ci, par la suite nommé B-CRPS-PP, minimise la somme des  $F$  pour chaque saison ( $F_{ete} + F_{automne} + F_{hiver}$ ) et, après calcul, correspond exactement au Ba-CRPS-PP et Bh-CRPS-PP. L'automne étant une saison intermédiaire, celle-ci peut biaiser l'estimation du PP optimal sur l'année. Le même calcul a alors été réalisé sur les deux saisons extrêmes ( $F_{ete} + F_{hiver}$ ) mais le PP désigné comme le B-CRPS-PP est aussi le n°426 comme pour le Ba-CRPS-PP et le Bh-CRPS-PP.

Notons qu'avec une deuxième méthode de calcul de  $F$  définie comme le produit des CRPS de pluie et vent ( $I_{CRPS_{ff10m}} \times I_{CRPS_{prec03}}$ ) mais non retenue pour l'étude (car nous supposons qu'elle ne prend pas autant en compte le CRPS de la force du vent et celui de la pluie), le PP optimal sur l'année correspondait aussi à celui de l'automne (n°636), mais était différent de celui de l'été (n°966) et de l'hiver (n°426). Ces résultats semblent suggérer que l'estimation du PP optimal sur l'année peut être réalisée uniquement sur l'automne, saison intermédiaire où les épisodes de fortes précipitations et de vent sont nombreux. Cependant, il faudrait vérifier cette supposition en testant sur plusieurs années différentes, mais ceci n'est pas l'objet de ce travail.

### 7.3 Optimisation de la méthode RPP

Afin de prendre également en compte la variabilité de l'incertitude d'une date à l'autre, la méthode Random Perturbed Parameter (RPP) a aussi été implémentée dans la PEARO. Celle-ci tire aléatoirement, pour chaque prévision, de nouvelles valeurs de paramètres dans les intervalles de perturbations définis par les physiciens (voir tableau 5.1). Une première configuration de la méthode utilise un tirage aléatoire suivant une distribution uniforme afin qu'aucune valeur ne soit privilégiée. Nous appellerons cette configuration simplement RPP.

Pour optimiser cette méthode de représentation de l'erreur de modélisation, d'autres configurations peuvent être employées avec des distributions de probabilité des paramètres privilégiant les valeurs de paramètres donnant les meilleurs scores. Nous appellerons par la suite ces ensembles optimisés, « oRPP ». Tout comme pour le B-CRPS-PP, nous choisissons ici d'optimiser la méthode selon le score de CRPS.

Ainsi, au total, sept configurations différentes de RPP ont été testées, utilisant chacune des distributions de paramètres différentes. De la plus simple à la plus compliquée, ces configurations, présentées par la suite, utilisent les distributions suivantes :

- RPP : distributions uniformes pour tous les paramètres et membres de l'ensemble ;
- oRPP gaussien : distributions gaussiennes centrées autour des valeurs des paramètres du B-CRPS-PP pour chaque membre. Ces distributions sont tronquées pour rester dans les intervalles de perturbation de paramètres proposés par les physiciens ;
- oRPP B-CRPS-PP : distributions à partir de l'ensemble des valeurs prises par les paramètres du B-CRPS-PP. Ces distributions sont identiques pour chaque membre ;
- oRPP 50B-CRPS-PP : distributions à partir de l'ensemble des valeurs prises par les paramètres des 50 B-CRPS-PP. Ces distributions sont identiques pour chaque membre ;
- oRPP Morris : distributions à partir de l'ensemble des valeurs prises par les paramètres dans le plan d'expérience de Morris. Ces distributions sont identiques pour chaque membre ;
- oRPP 50B-eCRPS-PP : distributions à partir de l'ensemble des valeurs prises par les paramètres des 50 B-CRPS-PP **émulés**. Ces distributions sont identiques pour chaque membre.

### 7.3.1 A partir de distributions gaussiennes

Le B-CRPS-PP (optimisé sur l'année) correspondant à des valeurs de paramètres bien précises pour chaque membre, et les paramètres interagissant ensemble, il est nécessaire de garder ces jeux précis de valeurs pour chaque membre. En supposant que le B-CRPS-PP n'est pas le meilleur possible puisque toutes les valeurs de paramètres n'ont pas été testées en raison de la discrétisation du plan d'expérience de Morris à l'origine de ce PP, nous pouvons faire l'hypothèse qu'un meilleur PP que le B-CRPS-PP peut être proche du B-CRPS-PP déjà identifié. Une solution d'optimisation du RPP est donc d'échantillonner des valeurs de paramètres proches de celles du B-CRPS-PP.

Ainsi, nous appelons par la suite « oRPP Gaussien », la méthode RPP qui tire aléatoirement à chaque prévision, des valeurs de paramètres selon des distributions de probabilité gaussiennes centrées autour des valeurs du B-CRPS-PP (voir Figures 7.2). Nous avons choisi une distribution de probabilité gaussienne dont

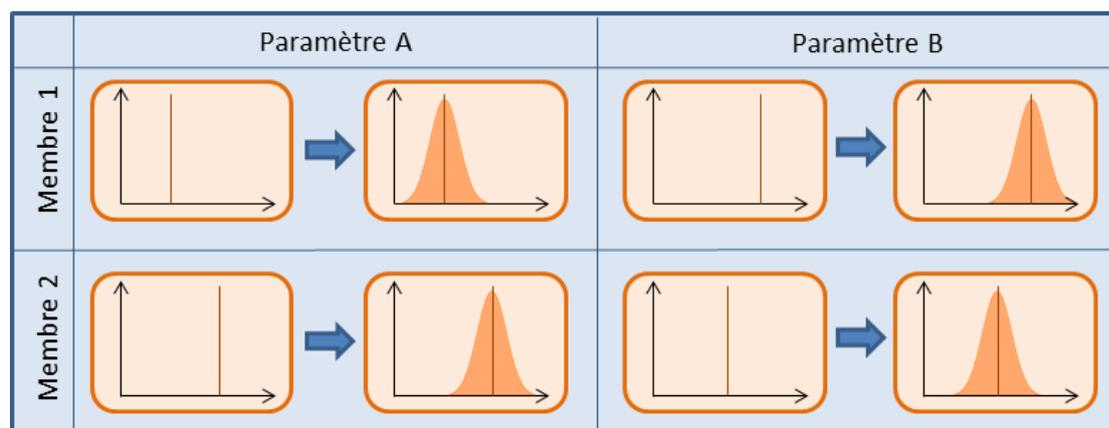


FIGURE 7.2 – Schéma explicatif du oRPP gaussien.

la moyenne correspond aux valeurs des paramètres du B-CRPS-PP. Concernant son écart-type, cette distribution doit être suffisamment large pour apporter une variabilité non négligeable, mais suffisamment fine pour ne pas trop s'éloigner du B-CRPS-PP et donc perdre l'optimisation ainsi que l'effet des interactions entre paramètres. La discrétisation des paramètres étant en  $p = 8$  points (discrétisation du plan d'expérience de Morris) et chaque valeur de paramètre étant séparée par un intervalle  $I$ , nous avons choisi un écart-type de  $I/2$ . La figure 7.3 montre un tirage aléatoire de 1000 valeurs selon ces distributions pour les paramètres des membres 1 et 2 du oRPP gaussien.

### 7.3.2 A partir des distributions du B-CRPS-PP

Le B-CRPS-PP étant composé de 12 membres dont chacun correspond à des valeurs spécifiques de paramètres, une distribution de probabilité peut être établie pour chaque paramètre à partir de leurs 12 valeurs. Contrairement aux distributions de probabilité du oRPP gaussien, les distributions obtenues à partir de ces 12 valeurs sont uniques pour chaque paramètre et identiques pour chaque membre de l'ensemble. Par ailleurs, celles-ci ne suivent pas de distributions gaussiennes : la Figure 7.5 montre ces histogrammes pour chaque paramètre et aucun ne correspond à un type de distribution de probabilité classique. Notons que la verticale rouge indique la valeur par défaut des paramètres dans le modèle AROME, qui n'est pas nécessairement situées dans le maximum de probabilité.

Afin d'optimiser le RPP à partir des distributions du B-CRPS-PP, il faut alors créer un générateur de nombre aléatoire spécifique pour chacune de ces distributions. Cependant, pour ce faire, il est nécessaire d'obtenir des distributions de probabilité continues. Le jeu de données à l'origine des PP utilisant le plan d'expérience de Morris n'est pas continu : seules  $p = 8$  valeurs de paramètres sont échantillonnées ce qui rend les distributions de probabilité des paramètres fortement discrétisées. Une estimation par noyau (dite Méthode de Parzen) est alors utilisée afin d'évaluer au plus juste la distribution de probabilité continue correspondante :

$$f_k(x) = \frac{1}{N_{Mb}h} \sum_{i=1}^{N_{Mb}} K\left(\frac{x - x_i}{h}\right) \quad (7.5)$$

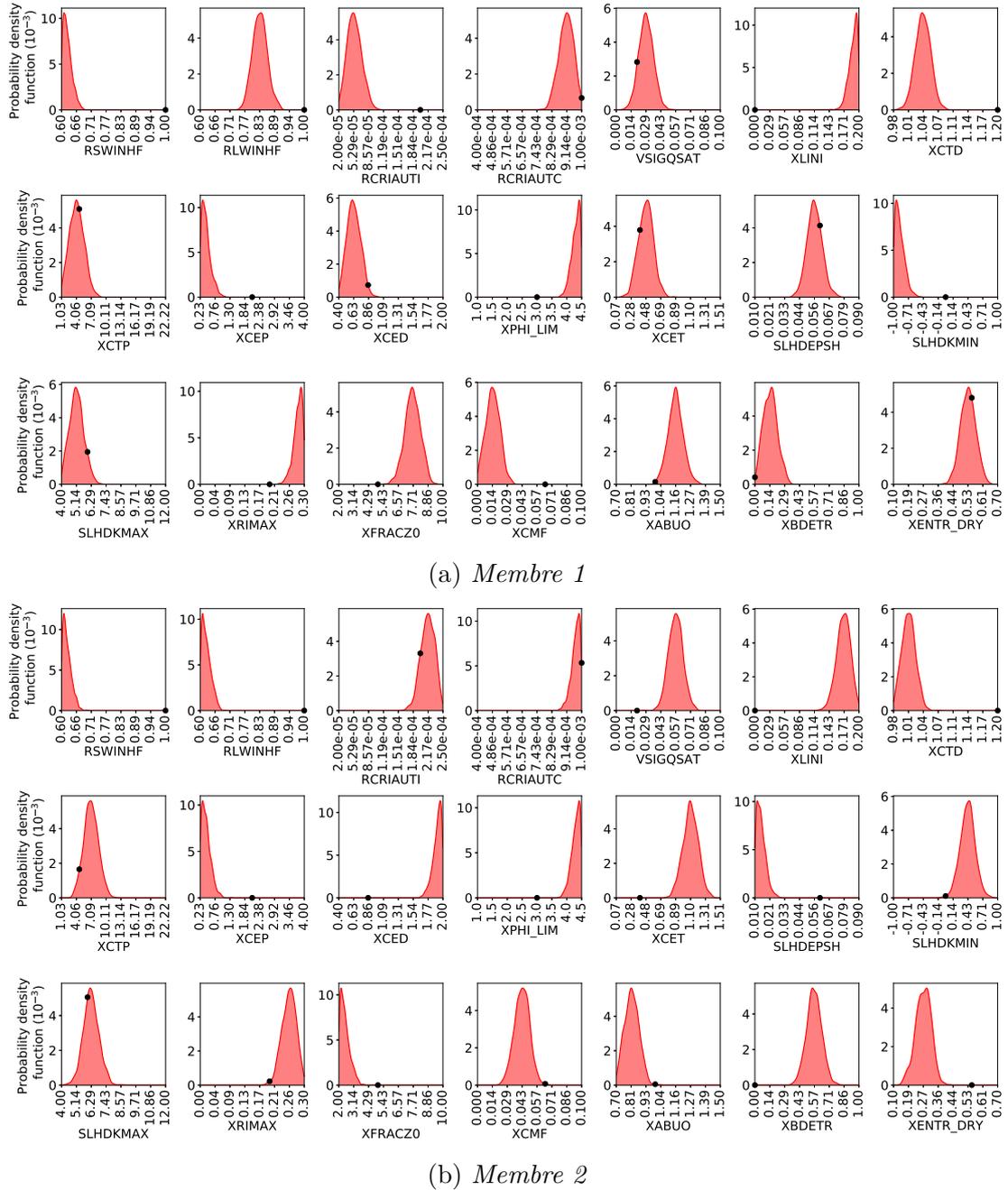


FIGURE 7.3 – Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le membre 1 (a) et le membre 2 (b) du oRPP gaussien. Le point noir représente la valeur par défaut des paramètres.

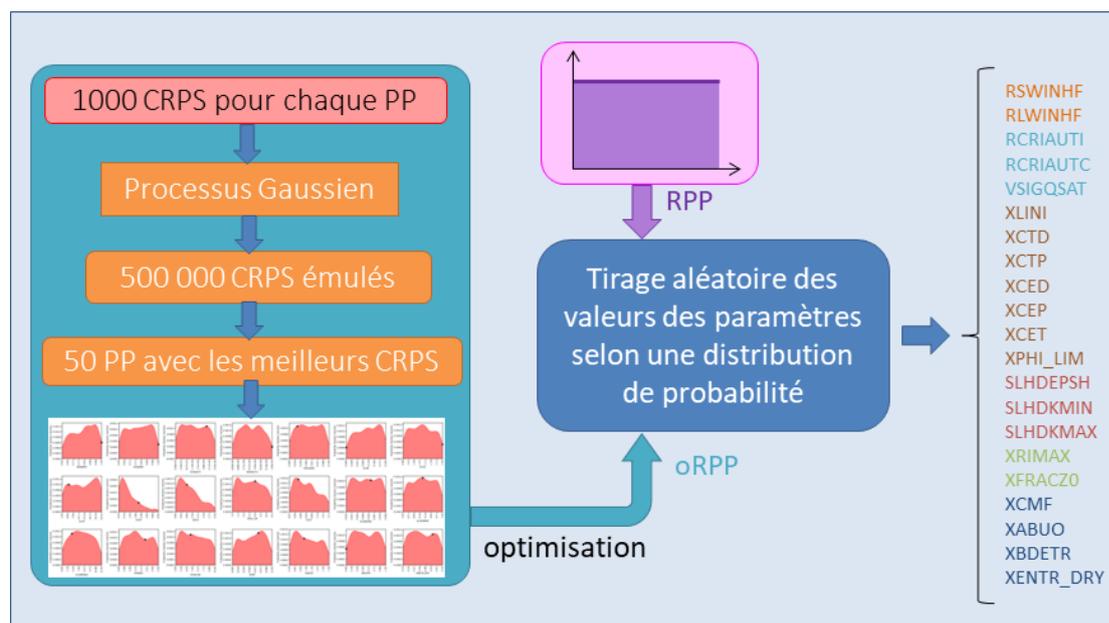


FIGURE 7.4 – Schéma comparant les tirages aléatoires des 21 paramètres perturbés, effectués pour chaque membre d'une prévision d'ensemble avec les méthodes RPP et oRPP (ex pour 50 B-eCRPS-PP)

où :

- $f_k(x)$  est la distribution continue pour toute valeur  $x$  d'un paramètre  $k$  ;
- $x_i$  correspond aux valeurs déjà connues de  $x$  (en l'occurrence, les 12 valeurs de chaque paramètre du B-CRPS-PP) ;
- $N_{Mb}$  représente le nombre de valeurs de chaque paramètre ( $N_{Mb} = 12$ ) ;
- $K(x)$  est un noyau gaussien :

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right) \quad (7.6)$$

- $h$  un paramètre de lissage appelé bande passante. Différentes formulations de  $h$  existent ; nous avons utilisé celle de Scott (1979), à savoir :

$$h = N_{mb}^{-1/5} \quad (7.7)$$

Afin de réaliser un tirage aléatoire selon ces distributions, l'inverse de la fonction de distribution cumulée de chacune d'entre elles est calculé. La Fonction de Distribution Cumulée (CDF) étant comprise entre 0 et 1, un tirage aléatoire uniforme sur ce même intervalle est transformé en tirage aléatoire selon la distribution de probabilité originale. Le schéma 7.7 synthétise comment cette méthode passe de l'histogramme des valeurs d'un paramètre au générateur de nombre aléatoire suivant la distribution de probabilité correspondante. L'exemple est donné pour le paramètre VSIQSAT et sa distribution dans l'échantillon du plan d'expérience de Morris (choix pour des raisons esthétiques). Cette méthode, appliquée aux

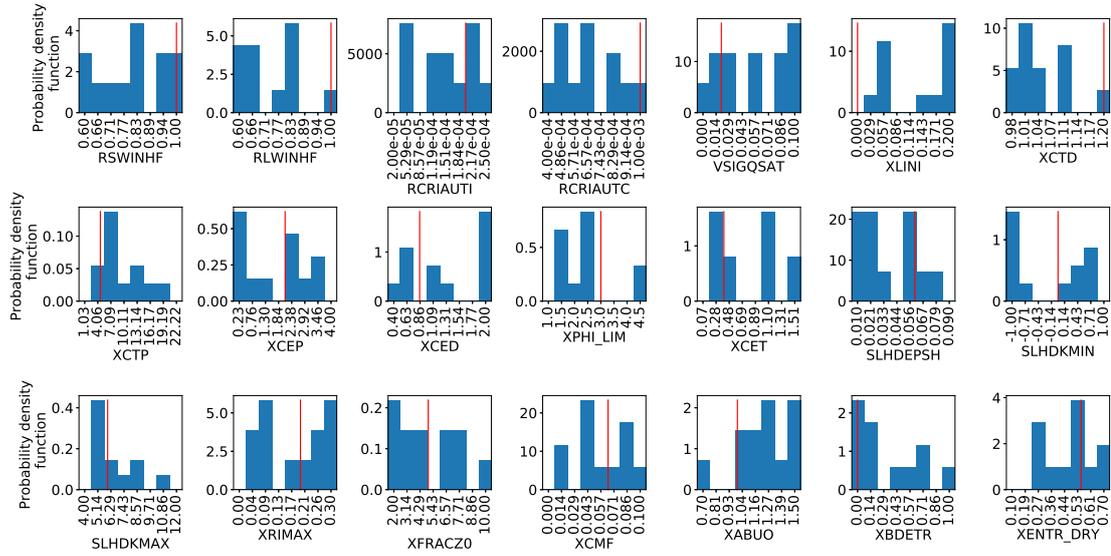


FIGURE 7.5 – Histogramme normalisé (intégrale unitaire) de la distribution de probabilité sur les 8 valeurs des paramètres perturbés pour le B-CRPS-PP. La verticale rouge représente la valeur par défaut des paramètres.

histogrammes de la figure 7.5, permet d’obtenir les distributions de probabilités correspondantes présentées en figure 7.6.

### 7.3.3 A partir des distributions des 50 meilleurs B-CRPS-PP et du plan d’expérience de Morris

Tout comme pour le B-CRPS-PP, des distributions de probabilité des paramètres permettant d’optimiser le RPP sont obtenues à partir des 50 meilleurs B-CRPS-PP en utilisant la même méthode que précédemment (voir schéma 7.7). Alors que le B-CRPS-PP fournit uniquement 12 valeurs pour chaque paramètre, les 50 B-CRPS-PP fournissent un échantillon plus robuste des valeurs privilégiées pour chaque paramètre. Leurs distributions apparaissant sur la Figure 7.8 semblent cependant assez similaires à celles du plan d’expérience de Morris, c’est-à-dire, les distributions issues des 264 valeurs de paramètres échantillonnées dans le plan d’expérience de Morris (Figure 7.9).

Ceci s’explique par la loi des grands nombres : plus un échantillon est grand, plus il tend vers la distribution de probabilité auquel il appartient. Utiliser un grand nombre de B-CRPS-PP a donc l’inconvénient de tendre vers les distributions du plan d’expérience de Morris. Un test de Kolmogorov-Smirnov a donc été réalisé pour chaque paramètre afin de vérifier la similarité entre les distributions issues des 50 B-CRPS-PP et celles du plan d’expérience de Morris. La p-value de ces tests étant majoritairement supérieure à 5% (excepté pour XCEP, XCET et SLHDKMIN, voir valeur dans le tableau 7.1), l’hypothèse selon laquelle les deux distributions sont égales ne peut pas être rejetée. Ce test ne montre alors pas, pour la plupart des paramètres, que les deux distributions sont issues de la même loi de probabilité, ni même qu’elles sont différentes. Seules les distributions de XCED, XCET et SLHDKMIN sont significativement différentes de leur homologue du plan de Morris.

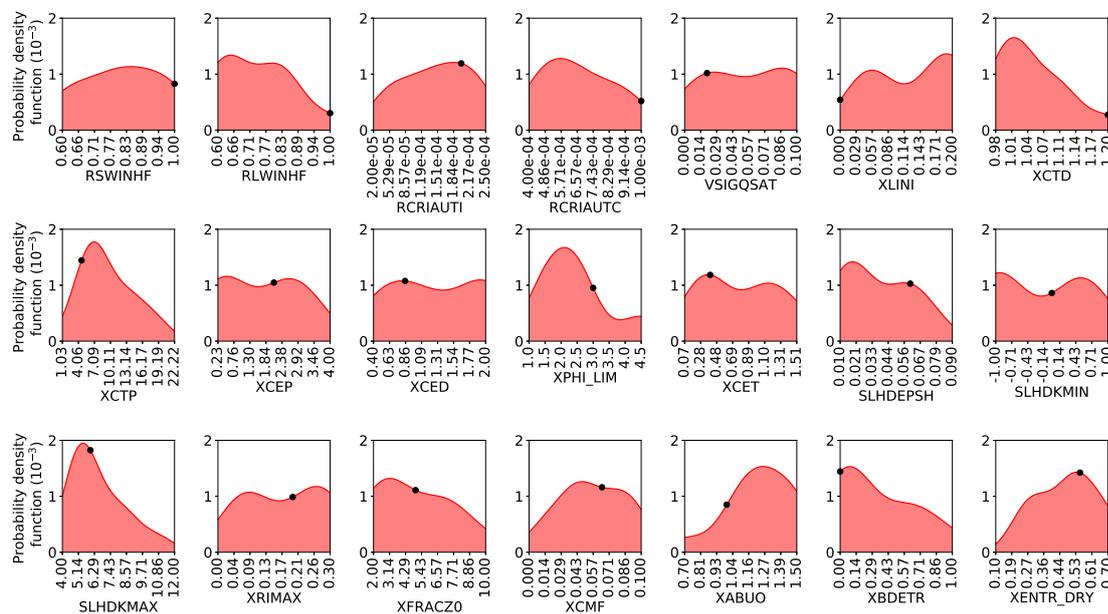


FIGURE 7.6 – Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le B-CRPS-PP. Ces distributions sont obtenues à partir de celles de la figure 7.3 et d'une estimation par noyau. Le point noir représente la valeur par défaut des paramètres.

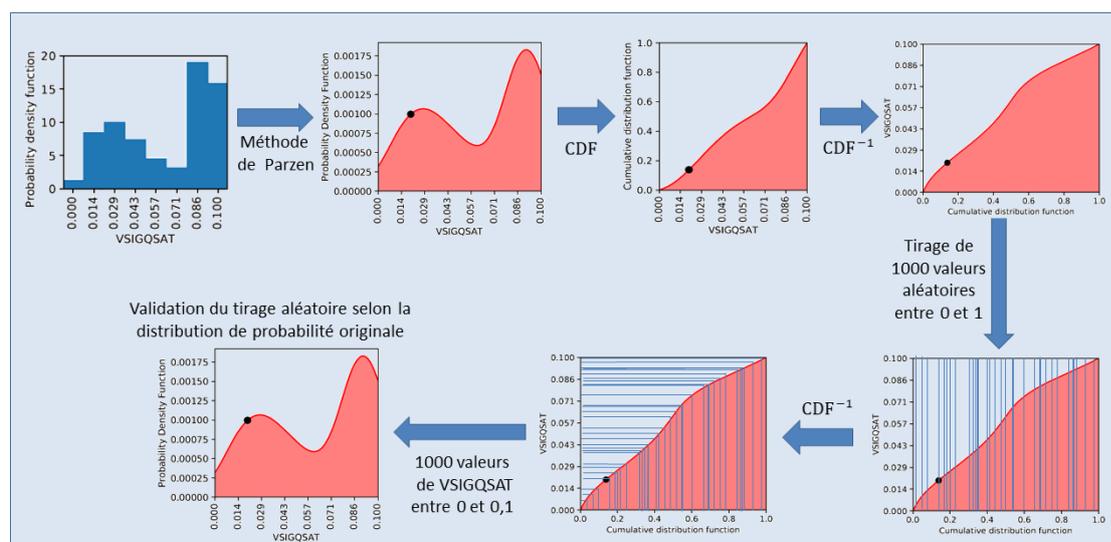


FIGURE 7.7 – Schéma expliquant la génération de nombre aléatoire suivant une loi de probabilité continue issue d'histogramme et validation de la méthode (exemple à partir de l'histogramme de VSIGQSAT du plan d'expérience de Morris)

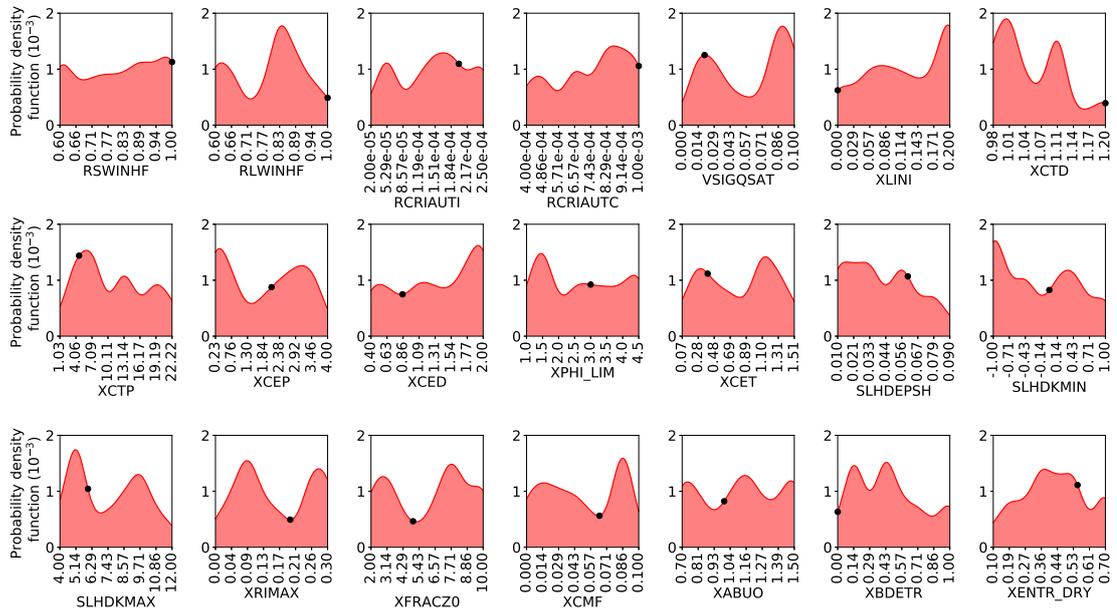


FIGURE 7.8 – Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé à partir des 50B-CRPS-PP. Le point noir représente la valeur par défaut des paramètres.

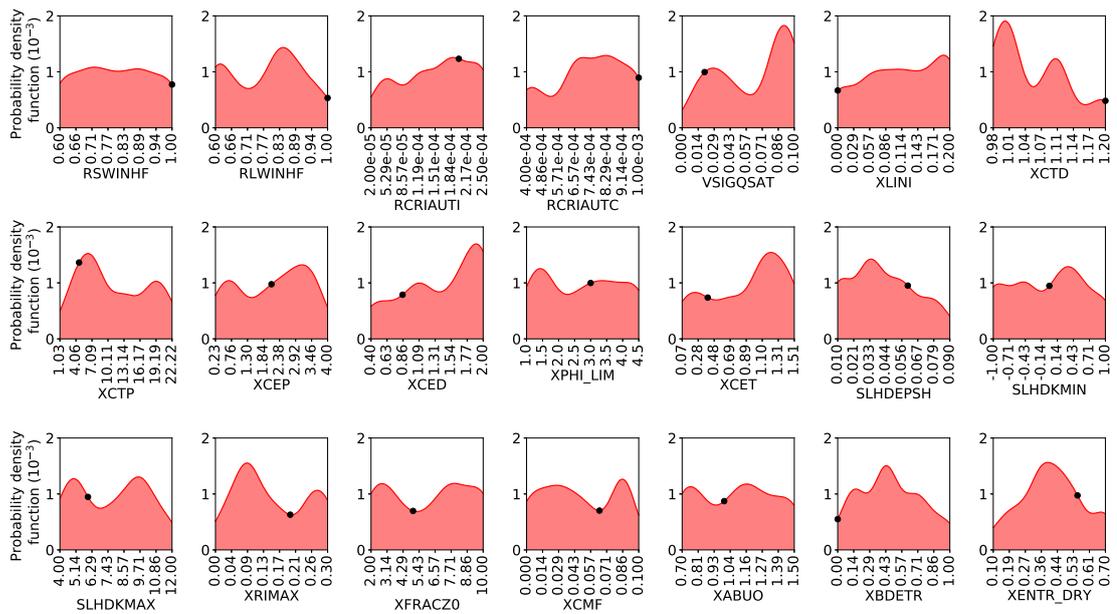


FIGURE 7.9 – Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le plan d'expérience de Morris. Le point noir représente la valeur par défaut des paramètres.

Paramètres	p-value (%)
RSWINHF	44.19
RLWINHF	53.84
RCRIAUTI	62.77
RCRIAUTC	49.23
VSIGQSAT	30.45
XLINI	11.85
XCTD	96.32
XCTP	96.32
XCEP	0.14
XCED	47.62
XPHI_LIM	99.98
XCET	3.79
SLHDEPSH	36.64
SLHDKMIN	0.76
SLHDKMAX	63.47
XRIMAX	25.66
XFRACZ0	89.47
XCMF	84.37
XABUO	39.44
XBDETR	49.23
XENTR_DRY	60.34

TABLE 7.1 – *p*-value d'un test de Kolmogorov-Smirnov réalisé entre les distributions des 50B-CRPS-PP et de Morris pour chaque paramètre

### 7.3.4 A partir des distributions obtenues avec un émulateur de CRPS

Les distributions de probabilités des paramètres avec les 50 B-CRPS-PP étant issues du plan d'expérience de Morris, seules  $p = 8$  valeurs de paramètres ont été échantillonnées. De ce fait, les distributions ont été obtenues avec une forte discrétisation des paramètres, ce qui peut être insuffisant pour une estimation correcte de ces distributions de probabilité. De plus, la gamme de variation des paramètres n'est pas pleinement échantillonnée et de meilleurs B-CRPS-PP pourraient exister pour des valeurs de paramètres non testées. L'idée est donc de créer de nouveaux ensembles PP avec d'autres valeurs de paramètres non échantillonnées par le plan d'expérience de Morris. Cependant, ceci demande la réalisation d'un nombre considérable de nouvelles prévisions. Or, en raison du coût de calcul des modèles de prévision numérique du temps, ceci n'est pas faisable. Une solution est donc d'utiliser un émulateur simulant la prévision. Cet émulateur étant une approximation statistique du modèle réel, il peut fournir, avec un coût de calcul très faible, un nombre suffisant de sorties du modèle pour chercher de « meilleures configurations » PP.

**Régression par Processus Gaussiens :** Nous avons choisi de construire un émulateur de CRPS ayant pour données d'entrée les valeurs des  $k = 21$  paramètres pour chacun des membres de l'ensemble. Nous noterons par la suite  $\Omega$  l'espace des

paramètres de tous les membres comprenant  $N_\Omega = k \times N_{Mb} = 252$  dimensions. Ce métamodèle a donc pour entrée 252 paramètres issus de l'espace  $\Omega$  et comme sortie l'unique valeur de  $F$  (combinaison du CRPS de pluie et vent) correspondante. Celui-ci est construit à partir des 1000 valeurs de  $F$  de chacun des 1000 PP déjà connus. Différents émulateurs peuvent convenir. Ici, nous avons choisi d'utiliser la régression par Processus Gaussiens (PG ; Rasmussen and Williams, 2006) car elle a déjà montré une réelle efficacité lors de l'émulation de scores sur d'autres modèles météorologiques (par ex. : Gong et al., 2015; Di et al., 2015, ...).

Le principe de cette régression par processus gaussien est le suivant. Tout d'abord, définissons  $X = [x_1, \dots, x_{N_\Omega}]$ , le vecteur des paramètres d'entrée définissant un point de l'espace  $\Omega$  et  $Y(X)$ , la sortie scalaire du modèle réel que l'on cherche à approximer. Dans notre cas, la sortie scalaire est la combinaison des CRPS de vent et pluie  $F$ . En supposant que  $Y(X)$  est un champ aléatoire gaussien, on pose :

$$Y(X) = M(X) + W(X), \forall X \in \Omega \quad (7.8)$$

où :

- $M(X)$  est la moyenne du champ aléatoire gaussien, définie ici comme une constante. Dans notre cas, il s'agit de la valeur moyenne de  $F$ . Notons que d'autres fonctions moyenne peuvent être utilisées, par exemple, un polynôme de degré 1 ;
- $W(X)$  est un processus gaussien centré stationnaire de variance  $\sigma^2$  et de fonction de corrélation  $R$ , donc de covariance  $\Sigma = cov(W(X), W(X')) = \sigma^2 R(X, X')$ .

Pour une base d'apprentissage de  $N_l = 1000$  données, chacune des composantes de  $Y_l = [Y(X_1), \dots, Y(X_{N_l})]^T$  correspondants aux entrées  $X_l = [X_1^T, \dots, X_{N_l}^T]^T$  du modèle réel est donc considérée comme un champ aléatoire gaussien. La loi de probabilité dite jointe, des variables composants  $Y_l$ , suit une loi Normale de moyenne  $M_l$  et de matrice de covariance  $\Sigma_l$  définis tels que :

$$M_l = M(X_l) = [M(X_1), \dots, M(X_{N_l})]^T \quad (7.9)$$

$$\Sigma_{l,(i,j)} = \sigma^2 R(X_i, X_j) \quad (7.10)$$

Notons que le long de la diagonale de la matrice de covariance  $\Sigma_l$ , un bruit  $\alpha$ , dit effet pépite, peut être rajouté afin de prendre en compte l'erreur sur la variance des données d'apprentissage. Ceci permet entre autres d'aider la régression par Processus Gaussien à mieux s'ajuster aux données d'apprentissage.

Si on ajoute ensuite un nouvel échantillon  $X^*$  dont on souhaite connaître la sortie  $Y(X^*)$ , la loi jointe suivra alors la loi Normale suivante :

$$[Y(X_l), Y(X^*)] \sim \mathcal{N} \left( \begin{pmatrix} M_l \\ M(X^*) \end{pmatrix}, \begin{pmatrix} \Sigma_l & r(X^*) \\ r(X^*) & \sigma^2 \end{pmatrix} \right) \quad (7.11)$$

où  $r(X^*) = [R(X_1, X^*), \dots, R(X_{N_l}, X^*)]^T$ .

Enfin, la loi conditionnelle de  $Y(X^*)|_{Y(X_l)=Y_l}$  suivra donc une loi normale de moyenne  $M^*(X^*)$  et de variance  $\Sigma^{*2}(X^*)$  où :

$$M^*(X^*) = \mathbb{E}(Y(X^*)|Y(X_l) = Y_l) = M(X^*) + r(X^*)^T \Sigma_l^{-1} (Y_l - M_l) \quad (7.12)$$

$$\Sigma^{*2}(X^*) = \mathbb{V}[Y(X^*)|Y(X_l) = Y_l] = \sigma^2 - r(X^*) \Sigma_l^{-1} r(X^*) \quad (7.13)$$

C'est à partir de cette loi conditionnelle, qu'un métamodèle, basé sur la régression par Processus Gaussien, prédit la valeur de sortie du modèle la plus probable pour  $X^*$ , c'est-à-dire  $M^*(X^*)$ . Notons qu'un tel métamodèle peut aussi estimer l'erreur  $\Sigma^*$  de l'estimation de sa prédiction  $M^*$ .

**Test des noyaux de covariances :** La régression par Processus Gaussien estimant la sortie d'une entrée  $X^*$  à partir de la covariance de cette entrée et avec celle de la base d'apprentissage  $X_l$ , une fonction de covariance doit être imposée à l'émulateur. Différents noyaux de covariance ont été testés afin de réduire au maximum l'erreur de validation du métamodèle. Utilisant le package python sklearn, nous avons testé 10 noyaux différents apparaissant en figure 7.10 et utilisant des combinaisons de noyau de covariances parmi la liste suivante :

- noyau constant :

$$k_c(X^*, X_l) = \theta \quad (7.14)$$

où  $\theta$  est le paramètre réglé par une méthode de maximisation de la vraisemblance pour que la régression par Processus Gaussien s'ajuste aux données d'apprentissage.

- noyau composé uniquement de bruit (White Kernel) :

$$k_{White}(X^*, X_l) = \alpha \quad (7.15)$$

où  $\alpha$  est un nombre aléatoire si  $X^* = X_l$ , sinon est nul,

- la fonction radial-basis :

$$k_{RBF}(X^*, X_l) = \exp\left(-\frac{d(X^*, X_l)^2}{2\theta^2}\right) \quad (7.16)$$

où  $d(X^*, X_l)$  est la distance euclidienne entre  $X^*$  et  $X_l$ ,

- le noyau de Matérn où  $\nu = 1/2, 3/2$  et  $5/2$  :

$$k_{Matern,1/2}(X^*, X_l) = \exp\left(-\frac{1}{\theta}d(X^*, X_l)\right) \quad (7.17)$$

$$k_{Matern,3/2}(X^*, X_l) = \left(1 + \frac{\sqrt{3}}{\theta}d(X^*, X_l)\right) \exp\left(-\frac{\sqrt{3}}{\theta}d(X^*, X_l)\right) \quad (7.18)$$

$$k_{\text{Matern},5/2}(X^*, X_l) = \left(1 + \frac{\sqrt{5}}{\theta}d(X^*, X_l) + \frac{5}{3\theta}d(X^*, X_l)^2\right) \exp\left(-\frac{\sqrt{5}}{\theta}d(X^*, X_l)\right) \quad (7.19)$$

- le noyau Rational Quadratic :

$$k_{RQ}(X^*, X_l) = \left(1 + \frac{d(X^*, X_l)^2}{2\beta\theta^2}\right)^{-\beta} \quad (7.20)$$

où  $\beta$  est un paramètre d'échelle de mélange ( $\beta = 1$ )

- le noyau Exp-Sine-Squared :

$$k_{ESS}(X^*, X_l) = \exp\left(-\frac{2 \sin^2\left(\frac{\pi d(X^*, X_l)}{p}\right)}{\theta^2}\right) \quad (7.21)$$

où  $p$  est la périodicité ( $p = 1$ )

- le noyau DotProduct :

$$k_{DP}(X^*, X_l) = \sigma^2 + X^* \cdot X_l \quad (7.22)$$

où  $\sigma$  est un facteur d'inhomogénéité ( $\sigma = 1$ ).

Afin d'évaluer la performance des différents noyaux de covariance, nous avons calculé l'erreur relative de validation et entraînement du métamodèle pour des échantillons correspondant à 30% et 70% respectivement de la base de données initiale. Ces sous-échantillons étant tirés aléatoirement, ce processus est réitéré 100 fois afin d'obtenir des erreurs relatives moyennes robustes. Les erreurs obtenues sont présentées sur la Figure 7.10 et montrent une faible erreur de validation pour le noyau de Matérn  $3/2$  (0.314%) pour une erreur d'entraînement négligeable (0.034%). Le noyau DotProduct est aussi retenu : il offre certes une erreur de validation plus faible (0.26%), cependant son erreur d'entraînement étant plus importante (0.17%), il semble moins provoquer de sur-apprentissage.

**Augmentation de la base d'apprentissage :** Les 1000 échantillons de  $F$  servant de base d'apprentissage pour l'émulateur étant peu nombreux pour un espace d'entrée  $\Omega$  à 252 dimensions, nous souhaitons augmenter la base d'apprentissage pour réduire l'erreur de validation du métamodèle. Puisque l'ordre des membres d'une prévision d'ensemble n'a pas d'impact sur la valeur du CRPS et donc de  $F$  de cet ensemble, une prévision d'ensemble  $A$  aura le même CRPS (et donc même  $F$ ) qu'une prévision d'ensemble  $B$  si  $A$  et  $B$  contiennent les mêmes membres mais rangés dans un ordre différent. Nous allons donc nous baser sur ce fait pour augmenter la base d'apprentissage.

En effet, en changeant l'ordre des membres et donc des paramètres correspondants, on crée un nouveau point dans l'espace  $\Omega$  des 252 paramètres d'entrées de

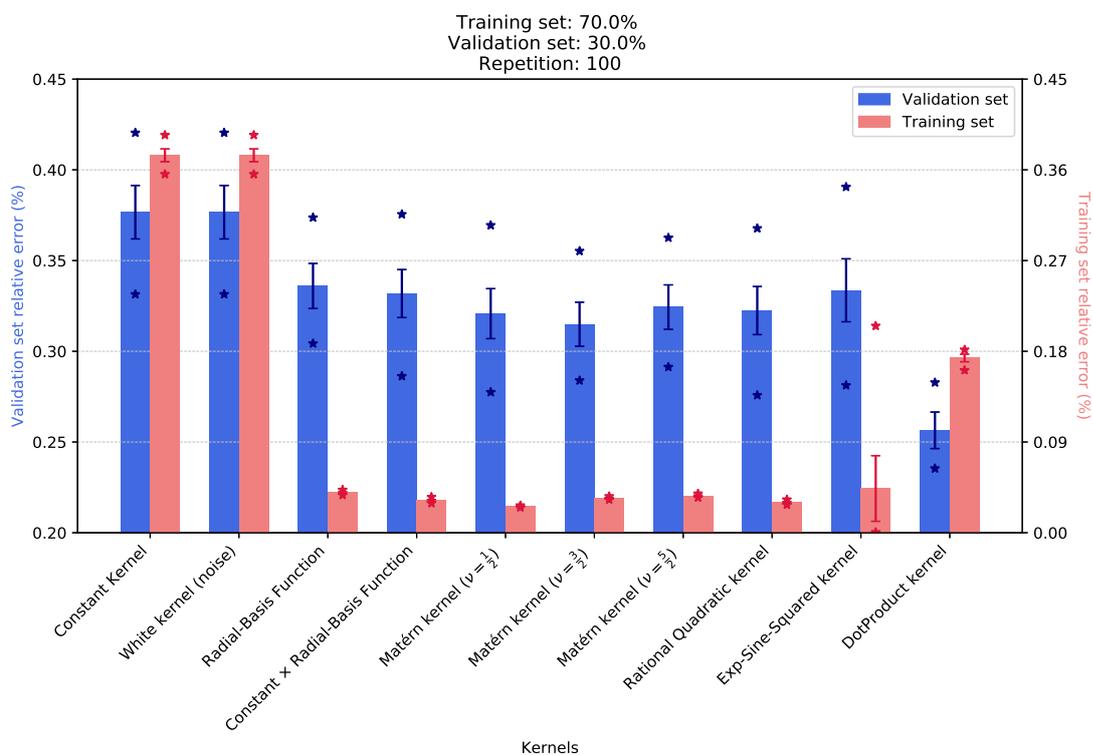


FIGURE 7.10 – Erreurs relatives moyennes de validation (en bleu) et d’entraînement (en rouge) sur des échantillons correspondants respectivement à 30% et 70% de la base d’apprentissage. Les barres d’erreur indiquent les quantiles 25% et 75% alors que les minima et maxima apparaissent sous la forme d’astérisques. Notons deux axes des ordonnées, à gauche pour l’erreur relative moyenne de validation et à droite pour celle de l’entraînement.

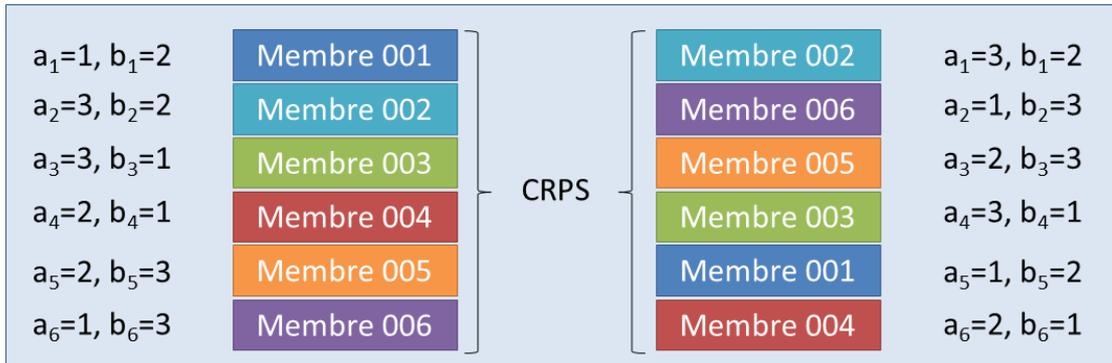


FIGURE 7.11 – Schéma d'explication du rééchantillonnage et augmentation de la base de données d'apprentissage

l'émulateur (voir schéma 7.11). Par exemple, pour simplifier, considérons une prévision d'ensemble à  $N_{Mb} = 6$  membres perturbant  $k = 2$  paramètres nommés  $a$  et  $b$  et ayant pour score (et donc sortie du modèle)  $F$ . Le membre  $n^{\circ}i$  de cette prévision d'ensemble utilise comme valeur des paramètres  $a$  et  $b$ , les valeurs  $a_i$  et  $b_i$  (pour  $i$  allant de 1 à  $N_{Mb} = 6$ ). Ainsi, le point dans l'espace  $\Omega$  correspondant au score  $F$  sera donc  $(a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5, a_6, b_6)$ . Si l'on intervertit aléatoirement l'ordre des membres, on crée une nouvelle prévision d'ensemble correspondant, par exemple, au point  $(a_2, b_2, a_6, b_6, a_5, b_5, a_3, b_3, a_1, b_1, a_4, b_4)$  dont la valeur de  $F$  est la même que celle de la prévision originale (voir Figure 7.11). De cette façon, on peut augmenter le nombre de données d'apprentissage pour l'émulateur.

Différentes tailles d'augmentation de la base de données ont été testées. En détail, nous avons calculé l'erreur de validation pour des tailles allant de 1000 à 40 000. L'augmentation de la taille de la base d'apprentissage a été limitée à 40 000 pour des raisons de coût de calcul. Ces différentes erreurs de validation ont été estimées sur la moyenne de 30 tests utilisant des échantillons de validation tirés aléatoirement et correspondant à 30% de la base d'apprentissage initiale. La Figure 7.12 présente alors ces erreurs de validation en fonction de la taille de la base d'apprentissage, pour une régression par Processus Gaussien utilisant la fonction de Matérn  $3/2$  comme noyau de covariance. L'erreur du métamodèle étant décroissante, selon une loi logarithmique (Figure 7.12), avec la taille de l'échantillon d'apprentissage, nous avons donc choisi d'augmenter notre base d'apprentissage, comprenant initialement de  $1000F$ , au maximum possible, soit 40 000  $F$ .

**Recherche des distributions de probabilité de PP optimaux :** Une fois l'émulateur entraîné sur les 40 000 données d'apprentissage, un tirage aléatoire de Monte-Carlo de 500 000 points dans l'espace  $\Omega$  a été effectué, représentant ainsi les valeurs des paramètres de chaque membre pour 500 000 PP. Pour chacun de ces points, l'émulateur a estimé les valeurs de  $F$  correspondantes.

Les 50 jeux de paramètres présentant les meilleurs scores  $F$  émulés sont appelées les 50 « Best emulated CRPS-PP » (B-eCRPS-PP). Tout comme pour les 50 B-CRPS-PP, ces 50B-eCRPS-PP permettent d'aboutir à des distributions de probabilité pour chaque paramètre afin d'optimiser le RPP.

Puisque deux noyaux de covariances ont été retenus, deux émulateurs ont donc été réalisés fournissant ainsi deux distributions de probabilité différentes pour la méthode oRPP. Nous noterons donc par la suite le RPP optimisé selon les

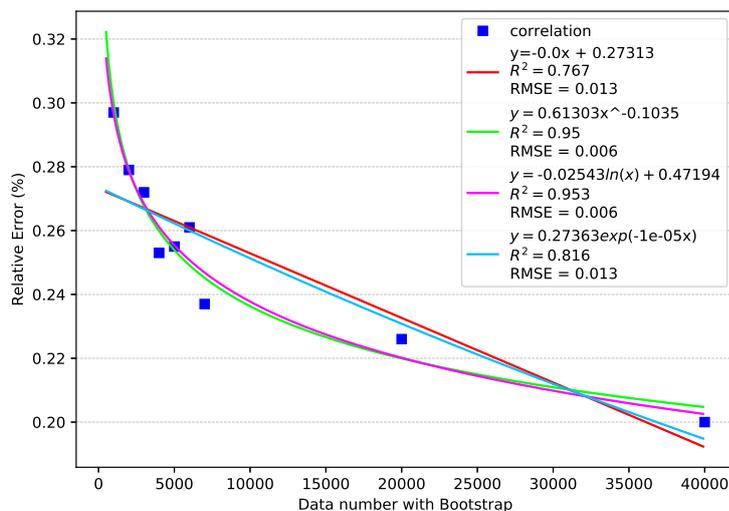


FIGURE 7.12 – Erreur de validation pour différentes tailles d'augmentation de la base d'apprentissage

distributions des 50 B-eCRPS-PP utilisant le noyau de Matérn, le oRPPMat, et celui utilisant le noyau DotProduct, le oRPPDot. Au vu des Figures 7.13 et 7.14, les deux noyaux de covariance n'aboutissent pas aux mêmes distributions de probabilité, en particulier pour XFRACZ0 ou XCMF par exemple.

Notons que la méthode employée ici s'adresse à un modèle ensembliste. Des méthodes semblables d'optimisation de paramètres ont déjà été mises en place pour des modèles déterministes. Celle utilisée par Gong et al. (2015) pour un modèle de surface et Duan et al. (2017) pour un modèle de prévisions du temps (WRF) s'appuie sur la méthode Adaptive Surrogate Modelling-based Optimization (ASMO) de Wang et al. (2014) : (1) un premier émulateur (cherchant à imiter un modèle météorologique réel) est construit à partir d'un échantillon initial de sorties du modèle réel (la sortie étant un score déterministe) ; (2) les sorties sont émulées sur l'ensemble du domaine de perturbation des paramètres ; (3) un sous-domaine de perturbation des paramètres donnant les meilleurs scores est identifié ; (4) de nouveaux échantillons sont obtenus avec le modèle réel pour le sous-domaine considéré ; (5) l'émulateur est à nouveau construit en rajoutant le nouvel échantillon à la base de donnée d'apprentissage. Le processus est réitéré jusqu'à ce qu'il converge vers un optimum. Duan et al. (2017) se sont appuyés, en particulier, sur l'analyse de sensibilité réalisée par Di et al. (2015) et Quan et al. (2016) pour établir les différents émulateurs.

## 7.4 Réduction à la perturbation de 8 paramètres

Enfin, d'après l'analyse de sensibilité de Morris, parmi les 21 paramètres initialement identifiés, 8 ont une influence prépondérante sur les prévisions de plusieurs variables météorologiques près de la surface. Afin de vérifier qu'ils sont suffisants pour représenter en grande partie l'erreur de modélisation avec la méthode de paramètres perturbés, nous définissons, pour chacun des systèmes de prévision d'ensemble précédemment définis, un équivalent où seuls les 8 paramètres iden-

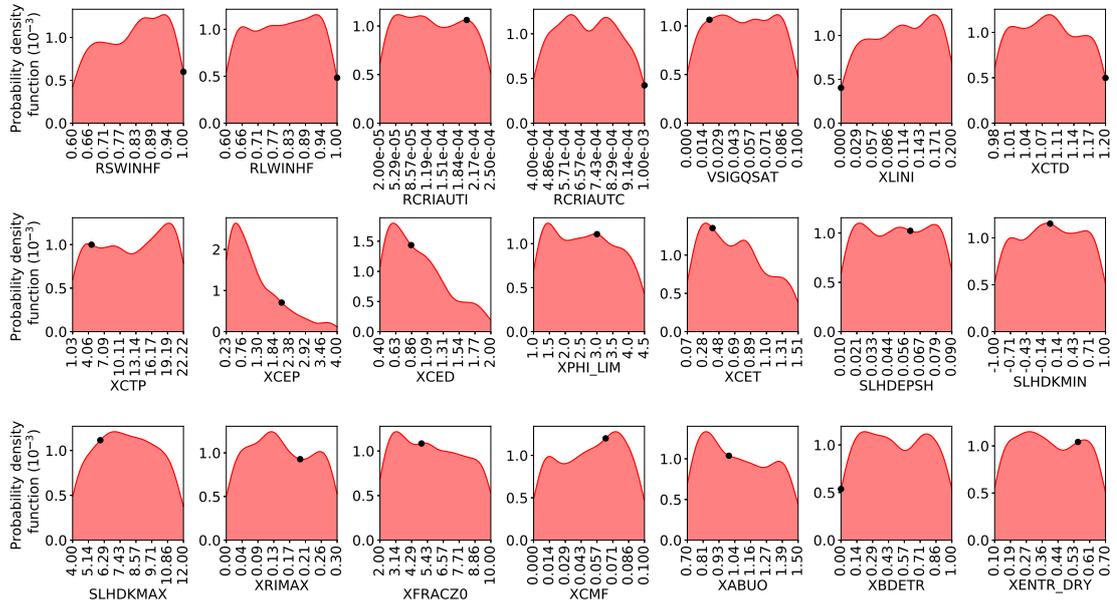


FIGURE 7.13 – Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le oRPPMat. Le point noir représente la valeur par défaut des paramètres.

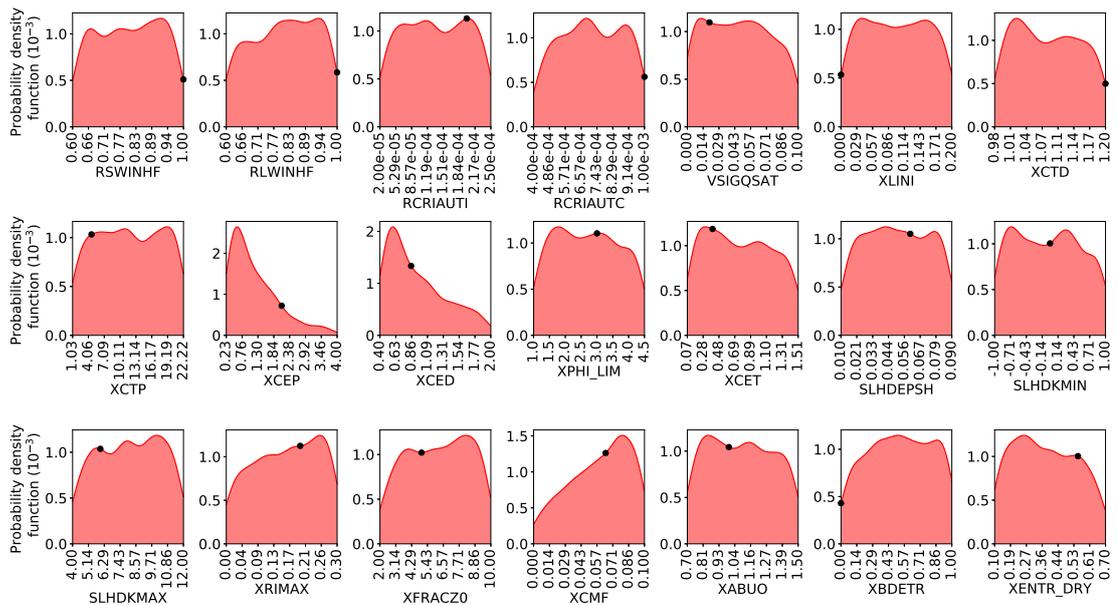


FIGURE 7.14 – Distribution de probabilité le long de l'intervalle de perturbation de chaque paramètre perturbé pour le oRPPDot. Le point noir représente la valeur par défaut des paramètres.

tifiés comme influents sont perturbés. Le tableau 7.2 liste ainsi la totalité des systèmes de prévision d'ensemble présentés dans la suite du manuscrit.

Système de Prédiction d'Ensemble	Type de perturbations	Nombre de paramètres perturbés
Mean PP	PP moyen	21
B-CRPS-PP	PP optimisé selon le CRPS sur l'année	21
B-CRPS-P8P	PP optimisé selon le CRPS sur l'année	8
Be-CRPS-PP	PP optimisé selon le CRPS sur l'été	21
Be-CRPS-P8P	PP optimisé selon le CRPS sur l'été	8
Ba-CRPS-PP	PP optimisé selon le CRPS sur l'automne	21
Ba-CRPS-P8P	PP optimisé selon le CRPS sur l'automne	8
Bh-CRPS-PP	PP optimisé selon le CRPS sur l'hiver	21
Bh-CRPS-P8P	PP optimisé selon le CRPS sur l'hiver	8
RPP	RPP utilisant une distribution uniforme	21
RP8P	RPP utilisant une distribution uniforme	8
oRPP Gaussien	oRPP utilisant une distribution gaussienne autour du B-CRPS-PP	21
oRP8P Gaussien	oRPP utilisant une distribution gaussienne autour du B-CRPS-PP	8
oRPP-B-CRSP-PP	oRPP utilisant les distributions issues du B-CRSP-PP	21
oRP8P-B-CRSP-PP	oRPP utilisant les distributions issues du B-CRSP-PP	8
oRPP-Morris	oRPP utilisant les distributions issues du plan d'expérience de Morris	21
oRP8P-Morris	oRPP utilisant les distributions issues du plan d'expérience de Morris	8
oRPP-50B-CRSP-PP	oRPP utilisant les distributions issues des 50 meilleurs B-CPPS-PP	21
oRP8P-50B-CRSP-PP	oRPP utilisant les distributions issues des 50 meilleurs B-CPPS-PP	8
oRPPMat	oRPP utilisant les distributions issues des 50 meilleurs CRPS émulsés avec un PG utilisant le noyau de Matern 3/2	21
oRP8PMat	oRPP utilisant les distributions issues des 50 meilleurs CRPS émulsés avec un PG utilisant le noyau de Matern 3/2	8
oRPPDot	oRPP utilisant les distributions issues des 50 meilleurs CRPS émulsés avec un PG utilisant le noyau DotProduct	21
oRP8PDot	oRPP utilisant les distributions issues des 50 meilleurs CRPS émulsés avec un PG utilisant le noyau DotProduct	8

TABLE 7.2 – Liste des différents systèmes de prévisions d'ensemble testés avec différentes optimisations. Les versions perturbant 8 paramètres sont en gris.

## Bibliographie

- Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, J. Quan, J. Li, C. Miao, A. Ye, and C. Tong. Assessing wrf model parameter sensitivity : A case study with 5 day summer precipitation forecasting in the greater beijing area. *Geophysical Research Letters*, 42(2) :579–587, 2015. doi : <https://doi.org/10.1002/2014GL061623>.
- L. I. Díaz-Isaac, T. Lauvaux, M. Bocquet, and K. J. Davis. Calibration of a multi-physics ensemble for estimating the uncertainty of a greenhouse gas atmospheric transport model. *Atmospheric Chemistry and Physics*, 19(8) :5695–5718, 2019. doi : [10.5194/acp-19-5695-2019](https://doi.org/10.5194/acp-19-5695-2019).
- Q. Duan, Z. Di, J. Quan, C. Wang, W. Gong, Y. Gan, A. Ye, C. Miao, S. Miao, X. Liang, and S. Fan. Automatic model calibration : A new way to improve numerical weather forecasting. *Bulletin of the American Meteorological Society*, 98(5) :959 – 970, 2017. doi : [10.1175/BAMS-D-15-00104.1](https://doi.org/10.1175/BAMS-D-15-00104.1).
- M. Eklblom, L. Tuppi, V. Shemyakin, M. Laine, P. Ollinaho, H. Haario, and H. Järvinen. Algorithmic tuning of spread–skill relationship in ensemble forecasting systems. *Quarterly Journal of the Royal Meteorological Society*, 146(727) :598–612, 2020. doi : <https://doi.org/10.1002/qj.3695>.
- W. Gong, Q. Duan, J. Li, C. Wang, Z. Di, Y. Dai, A. Ye, and C. Miao. Multi-objective parameter optimization of common land model using adaptive surrogate modeling. *Hydrology and Earth System Sciences*, 19(5) :2409–2425, 2015. doi : [10.5194/hess-19-2409-2015](https://doi.org/10.5194/hess-19-2409-2015).
- P. L. Houtekamer, B. He, D. Jacques, R. McTaggart-Cowan, L. Separovic, P. A. Vaillancourt, A. Zadra, and X. Deng. Use of a genetic algorithm to optimize a numerical weather prediction system. *Monthly Weather Review*, 149(4) :1089 – 1104, 2021. doi : [10.1175/MWR-D-20-0238.1](https://doi.org/10.1175/MWR-D-20-0238.1).
- J. Quan, Z. Di, Q. Duan, W. Gong, C. Wang, Y. Gan, A. Ye, and C. Miao. An evaluation of parametric sensitivities of different meteorological variables simulated by the wrf model. *Quarterly Journal of the Royal Meteorological Society*, 142(700) :2925–2934, 2016. doi : <https://doi.org/10.1002/qj.2885>.
- C. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. Adaptive computation and machine learning series. Cambridge, MA : MIT Press, 2006. ISBN 9780262182539.
- D. W. Scott. On optimal and data-based histograms. *Biometrika*, 66(3) :605–610, 12 1979. ISSN 0006-3444. doi : [10.1093/biomet/66.3.605](https://doi.org/10.1093/biomet/66.3.605).
- C. Wang, Q. Duan, W. Gong, A. Ye, Z. Di, and C. Miao. An evaluation of adaptive surrogate modeling based optimization with two benchmark problems. *Environmental Modelling and Software*, 60 :167–179, 2014. ISSN 1364-8152. doi : <https://doi.org/10.1016/j.envsoft.2014.05.026>.

# Chapitre 8

## Évaluation des nouvelles méthodes de représentation des erreurs de modélisation

Dans ce chapitre, nous allons évaluer chacune des méthodes de représentation des erreurs de modélisation présentées dans le chapitre 7 et les comparer à la méthode actuellement implémentée dans la PEARO, le SPPT. Une première partie s'attachera aux scores probabilistes de ces différentes méthodes puis nous nous intéresserons aux caractéristiques des méthodes de perturbations de paramètres (biais et dispersion). Enfin, d'un point de vue plus subjectif, nous nous pencherons sur le cas des orages du 01 Juillet 2018 ainsi que la tempête de neige Gabriel traversant la France les 29 et 30 Janvier 2019.

Notons que, pour l'ensemble des scores présentés, la significativité à 95% des résultats a été testée avec la méthode de ré-échantillonnage expliquée au paragraphe 2.6.7 et ces scores ont été calculés à partir de prévisions initialisées à 21h UTC sur les périodes définies au tableau 5.2.

### 8.1 Evaluation globale des méthodes PP et RPP

Dans un premier temps, nous allons nous focaliser sur les taux d'amélioration (ou de dégradation) des scores globaux de rapport Dispersion/Erreur et CRPS des méthodes de perturbations de paramètres par rapport à la méthode SPPT. La Figure 8.1 présente ces taux, moyennés selon les échéances, pour le CRPS (Figure 8.1a) et le rapport Dispersion/Erreur (Figure 8.1b) sur la période estivale allant du 1 Mai 2018 au 30 Juillet 2018, pour différentes variables météorologiques.

Notons tout d'abord que de manière générale, le rapport Dispersion/Erreur ainsi que le CRPS sont améliorés, de manière significative, pour la plupart des variables météorologiques de surface et méthodes de perturbations des paramètres considérées, comparées au SPPT (Figure 8.1b). Concernant le rapport Dispersion/Erreur, l'amélioration va jusqu'à plus de 25% pour la plupart des variables météorologiques. Elle atteint en revanche seulement 20% pour l'humidité relative et température à 2m.

Pour le CRPS, l'amélioration est plus faible que celle du rapport Dispersion/Erreur et plus hétérogène selon le champ météorologique étudié, allant d'un maximum à 5% pour la température à 2m jusqu'à un maximum à 20% pour les rafales. Seules les précipitations ont un CRPS similaire à celui du SPPT ou très légèrement moins bon.

### 8.1.1 Impact de la perturbation de paramètres

Dans un premier temps, nous nous focalisons sur les méthodes de perturbations de paramètres simples, sans optimisation, c'est-à-dire « Mean PP » (rouge) et RPP (rose). Pour la quasi totalité des variables météorologiques, les deux méthodes de perturbations de paramètres améliorent le rapport Dispersion/Erreur par rapport au SPPT (Figure 8.1b). Seul le rapport Dispersion/Erreur pour la couverture nuageuse et le rayonnement sont dégradés avec le « Mean PP ». Ceci montre une réelle efficacité de la méthode de perturbations de paramètres sur les scores proches de la surface par rapport à la méthode SPPT. Ceci est certainement dû, en partie, au fait que le SPPT ne perturbe pas les tendances physiques dans les basses couches de l'atmosphère pour des raisons de stabilité numérique, contrairement à la méthode de perturbation de paramètres. Cette dernière peut donc représenter l'erreur de modélisation dans cette zone de l'atmosphère et produire de meilleures dispersions.

Concernant le CRPS (Figure 8.1a), les deux méthodes de perturbations de paramètres l'améliorent pour la plupart des variables météorologiques (exceptées les précipitations), malgré l'absence d'optimisation. Ainsi, elles ne permettent pas seulement d'augmenter la dispersion mais aussi d'améliorer les performances de la PEARO telles que le CRPS.

Par ailleurs, la méthode RPP (rose) propose en général, des rapport Dispersion/Erreur et des CRPS moins performants que le « Mean PP » (rouge) et dégrade même plus ces deux scores lorsque le « Mean PP » est déjà moins performant que le SPPT (précipitations). Les deux principales différences entre ces deux méthodes de perturbations de paramètres résident en l'ajout d'une perturbation journalière des paramètres dans le RPP et dans le fait que le RPP échantillonne toutes les valeurs de paramètres alors que le « Mean PP » n'utilise que les 8 valeurs de paramètres définies par le plan d'expérience de Morris. Ainsi, le « Mean PP » tire aléatoirement des valeurs de paramètres selon une distribution relevant plus des distributions du plan d'expérience de Morris que de distributions uniformes. Ce résultat laisse donc supposer que les 8 valeurs de paramètres utilisées dans le « Mean PP » semblent déjà légèrement privilégier de meilleurs scores que le RPP.

### 8.1.2 Impact de l'optimisation selon le CRPS

Concernant le Be-CRPS-PP optimisé sur l'été 2018 (noir), celui optimisé sur l'année (vert) et le RPP utilisant des distributions gaussiennes autour du B-CRPS-PP (oRPP Gaussien; orange), ces méthodes étant optimisées sur le CRPS de vent et de pluie, ceux-ci offrent une plus grande amélioration du CRPS pour ces deux variables météorologiques comparée à leur équivalent sans optimisation. Ceci permet, entre autres, de corriger la dégradation du CRPS des pluies constatée avec la méthode de perturbations de paramètres non optimisée. Cependant, l'effet de

l'optimisation du CRPS sur les précipitations est beaucoup plus modéré que pour le vent.

Notons tout de même que cette optimisation, bien qu'elle se focalise sur ces deux variables de temps sensible, conduit également à une amélioration du CRPS pour les autres variables météorologiques par comparaison avec la version sans optimisation. Ainsi, concernant le taux d'amélioration du CRPS pour le rayonnement solaire (Sol01) et la couverture nuageuse (cloud), l'écart entre les versions avec et sans optimisation est important.

Bien qu'elles soient optimisées pour le CRPS, les méthodes Be-CRPS-PP (noir), B-CRPS-PP (vert) et oRPP Gaussien (orange) améliorent aussi le rapport Dispersion/Erreur, de manière plus importante que leur équivalent non optimisé pour toutes les variables météorologiques. Elles permettent, entre autres, de corriger la dégradation de ce score pour la couverture nuageuse et le rayonnement solaire par rapport à la méthode « Mean PP ». Ces résultats montrent donc que l'optimisation selon le CRPS n'améliore pas uniquement ce score, mais aussi d'autres scores probabilistes tels que le rapport Dispersion/Erreur.

Enfin, que ce soit pour le CRPS ou le rapport Dispersion/Erreur, les B-CRPS-PP et oRPP Gaussien autour du B-CRPS-PP proposent des taux d'amélioration du même ordre, voire généralement avec une légère diminution de l'amélioration de 2% pour le oRPP Gaussien. Seul le rapport de Dispersion/Erreur des rafales pour le oRPP Gaussien se distingue par un taux d'amélioration plus important d'une dizaine de pourcent.

Notons cependant que ce score reste inférieur à 1 quelles que soient les échéances (non montré). Il n'y a donc pas surdispersion.

### 8.1.3 Optimisation selon une ou 3 saisons

La figure 8.2 présente les mêmes taux d'amélioration que la figure 8.1 mais pour la période hivernale allant du 01 décembre 2018 au 30 janvier 2018. Notons cette fois-ci que l'ensemble des méthodes de perturbations de paramètres conduisent à des améliorations plus fortes, en terme de scores CRPS et rapport Dispersion/Erreur, que pour la période estivale. En particulier, on constate désormais une amélioration du CRPS des précipitations en hiver contrairement à l'été.

Par ailleurs, le B-CRPS-PP optimisé sur la période considérée donne en général les meilleurs scores de CRPS et rapport Dispersion/Erreur pour la plupart des variables météorologiques. Cependant, quelle que soit la saison considérée, le Be-CRPS-PP (optimisé sur l'été) offre toujours de meilleurs scores de CRPS et rapport Dispersion/Erreur pour l'humidité relative. A contrario, le B-CRPS-PP (optimisé sur l'année et correspondant à celui optimisé sur l'automne et l'hiver) donne de meilleurs scores en rafales et couverture nuageuse totale. Notons que malgré une augmentation de la Dispersion/Erreur de près de 100% en vent et rafales en hiver, les B-CRPS-PP sont toujours sous dispersifs (non montré). Enfin, les différences restant faibles entre les optimisations selon les deux périodes, une unique optimisation peut être utilisée pour l'ensemble de l'année.

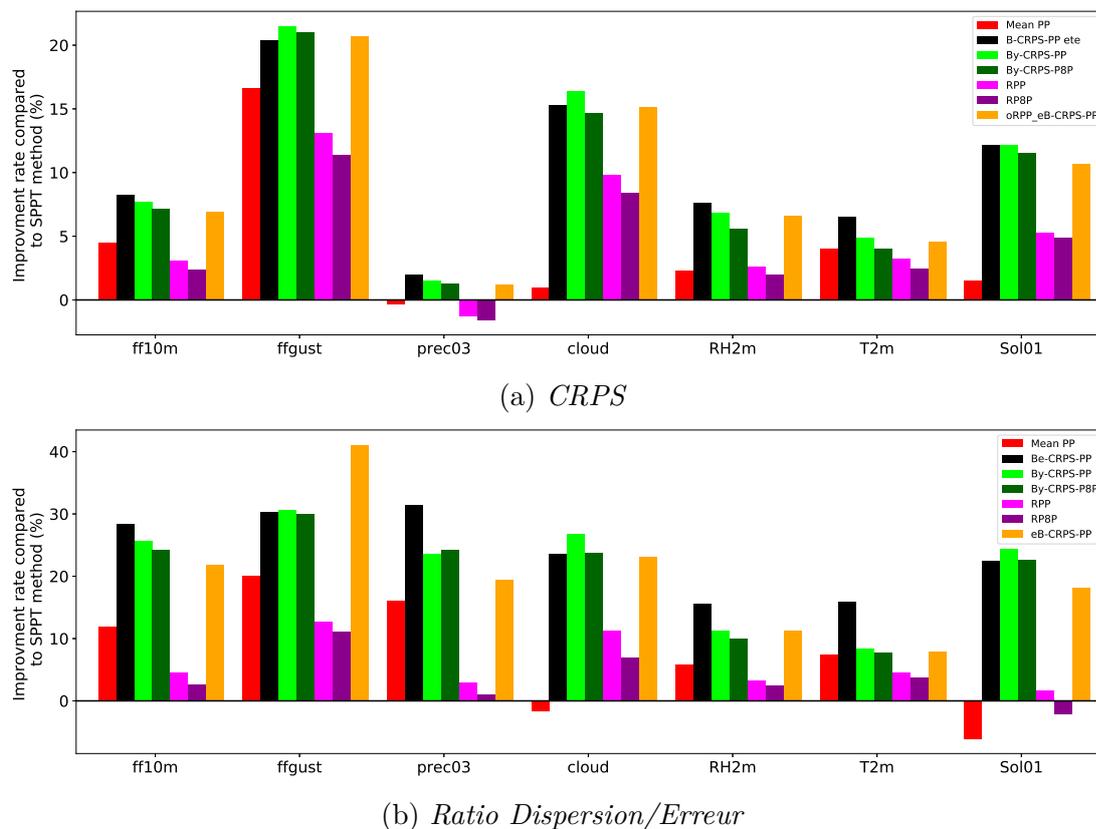


FIGURE 8.1 – Taux d'amélioration par rapport au SPPT du (a) CRPS et du (b) rapport Dispersion/Erreur pendant la période estivale du 01 Mai 2018 au 30 Juillet 2018

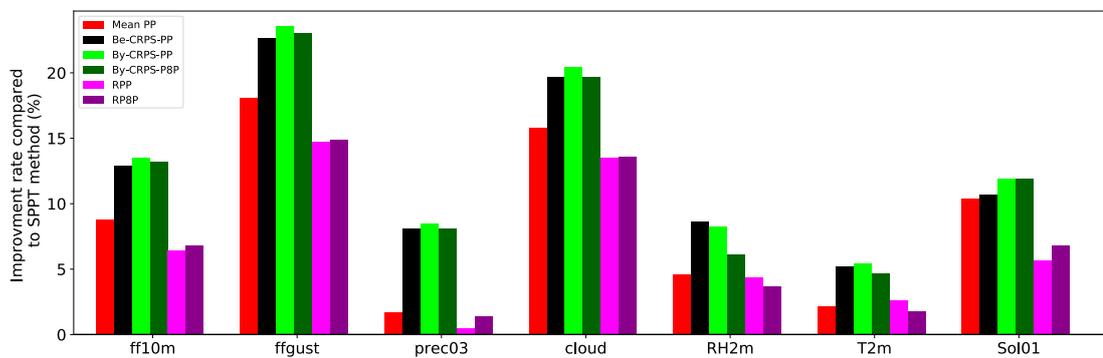
### 8.1.4 Réduction à 8 paramètres

La réduction de la liste des paramètres à perturber, de 21 aux 8 paramètres identifiés comme influents par l'analyse de sensibilité, engendre très peu de variation du CRPS et du rapport Dispersion/Erreur. Que ce soit pour une version optimisée du PP (B-CRPS-P8P ; vert foncé) ou une version non optimisée du RPP (RP8P ; violet foncé), le passage de 21 à 8 paramètres diminue les taux d'amélioration des scores généralement de moins de 1%, et de manière non significative. Cette dégradation peut aller jusqu'à 3% pour le rapport Dispersion/Erreur de la couverture nuageuse. Cette très légère dégradation indique donc que les 8 paramètres sélectionnés par l'analyse de sensibilité semblent suffisants pour représenter la majeure partie de l'erreur de modélisation.

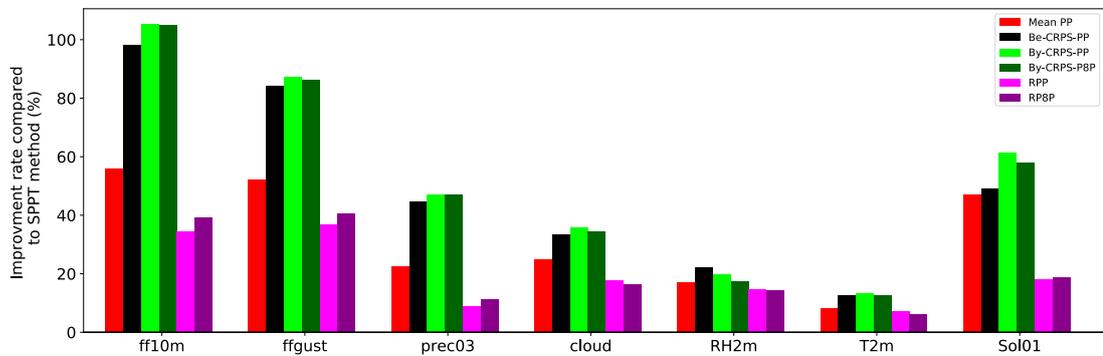
## 8.2 Optimisation de la méthode PP

La PEARO étant particulièrement utilisée pour les prévisions de vent près de la surface et la pluie, nous allons nous intéresser dorénavant, dans toute la suite de ces travaux, à ces deux variables météorologiques.

Afin de regarder plus en détails les scores mesurant la résolution et la fiabilité de ces systèmes de prévision d'ensemble PP, la figure 8.3a présente pour la période estivale l'aire sous la courbe ROC et le diagramme de fiabilité pour la force du vent



(a) *CRPS*



(b) *Ratio Dispersion/Erreur*

FIGURE 8.2 – Taux d'amélioration par rapport au SPPT du (a) *CRPS* et du (b) rapport *Dispersion/Erreur* pendant la période hivernale du 01 Décembre 2018 au 31 Janvier 2019

à 10m excédant le seuil de 20km/h et les précipitations accumulées pendant 3h dépassant 1mm, pour le « Mean PP » (rouge), le B-CRPS-PP (vert), le Be-CRPS-PP (noir) et le SPPT (bleu). Des points apparaissent sur les courbes lorsque la différence avec le SPPT est significative.

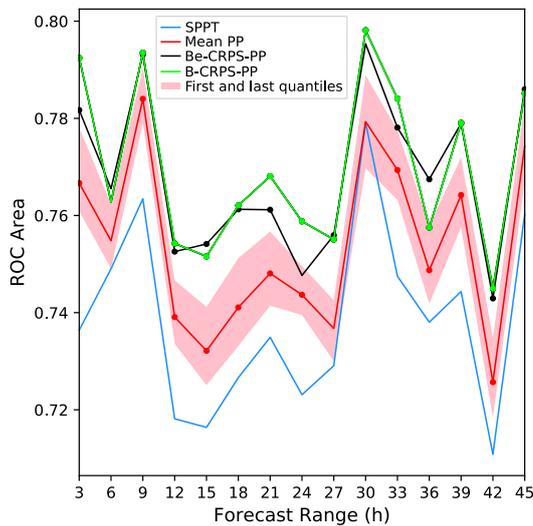
Les résultats sur le vent à 10m montrent une claire amélioration des scores. Concernant l'aire sous la courbe de ROC, l'ensemble des méthodes perturbant des paramètres ont une aire plus grande qu'avec le SPPT, dénotant ainsi une plus grande capacité à discriminer deux évènements. Cette amélioration est significative pour la plupart des échéances (apparition de point le long des courbes). Le « Mean PP » améliorant ce score significativement, la méthode même de perturbation des paramètres sans optimisation semble être mieux adaptée pour discriminer deux évènements que le SPPT. Lorsque cette approche est optimisée selon le CRPS, le score ROC n'est pas pour autant dégradé, il est même augmenté, en particulier pour les échéances correspondant aux après-midi (ex : de 12h à 24h d'échéance, rappelons que l'heure initiale est 21hUTC). Notons que le B-CRPS-PP optimisé sur les 3 saisons améliore d'avantage le score ROC de vent que le Be-CRPS-PP optimisé sur la période considérée.

Pour ce qui est du diagramme de fiabilité pour le vent à 10m (Figure 8.3b), celui-ci montre une sous-dispersion ainsi qu'un biais positif pour les quatre systèmes de prévisions d'ensemble, supposant que la PEARO prévoit de trop forts vents. Cependant, ce biais est plus prononcé pour le SPPT, dénotant ainsi une sur-estimation plus nette du vent avec cette méthode. Ainsi, les « Mean PP », B-CRPS-PP et Be-CRPS-PP semblent être plus fiables et corrigent en partie le biais de vent avec la méthode SPPT. Le « Mean PP » ayant une pente similaire à celle du SPPT, ces deux ensembles présentent donc une sous dispersion du même ordre (résultat confirmé par des calculs de dispersion, non montrés). En revanche, les B-CRPS-PP et Be-CRPS-PP ayant une pente se rapprochant plus de celle de la bissectrice que le « Mean PP », l'optimisation de la méthode PP par le CRPS tend à améliorer aussi la dispersion de l'ensemble.

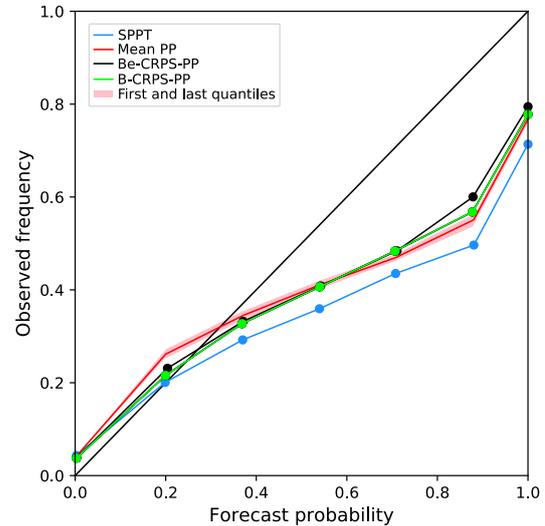
Concernant les précipitations cumulées pendant 3h, l'aire sous la courbe ROC est aussi améliorée (augmentée) avec les méthodes de perturbation de paramètres. Cependant, le gain est beaucoup plus modéré que pour le vent (Figure 8.3c). La perturbation de paramètres a donc un impact plus modéré sur la pluie que sur le vent, et ce même pour d'autres scores que le CRPS. Le Be-CRPS-PP en revanche, se distingue bien plus, de manière positive, de son homologue sur l'année (B-CRPS-PP).

A contrario, la différence entre les méthodes de perturbations de paramètres et le SPPT est bien plus marquée pour le diagramme de fiabilité des précipitations (Figure 8.3d). Le diagramme de fiabilité est dégradé avec les méthodes de perturbations de paramètre : la pente étant moins proche de la bissectrice. En revanche, quelle que soit la méthode, un biais positif non négligeable est présent, supposant que la PEARO prévoit trop de pluie. Cependant, notons que les PP optimisés tendent aussi à corriger ce biais.

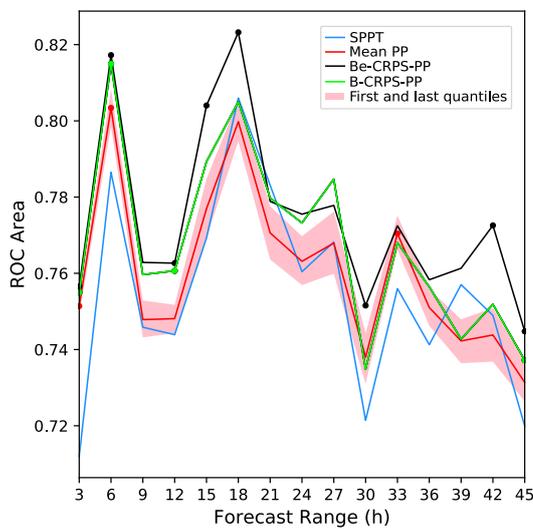
Comme constaté sur les résultats des taux d'amélioration du CRPS et du rapport Dispersion/Erreur, d'autres scores sont améliorés sur la période hivernale, par exemple la différence d'aire sous la courbe de ROC pour le vent à 10m est bien plus prononcée (Figure 8.4a). Contrairement aux résultats sur la période estivale, les scores de précipitations accumulées sur 3h sont meilleurs pour les méthodes de



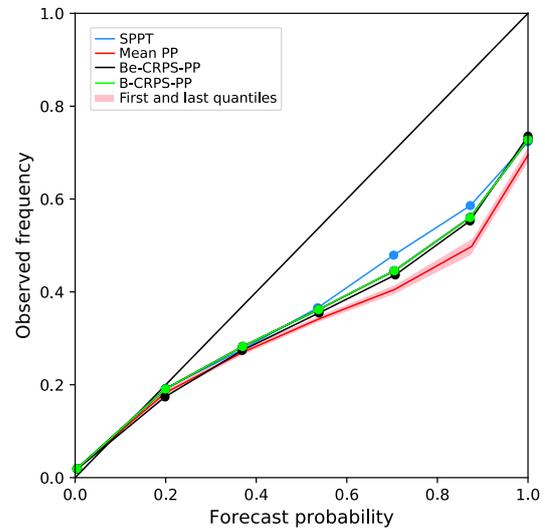
(a) Aire ROC de vent



(b) Diag. de fiab. de vent

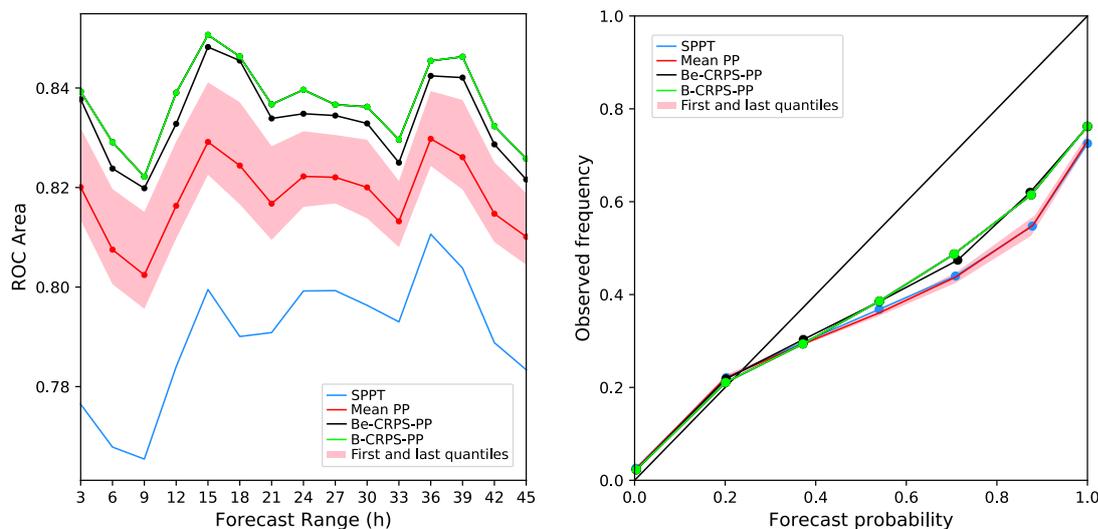


(c) Aire ROC de pluie



(d) Diag. de fiab. de pluie

FIGURE 8.3 – Impact de l'optimisation, pendant la période estivale, de la méthode PP sur les scores de la force de vent à 10m : (a) Aire sous la courbe ROC, (b) Diagramme de fiabilité (seuil : 20km/h) ; et de pluie accumulée pendant 3h : (c) Aire sous la courbe de ROC, (d) Diagramme de fiabilité (seuil : 1mm). Les scores du SPPT (bleu), Mean PP (Rouge), des scores maximaux et minimaux des 1000 PP (enveloppe rose), du Be-CRPS-PP (noir) et du B-CRPS-PP (vert) sont tracés en fonction des échéances.



(a) Aire ROC de vent

(b) Diag. de fiab. de pluie

FIGURE 8.4 – Impact de l'optimisation, pendant la période hivernale, de la méthode PP sur (a) l'aire sous la courbe ROC pour la force de vent à 10m (seuil : 20km/h) et (b) le diagramme de fiabilité de pluie accumulée pendant 3h (seuil : 1mm). Les scores du SPPT (bleu), Mean PP (Rouge), des scores maximaux et minimaux des 1000 PP (enveloppe rose), du Be-CRPS-PP (noir) et du B-CRPS-PP (vert) sont tracés le long des échéances.

perturbations de paramètres optimisées (Figure 8.4b). Par ailleurs, notons que si même le Be-CRPS-PP est optimisé sur la période estivale, celui-ci propose tout de même des scores du même ordre que son équivalent annuel.

**Bilan de l'optimisation du PP** Ainsi, les méthodes PP améliorent l'ensemble des scores de vent. L'optimisation permet un gain plus important même sur les autres scores. En revanche, en ce qui concerne la pluie, l'amélioration est plus modérée en été qu'en hiver et l'optimisation selon le CRPS semble nécessaire.

L'amélioration de l'ensemble des scores étant plus limitée sur la période estivale, nous nous intéresserons donc par la suite aux scores sur cette période.

### 8.3 Optimisation de la méthode RPP

Après s'être intéressé à l'effet de l'optimisation des méthodes PP sur les scores principaux d'une prévision d'ensemble, nous allons nous focaliser sur l'impact de l'optimisation de la méthode RPP.

#### Utilisation de distributions uniques ou personnalisées selon les membres :

La figure 8.5 présente le CRPS de vent à 10m, l'aire sous la courbe ROC selon les échéances et le diagramme de fiabilité de précipitations dépassant 1mm pour le B-CRPS-PP ainsi que les RPP utilisant des distributions de paramètres uniformes (RPP ; magenta), les distributions gaussiennes autour du B-CRPS-PP (oRPP Gaussien ; orange) et les distributions du B-CRPS-PP (oRPP-B-CRPS-PP ; gris).

La totalité des méthodes de perturbation de paramètres proposent des CRPS de vent à 10m plus faibles que ceux du SPPT (bleu) quelle que soit l'échéance (Figure 8.5a), et ce de manière significative (point apparaissant pour quasiment toutes les échéances). Cependant cette amélioration est beaucoup plus modérée pour les méthodes RPP utilisant une unique distribution de paramètre pour tous les membres, qu'elle soit uniforme (RPP ; magenta) ou personnalisée (oRPP-B-CRPS-PP ; gris). Cette dernière utilisant les distributions issues du B-CRPS-PP tend même vers la méthode SPPT pour les longues échéances.

En revanche, on constate une amélioration du CRPS, comparé au SPPT, de près de 6% avec le RPP utilisant des distributions gaussiennes autour du B-CRPS-PP (oRPP Gaussien ; orange). Le CRPS de vent étant très peu dégradé (ainsi que celui de pluie, non montré) entre le B-CRPS-PP et le oRPP Gaussien, ceci suppose que les distributions gaussiennes autour du B-CRPS-PP permettent de ne pas trop s'éloigner de cet optimum de CRPS de vent et pluie. Par ailleurs, la meilleure performance de cette optimisation gaussienne du RPP comparée à celle utilisant des distributions uniques de paramètre pour tous les membres, suggère une nécessité d'optimiser des distributions de paramètres pour chacun des membres de la prévision d'ensemble plutôt que d'utiliser une distribution de paramètres valide pour l'ensemble des membres.

En ce qui concerne les précipitations, la courbe ROC (Figure 8.5b) montre une similitude entre les méthodes RPP utilisant une unique distribution pour tous les membres et le SPPT, alors que ceux utilisant une optimisation selon chaque membre (B-CRPS-PP et oRPP Gaussien) se distinguent par une aire sous la courbe ROC souvent significativement meilleure que celle du SPPT. B-CRPS-PP et oRPP Gaussien semblent être donc plus pertinents pour la discrimination de deux événements contrairement aux méthodes RPP utilisant des distributions uniques pour chaque membre. Le diagramme de fiabilité de ces ensembles RPP avec distributions uniques (RPP et oRPP-B-CRPS-PP) montrent une claire dégradation de la fiabilité comparée à celle du SPPT alors que les B-CRPS-PP et oRPP Gaussien gardent un diagramme de fiabilité similaire à celui du SPPT (Figure 8.5c).

**Comparaison aux distributions du plan d'expérience de Morris :** Concernant les RPP optimisant des distributions de paramètres uniques pour tous les membres de la prévision d'ensemble, différentes distributions ont été élaborées à partir du plan d'expérience de Morris. Afin de vérifier que les distributions issues du B-CRPS-PP et des 50B-CRPS-PP diffèrent de celles du plan d'expérience de Morris, nous comparons ici les scores de ces prévisions d'ensemble à ceux du oRPP utilisant les distributions issues du plan d'expérience de Morris. La figure 8.6a représente le CRPS de vent à 10m en fonction des échéances du RPP utilisant des distributions uniformes (RPP ; magenta), celui avec le plan d'expérience de Morris (oRPP-Morris ; bleu clair), avec les distributions du B-CRPS-PP (oRPP-B-CRPS-PP ; gris) et des 50B-CRPS-PP (oRPP-50B-CRPS-PP ; jaune). Quelle que soit l'échéance considérée, une amélioration significative comparée au SPPT (bleu) est constatée. Cependant, cette différence est plus prononcée pour le oRPP utilisant les distributions issues des 50B-CRPS-PP (oRPP-50B-CRPS-PP ; jaune). Ce dernier étant plus performant que la version utilisant des distributions uniformes (RPP ; magenta) et celle utilisant les distributions issues du plan d'ex-

périence de Morris (oRPP-Morris ; bleu clair), ce résultat prouve une distinction entre les distributions des 50B-CRPS-PP et celles du plan d'expérience de Morris. En revanche, le fait que le oRPP utilisant les distributions du B-CRPS-PP (oRPP-B-CRPS-PP ; gris) ait un CRPS plus élevé que ces homologues pose question quant à sa réelle optimisation. Pour ce qui est des précipitations, aucun des oRPP n'offre une meilleure fiabilité que le SPPT (Figure 8.6b).

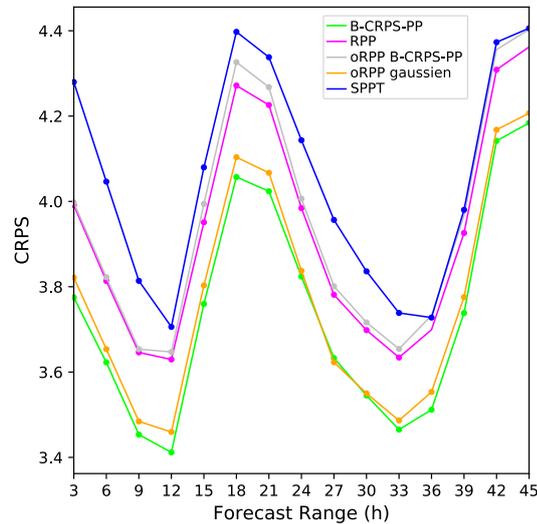
**Évaluation des émulateurs de CRPS :** En parallèle des optimisations de distributions de paramètres à partir du plan d'expérience de Morris, d'autres optimisations basées sur un émulateur de CRPS ont été testées. Puisque le CRPS varie très peu pour les précipitations entre les différents ensembles, nous nous basons uniquement sur ce score du vent pour valider l'utilisation d'émulateur.

La figure 8.7 représente le CRPS de vent à 10m en fonction des échéances pour les RPP utilisant des distributions uniformes (RPP ; magenta), celles du B-CRPS-PP (oRPP-B-CRPS-PP ; gris), celles issues des CRPS émulés par un processus gaussien utilisant un noyau de Matern (oRPPMat ; bleu clair) et celui utilisant le noyau DotProduct (oRPPDot ; vert). Nous constatons que quelles que soient les distributions de paramètres utilisées, ces oRPP peinent à reproduire des CRPS aussi faibles que la version utilisant des distributions uniformes (RPP ; magenta). Le RPP utilisant les distributions issues du processus gaussien avec le noyau de covariance DotProduct (oRPPDot) présente même un CRPS plus élevé que celui du SPPT, questionnant la réelle efficacité de l'émulateur avec le noyau de covariance DotProduct. A l'inverse, notons que celui utilisant le noyau de Matern (bleu) a un CRPS proche voire meilleur en fin d'après-midi que celui du oRPP-B-CRPS-PP. Ce résultat montre que l'utilisation d'un émulateur ne permet de faire mieux que le RPP non optimisé.

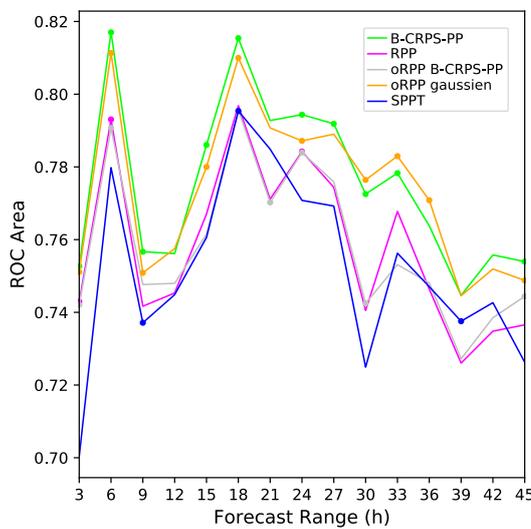
**Bilan de l'optimisation du RPP** Ainsi, de ces différents résultats, nous pouvons en conclure que le RPP présentant les meilleurs scores est celui utilisant des distributions gaussiennes autour du B-CRPS-PP (oRPP Gaussien). Ce résultat suggère donc que l'optimisation des distributions de probabilité des paramètres doit se faire pour chaque membre et non sur une même distribution valide pour tous les membres de la prévision d'ensemble. Ceci permet d'optimiser les interactions non négligeables entre paramètres.

## 8.4 Configuration opérationnelle

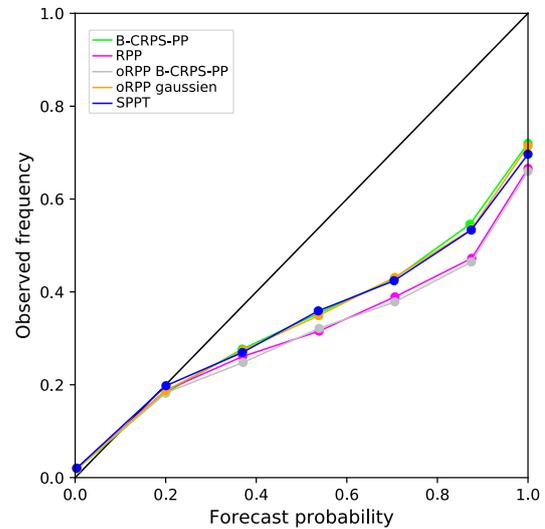
Les résultats présentés jusqu'à présent s'intéressaient aux méthodes de représentations des erreurs de modélisation seules. Or, il est nécessaire de vérifier que ces résultats sont toujours valides dans une configuration opérationnelle de la PEARO, c'est-à-dire intégrant la perturbation des conditions initiales, latérales et de surface. Par ailleurs, notons qu'en plus de l'ajout de la représentation de ces autres erreurs, la version opérationnelle de la PEARO utilise une version non hydrostatique d'AROME avec activation du schéma prédicteur-correcteur. La figure 8.8 présente la dispersion pour la force du vent à 10m et les pluies accumulées pendant 3h, ainsi que le CRPS de vent à 10m en fonction des échéances pour les



(a) *CRPS de vent*



(b) *Aire ROC de pluie*



(c) *Diag. de fiab. de pluie*

FIGURE 8.5 – Impact de l'optimisation, pendant la période estivale, de la méthode RPP sur le CRPS de la force de vent à 10m (a) CRPS, (b) l'aire sous la courbe ROC des précipitations accumulées pendant 3h (seuil : 1mm) et (c) le diagramme de fiabilité des précipitations sur 3h (seuil : 1mm). Les scores du SPPT (bleu), B-CRPS-PP (vert), RPP (magenta), oRPP-B-CRPS-PP (gris) et oRPP (jaune) sont tracés le long des échéances.

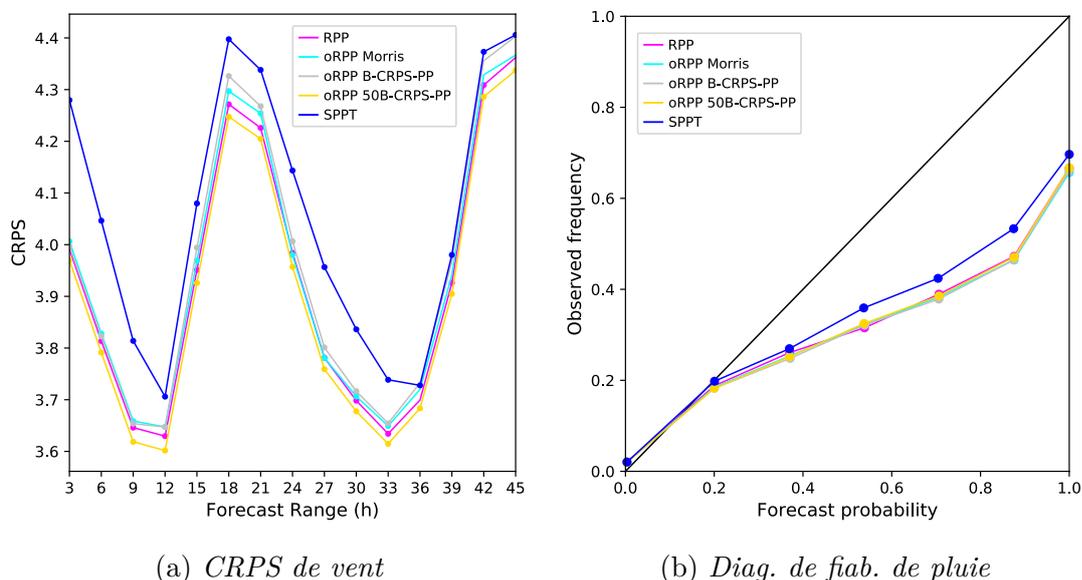


FIGURE 8.6 – Impact de l’optimisation, pendant la période estivale, de la méthode RPP sur (a) le CRPS de la force de vent à 10m et (b) sur le diagramme de fiabilité de pluie accumulée pendant 3h (seuil : 1mm) . Les scores du SPPT (bleu) RPP (magenta), oRPP-Morris (bleu clair), oRPP-B-CRPS-PP (gris) et oRPP-50B-CRPS-PP (jaune) sont tracés le long des échéances.

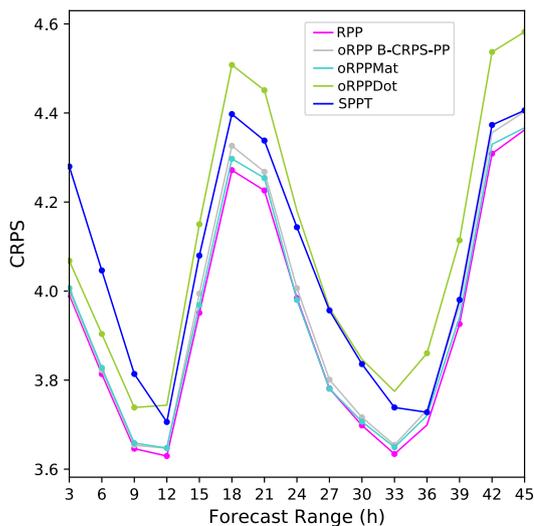


FIGURE 8.7 – Impact de l’optimisation, pendant la période estivale, de la méthode RPP sur le CRPS de la force de vent à 10m. Les scores du SPPT (bleu), RPP (magenta), oRPP-B-CRPS-PP (gris), oRPPMat (bleu clair) et oRPPDot (vert) sont tracés le long des échéances.

méthodes B-CRPS-PP (vert foncé) et SPPT (bleu clair) dans leur version opérationnelle. Afin d'estimer l'impact de l'ajout de ces perturbations des conditions initiales, de surface et latérales, les scores des ensembles B-CRPS-PP (vert) et SPPT (bleu) où seule l'erreur de modélisation est représentée, sont aussi tracés.

Comme attendu, quelle que soit la variable météorologique étudiée, l'ajout de perturbations des conditions initiales, de surface et latérale, augmente la dispersion de la PEARO. Là où le B-CRPS-PP seul avait déjà une dispersion plus élevée que celle du SPPT, sa configuration opérationnelle est aussi plus dispersive que celle du SPPT. La dispersion apportée par la perturbation des conditions initiales, de surface et latérales est plutôt constante entre les deux méthodes pour les précipitations (Figure 8.8b). En revanche, pour le vent à 10m (Figure 8.8a), cet apport diffère selon les échéances et les deux méthodes. Avec la méthode SPPT, l'apport de la dispersion est quasiment constant de 1.5km/h pour toutes les échéances, alors qu'avec la méthode B-CRPS-PP, il passe de 0.5km/h à 1.5km/h entre les premières et dernières échéances. Ceci est en partie dû aux fortes valeurs de dispersion en vent, déjà présentes avec la méthode B-CRPS-PP seule pour les premières échéances. En effet, ceci limite la contribution de la représentation des erreurs de conditions initiales, latérales et surfaciques sur la dispersion de vent en début de prévision.

Notons, par ailleurs, que la configuration opérationnelle du SPPT offre une dispersion de vent à 10m proche de celle du B-CRPS-PP seul aux premières échéances. De ce fait, la représentation de l'erreur de modélisation par la perturbation de paramètres seule semble disperser autant le vent qu'un ensemble représentant à la fois les erreurs de conditions initiales, de surface, latérales, mais aussi de modélisation par la méthode SPPT. Ce même phénomène (même dispersion pour le B-CRPS-PP seul que SPPT opérationnel) apparaît aussi pour la dispersion de pluie pour les échéances au delà de 15h. Notons que ce résultat n'est vérifié que pour la dispersion et non sur d'autres scores de pluie, alors qu'ils sont aussi vérifiés pour la totalité des scores de vent à 10m pour les premières échéances. Par exemple, sur la figure 8.8c, le B-CRPS-PP seul présente un CRPS de vent meilleur (plus petit) que celui du SPPT utilisé dans une configuration opérationnelle.

Les méthodes B-CRPS-PP et SPPT ayant des caractéristiques différentes (amplitude différente de la perturbation selon la verticale), il peut être aussi intéressant d'étudier leur combinaison. Ainsi, sur la figure 8.8, les courbes rouges correspondent aux scores pour la configuration opérationnelle de la PEARO où l'erreur de modélisation est représentée à la fois par la méthode SPPT et à la fois par le B-CRPS-PP. Notons que, quelle que soit la variable météorologique considérée (vent à 10m ou précipitation) ou le score (CRPS, dispersion, ou aussi d'autres scores non montrés), l'ensemble opérationnel combinant les deux approches (courbe rouge) est relativement similaire à l'ensemble opérationnel n'utilisant que la perturbation de paramètres (vert foncé). La méthode SPPT disperse donc moins la prévision d'ensemble sur les variables météorologiques de surface. Comme déjà expliqué, ceci s'explique probablement par l'absence de perturbation des tendances physiques en basses couches pour des raisons de stabilités numériques.

**Bilan de la configuration opérationnelle :** Les gains obtenus sur les ensembles considérant seulement l'erreur de modélisation restent valables dans une configuration opérationnelle. Par ailleurs, la combinaison des deux méthodes montre

quelques limites. D'un point de vue du B-CRPS-PP, l'ajout du SPPT est légèrement positif mais non significatif, voire dégrade la dispersion de précipitations entre 30 et 39h d'échéances (Figure 8.8b). En revanche, du point de vue du SPPT, ajouter une perturbation de paramètres permet d'augmenter la dispersion de l'ensemble, a minima, au niveau de la surface. Cet ajout permet de palier le manque de dispersion du SPPT dans les basses couches de l'atmosphère. Cependant, les scores n'ayant été étudiés que sur des variables proches de la surface, une étude plus approfondie des différences entre SPPT et B-CRPS-PP sur la verticale pourrait permettre de mieux comprendre la complémentarité (ou peut-être redondance) des deux méthodes en altitude où le SPPT peut mieux disperser.

## 8.5 Caractéristique des différentes représentations de l'erreur de modélisation

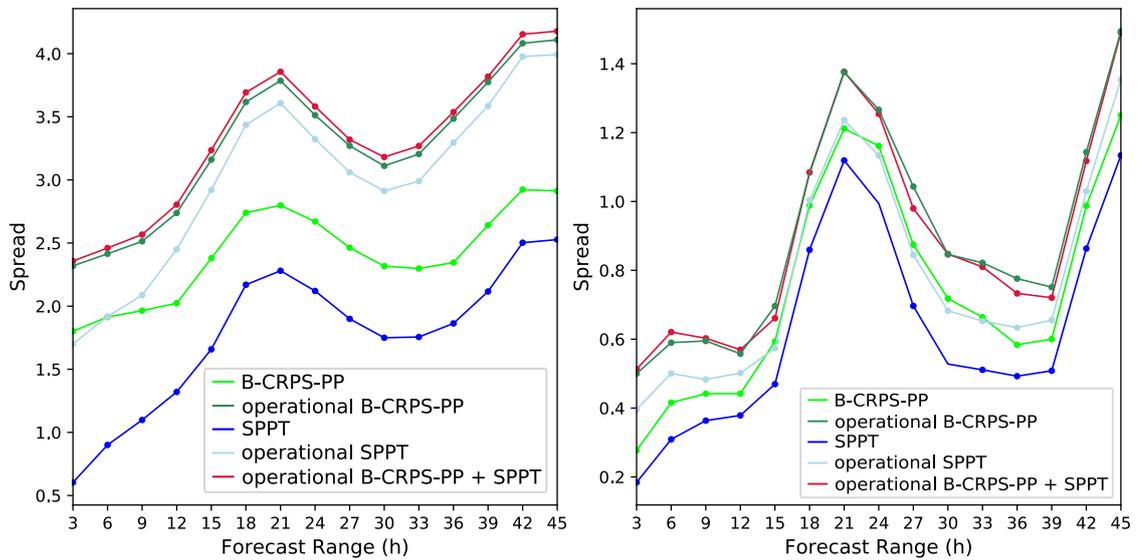
### 8.5.1 Biais des membres du B-CRPS-PP

La méthode PP et sa version optimisée utilisant des valeurs de paramètres différentes pour chacun de ses membres, cette technique peut être vue comme une approche multi-physiques où chaque membre comprend des paramétrisations avec des réglages différents. Ainsi, chaque membre a une configuration physique unique et peut présenter un biais différent. Même si une telle prévision d'ensemble présente un faible biais en moyenne sur les membres, il est nécessaire de vérifier que chaque membre ne présente pas des biais compensatoires trop importants.

Le biais de vent, rafales et précipitations accumulées pendant 3h sont présentés, pour la période estivale, sur la figure 8.9, en fonction des échéances pour chacun des membres, ainsi que pour un membre utilisant les valeurs par défaut des paramètres du modèle AROME (ligne noire tiretée).

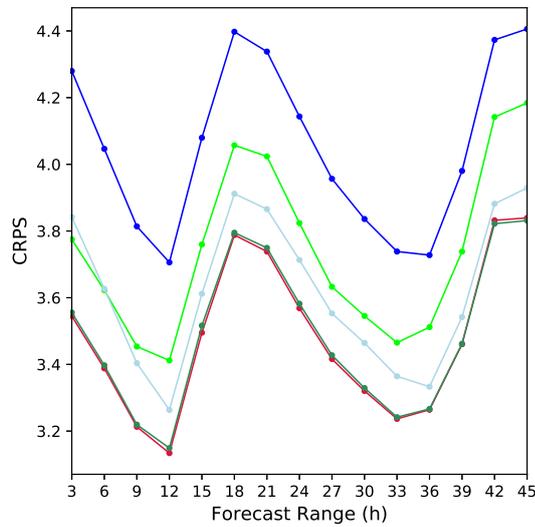
Concernant la force du vent à 10m (Figure 8.9a), remarquons que le biais du membre avec les valeurs par défaut est fortement lié au cycle diurne. Celui-ci est positif la nuit, supposant une sur-estimation du vent la nuit, alors que le biais négatif en journée suggère une sous-estimation. Ce même cycle diurne est présent pour l'ensemble des membres du B-CRPS-PP. Cependant, comparé au membre utilisant les valeurs par défaut, les membres 1 et 10 accentuent la sur-estimation du vent la nuit, mais sous-estiment le vent la journée, avec une amplitude similaire à ce membre par défaut. Notons que ces deux membres présentent la même valeur élevée de XFRACZ0, coefficient de réglage de la traînée orographique, à 7.7 (rappelons que XFRACZ0 est perturbé entre 2 et 10). Ceci suppose alors une longueur de rugosité plus faible pouvant expliquer la sur-estimation du vent la nuit. Cependant, il s'agit aussi de deux des trois membres à avoir les plus faibles valeurs de XCED (0.4 - 0.6 ; défaut : 0.85) dans le B-CRPS-PP. Rappelons que puisque XCED contrôle la dissipation de l'énergie cinétique turbulente, celui-ci a plus d'impact la nuit, comme montré par l'analyse de sensibilité. La valeur de XCED étant faible, l'énergie cinétique turbulente est moins dissipée, ce qui implique de plus forts vents et explique le biais positif de vent pour ces membres durant la nuit.

A l'inverse, l'ensemble des autres membres exceptés les membres 4 et 8, utilisent



(a) Force du vent à 10m (km/h)

(b) Pluie cumulée pendant 3h (mm)



(c) CRPS de vent à 10m

FIGURE 8.8 – Dispersion de (a) la force de vent à 10m et (b) des précipitations accumulées pendant 3h ainsi que (c) CRPS de vent à 10m, en fonction des échéances pendant la période estivale 2018 pour le B-CRPS-PP (vert), le B-CRPS-PP dans la configuration opérationnelle (vert foncé), SPPT (bleu), SPPT dans la configuration opérationnelle (bleu clair) ainsi que le B-CRPS-PP et SPPT combinés dans la configuration opérationnelle (rouge)

la valeur de XCED maximale de 2, ce qui a pour conséquence de mieux dissiper l'énergie cinétique turbulente et d'améliorer le biais de vent la nuit. Notons que le membre 4 a un biais se rapprochant le plus de celui utilisant les valeurs par défaut. Celui-ci utilise une valeur de 1.5 pour XCED bien supérieure à la valeur par défaut (0.85). Ainsi, augmenter la valeur de XCED n'est pas l'unique solution pour améliorer le biais de vent la nuit.

En journée, les membres 2, 5, 7, 8 et 12 présentent en revanche, un très fort biais négatif de vent. Ils ont tous en commun d'avoir de fortes valeurs de XCEP et XCTP, deux paramètres identifiés comme ayant un fort impact sur le vent en journée par l'analyse de sensibilité. Cependant, le membre 3 ayant aussi des valeurs fortes pour ces deux paramètres, présente un meilleur biais en vent. En revanche, ce membre utilise une valeur de XFRACZ0 beaucoup plus élevée que les membres 2, 5, 7, 8 et 12. Des interactions doivent alors avoir lieu entre ces paramètres, expliquant le meilleur biais de ce membre.

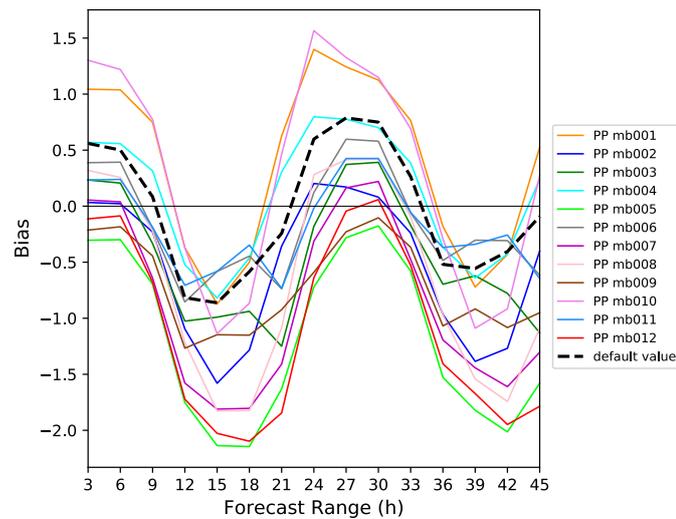
Concernant les rafales (Figure 8.9b), la plupart des membres présentent un meilleur biais que le membre utilisant les valeurs par défaut. Seuls les trois membres utilisant les trois plus faibles valeurs de XCED (membres 1, 8 et 10) ont un biais en rafales du même ordre que le membre par défaut. Rappelons que les rafales sont calculées à partir de l'énergie cinétique turbulente, elle-même contrôlée par XCED. L'ensemble des autres membres ayant des valeurs plus élevées de XCED que le membre par défaut, ceci suggère qu'une augmentation de XCED serait bénéfique pour le biais de rafales du modèle AROME.

Du point de vue du biais des précipitations (Figure 8.9c), la plupart des membres du B-CRPS-PP ont un biais légèrement négatif. Le seul membre présentant un fort biais positif en fin d'après-midi est le membre 9. Cette caractéristique de ce membre semble clairement liée à la convection et à sa très faible valeur de XCMF, contrôlant le flux de masse au premier niveau de l'atmosphère.

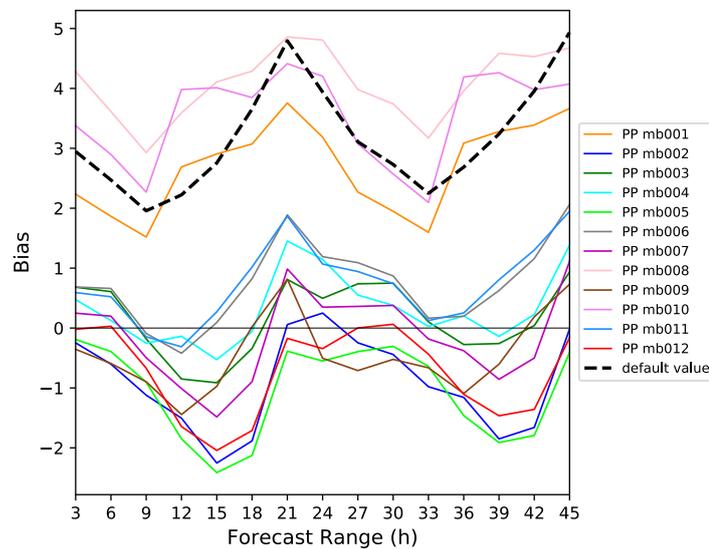
D'un point de vue ensembliste, que ce soit pour les précipitations ou le vent ou les rafales, les différents membres du B-CRPS-PP présentent chacun un biais spécifique. Ceci est parfois un avantage comme pour les rafales où la plupart des membres corrigent le biais positif de rafales constaté dans le modèle AROME. Cependant, pour les trois variables météorologiques étudiées, des effets compensatoires ont lieu. Par exemple, pour les précipitations, ces compensations suggèrent un biais majoritairement négatif pour l'ensemble B-CRPS-PP alors qu'un membre (n°9) propose des précipitations trop élevées. Ainsi, ces résultats montrent alors que les membres du B-CRPS-PP ne sont pas équiprobables. Notons que même si la moyenne du biais de précipitation est légèrement négative, la sous-estimation des pluies est plus prononcée pour le biais moyen du SPPT (non montré), suggérant une amélioration du biais avec le B-CRPS-PP malgré les effets compensatoires des biais de chaque membre.

Cependant, les résultats montrés ici, ne s'intéressent qu'à la variation selon les échéances et ne montrent pas si ces biais sont systématiques ou correspondent plutôt à une journée particulière. La figure 8.10 présente donc le biais de vent à 10m tout au long de la période estivale, à l'échéance 21h. Quel que soit le jour considéré, le biais de chaque membre reste plutôt uniforme, supposant un biais systématique des membres du B-CRPS-PP. Ceci représente donc un inconvénient de cette méthode.

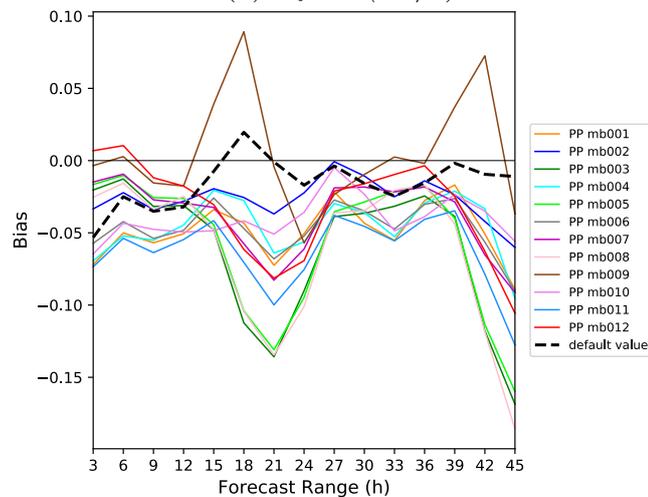
## CHAPITRE 8. ÉVALUATION DES NOUVELLES MÉTHODES DE REPRÉSENTATION DES ERREURS DE MODÉLISATION



(a) force du vent à 10m (km/h)



(b) rafales (km/h)



(c) pluie accumulée pendant 3h (mm)

FIGURE 8.9 – Biais selon les échéances, pendant la période estivale, du (a) vent à 10m (b), rafales, et (c) pluie accumulée pendant 3h, pour les différents membres du B-CRPS-PP. La ligne noire tiretée correspond au biais pour un membre utilisant les valeurs par défaut des paramètres (correspondant au modèle AROME)

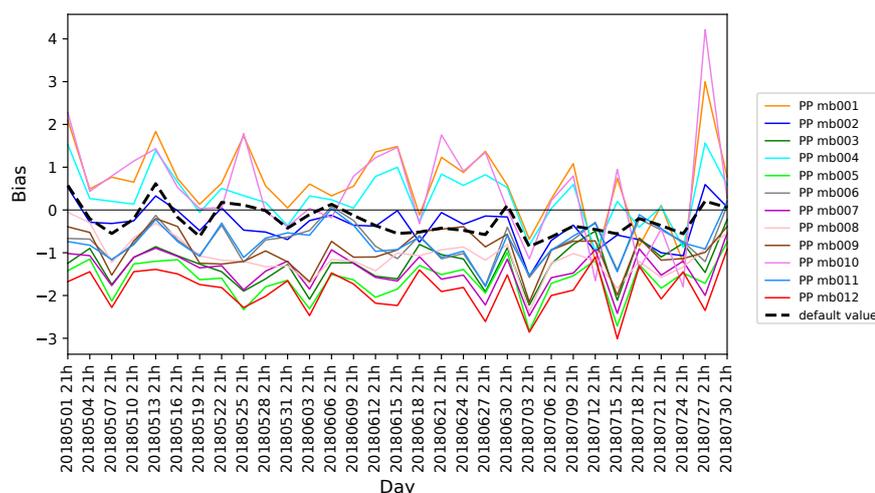


FIGURE 8.10 – Série temporelle du biais de vent à 10m (km/h), à 21hUTC pendant la période estivale pour les différents membres du B-CRPS-PP. La ligne noire tiretée correspond au biais pour un membre utilisant les valeurs par défaut des paramètres (autrement dit correspondant au modèle AROME).

### 8.5.2 Différence de dispersion entre l’été 2018 et 2020

L’optimisation des PP et des distributions des RPP ayant été réalisée sur l’année 2018, nous avons aussi vérifié que les améliorations par rapport au SPPT constatées sur cette période sont aussi valides sur une autre période, l’été 2020. Ainsi, l’ensemble des scores des différentes méthodes de perturbations des paramètres ont été calculées sur 31 jours allant du 1er mai 2020 au 30 Juillet 2020 (1 jour sur 3). Nous ne détaillons pas ici la totalité des résultats obtenus car de manière générale, les scores et leur conclusion sont proches de ceux obtenus sur l’été 2018. Cependant, nous illustrons ici les résultats principaux sur la dispersion et la RMSE du oRPP Gaussien. Ainsi, la figure 8.11a montre le taux d’amélioration, par rapport au SPPT, de la dispersion et de la RMSE sur les étés 2018 et 2020 pour les différentes variables météorologiques. Quelle que soit la variable météorologique considérée, le taux d’amélioration de la RMSE (vert) reste relativement constant entre les deux étés et augmente au maximum de 2.5% pour l’humidité relative lors de l’été 2020. Alors que la dispersion (orange) voit son taux d’amélioration presque doubler pour le vent et les rafales à 10m ainsi que la température à 2m pour l’été 2020. Pour les autres variables météorologiques, en revanche, l’amélioration est majoritairement limitée aux environs de 10%. Ces résultats montrent alors une plus grande dispersion du oRPP Gaussien, comparée au SPPT, lors de l’été 2020.

Cette différence entre les deux étés intervenant uniquement sur la dispersion et non la RMSE, des séries temporelles de dispersion de vent à 10m le long de ces étés sont présentés en figures 8.11b (2018) et 8.11c (2020), pour le SPPT (bleu) et le oRPP Gaussien (rose). Notons que pour l’été 2018, la dispersion du oRPP est beaucoup plus proche du SPPT que pour l’été 2020 où le SPPT peine à rattraper la dispersion du oRPP. D’un autre point de vue, le SPPT a une dispersion beaucoup plus variable en 2018 que le oRPP qui reste assez constant sur la période. Ceci suggère alors une plus grande sensibilité du SPPT aux situations météorologiques de 2018.

En effet, l'été 2018 ayant une activité convective plus importante que celui de 2020, les phénomènes sous-maîles ont eu un impact plus important sur les prévisions. Les contributions des paramétrisations physiques aux tendances de température, vent et humidité ont donc été plus grandes. Le SPPT, représentant l'erreur de modélisation en appliquant une perturbation multiplicative aux tendances physiques, génère une dispersion proportionnelle aux tendances physiques. Ainsi, le SPPT est plus à même de disperser les membres de la PEARO sur des étés avec une forte instabilité convective et donc avec de plus fortes tendances physiques, comme l'été 2018.

Le SPPT étant plus dispersif sur les étés convectifs, il rattrape alors les performances du oRPP sur la dispersion. En revanche, pour des étés peu convectifs tels que celui de 2020, il sera moins performant. Ainsi, l'amélioration plus importante de la dispersion constatée pour l'été 2020 avec le oRPP n'est pas liée à une meilleure performance de ce dernier sur cet été 2020, mais plutôt à de moins bons scores du SPPT sur cette période.

Afin d'étudier plus en détail les caractéristiques de la dispersion de vent entre le SPPT et les méthodes de perturbation de paramètres, nous sommes intéressés à la répartition spatiale de la dispersion pour la journée du 4 Juillet 2018. En effet, sur la figure 8.11b, le SPPT et le oRPP Gaussien présentent tous deux une dispersion équivalente à 15hUTC. Cependant, pour une dispersion équivalente, les autres scores sont améliorés pour le oRPP (CRPS, ROC,... non montrés), suggérant une meilleure représentation de la dispersion avec la méthode des paramètres perturbés. Des cartes de dispersion des membres du SPPT et du oRPP, pour le 4 Juillet 2018 à 15hUTC, sont alors présentées en figures 8.11d et 8.11e respectivement.

Tout d'abord notons une plus forte dispersion sur l'ensemble du domaine AROME-France avec la méthode oRPP. Remarquons que ce résultat semble être en contradiction avec la dispersion équivalente entre les deux méthodes de représentation de l'erreur de modélisation observée dans la figure 8.11b. Cependant, sur cette figure, il s'agit de la dispersion calculée aux stations d'observations, principalement situées sur terre et en France. Or, sur ces zones, la différence de dispersion entre les deux méthodes est moins prononcée que sur la Méditerranée ou encore l'Espagne. Ceci revient donc bien à une dispersion proche entre les deux méthodes sur terre et en France alors que lorsque l'on prend en compte la dispersion en mer et sur le reste de l'Europe, le oRPP présente une dispersion plus importante. Ce résultat montre, par ailleurs, les limites des calculs de scores à partir de données de stations au sol.

La dispersion étant certes plus élevée sur l'ensemble du domaine avec le oRPP, le SPPT offre cependant une dispersion plus étalée dans l'espace que le oRPP. En effet, la large zone de forte dispersion située sur le Golfe de Gascogne est plus étendue avec le SPPT qu'avec le oRPP. Un autre exemple se situe sur des petites zones de fortes valeurs de dispersion sur le Sud-Ouest de l'Allemagne qui sont plus étendues encore une fois avec le SPPT. Ainsi, le oRPP a tendance à plus disperser le vent sur la globalité du domaine que le SPPT. Mais en ce qui concerne les zones de forte sensibilité au vent, oRPP et SPPT donnent des dispersions équivalentes mais plus ciblées dans l'espace avec oRPP.

De ce fait, que ce soit d'un point de vue temporel ou spatial, nous pouvons en conclure que le oRPP disperse les membres de la prévision d'ensemble de manière

moins sensible à la situation météorologique que le SPPT.

## 8.6 Evaluation des méthodes de représentation des erreurs de modélisation sur différents cas d'étude

### 8.6.1 Cas de convection sur la France

Afin d'illustrer les différences entre le SPPT, le B-CRPS-PP et le oRPP Gaussien, nous étudions ici la prévision de l'épisode orageux du 1 Juillet 2018 par ces trois méthodes de représentation des erreurs de modélisation.

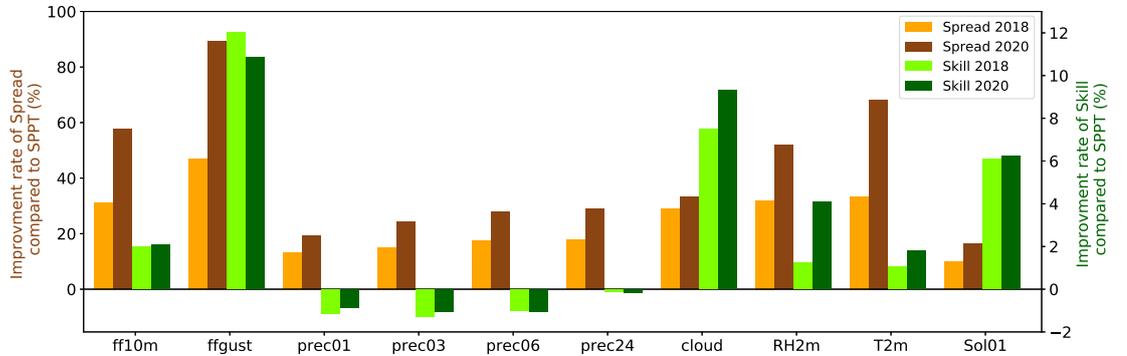
#### Situation météorologique du 01 Juillet 2018

Après une période sous influence anticyclonique, une goutte froide liée à un bas géopotentiel (voir anomalie d'hauteur du géopotentiel à 500hPa sur la Figure 8.12a), s'installe sur l'ouest de la péninsule ibérique alors que le Jet Stream est renforcé sur l'Espagne. Elle est notée par un D rouge sur la représentation schématique, dite ANASYG, d'une analyse par un prévisionniste, du 1 juillet 2018 à 12h UTC (Figure 8.12b) (signification des symboles en annexes). Un marais barométrique est alors présent sur la France, générant une situation propice au développement d'instabilité. Par ailleurs, une anomalie dynamique de tropopause (trait continu noir) et une divergence du Jet Stream sont présentes sur le Sud-Ouest de la France alors qu'une anomalie latente de tropopause (trait noir pointillé) est située sur le quart nord-ouest de la France. Ces forçages d'altitude favorisent alors les ascendances et donc la formation d'orages sur ces zones.

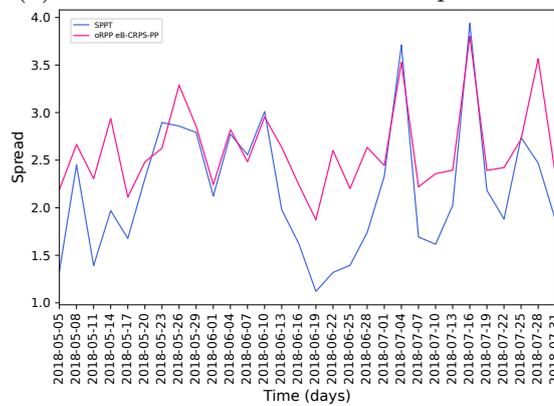
De ce fait, de nombreux orages ont eu lieu lors de la journée du 01 Juillet 2018 amenant à une vigilance orange sur l'ensemble de la côte française atlantique, du pays basque à la Bretagne (Figure 8.12c). Tout d'abord, la convection s'est enclenchée vers 15hUTC sur la Bretagne provoquant des précipitations cumulées sur l'évènement entre 40 et 80mm en moins de 2h. Des cumuls atteignant jusqu'à 42mm en 1h ont été relevés à Saint-Ségal (Finistère) (Figure 8.13a). Puis des orages ont pris plus d'importance sur l'Orne et le nord de l'Eure-et-Loir vers 18hUTC (Figure 8.13b) et se sont déplacés jusque vers le Loiret en soirée (Figure 8.13c). Ils ont provoqué des cumuls de pluie allant jusqu'à 85mm comme à St Evroult-Notre-Dame-du-Bois (Orne) et de la grêle a également été observée.

En parallèle, de violents orages supercellulaires, en provenance du Golfe de Gascogne, se sont développés sur le sud-ouest à partir de 18hUTC et ont perduré jusqu'en soirée (Figures 8.13b et 8.13c). Ils ont été responsables de forts cumuls de précipitations allant jusqu'à plus de 92 mm relevés à Soorts-Hossegor dans les Landes. Ces orages ont été particulièrement violents. Des grêlons de plus de 4cm ont été observés. On relève par ailleurs de fortes rafales entre 80 et 90km/h, allant même jusqu'à 111km/h localement à Sabres (Landes). Ces orages ont aussi été caractérisés par une forte activité électrique puisque près de 10 000 éclairs ont été comptés en 1h dont 8 000 dans les Landes en début de soirée. Ces orages isolés se

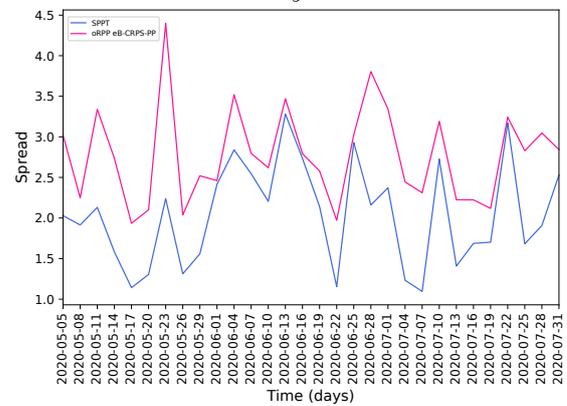
## CHAPITRE 8. ÉVALUATION DES NOUVELLES MÉTHODES DE REPRÉSENTATION DES ERREURS DE MODÉLISATION



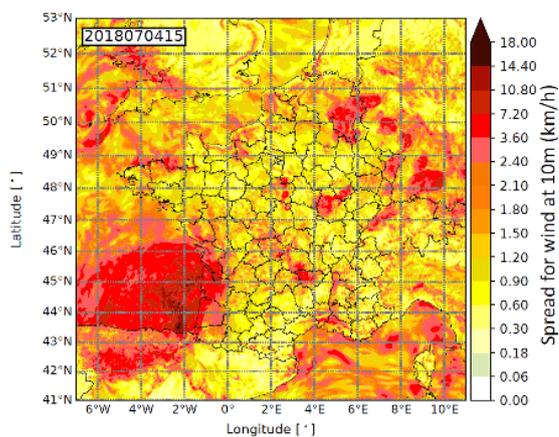
(a) Taux d'amélioration de la dispersion et de la RMSE de la moyenne de l'ensemble



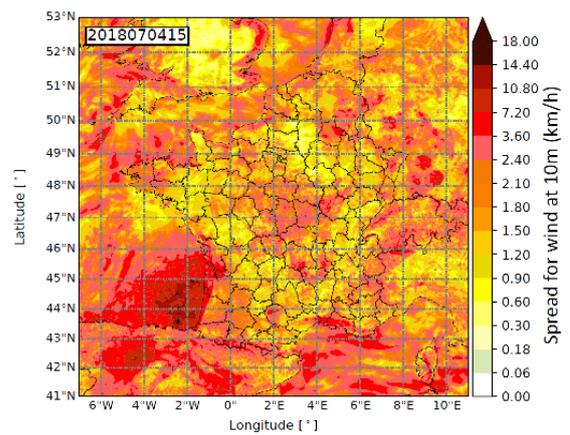
(b) Série temporelle, à 15hUTC, de la dispersion du vent à 10m sur l'été 2018 : (bleu : SPPT, rose RPP avec distributions gaussiennes)



(c) Série temporelle, à 15h UTC, de la dispersion du vent à 10m sur l'été 2020 : (bleu : SPPT, rose RPP avec distributions gaussiennes)



(d) Carte de dispersion de vent avec la méthode SPPT pour le 04/07/2018 à 15hUTC



(e) Carte de dispersion de vent avec le oRPP Gaussien pour le 04/07/2018 à 15hUTC

FIGURE 8.11 – Étude de la différence d'amélioration de la dispersion entre les étés 2018 et 2020

sont ensuite organisés sous forme d'un système convectif de méso-échelle vers le nord de l'Aquitaine alors que d'autres orages ont éclaté sur la Loire-Atlantique et la Vendée.

### **Evaluation des différentes méthodes de représentation de l'erreur de modélisation**

Afin de mettre en évidence le comportement des méthodes de représentation des erreurs de modélisation SPPT, B-CRPS-PP et oRPP Gaussien sur la prévision de ces événements orageux, nous comparons les cumuls de pluie observés sur l'intégralité de la journée du 1 juillet 2018 à ceux prévus par les ensembles utilisant ces trois méthodes. Cette observation de référence dite ANTILOPE consiste en des données de réflectivité radar converties en taux de précipitations et corrigées à partir des précipitations mesurées au sol par les pluviomètres (Laurantin, 2008). Les orages étant des phénomènes rares et intenses, l'étude des forts quantiles de la prévision d'ensemble est souvent utilisée par les prévisionnistes. Le cumul de précipitations observé sur 24h (Figure 8.13d) est alors comparé aux maxima de précipitations cumulées sur 24h prévus par la prévision d'ensemble (dit quantile maximum,  $Q_{max}$ ) ainsi qu'aux quantiles 90% (dit  $Q_{90}$ ) en chaque point de grille, tous membres confondus (Figure 8.14). Afin de ne comparer que l'impact de la méthode de représentation des erreurs de modélisation, les trois ensembles utilisent les mêmes conditions initiales, latérales et surfaciques du 30 Juin 2018 à 21hUTC.

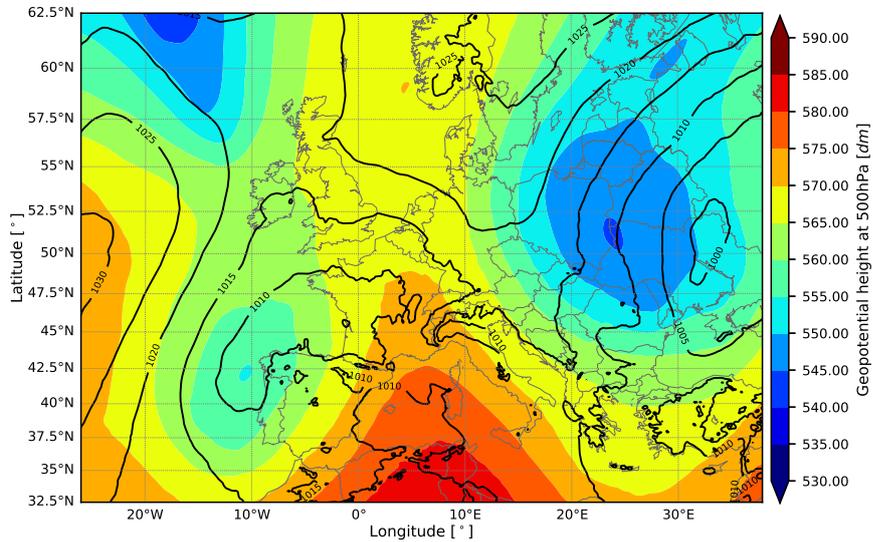
De manière générale, que ce soit pour le  $Q_{90}$  ou  $Q_{max}$ , nous notons un plus large étalement des précipitations avec le SPPT (notez, l'étalement des fortes précipitations de l'axe Bretagne-Loiret jusque vers l'Indre-et-Loire). Ceci est donc à l'avantage du B-CRPS-PP et plus particulièrement du oRPP qui semblent alors mieux cibler deux zones de précipitations : axe Bretagne/Loiret ainsi que les Landes et Pyrénées-Atlantiques. Notons, en revanche, que les trois méthodes ratent les forts cumuls de précipitations apparaissant en début de nuit sur la Loire-Atlantique.

**Convection sur la Bretagne :** Les précipitations dues aux premiers orages ayant lieu en Bretagne et se déplaçant vers l'Est sont plutôt bien représentées par les trois types de représentation des erreurs de modélisation, avec des cumuls dépassant les 80mm localement le long de l'axe.

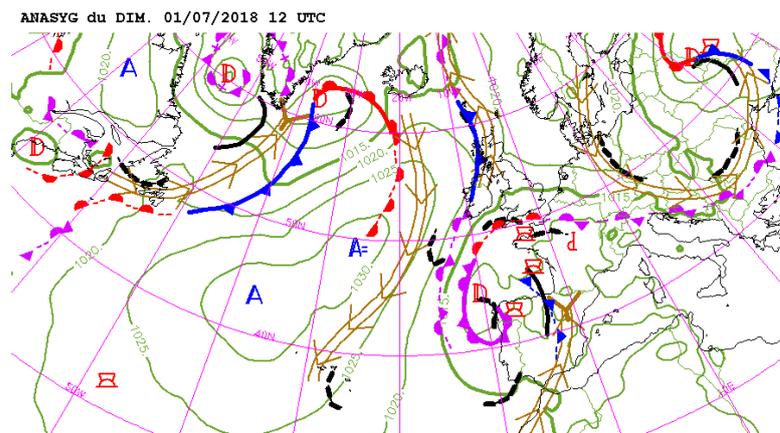
Pour ce qui est du début du phénomène, le  $Q_{90}$  de chacune des prévisions d'ensemble peine à atteindre des cumuls de précipitations se rapprochant de l'observation sur le Finistère. Le  $Q_{max}$  du SPPT présente bien des cumuls de 90 mm mais ceux-ci sont trop à l'Est. Il en est de même pour le  $Q_{max}$  du B-CRPS-PP et oRPP, mais ceux-ci proposent en revanche des  $Q_{max}$  avec des cumuls plus importants sur le bord Est du Finistère, jusqu'à 70mm pour le oRPP voire quelques zones à 90mm pour le B-CRPS-PP, contre 60mm maximum pour le SPPT.

Les trois représentations produisent, à tort, des  $Q_{max}$  et  $Q_{90}$  trop élevés sur les Côtes-d'Armor et le Morbihan allant jusqu'à 100mm pour le  $Q_{max}$  du oRPP. Les deux quantiles des B-CRPS-PP et oRPP étant supérieurs à ceux du SPPT, ceux-ci suggèrent de fortes pluies. A l'inverse, le  $Q_{90}$  de chacune des méthodes indique de plus faibles précipitations sur l'Ille-Et-Vilaine. Seul le B-CRPS-PP a un  $Q_{max}$  plus

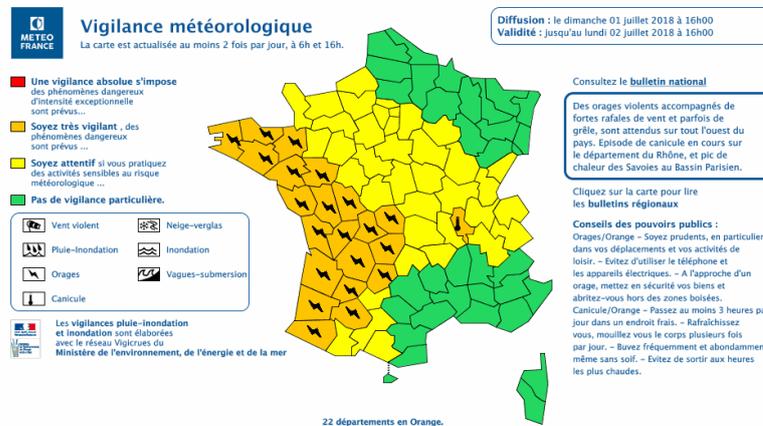
# CHAPITRE 8. ÉVALUATION DES NOUVELLES MÉTHODES DE REPRÉSENTATION DES ERREURS DE MODÉLISATION



(a) Analyse ARPEGE du 01 Juillet 2018 à 0h UTC : Hauteur du géopotiel à 500hPa (dm ;couleur) et Pression au niveau de la mer (hPa ; contour)



(b) ANASYG du 01 Juillet 2018 à 12hUTC



(c) Bulletin de Vigilance à 16h

FIGURE 8.12 – Cas d'étude du 01 Juillet 2018

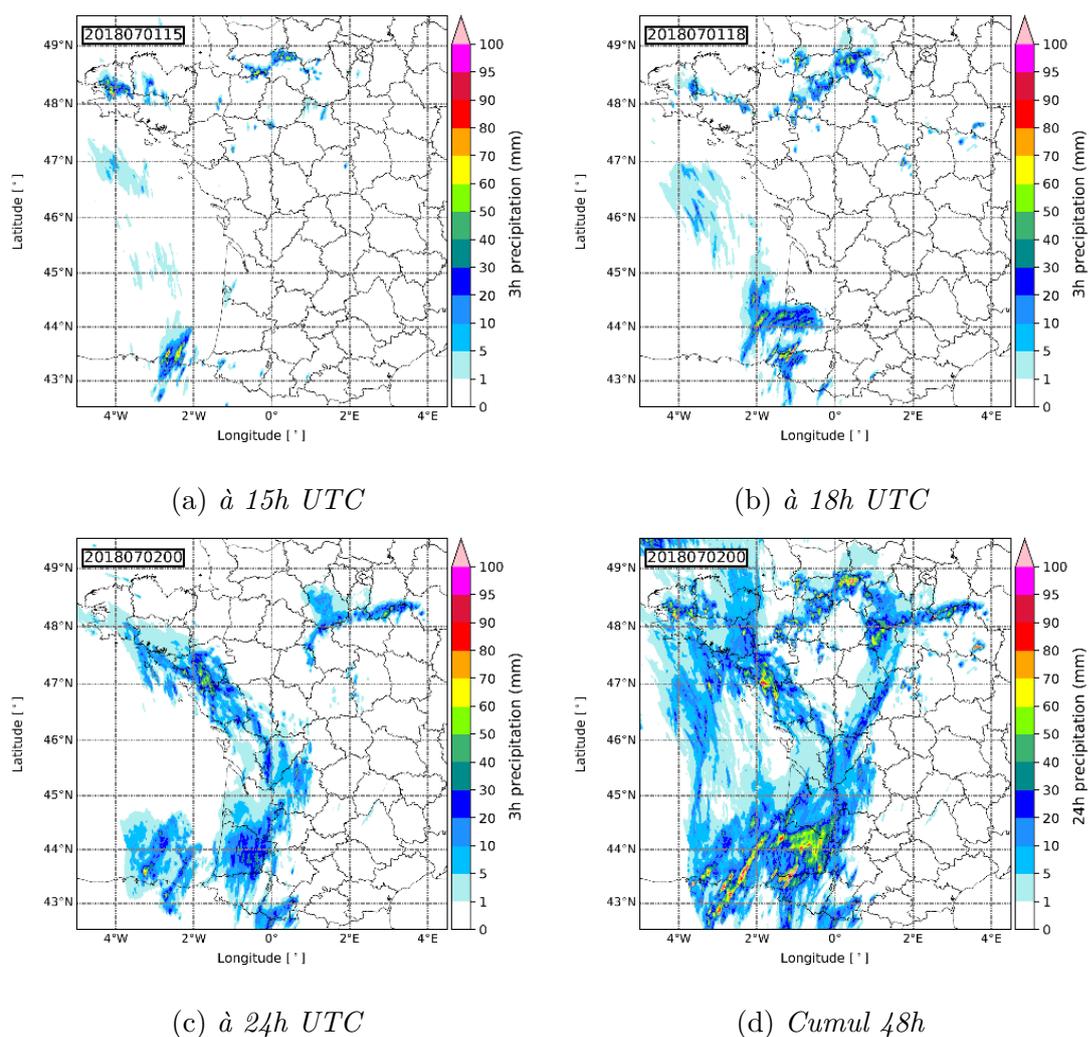


FIGURE 8.13 – Observations de précipitations cumulées sur 3h, le 01 Juillet 2018 à (a) 15h UTC, (b) 18h UTC et (c) 24h UTC ainsi qu'observations de précipitations cumulées sur 24h (d) pendant la journée du 01 Juillet 2018 (Produit ANTILOPE combinant données radar et pluviomètres)

modéré sur cette zone que ces homologues et semble mieux représenter l'absence de précipitations.

Concernant l'Orne où des cumuls de précipitations de plus de 100mm ont été observées, le SPPT peine à produire plus de 70mm, même pour le  $Q_{max}$ . Alors que le  $Q_{max}$  des B-CRPS-PP et oRPP propose des cumuls bien plus élevés, autour de 100mm. Leur  $Q_{90}$  propose des cumuls assez élevés (jusqu'à 80mm un peu trop à l'Ouest, sur la Manche, avec le B-CRPS-PP).

Enfin, seul le oRPP propose un  $Q_{max}$  jusqu'à 80mm et un  $Q_{90}$  à plus de 30mm sur le Loiret alors que les  $Q_{max}$  des SPPT et B-CRPS-PP proposent des cumuls atteignant à peine 50mm très localement. Ces derniers n'indiquent clairement pas de forts cumuls sur ce département contrairement au oRPP.

De manière générale, nous notons de très forts cumuls de précipitations avec le  $Q_{max}$  du oRPP (supérieur à 100mm) alors que son  $Q_{90}$  est plus modéré (maximum 70mm) suggérant ainsi qu'un seul membre du RPP optimisé sur-estime ces précipitations. Le  $Q_{90}$  correspondant étant moins alarmant, il permet de limiter le nombre de fausses alarmes pouvant être produites avec cette méthode.

**Supercellules sur le Sud-Ouest :** Concernant maintenant les super-cellules nées sur le Sud-Ouest, des cumuls de près de 90-100mm en 24h sont constatés sur les Landes et Pyrénées-Atlantiques avec les données ANTILOPE. Le SPPT peine à reproduire de telles précipitations. Le  $Q_{max}$  atteint difficilement les 70mm mais ceux-ci sont situés trop au Nord de la zone d'intérêt (Landes au lieu de sud de la Gironde). Le  $Q_{90}$  suggère par ailleurs des précipitations décalées vers l'Est.

Tout comme le SPPT, le  $Q_{max}$  du B-CRPS-PP propose des cumuls de pluie jusqu'à 90mm maximum sur la Gironde, mais prévoit correctement 90mm sur les Landes. Son  $Q_{90}$  est par ailleurs mieux localisé sur les Landes que le SPPT.

Le RPP optimisé, quant à lui, propose des  $Q_{max}$  et  $Q_{90}$  sur une zone correctement ciblée sur les Landes et les maximums de précipitations atteignent les 90mm comme observés. Quelle que soit la méthode de représentation de l'erreur de modélisation, les maxima de précipitations sont sous-estimés sur les Pyrénées-Atlantiques mais les  $Q_{90}$  et  $Q_{max}$  des méthodes de perturbations de paramètres restent plus importants que ceux du SPPT. Les B-CRPS-PP et oRPP semblent donc mieux représenter cet événement intense et sont plus à même d'alerter quant à sa sévérité.

**Convection sur la Loire-Atlantique :** A l'inverse, pour les orages se développant sur la Loire-Atlantique en début de nuit, aucune des trois méthodes de représentation de l'erreur de modélisation ne suggère de forts cumuls de pluie comme ceux observés (90mm). Cependant, le SPPT ayant une plus grande dispersion spatiale des précipitations, il propose un  $Q_{90}$  et  $Q_{max}$  un peu plus élevé sur cette zone que les B-CRPS-PP et oRPP, mais pas suffisant pour alerter sur un risque potentiel de convection sévère.

**Bilan du cas de convections sévères :** Notons que d'autres cas de convections sévères ont été étudiés (8 autres cas sur l'été 2018) et en général, les mêmes conclusions ont été tirées. Le SPPT a tendance à produire des précipitations plus étalées dans l'espace alors que les B-CRPS-PP et oRPP proposent des plus forts cumuls de pluie. Les cumuls de précipitations étant plus importants pour les  $Q_{90}$  et

$Q_{max}$  pour ces deux méthodes que le SPPT, ceci suppose un décalage de la distribution de l'état de l'atmosphère avec ces méthodes, vers un état plus pluvieux que celui du SPPT. Ceci est confirmé par un biais de précipitations cumulées sur 24h bien plus négatif avec le SPPT (-0.25mm) qu'avec les méthodes de perturbations de paramètres (-0.07mm).

## 8.6.2 Cas de la tempête Gabriel

Après avoir étudié un cas de convections sévères pendant l'été 2018, nous nous intéressons maintenant à un cas de dépression des moyennes latitudes balayant la France pendant l'hiver 2018-2019, la tempête Gabriel.

### Situation météorologique des 29 et 30 Janvier 2019

Le 28 Janvier 2019, une perturbation se forme le long d'un front quasi-stationnaire situé sur l'Océan Atlantique et amène à la formation de la dépression Gabriel à 12hUTC ce même jour. Cette dernière se creuse rapidement jusqu'à moins de 985hPa et arrive à maturité le 29 à 00h UTC au large de la France (Figure 8.15a). Douze heures plus tard, lorsque son centre dépressionnaire est situé sur la pointe bretonne, son front chaud va alors commencer à balayer la France. A 18h UTC, le 29 Janvier, son front froid rattrape le front chaud, un front occlus balaye alors la France d'Ouest en Est et produit des chutes de neige allant jusqu'à 12cm à Lille ou encore 10 cm à Toussus-le-Noble dans les Yvelines. Cette tempête arrive ensuite en Allemagne à 12h UTC, le 30 Janvier. Derrière son front occlus, un ciel de traîne assez actif (zone bleu sur la Figure 8.15b) produit alors de nouveaux cumuls de pluie sur le Sud-Ouest de la France dans la nuit du 29 au 30 Janvier 2019.

La tempête a par ailleurs créé de fortes rafales de plus de 100km/h tout le long de la côte française occidentale et un maximum de 169km/h a été mesuré à Prat de Bouc dans le Cantal. Ainsi, plusieurs vigilances oranges ont été lancées (Figure 8.15c) : pour le vent le long de la côte Atlantique entre la Loire-Atlantique et les Pyrénées-Atlantiques, ainsi que pour neige et verglas sur le centre-nord de la France et quelques départements sur le Massif Central.

### Evaluation des différentes méthodes de représentation de l'erreur de modélisation

Pour ce cas de tempête, nous allons nous intéresser aux méthodes SPPT et B-CRPS-PP. Les cumuls de précipitations étant moins extrêmes que pour les cas de convection profonde estivale, les quantiles 75% et 90% de précipitations (pluie + neige) cumulées sur 48h (du 28 Janvier à 21hUTC au 30 Janvier 2019 à 21hUTC) sont comparés aux cumuls observés des produits ANTILOPE pour les deux méthodes. Encore une fois, les prévisions d'ensemble (SPPT et B-CRPS-PP) utilisent les mêmes conditions initiales, latérales et surfaciques du 28 Janvier 2019 à 21hUTC, pour faciliter la comparaison de la représentation de l'erreur de modélisation.

Quatre zones de précipitations sont relevées dans les cumuls de pluie et neige mesurés (Figure 8.16.a).

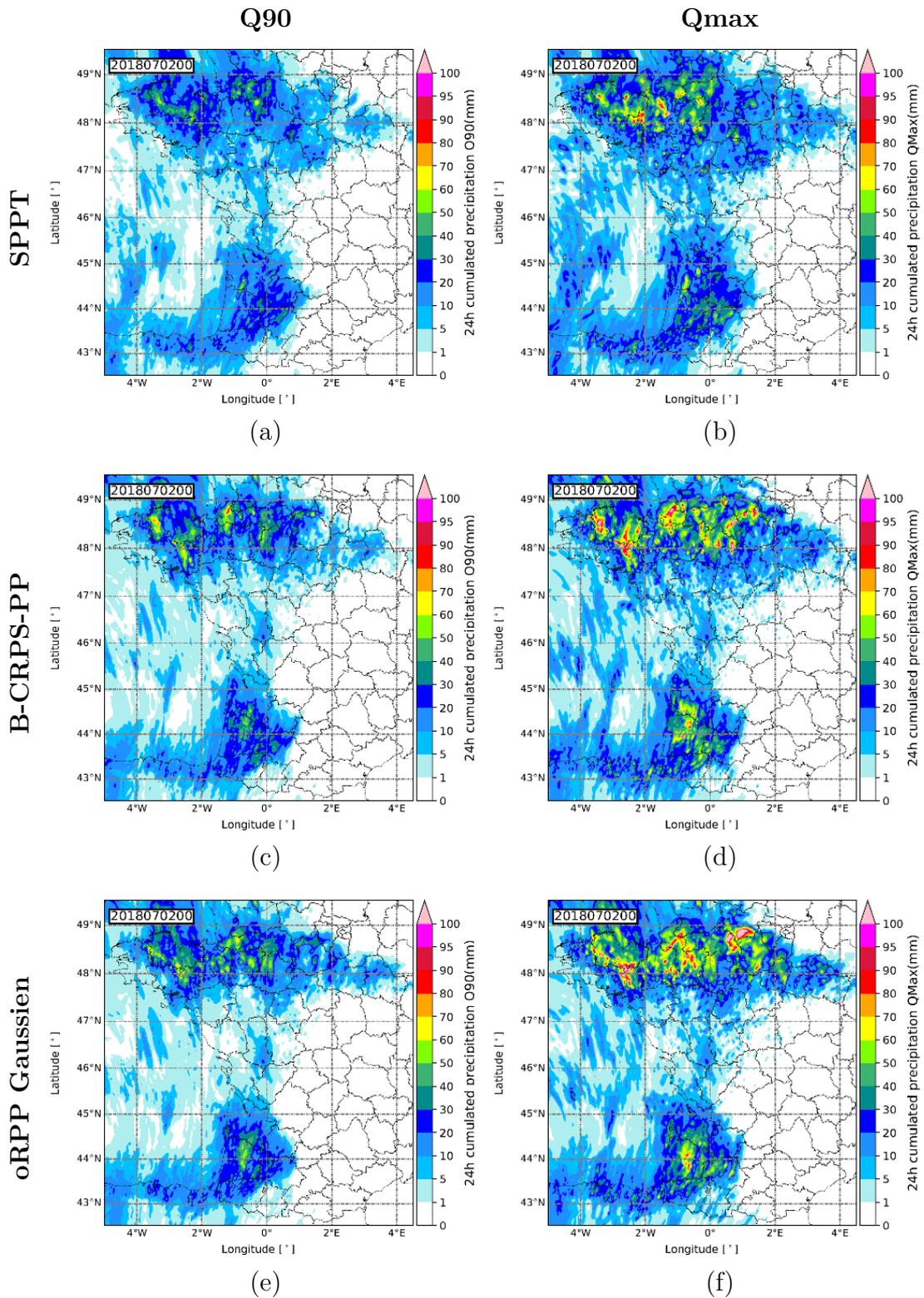
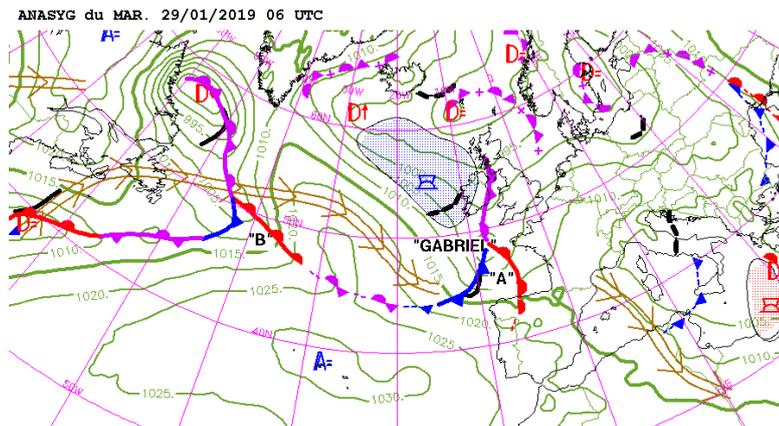
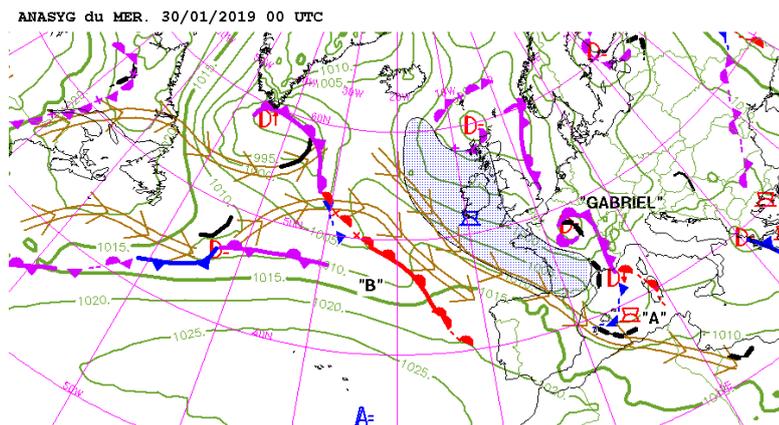


FIGURE 8.14 – Prévisions de la PEARO du 30 juin 2018 à 21hUTC du cumul de précipitation sur 24h prévu pour la journée du 01 Juillet 2018 (de 3 à 27h d'échéance) : Colonne de gauche Q90, Colonne de droite Qmax; en haut : SPPT, au milieu : B-CRPS-PP, en bas : oRPP utilisant des distributions gaussiennes.

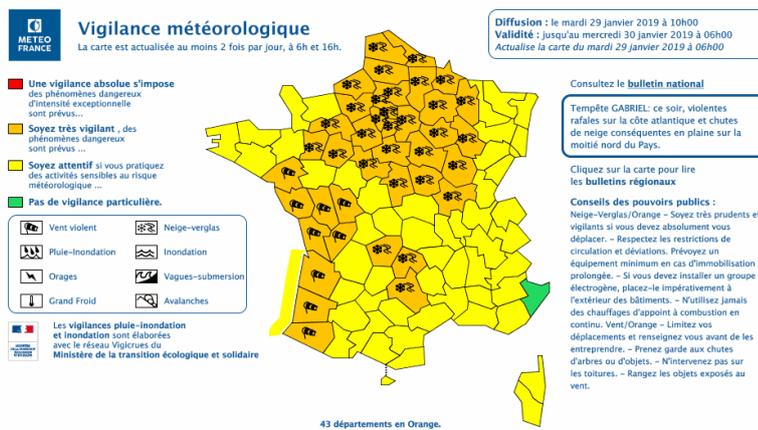
# PARTIE III. REPRÉSENTATION DES ERREURS DE MODÉLISATION DANS LA PEARO



(a) ANASYG du 29 Janvier 2019 à 6hUTC



(b) ANASYG du 30 Janvier 2019 à 0hUTC



(c) Bulletin de Vigilance à 10h le 29 Janvier 2019

FIGURE 8.15 – Cas d'étude des 29 et 30 Janvier 2019

La première, située sur le Nord de la Bretagne, comprend les plus forts cumuls allant jusqu'à 50mm au large des Côtes-d'Armor et en Ille-et-Vilaine. Celle-ci est prévue un peu plus au Sud, dans les terres pour toutes les prévisions d'ensemble (Q75 et Q90). Les Q75 des SPPT (Figure 8.16.b) et B-CRPS-PP (Figure 8.16.d) proposent tous deux plus de 50mm de précipitations, alors que le Q90 du B-CRPS-PP (Figure 8.16.e), contrairement à son homologue du SPPT (Figure 8.16.c), prévoit des cumuls plus importants, dépassant les 60mm en Côtes d'Armor. Ainsi, pour des Q75 proches, la méthode B-CRPS-PP a un Q90 plus élevé, suggérant une plus forte dispersion de l'ensemble avec cette approche sur cette zone. A l'inverse, dans la Sarthe, le Q90 du SPPT dépasse plus facilement les 30mm que celui du B-CRPS-PP.

Sur la région parisienne, les cumuls de précipitations observés sont en général autour de 15mm et dépassent localement 30mm dans les Yvelines. Les Q75 et Q90 du B-CRPS-PP représentent bien des cumuls de cet ordre de grandeur avec quelques zones à plus de 20mm, alors que ceux du SPPT proposent des cumuls trop élevés dépassant plus de 20mm sur l'ensemble de la région. Dans cette région, le SPPT semble produire, à tort, même pour de faibles quantiles, de trop fortes précipitations.

Sur les contreforts du Massif central, des cumuls allant jusqu'à plus de 30mm sont observés dans le Cantal. Les Q75 et Q90 des SPPT et B-CRPS-PP sont très proches et proposent tous deux des précipitations cumulées dépassant les 40mm très localement. Ceci suggère un biais positif de la distribution de l'état de l'atmosphère équivalent entre les deux méthodes sur cette région. Cette sur-estimation étant commune aux deux méthodes, elle n'est donc pas due à la représentation de l'erreur de modélisation par l'ajout de perturbation, mais aux paramétrisations même d'AROME (notons qu'elle apparaît aussi dans la version déterministe d'AROME, non montré).

Enfin, la dernière zone de précipitations concerne le Sud-Ouest de la France. Des cumuls allant jusqu'à plus de 30mm sont mesurés en Gironde ainsi que le long de la côte septentrionale espagnole et des Pyrénées. Que ce soit pour le Q75 ou le Q90, les deux méthodes proposent des cumuls dépassant les 70mm le long des Pyrénées, surestimant ainsi les chutes de neige dans ce massif montagneux. En revanche, dans les Landes et en Gironde, le long de la côte Atlantique, le B-CRPS-PP produit de forts cumuls de précipitations, même pour de faibles quantiles comme le Q75. Ceci est dû à une trop forte convection dans le ciel de traîne qui provoque des cumuls de précipitations trop importants lors de son arrivée sur les terres. A l'inverse, le Q75 du SPPT est beaucoup plus modéré et propose des cumuls de pluie de plus de 30mm comme observé. Le Q90 en revanche, tend à proposer une plus forte convection le long de la côte, tout comme le B-CRPS-PP mais tout en restant dans des cumuls du même ordre que ceux observés.

**Bilan de la tempête Gabriel :** Ainsi, le B-CRPS-PP semble mieux correspondre à la réalité que le SPPT pour ce qui est de l'épisode neigeux sur la région parisienne. En revanche, de nombreux membres semblent surestimer la convection en mer et produire de fortes précipitations lorsque celle-ci atteint la côte atlantique.

### 8.6.3 Bilan des études de cas

Pour conclure, ces deux études de cas permettent de montrer la tendance des méthodes de perturbations de paramètres à produire de la convection plus sévère, que ce soit en été ou en hiver. Elles proposent ainsi de plus forts cumuls de précipitations, alors que le SPPT semble moins bien cibler les zones de précipitations convectives. Pour des phénomènes de plus grande échelle comme les dépressions, en revanche, les deux méthodes restent assez similaires puisque moins de paramétrisations sont en jeu dans le calcul des précipitations stratiformes résolues explicitement.

## 8.7 Synthèse des résultats

Pour conclure ce chapitre, nous proposons une synthèse des résultats obtenus.

A partir des scores probabilistes calculés pour les différents ensembles testés et présentés au chapitre 7, nous avons pu constater le réel potentiel des méthodes de perturbations de paramètres à améliorer les scores de champs météorologiques près de la surface par rapport à la méthode actuellement employée dans la PEARO. Ce gain concerne la quasi-totalité des méthodes de perturbations de paramètres, mais est plus important pour les méthodes PP en particulier sur les rafales. Cette amélioration reste cependant plus modérée sur les précipitations ou encore avec la méthode RPP ajoutant une variabilité journalière à la perturbation de paramètres.

L'optimisation de la méthode PP réalisée sur le CRPS de vent et pluie a montré une amélioration importante de ce score mais pas uniquement, montrant ainsi la capacité du CRPS à prendre en compte les différentes qualités attendues d'une prévision d'ensemble. Elle a aussi montré une forte proximité entre les ensembles optimisés sur la période estivale ou ceux optimisés sur les périodes hivernales et automnales, correspondant à l'optimisation sur l'année. L'optimisation de la méthode RPP a montré qu'une unique distribution de probabilité optimisée pour chaque paramètre n'est pas suffisante. L'approche oRPP gaussienne, qui optimise les distributions pour chaque membre selon des jeux de paramètres spécifiques prenant en compte les interactions entre paramètres, semble plus pertinente. Par ailleurs, l'emploi d'émulateurs n'a pas montré de réelle efficacité sur l'optimisation des distributions de probabilités des paramètres.

La similarité entre la perturbation des 21 ou 8 paramètres identifiés comme influents par l'analyse de sensibilité de Morris, a montré que la majeure partie de l'erreur de modélisation représentée par la perturbation de paramètres peut être limitée à la perturbation de seulement 8 paramètres. Ce résultat peut permettre de limiter le coût de maintenance de cette méthode.

En outre, l'évaluation dans une configuration opérationnelle, incluant la représentation des autres sources d'incertitude (conditions initiales, surfaciques et latérales), a confirmé les gains substantiels obtenus avec les méthodes de perturbations des paramètres. Cependant, la combinaison des méthodes PP et SPPT n'a pas montré de réelle plus value sur les scores de surface, comparée à l'utilisation de la méthode PP seule.

Une analyse du biais des membres du B-CRPS-PP a mis en évidence des biais systématiques pour chacun des membres de cet ensemble. Ceci peut à la fois

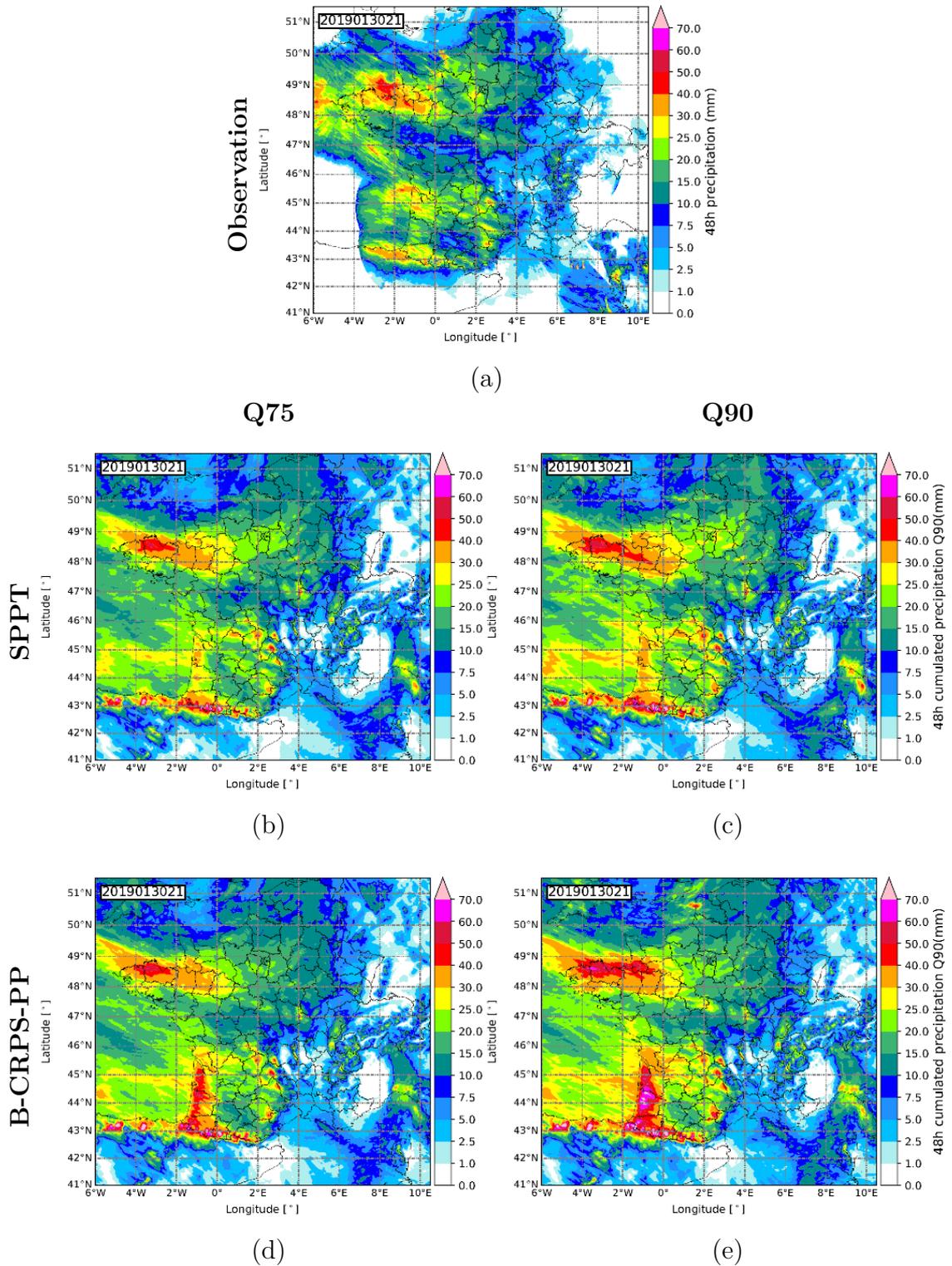


FIGURE 8.16 – Cumul de précipitation sur 48h entre le 28 Janvier 2019, 21h UTC et le 30 Janvier 2019, 21h UTC : (a) observation ANTILOPE, (b) Q75 pour le SPPT, (c) Q90 pour le SPPT, (d) Q75 pour le B-CRPS-PP et (e) Q90 pour le B-CRPS-PP

permettre de corriger des biais du modèle AROME (ex : rafales) ou encore de cacher des biais compensatoires au sein de la PEARO (ex : pluie).

Une étude réalisée sur deux étés différents a montré que l'optimisation réalisée sur l'année 2018 est toujours valide sur une autre période (été 2020). Cependant, un gain différent en dispersion du oRPP Gaussien au SPPT, a été constaté entre les deux étés. Le oRPP Gaussien présente une dispersion plus homogène selon et pendant les saisons, mais aussi plus forte que la méthode SPPT. Ceci suppose alors une plus forte dépendance du SPPT à la situation météorologique, celui-ci étant plus actif lors d'étés convectifs.

Enfin, des études de cas ont été analysées. Pour un évènement de convection profonde sévère, les B-CRPS-PP et oRPP Gaussien proposent de plus forts cumuls de précipitations que la méthode SPPT en raison d'un décalage des distributions de probabilité des états de l'atmosphère entre ces méthodes. Le SPPT propose en revanche des cumuls de pluie plus étalés dans l'espace, ciblant moins bien les zones à risque. Par ailleurs, l'étude de la tempête Gabriel a montré de meilleurs résultats sur l'épisode neigeux en Île-de-France avec la méthode B-CRPS-PP que le SPPT. Cependant, celle-ci semble surestimer la convection ayant lieu dans le ciel de traîne de la dépression.

## Bibliographie

O. Laurantin. Antilope : hourly rainfall analysis merging radar and rain gauges data. In *Proceedings of Weather Radar and Hydrology Conference 2008*, Grenoble, France, 2008.

# Conclusion

Les systèmes de prévision d'ensemble permettent une approche de la prévision météorologique sous forme probabiliste. Au contraire de l'approche déterministe classique, ils visent à évaluer l'incertitude inhérente à toute prévision. Dans cette optique, l'ensemble des incertitudes doit être représenté, et plus particulièrement une des plus importantes, à savoir celle liée à la représentation des processus physiques sous-maille. A ce jour, cette erreur de modélisation est représentée dans la PEARO par la perturbation des tendances temporelles calculées par ces paramétrisations physiques, approche dite SPPT. L'objectif de ce travail était de mettre en place et évaluer d'autres méthodes, plus physiques, reposant sur la perturbation des paramètres incertains des schémas de paramétrisations.

Dans un premier chapitre, nous avons présenté, dans les grandes lignes, la structure des modèles de prévision atmosphériques afin de donner un aperçu des sources d'incertitude présentes. En particulier, une illustration des incertitudes des modèles atmosphériques sur le cas d'une tempête a été étudiée. Nous avons vu que le choix de paramétrisations pour représenter la convection ne se limite pas uniquement à une incertitude des phénomènes orageux mais peut influencer la structure des vents d'altitude et impacter, par là même, la prévisibilité sur une région aussi grande que l'Europe. Par ailleurs, nous avons aussi illustré l'incertitude de paramètres avec un exemple sur la vitesse de chute de la neige dans les paramétrisations de la microphysique nuageuse. Ainsi, nous avons montré la nécessité d'utiliser des méthodes de prévisions d'ensemble afin de rendre compte de ces incertitudes.

Nous avons ensuite donné un aperçu des différentes méthodes pour représenter les sources d'incertitudes dans les prévisions d'ensemble, que ce soit celles dues aux conditions initiales, de surface, latérales dans le cas des modèles à aire limitée ou encore celles dues aux erreurs de modélisation. Les différents scores probabilistes permettant l'évaluation de ces prévisions d'ensemble ont été aussi abordés.

Puis, nous nous sommes intéressés à la prévision d'Ensemble AROME, à sa construction et en particulier à sa représentation de l'erreur de modélisation. Nous avons plus particulièrement étudié les limites de la méthode SPPT et l'intérêt d'étudier les méthodes basées sur la perturbation de paramètres des schémas physiques.

Dans la partie II, nous avons identifié 21 paramètres dont les valeurs sont incertaines, et qui sont ainsi susceptibles de contribuer de manière importante à l'erreur de modélisation. Ceux-ci sont issus de six paramétrisations physiques et dynamiques différentes. Cependant, tous ces paramètres n'engendrent pas nécessairement une incertitude des prévisions proportionnelles à leur incertitude même. Afin d'éviter des perturbations inutiles de paramètres, engendrant des coûts de

calcul plus ou moins conséquents ou encore pour alléger la maintenance des codes informatiques de la PEARO, il faut alors centrer la représentation de l'erreur de modélisation sur les paramètres les plus à même de représenter l'incertitude. Pour les identifier, nous avons réalisé différentes analyses de sensibilité utilisant les méthodes de Morris et de Sobol'. Celles-ci ont permis d'identifier 8 paramètres plus influents et d'étudier selon différents points de vue spatio-temporels, leurs effets sur les prévisions du modèle AROME. Il a été constaté un véritable potentiel de l'ensemble des paramètres à représenter l'incertitude de modélisation.

Dans la partie III, nous nous sommes attachés aux méthodes de perturbations de paramètres. A partir des prévisions réalisées pour les analyses de sensibilité, une première méthode utilisant des valeurs différentes de paramètres pour chacun des membres de la prévision d'ensemble, mais constantes d'un jour à l'autre (méthode dite PP), a été mise en place. Une optimisation selon le score CRPS a été réalisée et a montré un réel potentiel pour cette approche. Cette optimisation ne se limite pas uniquement à ce score probabiliste, la plupart des scores ont vu un net progrès avec la méthode de perturbations de paramètres par rapport à la méthode SPPT.

Par ailleurs, une autre méthode de perturbation des paramètres (dite RPP) a été mise en place : elle comprend une variabilité temporelle pour les perturbations des paramètres, changeant ainsi de valeurs à chaque initialisation. Différentes distributions de probabilité des paramètres ont été testées afin d'optimiser, toujours selon le CRPS, les valeurs prises par les paramètres. La méthode la plus performante a été celle utilisant des distributions gaussiennes autour de l'optimum obtenu avec la méthode PP. L'optimisation de la méthode RPP tendant vers celle du PP, ceci montre alors que la perturbation de paramètres pourrait se limiter uniquement à une perturbation fixe pour toutes dates.

De plus, la perturbation des 8 paramètres identifiés par les analyses de sensibilité, proposant des résultats similaires à la version avec les 21 paramètres, démontre alors la possibilité de représenter l'erreur de modélisation avec seulement quelques paramètres. Cela souligne l'intérêt de telles analyses de sensibilité pour limiter le coût de maintenance des approches par perturbations de paramètres.

Des études des caractéristiques des prévisions d'ensemble ont montré, en revanche, certaines limites des méthodes utilisant la perturbation de paramètres. Celles-ci peuvent présenter des biais uniques et systématiques pour chaque membre. De ce fait, les membres ne peuvent être équiprobables puisque certains ont de meilleures performances que d'autres.

Une étude de la dispersion a également montré une plus forte sensibilité de la méthode SPPT aux situations météorologiques, ce qui peut aussi refléter en partie une sous-estimation des incertitudes par cette méthode SPPT, lorsque les tendances physiques sont faibles. Enfin, des études de cas ont montré des cumuls de précipitations convectives plus importants avec les méthodes de perturbations de paramètres, par opposition à la méthode SPPT qui cible moins bien les zones à risque.

Dans le présent travail, nous avons testé uniquement la variabilité des paramètres perturbés selon les différentes dates initiales des prévisions lancées. Un prolongement naturel de ce travail serait de considérer des variations temporelles des perturbations au cours de la prévision, tel que réalisé dans la prévision d'ensemble régionale du Royaume-Uni. Des variations spatio-temporelles pourraient aussi être envisagées, comme dans la méthode dite SPP.

Les cartes de sensibilité obtenues à partir de l'analyse de sensibilité de Morris ainsi que les résultats selon les échéances seraient en capacité de fournir des longueurs et temps caractéristiques pour ces perturbations stochastiques. Par exemple, la sensibilité des paramètres présente des variations sur une échelle de 6h, un tel temps caractéristique pourrait alors être utilisé. Les cartes spatiales seraient par ailleurs à même de servir de base pour évaluer les longueurs de corrélations spatiales à utiliser.

L'étude du biais a permis d'établir des valeurs de paramètres qu'il serait possible de privilégier pour limiter le biais individuel des membres. En étudiant le biais d'un grand nombre de prévisions, des distributions de paramètres obtenus à partir de prévisions non biaisées pourraient aussi être établis et utilisés afin de corriger les biais des méthodes de perturbation de paramètres.

En outre, notons que la prochaine version opérationnelle de l'assimilation d'ensemble AROME reposera elle aussi sur la méthode SPPT. Il serait donc ultérieurement pertinent de compléter ou remplacer cela par une méthode de perturbation de paramètres.

D'un point de vue plus pratique, la PEARO possédant maintenant 16 membres et des résolutions spatiales et temporelles plus fines, une optimisation des méthodes de perturbation de paramètres dans ce cadre pourrait alors être réalisée afin de s'adapter à cette nouvelle configuration de la PEARO.

Enfin, l'intégration de la stochasticité directement dans les paramétrisations des processus physiques est une possibilité de représentation de l'erreur de modélisation. Celle-ci est plus à même de représenter cette erreur de manière physique, mais elle reste pour l'instant à l'état de recherche<sup>1</sup>.

---

1. par ex : Kober, K., et Craig, G. C. *Physically Based Stochastic Perturbations (PSP) in the Boundary Layer to Represent Uncertainty in Convective Initiation*, Journal of the Atmospheric Sciences, 73(7), 2893-2911, 2016.

## CONCLUSION

---

# **Annexes**



# Annexe A

## Impact des schémas de convection profonde sur la dynamique de la dépression Stalactite

Les articles suivants ont été soumis au *Weather and Climate Dynamics*. Le premier, publié et nommé « *The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during NAWDEX IOP6* » concerne l'impact des schémas de convection profonde sur la double structure du Jet Stream en sortie de la bande transporteuse d'air chaud de la dépression Stalactite. Tandis que le second article, intitulé « *Diabatic processes modulating the vertical structure of the jet stream above the cold front of an extratropical cyclone : sensitivity to deep convection schemes* », s'intéresse plutôt à l'impact sur la structure verticale du Jet Stream dans la zone d'ascendance de la bande transporteuse d'air chaud.



# The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during NAWDEX IOP6

Gwendal Rivière<sup>1</sup>, Meryl Wimmer<sup>2</sup>, Philippe Arbogast<sup>3</sup>, Jean-Marcel Piriou<sup>2</sup>, Julien Delanoë<sup>4</sup>, Carole Labadie<sup>2</sup>, Quitterie Cazenave<sup>4</sup>, and Jacques Pelon<sup>4</sup>

<sup>1</sup>Laboratoire de Météorologie Dynamique/IPSL, École Normale Supérieure, PSL Research University, Sorbonne University, École Polytechnique, IP Paris, CNRS, Paris, France

<sup>2</sup>CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

<sup>3</sup>Direction des Opérations pour la prévision, Météo-France, Toulouse, France

<sup>4</sup>LATMOS-IPSL, CNRS/INSU, University of Versailles, Guyancourt, France

**Correspondence:** Gwendal Rivière (griviere@lmd.ens.fr)

Received: 11 June 2021 – Discussion started: 15 June 2021

Revised: 19 August 2021 – Accepted: 29 September 2021 – Published: 29 October 2021

**Abstract.** The effect of parameterized deep convection on warm conveyor belt (WCB) activity and the jet stream is investigated by performing simulations of an explosively developing large-scale cyclone that occurred during the North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX) field campaign using the Météo-France global atmospheric model ARPEGE. Three simulations differing only from their deep convection representation are analysed. The first one was performed with the Bougeault (1985) scheme (B85), the second one with the Prognostic Condensates Microphysics and Transport (PCMT) scheme of Piriou et al. (2007), and the third one without any parameterized deep convection. In the latter simulation, the release of convective instability at the resolved scales of the model generates localized cells marked by strong heating with few degrees extent in longitude and latitude along the fronts. In runs with active parameterized deep convection (B85, PCMT), the heating rate is more homogeneously distributed along fronts as the instability release happens at subgrid scales. This difference leads to more rapid and abrupt ascents in the WCB without parameterized deep convection and more moderate but more sustained ascents with parameterized deep convection. While the number of WCB trajectories does not differ much between the three simulations, the averaged heating rates over the WCB trajectories exhibits distinct behaviour. After 1 d of simulations, the upper-level heating rate is on av-

erage larger, with the B85 scheme leading to stronger potential vorticity (PV) destruction. The difference comes from the resolved sensible and latent heating and not the parameterized one. A comparison with (re)analyses and a large variety of airborne observations from the NAWDEX field campaign (Doppler radar, Doppler lidar, dropsondes) made during the coordinated flights of two aircraft in the WCB outflow region shows that B85 performs better in the representation of the double jet structure at 1 d lead time than the other two simulations. That can be attributed to the more active WCB at upper levels. However, this effect is too strong and that simulation becomes less realistic than the other ones at forecast ranges beyond 1.5 d. The simulation with the PCMT scheme has an intermediate behaviour between the one with the B85 scheme and without parameterized deep convection, but its impact on the jet stream is closer to the latter one. Finally, additional numerical experiments show that main differences in the impact on the jet between PCMT and B85 largely come from the chosen closure, with the former being based on CAPE and the latter on moisture convergence.

## 1 Introduction

Despite significant and continuous improvements of numerical weather forecasts during the last decades (Simmons and Hollingsworth, 2002; Bauer et al., 2015), mid-latitude weather forecast errors still occur at synoptic and planetary scales. For instance, at short-range (days 1–3) weather forecasts still struggle to accurately predict the intensity, timing, and location of extratropical storms (Korfe and Colle, 2018) and in particular their wind gusts footprints (Hewson et al., 2014). At medium range (days 4–8), forecast busts may occur in the prediction of large-scale atmospheric circulation patterns like blockings (Rodwell et al., 2013; Grams et al., 2018; Maddison et al., 2020). Forecast errors are usually characterized by an upscale and downstream error growth (Zhang et al., 2007; Selz and Craig, 2015), which start at the convective scale and mesoscale (Rodwell et al., 2013). They then often propagate downstream and amplify along upper-tropospheric Rossby wave trains (Gray et al., 2014; Lillo and Parsons, 2017) where nonlinear effects are strong (Baumgart et al., 2018). They finally form large-scale forecast error patterns further downstream, typically at the end of storm tracks like over western Europe (Grams et al., 2018; Maddison et al., 2020).

These forecast errors originate from two main sources of uncertainties: uncertainties in determining the initial state of the atmosphere and uncertainties due to the imperfections of numerical weather prediction (NWP) models (Tribbia and Baumhefner, 1988). But it is somewhat difficult to determine the origins of forecast busts among these two sources as they intertwine and can compensate for or reinforce each other (Rodwell et al., 2013). Among the model deficiencies, the representation of subgrid-scale physical processes (e.g. cloud microphysics, convection, radiation, turbulence) by different parametrization schemes is a key question and such schemes are constantly upgraded.

The question of the impact of diabatic processes on the predictability of the atmospheric flow was central to the international NAWDEX (North Atlantic Waveguide Downstream and Impact Experiment) project and its field campaign that occurred in September–October 2016 (Schäfler et al., 2018). In particular, one main hypothesis of the project is that the potentially misrepresented diabatic processes embedded in the so-called warm conveyor belts (WCBs) play a key role in triggering forecast errors along the North Atlantic upper-tropospheric waveguide. They may significantly deteriorate the forecast of high-impact weather events more downstream. WCBs correspond to air masses flowing within the boundary layer in the warm sector of extratropical cyclones and ascending to the tropopause level east of the cyclone centre (Browning, 1990; Wernli and Davies, 1997). These coherent air masses make the connections between the different layers of the troposphere as they gain moisture in the boundary layer, undergo heat release by condensation when they ascend and have an impact on the upper-

tropospheric circulation. When they reach the tropopause level, negative potential vorticity (PV) is generated in the WCB outflow region that reinforces the ridge downstream of the cyclone but also the upper-tropospheric PV gradient and hence the jet stream (Pomroy and Thorpe, 2000; Grams et al., 2011; Schemm et al., 2013; Chagnon et al., 2013; Gray et al., 2014). Uncertainties in the heat release within WCBs due to misrepresented cloud microphysics or convection could trigger forecast errors at the tropopause level that may then propagate downstream along the waveguide (Martinez-Alvarado et al., 2016; Grams et al., 2018; Maddison et al., 2020). The objective of the present study is to investigate the effect of parameterized deep convection in a global NWP model on the WCB activity and the jet stream.

Embedded convective activity within WCBs was initially identified using airborne and satellite-derived data within an extratropical cyclone during the Experiment on Rapidly Intensifying Cyclones over the Atlantic (ERICA) field experiment (Neiman et al., 1993). Various ground-based, spaceborne and airborne remote sensing measurements confirm the regular presence of convective activity in the warm sector of extratropical cyclones and WCBs in particular (Crespo and Posselt, 2016; Flaounas et al., 2016, 2018; Oertel et al., 2019; Blanchard et al., 2020; Jeyaratnam et al., 2020). This was detected in various regions of the Northern Hemisphere: over the United States (Jeyaratnam et al., 2020), the North Atlantic ocean (Crespo and Posselt, 2016; Oertel et al., 2019, 2020, 2021; Blanchard et al., 2020, 2021) and in the Mediterranean sector (Flaounas et al., 2016, 2018). Recent numerical studies relying on convection-permitting model simulations of various extratropical cyclones also highlighted the presence of embedded convection within WCBs (Rasp et al., 2016; Oertel et al., 2019, 2020, 2021; Blanchard et al., 2020, 2021). The emerging picture is that WCBs are composed of various ascending air streams in which convective ascents may intermittently happen within the main slantwise ascending airflow. Such an intermittent convective activity may have non-negligible effects on the upper-level jet stream. Oertel et al. (2020) and Blanchard et al. (2021) showed that the WCB-embedded convective cells form dipolar PV anomalies with the negative pole being closer to the jet which tend to reinforce the PV gradient and hence the jet.

Because WCB-embedded convection plays an important role in the large-scale circulation, it is worth investigating the effect of parameterized deep convection on WCBs in models that do not have enough resolution to explicitly resolve that phenomenon. The contribution of parameterized convection in the total heating and precipitation rates within extratropical cyclones varies a lot by changing schemes as shown by Martinez-Alvarado et al. (2014b) and Booth et al. (2018). Differences in total precipitation rates between models usually vary much less than differences in the contribution of parameterized convection to the total. Done et al. (2006) showed that the upper-level large-scale circu-

lation is sensitive to the relative amount of parameterized vs. explicit convection, but their focus was on the effect of mesoscale convective systems and not WCBs of deep cyclones. More relevant to the present study, the impact of convection parametrization on WCB dynamics and its effect on the upper-level flow was studied by Martinez-Alvarado and Plant (2014) in a case study of a North Atlantic extratropical cyclone. The authors show that parameterized convection “regulates the action of large-scale heating” by releasing convective instability at subgrid scales which would have otherwise been released at larger and resolved scales. Because of this regulation, WCB trajectories more gradually ascend when parameterized convection is more active than when it is reduced. Furthermore, in the presence of reduced parameterized convection, the more abrupt ascents lead to more important impact on upper-level circulation even though differences with active parameterized convection are small at 24 h lead time (Martinez-Alvarado et al., 2014a). In the same vein, the present study aims to address the following questions.

- How different are WCBs between simulations with active parameterized deep convection and those without?
- What is the impact of parameterized deep convection on the jet stream at short range (less than 2 d)?
- How different are WCBs and their impact on the jet stream between simulations performed with distinct deep convection schemes?

Our approach is based on the case study of an extratropical cyclone that occurred during NAWDEX called the Stalactite cyclone (29 September–3 October 2016) and corresponding to Intensive Observing Period (IOP) 6. It was an explosively deepening cyclone at a rate of roughly 24 hPa in 24 h (Flack et al., 2021) that formed off the east coast of Newfoundland and travelled over the North Atlantic toward Iceland and Greenland. This large-scale cyclone participated in the formation of a block over Scandinavia (Schäfler et al., 2018). The Stalactite cyclone was already the subject of several studies. Maddison et al. (2019, 2020) showed that the low predictability of the block onset at medium range was associated with uncertainties in the timing, location and intensity of that cyclone. Blanchard et al. (2020) and Blanchard et al. (2021) studied the WCB-embedded convection by performing convection-permitting model simulations, while Mazoyer et al. (2021) analysed the impact of cloud microphysics on the WCB activity of the Stalactite cyclone with similar convection-permitting simulations. Flack et al. (2021) analysed the whole life cycle of the cyclone in simulations of two climate models. Maddison et al. (2020) showed that medium-range forecast skills of the cyclone are sensitive to the choice of the deep convection schemes, but none of these studies systematically investigated the role of parameterized deep convection on the WCB of the cyclone at short range. In addition to its dynamical interest, the Sta-

lactite cyclone is worth studying because numerical simulations can be compared and validated against airborne observations: two flights of the French Falcon 20 from the Service des Avions Français Instrumentés pour la Recherche en Environnement (SAFIRE) and one flight from the Deutsches Zentrum für Luft- und Raumfahrt (DLR) Dassault Falcon were conducted on 2 October 2016.

The paper is structured as follows. Section 2 is dedicated to the method where the global atmospheric model ARPEGE (Action de Recherche Petite Echelle Grande Echelle) and the three numerical simulations are presented. Two simulations are run with two distinct deep convection schemes used and developed within ARPEGE. The first one is used in the deterministic operational forecasts of the NWP version of the model and was developed by Bougeault (1985), while the second, called the Prognostic Condensates Microphysics and Transport scheme and developed by Piriou et al. (2007), was implemented in the CMIP6 version of the ARPEGE climate model (Roehrig et al., 2020). A third simulation is performed in which the deep convection scheme is turned off. The Lagrangian trajectory algorithm, PV diagnostics, and the re-analysis and airborne observation datasets are also summarized in Sect. 2. In Sect. 3, the impact of deep convection schemes on heating rates and WCB trajectories is analysed, while Sect. 4 is focused on their effect on the upper-level circulation and the jet stream. Finally, Sect. 5 provides concluding remarks.

## 2 Data and method

### 2.1 Model and simulations set-up

The Météo-France global forecast model ARPEGE (Courtier et al., 1991) is used at a T798 resolution with full physics and 90 vertical levels. The stretched horizontal grid (mapping factor of 2.4) corresponds to about 10 km horizontal resolution over France, 15 km near Iceland and 60 km on the opposite side of the globe relative to France. The temporal resolution is 7.5 min. Simulations analysed in the present study are based on the 2016 operational version of the ensemble prediction system (EPS) of ARPEGE called the Prévision d'Ensemble ARPEGE (PEARP; Descamps et al., 2015). It is documented in Ponzano et al. (2020) (see their Table 1) and was already used to study the predictability of a heavy precipitation event that occurred over France during NAWDEX (Binder et al., 2021). It includes 10 members that have the same initial state, which is the ARPEGE operational analysis made with a 4D-Var data assimilation scheme. They only differ in the representation of at least one of the following physical packages: deep convection, turbulence, shallow convection and surface oceanic fluxes. Two largely distinct deep convection schemes are used: one is the scheme developed by Bougeault (1985), hereafter called B85, and the other one is the Prognostic Condensates Microphysics and Trans-

port scheme of Piriou et al. (2007), hereafter called PCMT. For some members, modified versions of B85 are activated. Two turbulent schemes are used: the turbulent kinetic energy scheme of Cuxart et al. (2000) and the turbulence scheme of Louis (1979). Four shallow convection schemes are considered: the mass flux scheme of Kain and Fritsch (1993) and Bechtold et al. (2001), PCMT, the eddy diffusivity and Kain–Fritsch scheme, and the Pergaud et al. (2009) scheme. Surface oceanic fluxes are represented by the Belamari (2005) scheme or by an alternate version in which the evaporative fluxes are enhanced. The paper is focused on two particular members of this EPS, corresponding to the REF member and seventh member respectively as it appears in Table 1 of Ponzano et al. (2020). They only differ in the representation of deep convection, one using the B85 scheme and the other the PCMT scheme. The two simulations share the same physical parameterizations for turbulence (turbulent kinetic energy scheme; Cuxart et al., 2000), shallow convection (Kain and Fritsch, 1993; Bechtold et al., 2001), large-scale microphysics (Lopez, 2002) and oceanic flux (Exchange Coefficients from Unified Multi-campaigns Estimates; Belamari, 2005). Finally, a third simulation is performed without any active deep convection scheme (hereafter called NoConv) but with the same other physical packages as the other two simulations and will serve as a reference to assess the impact of B85 and PCMT schemes.

The three simulations are hereafter systematically compared. The starting time and date are 12:00 UTC on 1 October 2016 when the surface cyclone is already in its mature stage and located in the middle of the North Atlantic (see the position of the minimum sea level pressure at 1 d lead time in Fig. 1).

### 2.1.1 The Bougeault (1985) deep convection scheme

This mass flux scheme is triggered when the resolved plus subgrid-scale moisture convergence is positive in the low levels and the atmospheric profile is unstable. So the scheme is closed with moisture convergence. Following Kuo (1965) the total moisture convergence is either detrained in the convective environment or precipitated. This scheme was further developed by Ducrocq and Bougeault (1995) for downdraughts. It is part of the global operational NWP (numerical weather prediction) system at Météo-France and is currently used to perform ARPEGE deterministic operational simulations.

### 2.1.2 The Prognostic Condensates Microphysics and Transport scheme of Piriou et al. (2007)

This convection scheme separates microphysics and transport in grid-scale equations to overcome stationary cloud budget assumptions, as proposed by Piriou et al. (2007). Liquid and ice cloud condensates, as well as rain and snow, have prognostic mixing ratios to deal with the same level of so-

phistication inside convective updraught as in the resolved-scale microphysics (Lopez, 2002; Bouteloup et al., 2011), therefore including autoconversion, aggregation, collection, riming, melting, etc. The closure of all experiments run with PCMT in the present study is based on CAPE (convective available potential energy). As previously said, this scheme is used in PEARP but also is part of the CNRM Earth System Model for CMIP6 (Roehrig et al., 2020).

## 2.2 Model output and diagnostics

The output datasets of the simulations are provided on a regular longitude–latitude grid of  $0.5^\circ$ , a pressure grid spacing of 50 hPa in the vertical and a frequency of 15 min.

### 2.2.1 WCB trajectories

The Lagrangian trajectory code is designed to work with latitude  $\times$  longitude  $\times$  pressure files of zonal wind  $u$ , meridional wind  $v$ , vertical velocity  $\omega$ , and other variables such as temperature and diabatic tendencies. The algorithm is based on a prediction-correction method of the advection at the midpoint of the trajectory. Let  $DT$  be the time interval between two model outputs. To account for curvature effects, the trajectory model has higher resolution than the model outputs, and its time step  $dt$  is such that  $n \times dt = DT$ . At  $t = i \times DT + j \times dt$  (where  $j = 1, \dots, n-1$ ) and for each point  $(x(t), y(t), p(t))$  ( $p$  denotes pressure,  $x$  and  $y$  are horizontal coordinates) belonging to a trajectory, we look for the previous position  $(x(t-dt), y(t-dt), p(t-dt))$ , with the advection being made with  $u$ ,  $v$  and  $\omega$  in the middle of the trajectory portion at  $t-dt/2$ , which is not known a priori. We first apply a time interpolation to compute the 3D wind field at time  $t$  using the two closest model outputs for  $(u, v, \omega)$  at time  $i \times DT$  and  $(i+1) \times DT$  and referred to as  $\mathbf{U}(x, y, p, t)$ . We then apply an iterative method starting with the wind at the point of the trajectory at time  $t$  (i.e.  $\mathbf{U}(x(t), y(t), p(t))$ ) to build up a first estimation of the previous position of the trajectory  $(x_1(t-dt), y_1(t-dt), p_1(t-dt))$  and extract the wind along this first estimated trajectory at  $t-dt/2$  (called  $\mathbf{U}_1(x(t-dt/2), y(t-dt/2), p(t-dt/2))$ ) by horizontal, vertical and time interpolation. Horizontal interpolation is bilinear (four neighbouring points are used). A second estimated trajectory is calculated using  $\mathbf{U}_1$ , leading to a second estimation of the previous position  $(x_2(t-dt), y_2(t-dt), p_2(t-dt))$ . The process can be repeated several times, but practical tests show it converges after about two iterations. If a trajectory goes beyond 975 hPa, its position is shifted to 975 hPa. Note that such an algorithm allows the computation of both backward and forward trajectories. In what follows,  $DT$  and  $dt$  are equal to 15 and 7.5 min respectively ( $n = 2$ ).

Forward trajectories are initialized at 12:00 UTC on 1 October in the warm sector of the extratropical cyclone and computed during 48 h. To select WCB trajectories, a criterion of ascent exceeding 300 hPa within 1 d during the period be-

tween 12:00 UTC on 1 October and 12:00 UTC on 3 October is applied. This is a less strict criterion than the more usual criterion of 600 hPa ascent within 2 d (Madonna et al., 2014; Binder et al., 2016) but allows a broad view of the different ascents in the WCB.

### 2.2.2 Heating and PV tendencies

The Ertel potential vorticity  $q$  and its Lagrangian derivative can be expressed as (Ertel, 1942)

$$q = \frac{1}{\rho} \boldsymbol{\zeta}_a \cdot \nabla \theta, \quad (1)$$

$$\frac{Dq}{Dt} = \frac{1}{\rho} (\boldsymbol{\zeta}_a \cdot \nabla \dot{\theta} + \nabla \times \mathbf{F} \cdot \nabla \theta), \quad (2)$$

where  $D/Dt = \partial/\partial t + u\partial/\partial x + v\partial/\partial y + \omega\partial/\partial p$  is the Lagrangian derivative,  $\theta$  is the potential temperature,  $\rho$  is density,  $\boldsymbol{\zeta}_a$  is the three-dimensional absolute vorticity vector,  $\nabla$  is the three-dimensional gradient and  $\mathbf{F}$  is the three-dimensional frictional acceleration. The diabatic heating is denoted as  $\dot{\theta} = D\theta/Dt$ . These two equations can be rewritten under the hydrostatic balance in pressure coordinates as

$$q = -g \left[ (\zeta + f) \frac{\partial \theta}{\partial p} + \frac{\partial u}{\partial p} \frac{\partial \theta}{\partial y} - \frac{\partial v}{\partial p} \frac{\partial \theta}{\partial x} \right], \quad (3)$$

$$\begin{aligned} \frac{Dq}{Dt} = -g \left[ (\zeta + f) \frac{\partial \dot{\theta}}{\partial p} + \frac{\partial u}{\partial p} \frac{\partial \dot{\theta}}{\partial y} - \frac{\partial v}{\partial p} \frac{\partial \dot{\theta}}{\partial x} \right. \\ \left. + \left( \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \frac{\partial \theta}{\partial p} + \frac{\partial F_x}{\partial p} \frac{\partial \theta}{\partial y} - \frac{\partial F_y}{\partial p} \frac{\partial \theta}{\partial x} \right], \quad (4) \end{aligned}$$

where  $F_x$  and  $F_y$  are the zonal and meridional components of the frictional acceleration. Even though the friction-induced PV tendency is not negligible and generally generates a positive PV anomaly in lower troposphere (Stoelinga, 1996), it is not shown in the present study which is focused on the upper-level PV modification. PV modification along WCB has been generally attributed to the term involving the vertical gradient of the diabatic heating (Wernli and Davies, 1997; Joos and Wernli, 2012) – that is, the first term on the right-hand side of Eq. (4). The second and third terms involving the horizontal gradient of the heating have been generally considered to be second-order effects. Very recently the horizontal gradient has been shown to be very important in high-resolution convection-resolving simulations (Oertel et al., 2020).

The first method to compute the diabatic heating  $\dot{\theta}$  consists of approximating the time and space derivatives of potential temperature using centred finite-difference schemes applied to the potential temperature over the  $0.5^\circ \times 0.5^\circ$  horizontal grid and 50 hPa vertical grid and with a frequency of 15 min. An alternative to compute the total heating is to sum all the temperature tendencies due to each physical process (resolved and parameterized sensible and latent heating, radiation, turbulence) that are made available as model outputs on a non-regular grid. Spatial interpolations were necessary

to get them on the same regular  $0.5^\circ \times 0.5^\circ$  grid of the model outputs. The total heating obtained with this second method was found to be less accurate to represent the change in potential temperature undergone by the Lagrangian trajectories. Hence the first method based on centred finite-difference schemes is used to make all the main figures of the paper. The second method is only used in some Supplement to show the robustness of the results as well as to attribute the total heating to some specific physical processes, which cannot be provided by the first method. It is not surprising that the first method accurately approximates the potential temperature variations along trajectories because the computations of both the trajectories and the finite differences are made on the same grid and, in that sense, are self consistent. The second method less well approximates the potential temperature variations along trajectories because the post-process treatment to get the individual diabatic temperature tendencies and the underlying interpolations were not the same as for the other variables (wind, temperature). Finally, to get the PV tendencies due to the heating (first three terms on the right-hand side of Eq. 4), centred finite-difference schemes are applied to the heating.

## 2.3 Reanalysis and observations

### 2.3.1 Reanalysis and operational analysis

Operational analyses from ARPEGE and the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) are used on the same regular latitude–longitude grid with  $0.5^\circ$  resolution and the same pressure levels with 50 hPa spacing as the model simulations outputs. ERA5 reanalysis datasets (Hersbach et al., 2020) are also used with the same  $0.5^\circ$  grid spacing – that is, keeping one grid point every two grid points from the  $0.25^\circ$  original grid of the reanalysis.

### 2.3.2 Airborne observations and comparison to model outputs

Two coordinated flights of the SAFIRE and DLR Falcon aircraft were conducted from 09:00 to 12:00 UTC on 2 October between Iceland and Greenland to sample the WCB outflow region of the Stalactite cyclone. The two aircraft followed each other with a 10 min interval between 09:00 and 10:15 UTC. The 95 GHz Doppler cloud radar RASTA (Radar Airborne System Tool for Atmosphere; Delanoe et al., 2013) on board the SAFIRE Falcon measured both reflectivity and Doppler velocity along three antennas (nadir, backward and transverse) that allow for reconstructing the three wind components. The temporal resolution between two consecutive nadir measurements being 750 ms and the typical Falcon 20 speed  $200 \text{ m s}^{-1}$ , the horizontal resolution is near 300 m. Additional wind measurements were made during the flight of the SAFIRE Falcon by launching nine dropsondes and by in

situ sensors at the aircraft level. The 2  $\mu\text{m}$  scanning coherent/heterodyne detection Doppler wind lidar (DWL; Weissmann et al., 2005; Witschas et al., 2017) on board the DLR Falcon measured vertical profiles of line-of-sight wind speed, horizontal wind vectors and detected aerosol/cloud layers.

Since the model output grid is  $0.5^\circ \times 0.5^\circ$ , which corresponds to  $28 \text{ km} \times 55 \text{ km}$  spacing in longitude and latitude respectively, the radar RASTA and DWL wind measurements are averaged over intervals of 180 s to get the same horizontal resolution as the model outputs (the Falcon travels a distance of 36 km in 180 s).

### 3 Impact of convection representation on heating rates and warm conveyor belt trajectories

#### 3.1 General overview

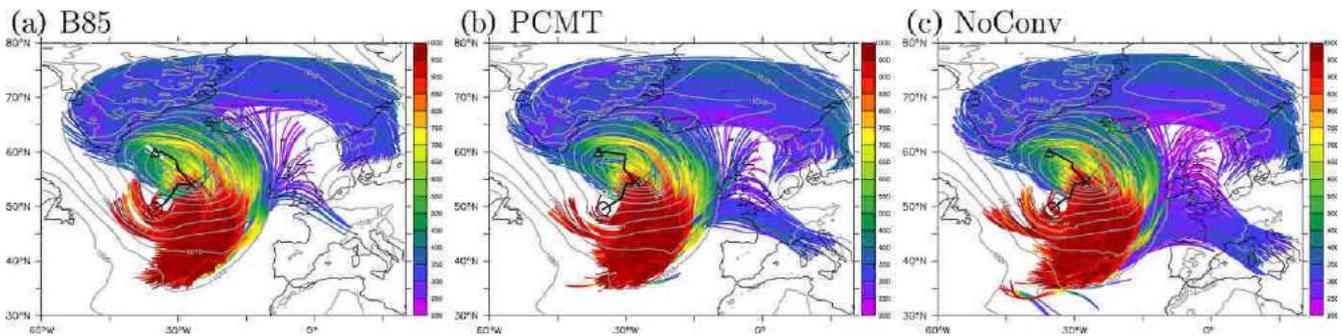
Differences in sea level pressure (SLP) between the three integrations are rather small (Fig. 1). The tracks of the minimum SLP are rather similar, and differences in the minimum SLP do not go beyond 2 hPa at 1 d lead time (962.9 hPa in B85, 962.0 hPa in PCMT and 961.2 hPa in NoConv) or at 2 d lead time (959.5 hPa in B85, 959.8 hPa in PCMT and 958 hPa in NoConv). This subsection is more particularly dedicated to presenting differences in the main characteristics of the WCB between the three runs. The number of trajectories satisfying the WCB criterion of the present study (ascent rate of at least 300 hPa in 24 h) does not differ much between the three runs: 9876, 10 086 and 11 421 for B85, PCMT and NoConv simulations respectively. The WCB trajectories share many common features between the runs. A majority of trajectories exhibit a well-marked cyclonic curvature during their ascents to the east of the surface cyclone and then change their curvature in the upper troposphere over Greenland to become more zonally oriented and slightly anticyclonically curved toward Scandinavia (Fig. 1). These trajectories in the upper troposphere span a larger latitudinal band in NoConv than PCMT and B85 (compare the banded areas formed by the blue segments of the trajectories). The largest difference between the three runs concerns another set of trajectories that has a more zonal orientation from the mid-Atlantic to western Europe near  $45^\circ \text{ N}$ . There are very few such trajectories in the B85 run, while they are more abundant in the PCMT run and are very numerous in the NoConv run.

Figure 2 presents averaged quantities along the WCB trajectories of the three runs. The mean pressure is about 894 hPa at 12:00 UTC on 1 October for the three runs and reaches 415 hPa for B85, 427 hPa for PCMT and 435 hPa for NoConv at 12:00 UTC on 3 October. Some differences also appear in the 90th percentile, with the WCB trajectories of the B85 run having systematically lower pressure levels than the other two runs (Fig. 2a). There are not so many differences between the mean potential temperatures of the three runs. The only notable difference is the slightly higher

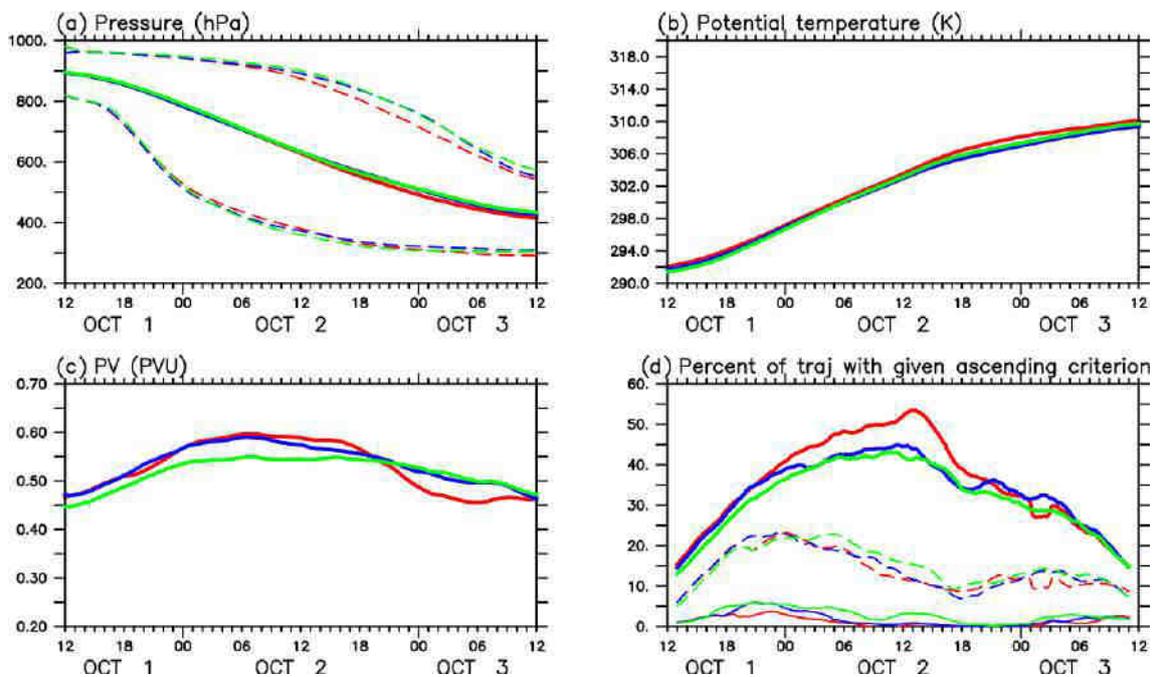
mean temperature in the B85 run than in the other two runs from 12:00 UTC on 2 October to 12:00 UTC on 3 October (Fig. 2b). This difference between B85 and PCMT is significant at the 99 % level following a Welch test, while differences with NoConv are not significant. The time evolution of the PV along WCB trajectories is characterized by an increase followed by a decrease (Fig. 2c), which is a classical picture of WCBs as they undergo PV production at low levels below the heating layer and PV destruction at upper levels above the heating layer (e.g. Wernli and Davies, 1997; Schemm et al., 2013). Differences in PV between the three runs are more important than differences in pressure or potential temperature. Trajectories of B85 and PCMT runs undergo a more important increase and decrease of PV than those of the NoConv run on average. While the two runs with activated parameterized convection exhibit a similar increase in PV during the first 24 h, they differ in the subsequent decrease in PV. B85 trajectories exhibit a more rapid decrease in PV between 18:00 UTC on 2 October and 00:00 UTC on 3 October, while PCMT trajectories show a smoother decrease in PV.

To get a better insight on the heating and PV evolution along trajectories, Fig. 3 shows the time evolution of the mean vertical profile of the heating and PV tendency due to the vertical gradient of the heating along WCB trajectories. While the magnitudes of the heating and PV tendency are rather similar between the three runs, some significant differences also appear. The NoConv case exhibits the strongest heating values at the early stage (18:00 UTC on 1 October) but the lowest ones at the later stage (18:00 UTC on 2 October to 06:00 UTC on 3 October) compared to the other two cases, suggesting a more rapid destabilization of the atmosphere in NoConv. In terms of PV, the expected picture of PV gain and loss below and above the main heating layer respectively is visible for the three cases. We observe that the PV gain region presents two peaks between 12:00 UTC on 1 October and 12:00 UTC on 2 October, one in the boundary layer and another near 700–800 hPa. But more importantly, the negative PV tendency above the heating layer is stronger for B85 than PCMT and NoConv during the whole period and in particular between 18:00 UTC on 2 October and 00:00 UTC on 3 October when WCB trajectories reach these upper levels. Such a difference is also seen by computing the total PV tendency due to the heating, i.e. by summing the first three terms on the right-hand side of Eq. (4) (not shown), and is in agreement with the more rapid decrease in PV seen in Fig. 2c for B85. This difference is mainly due to the fact that the heating reaches higher values in the upper troposphere between 2 and 3 October in B85 and thus has stronger vertical gradients (e.g. see the more important tightening of black contours between 00:00 UTC on 2 October and 00:00 UTC on 3 October in Fig. 3a than Fig. 3b, c).

Let us now document differences in the instantaneous ascending motion of the trajectories between the three runs. The proportion of trajectories ascending by more than 100,



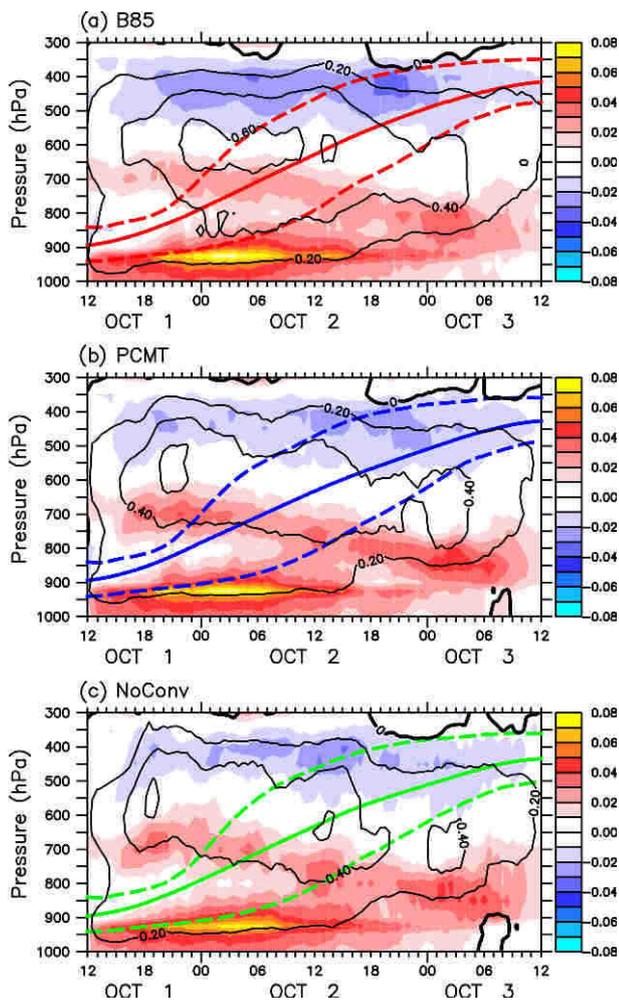
**Figure 1.** WCB trajectories from 12:00 UTC on 1 October to 12:00 UTC on 3 October coloured according to pressure (in hPa) for (a) B85, (b) PCMT and (c) NoConv simulations. The sea level pressure at 12:00 UTC on 2 October is shown in grey contours (interval: 5 hPa). Only WCB trajectories with an ascent criterion of at least 600 hPa in 48 h are shown for clarity purposes. In each panel, the black thick line represents the track of the minimum sea level pressure of the corresponding run between 12:00 UTC on 1 October and 12:00 UTC on 3 October. The circle, cross and triangle correspond to the position of the minimum sea level pressure at 12:00 UTC on 1 October, 12:00 UTC on 2 October and 12:00 UTC on 3 October respectively.



**Figure 2.** Time evolution along WCB trajectories of (a) mean pressure (solid lines, units: hPa) and the 10th and 90th percentiles of pressure (dashed lines), (b) mean potential temperature (solid lines; units: K), and (c) mean PV (units: PVU). Panel (d) indicates the percentage of WCB trajectories in each run satisfying a given ascending criterion at a given time: 100 hPa in 2 h (light solid lines), 50 hPa in 2 h (dashed lines) and 25 hPa in 2 h (bold solid lines). The red, blue and green curves correspond to B85, PCMT and NoConv simulations respectively.

50 and 25 hPa in 2 h between  $T - 1$  h and  $T + 1$  h is indicated as function of time  $T$  in Fig. 2d. NoConv has the largest number of trajectories ascending most rapidly (i.e. 100 hPa within 2 h and 50 hPa within 2 h), B85 is the run having the lowest number, and PCMT is in between regardless of the time. In the category of more modest ascents (25 hPa in 2 h), B85 has more such trajectories in proportion than the other two runs from 00:00 to 18:00 UTC on 2 October. Beyond 18:00 UTC on 2 October, the proportions of moderately ascending trajec-

tories are closer to each other between the three runs. These differences in the ascending motion properties are synthesized in Fig. 4. The NoConv trajectories undergo the most rapid vertical motion with a maximum vertical displacement of about 104 hPa in 2 h when averaged over all trajectories. In comparison, the averaged maximum ascending rate is 94 and 86 hPa for PCMT and B85 runs respectively (see the zero lag). However, the three curves cross each other near the 4 h lag. After that lag the order reverses and B85 shows the high-

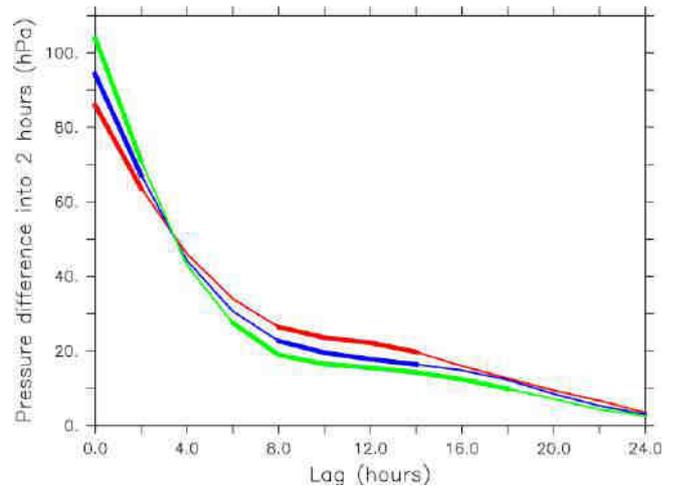


**Figure 3.** Time evolution of the mean vertical profiles along all WCB trajectories of the heating rate (interval:  $0.20 \text{ K h}^{-1}$ ) and the PV tendency part due to the vertical derivative of the heating (interval:  $0.01 \text{ PVU h}^{-1}$ ) for (a) B85, (b) PCMT and (c) NoConv simulations. The mean pressure of the WCB trajectories is shown in solid lines and the 25th and 75th percentiles in dashed lines.

est ascending rate. Therefore, even though B85 has a weaker ascending rate maxima, it is the run for which trajectories exhibit more sustained ascents. To conclude, parameterized deep convection tends to induce more sustained and longer-lasting ascending motion than when explicit deep convection occurs at grid scales of the model. The two schemes lead to the same qualitative effect, but B85 is the one whose behaviour distinguishes the most from explicit deep convection and presents more sustained ascents.

### 3.2 Differences in the rapid ascending trajectories during the first 12 h of the simulations

To better visualize the effect of parameterized deep convection on heating rates in physical space, horizontal maps and vertical cross sections are shown at 21:00 UTC on 1 Octo-



**Figure 4.** Time lag composite of the pressure difference in 2 h over WCB trajectories for B85 (red), PCMT (blue) and NoConv (green) simulations. The zero lag corresponds to the time of maximum pressure difference. Only WCB trajectories having their maximum pressure difference reached before 12:00 UTC on 2 October are considered to get the same number of trajectories at all time lags. Thick segments correspond to time lags for which each composite is significantly different from the other two composites at 99.9 % following a Welch test.

ber in Fig. 5. This time was chosen because it corresponds to large differences in the proportion of the most rapidly ascending trajectories (i.e. those exceeding 100 hPa ascent within 2 h) between the three runs (Fig. 2d). Only the positions of WCB trajectories belonging to this category are shown in Fig. 5. The cold and bent-back warm fronts bring strong similarities between the three runs. Additionally, heating rates are strong for all runs ahead of the cold front and in the vicinity of the bent-back warm front (upper left side of Fig. 5a, c, e). Not surprisingly all the positions of the most rapidly ascending trajectories at that time are located within these regions of strong heating rate. However, the largest differences between the three runs appear in the heating rates ahead of the cold front ( $42\text{--}52^\circ \text{ N}$ ,  $26\text{--}22^\circ \text{ W}$ ) and not near the bent-back warm front ( $52\text{--}56^\circ \text{ N}$ ,  $34\text{--}30^\circ \text{ W}$ ). Ahead of the cold front, NoConv is characterized by four regions of strong heating rates with a few degrees extent in latitude and longitude and whose peak values are beyond  $3 \text{ K h}^{-1}$ . In the same area, PCMT shows similar peak values, but the heating rates are more homogeneously distributed along the WCB than NoConv. In contrast, the peak values for B85 are much weaker (no more than  $2.4 \text{ K h}^{-1}$ ), and the large values of the heating rate are even more homogeneously distributed than in PCMT.

Vertical cross sections of the heating rates averaged between  $42$  and  $48^\circ \text{ N}$  – that is, ahead of the cold front – are shown in Fig. 5b, d and f. In that region, the heating rate is much stronger in NoConv and PCMT than in B85 over

the whole tropospheric column. In NoConv, strong heating rates appear in the upper troposphere at pressure levels lower than 350 hPa that do not exist for the two runs with parameterized convection. This yields slightly lower isentropic levels in the upper troposphere for NoConv compared to the other two runs. However, it is worth mentioning there also exist regions where NoConv tends to get much smaller heating values ahead of the cold front than the other two runs (see for instance the area centred at 50° N, 22° W). These results show similar findings to those of Martinez-Alvarado and Plant (2014), who compared two runs with full and reduced parameterized convection.

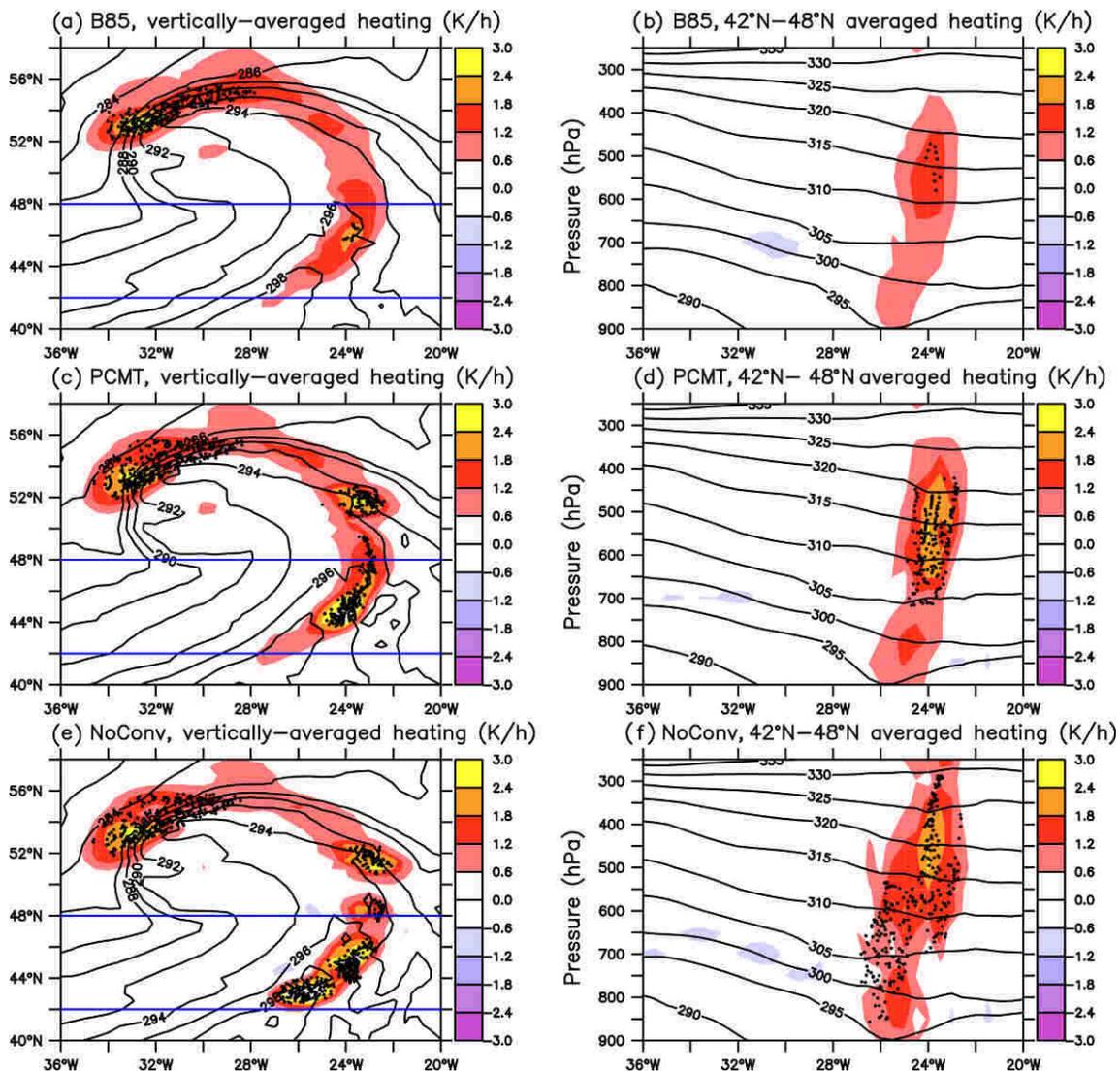
### 3.3 Differences in the moderate ascending trajectories after 24 h simulations

Following Fig. 2d another particular interesting period is between 12:00 and 18:00 UTC on 2 October during which the proportion of trajectories with moderate ascents (25 hPa in 2 h) becomes much larger for B85 than the other two runs. Heating rates and trajectories having moderate ascents at 12:00 UTC on 2 October are shown in Fig. 6. Compared to 15 h earlier (Fig. 5), the peak values of the vertically averaged heating rates are weaker, and significant differences between the runs appear along the bent-back warm front. However, the qualitative picture is roughly the same. The NoConv heating rate is more bumpy along the front and marked by more pronounced regions of high and low heating rates than the other two runs. B85 is the run with smoother heating rate (Fig. 6a, c, e). There are two particular regions of inactivity in NoConv: one located along the front at 62° N, 26° W and another one further north between 63 and 65° N and from 32 to 24° W. PCMT exhibits slightly stronger heating rates and more trajectories with moderate ascents than NoConv in those regions. But B85 has an even stronger heating rate and many more trajectories. The difference in the number of trajectories seen in Fig. 2d mainly comes from these two regions.

Vertical cross sections of heating rates averaged between 63 and 65° N show that the second region of marked WCB inactivity in NoConv occurs at upper levels between 350 and 550 hPa in the longitudinal sector 34–24° W. We do see many more WCB trajectories with moderate ascents in B85 than in NoConv in association, with the heating rate peak values of the former being more than twice that of the latter (Fig. 6b, f). As already seen in Figs. 1–5, PCMT has an intermediate behaviour between B85 and NoConv (Fig. 6d). A decomposition of the heating rate into various physical processes shows that differences in the heating rates in that sector come from the resolved heating (sensible plus latent) and not the parameterized heating, which is mainly negligible at upper levels at that time (see Figs. S1 and S2). Parameterized convection and turbulence have similar amplitudes to the resolved heating in the lower troposphere but are negligible compared to the resolved heating at upper levels. The radiative term is

found to be much smaller than the other terms in that sector at that time over the whole troposphere. Therefore it is not a direct effect of parameterized heating that is observed here but rather an indirect effect in which parameterized deep convection interacts with the resolved flow so that it leads to more intense WCB activity north of the front in the upper troposphere in B85 than in the other two runs.

Figure 7 provides a connection between the WCB activity, heating rates and PV patterns north of the front where large differences between the three runs have been identified. At lower pressure than 450 hPa, two main regions of WCB activity appear: one near 42–40° W in the vicinity of a well-marked localized negative PV region seen in all three runs (Fig. 7a, c, e) and another one along the band 34–24° W discussed previously. Areas of negative PV are visible slightly above and west of the latter region of WCB activity. These areas are small and centred at 32° W for PCMT and NoConv, whereas the area is much larger and extends from 36 to 32° W in B85. More WCB trajectories in B85 than in the other two runs would be a good explanation for the more extended negative PV area in B85. However, we observe that this negative PV area is not exactly co-located with the positions of the WCB trajectories which are a bit below and more to the east relative to the negative PV area (Fig. 7a). In other words, the PV loss undergone by WCB trajectories is not co-located with the negative PV area. To explain this shift of the negative PV area relative to the PV destruction region, it is important to consider spatial redistribution of PV by advection terms and in particular the indirect diabatic effects of advection by the divergent wind which tends to displace the PV destruction outward of the WCB outflow region (Archambault et al., 2013; Steinfeld and Pfahl, 2019). Ageostrophic winds that can be considered a proxy for the divergent winds are shown in Fig. 7b, d and f. West of the main WCB outflow region, i.e. west of 30° W in the 300–400 hPa layer, winds are westward and upward. Moving to the west they become more and more horizontal and align more with the isentropic slopes (compare the orientation of the vectors with the slope of the 315 K isentropic surface in Fig. 7a). The PV advection by the ageostrophic wind is generally negative because the stratospheric high PV values are west of the tropospheric low PV values. This leads to a westward expansion of the ridge anomaly, consistent with other studies (Steinfeld and Pfahl, 2019). Still looking at the same region, the ageostrophic winds are stronger in B85 than in PCMT or NoConv as expected from its more intense WCB activity there. While the advection terms are similar between the three runs between 44 and 40° W, they significantly differ between 38 and 36° W, where negative PV advection is clearly stronger for B85. Therefore, stronger WCB activity in connection with the higher heating rate in the middle and upper troposphere leads to stronger PV loss in the WCB outflow region in B85 compared to the other two runs, which is then advected upward and westward by the stronger diver-

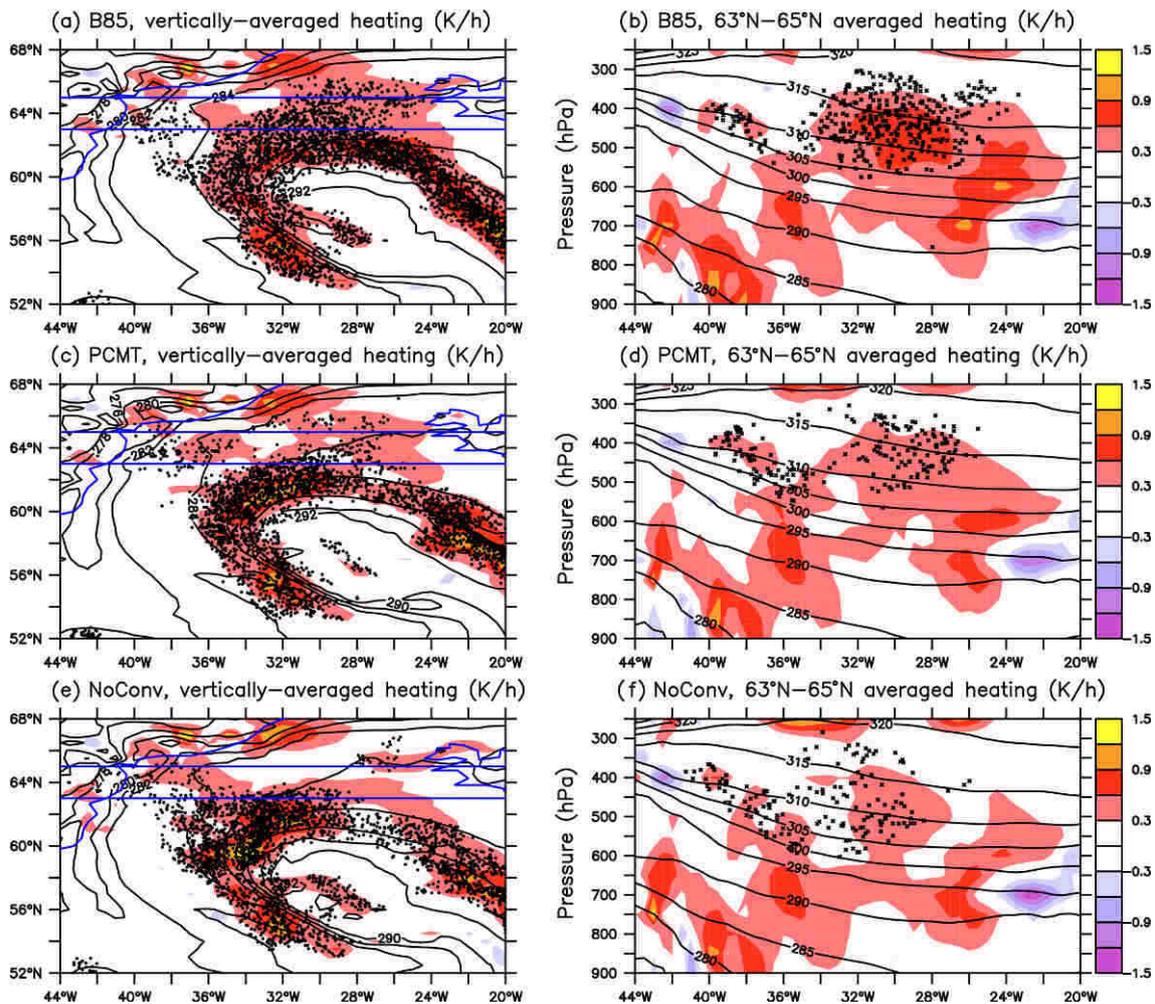


**Figure 5.** (a, c, e) Vertically averaged heating rate between 300 and 800 hPa (shadings; units:  $\text{K h}^{-1}$ ), potential temperature at 850 hPa (interval: 2 K) and WCB trajectories satisfying 100 hPa ascent in 2 h at 21:00 UTC on 1 October for (a) B85, (c) PCMT and (e) NoConv simulations. (b, d, f) Latitudinally averaged ( $42\text{--}48^\circ\text{N}$ ; see blue lines in panels a, c and e) heating rate (shadings; units:  $\text{K h}^{-1}$ ), potential temperature (contours; interval: 5 K) and WCB trajectories satisfying 100 hPa ascent within 2 h and located in the same latitudinal band at 21:00 UTC on 1 October for (b) B85, (d) PCMT and (f) NoConv simulations.

gent winds to form a more important zone of negative PV for the former run.

The findings of this section can be summarized as follows. With parameterized deep convection, heating rates are more homogeneously distributed along the cold and warm fronts, while without parameterized deep convection, heating rates are marked by strong spatial variations along the fronts. This leads to more rapid instantaneous ascents for the latter and more moderate but more sustained ascents for the former. Martinez-Alvarado and Plant (2014) also emphasized the smooth and regulating effect of parameterized deep convection within WCB. This is due to the fact that parameterized convection releases convective instability at subgrid

scales, while explicit convection releases convective instability at resolved scales. Among the two schemes, B85 is the one having the greatest difference with NoConv, while PCMT has a more intermediate behaviour. The three runs significantly differ in terms of the impact of WCB activity on PV. The B85 scheme generally leads to higher heating rates at upper levels and more PV destruction in the upper troposphere during the second half of the simulations.



**Figure 6.** (a, c, e) Vertically averaged heating rate between 300 and 800 hPa (shadings; units:  $\text{K h}^{-1}$ ), potential temperature at 850 hPa (interval: 2 K) and WCB trajectories satisfying 25 hPa ascent in 2 h at 12:00 UTC on 2 October for (a) B85, (c) PCMT and (e) NoConv simulations. (b, d, f) Latitudinally averaged ( $63\text{--}65^\circ\text{N}$ ; see blue lines in panels a, c and e) heating rate (shadings; units:  $\text{K h}^{-1}$ ), potential temperature (contours; interval: 5 K) and WCB trajectories satisfying 25 hPa ascent in 2 h and located in the same latitudinal band at 12:00 UTC on 2 October for (b) B85, (d) PCMT and (f) NoConv simulations.

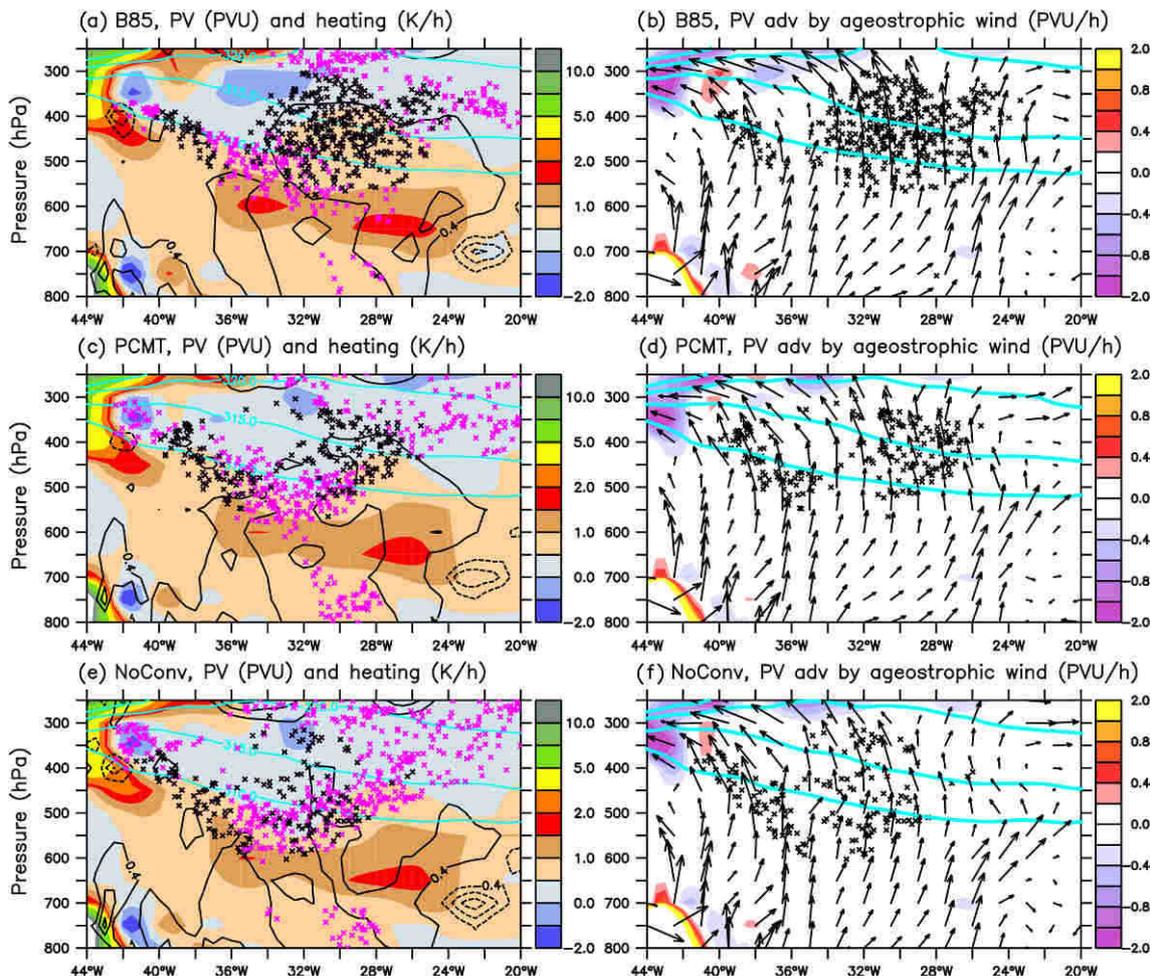
#### 4 Impact of deep convection representation on the jet stream

The impact on the jet stream is investigated in the present section, and forecast skills of the three runs are assessed by comparing to the reanalyses and airborne observations.

##### 4.1 Comparison to (re)analyses

PV horizontal maps at 300 hPa are shown for the three runs in Fig. 8 and compared with ECMWF-IFS and ARPEGE analyses and the ERA5 reanalysis. The separation between stratospheric and tropospheric air is well marked in all panels by an abrupt jump in PV from near zero and negative values to large positive values close to 8–10 PVU. This boundary is associated with a tropopause fold as shown in Fig. 7a, c and e. Despite this well-defined limit being located more or

less at the same place for the six datasets (three simulations, two analyses and one reanalysis), the region of ridge building is not characterized by homogeneously distributed weak and negative PV values. For instance, at  $64^\circ\text{N}$ , going from  $45$  to  $25^\circ\text{W}$  (i.e. roughly along the main legs of the two Falcon flights), B85 exhibits large positive PV associated with stratospheric air, a sudden decrease to slightly negative values, then another area of positive PV values and finally negative values (Figs. 7a and 8d). The area of positive PV within the ridge forms a band of PV with values varying between 0.5 and 1.5 PVU largely covering the Greenland eastern coast (Fig. 8d) but whose vertical extent is rather limited between 300 and 350 hPa at  $39^\circ\text{W}$  (Fig. 7a). A similar narrow band of positive PV exists in the other datasets but with different values of PV and different locations. In the ECMWF-IFS analysis and ERA5, the positive PV values are smaller but have the

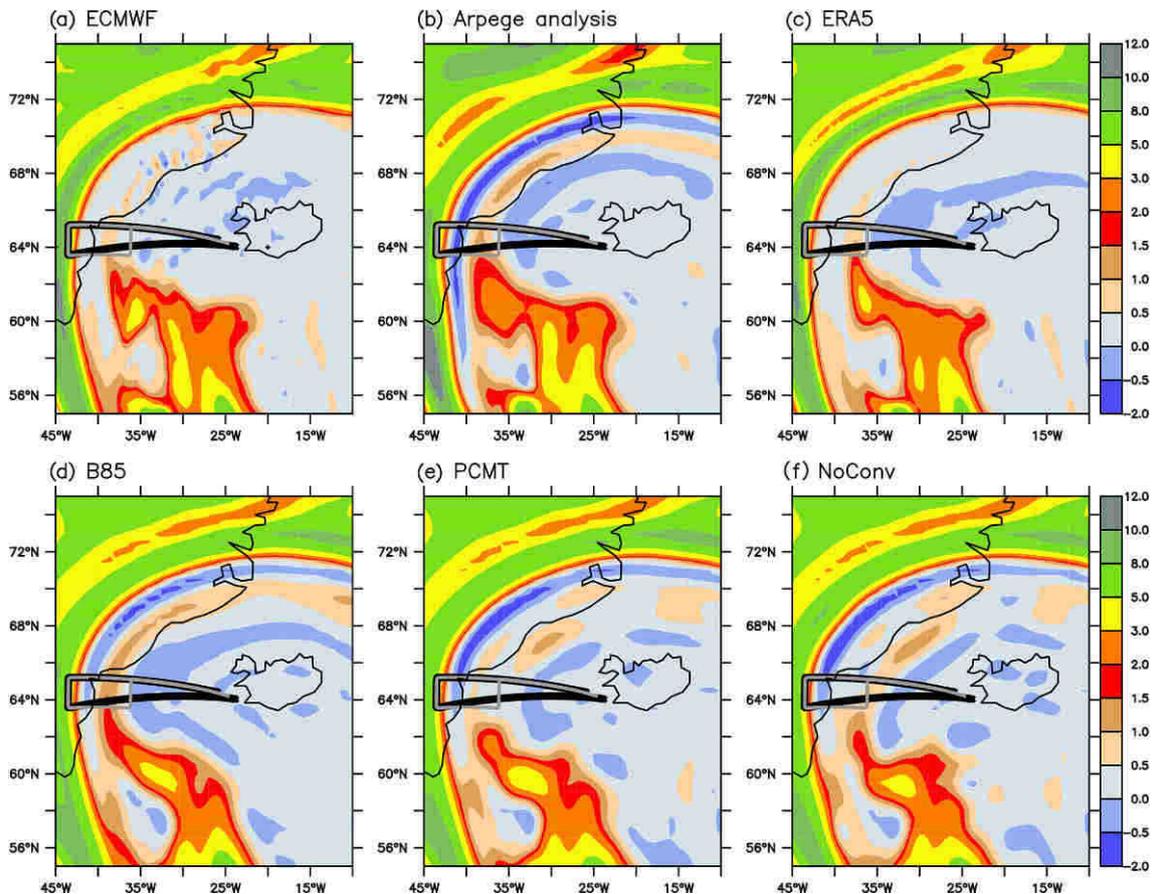


**Figure 7.** (a, c, e) Latitudinally averaged ( $63\text{--}65^\circ\text{N}$ ) PV (shadings; units: PVU), heating rate (black contours; interval:  $0.4\text{ K h}^{-1}$ ), potential temperature (light blue contours for 310, 315, 320 and 325 K) and WCB trajectories satisfying 25 hPa ascent in 2 h (black crosses) and the other WCB trajectories (purple crosses) and located in the same latitudinal band at 12:00 UTC on 2 October for (a) B85, (c) PCMT and (e) NoConv simulations. (b, d, f) Latitudinally averaged ( $63\text{--}65^\circ\text{N}$ ) PV advection by the zonal and vertical components of the ageostrophic wind (shadings; units:  $\text{PVU h}^{-1}$ ), and the associated ageostrophic wind vectors (black arrows). The light blue contours represent potential temperature as in left panels. The ageostrophic wind vectors are represented by multiplying respectively the zonal and vertical components by the distances occupied by 1 m in longitude and 1 Pa in pressure on the diagram.

same location as in B85 (Fig. 8a, c). In the ARPEGE analysis, the band is rather similar in intensity and location to B85, but this is not surprising since the ARPEGE analysis is made using a deterministic forecast whose deep convection scheme is B85. In contrast, in NoConv the band is eastward shifted by a few degrees in longitude compared to B85 (Figs. 7e and 8f). In PCMT, the band is less well defined, and the pattern resembles a mix between those of B85 and NoConv patterns.

Because there are two jumps in PV at 300 hPa, a double jet structure is well visible in Fig. 9 for all datasets, with the main jet being the one more to the northwest separating the tropospheric air from the stratospheric air. While the secondary jet lies just above the Greenland eastern coastline for ECMWF-IFS, ARPEGE analyses, ERA5 and B85 (Fig. 9a–d), it is located further east in PCMT and NoConv (Fig. 9e–f).

This is consistent with the PV gradient deduced from Fig. 8. In the immediate vicinity of the eastern coast of Greenland north of  $64^\circ\text{N}$ , PV values are higher to the west and lower to the east in ECMWF-IFS, ARPEGE analyses, ERA5 and B85, which explains the co-location of the secondary jet for these datasets. Anomalies with respect to the ECMWF-IFS analysis are shown in black contours in Fig. 9d–f. The B85 run produces a jet that is slightly too strong with positive anomalies of  $4\text{ m s}^{-1}$  amplitude (Fig. 9d), consistent with the higher PV values along the narrow ribbon of 0.5–1.5 PVU values shown in Fig. 8d compared to ECMWF (Fig. 8b). The other two runs are marked by an eastward shift of the jet and negative anomalies of the order of  $8\text{ m s}^{-1}$  amplitude over the Greenland eastern coastline (Fig. 9e–f), consistent with the eastward shift of the ribbon of 0.5–1.5 PVU values com-



**Figure 8.** Potential vorticity (units: PVU) at 300 hPa at 12:00 UTC on 2 October for the (a) ECMWF-IFS operational analysis, (b) ARPEGE operational analysis, (c) ERA5 reanalysis, (d) B85 run, (e) PCMT run and (f) NoConv run. The black and grey lines correspond to the SAFIRE and DLR Falcon flights that occurred between 09:00 and 12:00 UTC on 2 October in an anticlockwise direction.

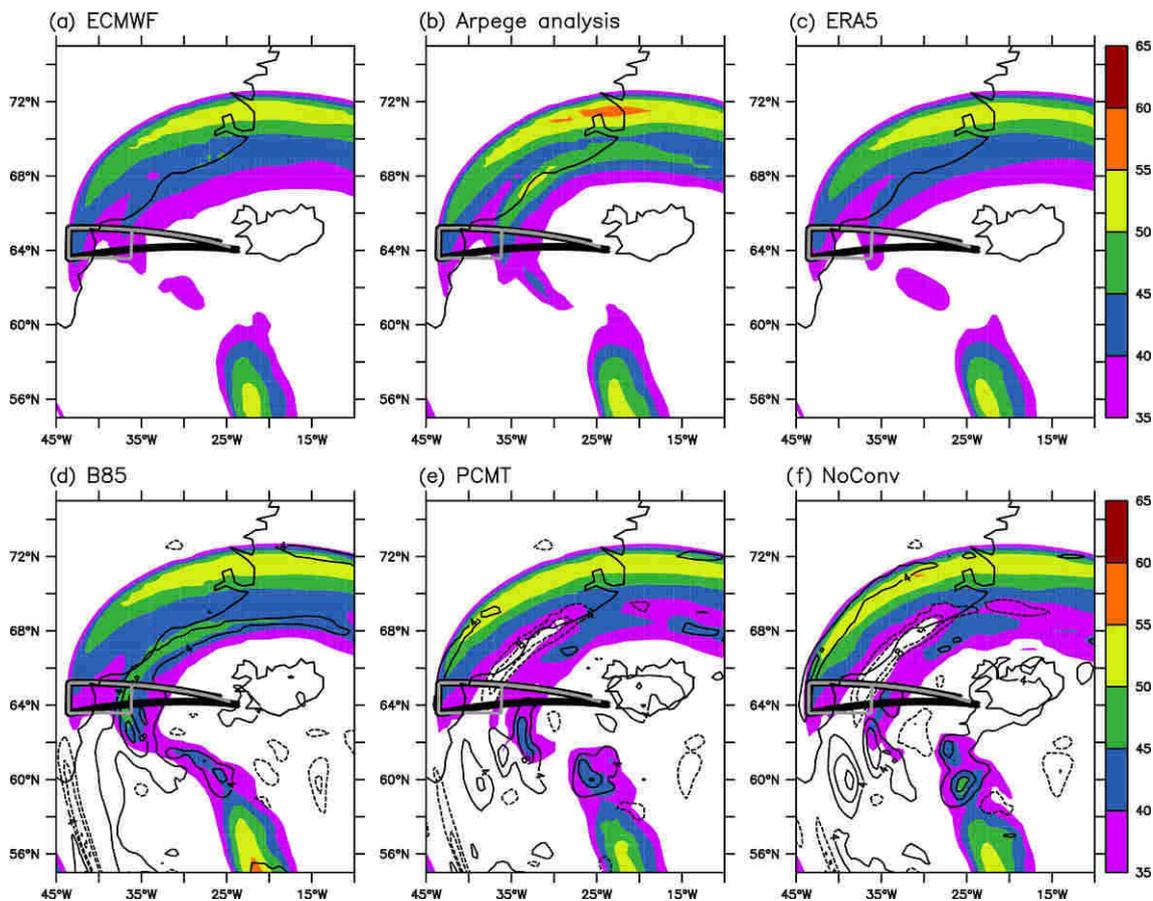
pared to the other datasets. At that time, B85 performs better than PCMT or NoConv in capturing the distance between the two jets, and this is attributed to the more active WCB in that region in B85 that reinforces the PV gradient further west and pushes the secondary jet closer to the main one. Another interesting feature is the stronger wind magnitude of the main jet in PCMT and NoConv compared to B85. This is due to more negative PV values just east of the tropopause fold (Fig. 8d–f) and more active WCB in that area for the former runs than the latter (see 42° W in Fig. 7a, c and e). Here, B85 is also slightly closer to ECMWF-IFS than the other two runs. In other words, the WCB outflow region is separated into two branches: one reinforcing the main jet and the other one the secondary jet. The runs are distinguished from each other in the relative importance of the two branches.

Six hours later, at 18:00 UTC on 2 October, the double jet structure is much less pronounced in ECMWF-IFS, ARPEGE analyses and ERA5 as well as in PCMT and NoConv, but it is still there in B85 (Fig. 10). Furthermore, the anomalies with respect to ECMWF-IFS are larger north of 64° N in B85 than PCMT or NoConv. In that case, B85 does

not perform better than the other two runs and is even less skilful. Further south, the jet stream is too weak in NoConv compared to the other datasets (see the lower right side of the panels in Fig. 10d–f), but this sector is not the focus of the present study as no flights were conducted there.

#### 4.2 Comparison to airborne observations

For the flights of the two aircraft occurring in the region of the double jet structure (Fig. 9), it is worth comparing the three forecasts to airborne observations to determine which one performs better in representing the jet stream structure and intensity. The two aircraft followed each other with 10 min lag and observed the same meteorological features during more than half of the flights' duration. Besides, since the Doppler radar on board the SAFIRE Falcon and the Doppler lidar on board the DLR falcon are not sensitive to the same particles of the atmosphere, they provide complementary datasets as seen in Fig. 11a and b. The data have been interpolated at the model grid in the present study (Sect. 2.3.2), while higher-resolution profiles of the same datasets are shown in Fig. 9 of Schäfer et al. (2018).



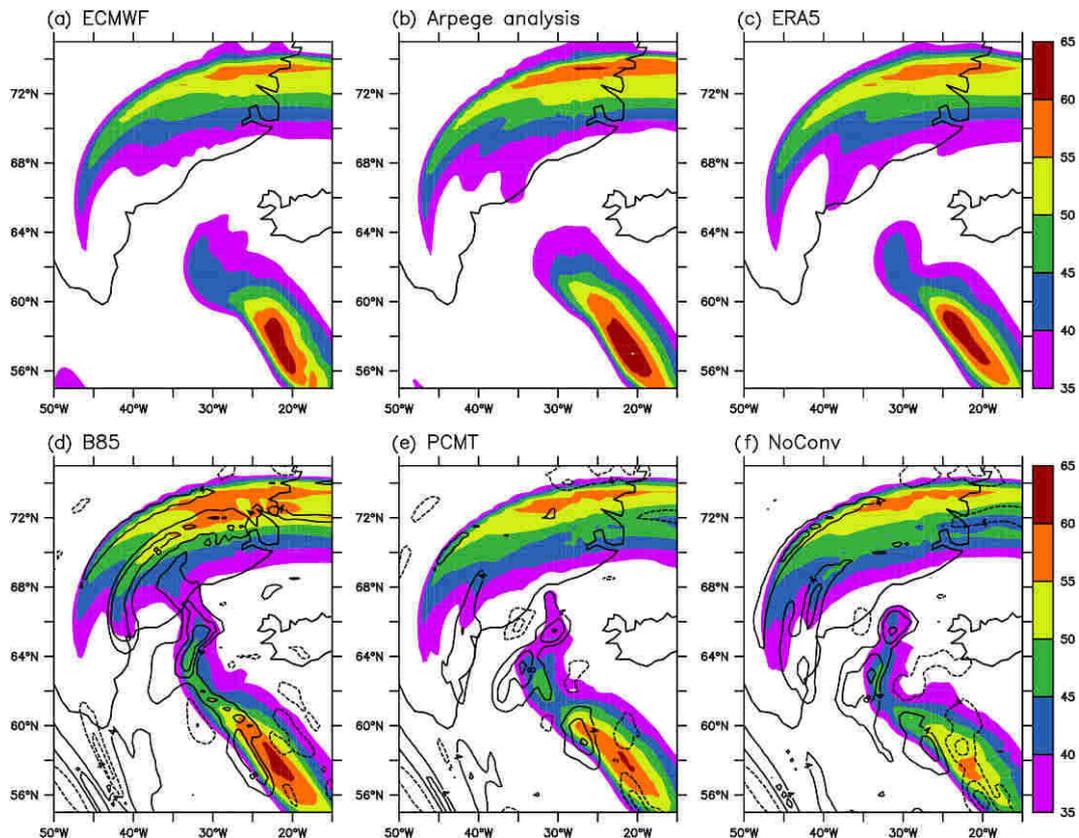
**Figure 9.** Wind speed (shadings; units:  $\text{m s}^{-1}$ ) at 300 hPa at 12:00 UTC on 2 October for the (a) ECMWF-IFS operational analysis, (b) ARPEGE operational analysis, (c) ERA5 reanalysis, (d) B85 run, (e) PCMT run and (f) NoConv run. In panels (d)–(f), the black contours represent the wind speed anomalies (interval:  $4 \text{ m s}^{-1}$ ) of the three runs with respect to the ECMWF-IFS operational analysis. The black and grey lines correspond to the SAFIRE and DLR Falcon flights that occurred between 09:00 and 12:00 UTC on 2 October in an anticlockwise direction.

In Fig. 11a, a good correspondence generally appears between the three kinds of measurements: Doppler radar, airborne in situ and the nine dropsondes. The intensities of the lower- and upper-level jets detected close to the Greenland coastline are similar in the dropsondes and radar data. The three datasets also agree on the intensities of the upper-level wind magnitude; two main regions of high values are distinguishable: one between 9.5 and 9.7 h and the other between 9.8 and 10 h. Some local discrepancies also occur for instance between the dropsonde data and radar data at 10.1 h or between the dropsonde data and in situ aircraft measurements at 10.5 h. The presence of a double jet structure is confirmed by the lidar measurements of the DLR Falcon (Fig. 11b), with the two jets being quite close to each other.

Figure 11c–e show the wind magnitude of the three forecasts. The double jet structure is present in all three runs as already shown in Fig. 9, but once again we do see that the distance between the two jets is significantly larger for PCMT and NoConv than B85. B85 is the forecast with the

closest patterns to the observations, even though the distance between the two jets is still slightly too large in that run (e.g. compare Fig. 11b and c).

To better visualize differences between forecasts and observations, forecast errors with respect to the observations made during the SAFIRE Falcon flight are shown in Fig. 12. While the three forecasts share the same errors at low levels with a too strong low-level jet in the vicinity of Greenland (see near 9.8 and 10.6 h), errors in the upper troposphere do not have the same magnitude among the three forecasts. Between 200 and 400 hPa, PCMT and NoConv forecast errors exhibit tripolar anomalies (positive–negative–positive) between 9.2 and 9.9 h and between 10.7 and 11.1 h, and their magnitude is about  $12 \text{ m s}^{-1}$ . The negative forecast errors shown at 9.4 and 10.9 h are consistent with the negative anomalies shown in Fig. 9e and f in which the reference is the ECMWF-IFS analysis. In contrast, B85 errors fluctuate between 6 and  $9 \text{ m s}^{-1}$  only. The root-mean-square error computed in the upper troposphere (pressure lower than



**Figure 10.** As in Fig. 9 but at 18:00 UTC on 2 October.

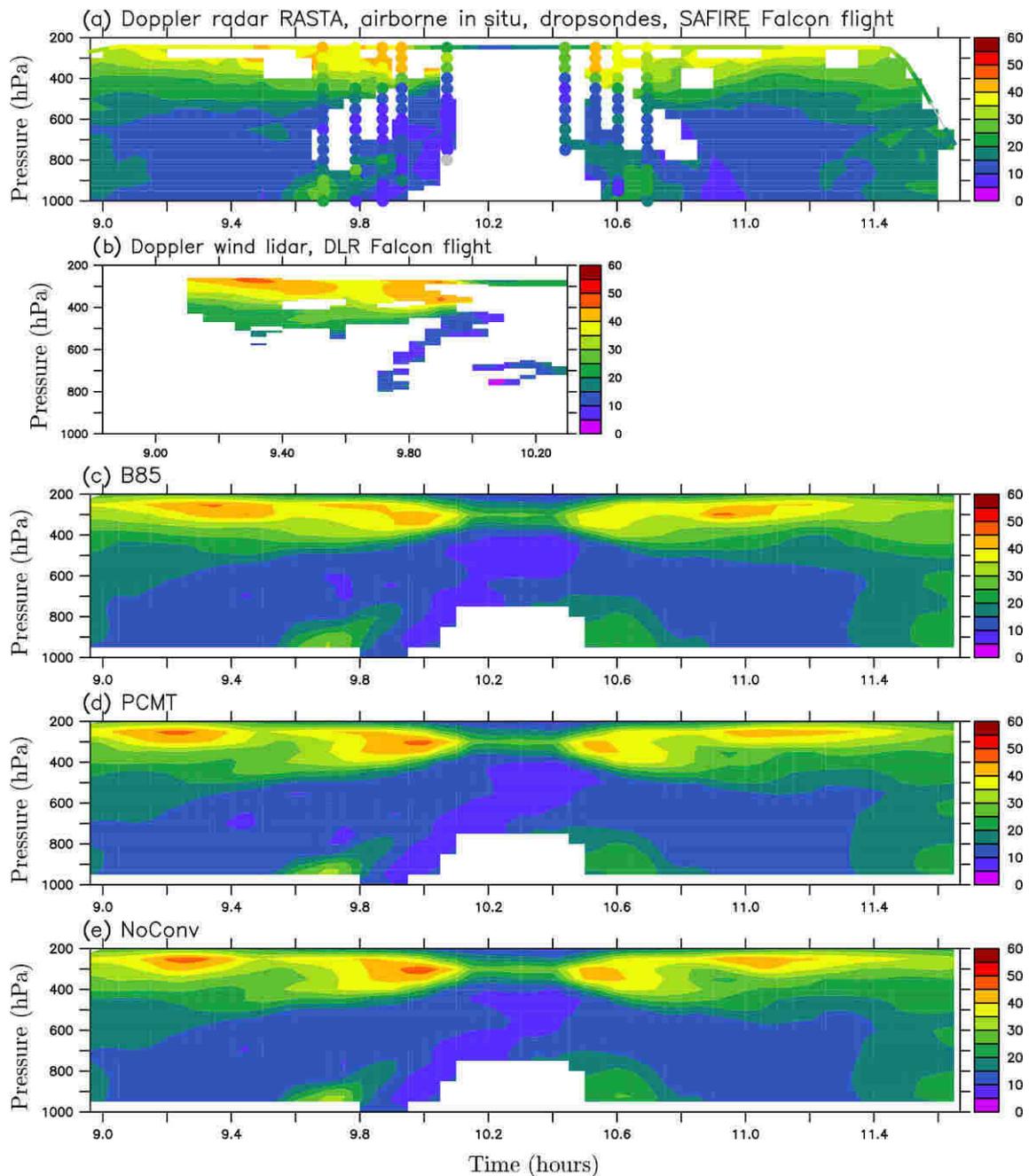
500 hPa) for the difference between each run and the radar observations is  $4.36 \text{ m s}^{-1}$  for B85,  $5.00 \text{ m s}^{-1}$  for PCMT and  $4.81 \text{ m s}^{-1}$  for NoConv. Doing the same computation but using in situ airborne measurements, we get  $4.27 \text{ m s}^{-1}$  for B85,  $5.43 \text{ m s}^{-1}$  for PCMT and  $5.65 \text{ m s}^{-1}$  for NoConv, corresponding to an improvement of 10 % to 30 % in the representation of the wind speed in that region in B85 compared to the other two runs. To conclude, comparison with airborne observations confirms that B85 performs better than the other two forecasts in the location of the jets at 24 h lead time.

## 5 Conclusions

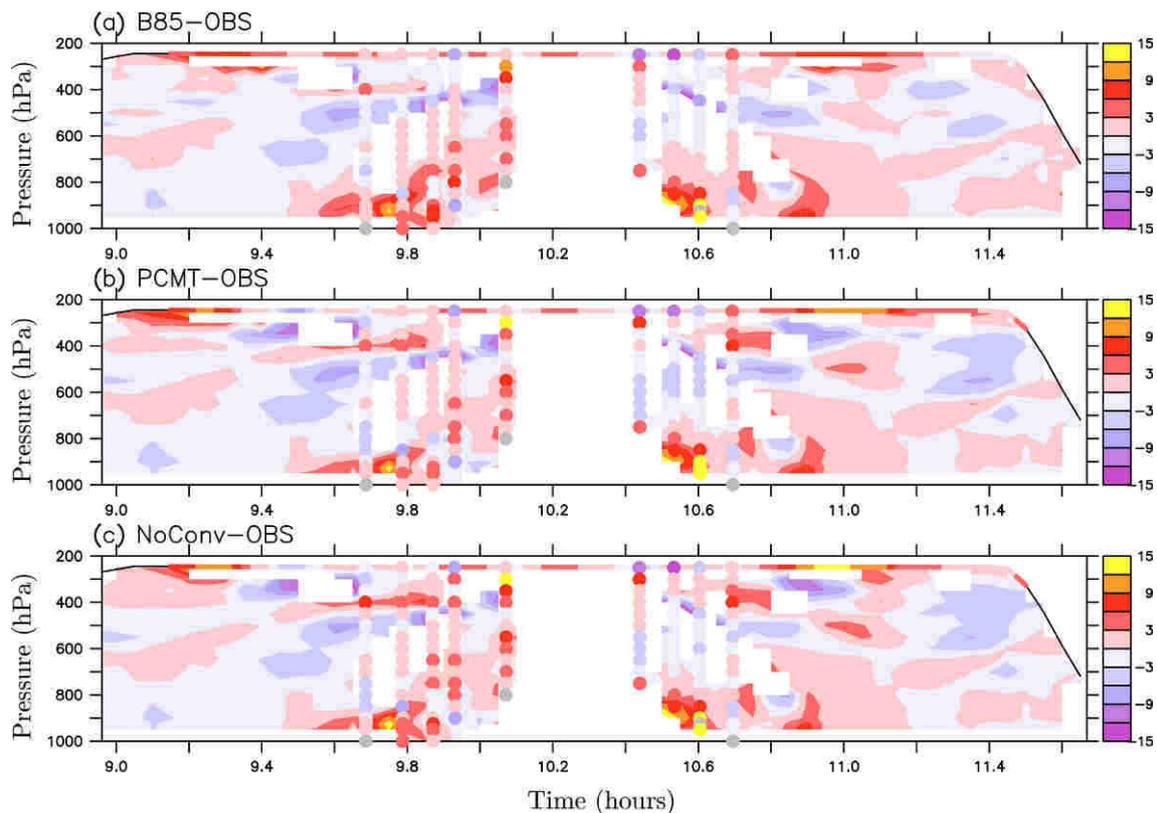
The effect of parameterized deep convection on WCB activity and jet stream was investigated by performing simulations of an explosively developing large-scale cyclone, which occurred from 29 September to 3 October 2016 during NAWDEX and is called the Stalactite cyclone, using the Météo-France global atmospheric model ARPEGE. Three simulations differing only from their deep convection representation are analysed. For two of them, parameterized deep convection was activated with distinct schemes (B85 corresponding to the Bougeault, 1985, scheme and PCMT corresponding to the Piriou et al., 2007, scheme), while for the

third one, parameterized convection was turned off. The main findings can be summarized as follows and correspond to answers to the three main questions raised in the introduction.

- How different are WCBs between simulations with active parameterized deep convection and those without? When the parameterized deep convection scheme is turned off, convective instability is released at the resolved scales such that few localized cells of a few degrees extent in longitude and latitude appear along the cold and warm fronts of the cyclone. These localized cells are characterized by strong heating and fast ascending motion. In contrast, when parameterized deep convection is active, the heating rate is more homogeneously distributed along the fronts, and its large values are more spread out while having weaker maxima than when deep convection is explicit at the model grid scales. This results in different behaviours in WCB ascents: without parameterized deep convection, ascents are rapid and abrupt, while with parameterized deep convection, ascents are more moderate but more sustained. These results confirm the regulating effect of parameterized deep convection emphasized by Martínez-Alvarado and Plant (2014).



**Figure 11.** Wind speed (units:  $\text{m s}^{-1}$ ) as function of time: **(a)** derived from the SAFIRE Falcon Doppler radar RASTA, airborne in situ measurements (upper coloured line) and dropsondes (coloured circles) along the whole flight track; **(b)** derived from the DLR Falcon  $2 \mu\text{m}$  Doppler wind lidar along part of the flight common to the SAFIRE flight; and derived from **(c)** B85, **(d)** PCMT and **(e)** NoConv runs along the SAFIRE flight track. In panels **(a)** and **(b)**, the measured wind speeds have been interpolated on the model grid. Note that all the panels can be compared to each other and the position and span of panel **(b)** are precisely chosen to have co-location in space with the other panels. The difference in time is there because the DLR Falcon was ahead of the SAFIRE Falcon with roughly a 20 min lag. In panels **(c)**–**(e)**, the wind speed data are derived from 15 min output centred on the time of interest, which is that of the SAFIRE Falcon flight. See Sect. 2.3.2 for further details on the interpolations procedure.



**Figure 12.** Difference (experiment – observations) in wind speed (units:  $\text{m s}^{-1}$ ) for (a) B85, (b) PCMT and (c) NoConv simulations. The wind observations are composed of dropsondes, airborne in situ measurements and Doppler radar measurements along the SAFIRE Falcon track.

– What is the impact of parameterized deep convection on the jet stream at short range (less than 2 d)? Explicit convection yields strong localized heating over the whole troposphere that may potentially have a more important impact on the upper-level circulation locally than parameterized deep convection. However, the heating rate averaged over all WCB trajectories was found to be weaker for explicit deep convection than parameterized deep convection. Moreover, between 1 and 2 d lead times, one of the schemes (B85) shows more important averaged heating in the upper troposphere and more important vertical gradients of heating that lead to more PV destruction. This stronger heating can be attributed to the resolved heating and not the parameterized heating which is mainly localized at lower levels. These results should be contrasted with Done et al. (2006), who found that explicit convection yielded more PV destruction than parameterized convection in their case study. The present results also show that differences in WCB activity between explicit and parameterized deep convection may lead to large differences in the jet stream position and intensity at 1 d lead time. This is to be contrasted with Martinez-Alvarado and Plant (2014), who found rather modest impacts at the same short range.

The impact on the upper-tropospheric circulation may depend on case studies. Here the extratropical cyclone is deeper than the one studied by Martinez-Alvarado and Plant (2014).

– How different are WCBs and their impact on the jet stream between simulations performed with distinct deep convection schemes? The effects of the two deep convection schemes on WCB and their impact on the jet stream significantly differ from each other. B85 is the scheme inducing the most drastic differences compared to the run with explicit deep convection, while PCMT has a more intermediate behaviour. In terms of the impact on the jet stream, PCMT is closer to the run with explicit deep convection than to B85. Comparison with operational analyses and airborne observations of different types helped us to unambiguously determine the most skillful forecasts. At 1 d lead time, B85 performed better than the other two runs to represent the double jet structure. The shorter distance between the main and secondary jets in B85 than in other two forecasts was found to be more realistic by comparing to airborne observations and (re)analysis datasets. This was attributed to a more active branch of the WCB in a region of the

upper troposphere that pushed the secondary jet closer to the main one. However, at the longer forecast range, this more active branch of the WCB was too strong and led to less realistic behaviour in B85 than the other two runs.

An analysis of the jet stream representation by the other members of the ARPEGE EPS shows that members 1, 2, 4, 5 and 9 behave similarly to member 0, which is the B85 run discussed in the present study (Fig. S3). All these members share the same deep convection scheme as B85 while they differ in the representation of other processes such as turbulence, shallow convection or oceanic flux. In contrast, members 3, 6, 7 and 8 resemble each other and are marked by a larger distance between the two jets than for the other set of members. The deep convection scheme of members 6, 7 and 8 is PCMT, while that of member 3 is the B85 scheme in which the humidity convergence closure was replaced by the CAPE closure. Since PCMT is also based on the CAPE closure, it indicates that the main difference in the jet representation between PCMT and B85 largely comes from the closure.

The humidity flux convergence used by the B85 closure has two components: one coming from the resolved-scale horizontal fluxes and the other from turbulent fluxes. The resolved-scale fluxes are expected to be strong in the presence of synoptic-scale forcing, like in the inflow regions of the warm conveyor belts. Therefore, it is not surprising to get more triggered convection in such cases with strong synoptic-scale forcing, when the closure is based on moisture convergence rather than when it is based on CAPE (PCMT convection scheme). In pure convective situations, when there is no significant synoptic-scale forcing, as for instance during summertime convection over land, CAPE is expected to get higher values, and in that case convection is less sensitive to humidity convergence (Yano et al., 2013).

Further investigations would be needed to better estimate the role of the convective closure type (such as CAPE closures versus moisture convergence closures) on the difference in cyclogenesis, WCB and jet stream dynamics. One way would be to analyse the mirror experiment – that is, PCMT closed with moisture convergence. The other way would be to run the B85 scheme by arbitrarily multiplying the intensity of the humidity convergence by a given factor in the closure, to investigate the sensitivity of WCB to convective intensity in the different regions of the WCB. These sensitivity experiments would necessitate a full analysis and could be an interesting aspect to investigate in future studies.

The sensitivity to initial conditions was analysed to check the robustness of the results. Starting the simulations 12 h earlier leads to similar findings: a too large distance between the two jets and a weaker secondary jet for PCMT compared to B85 (not shown). For hindcasts starting at even earlier dates (e.g. 30 September), members share similar forecast errors in the representation of the jet stream, and it is more

difficult to state which one performs better, but we still observe a more intense secondary jet in B85.

Similar numerical simulations of the subsequent cyclone that followed the Stalactite cyclone 2 d after have been also analysed (IOP7; Schäfler et al., 2018). In that case, generation of negative PV at the tropopause level was found to be more pronounced in B85 than PCMT, leading to a stronger jet in the former than in the latter. The intermediate case based on the B85 scheme but activated with the CAPE closure resembles simulations based on the PCMT scheme (not shown). These results corroborate the case study of the Stalactite cyclone and support the idea that WCB in B85 is on average more active in the upper troposphere in connection with the humidity convergence closure.

While the focus of the present study was on the impact of parameterized deep convection in the WCB outflow region and horizontal structure of the jet stream, a companion paper will follow up to provide a more detailed analysis on the effect of parameterized deep convection in the ascending branch of the WCB and the vertical structure of the jet stream (Wimmer et al., 2021). It will rely on observations made during the second flight of the SAFIRE Falcon on 2 October, which were not shown here.

*Data availability.* Data are available by contacting the corresponding author. ERA5 data are accessible via the climate data store (<https://doi.org/10.24381/cds.bd0915c6>, Hersbach et al., 2018).

*Supplement.* The supplement related to this article is available online at: <https://doi.org/10.5194/wcd-2-1011-2021-supplement>.

*Author contributions.* GR and PA designed the initial study. GR and MW performed the data analysis, MW computed the Lagrangian trajectories and GR made the figures. CL performed the ARPEGE simulations with the help of JMP. PA developed the Lagrangian trajectory algorithm. JD, QC and JP provided the observational datasets. All authors contributed to the scientific discussions.

*Competing interests.* Some authors are members of the editorial board of *Weather and Climate Dynamics*. The peer-review process was guided by an independent editor, and the authors have also no other competing interests to declare.

*Disclaimer.* Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

*Acknowledgements.* The authors thank the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt e.V., DLR) and Andreas Schäfler in particular for granting access to the DLR

Falcon wind lidar data. The study benefited from discussions with the participants of the DIP-NAWDEX (Diabatic Processes in the North Atlantic Waveguide and Downstream impact Experiment) project, which is supported and funded by the Agence Nationale de la Recherche (ANR). It also benefited from discussions with our international NAWDEX partners during annual workshops. We thank Heini Wernli for downloading the ECMWF-IFS operational analysis data and Hanin Binder for sending them to us and for fruitful discussions on warm conveyor belts. Finally, two anonymous referees are acknowledged for their helpful comments during the review process.

*Financial support.* This research has been supported by the Agence Nationale de la Recherche (grant no. ANR-17-CE01-0010-01). The airborne measurements and the SAFIRE Falcon flights received direct funding from IPSL, Météo-France, INSU-LEFE, EUFAR-NEAREX and ESA (EPATAN, contract no. 4000119015/16/NL/CT/gp). The DLR Falcon contribution to NAWDEX was supported by DLR, the European Space Agency (ESA, contract no. 4000114053/15/NL/FF/gp), NRL Monterrey and the European Facility for Airborne Research (EUFAR, project NAWDEX Influence).

*Review statement.* This paper was edited by Juliane Schwendike and reviewed by two anonymous referees.

## References

- Archambault, H. M., Bosart, L. F., Keyser, D., and Cordeira, J. M.: A Climatological Analysis of the Extratropical Flow Response to Recurring Western North Pacific Tropical Cyclones, *Mon. Weather Rev.*, 141, 2325–2346, <https://doi.org/10.1175/MWR-D-12-00257.1>, 2013.
- Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, *Nature*, 525, 47–55, <https://doi.org/10.1038/nature14956>, 2015.
- Baumgart, M., Riemer, M., Wirth, V., and Teubler, F.: Potential Vorticity Dynamics of Forecast Errors: A Quantitative Case Study, *Mon. Weather Rev.*, 146, 1405–1425, <https://doi.org/10.1175/MWR-D-17-0196.1>, 2018.
- Bechtold, P., Bazile, E., Guichard, F., Mascart, P., and Richard, E.: A mass flux convection scheme for regional and global models, *Q. J. Roy. Meteor. Soc.*, 127, 869–886, <https://doi.org/10.1002/qj.49712757309>, 2001.
- Belamari, S.: Report on uncertainty estimates of an optimal bulk formulation for surface turbulent fluxes, (Deliverable No. D.4.1.2), Toulouse, France, Marine environment and security for the European area – Integrated Project (MERSEA IP), 2005.
- Binder, H., Boettcher, M., Joos, H., and Wernli, H.: The Role of Warm Conveyor Belts for the Intensification of Extratropical Cyclones in Northern Hemisphere Winter, *J. Atmos. Sci.*, 73, 3997–4020, <https://doi.org/10.1175/JAS-D-15-0302.1>, 2016.
- Binder, H., Rivière, G., Arbogast, P., Maynard, K., Bossert, P., Joly, B., and Labadie, C.: Dynamics of forecast-error growth along cut-off Sanchez and its consequence for the prediction of a high-impact weather event over southern France, *Q. J. Roy. Meteor. Soc.*, 147, 3263–3285, <https://doi.org/10.1002/qj.4127>, 2021.
- Blanchard, N., Pantillon, F., Chaboureau, J. P., and Delanoë, J.: Organization of convective ascents in a warm conveyor belt, *Weather Clim. Dynam.*, 1, 617–634, <https://doi.org/10.5194/wcd-1-617-2020>, 2020.
- Blanchard, N., Pantillon, F., Chaboureau, J. P., and Delanoë, J.: Mid-level convection in a warm conveyor belt accelerates the jet stream, *Weather Clim. Dynam.*, 2, 37–53, <https://doi.org/10.5194/wcd-2-37-2021>, 2021.
- Booth, J. F., Naud, C. M., and Willison, J.: Evaluation of extratropical cyclone precipitation in the North Atlantic Basin: An analysis of ERA-Interim, WRF, and two CMIP5 models, *J. Climate*, 31, 2345–2360, <https://doi.org/10.1175/JCLI-D-17-0308.1>, 2018.
- Bougeault, P.: A simple parameterization of the large-scale effects of cumulus convection., *Mon. Weather Rev.*, 113, 2105–2121, 1985.
- Bouteloup, Y., Seity, Y., and Bazile, E.: Description of the sedimentation scheme used operationally in all Météo-France NWP models, *Tellus A*, 63, 300–311, <https://doi.org/10.1111/j.1600-0870.2010.00484.x>, 2011.
- Browning, K. A.: Organization of clouds and precipitation in extratropical cyclones, vol. : extratropical cyclones, Erik Palmén memorial volume, chap. 8, 129–153, American Meteorological Society, 1990.
- Chagnon, J., Gray, S. L., and Methven, J.: Diabatic processes modifying potential vorticity in a North Atlantic Cyclone, *Q. J. Roy. Meteor. Soc.*, 139, 1270–1282, 2013.
- Courtier, P., Freydier, C., Geleyn, J., Rabier, F., and Rochas, M.: The ARPEGE project at Météo-France., in: ECMWF Seminar Proceedings, Reading, volume II, 193–231, 1991.
- Crespo, J. A. and Posselt, D. J.: A-Train-Based Case Study of Stratiform–Convective Transition within a Warm Conveyor Belt, *Mon. Weather Rev.*, 144, 2069–2084, <https://doi.org/10.1175/MWR-D-15-0435.1>, 2016.
- Cuxart, J., Bougeault, P., and Redelsperger, J. L.: A Turbulence Scheme Allowing for Mesoscale and Large-Eddy Simulations., *Q. J. Roy. Meteor. Soc.*, 126, 1–30, <https://doi.org/10.1002/qj.49712656202>, 2000.
- Delanoë, J., Protat, A., Jourdan, O., Pelon, J., Paparoni, M., Dupuy, R., Gayet, J.-F., and Jouan, C.: Comparison of Airborne In Situ, Airborne Radar-Lidar, and Spaceborne Radar-Lidar Retrievals of Polar Ice Cloud Properties Sampled during the POLARCAT Campaign, *J. Atmos. Ocean. Tech.*, 30, 57–73, 2013.
- Descamps, L., Labadie, C., Joly, A., Bazile, E., Arbogast, P., and Cébron, P.: PEARP, the Météo-France short-range ensemble prediction system., *Q. J. Roy. Meteor. Soc.*, 141, 1671–1685, 2015.
- Done, J. M., Craig, G. C., Gray, S. L., Clark, P. A., and Gray, M. E. B.: Mesoscale simulations of organized convection: Importance of convective equilibrium, *Q. J. Roy. Meteor. Soc.*, 132, 737–756, 2006.
- Ducrocq, V. and Bougeault, P.: Simulations of an observed squall line with a meso-beta scale hydrostatic model, *Weather Forecast.*, 10, 380–399, 1995.
- Ertel, H.: Ein neuer hydrodynamischer Wirbelsatz, *Meteorol. Z.*, 59, 271–281, 1942.
- Flack, D. L. A., Rivière, G., Musat, I., Roehrig, R., Bony, S., Delanoë, J., Cazenave, Q., and Pelon, J.: Representation by two climate models of the dynamical and diabatic processes

- involved in the development of an explosively deepening cyclone during NAWDEX, *Weather Clim. Dynam.*, 2, 233–253, <https://doi.org/10.5194/wcd-2-233-2021>, 2021.
- Flaounas, E., Lagouvardos, K., Kotroni, V., Claud, C., Delanoë, J., Flamant, C., Madonna, E., and Wernli, H.: Processes leading to heavy precipitation associated with two Mediterranean cyclones observed during the HyMeX SOP1, *Q. J. Roy. Meteor. Soc.*, 142, 275–286, <https://doi.org/10.1002/qj.2618>, 2016.
- Flaounas, E., Kotroni, V., Lagouvardos, K., Gray, S. L., Rysman, J.-F., and Claud, C.: Heavy rainfall in Mediterranean cyclones. Part I: contribution of deep convection and warm conveyor belt, *Clim. Dynam.*, 50, 2935–2949, <https://doi.org/10.1007/s00382-017-3783-x>, 2018.
- Grams, C. M., Wernli, H., Bottcher, M., Campa, J., Corsmeier, U., Jones, S. C., Keller, J. H., Lenz, C.-J., and Wiegand, L.: The key role of diabatic processes in modifying the upper-tropospheric wave guide: a North Atlantic case-study, *Q. J. Roy. Meteor. Soc.*, 137, 2174–2193, <https://doi.org/10.1002/qj.891>, 2011.
- Grams, C. M., Magnusson, L., and Madonna, E.: An atmospheric dynamics perspective on the amplification and propagation of forecast error in numerical weather prediction models: a case study, *Q. J. Roy. Meteor. Soc.*, 144, 2577–2591, <https://doi.org/10.1002/qj.3353>, 2018.
- Gray, S. L., Dunning, C. M., Methven, J., Masato, G., and Chagnon, J. M.: Systematic model forecast error in Rossby wave structure, *Geophys. Res. Lett.*, 41, 2979–2987, 2014.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on pressure levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], <https://doi.org/10.24381/cds.bd0915c6>, 2018.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, *Q. J. Roy. Meteor. Soc.*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Hewson, T., Magnusson, L., Breivik, O., Prates, F., Tsonevsky, I., and Vries, H. J. W. D.: Windstorms in northwest Europe in late 2013, *ECMWF Newsletter*, Spring 2014, 22–28, 2014.
- Jeyaratnam, J., Booth, J. F., Naud, C. M., Luo, Z. J., and Homeyer, C. R.: Upright convection in extratropical cyclones: A survey using ground based radar data over the United States., *Geophys. Res. Lett.*, 47, e2019GL086620, <https://doi.org/10.1029/2019GL086620>, 2020.
- Joos, H. and Wernli, H.: Influence of microphysical processes on the potential vorticity development in a warm conveyor belt: a case-study with the limited-area model COSMO, *Q. J. Roy. Meteor. Soc.*, 138, 407–418, <https://doi.org/10.1002/qj.934>, 2012.
- Kain, J. and Fritsch, J.: Convective Parameterization for Mesoscale Models: The Kain-Fritsch Scheme., vol. The Representation of Cumulus Convection in Numerical Models of Meteorol. Monogr., chap. 16, 165–170, American Meteorological Society, Boston, MA, [https://doi.org/10.1007/978-1-935704-13-3\\_16](https://doi.org/10.1007/978-1-935704-13-3_16), 1993.
- Korfe, N. G. and Colle, B. A.: Evaluation of Cool-Season Extratropical Cyclones in a Multimodel Ensemble for Eastern North America and the Western Atlantic Ocean, *Weather Forecast.*, 33, 109–127, <https://doi.org/10.1175/WAF-D-17-0036.1>, 2018.
- Kuo, H.: On formation and intensification of tropical cyclones through latent heat release by cumulus convection, *J. Atmos. Sci.*, 22, 40–63, 1965.
- Lillo, S. P. and Parsons, D. B.: Investigating the dynamics of error growth in ECMWF medium-range forecast busts, *Q. J. Roy. Meteor. Soc.*, 143, 1211–1226, <https://doi.org/10.1002/qj.2938>, 2017.
- Lopez, P.: Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes, *Q. J. Roy. Meteor. Soc.*, 128, 229–257, <https://doi.org/10.1256/00359000260498879>, 2002.
- Louis, J.: A parametric model of vertical eddy fluxes in the atmosphere, *Bound.-Lay. Meteorol.*, 17, 187–202, <https://doi.org/10.1007/BF00117978>, 1979.
- Maddison, J. W., Gray, S. L., Martinez-Alvarado, O., and Williams, K. D.: Upstream cyclone influence on the predictability of block onsets over the Euro-Atlantic region, *Mon. Weather Rev.*, 147, 1277–1296, <https://doi.org/10.1175/MWR-D-18-0226.1>, 2019.
- Maddison, J. W., Gray, S. L., Martinez-Alvarado, O., and Williams, K. D.: Impact of model upgrades on diabatic processes in extratropical cyclones and downstream forecast evolution, *Q. J. Roy. Meteor. Soc.*, 146, 1322–1350, <https://doi.org/10.1002/qj.3739>, 2020.
- Madonna, E., Wernli, H., Joos, H., and Martius, O.: Warm Conveyor Belts in the ERA-Interim Dataset (1979–2010). Part I: Climatology and Potential Vorticity Evolution, *J. Climate*, 27, 3–26, <https://doi.org/10.1175/JCLI-D-12-00720.1>, 2014.
- Martinez-Alvarado, O. and Plant, R. S.: Parametrized diabatic processes in numerical simulations of an extratropical cyclone, *Q. J. Roy. Meteor. Soc.*, 140, 1742–1755, <https://doi.org/10.1002/qj.2254>, 2014.
- Martinez-Alvarado, O., Baker, L. H., Gray, S. L., Methven, J., and Plant, R. S.: Distinguishing the Cold Conveyor Belt and Sting Jet Airstreams in an Intense Extratropical Cyclone, *Mon. Weather Rev.*, 142, 2571–2595, 2014a.
- Martinez-Alvarado, O., Joos, H., Chagnon, J., Boettcher, M., Gray, S. L., Plant, R. S., Methven, J., and Wernli, H.: The dichotomous structure of the warm conveyor belt, *Q. J. Roy. Meteor. Soc.*, 140, 1809–1824, 2014b.
- Martinez-Alvarado, O., Madonna, E., Gray, S., and Joos, H.: A route to systematic error in forecasts of Rossby waves, *Q. J. Roy. Meteor. Soc.*, 142, 196–210, <https://doi.org/10.1002/qj.2645>, 2016.
- Mazoyer, M., Ricard, D., Rivière, G., Delanoë, J., Arbogast, P., Vié, B., Lac, C., Cazenave, Q., and Pelon, J.: Microphysics impacts on the warm conveyor belt and ridge building of the NAWDEX IOP6 cyclone, *Mon. Weather Rev.*, <https://doi.org/10.1175/MWR-D-21-0061.1>, in press, 2021.
- Neiman, P., Shapiro, M., and Fedor, L.: The life cycle of an extratropical marine cyclone. Part II: Mesoscale structure and diagnostics, *Mon. Weather Rev.*, 121, 2177–2199, 1993.
- Oertel, A., Boettcher, M., Joos, H., Sprenger, M., and Wernli, H.: Convective activity in an extratropical cyclone and its warm

- conveyor belt – a case-study combining observations and a convection-permitting model simulation, *Q. J. Roy. Meteor. Soc.*, 145, 1406–1426, <https://doi.org/10.1002/qj.3500>, 2019.
- Oertel, A., Boettcher, M., Joos, H., Sprenger, M., and Wernli, H.: Potential vorticity structure of embedded convection in a warm conveyor belt and its relevance for large-scale dynamics, *Weather Clim. Dynam.*, 1, 127–153, <https://doi.org/10.5194/wcd-1-127-2020>, 2020.
- Oertel, A., Sprenger, M., Joos, H., Boettcher, M., Konow, H., Hagen, M., and Wernli, H.: Observations and simulation of intense convection embedded in a warm conveyor belt – how ambient vertical wind shear determines the dynamical impact, *Weather Clim. Dynam.*, 2, 89–110, <https://doi.org/10.5194/wcd-2-89-2021>, 2021.
- Pergaud, J., Masson, V., Malardel, S., and Couvreur, F.: A Parameterization of Dry Thermals and Shallow Cumuli for Mesoscale Numerical Weather Prediction, *Bound.-Lay. Meteorol.*, 132, 83, <https://doi.org/10.1007/s10546-009-9388-0>, 2009.
- Piriou, J.-M., Redelsperger, J.-L., Geleyn, J.-F., Lafore, J.-P., and Guichard, F.: An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations, *J. Atmos. Sci.*, 64, 4127–4139, <https://doi.org/10.1175/2007JAS2144.1>, 2007.
- Pomroy, H. and Thorpe, A.: The evolution and dynamical role of reduced upper-tropospheric potential vorticity in intensive observing period one of FASTEX, *Mon. Weather Rev.*, 128, 1817–1834, 2000.
- Ponzano, M., Joly, B., Descamps, L., and Arbogast, P.: Systematic error analysis of heavy-precipitation-event prediction using a 30-year hindcast dataset, *Nat. Hazards Earth Syst. Sci.*, 20, 1369–1389, <https://doi.org/10.5194/nhess-20-1369-2020>, 2020.
- Rasp, S., Selz, T., and Craig, G. C.: Convective and Slantwise Trajectory Ascent in Convection-Permitting Simulations of Midlatitude Cyclones, *Mon. Weather Rev.*, 144, 3961–3976, <https://doi.org/10.1175/MWR-D-16-0112.1>, 2016.
- Rodwell, M. J., Magnusson, L., and co authors: Characteristics of occasional poor medium-range weather forecasts for Europe, *B. Am. Meteorol. Soc.*, 94, 1393–1405, 2013.
- Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J.-F., Voldoire, A., Younous, A., Bazile, E., Belamari, S., Blein, S., Bouniol, D., Bouteloup, Y., Cattiaux, J., Chauvin, F., Chevallier, M., Colin, J., Douville, H., Marquet, P., Michou, M., Nabat, P., Oudar, T., Peyrillé, P., Piriou, J.-M., Salas y Melia, D., Sférian, R., and Sénési, S.: The CNRM global atmosphere model ARPEGE-Climat 6.3: description and evaluation, *J. Adv. Model. Earth Sy.*, 12, e2020MS002075, <https://doi.org/10.1029/2020MS002075>, 2020.
- Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., Methven, J., Rivière, G., Ament, F., Boettcher, M., Bramberger, M., Cazenave, Q., Cotton, R., Crewell, S., Delanoë, J., Dörnbrack, A., Ehrlich, A., Ewald, F., Fix, A., Grams, C. M., Gray, S. L., Grob, H., Groß, S., Hagen, M., Harvey, B., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Lemmerz, C., Lux, O., Magnusson, L., Mayer, B., Mech, M., Moore, R., Pelon, J., Quinting, J., Rahm, S., Rapp, M., Rautenhaus, M., Reitebuch, O., Reynolds, C. A., Sodemann, H., Spengler, T., Vaughan, G., Wendisch, M., Wirth, M., Witschas, B., Wolf, K., and Zinner, T.: The North Atlantic Waveguide and Downstream Impact Experiment., *B. Am. Meteorol. Soc.*, 99, 1607–1637, 2018.
- Schemm, S., Wernli, H., and Papritz, L.: Warm conveyor belt in idealized moist baroclinic wave simulation, *J. Atmos. Sci.*, 70, 627–652, <https://doi.org/10.1175/JAS-D-12-0147.1>, 2013.
- Selz, T. and Craig, G.: Upscale Error Growth in a High-Resolution Simulation of a Summertime Weather Event over Europe, *Mon. Weather Rev.*, 143, 813–827, <https://doi.org/10.1175/MWR-D-14-00140.1>, 2015.
- Simmons, A. and Hollingsworth, A.: Some aspects of the improvement in skill of numerical weather prediction, *Q. J. Roy. Meteor. Soc.*, 128, 647–677, 2002.
- Steinfeld, D. and Pfahl, S.: The role of latent heating in atmospheric blocking dynamics: a global climatology, *Clim. Dynam.*, 53, 6159–6180, <https://doi.org/10.1007/s00382-019-04919-6>, 2019.
- Stoelinga, M. T.: A Potential Vorticity-Based Study of the Role of Diabatic Heating and Friction in a Numerically Simulated Baroclinic Cyclone, *Mon. Weather Rev.*, 124, 849–874, 1996.
- Tribbia, J. J. and Baumhefner, D. P.: The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modeling deficiencies, *Mon. Weather Rev.*, 116, 2276–2288, 1988.
- Weissmann, M., Busen, R., Dörnbrack, A., Rahm, S., , and Reitebuch, O.: Targeted observations with an airborne wind lidar, *J. Atmos. Tech.*, 22, 1706–1719, 2005.
- Wernli, H. and Davies, H. C.: A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications, *Q. J. Roy. Meteor. Soc.*, 123, 467–489, 1997.
- Wimmer, M., Rivière, G., Arbogast, P., Piriou, J.-M., Delanoë, J., Labadie, C., Cazenave, Q., and Pelon, J.: Diabatic processes modulating the vertical structure of the jet stream: sensitivity to deep convection schemes, *Weather Clim. Dynam.*, in preparation, 2021.
- Witschas, B., Rahm, S., Dörnbrack, A., Wagner, J., and Rapp, M.: Airborne wind lidar measurements of vertical and horizontal winds for the investigation of orographically induced gravity waves., *J. Atmos. Ocean. Tech.*, 34, 1371–1386, <https://doi.org/10.1175/JTECH-D-17-0021.1>, 2017.
- Yano, J.-I., Bister, M., Fuchs, Ž., Gerard, L., Phillips, V. T. J., Barkidija, S., and Piriou, J.-M.: Phenomenology of convection-parameterization closure, *Atmos. Chem. Phys.*, 13, 4111–4131, <https://doi.org/10.5194/acp-13-4111-2013>, 2013.
- Zhang, F., Bei, N., Rotunno, R., Snyder, C., and Epifanio, C.: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics, *J. Atmos. Sci.*, 64, 3579–3594, <https://doi.org/10.1175/JAS4028.1>, 2007.

Supplement of Weather Clim. Dynam., 2, 1011–1031, 2021  
<https://doi.org/10.5194/wcd-2-1011-2021-supplement>  
© Author(s) 2021. CC BY 4.0 License.



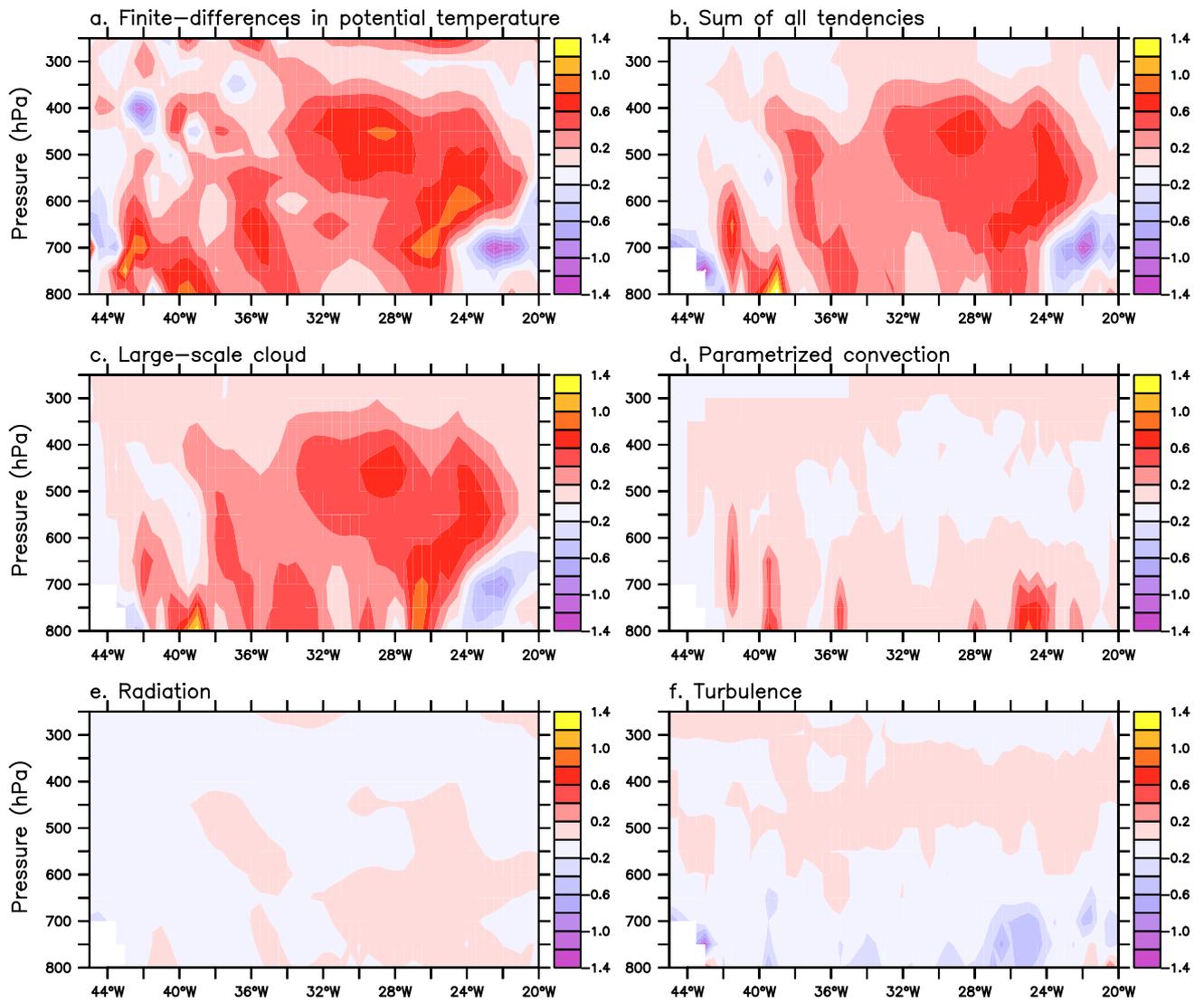
*Supplement of*

## **The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during NAWDEX IOP6**

**Gwendal Rivière et al.**

*Correspondence to:* Gwendal Rivière ([griviere@lmd.ens.fr](mailto:griviere@lmd.ens.fr))

The copyright of individual parts of the supplement might differ from the article licence.



**Figure S 1.** Latitudinally-averaged ( $63^{\circ}$ - $65^{\circ}$ ) heating rates (int:  $0.2 \text{ K h}^{-1}$ ) in B85 run for (a) computed with finite differences in time and space as in the whole paper and (b) resulting from the sum of all diabatic tendencies issued from model outputs. Panel (b) is the sum of (c) large-scale cloud (resolved) sensible and latent heating, (d) parametrized convection sensible and latent heating, (e) radiative heating and (f) turbulence heating.

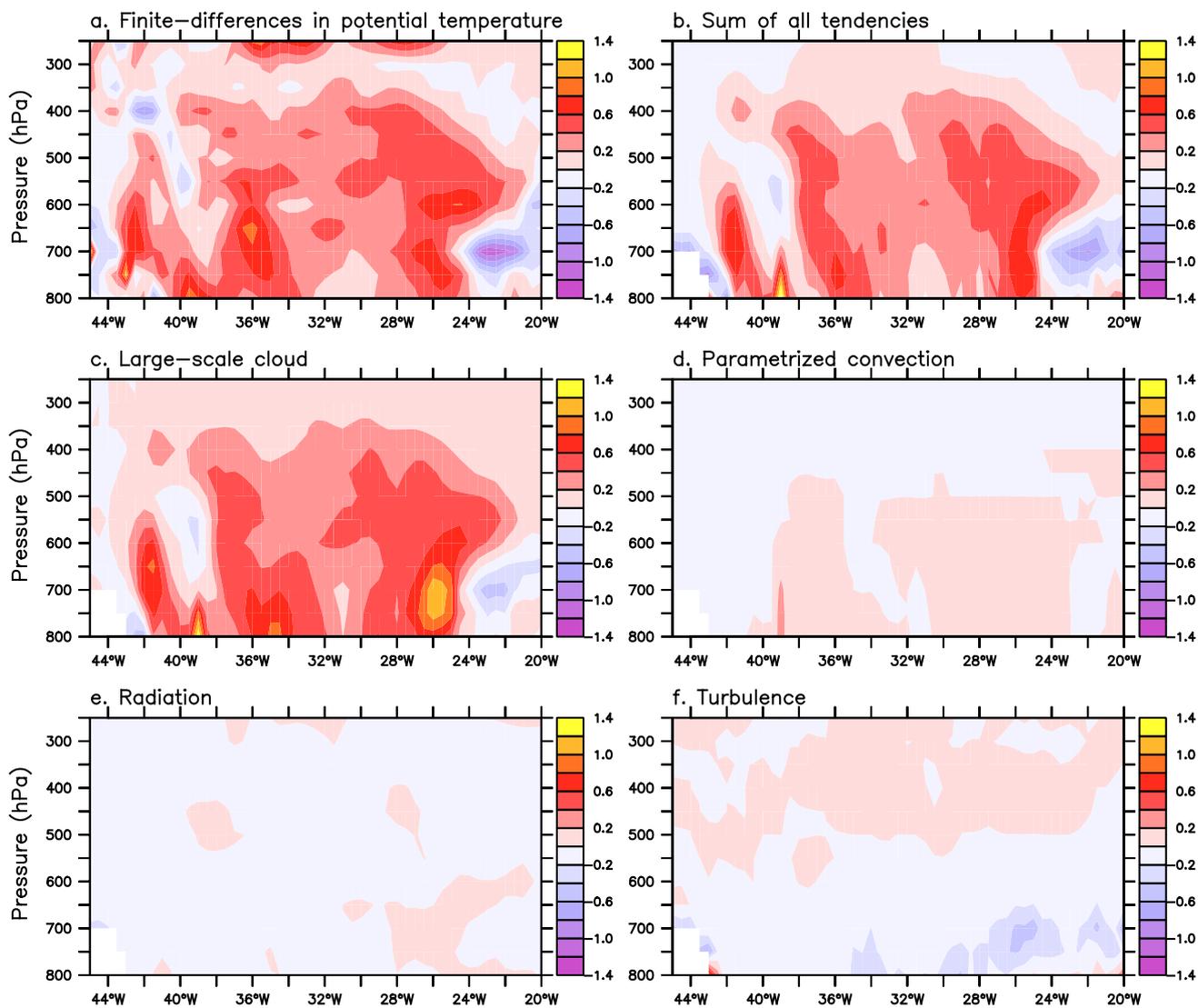
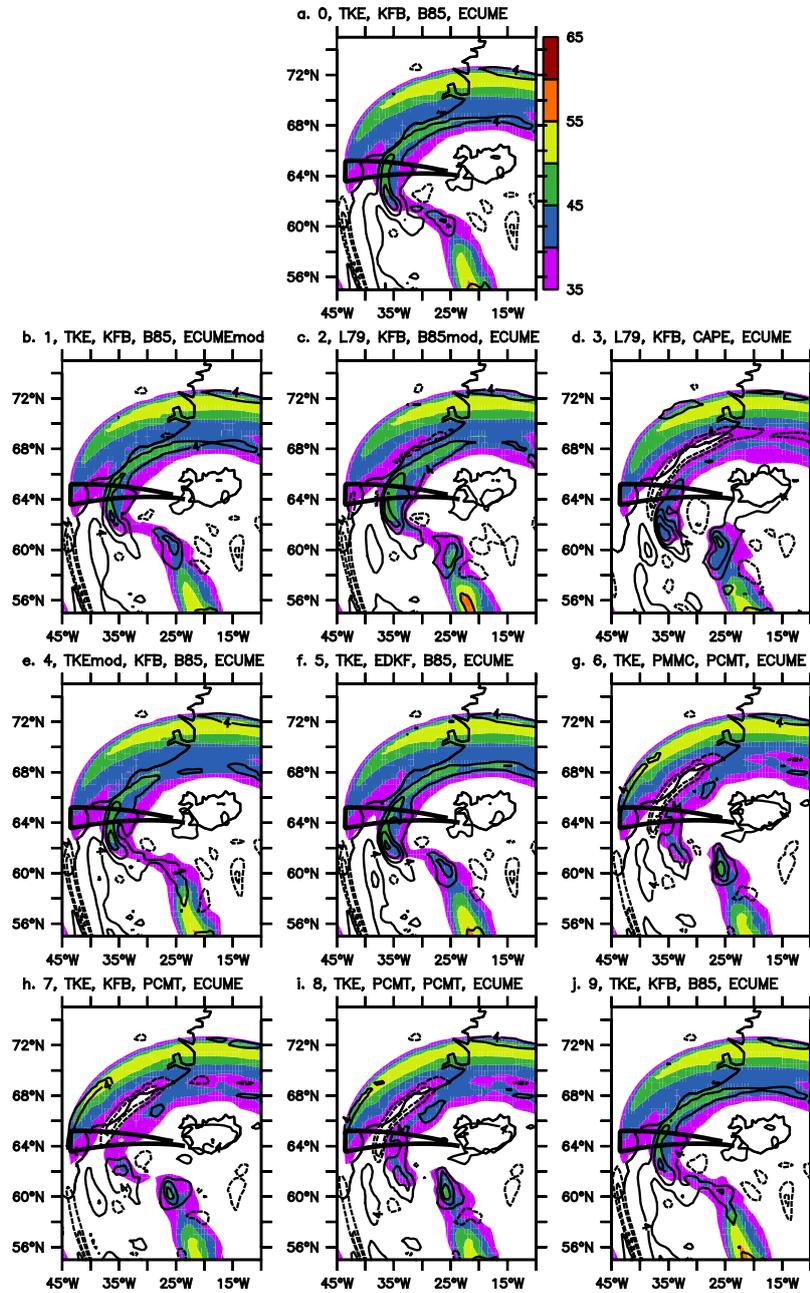


Figure S 2. As in Fig. 2 but for PCMT run.



**Figure S 3.** Wind speed at 300 hPa (shadings; int:  $5 \text{ m s}^{-1}$ ) and difference with respect to ECMWF operational analysis (contours; int:  $4 \text{ m s}^{-1}$ ) at 12 UTC 2 October (i.e corresponding to 24-hours forecast) for 10 members starting with the same initial conditions but differing from four main physics package: turbulence, shallow convection, deep convection and surface fluxes and whose schemes are indicated in that order in the subtitle of each panel. Turbulence schemes can be TKE (Turbulent Kinetic Energy scheme of Cuxart et al. 2000), L79 (Louis, 1979), or TKEmod (slightly modified version of TKE in which horizontal advection is ignored). Shallow convection can be KFB (Kain and Fritsch, 1993; Bechtold et al., 2001), PCMT, PPMC (Pergaud et al. 2009) or EDKF (eddy diffusivity and Kain-Fritsch scheme). Deep convection schemes can be B85, PCMT, CAPE (B85 where the closure is based on the CAPE), B85mod (B85 in which deep convection is only triggered if cloud top exceeds 3000 m). The surface oceanic fluxes are based on Belamari et al. (2005)'s scheme (ECUME) and an alternative in which evaporative fluxes are enhanced (ECUMEmod). [Members 0 and 9 are characterized by the same parametrization set-up, but member 9 differs in the modelling of orographic waves.](#) See Ponzano et al. (2020) for more details on the setup of the different members.

# Diabatic processes modulating the vertical structure of the jet stream above the cold front of an extratropical cyclone: sensitivity to deep convection schemes

Meryl Wimmer<sup>1</sup>, Gwendal Rivière<sup>2</sup>, Philippe Arbogast<sup>3</sup>, Jean-Marcel Piriou<sup>1</sup>, Julien Delanoë<sup>4</sup>, Carole Labadie<sup>1</sup>, Quitterie Cazenave<sup>4</sup>, and Jacques Pelon<sup>4</sup>

<sup>1</sup>CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

<sup>2</sup>Laboratoire de Météorologie Dynamique/IPSL, Ecole Normale Supérieure, PSL Research University, Sorbonne University, École Polytechnique, IP Paris, CNRS, Paris, France

<sup>3</sup>Direction des Opérations pour la prévision, Météo-France, Toulouse, France

<sup>4</sup>LATMOS-IPSL, CNRS/INSU, University of Versailles, Guyancourt, France

**Correspondence:** Meryl Wimmer (meryl.wimmer@umr-cnrm.fr)

**Abstract.** The effect of deep convection parameterization on the jet stream above the cold front of an explosive extratropical cyclone is investigated in the global numerical weather prediction model ARPEGE, operational at Météo-France. Two hindcast simulations differing only in the deep convection scheme used are systematically compared with each other, with (re)-analysis datasets and with NAWDEX airborne observations.

- 5 The deep convection representation has an important effect on the vertical structure of the jet stream above the cold front at one-day lead time. The simulation with the less active scheme shows a deeper jet stream, associated with a stronger potential vorticity (PV) gradient in the jet core in middle troposphere. This is due to a larger deepening of the dynamical tropopause on the cold-air side of the jet and a higher PV destruction on the warm-air side, near 600 hPa. To better understand the origin of this stronger PV gradient, Lagrangian backward trajectories are computed.
- 10 On the cold-air side of the jet, numerous trajectories undergo a rapid ascent from the boundary layer to the mid levels in the simulation with the less active deep convection scheme, whereas they stay at mid levels in the other simulation. This ascent explains the higher PV noted on that side of the jet in the simulation with the less active deep convection scheme. These ascending air masses form mid-level ice clouds that are not observed in the microphysical retrievals from airborne radar-lidar measurements.
- 15 On the warm-air side of the jet, in the warm conveyor belt (WCB) ascending region, the Lagrangian trajectories with the less active deep convection scheme undergo a higher PV destruction due to a stronger heating occurring in the lower and middle troposphere. In contrast, in the simulation with the most active deep convection scheme, both the heating and PV destruction extend further up in the upper troposphere.



## 1 Introduction

20 Midlatitude high-impact weather (HIW) events are usually dynamically forced by near-tropopause disturbances and by specific configurations of the jet stream. Their surface imprints largely depend on the structure and intensity of the jet stream aloft. For instance, the rapid deepening of wind storms depends on the intensity of the jet stream (Wernli et al., 2002; Pinto et al., 2009; Rivièrè et al., 2010) and is favoured by the presence of a jet streak (Uccellini, 1990; Fink et al., 2009). As a second example, heavy precipitation and flood events are often forced by an elongated trough along the jet stream or a cut-off that just separated  
25 from the jet stream following wave breaking (Massacand et al., 1998; Martius et al., 2008; Nuissier et al., 2011; Grams et al., 2014). Since the near-tropopause disturbance triggering the HIW event is often part of a Rossby wave train, the skills of NWP forecasts to accurately predict a HIW event depends on their ability to represent the troughs/ridges propagating along the jet stream during the days prior to the event (Parsons et al., 2017; Wirth et al., 2018). Consequently, there is a growing body of literature identifying systematic NWP biases along these downstream propagating near-tropopause wavelike disturbances  
30 (Rodwell et al., 2013; Gray et al., 2014) and investigating dynamics of forecast errors along the midlatitude waveguide (Davies and Didone, 2013; Baumgart et al., 2018).

Looking at different NWP models, Gray et al. (2014) found systematic forecast errors in the jet representation increasing with forecast lead times. In particular, the potential vorticity (PV) gradient gets smoother on the poleward flank of ridges and Rossby wave amplitudes get smaller, the two being closely related (Harvey et al., 2016). Another consequence of the too smooth  
35 PV gradient is the slowdown of phase speed with forecast lead time (Harvey et al., 2018). This is in that particular context, that the international field campaign NAWDEX (North Atlantic Waveguide Downstream and impact EXperiment) occurred in September-October 2016 (Schäfler et al., 2018). NAWDEX objective was to investigate the diabatic origin of forecast errors in the ascending part of extratropical cyclones along the so-called warm conveyor belts (WCBs), to analyse their downstream propagation along the waveguide and how they may affect the predictability of HIW events. Using NAWDEX observations  
40 as a reference, Schäfler et al. (2020) showed underestimation of vertical wind shear in the vicinity of the tropopause in very short-term forecasts and analysed this could affect Rossby wave propagation.

Regions of strongest forecast errors and systematic analysis of forecast busts suggest that forecast errors originate from diabatic processes (Rodwell et al., 2013; Gray et al., 2014). Because of the PV invertibility properties and its conservation under adiabatic and frictionless processes, the PV perspective offers a classical and useful framework to investigate the influence of  
45 diabatic processes on the atmospheric flow. The PV tracer technique that decomposes the PV rate of change into different model processes has been widely used during the last decade, mainly to study the near-tropopause PV anomalies associated with the jet stream (Chagnon et al., 2013; Martinez-Alvarado et al., 2014; Saffin et al., 2017; Spreitzer et al., 2019; Harvey et al., 2020) but also to study low-level PV anomalies associated with a surface cyclone (Crezee et al., 2017). The former cited studies found that near-tropopause PV is strongly affected by diabatic processes, mainly by latent heating, turbulence and  
50 longwave radiation, and these processes maintain the strong PV gradient there. Saffin et al. (2017) showed that the decrease in tropopause sharpness with forecast lead time, originally diagnosed by Gray et al. (2014), is mainly due to the advection scheme and is only partially compensated by the increase in tropopause sharpness due to nonconservative processes.



The PV and potential temperature ( $\theta$ ) Lagrangian framework can be used to explain atmospheric circulation differences between simulations performed with distinct models (Martinez-Alvarado et al., 2014) or between sensitivity numerical experiments made with the same model but using different parameterization schemes (Martinez-Alvarado and Plant, 2014; Joos and Forbes, 2016; Mazoyer et al., 2021; Rivière et al., 2021). Joos and Forbes (2016) used this approach to compare two simulations of the ECMWF-IFS global model with distinct cloud microphysics schemes. They found slight PV differences in the WCB outflow region amplifying on the downstream side of the ridge at 24 h-72 h lead time. Following a similar approach, Mazoyer et al. (2021) compared two simulations of a regional convection permitting model performed with two distinct cloud microphysics schemes. At 24 h-48 h lead time, the PV and wind speed differences were already well marked and anomalies span an extended band along the western edge of the ridge due to different heating rates within the WCB. The PV- $\theta$  framework has been also used to analyze WCB differences and impact on the tropopause with different deep convection parameterization schemes (Martinez-Alvarado and Plant, 2014; Rivière et al., 2021) but the amplitude of the impact on the upper-level circulation varies from case to case. Martinez-Alvarado and Plant (2014) found rather modest differences in the tropopause location after 24 h for a moderate cyclone between reduced and intense parameterized convection while Rivière et al. (2021) found important differences with a jet stream shift of a few hundred kilometers after 24 h for an explosive cyclone.

Recent NAWDEX-related studies have emphasized the importance of embedded convection within WCBs by comparing satellite observations and convective-permitting simulations to airborne radar measurements gathered during NAWDEX (Oertel et al., 2019, 2020, 2021; Blanchard et al., 2020, 2021). Oertel et al. (2020) and Blanchard et al. (2021) showed that the heating associated with embedded convection generates dipolar PV anomalies near the tropopause that reinforce the PV gradient and hence the jet stream. The ability of convectively created PV dipole to reinforce the jet depends on the region where convection occurs and on the vertical wind shear (Chagnon and Gray, 2009; Harvey et al., 2020; Oertel et al., 2021).

Following the same approach as in a companion paper (Rivière et al., 2021, hereafter RW21), the present study investigates the effect of parameterized deep convection on WCB and jet stream. RW21 compared three simulations of the Météo-France global model ARPEGE: two simulations were performed with two distinct deep convection parameterization schemes developed within ARPEGE, the one described in Bougeault (1985, thereafter B85) and the Prognostic Condensates Microphysics and Transport scheme of Piriou et al. (2007, thereafter PCMT). The third simulation was performed without any deep convection parameterization. Without deep convection parameterization the heating ahead of the cold front is organized in distinct cells of high values with a few degrees extent in longitude and latitude because convective instability is released at the resolved scales. In contrast, in presence of parameterized deep convection, the heating is much smoother because convective instability is released at subgrid scales. The consequence is that WCB ascents are more sustained with parameterized deep convection while they are more abrupt without. This regulating effect of deep convection parameterization on WCB was already noticed by Martinez-Alvarado and Plant (2014). However, this does not mean that the impact on the tropopause is stronger without deep convection parameterization. RW21 showed that the run in which deep convection is more active (B85) is also the one for which the heating extends further upward above the warm front of the extratropical cyclone and has a stronger PV destruction at upper levels. This leads to a distinct location of the jet stream compared to the other two runs, the one with PCMT deep convection scheme and the other without any active scheme. In RW21, the focus was on the WCB outflow region above the

bent-back warm front and the horizontal structure of the jet stream. In the present study, our aim is to analyse differences between the same simulations within the WCB ascending region above the cold front and to highlight differences in the vertical structure of the jet stream.

The extratropical cyclone hereafter studied, called the "Stalactite cyclone" (1-4 October 2016), is of particular interest in a number of respects. It was formed off the Newfoundland coast and intensified over the North Atlantic with a deepening rate of 24 hPa in 24 h (Flack et al., 2021) as a classical bomb event (Sanders and Gyakum, 1980). It has been intensively observed during NAWDEX (Schäfler et al., 2018) by three flights: two flights of the French Falcon 20 from the Service des Avions Français Instrumentés pour la Recherche en Environnement (SAFIRE) and one flight from the Deutsches Zentrum für Luft- und Raumfahrt (DLR) Dassault Falcon. Its development was accompanied by a burst in latent heating (Steinfeld et al., 2020) and the ridge building aloft led to the onset stage of the "Thor" block (Maddison et al., 2019) that last until the end of NAWDEX in mid-October 2016.

The paper is organized as follows. Section 2 presents the data and methods. It includes the description of the model simulations and the main characteristics of the two deep convection schemes B85 and PCMT. It also provides information on various (re)-analysis datasets and on airborne observations made during the flight of the SAFIRE Falcon aircraft on 2 October over the ascending WCB region of the Stalactite cyclone. Finally, section 2 details the computation of PV- $\theta$  Lagrangian budgets. Section 3 shows differences in the jet stream representation between the different simulations and (re)-analysis datasets. Section 4 provides an explanation for these differences in terms of the PV- $\theta$  framework. Section 5 compares model simulations to airborne observations to highlight the more realistic forecasts in the different regions. Finally section 6 is dedicated to the concluding remarks and discussion.

## 2 Data and method

### 2.1 Model and simulations set-up

As in the companion paper RW21, the study relies on simulations of the operational Météo-France global model, ARPEGE (Action de Recherche Petite Echelle Grande Echelle; Courtier et al., 1991) and in particular on different members of its Ensemble Prediction System associated, called PEARP (Prévision d'Ensemble ARPEGE; Descamps et al., 2015).

For all PEARP members, the vertical resolution has 90 levels while the horizontal grid corresponding to T798 resolution, is stretched by a factor 2.4 and centred on France. Consequently, this resolution is about 10 km on France, 15 km on the zone of interest of this study and 60 km on the antipode of France. The time step of the model is 7.5 minutes. Models outputs are available with a temporal resolution of 15 minutes and an horizontal resolution of  $0.5^\circ$ , while the vertical resolution is 50 hPa.

While operational PEARP members include perturbations in both model physics and initial conditions, the present study is based on PEARP reforecast dataset corresponding to ten members that have the same initial conditions (ARPEGE operational 4D-Var analysis at 12 UTC 1 October 2016) and only differ in their physics as in Ponzano et al. (2020), Binder et al. (2021) and RW21.

120 Two members amongst the ten ones are hereafter more deeply investigated. They only differ in their deep convection scheme: one with the Bougeault (1985) scheme (B85), the other with Prognostic Condensates Microphysics and Transport (PCMT; Piriou et al., 2007). They are referred as the REF and member 7 in Ponzano et al. (2020) and as B85 and PCMT in RW21. In addition, a third simulation, called NoConv, has no deep convection scheme activated. For more details on these three simulations, particularly concerning physical parameterization, reader is referred to RW21.

## 125 2.2 Differences between the two deep convection schemes

The two simulations studied in the present paper use two distinct parameterization schemes of deep convection, both being based on the mass-flux approach. Detrainment in the environment, precipitation and downdrafts phenomena are modeled in both B85 and PCMT. But in contrast to B85, the PCMT scheme is able to estimate the prognostic mixing ratios of the different hydrometeors inside the mass flux (updraft). It includes the same four hydrometeors (liquid cloud water, pristine ice, rain and snow) involved in the large-scale cloud microphysical scheme of Lopez (2002), and the same microphysical phenomena, such as accretion, autoconversion, riming.

130 The two schemes are hereafter activated with two distinct closures. Bougeault (1985)'s scheme is activated with the convergence of total moisture fluxes (including both resolved and turbulent moisture fluxes) integrated from the surface to the considered level and when the atmosphere presents an unstable profile. In contrast, the Prognostic Condensates Microphysics and Transport scheme (PCMT; Piriou et al., 2007) considers a CAPE closure.

## 135 2.3 Airborne observations and (re)-analyses

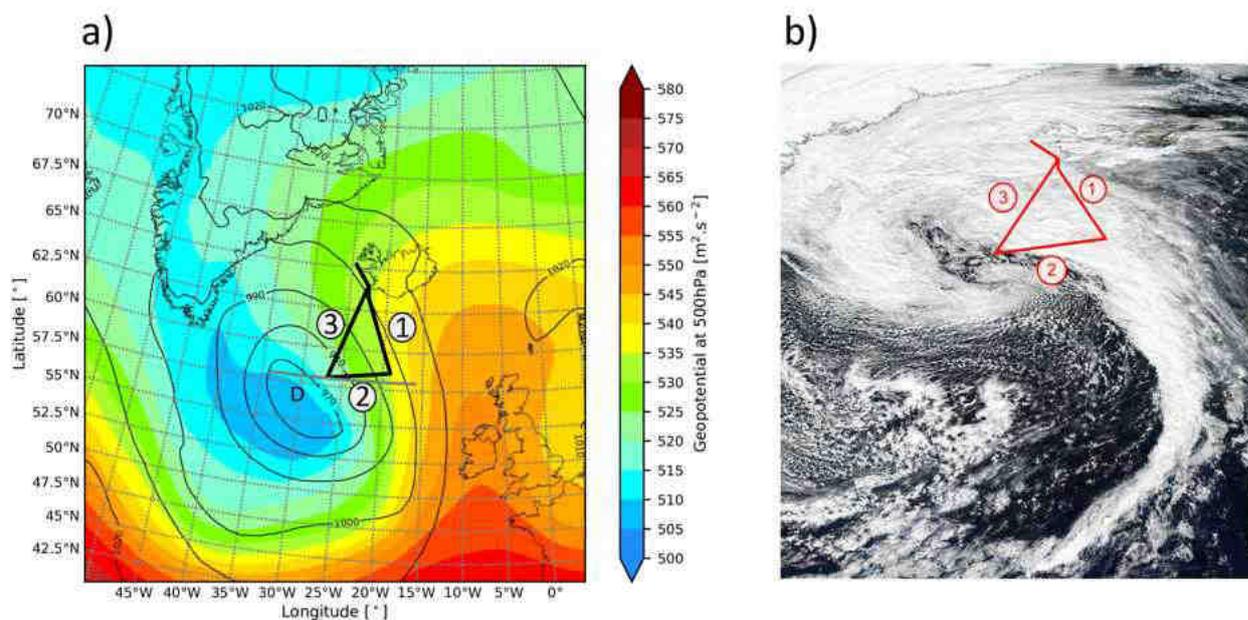
To better compare the effects of the deep convection scheme and to better estimate their realism, two types of references are used: observations from the NAWDEX IOP6, as well as operational analyses and reanalysis.

### 2.3.1 Observations from NAWDEX IOP6

140 The flight of the French Falcon 20 of SAFIRE, studied in the present study, occurred between 1301 UTC and 1616 UTC on 2 October during the NAWDEX field campaign (Schäfler et al., 2018). Figures 1a and b show the position of the flight according to the Stalactite cyclone. The aircraft took off at Keflavik, went south, realized a clockwise loop triangular in shape to the northeast of the Stalactite cyclone.

145 Figure 1b gives an insight of the cloudy region sampled by the flight. The major part of the flight occurred in the cloudy region ahead of the cold front close to the cyclone center, which likely corresponds to the ascending part of the WCB. But the clear-sky zone appearing near the westernmost vertex of the triangle suggests that the flight crossed the cold front from east to west near 58° N.

150 During this flight, different measurements have been made. In-situ sensors on board the SAFIRE Falcon 20 measured pressure, wind and temperature at the flight level near 300 hPa. In addition, the RALI (RADar-LIdar) platform developed at LATMOS and DT-INSU, was on board the aircraft. This platform includes a multi-beam 95 GHz Doppler cloud radar



**Figure 1.** Visualisation of the Stalactite cyclone during its mature stage, the 2nd of October 2016: a) Geopotential at 500 hPa (shading), sea level pressure (black thin contour) at 12 UTC, Falcon flight (black bold line) and vertical cross section at 58° N (grey bold line) b) visible picture from VIIRS of the Suomi NPP satellite (NASA Worldview) with the Falcon flight in red.

(RASTA) (Radar Airborne System; Delanoe et al., 2013) and a Doppler high-spectral-resolution lidar (LNG; Bruneau et al., 2015). RASTA measures both reflectivity and Doppler velocity along three non colinear directions thanks to three downward antennas (nadir, backward and transverse). This configuration allows one to retrieve the 3D wind field in the vertical below the Falcon with a range resolution of 60 m and every 0.75 s leading to an horizontal resolution of about 300 m at the speed of  
155 the aircraft. The lidar operates at 532 nm and 1064 nm in backscatter mode only but measures Doppler velocity, polarisation and the backscattered light from molecules and particles separately at 355 nm. It gives information about optical parameters of aerosol and thin clouds together with along-sight wind below the aircraft at 15 m and 5 s range and time resolution respectively. Additional wind measurements were made by in-situ sensors at the aircraft altitude and dropsondes.

To better compare observations with model outputs, in-situ and RALI measurements have been averaged over intervals of  
160 180 s to be close to the model grid spacing. Indeed, as the Falcon 20 has a mean speed of  $200 \text{ m s}^{-1}$ , it travels a distance of 36 km during 180 s, which is close to the  $0.5^\circ$  horizontal grid spacing of the model outputs.

### 2.3.2 Operational analysis and reanalysis

Operational analyses of the ARPEGE and Integrated Forecasting System (IFS) models are considered. The same vertical resolution of 50 hPa and horizontal resolution of  $0.5^\circ$  than the ARPEGE simulations outputs are used. ERA5 reanalyses

165 dataset (Hersbach et al., 2020) is also considered. However, as the original horizontal is about  $0.25^\circ$ , one grid point every two  
grid points is kept in order to get the same  $0.5^\circ$  resolution than the other simulations and analyses.

## 2.4 Lagrangian warm conveyor belt trajectories

### 2.4.1 Initialization in the warm sector

The same forward trajectories as those shown in RW21 are used. They are initialized at 12 UTC on 1 October in the warm  
170 sector of the extratropical cyclone and last 48 hours. To select WCB trajectories, a criterion of an ascent exceeding 300 hPa  
within 1 day during the period between 12 UTC on 1 October and 12 UTC on 3 October is applied.

### 2.4.2 Initialization along the flight

To better characterize properties of the WCB air masses of the Stalactite Cyclone crossing the flight F7, another set of tra-  
jectories has been computed with the same trajectory algorithm as in RW21. It consists of 24 h backward and 24 h forward  
175 trajectories starting from the flight legs over the whole vertical. For each flight leg, the trajectories are seeded on a vertical  
regular grid spacing of 12.5 hPa from 975 hPa to 200 hPa and a horizontal grid spacing of about  $0.3^\circ$  in longitude and latitude.  
As the flight lasted more than 3 h, trajectories from each flight leg must be seeded at a different time. The time of seeding is the  
time when the aircraft is in the middle of each leg. Hence, the first leg trajectories are seeded at 26.25 h forecast range, namely  
1415 UTC on 2 October, the second leg trajectories at 27 h forecast range corresponding to 1500 UTC and finally the third leg  
180 trajectories at 27.75 h forecast range, so 1545 UTC.

Overall, 5292 trajectories lasting 48 h have been computed with 5292 seeding points along the flight path (63 in the vertical  
 $\times$  84 in the horizontal). To prevent trajectories to cross the surface, pressure is limited at 975 hPa. Trajectories with a minimum  
ascending rate of 300 hPa in 24 h are considered as belonging to the WCB. It leads to 1870 WCB trajectories for B85 and 1972  
WCB trajectories for PCMT.

## 185 2.5 Heating and PV tendencies

As in RW21, the heating  $\dot{\theta}$  is computed into two different manners. The first method uses instantaneous temperature tendency  
datasets associated with the different diabatic processes parameterized in the model (large-scale cloud microphysics, convec-  
tion, radiation, turbulence). These temperature tendencies are first provided on the ARPEGE stretched grid and model levels  
before being interpolated on the  $0.5^\circ \times 0.5^\circ$  horizontal grid and pressure levels. The second method computes the heating  
190 using centred finite-difference schemes applied to the potential temperature at the resolution of the model outputs chosen for  
this study:  $0.5^\circ \times 0.5^\circ$  in the horizontal, 50 hPa in the vertical and 15 minutes in time. Since the Lagrangian trajectories are  
also computed on the latter grid, the variations in  $\theta$  along trajectories is very close to the integrated heating  $\dot{\theta}$  obtained with the  
second method (not shown). The first method only roughly approximates the variations in  $\theta$  along trajectories for mainly two  
reasons: firstly, the dynamical core of ARPEGE does not strictly conserve  $\theta$  because of numerical diffusion in the advection  
195 scheme and secondly the various interpolation steps in the offline trajectory algorithm to get the temperature tendency terms

on the model output grid generate uncertainties. Both methods are hereafter used and information on the choice of the method is provided in the captions of the figures: the second method has the advantage to nearly close the heating budget while the first method has the advantage to provide a decomposition of the heating into various diabatic processes.

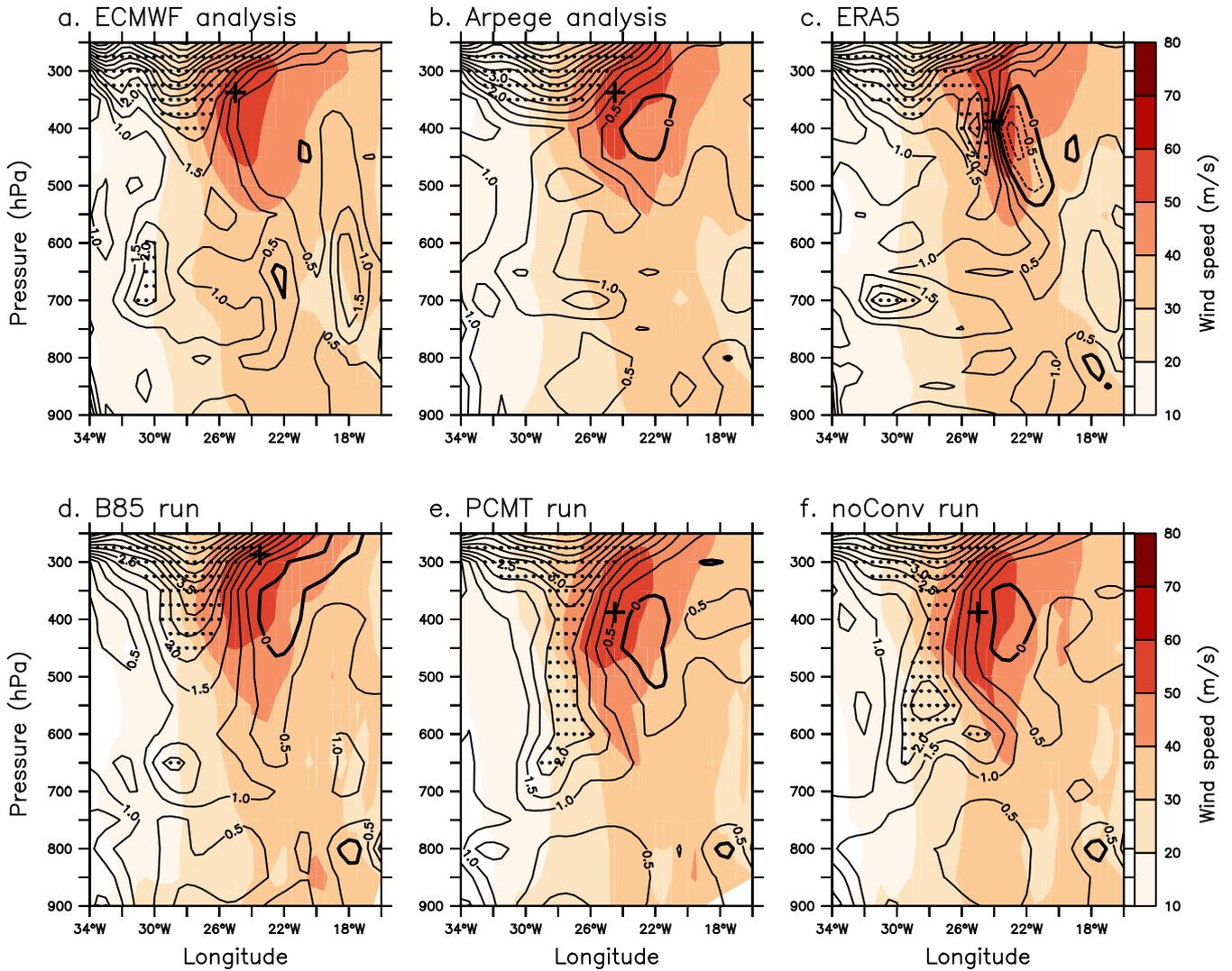
As the PV tendency depends on spatial variations of the heating and frictional terms, their computation is made by applying  
200 finite-difference schemes to the heating and frictional terms. The frictional terms in the zonal and meridional momentum equations are made available in the stretched/rotated Gaussian reduced model grid and at model levels. They are first interpolated on the  $0.5^\circ \times 0.5^\circ$  horizontal and 50 hPa vertical grids before applying the finite-difference schemes to them. The heating is computed following the second method described above as it leads to a much more accurate approximation of the total PV tendency than the first method.

### 205 3 Impact of deep convection representation on the Stalactite Cyclone dynamic

The vertical structure of the jet stream at  $58^\circ$  N is shown for the three simulations and different (re)-analysis datasets in Fig. 2. This latitude roughly corresponds to the southern leg of the flight (see grey line in Fig. 1a) and to the western edge of the upper-level ridge. In the three references (i.e. the two analyses and ERA5), the maximum wind speed is located near  $24^\circ$  W between 300 hPa and 400 hPa at the interface between stratospheric and tropospheric air and varies between  $55 \text{ m s}^{-1}$  and  
210  $60 \text{ m s}^{-1}$ . The height of maximum wind speed fluctuates from case to case: it is about 340 hPa in the two analyses while it is near 390 hPa in ERA5. Such differences in wind speed are accompanied by similar differences in PV: the dynamic tropopause (2 PVU isoline) descends until about 400 hPa in the two analyses while it descends further down to 500 hPa in ERA5. On the tropospheric side, east of the jet maximum, stronger negative values (more than  $-0.5$  PVU at  $23^\circ$  W) appear in ERA5 than in the two analyses. The jet is narrower and slightly deeper in ERA5 than in the analyses, consistent with stronger PV gradient  
215 between 400 hPa and 500 hPa in the former than in the latter datasets.

The three ARPEGE forecasts with distinct deep convection representation simulate the speed and position of the jet stream reasonably well in comparison with the three references: the jet is centred at  $24^\circ$  W with also a maximum between 50 and  $60 \text{ m s}^{-1}$  in both simulations. However, the vertical structure differs from one run to another. PCMT and NoConv simulate a deeper jet stream with a center located at 390 hPa with wind speed values up to  $40 \text{ m s}^{-1}$  reaching 650 hPa. In contrast, in B85, such  
220 high wind speed values do not go further down than 575 hPa. The deeper jet stream in PCMT and NoConv is associated with a lower tropopause to the west, as it goes down to 650 hPa in both PCMT and NoConv and only to 400 hPa in B85. It is also associated with negative PV values going further down to the east: the area of negative PV values extends from 350 hPa to 525 hPa in PCMT and from 350 hPa to 475 hPa in NoConv while it goes from 450 hPa to even higher than 250 hPa in B85.

As the largest differences between the two forecasts appear in the middle of the troposphere, horizontal cross sections of  
225 wind speed and potential vorticity are shown at 600 hPa in Fig. 3. In all simulations (references as well as forecasts), wind speed values higher than  $40 \text{ m s}^{-1}$ , corresponding to the lower part of the jet stream, are located above and along the cold front of the Stalactite cyclone which is noticeable by the change in curvature of the sea level pressure contours along an axis oriented from southeast to northwest.



**Figure 2.** Vertical cross section at  $58^\circ$  N (grey line in Fig. 1 of the zonal wind (shadings) and Potential Vorticity (black contours with hatched areas for values superior to 2PVU) at 15 UTC, 2 October 2016 for a) IFS analysis, b) ARPEGE analysis, c) ERA reanalysis, d) simulation with B85, e) simulation with PCMT and f) simulation without deep convection parameterization. The thick crosses represent the location of wind speed maxima.

Among the six datasets, PCMT and Noconv are the two simulations exhibiting the most intense jet with wind speed values beyond  $40 \text{ m s}^{-1}$  from  $18^\circ$  W to  $26^\circ$  W and between  $55^\circ$  N and  $58^\circ$  N, located east of a band of PV values exceeding 2 PVU. The other datasets do not exhibit PV values as large as in PCMT and NoConv at 600 hPa in that region close to the westernmost vertex of the triangular-shaped flight. NoConv is the run showing the highest PV values to the west of that primary jet. More to the east, near the easternmost vertex, a less well-defined secondary jet with values close to  $30 \text{ m s}^{-1}$  appears in some datasets.

**Table 1.** Root mean square of the difference with ERA5 reanalysis of the IFS and ARPEGE analyses and the three forecasts in PV and wind speed at 600 hPa over the domain shown in Fig. 3.

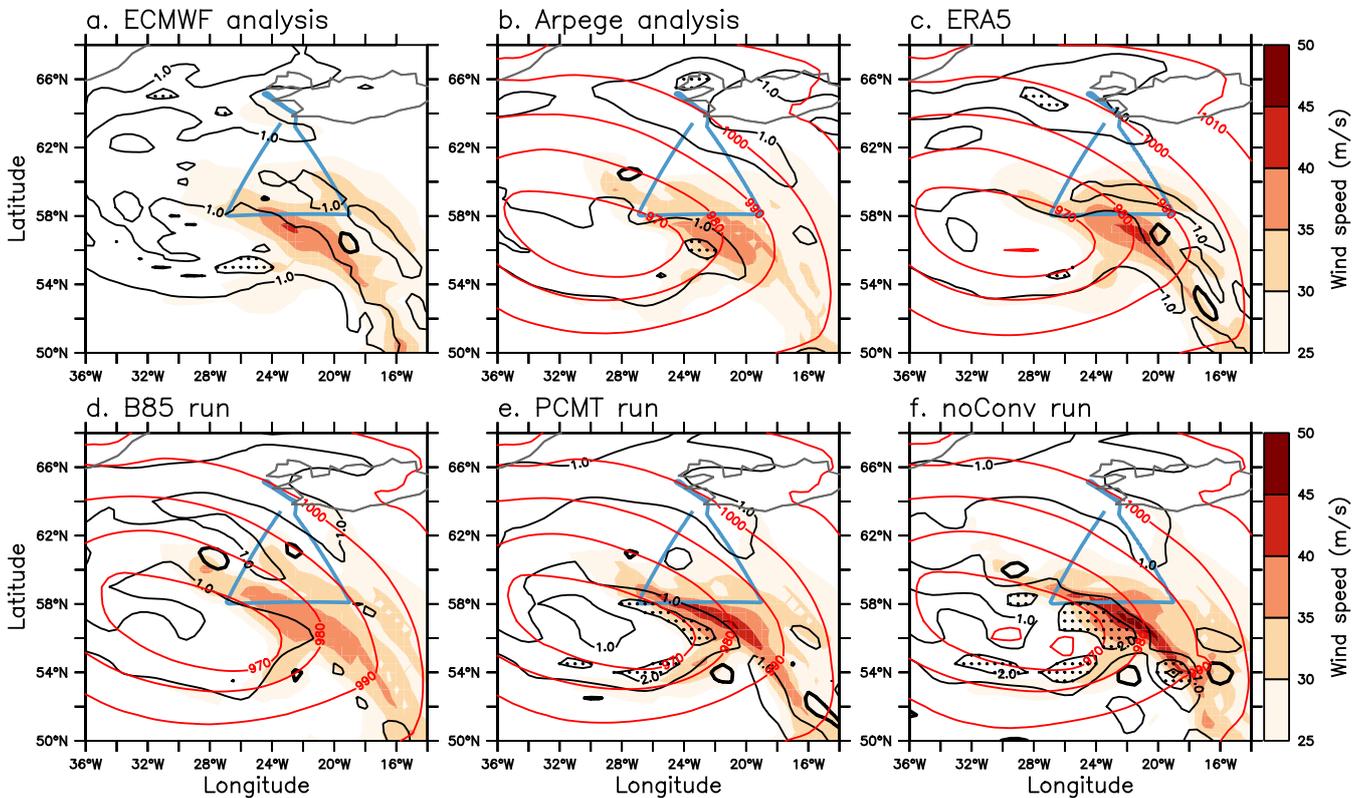
	ECMWF-IFS analysis	ARPEGE analysis	B85	PCMT	NoConv
wind speed ( $\text{m.s}^{-1}$ )	1.85	2.37	2.95	2.99	3.20
PV (PVU)	0.31	0.33	0.34	0.38	0.47

It has different shape and extension in the different datasets. For instance, this secondary jet in B85 extends further to the northwest than in PCMT or in ECMWF-IFS and ERA5. The ARPEGE analysis brings similarities with B85, which is not surprising as they both use the same model and the same deep convection scheme (B85). ERA5 and ECMWF-IFS analysis are similar too as they both use the IFS model.

This is confirmed by computing root-mean-square (RMS) of the differences of each dataset with ERA5: the lowest RMS is with ECMWF-IFS, the second lower is ARPEGE analysis, then B85, PCMT and NoConv in ascending order (table 1). Since the two analyses have lower RMS difference with respect to ERA5 than the forecasts, it gives confidence in assessing the performance of the three forecasts as the three references are closer to each other than to the forecasts. Among the three forecasts, NoConv is the one leading to the highest RMS and in that sense it is less skillful than the other two. PCMT and B85 have rather similar RMS values with those of PCMT being slightly higher. In terms of PV, it is clearly the band of high PV values to the west of the primary jet that increases the RMS error of PCMT with respect to ERA5, and more importantly that of NoConv. But PCMT does not perform so differently from B85 because B85 has other significant differences with ERA5 along the third leg of the flight and related to the too far northwestward extended secondary jet as confirmed in section 5.

The other PEARP members, differing only in their physical parameterization of deep convection, turbulence, shallow convection and surface oceanic fluxes, as described in RW21, are also compared in an additional sensitivity study (Figs. S1 and S2). Members 1, 2, 4, 5, and 9 predict a maximum of the jet stream near 300 hPa as the B85 simulation (member 0) (Fig. S1). In contrast, members 3, 6, 8 are marked by a jet maximum located, lower, between 350 hPa and 400 hPa and, in that sense, behave as in the PCMT simulation (member 7). Furthermore, the first group of members does not show any stratospheric air with a PV superior to 2 PVU in middle troposphere (600 hPa) whereas the second group shows systematic areas with PV higher than 2 PVU (Fig. S2).

Notice that the common point of each group is the type of closure used in the deep convection schemes. Members 0, 1, 2, 4, 5 and 9 share the same deep convection scheme, namely B85 with a moisture convergence closure. In the other group, all members share a CAPE closure with members 6, 7 and 8 using the PCMT scheme and member 3 using the B85 scheme. By analyzing hindcasts of heavy precipitation events, Ponzano et al. (2020) also emphasized a clustering of the 10 PEARP members into 2 groups but their partition was dependent on the deep convection used (B85 vs PCMT). In the present case as well as in RW21, the separation more clearly emerges according to the convection-parameterization closure (moisture vs CAPE).



**Figure 3.** Wind (shadings) and Potential Vorticity (black contours with hatched areas for values superior to 2 PVU) at 600 hPa with sea level pressure (red contours) at 15 UTC, 2 October 2016 for a) IFS analysis, b) ARPEGE analysis, c) ERA reanalysis, d) simulation with B85, e) simulation with PCMT and f) simulation without any deep convection representation. The Flight F7 of the SAFIRE Falcon appears in blue line.

To conclude, PCMT has a deeper jet stream than B85 over the cold front in association with higher positive PV values to the west and smaller negative PV values to the east in the mid troposphere around 600 hPa. At this stage, it is rather difficult to determine which deep convection scheme is more realistic because the height and vertical structure of the jet in the three references (two analyses and ERA5) are usually in between the two runs. Since the two simulations with activated deep convection scheme behave in opposite ways and exhibit a large difference in the vertical structure of the jet stream, it is worth investigating the reasons of this difference as done in next sections, in particular by analysing the properties of the WCB.

#### 4 Heating and Potential Vorticity differences in the Warm Conveyor Belt

Since the main difference in the jet stream highlighted in the previous section occurs during and near the SAFIRE flight on 2 October afternoon, the air streams crossing the flight are studied in the present section. They are modeled by 48 h Lagrangian trajectories centred on the flight (see subsection 2.4.2 for their definition). Trajectories satisfying the WCB criterion (300 hPa

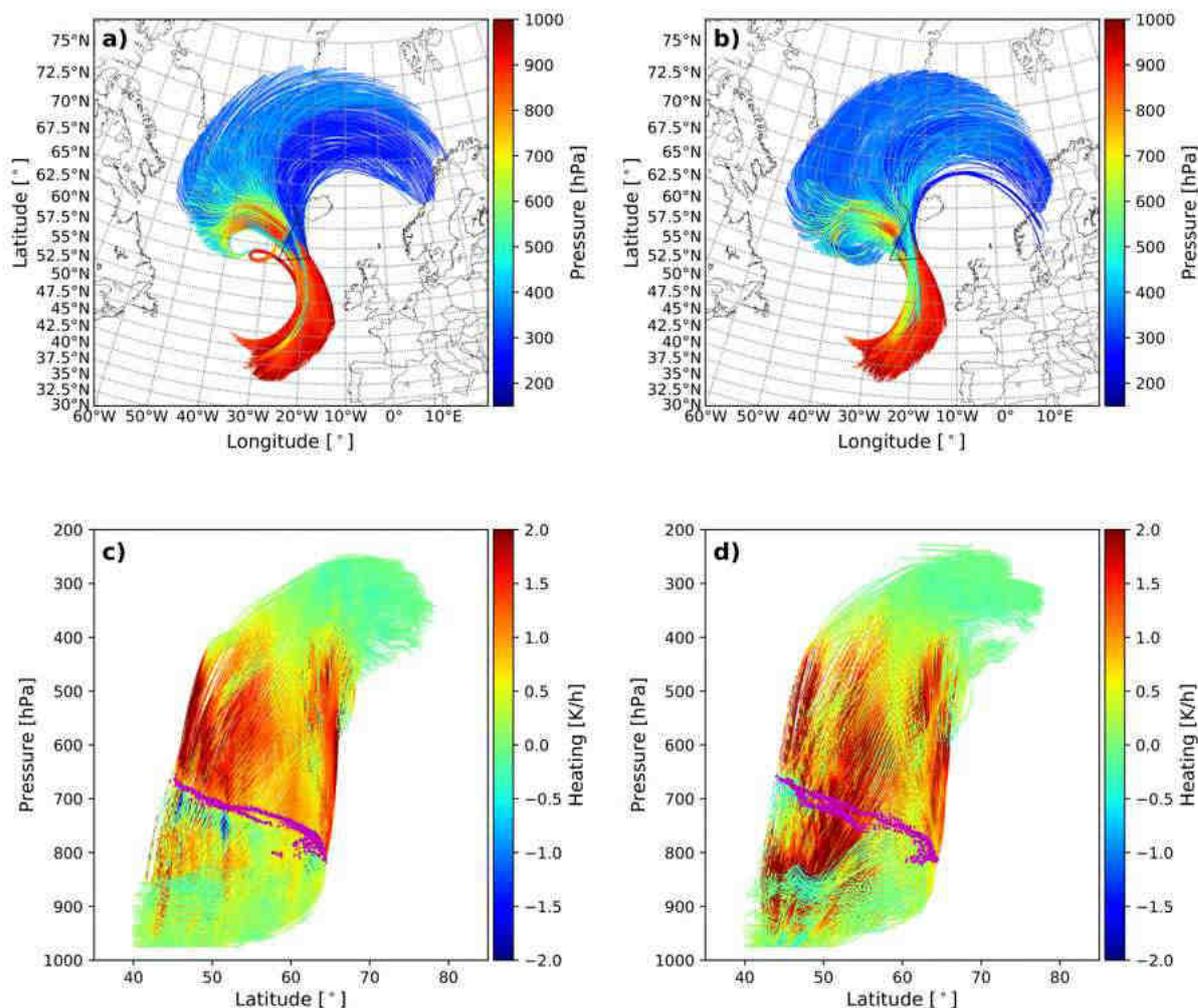
ascent in 24 h) are represented in Fig. 4a for B85 and Fig. 4b for PCMT. All WCB trajectories have a poleward direction along the cold front and then may turn cyclonically or anticyclonically as in the classical picture of the WCB (Schemm et al., 2013; Martinez-Alvarado et al., 2014). Those having a cyclonic curvature move towards the cyclone center, while the majority of them have a pronounced anticyclonic curvature and orient towards Scandinavia at the end. However, there is no  
275 separation between the cyclonic and anticyclonic branches of the WCB but rather a continuum exists between cyclonically- and anticyclonically-curved trajectories. Also, many trajectories first curve cyclonically before curving anticyclonically.

More cyclonic trajectories are present near the minimum of pressure in PCMT than B85. Computation of the curvature during the last 3 h of the trajectories confirms it: 76% of them have an anticyclonic curvature during that last 3 h and so 24% are cyclonic in B85 while the proportions are quite different in PCMT with 72% of anticyclonic trajectories and 28%  
280 of cyclonic ones. Finally, some slight differences are noticed in the anticyclonic branch with more trajectories moving further south of Scandinavia in PCMT than in B85.

The pressure along these trajectories represented in color, shows that trajectories ascend as they move northward. For both simulations, some ascending regions are located far south of the flight (e.g., near 25° W; 50° N), some others further north (40° W; 62.5° N) but most of them are in the vicinity of the flight. Hence, the flight clearly occurs in the main ascending region  
285 of the WCB.

Latitude-pressure plots of the trajectories colored by the heating rate  $\dot{\theta}$  are shown in Fig. 4c for B85 and Fig. 4d for PCMT. In both simulations, the maximum heating undergone by the trajectories is about 2 K h<sup>-1</sup> and logically occurs in the ascending part of the trajectories. Some cooling stage occurs in the lower troposphere before the trajectories reach the freezing point 0° C (purple dots) due to evaporative/melting processes. In B85, strong cooling is obvious just below the freezing point between  
290 45° N and 50° N that is likely due to snow melting. Another slight cooling area appears in both simulations: in the upper troposphere due to long wave radiation. Some large differences also exist between the two simulations. A large part of the PCMT trajectories present a strong heating of 2 K h<sup>-1</sup> in the positive temperature area, while this phenomenon is much more reduced in B85, questioning the different behaviour between these two convection schemes in the liquid phase. This more intense heating occurring sooner along the trajectories and at lower altitude in PCMT has some implications in terms of PV  
295 tendencies as shown later.

Figures 5a and 5b show the wind speed at 15 UTC on 2 October 2016 along the last half of the flight for B85 and PCMT respectively, together with the difference in PV between PCMT and B85 (PCMT-B85) in each panel. Additionally, the positions of the WCB trajectories initialized along the legs of the flight are represented by the grey circles. Only the second and third leg of the flight F7 are considered as very few trajectories satisfying the WCB criterion cross the first leg. Note that the abscissa  
300 is not the time but a trajectory index, which is the number of trajectory seeds, on the horizontal, along the flight. Between 300 hPa and 500 hPa, dipolar PV anomalies appear in the vicinity of the wind speed maxima with positive values to the east and anticyclonic to the west. It means that the PV gradient is stronger in B85 in the upper troposphere and is logically associated with stronger wind speed maxima at those levels. Between 500 hPa and 700 hPa, opposite-sign PV anomalies also appear on both sides of the jet, but here the positive values are to the west and negative ones to the east of the jet. It means  
305 there is a stronger PV gradient, which is associated with a stronger wind speed maxima in the mid troposphere and thus a



**Figure 4.** Upper panels: longitude-latitude plots of the pressure (shadings) along the warm conveyor belt trajectories crossing the flight F7 for a) B85 and b) PCMT. Lower panels: vertical point of view (Latitude-Pressure plots) heating (shadings; 1st method of computation) along the warm conveyor belt trajectories crossing F7 for c) B85 and d) PCMT. Intersection with the iso-0° C is represented by the purple dots.

deeper jet for PCMT. At the same levels but further away from the jet, the PV anomaly changes sign again (see trajectory index higher than 70). The opposite-sign PV anomalies centred at trajectory index 70 reinforce the PV gradient in B85 with respect to PCMT and leads to the presence of a secondary jet at those levels for the former run, which has been already discussed when commenting Fig. 3d. Therefore, between 500 hPa and 700 hPa, the PV difference exhibits a tripolar anomaly in each leg, which is systematic of all sections crossing the cold front from 12° W-50° N to 28° W-62° N (Figs.5c-d).

310



The positions of the WCB trajectories are located between 900 hPa and 300 hPa for both simulations, but they are more numerous in the upper layer between 400 hPa and 300 hPa in B85, particularly in leg 2 (compare Figs. 5a and b). The more numerous upper-level WCB trajectories in B85 are in a positive PV difference, which means a lower PV in B85. As a strong heating occurs between 800 hPa and 400 hPa followed by a rapid decrease above 400 hPa (see Figs. 4c and d), it indicates that the vertical gradient of the heating is negative at 400 hPa and above. Hence, the WCB trajectories at the time of the flight undergo a negative PV tendency above 400 hPa and the more numerous WCB trajectories in B85 at those levels induce more PV destruction than those in PCMT. It provides an explanation for the smaller PV east of the jet, stronger PV gradient, and stronger wind speed in B85 at pressure lower than 400 hPa. Between 500 hPa and 700 hPa along the cold front, most of the WCB trajectories initialized in the warm sector are located in the negative PV anomaly (Figs. 5c,d). Very few of them, initialized along the flight, are within the positive PV anomaly to the east of the main jet (see trajectory index between 45 and 60 in Figs. 5a-b).

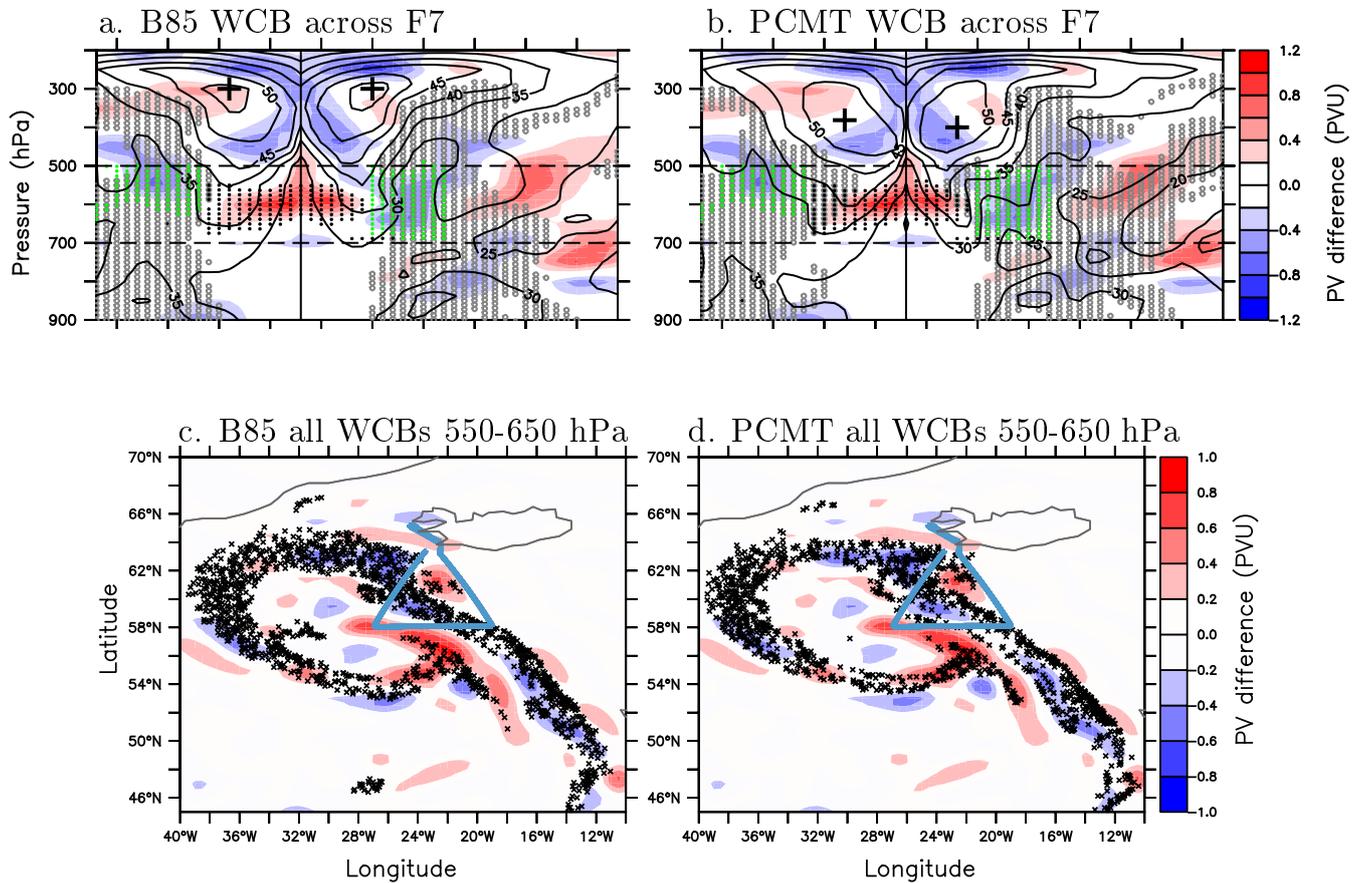
To better explain the deeper jet stream in PCMT, the next section analyses the origin of the positive PV anomaly (black dots in Figs. 5a-b) and negative PV anomaly (green dots) located on both sides of the jet stream.

#### 4.1 Positive PV anomaly on the cold-air side of the jet

To investigate the origin of the positive PV anomaly, the 24 h backward trajectories, whose seeding point is represented by black dots on Figs. 5a and b, are considered. Figure 6a shows the time evolution of PV averaged over all trajectories reaching the positive PV anomaly along legs 2 and 3 (red and orange respectively for B85; blue and cyan respectively for PCMT). The first striking result is that the averaged PV is almost the same at the initial time (16 UTC on 1 October) between the two simulations while they differ by about 0.3 PVU at the final time (16 UTC on 2 October). It clearly shows that the higher PV in PCMT than B85 at the time of the flight (1500 UTC and 1545 UTC on 2 October for leg 2 and 3 resp.) is solely due to diabatic PV modification along trajectories. More precisely it is between 00 UTC and 04 UTC on 2 October that the B85 and PCMT curves move away from each other (compare the red and blue curves or the orange and cyan ones). After 04 UTC, the PV difference B85 and PCMT is maintained whatever the leg, even though the separation distance between the curves may undergo some fluctuations, as at 11 UTC for leg 2.

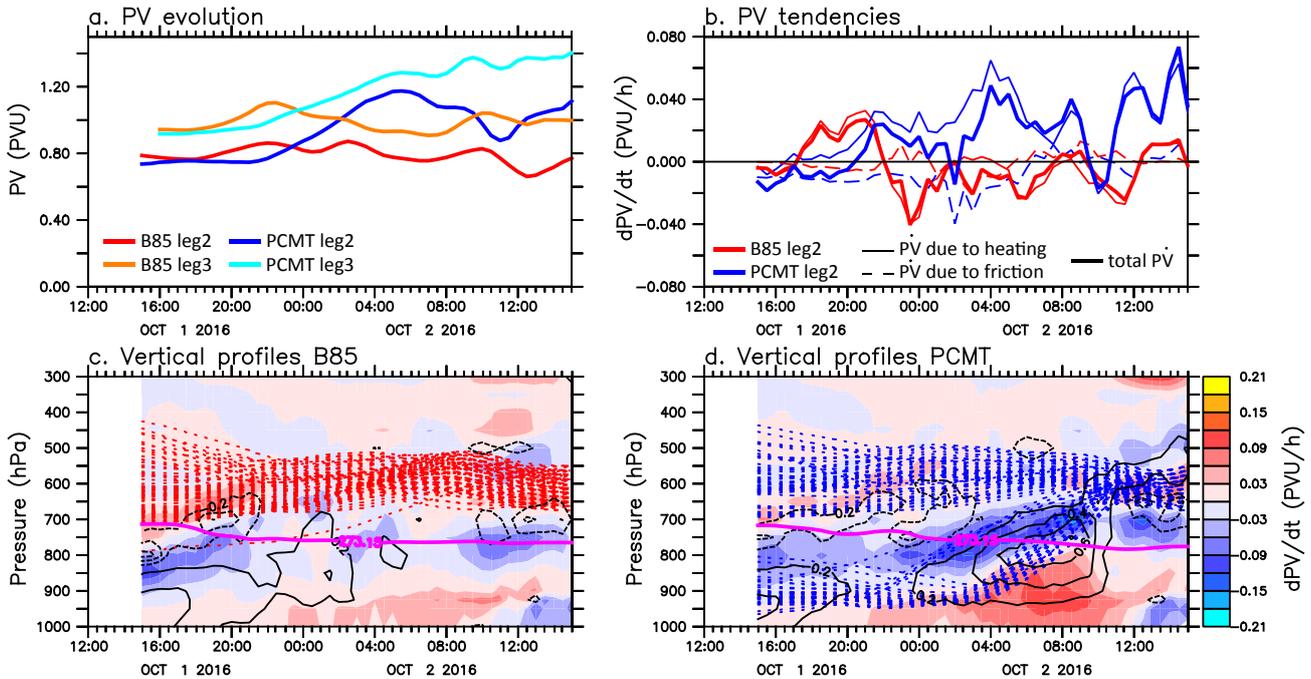
The time evolution of the PV tendencies computed by summing the tendencies due to heating and friction is shown in Fig. 6b for leg2 (bold solid lines). The good correspondence between the sign of that sum (Fig. 6b) and the slope of the PV evolution (Fig. 6a) shows that the budget is correctly done. For B85, between 16 UTC and 22 UTC, PV increases and the sum of all terms is positive while after 22 UTC, PV tends to slightly decrease consistent with near zero or negative PV tendency. Only at later times, after 12 UTC on 2 October, PV increases a little bit. For PCMT, the tendency is first near zero and the PV does not change much prior to 20 UTC on 1 October. But after that short period, the PV tendency is quasi-systematically positive and higher than that of B85 except near 03 UTC on 2 October or 10 UTC on 2 October.

The PV tendencies decomposition into heating (thin line) and friction (thin dashed line) parts clearly shows that the PV fluctuations are dominated by the heating. In particular, positive PV tendencies are solely due to the heating. Figures 6c and d represent the vertical profiles of PV tendencies and heating averaged over all grid points where there is a WCB trajectory from



**Figure 5.** Upper panels: Vertical cross section of the difference (PCMT-B85) in PV (shadings) along the second and third leg of flight F7 at 15 UTC, 2 October 2016. The wind speed (black contours) and intersection of WCB trajectories with an ascending of 300 hPa in 24 h with F7 (grey circles) are shown for a) B85 and b) PCMT. In panels (a) and (b), the thick crosses represent the wind speed maxima in legs 2 and 3 for B85 and PCMT respectively. Trajectories where PV anomalies are positive and negative between 500 hPa and 700 hPa are in black and green dots respectively. Lower panels: PV difference at 600 hPa (shadings) and WCB trajectories, initialized in the warm sector, positions between 550-650 hPa (crosses) at 15 UTC 2 October 2016 for c) B85 and d) PCMT. The flight is in blue line.

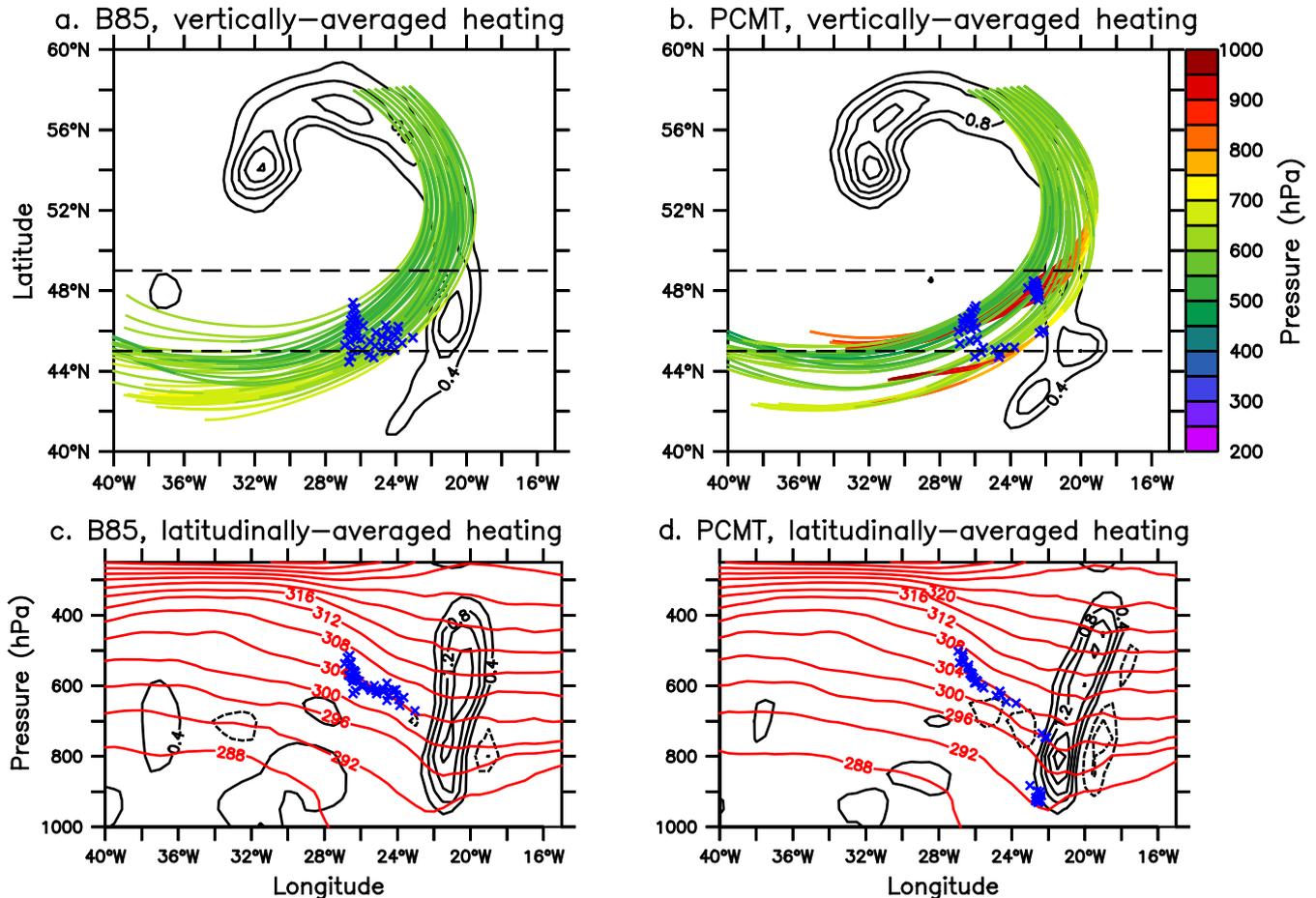
345 leg2 for B85 and PCMT respectively. One large difference between the two figures concerns the pressure distribution along the trajectories (red or blue dashed curves). For B85, they are clustered in a single group always transported in the middle of the troposphere during 24 hours before reaching the positive anomaly of leg 2 (see red dashed curves in Fig. 6c). Depending on how they are positioned relative to the heating/cooling regions, they may undergo PV increase or decrease. For instance, the slight PV increase between 16 UTC and 22 UTC on 1 October or after 12 UTC on 2 October is explained by the fact that  
 350 the trajectories are above a cooling region. In contrast, for PCMT, the trajectories are clustered into two well separated groups, one half transported in the middle of the troposphere as in B85 but the other half starting in the boundary layer and suddenly rising at 04 UTC 2 October (see blue dashed curves in Fig. 6d). Slightly before and during the beginning of the ascents of the



**Figure 6.** a) Time evolution of PV for WCB trajectories reaching the positive PV anomaly in leg 2 (red for B85 and blue for PCMT) and leg 3 (orange for B85 and light blue for PCMT); b) Time evolution of PV tendencies due to heating (thin solid lines), friction (thin dashed lines) and total (bold solid lines) for B85 (red) and PCMT (blue); Vertical profiles, according to time, of heating (contours, dashed and solid for negative and positive values resp.; second method of computation) and PV tendency due to heating (shadings) along WCB trajectories reaching the positive PV anomaly for c) B85 (red dashed curves) and d) PCMT (blue dashed curves) respectively. The iso-0°C is the purple line.

latter trajectories (i.e. from 00 UTC to 04 UTC on 2 October), PV rapidly increases because the trajectories are below a region of strong heating. This is precisely during this period, that the two averaged PV in B85 and PCMT get away from each other (Fig. 6a) and the PV tendencies are largely different (Fig. 6b). When the same trajectories go above the heating, they undergo a short period of PV decrease between 08 UTC and 10 UTC but they rapidly go above a cooling region after 12 UTC when they catch up the first group of trajectories and their PV increases once again.

Figure 7 helps to further visualize the position of the trajectories with respect to the heating when the two PV start to get away from each other (Fig. 6a). For that purpose, the chosen time is 03 UTC on 2 October. As previously observed in Fig. 6c, WCB trajectories from the positive anomaly do not present any ascent in B85 (Fig. 7a). They are all transported west of the heating area behind the cold front around 650 hPa and the iso-304 K (Figs. 7a,c). In PCMT, the group of trajectories around 650 hPa have more or less the same position relative to the main heating area as the one in B85 (Figs. 7b,d). However they are above a well marked cooling region near 24-26° W and 700 hPa and undergo a slightly larger PV increase than those for B85.



**Figure 7.** Pressure (shadings) along WCB trajectories reaching the positive PV anomaly in leg 2 for a) B85 and b) PCMT respectively, heating (black contours ; units :  $0.4 \text{ K h}^{-1}$ ) and WCB position (blue crosses) at 03 UTC, 2 October; Vertical cross sections of heating averaged between  $45^\circ \text{ N}$  and  $49^\circ \text{ N}$  (black contours; second method of computation), potential temperature (red contours) and WCB positions (blue crosses) at 03 UTC 2 October, for c) B85 and d) PCMT.

The heating budgets in Figs. S3 and S4 for B85 and PCMT respectively show that the cooling region is due to both radiation and turbulence at the top of mid-level convective clouds in the cold sector of the cyclone. This increase in PV for trajectories moving in the mid troposphere over the cold sector can be seen at different time intervals in both runs (between 15 UTC on 1 October and 00 UTC on 2 October or after 12 UTC on 2 October in Figs. 6c,d) but the PV increase appears to be more important in average in PCMT (compare the reddish colors between 700 hPa and 600 hPa in Figs. 6c and d). The other group of trajectories found in PCMT is located at  $22^\circ \text{ W}$  and 900 hPa. They are located in the lowest and most western part of the main heating area ahead of the cold front, which is dominated by large-scale cloud heating (Fig. S4), and will rapidly ascent during the following hours (Fig. 6d).



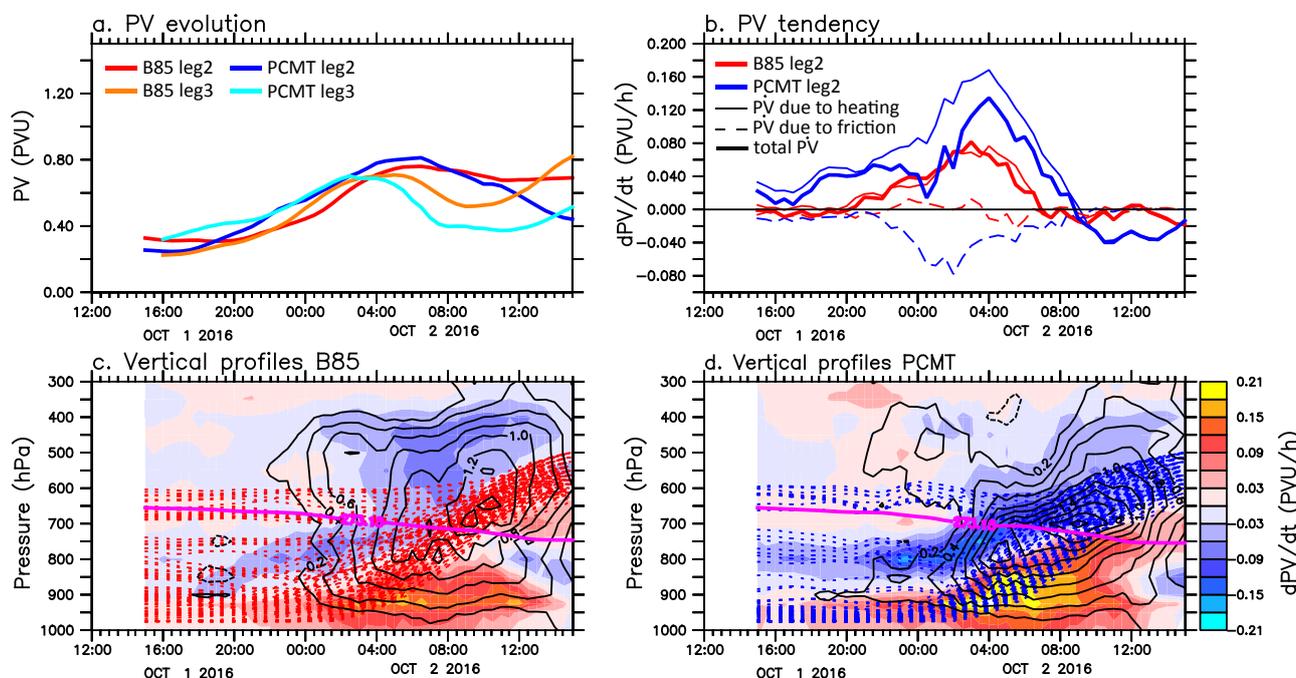
To conclude, the higher PV obtained in PCMT than B85 on the cold-air side of the jet stream is mainly due to diabatic processes occurring between 00 UTC and 04 UTC 2 October during which half of the PCMT trajectories rapidly ascend and undergo a PV increase below a strong heating area. This heating is mainly due to large-scale latent heating and, to a lesser extent, due to convection heating (Figs. S3, S4). Some of these trajectories exceeding 300 hPa ascent in 24 h satisfy the WCB criterion and are thus identified with grey circles in Fig. 5b near trajectory indexes 45-50. An additional factor to explain the difference in PV between the two runs concerns the group of trajectories evolving in the middle of the troposphere: in PCMT, they are more often subject to PV increase in presence of cooling areas below them at the top of convective mid-level clouds in the cold sector of the cyclone. In B85, this happens less regularly along similar trajectories.

#### 380 4.2 Negative PV anomaly on the warm-air side of the jet

The same approach based on backward trajectory is adopted to better understand the origin of the negative PV anomaly to the east of the jet, which is mainly embedded in the WCB region (Figs. 5a-b). At the time of the flight (1500 UTC and 1545 UTC on 2 October for leg 2 and 3 resp.), the averaged PV is about 0.3 PVU lower in PCMT than in B85 whatever the leg (Fig. 8a). For leg 3, the difference in PV rapidly increases from 04 UTC to 08 UTC on 2 October while for leg 2, it increases from 12 UTC to 16 UTC on 2 October. Even though the timing is different, for both legs the PV difference is small at the initial time (16 UTC on 1 October) and the PV difference has a diabatic origin occurring during the last 12 hours before reaching the flight legs.

Let us now focus on the negative PV anomaly of leg 2. For both PCMT and B85, PV first increases and then decreases. However, the PV variations are larger in PCMT than B85. This difference is due to the heating contribution in the PV tendency budget and not due to the friction contribution (Fig. 8b). The heating being stronger in PCMT than B85 (maxima are about 2 K h<sup>-1</sup> for PCMT and 1.4 K h<sup>-1</sup> for B85), its gradient is stronger leading to higher amplitude PV tendency in the former case (Figs. 8c-d). The heating is also more vertically stacked in the lower troposphere in PCMT in such a way that the trajectories are already all advected in a region above the heating maximum in PCMT after 10 UTC 2 October and undergo a negative PV tendency (Figs. 8b,d). During this later period, the B85 trajectories are advected near a region of maximum heating and have thus near zero PV tendency (Figs. 8b,c).

Figure 9 provides a 3D picture of the position of the trajectories with respect to the heating at 12 UTC on 2 October, i.e. the time when the PV difference between PCMT and B85 starts to increase. At the initial time, a majority of trajectories were located in the boundary layer of the warm sector of the cyclone whatever the simulation (Figs. 9a-b). At the time of the plot, all trajectories lie within the strong heating region ahead of the cold front. However, their positions relative to the heating vertical gradient largely differ between PCMT and B85. This is mainly due to two distinct features in the heating fields. The B85 heating extends more to the upper troposphere and is more vertical whereas PCMT heating is more confined in the middle troposphere and is marked by an eastward tilt with height. The latter tilt is rather systematic whatever the time chosen (see Figs. 7d and 9d). These two distinct features place the trajectories in a region of negative heating gradient and thus negative PV tendency in PCMT while the heating gradient is weak for the B85 trajectories. Thus, contrarily to B85, the PCMT trajectories are already passing over the main heating area ahead of the cold front and already loose PV before reaching the flight leg.



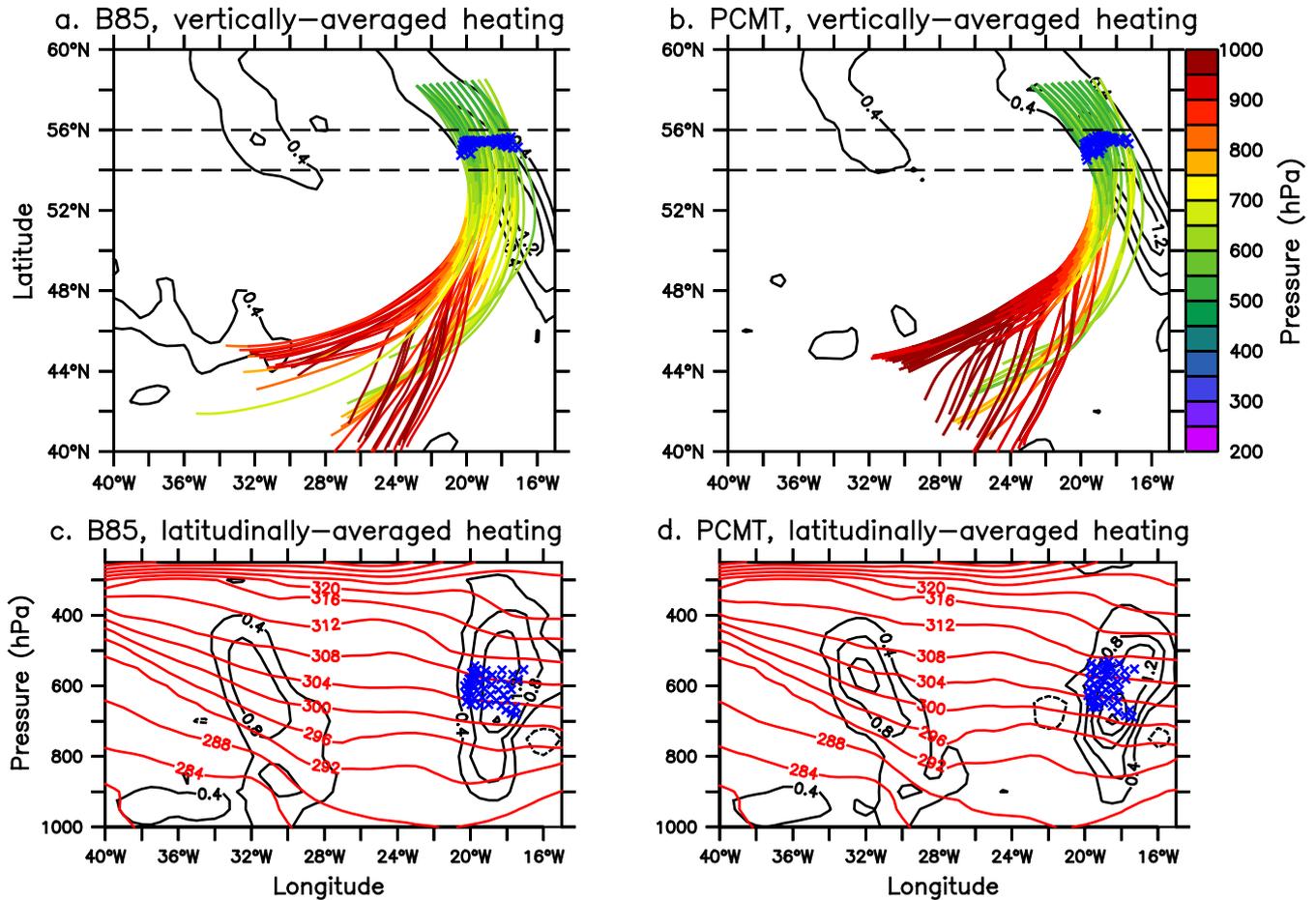
**Figure 8.** Same as Figure 6 but for the trajectories reaching the negative PV anomaly shown in Figures 5a,b.

To summarize, the deeper jet stream in PCMT than B85 can be explained by distinct diabatic processes occurring on both sides of the jet in the middle troposphere. On the cold-air side, half of the PCMT trajectories undergo some PV increase as they travel below the heating before reaching the middle troposphere while all B85 trajectories keep travelling in the middle troposphere. On the warm-air side, PCMT trajectories undergo a more rapid PV decrease because they already pass over the heating which is more confined at lower levels. This induces a PV difference between PCMT and B85 that reinforces the PV gradient near the jet core and thus the jet in PCMT relative to B85.

## 5 Comparison with observations from the NAWDEX IOP6

Since it is the difference in the heating structure that makes the difference in the vertical structure of the jet stream, and since the heating is linked to cloud formation and microphysics, a comparison is made between the ice water content (IWC) of the model simulations and the one retrieved from the radar-lidar observations in Fig. 10.

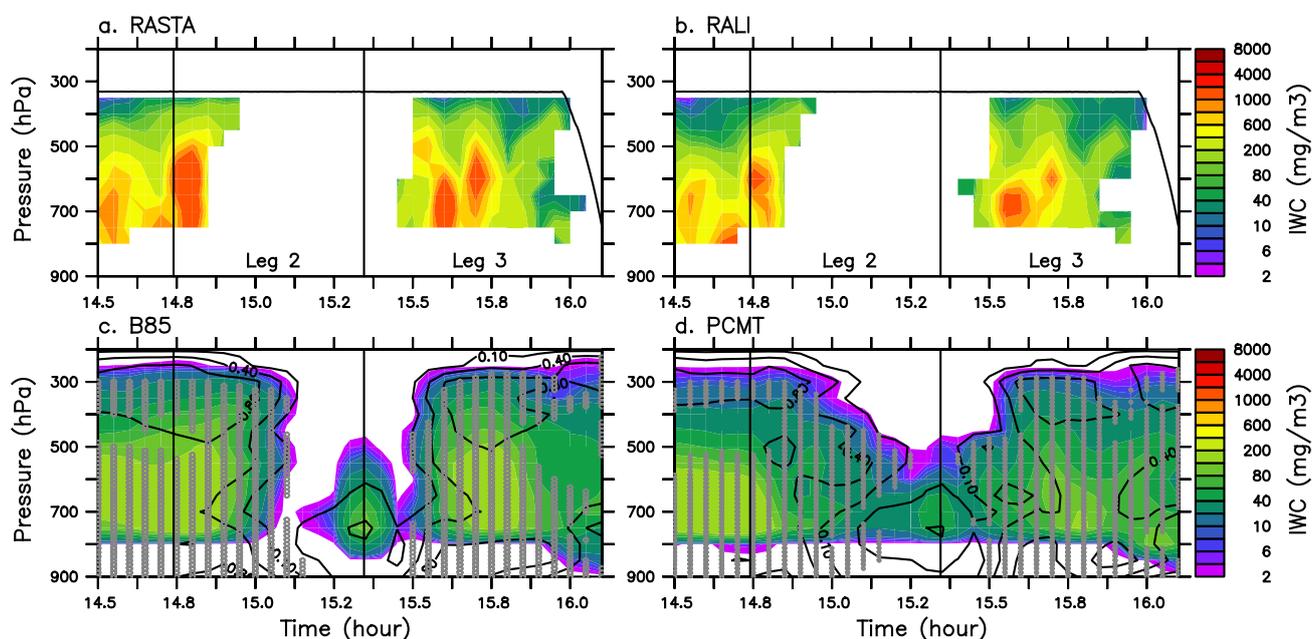
Only the second half of the flight is considered. To better compare to ARPEGE simulations, observations are interpolated at the model outputs resolution ( $0.5^\circ$  grid spacing roughly corresponding to 180 s at the aircraft speed) in Fig. 10. Two IWC products are retrieved from the observations using the Varcloud algorithm (Delanoë and Hogan, 2008; Cazenave et al., 2019): one is based on the radar RASTA measurements only (Fig. 10a) ingesting both reflectivity and Doppler velocity, and the other



**Figure 9.** Same as Fig. 7 but for the trajectories reaching the negative PV anomaly in leg 2 shown in Figs. 5a,b and at 12 UTC 2 October.

420 on RALI (radar and lidar) measurements, i.e. assimilating radar reflectivity and lidar backscatter (Fig. 10b). Since the lidar is more sensitive to small particles and small hydrometeors contents, the RALI retrieval usually leads to smaller IWC than the RASTA retrieval as confirmed by comparing Figs. 10a and b. Note that these two retrievals do not use the same microphysical assumptions due to the difference in sensitivity and penetration capability of these two instruments. As described in Cazenave et al. (2019), the main uncertainties in the retrieval come from the mass-size, area-size relationships. Therefore, the comparison  
 425 between these two retrievals gives an idea of the uncertainties related to those retrievals.

As the flight crosses the WCB region twice, two zones with high IWC are observed in each retrieval: one between 14.5 h and 15 h and the second between 15.4 h and 16 h. The peak values of the retrieved IWC are near  $2000 \text{ mg m}^{-3}$  (Fig. 10a,b) whereas those of the model simulations do not exceed  $400 \text{ mg m}^{-3}$  (Fig. 10c,d). The IWC values of the model simulations strongly depend on the snow fall speed which is a constant prescribed in the model. In the present simulations, its value is  $1.5 \text{ m s}^{-1}$ . Additional sensitivity experiments made by setting its value to  $0.6 \text{ m s}^{-1}$  led to IWC peak values near  $800 \text{ mg m}^{-3}$  (not  
 430

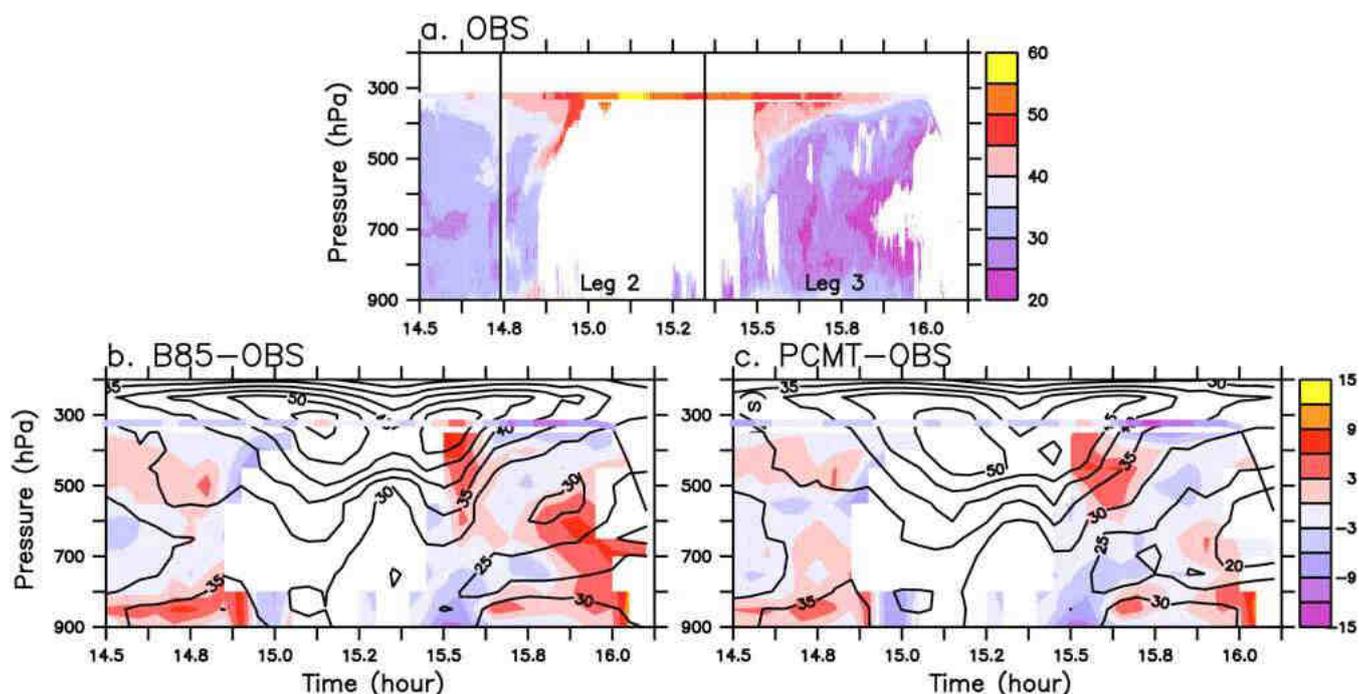


**Figure 10.** Ice water content ( $\text{mg m}^{-3}$ ) for a) retrieval from radar RASTA b) retrieval from radar-lidar RALI along legs 2 and 3 of F7 (airborne level in black line); Total ice water content (snow+cloud ice water;  $\text{mg m}^{-3}$ ; shading) and cloud fraction (black contours) for c) B85 and d) PCMT respectively. Grey circles represent WCB trajectories crossing F7.

shown). So even in the case of low snow fall velocity, the IWC is largely underestimated in the model. In a supplementary figure (Fig. S5), the IWC divided by the cloud fraction, which could be thought as being more relevant to compare to observations, also fails to reproduce the high IWC values detected in the observations. This result is not surprising following Mazoyer et al. (2021) who found similar underestimation in regional model simulations of the Stalactite cyclone. The comparison between the two simulations shows that the peak values are slightly higher in B85 than in PCMT but the difference is weak to be conclusive.

The difference in IWC spatial distribution is worth commenting, especially with regard to the previous sections. The flight crossed the separating area between the cloudy WCB region with high IWC values and the clear-sky region close to the cyclone center twice, at 14.9 h and 15.5 h. These two transitions are well visible at 15.0 h and 15.5 h in B85 but are much less clear in PCMT. In the latter run, in the observed clear-sky region, there are clouds (Fig. 10d) whose tops are about 500 hPa. Also many trajectories belonging to that area between 15.0 h and 15.5 h satisfy the WCB criterion (see grey circles in Fig. 10c) and correspond to the trajectories which undergo strong heating during their ascent between 900 hPa and 600 hPa (Fig. 6d). Figure S6 shows the same patterns but with model outputs and interpolation made over a  $0.1^\circ \times 0.1^\circ$  grid. The results are qualitatively the same but the scales of the clouds are more representative of the model resolution.

To conclude, even though both simulations fail to produce the high values of IWC seen in the observations, the spatial distribution of IWC differs between the two simulations. B85 better represents the abrupt transition between the cloudy region of the WCB and the clear-sky region behind the cold front.



**Figure 11.** (a) Wind speed (shadings) observations (RASTA+ aircraft) at full resolution; Wind speed (black contour) and wind speed anomaly respectively with observations interpolated at model resolution (shadings) for b) B85 (B85-Obs.), c) PCMT (PCMT-Obs.).

To determine which run better represents the jet stream, the wind speed of the simulations is compared with those measured by the radar RASTA and on board the aircraft in Fig. 11. There is generally a good correspondence between the wind speed measured by the aircraft and that measured by the radar (Fig.11a). However, as there are no clouds from 14.9 h and 15.5 h, no wind speed observations from the radar are available. Hence, the jet stream, which is crossed twice at 15.1 h and 15.5 h, is only very partially covered by the radar measurements (compare the color shadings with the contours). The radar measurements are useless to look precisely at the vertical structure of the jet stream and the aircraft measurements provide information at a given level only. At the aircraft level, the two runs give similar anomalies with some underestimation of winds between 15.6 h and 16.0 h. There is only one region where the two simulations behave very differently: it is between 15.6 h and 16 h and near the 500-700 hPa layer (Figs. 11b-c). While there is a minimum in wind speed in the observations as well as in PCMT in that region, a secondary jet is present in B85 with values near  $30 \text{ m s}^{-1}$ , which was already discussed in section 3. The wind speed in B85 is overestimated by about  $6-9 \text{ m s}^{-1}$ . According to Fig. 5a, this wind speed anomaly is linked to the stronger PV gradient in B85 than PCMT associated with the dipolar PV anomaly located at 600 hPa at the end of the flight (trajectory index from 65 to 75). Comparison with radar measurements leads to the same conclusion as the comparison made with (re)-analyses in section 3. The secondary jet in B85 along leg 3 is not present in any of these references.

## 6 Conclusions

The present study and our companion paper (Rivière et al., 2021, RW21) provide a general view of the impact of deep convection representation in a global numerical weather model on the WCB of an explosive extra-tropical cyclone observed during NAWDEX and on the jet stream aloft. Three simulations of the Météo-France global model ARPEGE, which only differ by their deep convection representation, are investigated. Two of them use the model with a distinct deep convection scheme activated: one with the scheme developed in Bougeault (1985, B85), the other one with the one from Piriou et al. (2007, PCMT). In the last ARPEGE simulation, called NoConv, no deep convection scheme is activated. The companion paper investigated the general behaviour of WCB activity in the three simulations and its impact on the jet stream in the WCB outflow region above the bent-back warm front. The present study was dedicated to the impact of these parameterization schemes on the jet stream in the WCB ascending region above the cold front.

The systematic comparison made between the three simulations led to the following conclusions:

- The deep convection representation has an important effect on the vertical structure of the jet stream above the cold front: the jet stream is deeper in NoConv simulation, i.e. without parameterized deep convection, and in PCMT simulation than in B85 simulation.
- The deeper jet stream in NoConv and PCMT compared to B85 is associated with a deepening of the dynamical tropopause (i.e. higher PV) behind the cold front and with more PV destruction ahead of the cold front in middle troposphere (600 hPa). The difference in PV between PCMT and B85 is marked by a dipolar PV anomaly centred on the jet core which reinforces the PV gradient and thus the jet in middle troposphere. The dipolar PV anomaly is due to differences in diabatic processes between the two simulations.
- The same tropopause deepening is observed for PEARP members sharing the same deep convection closure, suggesting, as in RW21, the key role played by that closure on the jet stream structure.
- On the cold-air side of the jet, the high PV area of the dipolar anomaly is due to different behaviours of the Lagrangian trajectories reaching that area. In B85, they form an homogeneous group of trajectories staying at the same level pressure in middle troposphere and undergoing modest PV fluctuations in the cold sector. In contrast, the PCMT trajectories are clearly separated into two groups. One group of trajectories behave like in B85 with weak pressure variations but are more subject to PV increase because they pass over a more marked cooling due to radiation and turbulence above cold-sector convective clouds. The second group behaves in a totally different manner; they come from the boundary layer, ascent on the western flank of the region of strong latent heating and undergo PV increase, before joining the first group at the same altitude. The strong latent heating is mainly due to large-scale cloud and to a lesser extent to convection. The trajectories of the second group satisfy the WCB criterion of 300 hPa ascent in 24 h chosen in the present study. The presence of ascending trajectories very near the core of the cold front is unexpected and is associated with a more important overlapping of the heating area and the horizontal temperature gradient in PCMT and NoConv than in B85 (not shown).

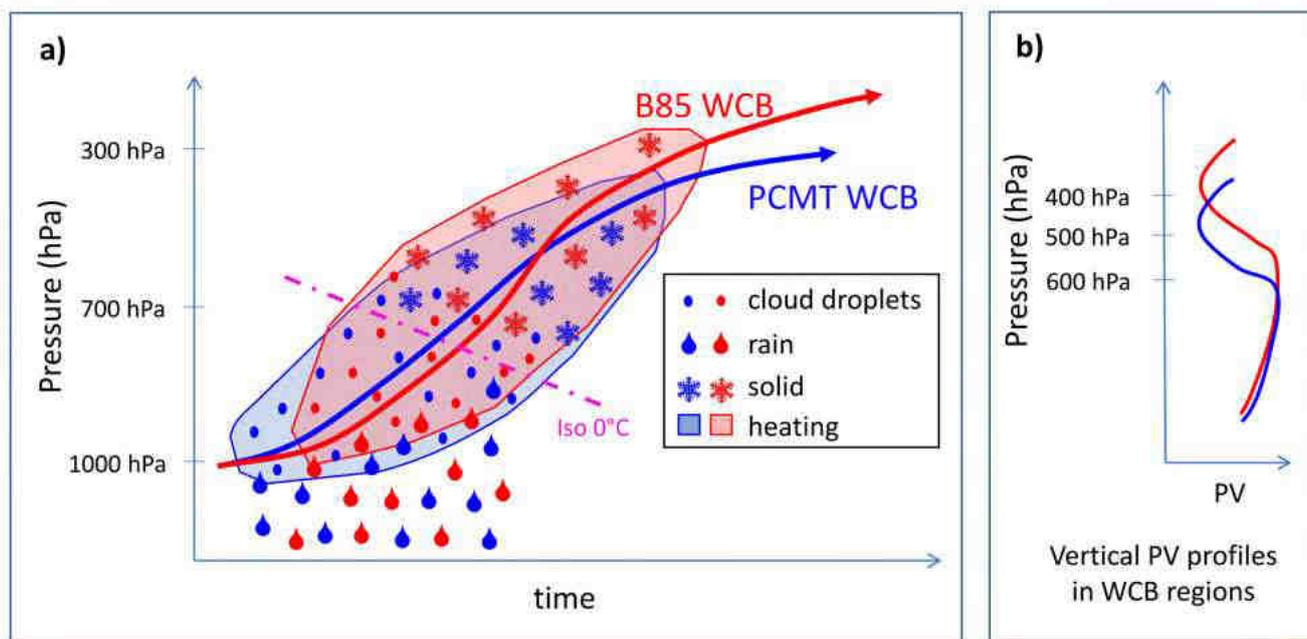


– On the warm-air side of the jet, WCB Lagrangian trajectories are quite similar but their behaviour, synthesized in Fig. 12, is clearly different between the two deep convection schemes. With PCMT, WCB trajectories pass sooner through the main heating area as it is located at lower altitude than in B85 (Fig. 12a). Hence, a sooner decrease of PV occurs for the trajectories in PCMT and negative PV tendency appears in middle troposphere (Fig. 12b). This is to be contrasted with B85 where the peak values of the heating extend further upward and much less PV destruction occurs at mid-tropospheric levels. This difference in the altitude of the heating maximum explains the difference in PV ahead of the cold front in middle troposphere.

Then, the question of the realism of the different hindcasts has been addressed by comparing them to different (re)-analyses datasets and to NAWDEX airborne observations. It led to the following conclusions:

- The jet stream structure in (re)-analyses datasets as provided by ERA5 and ECMWF / Météo-France operational analyses lies in between the B85 simulation on the one hand and the PCMT and NoConv simulations on the other hand. For instance, the altitude of jet maximum in B85 is located above those in (re)-analyses datasets which themselves are above those in PCMT and NoConv. Another example concerns the peak values of PV near 600 hPa: in descending order there are the PCMT and NoConv values followed by (re)-analyses datasets and finally B85 values.
- The ice water content retrieved from the radar-lidar measurements using the Varcloud algorithm (Delanoë and Hogan, 2008; Cazenave et al., 2019) shows a clear separation between the cloudy region ahead of the cold front and the ice-cloud free region behind it. This separation is well identified in B85 but not in PCMT. Behind the cold front, B85 is more realistic than PCMT which exhibits too much mid-level clouds.
- Since the main jet is largely embedded in clear-sky regions, the Doppler cloud radar observations are useless to determine which run is more realistic. However, analysis of the wind speed anomalies with respect to the observations in cloudy regions ahead of the cold front indicate that B85 creates a secondary jet in middle troposphere which does not appear in the observations nor in PCMT. Therefore, in that particular region, the PCMT simulation performs better and this is due to the sooner PV destruction in PCMT.

Therefore, the present analysis cannot state which hindcasts better represent the observations as the conclusion is strongly dependent on the regions we are looking at. However, it shows that PCMT and B85 have drastically different behaviours with the former being close to the simulation without parameterized convection. The overall picture provided by the present study and the companion paper is the following. In B85 simulation, the heating is more homogeneously distributed ahead of the cold front and along the bent-back warm front, it extends further up leading to stronger PV destruction in the upper troposphere that accelerates the ridge building in the WCB outflow region. In PCMT and NoConv simulations, the heating is more heterogeneous, especially in NoConv, it extends less in the upper troposphere and the loss of PV happens at a lower altitude ahead of the cold front and makes the jet deeper in that region. Finally, note that such a difference between the two deep convection schemes was also found above the cold front of the following extratropical cyclone on 4-5 October 2016 (NAWDEX IOP7; not shown).



**Figure 12.** Schematic representing differences between the two convection schemes in latent heating and PV tendencies.

The important question that follows up is: where do these different behaviours come from? A first answer found in this study is the different deep convection parameterization closure used, namely the CAPE for PCMT and moisture convergence closure for B85. However, further sensitivity studies will be planned in order to better identify effects of different parts of the deep convection schemes on mid-latitude cyclogenesis and jet stream.

*Data availability.* Data is available by contacting the corresponding author. ERA5 data are accessible via the climate data store (<https://cds.climate.copernicus.eu>; DOI: 10.24381/cds.bd0915c6; latest access: April 2021).

Worldview NASA picture is available at <https://go.nasa.gov/3xOhRhv>

*Author contributions.* GR and PA designed the initial study. MW and GR performed the data analysis and made the figures. MW computed the Lagrangian trajectories. CL performed the ARPEGE simulations with the help of JMP. PA developed the Lagrangian trajectory algorithm. JD, QC, and JP provided the observational datasets. All authors contributed to the scientific discussions.



*Competing interests.* The authors declare that they have no competing interests.

*Acknowledgements.* The study benefited from discussions with the participants of the DIP-NAWDEX (DIabatic Processes in the North Atlantic Waveguide and Downstream impact EXperiment) project which is supported and funded by the Agence Nationale de la Recherche (ANR). It also benefited from discussions with our international NAWDEX partners during annual workshops. ERA5 datasets that have been  
540 generated under the framework of the Copernicus Climate Change Service (C3S). We thank Heini Wernli for downloading the ECMWF-IFS operational analysis data and Hanin Binder for sending them to us. We acknowledge the use of imagery from the NASA Worldview application (<https://worldview.earthdata.nasa.gov>), part of the NASA Earth Observing System Data and Information System (EOSDIS).

*Financial support.* This research has been supported by the Agence Nationale de la Recherche (grant no. ANR-17-CE01-0010-01). The  
545 airborne measurements and the SAFIRE Falcon flights received direct funding from IPSL, Météo-France, INSU-LEFE, EUFAR-NEAREX, and ESA (EPATAN, contract no. 4000119015/16/NL/CT/gp).



## References

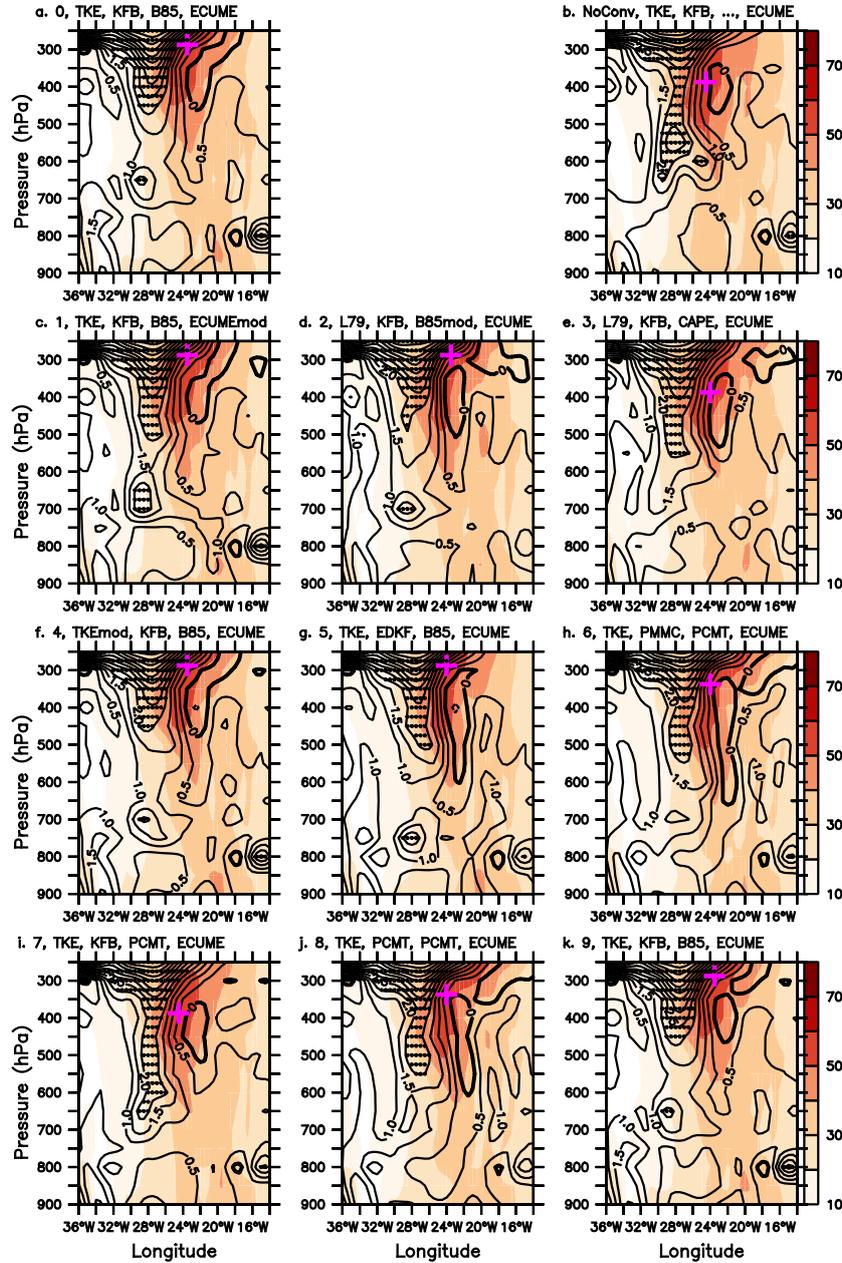
- Baumgart, M., Riemer, M., Wirth, V., and Teubler, F.: Potential Vorticity Dynamics of Forecast Errors: A Quantitative Case Study, *Mon. Wea. Rev.*, 146, 1405–1425, <https://doi.org/10.1175/MWR-D-17-0196.1>, 2018.
- 550 Binder, H., Rivière, G., Arbogast, P., Maynard, K., Bosser, P., Joly, B., and Labadie, C.: Dynamics of forecast-error growth along cut-off Sanchez and its consequence for the prediction of a high-impact weather event over southern France, *Quart. J. Roy. Meteor. Soc.*, 147, 3263–3285, <https://doi.org/10.1002/qj.4127>, 2021.
- Blanchard, N., Pantillon, F., Chaboureaud, J. P., and Delanoë, J.: Organization of convective ascents in a warm conveyor belt, *Weather Clim. Dynam.*, 1, 617–634, <https://doi.org/10.5194/wcd-1-617-2020>, 2020.
- 555 Blanchard, N., Pantillon, F., Chaboureaud, J. P., and Delanoë, J.: Mid-level convection in a warm conveyor belt accelerates the jet stream, *Weather Clim. Dynam.*, 2, 37–53, <https://doi.org/10.5194/wcd-2-37-2021>, 2021.
- Bougeault, P.: A simple parameterization of the large-scale effects of cumulus convection., *Mon. Wea. Rev.*, 113, 2105–2121, 1985.
- Bruneau, D., Pelon, J., Blouzon, F., Spatazza, J., Genau, P., Buchholtz, G., Amarouche, N., Abchiche, A., and Aouji, O.: 355-nm high spectral resolution airborne lidar LNG: system description and first results, *Appl. Opt.*, 54, 8776–8785, <https://doi.org/10.1364/AO.54.008776>,  
560 2015.
- Cazenave, Q., Ceccaldi, M., Delanoë, J., Pelon, J., Groß, S., and Heymsfield, A.: Evolution of DARDAR-CLOUD ice cloud retrievals: new parameters and impacts on the retrieved microphysical properties, *Atmospheric Measurement Techniques*, 12, 2819–2835, <https://doi.org/10.5194/amt-12-2819-2019>, 2019.
- Chagnon, J. and Gray, S. L.: Horizontal potential vorticity dipoles on the convective storm scale, *Quart. J. Roy. Meteor. Soc.*, 135, 1392–1408,  
565 2009.
- Chagnon, J., Gray, S. L., and Methven, J.: Diabatic processes modifying potential vorticity in a North Atlantic Cyclone, *Quart. J. Roy. Meteor. Soc.*, 139, 1270–1282, 2013.
- Courtier, P., Freyrier, C., Geleyn, J., Rabier, F., and Rochas, M.: The ARPEGE project at Météo-France., in: *ECMWF Seminar Proceedings*, pp. 193–231, Reading, volume II, 1991.
- 570 Crezee, B., Joos, H., and Wernli, H.: The Microphysical Building Blocks of Low-Level Potential Vorticity Anomalies in an Idealized Extratropical Cyclone, *J. Atmos. Sci.*, 74, 1403–1416, 2017.
- Davies, H. C. and Didone, M.: Diagnostics and dynamics of forecast error growth, *Mon. Wea. Rev.*, 141, 2483–2501, 2013.
- Delanoë, J., Protat, A., Jourdan, O., Pelon, J., Papazonni, M., Dupuy, R., Gayet, J.-F., and Jouan, C.: Comparison of Airborne In Situ, Airborne Radar-Lidar, and Spaceborne Radar-Lidar Retrievals of Polar Ice Cloud Properties Sampled during the POLARCAT Campaign,  
575 *J. Atmos. Oceanic. Technol.*, 30, 57–73, 2013.
- Delanoë, J. and Hogan, R. J.: A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer, *J. Geophys. Res.*, 113, <https://doi.org/10.1029/2007JD009000>, 2008.
- Descamps, L., Labadie, C., Joly, A., Bazile, E., Arbogast, P., and Cébron, P.: PEARP, the Météo-France short-range ensemble prediction system., *Quart. J. Roy. Meteor. Soc.*, 141, 1671–1685, 2015.
- 580 Fink, A. H., Brucher, T., Ermert, V., Kruger, A., and Pinto, J. G.: The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change, *Nat. Hazards Earth Syst. Sci.*, 9, 405–423, 2009.

- Flack, D. L. A., Rivière, G., Musat, I., Roehrig, R., Bony, S., Delanoë, J., Cazenave, Q., and Pelon, J.: Representation by two climate models of the dynamical and diabatic processes involved in the development of an explosively deepening cyclone during NAWDEX, *Weather Clim. Dynam.*, 2, 233–253, <https://doi.org/10.5194/wcd-2-233-2021>, 2021.
- 585 Grams, C. M., Binder, H., Pfahl, S., Piaget, N., and Wernli, H.: Atmospheric processes triggering the central European floods in June 2013, *Nat. Hazards Earth Syst. Sci.*, 14, 1691–1702, 2014.
- Gray, S. L., Dunning, C. M., Methven, J., Masato, G., and Chagnon, J. M.: Systematic model forecast error in Rossby wave structure, *Geophys. Res. Lett.*, 41, 2979–2987, 2014.
- Harvey, B., Methven, J., and Ambaum, M. H. P.: Rossby wave propagation on potential vorticity fronts with finite width, *J. Fluid Mech.*,  
590 794, 775–797, 2016.
- Harvey, B., Methven, J., and Ambaum, M. H. P.: An Adiabatic Mechanism for the Reduction of Jet Meander Amplitude by Potential Vorticity Filamentation, *J. Atmos. Sci.*, 75, 4091–4106, 2018.
- Harvey, B., Methven, J., Sanchez, C., and Schafler, A.: Diabatic generation of negative potential vorticity and its impact on the North Atlantic jet stream, *Quart. J. Roy. Meteor. Soc.*, 146, 1477–1497, 2020.
- 595 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G. D., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, *Quart. J. Roy. Meteor. Soc.*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- 600 Joos, H. and Forbes, R. M.: Impact of different IFS microphysics on a warm conveyor belt and the downstream flow evolution, *Quart. J. Roy. Meteor. Soc.*, 142, 2727–2739, <https://doi.org/10.1002/qj.2863>, 2016.
- Lopez, P.: Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes, *Quart. J. Roy. Meteor. Soc.*, 128, 229–257, <https://doi.org/10.1256/00359000260498879>, 2002.
- Maddison, J. W., Gray, S. L., Martínez-Alvarado, O., and Williams, K. D.: Upstream cyclone influence on the predictability of block onsets  
605 over the Euro-Atlantic region, *Mon. Wea. Rev.*, 147, 1277–1296, <https://doi.org/10.1175/MWR-D-18-0226.1>, 2019.
- Martínez-Alvarado, O. and Plant, R. S.: Parametrized diabatic processes in numerical simulations of an extratropical cyclone., *Quart. J. Roy. Meteor. Soc.*, 140, 1742–1755, <https://doi.org/10.1002/qj.2254>, 2014.
- Martínez-Alvarado, O., Joos, H., Chagnon, J., Boettcher, M., Gray, S. L., Plant, R. S., Methven, J., and Wernli, H.: The dichotomous structure of the warm conveyor belt, *Quart. J. Roy. Meteor. Soc.*, 140, 1809–1824, 2014.
- 610 Martius, O., Schwierz, C., and Davies, H.: Far upstream precursors of heavy precipitation events on the Alpine south side., *Quart. J. Roy. Meteor. Soc.*, 134, 417–428, 2008.
- Massacand, A., Wernli, H., and Davies, H.: Heavy precipitation on the Alpine southside: an upper-level precursor, *Geophys. Res. Lett.*, 25, 1435–1438, 1998.
- Mazoyer, M., Ricard, D., Rivière, G., Delanoë, J., Arbogast, P., Vié, B., Lac, C., Cazenave, Q., and Pelon, J.: Microphysics impacts on the  
615 warm conveyor belt and ridge building of the NAWDEX IOP6 cyclone, *Mon. Wea. Rev.*, p. in press, 2021.
- Nuissier, O., Joly, B., Joly, A., Ducrocq, V., and Arbogast, P.: A statistical downscaling to identify the large-scale circulation patterns associated with heavy precipitation events over southern France, *Quart. J. Roy. Meteor. Soc.*, 137, 1812–1827, 2011.

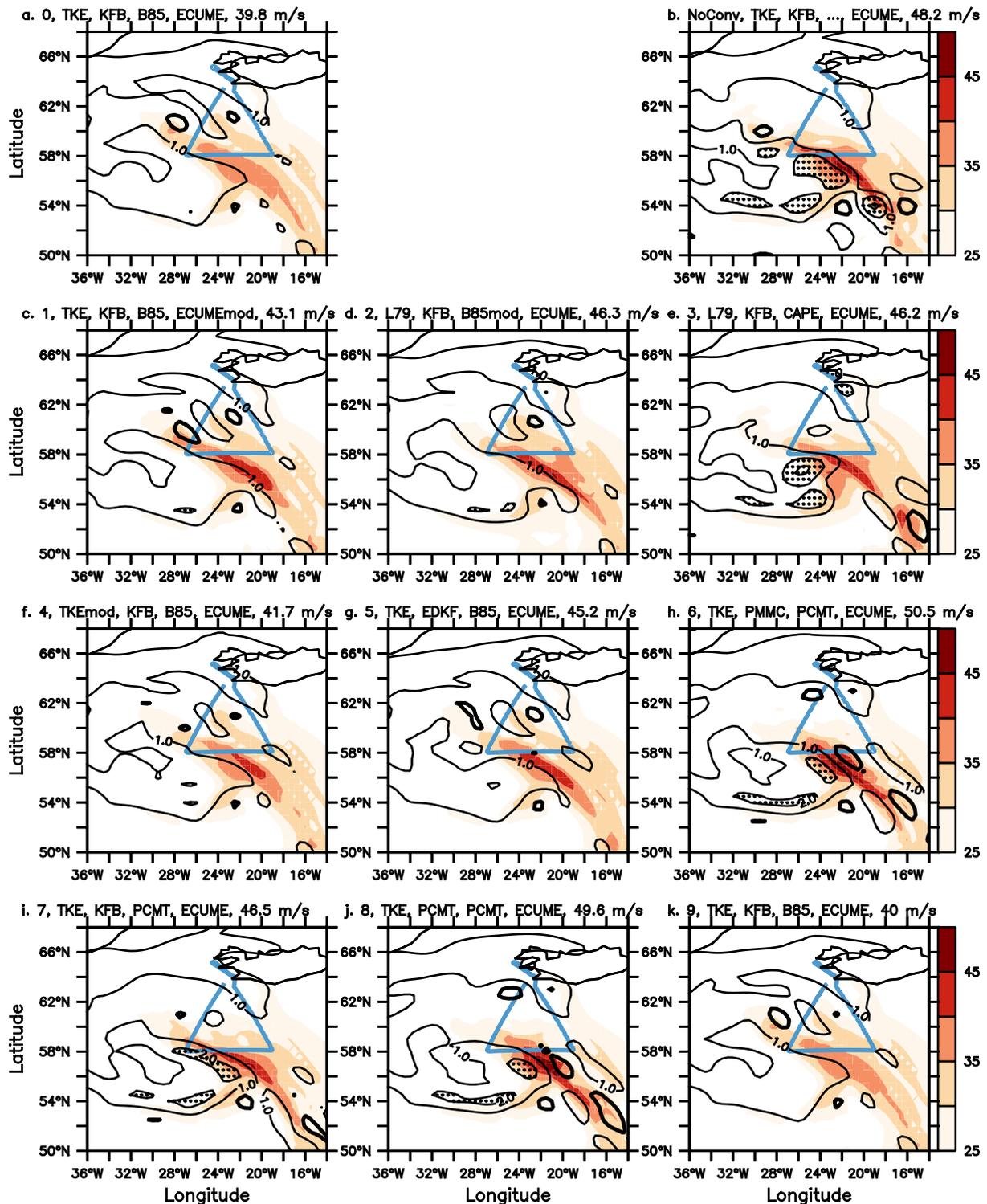
- Oertel, A., Boettcher, M., Joos, H., Sprenger, M., and Wernli, H.: Convective activity in an extratropical cyclone and its warm conveyor belt – a case-study combining observations and a convection-permitting model simulation, *Quart. J. Roy. Meteor. Soc.*, 145, 1406–1426, <https://doi.org/10.1002/qj.3500>, 2019.
- Oertel, A., Boettcher, M., Joos, H., Sprenger, M., and Wernli, H.: Potential vorticity structure of embedded convection in a warm conveyor belt and its relevance for large-scale dynamics, *Weather Clim. Dynam.*, 1, 127–153, <https://doi.org/10.5194/wcd-1-127-2020>, 2020.
- Oertel, A., Sprenger, M., Joos, H., Boettcher, M., Konow, H., Hagen, M., and Wernli, H.: Observations and simulation of intense convection embedded in a warm conveyor belt – how ambient vertical wind shear determines the dynamical impact, *Weather Clim. Dynam.*, 2, 89–110, <https://doi.org/10.5194/wcd-2-89-2021>, 2021.
- Parsons, D., Beland, M., Burridge, D., and co authors: THORPEX research and the science of prediction, *Bull. Amer. Meteor. Soc.*, pp. 807–830, 2017.
- Pinto, J., Zacharias, S., Fink, A. H., Leckebusch, G., and Ulbrich, U.: Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO, *Clim. Dyn.*, 32, 711–737, 2009.
- Piriou, J.-M., Redelsperger, J.-L., Geleyn, J.-F., Lafore, J.-P., and Guichard, F.: An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations, *J. Atmos. Sci.*, 64, 4127–4139, <https://doi.org/10.1175/2007JAS2144.1>, 2007.
- Ponzano, M., Joly, B., Descamps, L., and Arbogast, P.: Systematic errors analysis of heavy precipitating events prediction using a 30-year hindcast dataset., *Nat. Hazards Earth Syst. Sci.*, 20, 1369–1389, <https://doi.org/10.5194/nhess-20-1369-2020>, 2020.
- Rivière, G., Arbogast, P., Maynard, K., and Joly, A.: The essential ingredients leading to the explosive growth stage of the European wind storm “Lothar” of Christmas 1999, *Quart. J. Roy. Meteor. Soc.*, 136, 638–652, 2010.
- Rivière, G., Wimmer, M., Arbogast, P., Piriou, J.-M., Delanoë, J., Labadie, C., Cazenave, Q., and Pelon, J.: The impact of deep convection representation in a global atmospheric model on the warm conveyor belt and jet stream during NAWDEX IOP6, *Weather Clim. Dynam.*, 2021, 1–32, <https://doi.org/10.5194/wcd-2021-38>, 2021.
- Rodwell, M. J., Magnusson, L., and co authors: Characteristics of occasional poor medium-range weather forecasts for Europe, *Bull. Amer. Meteor. Soc.*, 94, 1393–1405, 2013.
- Saffin, L., Gray, S. L., Methven, J., and Williams, K. D.: Processes Maintaining Tropopause Sharpness in Numerical Models, *J. Geophys. Res.*, 122, 9611–9627, 2017.
- Sanders, F. and Gyakum, J.: Synoptic-dynamic climatology of the “Bomb”, *Mon. Wea. Rev.*, 108, 1589–1606, 1980.
- Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., Methven, J., Rivière, G., Ament, F., Boettcher, M., Bramberger, M., Cazenave, Q., Cotton, R., Crewell, S., Delanoë, J., Dörnbrack, A., Ehrlich, A., Ewald, F., Fix, A., Grams, C. M., Gray, S. L., Grob, H., Groß, S., Hagen, M., Harvey, B., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Lemmerz, C., Lux, O., Magnusson, L., Mayer, B., Mech, M., Moore, R., Pelon, J., Quinting, J., Rahm, S., Rapp, M., Rautenhaus, M., Reitebuch, O., Reynolds, C. A., Sodemann, H., Spengler, T., Vaughan, G., Wendisch, M., Wirth, M., Witschas, B., Wolf, K., and Zinner, T.: The North Atlantic Waveguide and Downstream Impact Experiment., *Bull. Amer. Meteor. Soc.*, 99, 1607–1637, 2018.
- Schäfler, A., Harvey, B., Methven, J., Doyle, J. D., Rahm, S., Reitebuch, O., Weiler, F., and Witschas, B.: Observation of Jet Stream Winds during NAWDEX and Characterization of Systematic Meteorological Analysis Errors, *Mon. Wea. Rev.*, 148, 2889–2907, 2020.
- Schemm, S., Wernli, H., and Papritz, L.: Warm conveyor belt in idealized moist baroclinic wave simulation, *J. Atmos. Sci.*, 70, 627–652, <https://doi.org/10.1175/JAS-D-12-0147.1>, 2013.



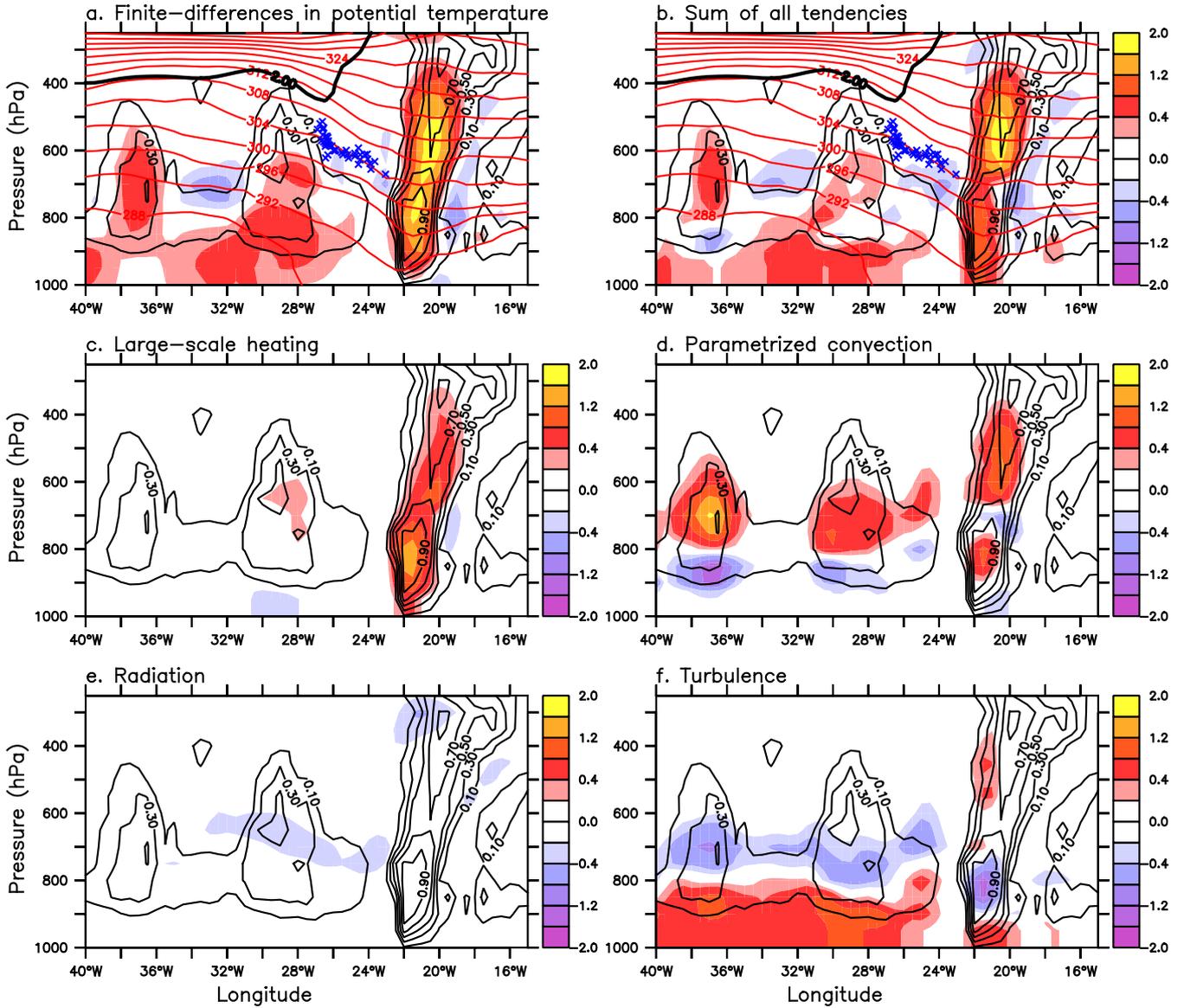
- 655 Spreitzer, E., Attinger, R., Boettcher, M., Forbes, R., Wernli, H., and Joos, H.: Modification of Potential Vorticity near the Tropopause by Nonconservative Processes in the ECMWF Model, *J. Atmos. Sci.*, 76, 1709–1726, 2019.
- Steinfeld, D., Boettcher, M., Forbes, R., and Pfahl, S.: The sensitivity of atmospheric blocking to upstream latent heating – numerical experiments, *Weather Clim. Dynam.*, 1, 405–426, <https://doi.org/10.5194/wcd-1-405-2020>, 2020.
- Uccellini, L. W.: Processes contributing to the rapid development of extratropical cyclones, vol. Extratropical cyclones, Newton and  
660 Holopainen eds, p. chapter 6, American Meteorological Society, 1990.
- Wernli, H., Dirren, S., Liniger, M. A., and Zillig, M.: Dynamical aspects of the life cycle of the winter storm 'Lothar' (24-26 December 1999), *Quart. J. Roy. Meteor. Soc.*, 128, 405–429, 2002.
- Wirth, V., Riemer, M., Chang, E. K. M., and Martius, O.: Rossby Wave Packets on the Midlatitude Waveguide—A Review, *Mon. Wea. Rev.*, 146, 1965–2001, 2018.



**Figure S 1.** Vertical cross section at  $58^\circ$  N (grey line in fig. 1) of the zonal wind (shadings, pink crosses localize wind speed maximum) and Potential Vorticity (black contours with hatched areas for values superior to 2PVU) at 15 UTC, 2 October 2016 for the NoConv and the 10 PEARP members starting with the same initial conditions but differing from four main physics package: turbulence, shallow convection, deep convection and surface fluxes and whose schemes are indicated in that order in the subtitle of each panel. Turbulence schemes can be TKE (Turbulent Kinetic Energy scheme of Cuxart et al. 2000), L79 (Louis, 1979), or TKEmod (slightly modified version of TKE in which horizontal advection is ignored). Shallow convection can be KFB (Kain and Fritsch, 1993; Bechtold et al., 2001), PCMT, PMMC (Pergaud et al. 2009) or EDKF (eddy diffusivity and Kain-Fritsch scheme). Deep convection schemes can be B85, PCMT, CAPE (B85 where the closure is based on the CAPE), B85mod (B85 in which deep convection is only triggered if cloud top exceeds 3000 m). The surface oceanic fluxes are based on Belamari et al. (2005)'s scheme (ECUME) and an alternative in which evaporative fluxes are enhanced (ECUMEmod). Members 0 and 9 are characterized by the same parametrization set-up, but member 9 differs in the modelling of orographic waves. See Ponzano et al. (2020) for more details on the setup of the different members.



**Figure S 2.** Wind (shadings) and Potential Vorticity (black contours with hatched areas for values superior to 2 PVU) at 600 hPa at 15 UTC, 2 October 2016 for the NoConv and the 10 PEARP members starting with the same initial conditions. The Flight F7 of the SAFIRE Falcon appears in blue line.



**Figure S 3.** Vertical cross sections of heating averaged between  $45^{\circ}\text{N}$  and  $49^{\circ}\text{N}$  (shading; units:  $\text{K h}^{-1}$ ), PV (2 PVU isoline in bold black contour), potential temperature (red contours) and WCB positions (blue crosses) at 03 UTC 2 October with B85. (a) heating computed with finite differences in time and space as in the whole paper and (b) resulting from the sum of all diabatic tendencies issued from model outputs. Panel (b) is the sum of (c) large-scale cloud (resolved) sensible and latent heating, (d) parametrized convection sensible and latent heating, (e) radiative heating and (f) turbulence heating. The thin black contours represent the cloud fraction.

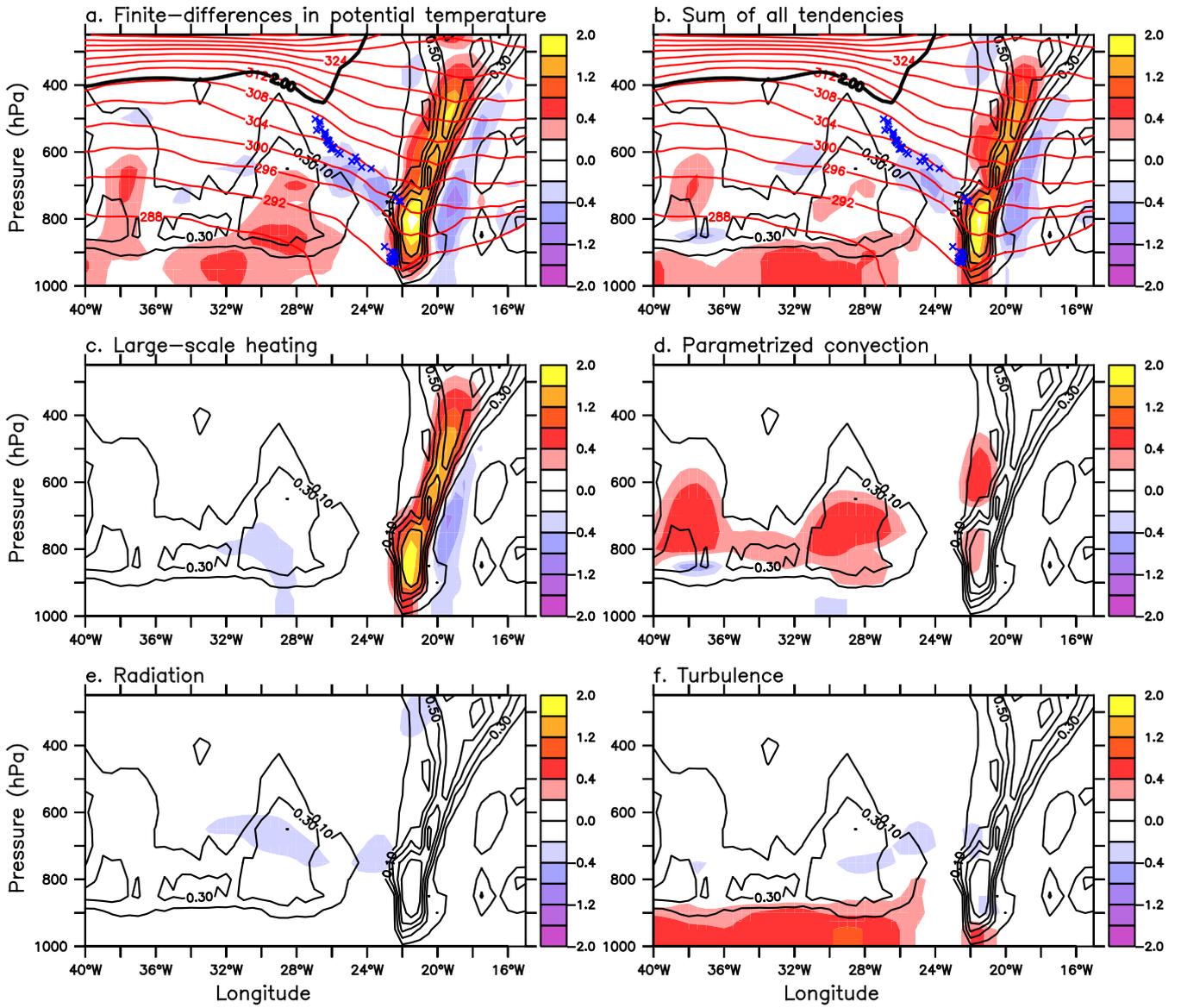


Figure S 4. Same as Figure S3 but for PCMT simulation

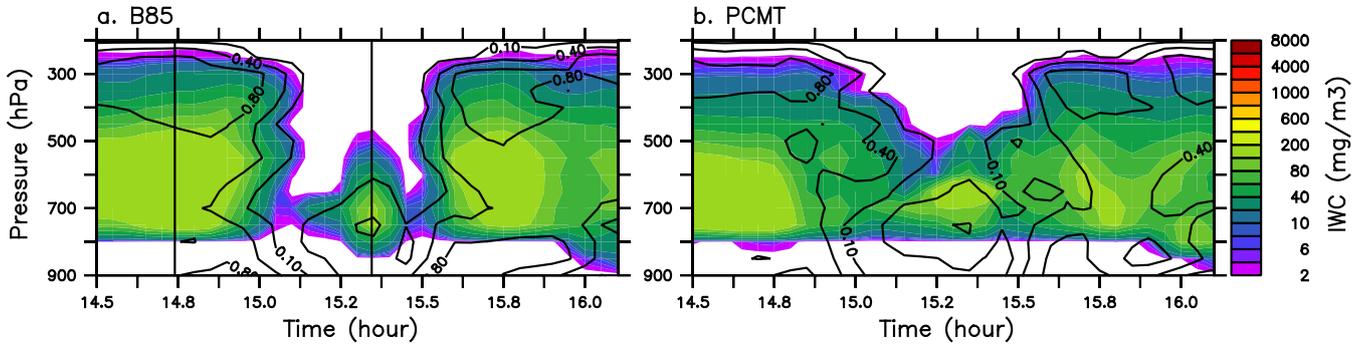


Figure S 5. Same as Figure 10c,d but divided by cloud fraction (cloud ice water/3D cloud fraction + snow/2D cloud fraction)

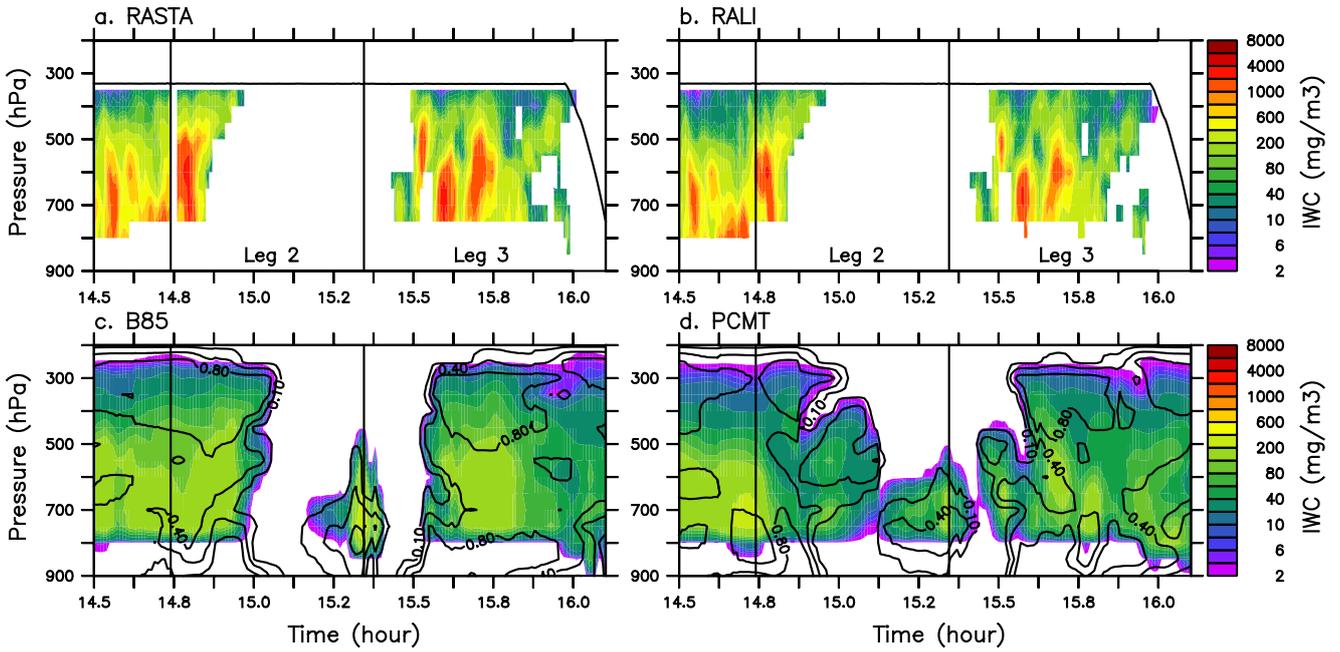


Figure S 6. Same as Figure 10 but with model outputs and radar observation interpolated over a  $0.1^\circ \times 0.1^\circ$  grid

---

---

## **Annexe B**

# **ANASYG - Représentation schématique des situations météorologiques**

Pour les études de cas présentées dans ce manuscrit, nous avons utilisé des ANALyses SYnoptiques Graphiques (ANASYG). Celles-ci sont réalisées par les prévisionnistes de Météo-France à partir du modèle global déterministe le mieux calibré (le plus proche des observations entre ARPEGE et IFS) ainsi qu'à partir d'observations (satellites, radar, stations au sol, ...). Elles permettent de résumer la situation météorologique en utilisant des symboles dont la signification est représentée dans la légende en Figure B.1.

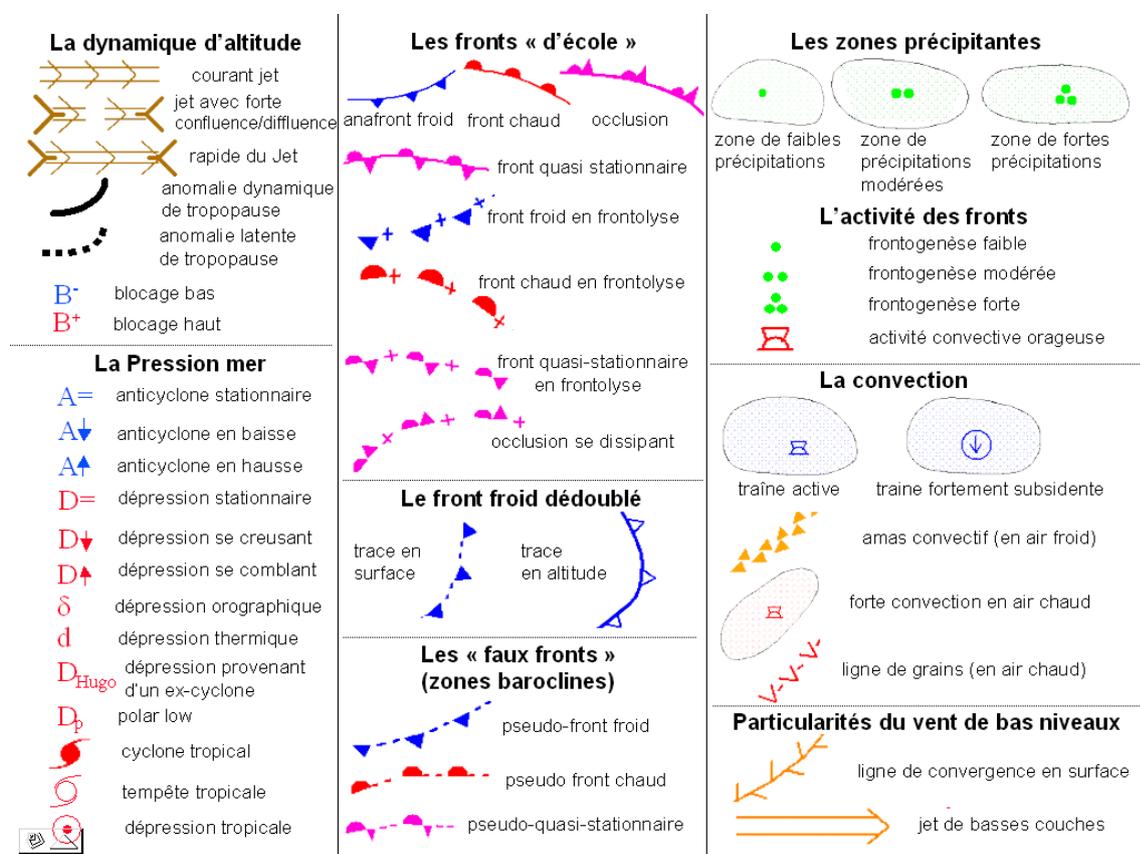


FIGURE B.1 – Légende des symboles utilisées dans les ANASYG







---

**Résumé** — Malgré une amélioration constante des modèles numériques de prévision du temps, ceux-ci restent toujours entachés d’erreurs. La représentation de ces sources d’erreurs est donc primordiale, en particulier dans les systèmes de Prévision d’Ensemble. La Prévision d’Ensemble AROME (PEARO) utilisée à Météo-France représente actuellement les incertitudes du modèle AROME en perturbant des tendances en sortie des paramétrisations physiques. Cependant, cette méthode présente de nombreux inconvénients dont une difficile interprétabilité physique des résultats. Le présent travail s’intéresse à des méthodes plus physiques, s’appuyant sur la perturbation de paramètres au sein de ces paramétrisations.

Sur les conseils d’experts en physique, 21 paramètres incertains à perturber ont été sélectionnés. Des analyses de sensibilité utilisant les méthodes de Morris et de Sobol’ ont permis de réduire cette liste à huit paramètres ayant une influence forte sur les prévisions du modèle AROME. Différentes techniques de perturbations des paramètres incertains ont ensuite été mises en place et évaluées. Celles-ci améliorent les performances de la PEARO pour la plupart des variables de temps sensible telles que le vent et les précipitations. Différentes méthodes d’optimisation se focalisant sur l’amélioration du score statistique CRPS ont été testées. Ainsi, un jeu de paramètres pour chaque membre de la PEARO a été identifié. Cependant, celles-ci engendrent un biais systématique des membres de la PEARO. La réduction aux huit paramètres les plus influents a montré des résultats similaires à la version perturbant l’ensemble des paramètres incertains, suggérant un potentiel coût de réglage des modèles atmosphériques plus faible.

**Mots-Clés :** Prévisibilité, Prévision d’Ensemble, Paramétrisation, Processus de mésoschelle, Analyse de Sensibilité, Erreur de modélisation, Paramètres Perturbés

---

**Abstract** — Despite a continuous improvement of numerical weather prediction models, some forecast busts still occur due to a presence of error in models. The representation of the different origins of model uncertainty is an important aspect, in particular in Ensemble Prediction Systems (EPS). The regional Ensemble Prediction System used at Météo-France, AROME-EPS, currently represents model uncertainties through the perturbation of global output tendencies of physical parameterization. However, this method presents some disadvantages such as a difficult physical interpretation of results. Thus, this PhD-thesis aims to study more physical model error representation methods, based on the perturbation of input parameters of the physical parameterization schemes.

Following advices of parameterization experts, 21 parameters to perturb, whose values are uncertain, have been selected. Sensitivity analyses using the Morris screening and Sobol’ sensitivity indices, have led to reduce this list to eight parameters with a high impact on AROME forecasts. Several perturbed parameters techniques have then been set up and evaluated over long periods. They largely improve AROME-EPS performances for most near-surface variables including wind speed and accumulated precipitation. Different optimizations improving the statistical CRPS score have also been tested. Thus, a set of parameters have been identified for each AROME-EPS member. However, they induce a systematic bias of AROME-EPS members. Reducing the perturbation to the eight most influential parameters has shown similar results as the version perturbing the full set of parameters, suggesting a possible cheaper setting of weather prediction models.

**Keywords :** Predictability, Ensemble Prediction System, Parameterization, Mesoscale processes, Sensitivity Analysis, Model error, Perturbed Parameters

---