



**Évaluer l'intérêt de pondérer un ensemble
hétérogène de modèles de climat afin
d'améliorer les projections hydroclimatiques
futures sur le Québec méridional**

Rapport final
Juillet 2023

Évaluer l'intérêt de pondérer un ensemble hétérogène de modèles de climat afin d'améliorer les projections hydroclimatiques futures sur le Québec méridional

Rapport final

Juillet 2023

ÉQUIPE DE RÉALISATION :

Martin Leduc, Ouranos

Anne-Marie Bégin, Ouranos

Dominique Paquin, Ouranos

Sébastien Biner, Ouranos (actuellement à Hydro-Québec)

Numéro du projet Ouranos : 701000

Citation suggérée : Leduc, M., Bégin, A.-M., Paquin, D., Biner, S. (2023). Évaluer l'intérêt de pondérer un ensemble hétérogène de modèles de climat afin d'améliorer les projections hydroclimatiques futures sur le Québec méridional. Rapport présenté à la Direction de l'expertise hydrique du ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs. Ouranos, Montréal, Canada. 84 p. + annexes.

Les résultats et opinions présentés dans cette publication sont entièrement la responsabilité des auteurs et n'engagent pas Ouranos ni ses membres. Toute utilisation ultérieure du document sera au seul risque de l'utilisateur sans la responsabilité ou la poursuite juridique des auteurs.

Remerciements

Ce projet a été rendu possible grâce au soutien financier du ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP) dans le cadre du projet gouvernemental Info-Crue. Nous tenons à remercier plus particulièrement Richard Turcotte et Charles Malenfant du MELCCFP et Gabriel Rondeau-Genesse d'Ouranos pour leur soutien scientifique et technique durant le projet. Merci aussi à David Huard, Dominic Matte et Louis-Philippe Caron d'Ouranos pour avoir révisé ce rapport.

*Environnement,
Lutte contre
les changements
climatiques,
Faune et Parcs*

Québec 

InfoCrue


Ouranos

Résumé

Le projet gouvernemental Info-Crue, sous la responsabilité du ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP), vise à outiller le Québec par rapport aux risques d'inondation dans un contexte de changements climatiques. Dans le cadre de cette initiative, un ensemble de scénarios hydroclimatiques est utilisé pour évaluer les crues de récurrence 2 à 350 ans en période courante, ainsi que leur évolution pour différents horizons dans le futur (2030, 2050 et 2080). Le présent projet s'inscrit dans une démarche de R&D en soutien à Info-Crue, et a pour objectif d'évaluer l'intérêt d'appliquer des méthodes de pondération aux ensembles de modèles de climat afin d'améliorer les projections climatiques qui sont elles-mêmes à la base des scénarios hydroclimatiques.

Il est connu que les modèles de climat montrent différents niveaux de performance pour simuler le climat historique. Il est donc naturel de penser qu'une pondération ou sélection des modèles par rapport à leur performance à reproduire le climat du passé pourrait permettre d'augmenter notre confiance dans les projections climatiques, d'obtenir une meilleure représentation de l'incertitude, et potentiellement de réduire la taille de l'ensemble afin d'en simplifier l'utilisation. Pour les besoins d'Info-Crue, un ensemble de 179 simulations climatiques (en versions sorties brutes et scénarios post-traitées) a été formé à partir des ensembles CMIP5, CORDEX et ClimEx. Dans ce projet, nous utilisons cet ensemble hétérogène pour examiner et comparer plusieurs méthodes de pondération, permettant ainsi de quantifier l'impact des choix méthodologiques dans l'application de ces méthodes.

Dans la première partie de l'analyse, nous évaluons les performances des données brutes et post-traitées à reproduire les caractéristiques principales des variables de température et de précipitation du climat passé, soit les moyennes climatologiques, la variabilité interannuelle, et les tendances historiques des changements climatiques, en utilisant comme référence le jeu de données sur grille de NRCan, également utilisé comme données de référence pour le post-traitement. Par construction, celui-ci améliore les moyennes cli-

matologiques de la température et de la précipitation. Cependant, de telles améliorations ne se produisent pas au niveau de la variabilité interannuelle et des tendances historiques des changements climatiques. La seconde partie de l'analyse consiste à évaluer l'impact de différentes méthodes de pondération sur la moyenne pondérée de l'ensemble. D'abord, nous comparons diverses méthodes de sélection *a priori*, qui consistent à choisir des modèles en fonction de certaines caractéristiques telles que le type de modèle (régional ou global), sa résolution spatiale, ou selon certains paradigmes de sélection comme "un modèle un vote", ou "un centre un vote". Nous explorons ensuite une pondération *a posteriori* basée sur les performances historiques, en analysant l'impact du choix de la fonction de poids, l'ajustement de son paramètre libre, et la combinaison de plusieurs critères pour évaluer les modèles. Cette analyse démontre l'importance significative des choix méthodologiques sur le calcul des poids de l'ensemble, et par conséquent sur la moyenne pondérée de l'ensemble. Dans un dernier temps, nous évaluons la possibilité d'utiliser des métriques de sensibilité climatique globale pour contraindre les projections régionales. Nous montrons que cette approche a un impact limité pour l'ensemble CMIP5, mais suggérons qu'elle pourrait avoir un impact notable pour contraindre les projections de températures à partir de l'ensemble CMIP6, dont certains modèles sont répertoriés dans le dernier rapport du GIEC comme ayant des sensibilités climatiques trop élevées, et donc moins probables.

Nous concluons ce rapport en rappelant que le paradigme de la démocratie des modèles, aussi connu sous le nom de "un modèle un vote" est difficile à discréditer lorsque les projections climatiques sont étudiées à l'échelle régionale, ce qui est en grande partie dû à la forte variabilité interne à ces échelles. Afin de changer de paradigme, il sera important au cours des prochaines années de réaliser des travaux de recherche visant à identifier les métriques les plus adéquates pour évaluer les performances des modèles en fonction de l'application, c'est-à-dire en fonction de la nature des processus physiques impliqués. En particulier, il y aurait un potentiel pour des travaux qui cherchent à établir des ponts entre les échelles, à partir de contraintes à l'échelle globale, vers des projections climatiques régionales, pour éventuellement atteindre l'échelle hydrologique.

Table des matières

Table des figures	7
Liste des tableaux	12
1 Contexte	14
2 Introduction	17
3 Objectifs	22
4 Méthodologie	23
4.1 Description des données	23
4.2 Traitement des données et statistiques climatiques	26
4.3 Calcul des poids basés sur les performances	27
5 Résultats	30
5.1 Évaluation des performances	30
5.1.1 Cycle annuel	30
5.1.2 Biais saisonnier	33
5.1.3 Variabilité interannuelle	36
5.1.4 Tendances de changement climatique	41
5.2 Pondération a priori	44
5.2.1 Méthodes de sélection des modèles	44
5.2.2 Sensibilité du changement climatique à la méthode de sélection	47
5.3 Pondération a posteriori	49
5.3.1 Choix de la fonction de pondération	49
5.3.2 Ajustement de la fonction de pondération	52
5.3.3 Sélection des métriques de performance	60
5.3.4 Combinaison des métriques	60

5.3.5	Impact de la pondération a posteriori	67
5.3.6	Contraindre les projections régionales à l'aide de mé- triques globales	68
6	Discussion	71
7	Conclusion	75
8	Bibliographie	77
A	Estimation de la variance d'une différence entre deux mo- dèles	85
B	Figures supplémentaires	87
C	Liste des modèles utilisant le scénario RCP8.5	91
D	Relation entre l'ECS et les projections régionales sur le Qué- bec à partir de l'ensemble CMIP6	93

Table des figures

1	Classification hiérarchique de l'ensemble de simulations climatiques (ESC). Du centre vers les extrémités du diagramme : le choix du scénario RCP, du MGC, du membre MGC, et s'il y a lieu du MRC et du domaine sur lequel il a été tourné. Pour les simulations MGC utilisées dans ce projet sans mise à l'échelle dynamique par un MRC, l'attribut "globe" est utilisé pour définir le domaine. Les 179 simulations climatiques formant l'ESC sont représentées aux extrémités du diagramme et incluent CMIP5 (mauve), CORDEX (rouge; incluant des simulations supplémentaires du MRCC5 produites par Ouranos) et ClimEx (bleu).	25
2	Domaine d'extraction (gauche) et domaine d'analyse du Québec méridional (droite) pour les données de l'ESC telles que représentées sur une grille de modèle à 0.44° de résolution. . .	26
3	Courbes de poids en fonction de la différence (D) pour la fonction inverse de la différence (Éq. 1, notée INV) et la fonction exponentielle (Éq. 2, notée EXP) avec différentes valeurs pour le rayon de similarité σ_D	28
4	Cycle annuel de la précipitation (haut) et la température (bas) pour la période 1971-2000, données brutes.	31
5	Cycle annuel de la précipitation (haut) et la température (bas) pour période 1971-2000, données post-traitées.	32
6	Biais saisonnier de la précipitation pour la période 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), pour les données brutes (gauche) et post-traitées (droite).	34

7	Biais saisonnier de la température moyenne 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), pour les données brutes (gauche) et post-traitées (droite).	35
8	Série annuelle de la précipitation (haut) et la température moyenne (bas) pour 1971-2000, données brutes.	37
9	Série annuelle de la précipitation (haut) et la température moyenne (bas) pour 1971-2000, données post-traitées.	38
10	Variabilité interannuelle de la moyenne saisonnière de la précipitation en hiver (DJF), printemps (MAM), été (JJA) et automne (SON) pour 1971-2000, données brutes (gauche) et post-traitées (droite).	39
11	Variabilité interannuelle de la moyenne saisonnière de la température moyenne en hiver (DJF), printemps (MAM), été (JJA) et automne (SON) pour 1971-2000, données brutes et post-traitées.	40
12	Tendance de changement climatique sur la période 1971-2000 (colonne de gauche) et le biais par rapport aux observations (colonne de droite) pour la précipitation pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes. La tendance calculée à partir des observations est donnée par la ligne rouge.	42
13	Tendance de changement climatique sur la période 1971-2000 (colonne de gauche) et le biais par rapport aux observations (colonne de droite) pour la température moyenne pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes. La tendance calculée à partir des observations est donnée par la ligne rouge.	43
14	Nombre de simulations dans l'ensemble ayant une résolution égale ou plus fine que $1\Delta x$, $2\Delta x$ et $5\Delta x$ pour une surface donnée en km^2 .	46
15	Delta saisonnier en 2071-2100 par rapport à 1981-2010 pour différentes méthodes de pondération pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), selon RCP 8.5, données brutes.	48
16	Poids calculés selon les fonctions INV (haut) et EXP (bas) pour la température moyenne sur la période 1971-2000 pour la saison hiver (DJF), en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La fonction EXP utilise la valeur de référence σ_{D^*} pour son paramètre σ_D . Données brutes.	51

17	Poids calculés pour la moyenne climatologique de la température moyenne en hiver (DJF) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est variée selon $1\sigma_{D^*}$, $2\sigma_{D^*}$, $5\sigma_{D^*}$ et $10\sigma_{D^*}$	55
18	Poids calculés pour la variabilité interannuelle de la température moyenne en hiver (DJF) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est variée selon $1\sigma_{D^*}$, $2\sigma_{D^*}$, $5\sigma_{D^*}$ et $10\sigma_{D^*}$	56
19	Poids calculés pour la tendance historique de la température moyenne en hiver (DJF) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est variée selon $1\sigma_{D^*}$, $2\sigma_{D^*}$, $5\sigma_{D^*}$ et $10\sigma_{D^*}$	57
20	Poids calculés pour la moyenne climatologique (haut), la variabilité interannuelle (centre) et la tendance historique (bas) de la température moyenne en hiver (DJF, gauche) et été (JJA, droite) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est fixée à $3\sigma_{D^*}$. Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2.	58
21	Poids calculés pour la moyenne climatologique (haut), la variabilité interannuelle (centre) et la tendance historique (bas) des précipitations en hiver (DJF, gauche) et été (JJA, droite) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est fixée à $3\sigma_{D^*}$. Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2.	59
22	Poids calculés à partir de combinaisons de huit métriques de moyennes climatologiques pour la période 1971-2000 en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Dans l'ordre sont combinés par multiplication les poids de hiver à automne pour la température moyenne (une à quatre métriques), en ajoutant ensuite les quatre saisons pour la précipitation (cinq à huit métriques). Données brutes pour CMIP5 (bleu), CORDEX (rouge) et ClimEX (gris). Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2. $\psi_D = 1$	63

23	Poids calculés à partir de combinaisons de huit métriques de moyennes climatologiques pour la période 1971-2000 en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Dans l'ordre sont combinés par multiplication les poids de hiver à automne pour la température moyenne (une à quatre métriques), en ajoutant ensuite les quatre saisons pour la précipitation (cinq à huit métriques). Données brutes pour CMIP5 (bleu), CORDEX (rouge) et ClimEX (gris). Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2. $\psi_D = 2$	64
24	Poids calculés à partir de combinaisons de huit métriques de moyennes climatologiques pour la période 1971-2000 en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Dans l'ordre sont combinés par multiplication les poids de hiver à automne pour la température moyenne (une à quatre métriques), en ajoutant ensuite les quatre saisons pour la précipitation (cinq à huit métriques). Données brutes pour CMIP5 (bleu), CORDEX (rouge) et ClimEX (gris). Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2. $\psi_D = 4$	65
25	Poids cumulés des modèles ordonnés en ordre décroissants de leur performance, en fonction du nombre de métriques combinées de 1 à 8 (de hiver à automne pour température, et ensuite en ajoutant les quatre saisons pour la précipitation), pour des valeurs de $\psi_D = 0.5, 1, 2$ et 4. Le premier membre de chacun des 40 modèles de la table 2 est utilisé. Chaque courbe pleine correspond à la moyenne des courbes obtenues en utilisant les 49 membres résiduels de ClimEx comme des pseudoréférences. Les courbes pointillées représentent le même calcul obtenu en utilisant les observations de NRCan comme référence. La droite pointillée (noire) correspond aux poids cumulés pour des modèles à poids égaux.	66
26	Delta saisonnier en 2071-2100 par rapport à 1981-2010 pour différentes méthodes de pondération a posteriori en utilisant les observations comme référence, pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), selon RCP 8.5, données brutes. Une simulation par modèle est utilisée pour chacun des 40 modèles de la table 2.	68

27	Delta de température (haut) et précipitation (bas) entre 1980-2009 et 2070-2099 pour les quatre saisons en fonction des valeurs de ECS (gauche) et de TCR (droite), pour les données brutes des simulations suivant le RCP8.5, avec une simulation par modèle (31) de la table 2 dont les valeurs d'ECS et de TCR du MCG étaient disponibles (Flynn et Mauritsen, 2020). . . .	70
28	Biais par rapport aux observations dans la tendance de changement climatique de la précipitation sur la période 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes.	88
29	Biais par rapport aux observations dans la tendance de changement climatique de la température moyenne sur la période 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes.	89
30	Delta saisonnier 2071-2100 différents ensembles pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), selon RCP 8.5, données brutes.	90
31	Relation entre ECS et TCR pour 27 modèles globaux de CMIP5. Les données proviennent de Flynn et Mauritsen (2020).	90
32	Delta de température (haut) et précipitation (bas) sur la province de Québec entre 1991-2020 et 2071-2100 pour les quatre saisons en fonction des valeurs d'ECS, en utilisant 21 modèles de CMIP6 pour le scénario SSP370 (ensemble ESPO-G6-E5L d'Ouranos, Lavoie <i>et al.</i> (2023)). Les valeurs d'ECS proviennent de Hausfather <i>et al.</i> (2022). Crédits : Juliette Lavoie, Ouranos.	94

Liste des tableaux

1	Valeurs de Γ (Éq. 7) représentant le rapport entre l'écart-type de l'ensemble ClimEx et l'écart-type multi-modèle (un membre par modèle, voir Tab. 2) pour deux variables (tmoy et pr), trois statistiques climatiques (moyenne climatologique (climato.), variabilité interannuelle (VIA) et tendance) et quatre saisons.	54
2	Liste des modèles et du nombre de membres disponibles pour le scénario RCP8.5.	92

Liste des abréviations, sigles et acronymes

CORDEX	Coordinated Regional Climate Downscaling Experiment
CMIP5/6	Model Intercomparison Project Phase 5 and 6
DEH	Direction de l'expertise hydrique
ECS	Equilibrium Climate Sensitivity
ESC	Ensemble de simulations climatiques
EXP	Fonction exponentielle
GES	Gaz à effet de serre
GIEC	Groupe d'experts intergouvernemental sur l'évolution du climat
INV	Fonction inverse
MELCCFP	Ministère de l'Environnement, de la Lutte contre les changements climatiques, des Forêts et des Parcs
MGC	Modèle global du climat
MRC	Modèle régional du climat
MRCC5	Modèle régional canadien du climat version 5
RCP	Representative Concentration Pathway / Profils représentatifs d'évolution de concentration
REA	Reliability Ensemble Averaging / Moyenne d'ensemble de fiabilité?
NRCan	Natural Resources Canada
TCR	Transient climate response / Réponse climatique transitoire
UQAM	Université du Québec à Montréal

Chapitre 1

Contexte

Le projet gouvernemental Info-Crue, sous la responsabilité du ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP), vise à outiller le Québec par rapport aux risques d'inondation. Dans ce contexte, Ouranos a été mandaté pour réaliser de la recherche et du développement en soutien direct à Info-Crue, et ce pour différents volets. L'un de ces volets porte sur l'évolution du climat pour des besoins en cartographie des zones inondables. Le plan de travail de ce volet prévoit répondre à l'un des besoins à court terme de la Direction de l'expertise hydrique (DEH), soit l'obtention de scénarios climatiques fiables qui seront utilisés pour simuler les crues de récurrence 2 à 350 ans en période de référence et pour différents horizons dans le futur (2030, 2050 et 2080). Ces scénarios sont aussi utilisés dans le cadre de la production par la DEH de l'Atlas hydroclimatique du Québec qui répondra directement aux besoins d'Info-Crue.

Or, prévoir l'évolution des crues dans un contexte de changements climatiques est un problème d'une grande complexité, nécessitant la mise en place d'une chaîne de modélisation allant du système climatique à l'échelle globale jusqu'aux impacts hydrologiques à l'échelle locale (Olsson *et al.*, 2016). Les résultats d'une telle chaîne dépendent de plusieurs choix méthodologiques qui se traduisent en incertitude dans les projections futures, autant d'un point de vue climatique (Hawkins et Sutton, 2009, 2011; Déqué *et al.*, 2007) qu'hydrologique (Giuntoli *et al.*, 2018). Les différentes sources d'incertitude correspondant aux principales étapes de la chaîne de modélisation des impacts des changements climatiques sur l'hydrologie comprennent 1) l'augmentation des concentrations atmosphériques de gaz à effet de serre (GES) dans le futur (scénarios RCP; Meinshausen *et al.* 2011), 2) la sensibilité cli-

matique des modèles globaux de climat (MGC), 3) la variabilité interne (ou naturelle) du climat (échantillonnée par de multiples réalisations d'un même MGC), 4) la réponse régionale des modèles régionaux de climat (MRC), qui inclut la variabilité interne des modèles régionaux, 5) la technique de post-traitement (et les données de référence utilisées pour ce post-traitement) appliquée aux simulations et 6) la réponse des modèles hydrologiques (sensibilité aux différentes constantes, paramétrages pour un même modèle, ainsi qu'aux différents modèles). L'étape #4 (parfois remplacée par une mise à l'échelle statistique) est préférable afin de réduire le saut d'échelle important entre les processus simulés par les MGC et le modèle hydrologique (Fowler *et al.*, 2007). L'étape #5, bien que sujette à des débats dans la communauté hydrologique (Chen *et al.*, 2017), est une pratique courante afin de réduire les biais des simulations climatiques par rapport aux observations, permettant ainsi de piloter le modèle hydrologique avec des intrants climatiques similaires aux observations utilisées pour calibrer les modèles hydrologiques (Muerth *et al.*, 2013).

Dans le contexte d'Info-Crue, un ensemble tenant compte des sources d'incertitude énumérées précédemment a été développé. Dans un premier temps, Ouranos a rassemblé 179 simulations de modèles de climat globaux et régionaux et créé un jeu de données tenant compte des sources d'incertitude #1 à 4 (voir rapport de Rondeau-Genesse et Braun 2020). La version post-traitée de cet ensemble de simulations climatiques a ensuite été utilisée par la DEH pour piloter le modèle hydrologique HYDROTEL sous différentes configurations (plateformes), permettant d'évaluer la source d'incertitude #6. Cet ensemble compte donc 179 simulations hydrologiques (une par simulation climatique) produites à l'aide du modèle HYDROTEL sur environ 28 000 tronçons de rivières du Québec méridional. Ces simulations couvrent la période 1955-2100 afin de simuler l'effet des changements climatiques sur les processus hydrologiques. Une première version de l'ensemble de simulations hydrologiques avait été produite et évaluée sur le bassin versant du Lac Champlain (Lucas-Picher *et al.*, 2020, 2021).

Dans le cadre du présent projet, nous évaluons l'intérêt et la faisabilité de pondérer l'ensemble de simulations climatiques utilisé par la DEH dans le contexte d'Info-Crue afin d'améliorer les projections hydroclimatiques futures sur le Québec méridional. Il est à noter que ce projet se concentre exclusivement sur la pondération de l'ensemble de simulations climatiques, tandis qu'un projet distinct (voir rapport Rondeau-Genesse *et al.* 2021) a exploré l'enjeu de la pondération de l'ensemble hydrologique. Ces deux projets sont complémentaires et permettent d'explorer l'enjeu de la pondération sous des angles différents.

Les modèles de l'ensemble climatique seront d'abord évalués en fonction de leur capacité à reproduire le climat historique. Ensuite, diverses approches de pondération seront comparées, d'abord en utilisant des critères de sélection a priori (p. ex. par rapport à la résolution des modèles de climat), et ensuite, en fonction de la performance des modèles, soit une pondération a posteriori. Ce rapport est structuré comme suit. L'introduction (Sect. 2) présente une revue des travaux associés aux méthodes de pondération des ensembles de modèles de climat, suivie d'une description des objectifs du projet (Sect. 3), de la méthodologie utilisée (Sect. 4), de la présentation des résultats (Sect. 5), de la discussion (Sect. 6) et de la conclusion (Sect. 7).

Chapitre 2

Introduction

L'analyse et la compréhension des changements climatiques sont en grande partie basées sur des simulations numériques générées à l'aide de modèles de climat. Ces modèles, développés par quelques dizaines de centres de recherche à travers le monde depuis plusieurs décennies, sont tous basés sur les équations fondamentales de la physique (équations de Navier-stokes), diffèrent par la formulation de leurs composantes ou dans l'ajustement de paramètres clés comme leur résolution spatiale. Afin d'estimer l'ampleur des changements climatiques qui sont attendus dans le futur, les scientifiques du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC) utilisent des ensembles composés de plusieurs dizaines de simulations produites par tous les modèles afin d'associer une mesure d'incertitude aux projections climatiques futures (Taylor *et al.*, 2012; O'Neill *et al.*, 2016).

Ces ensembles sont souvent interprétés comme des “ensembles d'opportunités” (Tebaldi et Knutti, 2007), c'est-à-dire que l'échantillonnage des simulations de modèles de même génération dépend principalement de leur disponibilité, qui elle est fonction des ressources et intérêts de chaque centre de recherche à contribuer par un certain nombre de réalisations, de modèles et de scénarios d'émissions de gaz à effet de serre (GES). Ces ensembles d'opportunités ont le principal avantage de regrouper de nombreuses simulations provenant des principaux modèles développés selon l'état de l'art, et ainsi d'offrir une grande diversité de projections climatiques possibles. Ce type d'ensemble n'est toutefois pas calibré ni en fonction de la crédibilité des modèles, ni d'une représentation optimale des trois principales sources d'incertitude affectant les projections climatiques (Hawkins et Sutton, 2009; Déqué *et al.*, 2007); soit l'incertitude due aux émissions futures de GES, l'incertitude modèle et la variabilité interne du climat. Il en découle que

toute interprétation probabiliste de ces ensembles de projections climatiques ne peut être établie directement (Tebaldi et Knutti, 2007). En l’absence de consensus scientifique sur une manière de contourner ce problème, l’approche utilisée par défaut dans les cinq premiers rapports d’évaluation du GIEC (p. ex. IPCC 2013, 2007) était celle de la “démocratie des modèles” (un modèle un vote), où tous les modèles de même génération sont considérés comme des représentations équivalentes (c.-à-d. équiprobables) du système climatique.

Dans un contexte de prise de décision, il est justifié d’attribuer des poids supérieurs aux modèles climatiques ayant la plus grande crédibilité pour des projections futures. Parmi les premiers travaux effectués en ce sens, la méthode “Reliability Ensemble Averaging (REA)” (Giorgi et Mearns, 2002; Tebaldi *et al.*, 2005) consiste à pondérer les projections par rapport à deux métriques, soit la performance des modèles en matière de représentation du climat présent et la convergence (ou consensus) des projections futures. La valeur du poids de performance associée à chaque modèle y est définie selon l’inverse du biais de sa moyenne climatologique par rapport aux observations, c’est à dire que le poids de ce modèle sera diminué par rapport aux autres si son biais est plus grand. Dans le projet Ensembles (van der Linden et Mitchell, 2009), Christensen *et al.* (2010) a utilisé plusieurs métriques en lien avec différents processus d’intérêt appliqués au domaine européen comme les patrons de circulation à grande échelle, la variabilité spatiale dans la méso-échelle, les distributions de probabilités de différentes variables, les tendances historiques, le cycle annuel et les extrêmes. À noter que toutes ces métriques, exception faite de celle sur la circulation, ne considèrent que les variables de température de surface et précipitation totale. En combinant ainsi plusieurs métriques représentant différents aspects du système climatique (p. ex. variables climatiques, saisons et régions), les performances générales des modèles ont été évaluées pour pondérer les projections climatiques (van der Linden et Mitchell, 2009). L’idée de combiner un grand nombre de métriques de performance afin d’identifier les “meilleurs modèles” peut toutefois dissimuler une certaine complexité dans les forces et faiblesses des différents modèles (Gleckler *et al.*, 2008; Christensen *et al.*, 2010). Tel que soulevé par Sanderson *et al.* (2017), il est peu probable qu’un modèle offre de bonnes performances dans un grand nombre de métriques, et donc de combiner plusieurs métriques de performance peut avoir comme effet de diminuer l’impact de la pondération. D’un autre côté, il peut être justifié de mettre une emphase particulière sur un plus petit nombre de métriques, mais qui soient plus pertinentes pour une application donnée (Christensen *et al.*, 2010).

Dans ce contexte, c’est tout récemment que le GIEC a utilisé pour la

toute première fois dans son plus récent rapport d'évaluation (IPCC, 2021) des méthodes de pondération. Il a en effet été reconnu (Hausfather *et al.*, 2022) que certains modèles de l'ensemble CMIP6 ont des sensibilités climatiques trop élevées par rapport à ce que suggèrent plusieurs évidences observationnelles et théoriques (Sherwood *et al.*, 2020). Cette approche se base sur le fait qu'il existe des liens physiques entre certaines caractéristiques du climat présent (prédicteurs) et celles projetées dans le futur (prédicants), justifiant ainsi l'application de contraintes observationnelles (ou "contraintes émergentes", voir Hall *et al.* 2019) pour réduire l'incertitudes des projections. Dans Lee *et al.* (2021), il a été possible de réduire les incertitudes dans les projections pour certaines variables climatiques comme la moyenne globale de la température de l'air près de la surface et la glace de mer. Plusieurs lignes d'évidence ont été utilisées afin de contraindre les projections globales, notamment des estimés améliorés de la sensibilité du système climatique réel obtenues à partir de données paléoclimatiques ou d'une meilleure compréhension des processus régissant la sensibilité climatique (Forster *et al.*, 2021). Pour la moyenne globale de la température, les projections provenant des modèles globaux de climat (MGC) de CMIP6 ayant les plus grandes sensibilités climatiques (c.-à-d. se réchauffant le plus rapidement) sont considérées comme étant très improbables à cause de leurs faibles performances à reproduire certaines caractéristiques du climat historique (Brunner *et al.*, 2020; Tokarska *et al.*, 2020). Pour la glace de mer, dont le signal de changement climatique dépend fortement de l'état initial (Knutti *et al.*, 2017), une sélection des meilleurs modèles permet d'augmenter notre niveau de confiance qu'un premier mois de septembre avec moins d'un million de km² de glace en Arctique adviendra avant 2050 (Notz et Community, 2020), la moyenne non-pondérée de l'ensemble de projections fournissant une occurrence plus tardive au cours du 21e siècle (Senftleben *et al.*, 2020).

Aux échelles régionale et locale, les avantages d'appliquer des contraintes observationnelles aux projections climatiques sont encore à établir étant donné la prédominance de la variabilité interne du climat. Cette approche a cependant démontré un potentiel dans les régions de l'Arctique où on détecte un retrait progressif de la glace de mer. Borodina *et al.* (2017a) a utilisé trois métriques d'évaluation des modèles en climat récent (le couvert de glace moyen ainsi que la variabilité interannuelle de la glace de mer et de la température) pour contraindre les projections futures de la variabilité interannuelle de la température. De manière similaire, Borodina *et al.* (2017b) a montré qu'en contraignant les modèles par rapport à leur représentation de la distribution historique de la température, il était possible de réduire l'incertitude des projections pour les extrêmes chauds de température sur

certaines régions comme l’Australie, l’Amérique centrale et le nord de l’Asie, mais sans résultat probant pour l’est de l’Amérique du Nord. Dans un même ordre d’idées, Lorenz *et al.* (2018) a montré qu’il existe certaines métriques prometteuses pour contraindre les projections des extrêmes chauds de température, dont les précipitations moyennes ainsi que la variabilité et la tendance de la température maximale journalière, mais que les lacunes dans la qualité des données d’observations viennent restreindre l’application de contraintes émergentes sur les projections régionales. Il est à noter que d’autres études permettent d’établir des liens entre prédicant et prédicteur, notamment Eum *et al.* (2016), où des corrélations significatives entre le biais de température en climat historique et les changements futurs pour la région sud du Québec, ont été trouvées en utilisant des modèles régionaux de climat (MRC). La relation entre le biais en climat présent et les projections en climat futur a aussi été mise en évidence à l’aide d’ensembles de MGC par Whetton *et al.* (2007) et Christensen et Boberg (2012).

Bien qu’indispensable, l’information sur la performance des modèles à reproduire le climat du passé s’avère difficile à utiliser pour contraindre les projections climatiques. Toutefois, d’autres avenues peuvent être considérées pour pondérer les projections. Par exemple, la métrique de convergence des projections futures tel que défini par Giorgi et Mearns (2002) accorde moins de poids à un modèle qui s’éloigne du consensus des autres modèles. Cette approche a donc pour effet de réduire l’influence des modèles se situant dans les queues de distributions sous prétexte qu’un fort consensus entre modèles devrait augmenter notre niveau de confiance dans les résultats des projections futures. Or la notion de convergence est parfois aussi interprétée comme le reflet d’un faux consensus dû aux liens de dépendances structurelles existant entre les modèles (Pirtle *et al.*, 2010). Les modèles de climat sont en effet à divers degrés dépendants les uns par rapport aux autres puisque développés par des institutions différentes mais qui partagent des connaissances sur la modélisation des processus climatiques, du code informatique, ainsi que des protocoles communs pour évaluer et améliorer les modèles (Masson et Knutti, 2011; Knutti *et al.*, 2013). Une conséquence directement reliée à ces dépendances structurelles entre modèles est que ces derniers peuvent partager certains types d’erreurs et donc induire des biais systématiques dans les projections (Knutti *et al.*, 2010). Il en découle aussi que la métrique de convergence comme proposé par Giorgi et Mearns (2002) tend à réduire exagérément l’incertitude, et donc à établir une surconfiance dans les résultats, ce qui peut ultimement conduire à une maladaptation aux changements climatiques.

Tel qu’illustré par Abramowitz *et al.* (2019), il n’y a aucune approche

communément acceptée pour traiter le problème de la dépendance des modèles de climat, mais quelques alternatives existent dans la littérature. Une méthode simple mais intuitive est celle de la démocratie institutionnelle (Leduc *et al.*, 2016), c’est-à-dire “un centre, un vote”, plutôt que l’approche classique de “un modèle, un vote” dans le calcul des statistiques d’ensemble. Cette approche est basée sur le fait que les modèles développés par un même centre de recherche tendent à avoir des caractéristiques communes, et donc qu’on devrait mettre moins d’emphase sur un consensus entre modèles d’une même famille. La démocratie institutionnelle a aussi été appliquée au niveau des composantes des modèles (p. ex. atmosphère, océan, glace de mer) par Boé (2018). Une autre approche consiste plutôt à pondérer les modèles par rapport à une mesure de dépendance basée sur la corrélation entre les biais des différents modèles (e.g. Sanderson *et al.* 2015; Dalelane *et al.* 2018). Bien que la dépendance des modèles de climat et leurs performances soient deux idées distinctes, Bishop et Abramowitz (2013) a démontré que la corrélation des erreurs par rapport aux observations est directement liée à la notion de dépendance des modèles.

L’approche de pondération développée par Sanderson *et al.* (2015) va en ce sens en utilisant des poids qui sont le produit de deux composantes : 1) la performance du modèle en climat historique, et 2) le caractère “ unique ” de ce modèle par rapport aux autres dans l’ensemble. En appliquant cette méthode, un modèle obtiendra un poids plus faible s’il montre une forte différence par rapport au climat observé, et si l’information fournie par ce dernier est fortement dupliquée dans l’ensemble. Cette méthode a été appliquée dans plusieurs études récentes (e.g. Knutti *et al.* 2017; Sanderson *et al.* 2017; Lorenz *et al.* 2018; Dalelane *et al.* 2018; Zhao *et al.* 2022). La métrique d’unicité utilisée par cette approche permet également une certaine flexibilité afin d’intégrer des ensembles multi-membres (comme le projet ClimEx; Leduc *et al.* 2019) aux ensembles multi-modèles (Merrifield *et al.*, 2020).

Chapitre 3

Objectifs

Les motivations pour pondérer un ensemble de modèles de climat sont d'augmenter notre confiance dans les projections climatiques, d'obtenir une meilleure représentation de l'incertitude, et potentiellement de réduire la taille de l'ensemble afin d'en simplifier l'utilisation. Le présent projet a pour objectif d'évaluer la faisabilité et l'intérêt d'appliquer une pondération à l'ensemble de simulations climatiques utilisé par la DEH dans le contexte d'Info-Crue. La réalisation de cet objectif repose sur les sous-objectifs suivants :

1. Évaluer la capacité des modèles de climat (simulations brutes et scénarios post-traités) à reproduire les principales caractéristiques du climat historique.
2. Identifier des métriques d'évaluation permettant de discriminer les performances des modèles.
3. Analyser l'effet de combiner plusieurs métriques dans un même schéma de pondération.
4. Quantifier l'impact des choix méthodologiques de la pondération sur les projections climatiques

Chapitre 4

Méthodologie

4.1 Description des données

L'ensemble de simulations climatiques (ESC) utilisé pour ce projet est décrit dans le rapport de Rondeau-Genesse et Braun (2020). Cet ensemble est composé de trois ensembles de simulations provenant de projets internationaux et connus sous les noms de CORDEX, CMIP5 et ClimEx. L'ESC est composé de simulations qui utilisent deux scénarios d'émissions décrivant l'évolution des concentrations atmosphériques de gaz à effet de serre dans le futur, soit 116 simulations pour le scénario RCP8.5 et 63 pour le RCP4.5. Cet ensemble, hautement hétérogène, est constitué de simulations de différentes résolutions passant de ClimEx à 0.11° aux simulations des modèles régionaux de CORDEX sur des grilles à 0.22° et 0.44° , jusqu'aux modèles globaux de CMIP5 à plus basse résolution (entre 0.75° et 3.75°). Le schéma de la figure 1 montre une classification hiérarchique de l'ESC, où nous voyons les deux branches de scénarios d'émissions RCP à partir du centre se ramifier selon les modèles MGC utilisés, la réalisation du MGC (les réalisations diffèrent par des conditions initiales seulement), le nom du MRC piloté, et le domaine ou la résolution du MRC. À noter que les sorties de certains MGC sont parfois utilisées directement, de sorte que les simulations en pourtour du schéma peuvent provenir autant d'une simulation globale que régionale.

L'ESC illustré est disponible en deux versions, soit les sorties brutes des modèles ainsi que celles ayant subi un post-traitement. Le post-traitement utilisé (voir Rondeau-Genesse et Braun 2020) se base sur la méthode de correction univariée par quantile en utilisant le jeu de données de référence produit par Ressources naturelles Canada (NRCan Canada Daily ; Hutchinson *et al.* 2009; McKenney *et al.* 2011). Ces données, ci-après nommées “données

NRCan” ou observations, sont des observations aux stations interpolées à l’aide de splines sur une grille à 10 km de résolution. Ces données sont également utilisées comme référence pour évaluer les simulations climatiques dans le cadre du présent rapport. Par construction, la méthode de post-traitement ne tente pas de corriger les tendances de changements climatiques sur la période historique, et préserve ces tendances simulées selon la sensibilité climatique des différents modèles. Au niveau du traitement des extrêmes, la méthode inclut un post-traitement paramétrique des queues de distributions (Roy *et al.*, 2023).

Les variables disponibles dans cet ensemble sont les températures minimales (tasmin) et maximales (tasmx) journalières de l’air à 2m ainsi que la précipitation totale (pr). Afin d’alléger l’analyse, ce rapport se concentre sur la température moyenne journalière (tmoy, obtenue par la moyenne de tasmin et tasmx). Lors de la préparation de cet ensemble, les données de toutes les simulations ont été extraites sur la province de Québec (Fig. 2). Les comparaisons sont faites sur les grille de chaque modèle, les données de référence ayant été interpolées au préalable.

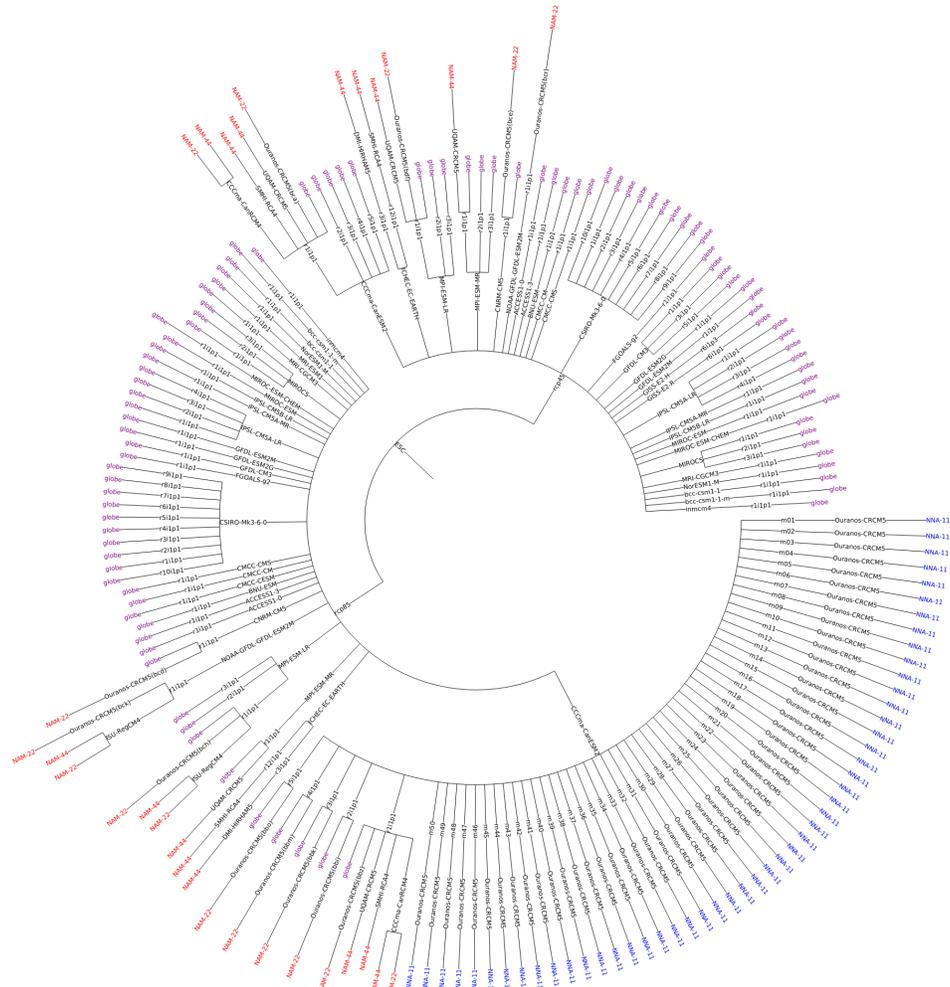


FIGURE 1 – Classification hiérarchique de l’ensemble de simulations climatiques (ESC). Du centre vers les extrémités du diagramme : le choix du scénario RCP, du MGC, du membre MGC, et s’il y a lieu du MRC et du domaine sur lequel il a été tourné. Pour les simulations MGC utilisées dans ce projet sans mise à l’échelle dynamique par un MRC, l’attribut “globe” est utilisé pour définir le domaine. Les 179 simulations climatiques formant l’ESC sont représentées aux extrémités du diagramme et incluent CMIP5 (mauve), CORDEX (rouge ; incluant des simulations supplémentaires du MRCC5 produites par Ouranos) et ClimEx (bleu).

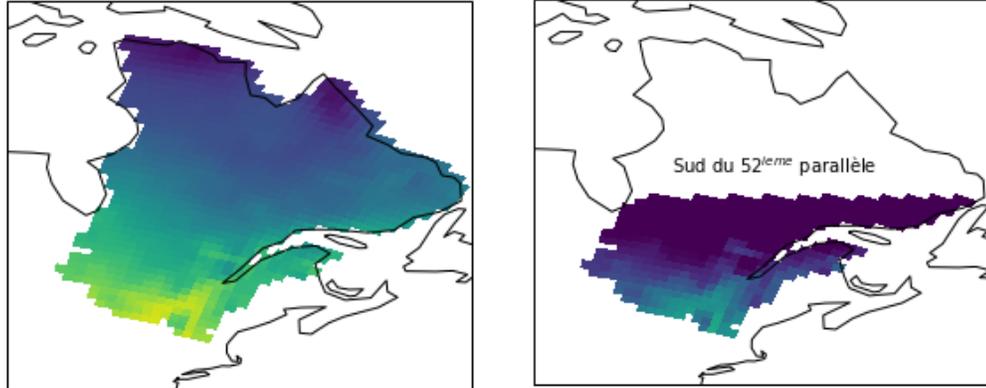


FIGURE 2 – Domaine d’extraction (gauche) et domaine d’analyse du Québec méridional (droite) pour les données de l’ESC telles que représentées sur une grille de modèle à 0.44° de résolution.

4.2 Traitement des données et statistiques climatiques

Les données brutes et post-traitées de chaque simulation ont été traitées en suivant les étapes suivantes. D’abord, les données sont extraites sur la région du Québec méridional (voir Fig. 2) pour ensuite être moyennées spatialement. Pour les différents ensembles de simulations (ClimEx, CORDEX, CMIP5) ainsi que pour les observations sur grille, les statistiques climatiques sont calculées sur 30 ans en utilisant les séries de moyennes annuelles ou saisonnières pour la période de 1971 à 2000¹. Pour chaque série, la moyenne climatologique, la variabilité interannuelle (écart-type temporel) et la tendance historique de changement climatique (régression linéaire) sont calculées. Le signal de changement climatique provient de la différence entre les périodes future 2071-2100 et de référence 1981-2010, sous forme de delta (différence absolue) pour la température et de changement relatif (%) pour la précipitation. L’analyse se concentre sur les simulations suivant le scénario RCP8.5, puisque ce dernier contient un plus grand nombre de simulations (voir la liste des 40 modèles utilisant le scénario RCP8.5 à la table 2).

1. Les années 1995-1996 sont omises de l’analyse à cause de données manquantes dans les données de référence utilisées pour le post-traitement.

4.3 Calcul des poids basés sur les performances

Il existe différentes façons de calculer le poids (w_i) des simulations en fonction de diverses mesures de performance, ou de la capacité à reproduire les observations. Dans ce rapport, deux fonctions de pondération sont comparées. La première consiste à convertir une mesure de performance en poids à l'aide d'une fonction inverse (notée INV) inspirée de la méthode "Reliability Ensemble Averaging" (Giorgi et Mearns, 2002; Tebaldi *et al.*, 2005) et décrite par l'équation suivante :

$$w_i = \frac{1}{|D_i|} \quad (1)$$

où D_i , la mesure de performance, représente la différence entre la simulation i et les observations pour une statistique climatique donnée, telles la moyenne climatologique ou la pente de la tendance. Une grande différence pour la simulation i correspond donc à un poids plus petit pour cette dernière.

La seconde méthode est basée sur une fonction exponentielle (notée EXP) qui s'inspire de travaux plus récents (Sanderson *et al.*, 2015, 2017; Knutti *et al.*, 2017; Lorenz *et al.*, 2018; Brunner *et al.*, 2020) et est décrite par l'équation suivante :

$$w_i = e^{-\left(\frac{D_i}{\sigma_D}\right)^2} \quad (2)$$

où σ_D est un paramètre libre connu sous le nom de "rayon de similarité" (Sanderson *et al.*, 2015). Sa valeur correspond au seuil des différences D_i en deçà duquel deux simulations, ou une simulation et les observations, sont considérées similaires. Puisque différentes approches existent afin d'attribuer une valeur à σ_D (p. ex. Brunner *et al.* 2020; Merrifield *et al.* 2020), nous évaluerons l'impact du choix de ce paramètre à la section 5.3.2. Pour ce faire, nous proposons d'abord de définir une valeur de référence (notée $\sigma_{D\star}$) basée sur la variabilité interne du climat. La variabilité interne étant une caractéristique intrinsèque du système climatique, cette approche permet de donner un sens physique au rayon de similarité. Afin de calculer cette valeur, l'ensemble ClimEx est tout indiqué puisqu'il permet de calculer directement l'incertitude due à la variabilité interne (provenant des MCG) pour n'importe quelle statistique climatique. Nous posons ici l'hypothèse que la variabilité interne de l'ensemble ClimEx est représentative des autres modèles ainsi que du système climatique réel. Il a d'ailleurs été montré par von Trentini *et al.* (2020) que la variabilité interannuelle simulée par l'ensemble ClimEx est comparable à d'autres grands ensembles régionaux et aux observations sur

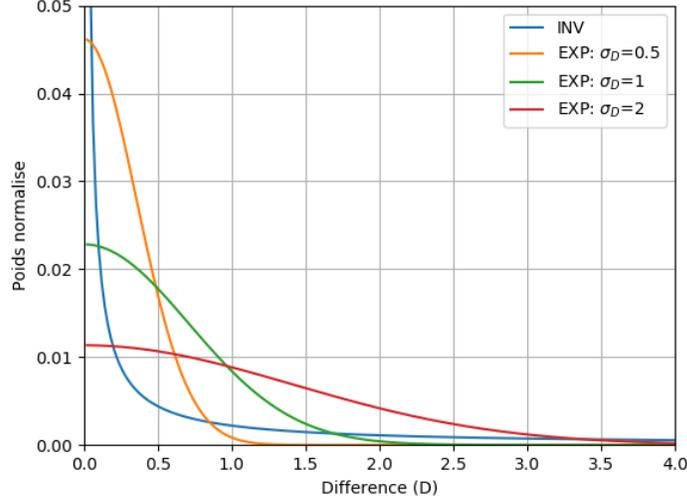


FIGURE 3 – Courbes de poids en fonction de la différence (D) pour la fonction inverse de la différence (Éq. 1, notée INV) et la fonction exponentielle (Éq. 2, notée EXP) avec différentes valeurs pour le rayon de similarité σ_D .

le domaine de l’Europe. Le σ_{D^*} sera donc défini comme l’écart-type de la différence entre des paires de membres de ClimEx, ce qui équivaut à la racine carrée de deux multipliée par l’écart-type de la distribution des membres (voir annexe A pour démonstration). Ainsi, nous définissons la valeur de référence comme :

$$\sigma_{D^*} = \sqrt{2} \sigma_{IV} \quad (3)$$

où σ_{IV} est l’écart-type entre les membres de l’ensemble ClimEx pour une statistique climatique donnée (p. ex. moyenne climatologique). La figure 3 montre des exemples de courbes de poids obtenues à partir des équations (1) et (2), où différentes valeurs de rayon de similarité sont utilisées pour la méthode EXP. Plus le rayon de similarité est grand, plus la fonction de pondération attribuera des poids uniformes entre les modèles.

Il est possible de généraliser l’équation (2) en combinant plusieurs métriques d’évaluation. Si nous redéfinissons le terme D_i comme une distance généralisée selon K dimensions (soit la différence entre la simulation i et les observations pour K métriques d’évaluation), ce terme peut s’écrire comme :

$$\frac{D_i^2}{\sigma_D^2} = \sum_{k=1}^K \frac{D_i^{(k)2}}{\sigma_D^{(k)2}} \quad (4)$$

où $\sigma_D^{(k)}$ est le rayon de similarité associé à chaque métrique. En remplaçant (4) dans (2), la transformation vers une distance généralisée équivaut à combiner par multiplication les poids de l'équation (2) pour K métriques comme :

$$w_i = e^{\frac{D_i^{(1)2}}{\sigma_D^{(1)2}}} \times e^{\frac{D_i^{(2)2}}{\sigma_D^{(2)2}}} \times \dots \times e^{\frac{D_i^{(K)2}}{\sigma_D^{(K)2}}} \quad (5)$$

où w_i doit être renormalisé avant d'être appliqué à un ensemble de modèles de climat.

Afin de combiner de manière équitable plusieurs métriques à l'aide de l'équation (4), Knutti *et al.* (2017) suggèrent d'abord de normaliser les $D_i^{(k)}$ en utilisant la distance médiane entre les différents modèles et la référence, et par la suite d'appliquer un rayon de similarité qui soit commun à toutes les métriques. En utilisant l'approche du modèle parfait pour pondérer les projections climatiques, Lorenz *et al.* (2018) montrent qu'une telle normalisation ainsi qu'un rayon de similarité d'environ 0.5 font en sorte que la projection future dans la référence tombera dans l'intervalle de confiance entre les 10^e et 90^e percentiles 80% du temps. Dans le contexte du présent rapport où différents rayons de similarité sont comparés, nous choisissons d'intégrer le facteur de normalisation (distance médiane) à notre définition du $\sigma_D^{(k)}$, celui-ci pouvant être redéfini comme :

$$\sigma_D^{(k)} = \psi_D \times D_{50}^{(k)} \quad (6)$$

où $\psi_D \approx 0.5$ représente le rayon de similarité (noté σ_D dans Knutti *et al.* 2017 et Lorenz *et al.* 2018), et $D_{50}^{(k)}$ la médiane des différences venant normaliser la métrique le $D_i^{(k)}$. Différentes valeurs de ψ_D seront utilisées dans l'analyse de la combinaison des métriques à la Sect. 5.3.4.

Chapitre 5

Résultats

5.1 Évaluation des performances

Dans cette section, les capacités des modèles à reproduire les principales caractéristiques du climat historique pour les variables de base (température de l'air près de la surface et précipitations totales), soit le cycle annuel, les moyennes saisonnières, la variabilité interannuelle et les tendances, sont évaluées. Tant les simulations brutes que post-traitées sont comparées aux observations.

5.1.1 Cycle annuel

Pour débiter l'analyse, nous comparons aux données NRCan le cycle annuel pour un sous-ensemble de simulations afin d'alléger les figures, soit en utilisant 10 simulations de CMIP5 et CORDEX et les 50 membres de ClimEx. Les 10 simulations ont été choisies afin de représenter un échantillonnage de l'effet du MGC, de la différence entre deux résolutions et de la différence entre deux MRCs. Sur la figure 4, le cycle annuel des données brutes montre des écarts par rapport aux observations pour les moyennes mensuelles des précipitations et de la température. Les mêmes cycles pour les données post-traitées (Fig. 5) sont plus semblables aux observations, montrant l'efficacité de la correction par le post-traitement. Les différences entre les courbes post-traitées est de l'ordre de la variabilité interne représentée par l'écart entre les membres de l'ensemble ClimEx.

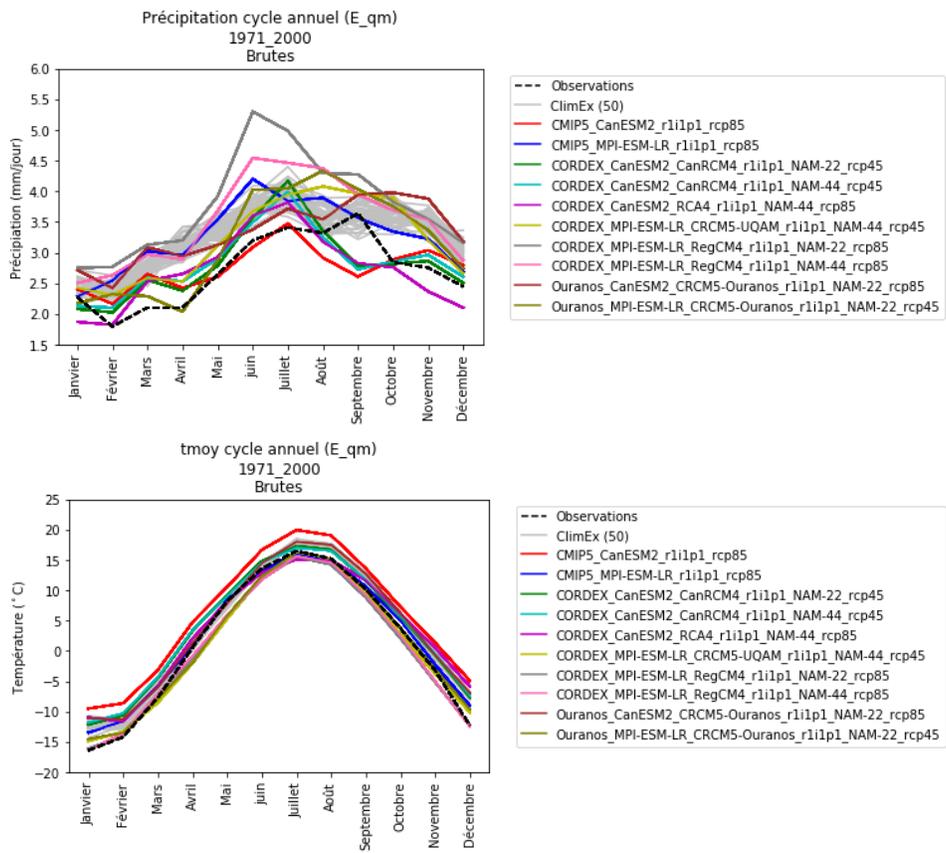


FIGURE 4 – Cycle annuel de la précipitation (haut) et la température (bas) pour la période 1971-2000, données brutes.

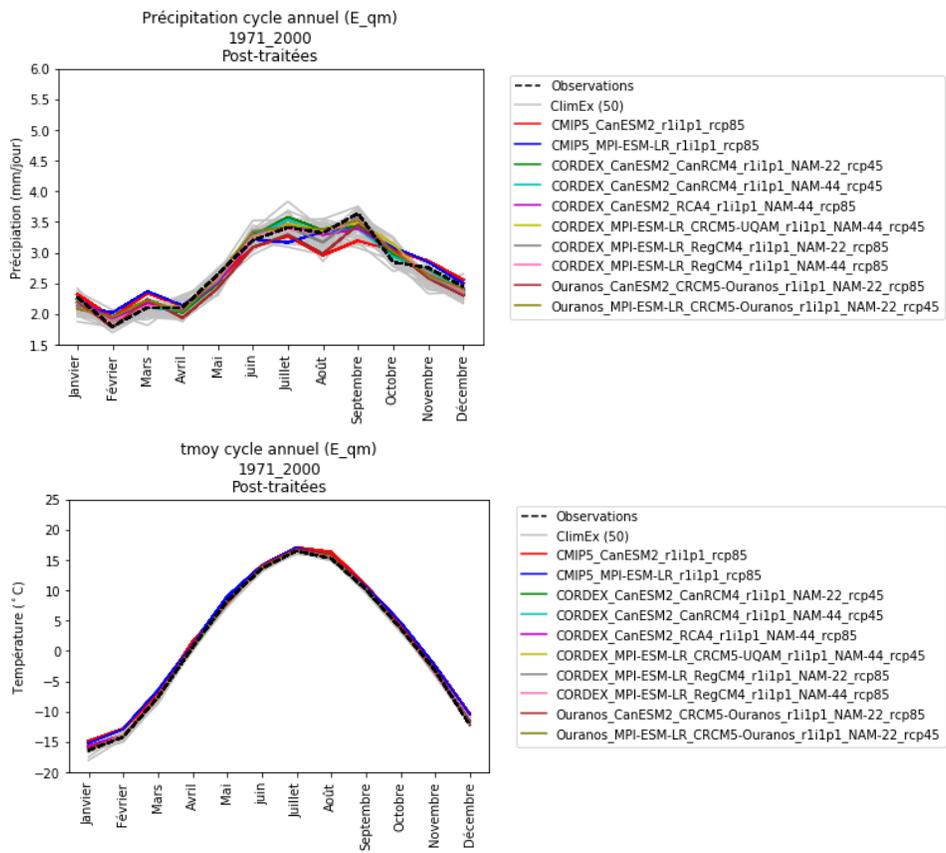


FIGURE 5 – Cycle annuel de la précipitation (haut) et la température (bas) pour période 1971-2000, données post-traitées.

5.1.2 Biais saisonnier

Le biais saisonnier du cycle annuel montre l'écart entre les simulations et les observations en période de référence. Les résultats sont présentés pour toutes les simulations historiques de l'ESC en séparant les différents ensembles (ClimEx, CORDEX, CMIP5) pour la précipitation (Fig. 6) et la température moyenne (Fig. 7).

Pour la précipitation, le biais est généralement positif pour les données brutes et près de zéro pour les données post-traitées. L'étendue des distributions des erreurs pour les données brutes représente l'incertitude due à l'utilisation de différents modèles climatiques pour les ensembles CORDEX et CMIP5, tandis que l'étendue de la distribution de ClimEx montre l'ampleur de la variabilité interne du climat telle que simulée par cet ensemble. Pour les données post-traitées, l'étendue de la distribution de ClimEx est préservée après l'application du post-traitement, celui-ci étant fait à l'échelle de l'ensemble et non des simulations individuelles. Les distributions post-traitées de CORDEX et CMIP5 deviennent similaires à celle de ClimEx après post-traitement. Les mêmes conclusions s'appliquent à la température moyenne.

En comparant l'étendue des distributions de biais obtenues dans les climatologies saisonnières entre les simulations de ClimEx avec l'étendue des biais provenant de différents modèles, ces résultats permettent de conclure que l'incertitude entre les modèles est en général beaucoup plus grande que la variabilité interne dans le cas des données brutes. Ceci suggère donc que les différences de performances entre les modèles peuvent être discriminées de manière assez robuste pour les données brutes, et donc que ces climatologies saisonnières pourraient potentiellement être utilisées dans un schéma de pondération basé sur les performances.

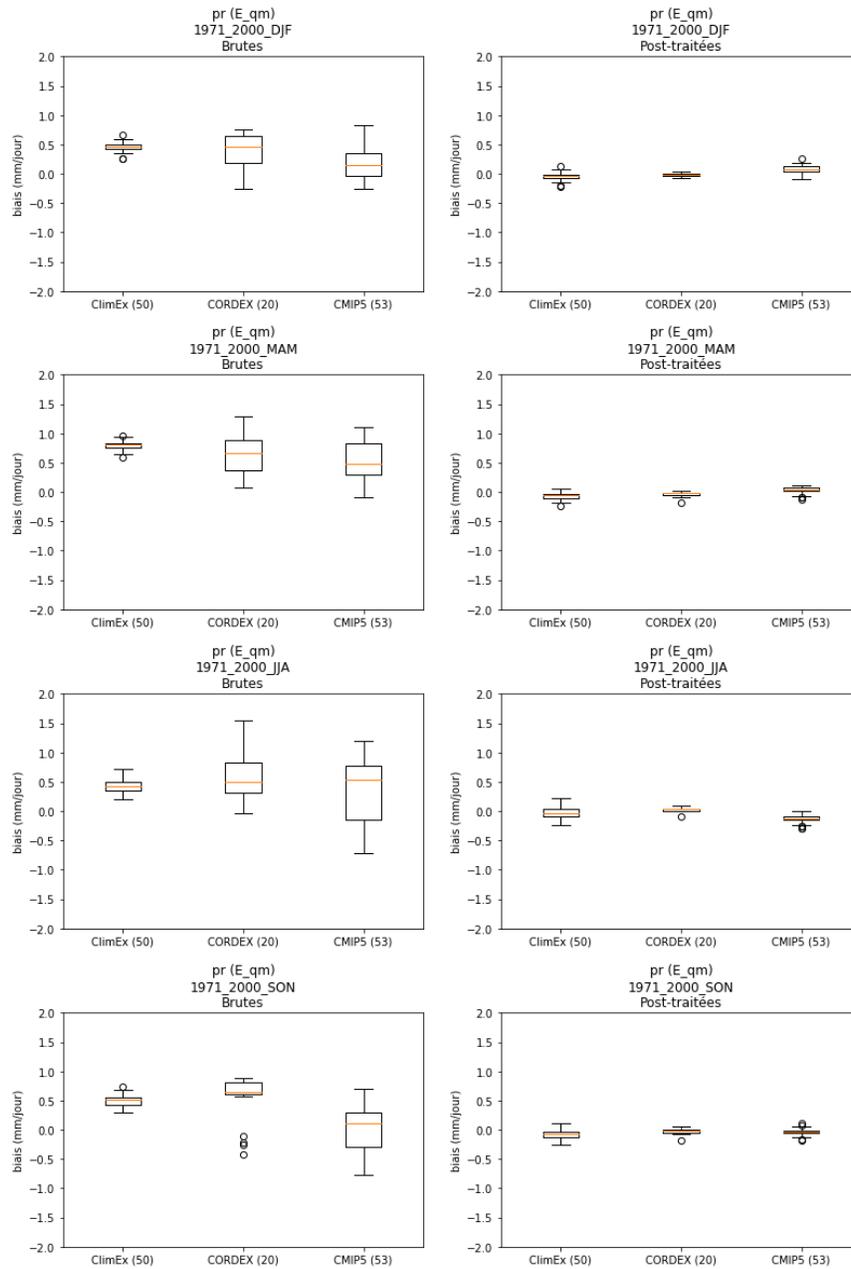


FIGURE 6 – Biases saisonnier de la précipitation pour la période 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), pour les données brutes (gauche) et post-traitées (droite).

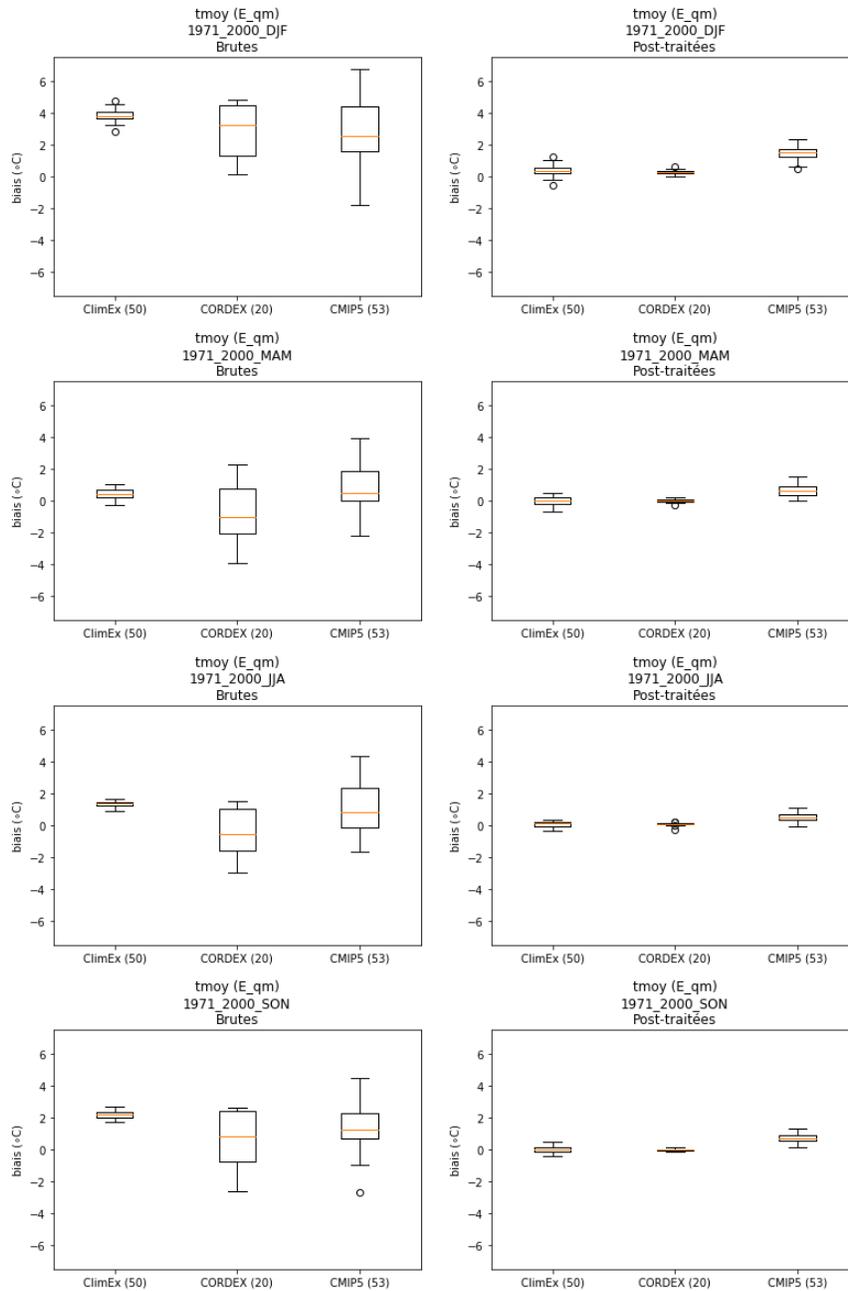


FIGURE 7 – Biais saisonnier de la température moyenne 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), pour les données brutes (gauche) et post-traitées (droite).

5.1.3 Variabilité interannuelle

Afin d'évaluer la variabilité interannuelle simulée, nous comparons d'abord les séries de moyennes annuelles de précipitation et température moyenne sur la période historique de 30 ans pour les 10 simulations choisies précédemment, ClimEx et les observations avec les simulations brutes, Fig. 8 et les données post-traitées, Fig. 9. Les deux figures montrent que le post-traitement réduit les biais tout en préservant la forme générale des séries temporelles.

Les figures 10 et 11 montrent le rapport des écarts types des séries de moyennes saisonnières entre toutes les simulations de l'ensemble et les observations, pour les précipitations et la température respectivement. Pour la précipitation, l'amplitude de la variabilité interannuelle des simulations brutes est généralement plus grande que celle des observations tandis que les centres des distributions des données post-traitées se rapprochent du rapport de valeur unitaire. Le post-traitement ne corrige toutefois pas la variabilité interannuelle de manière systématique. Pour la température (Fig. 11), la variabilité interannuelle des simulations brutes et post-traitées se situe généralement autour de 1. Nous pouvons aussi constater une amélioration limitée au niveau du centre des distributions qui se rapproche de 1 après post-traitement. Finalement, tant pour la température que les précipitations, la plage couverte par les distributions est peu affectée par le post-traitement et reste de l'ordre de la variabilité interne de ClimEx.

En somme, ces résultats montrent que le post-traitement corrige partiellement la variabilité interannuelle dans les simulations. De plus, puisque les écarts de performances entre les différents modèles de climat (ensembles CORDEX et CMIP5) couvrent une plage similaire aux différences dues à la variabilité interne, l'évaluation de la variabilité interannuelle reste une métrique offrant peu de robustesse pour discriminer les performances des modèles.

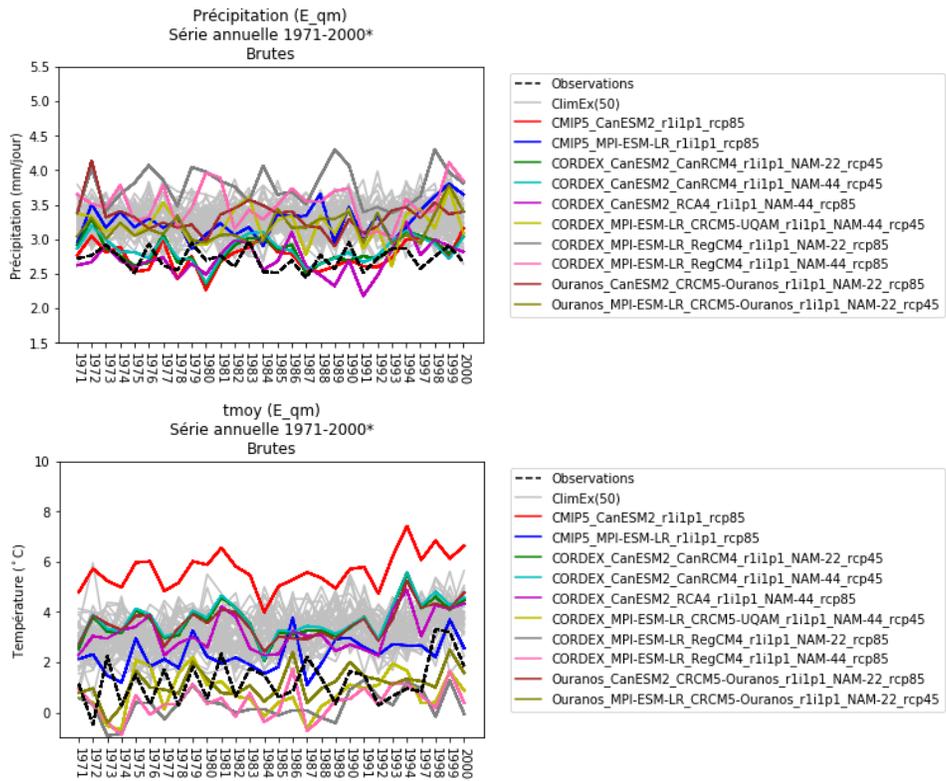


FIGURE 8 – Série annuelle de la précipitation (haut) et la température moyenne (bas) pour 1971-2000, données brutes.

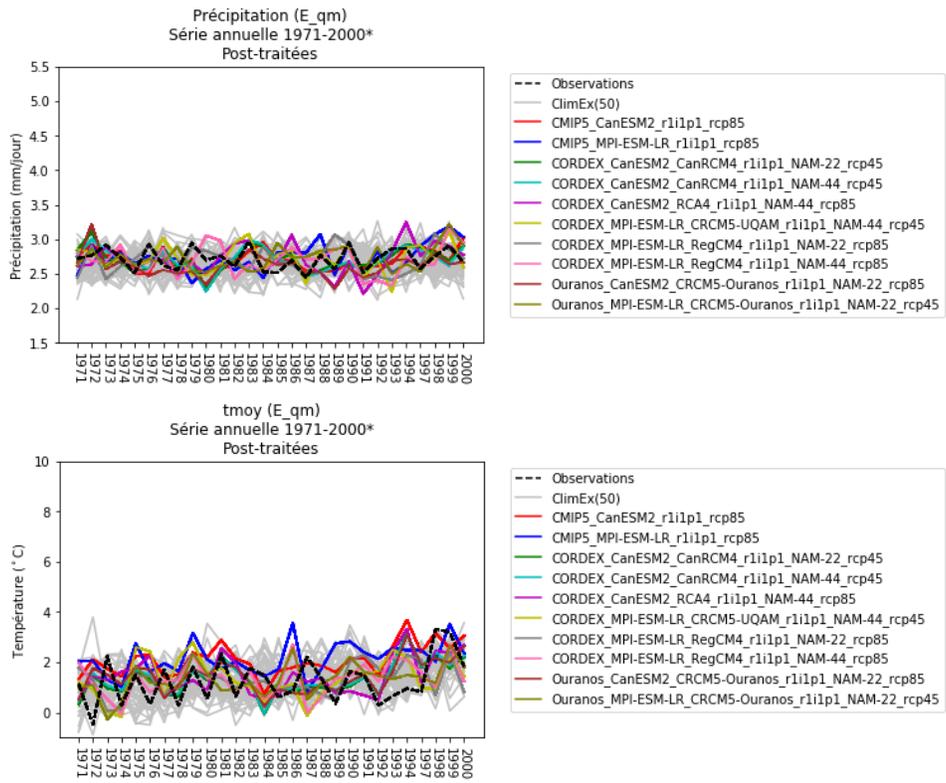


FIGURE 9 – Série annuelle de la précipitation (haut) et la température moyenne (bas) pour 1971-2000, données post-traitées.

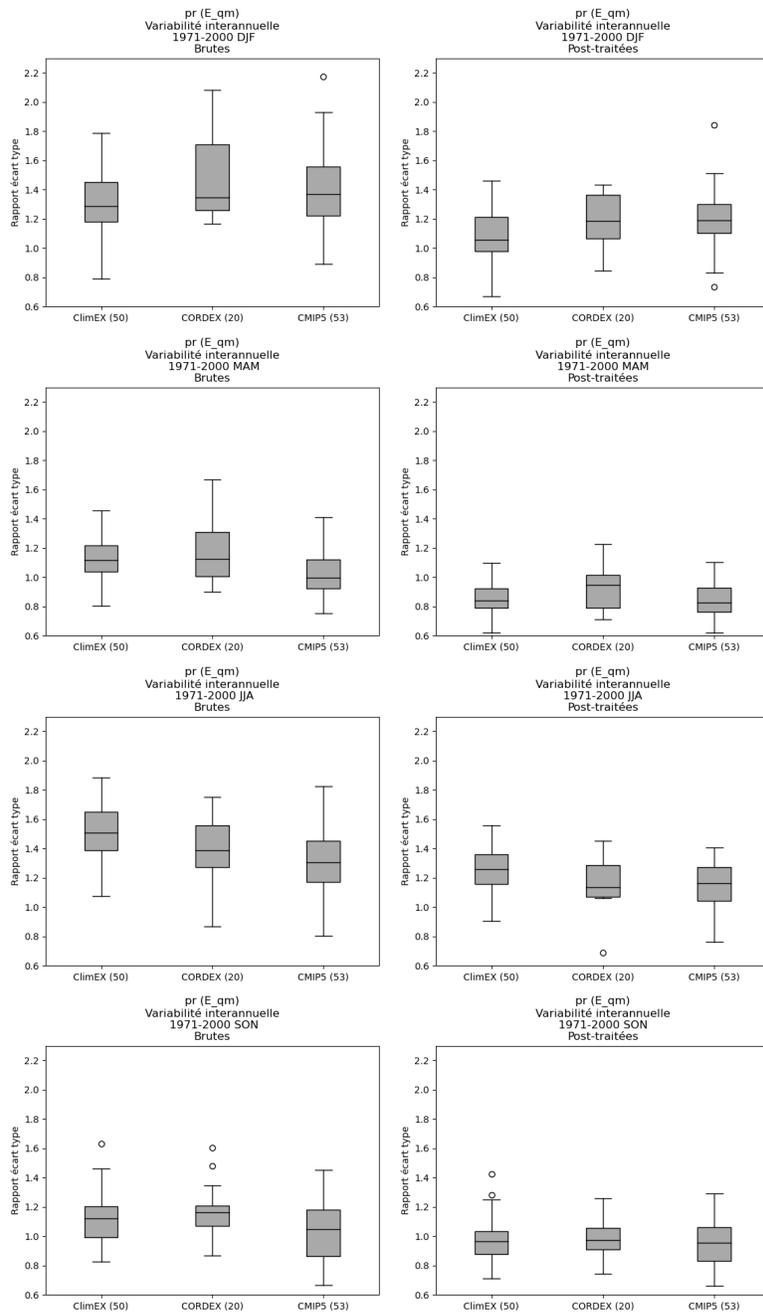


FIGURE 10 – Variabilité interannuelle de la moyenne saisonnière de la précipitation en hiver (DJF), printemps (MAM), été (JJA) et automne (SON) pour 1971-2000, données brutes (gauche) et post-traitées (droite).

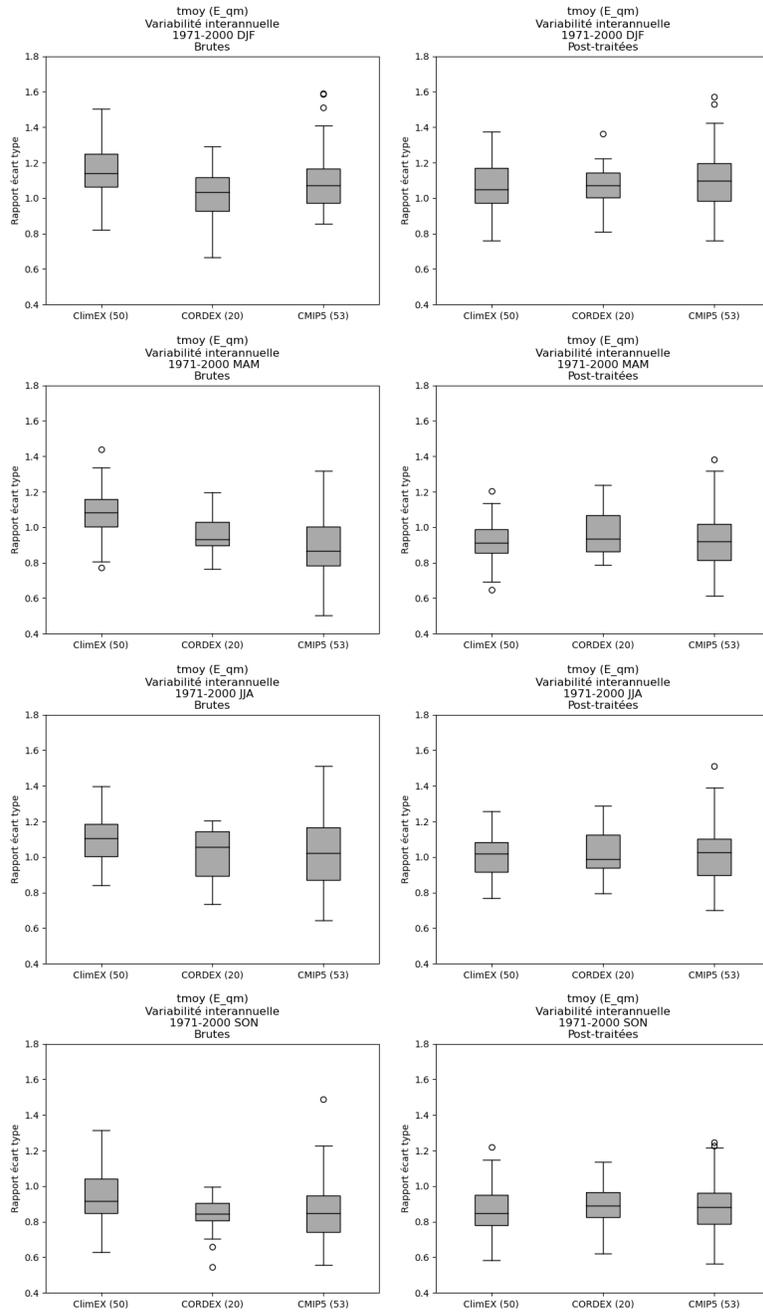


FIGURE 11 – Variabilité interannuelle de la moyenne saisonnière de la température moyenne en hiver (DJF), printemps (MAM), été (JJA) et automne (SON) pour 1971-2000, données brutes et post-traitées.

5.1.4 Tendance de changement climatique

La capacité des modèles climatiques à reproduire les tendances associées au changement climatique historique est évaluée dans cette section. Puisque par construction, la méthode de post-traitement appliquée aux simulations préserve les tendances, seules les données brutes sont présentées dans cette section.

Les tendances historiques de changement climatique sont calculées à l'échelle saisonnière pour les différents ensembles de simulations et les observations (données sur grille NRCan). Pour les tendances de précipitation (Fig. 12), les modèles surestiment généralement la tendance observée (ligne rouge) pour les saisons d'hiver et printemps, tandis qu'un biais majoritairement négatif est observé en été et automne. Pour la température (Fig. 13), les tendances observées sont sous-estimées en hiver et au printemps tandis qu'un biais positif apparaît en été. Pour l'automne, le biais dans la tendance est généralement près de zéro. Les figures 28 et 29 en annexe B montrent les mêmes résultats que sur les figures 12 et 13 mais sous forme de différences par rapport aux observations.

Si l'on compare la dispersion due à la variabilité interne entre les estimés de tendances dans l'ensemble ClimEx avec la dispersion entre les modèles de CORDEX et CMIP5, la variabilité interne semble représenter la majeure partie de l'incertitude inter-modèle. L'incertitude considérable due à la variabilité interne sur le calcul des tendances simulées suggère que les tendances observées ne peuvent être uniquement attribuées au forçage de l'augmentation des concentrations de gaz à effet de serre dans l'atmosphère, mais contiennent également une part importante de variabilité interne. Puisque par construction, les simulations des modèles globaux de climat de cet ensemble ne sont pas contraintes de reproduire l'évolution de la variabilité naturelle observée (p. ex. au niveau de la circulation atmosphérique et océanique), les simulations utilisées ici peuvent difficilement être évaluées sur la base de tendances fortement influencées par la variabilité interne. Ceci suggère donc que la présente évaluation des tendances simulées doit être interprétée avec de fortes réserves, et ne permet pas de distinguer de manière robuste les performances entre les modèles de manière robuste.

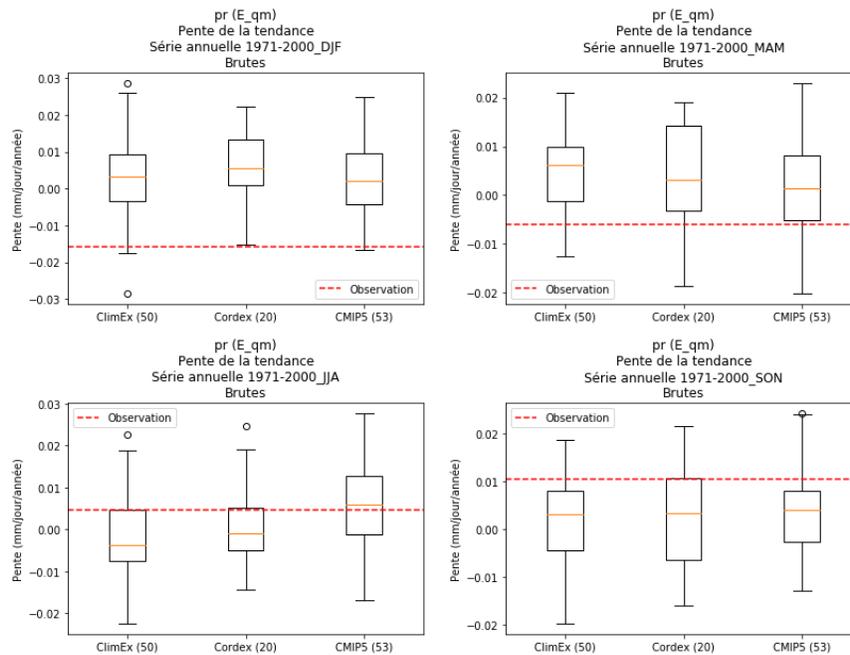


FIGURE 12 – Tendance de changement climatique sur la période 1971-2000 (colonne de gauche) et le biais par rapport aux observations (colonne de droite) pour la précipitation pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes. La tendance calculée à partir des observations est donnée par la ligne rouge.

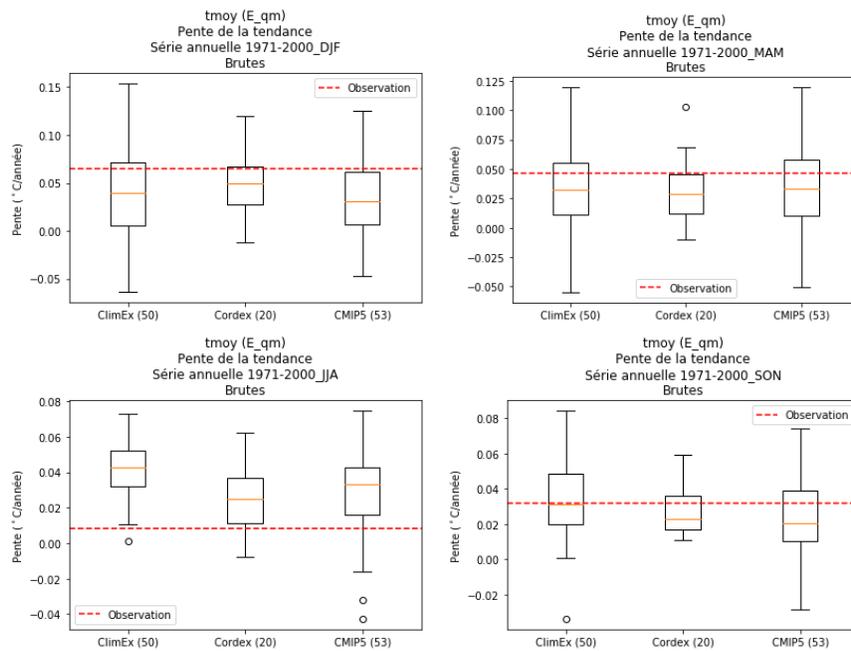


FIGURE 13 – Tendence de changement climatique sur la période 1971-2000 (colonne de gauche) et le biais par rapport aux observations (colonne de droite) pour la température moyenne pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes. La tendance calculée à partir des observations est donnée par la ligne rouge.

5.2 Pondération a priori

Dans cette section, nous présentons quelques méthodes simples pour pondérer un ensemble de projections climatiques. Ces méthodes de pondération, dites *a priori*, consistent en des méthodes de sélection de modèles basées strictement sur nos connaissances quant à la composition de l'ensemble, par exemple les noms et types de modèles utilisés ou leur résolution. Les méthodes de pondération a priori ne tiennent donc pas compte des résultats des modèles et de leurs performances telles que décrites à la section précédente. Une méthode de pondération spécifiquement basée sur les performances des modèles sera décrites à la section 5.3.

5.2.1 Méthodes de sélection des modèles

Les méthodes de pondération a priori choisies sont les suivantes :

- **Un modèle un vote uniforme**, aussi appelée démocratie des modèles, cette méthode implique que chaque modèle a un poids équivalent de 1. Pour les modèles ayant produit des simulations multi-membres, chaque membre compte pour une fraction totale de un, par exemple ClimEx où chaque membre vaudra 1/50. Il est à noter que nous utilisons ici, par simplicité, le nom du modèle et de son centre comme critère pour définir un modèle. Par exemple, le Modèle régional canadien du climat (MRCC5) sera considéré comme deux modèles distincts lorsqu'il est utilisé par Ouranos ou l'Université du Québec à Montréal (UQAM). Certaines différences (bien que mineures) sont connues dans les configurations utilisées par les deux centres.
- **Un modèle un vote binaire**, est une variante de l'approche précédente de la démocratie des modèles où nous utilisons seulement le premier membre de chaque MGC, et la grille de plus haute résolution pour les MRC.
- **Un centre, un vote**, aussi appelée démocratie institutionnelle (Leduc *et al.*, 2016), consiste à donner un poids égal à chaque centre de modélisation représenté dans l'ensemble. Cette approche est implémentée ici en n'utilisant qu'un seul modèle par centre, soit en donnant prépondérance à une simulation régionale ou au modèle supérieur en complexité.
- **Sélection par la résolution** : Lorsque la résolution spatiale est importante pour une application donnée, il peut être utile de considérer cette caractéristique comme critère de sélection des modèles. Or, il est important de noter que les modèles de climat ont généralement

une “résolution effective”, plus basse que leur résolution nominale (c.-à-d. telle que définie par la grille du modèle), qui est notamment dû à l’application de filtres qui limitent le développement d’instabilités numériques. Les études proposent différents estimés pour la résolution effective, soit entre 2.7 et 8 fois la résolution nominale de la grille du modèle (Klaver *et al.*, 2020; Skamarock, 2004; Abdalla *et al.*, 2013). Dans le cadre d’Info-Crue, les bassins versants utilisés en hydrologie ont des superficies qui se situent entre 69 et 42,657 km², ce qui correspond approximativement à des longueurs d’échelle de 8 à 207 km. Assumant une résolution effective de cinq fois la résolution nominale, la distance entre deux points de grille du modèle devrait être d’au plus 1.6 à 41 km, respectivement, afin de modéliser physiquement une variabilité spatiale sur de telles régions. Nous constatons que seuls les modèles régionaux de l’ensemble (dont les résolutions sont de 0.12°, 0.22° et 0.44°) sont en mesure de fournir une information spatiale adéquate pour les plus grands bassins, tandis que les plus petits bassins ne sont représentés par aucun modèle dans l’ensemble. La figure 14 montre le nombre de simulations dans l’ensemble pour une surface donnée en km², en utilisant des seuils de résolution de 1, 2 et 5 fois la résolution nominale des modèles. Dans ce chapitre, nous utilisons par simplicité un seuil de $\Delta x = 100$ km, permettant de représenter la variabilité spatiale (à 5 fois la résolution nominale) sur une zone de 250,000 km².

- **Sélection par ensemble** : Nous ajoutons finalement une méthode de sélection par ensemble, soit ClimEx, CORDEX et CMIP5.

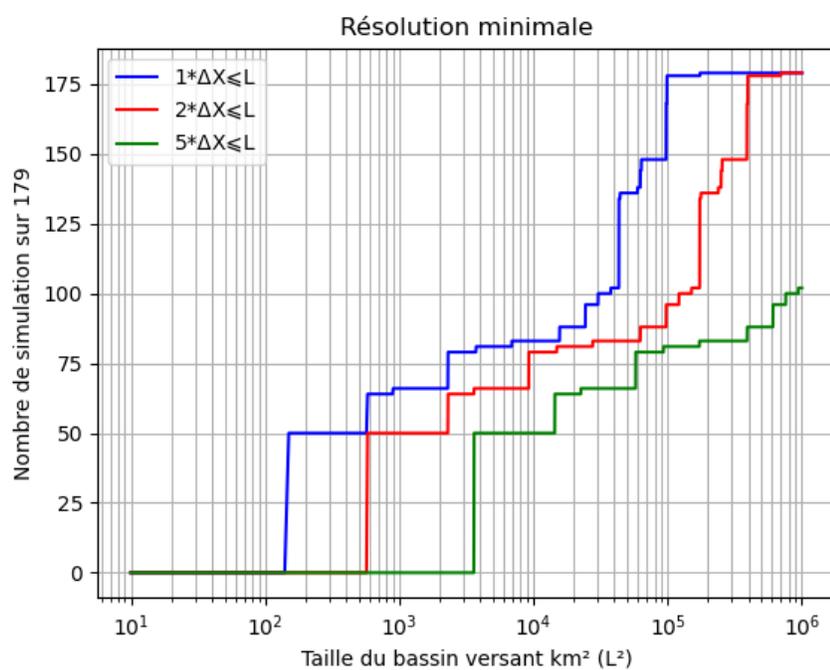


FIGURE 14 – Nombre de simulations dans l’ensemble ayant une résolution égale ou plus fine que $1\Delta x$, $2\Delta x$ et $5\Delta x$ pour une surface donnée en km^2 .

5.2.2 Sensibilité du changement climatique à la méthode de sélection

Dans cette section, le signal de changement climatique de la moyenne saisonnière de la précipitation en fonction de celui de la température est montré sous forme de diagrammes de dispersion pour la période 2071-2100 par rapport à 1981-2010 selon le scénario RCP8.5. Les méthodes de pondération a priori précédemment décrites sont utilisées pour calculer la moyenne d'ensemble des deltas. Les différentes méthodes de pondération a priori précédemment décrites sont utilisées pour calculer la moyenne d'ensemble des deltas à partir des données brutes uniquement afin d'alléger le rapport (Fig. 15).

Pour la température, la moyenne d'ensemble de ClimEx (point rouge) montre le plus grand changement pour toutes les saisons, tandis que l'ensemble CORDEX (point kaki) a généralement le delta le plus faible. Il est à noter que la sélection par rapport à la résolution (point bleu) inclut les simulations de ClimEx, CORDEX et un MGC, et donc ce point tend à se rapprocher de ClimEx et ses 50 membres. En combinant la méthode un modèle un vote avec la sélection par rapport à la résolution, le point bleu devrait donc se rapprocher du point kaki. Pour les précipitations, l'ensemble ClimEx (point rouge) montre le changement le plus important en hiver et le plus faible dans les autres saisons. La méthode un modèle un vote (en vert) se situe dans les changements les plus petits (importants) pour la température (précipitations) pour presque toutes les saisons. En résumé, l'écart entre les méthodes oscille en deçà de 1.6°C pour la température selon les saisons, et en deçà de 10% pour la précipitation.

Il est important de noter que malgré des hypothèses fortes dans la sélection a priori des modèles, l'impact sur le signal de changement climatique moyen reste somme toute modéré. Ce faible impact s'explique notamment par le grand nombre de simulations utilisées et sélectionnées selon les différentes méthodes. Donc, dans une situation où un petit échantillon de simulations serait utilisé, la moyenne pondérée et autres statistiques d'ensemble pourraient être fortement influencées par ces hypothèses de sélection. Afin de mettre en perspective l'impact de la sélection des simulations sur les moyennes d'ensemble, la figure 30 en annexe B propose une analyse des simulations individuelles associées aux ensembles CORDEX, ClimEx et CMIP5.

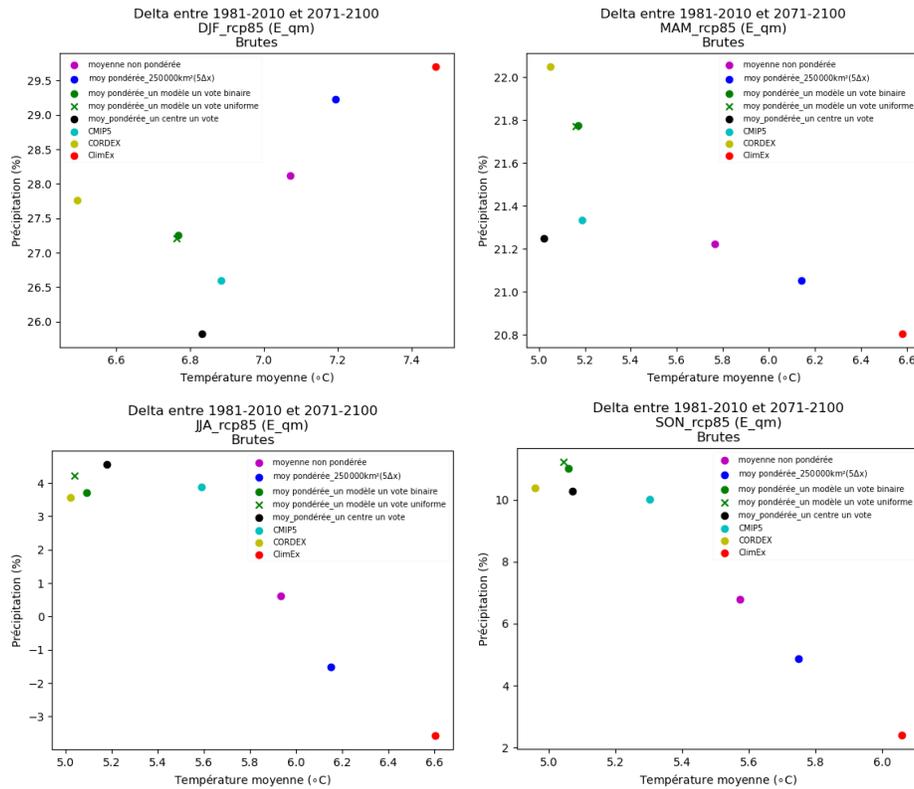


FIGURE 15 – Delta saisonnier en 2071-2100 par rapport à 1981-2010 pour différentes méthodes de pondération pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), selon RCP 8.5, données brutes.

5.3 Pondération a posteriori

À la lumière des résultats de la section 5.1 où les variables simulées par les modèles sont bien corrigées par le post-traitement dans le cas des moyennes climatologiques, corrigées partiellement pour la variabilité interannuelle, ou inchangées pour les tendances, les données post-traitées ne permettent pas de discerner des écarts notables de performances entre les modèles. Pour cette raison, seules les données brutes des modèles seront utilisées pour la pondération basée sur les performances.

Dans cette section, nous analysons certains des concepts sous-jacents à la construction d'un schéma de pondération basé sur les performances des modèles. Plus précisément, nous allons évaluer la sensibilité d'un schéma de pondération typique à certains de ses paramètres incertains, dont les valeurs permises devraient normalement être circonscrites en fonction de l'application où est utilisée la pondération. Parmi ces paramètres, notons d'abord le choix de la fonction de pondération elle-même (Sect. 5.3.1), l'ajustement paramétrique de la fonction de pondération (Sect. 5.3.2), la sélection des métriques d'évaluation des modèles (Sect. 5.3.3), et la combinaison de ces métriques en une valeur de poids par modèle (Sect. 5.3.4). Nous évaluons ensuite l'impact de ces facteurs sur la moyenne d'ensemble des projections climatiques (Sect. 5.3.5), et terminerons avec une discussion sur la possibilité d'appliquer des métriques globales pour contraindre les projections régionales (Sect. 5.3.6).

Afin de nous guider dans l'évaluation du schéma de pondération, les sections 5.3.1 à 5.3.4 utilisent l'approche du modèle parfait. Le principe de cette approche consiste à utiliser plusieurs simulations (membres) d'un même modèle comme des pseudoréférences. Nous utiliserons à tour de rôle les 50 membres de l'ensemble ClimEx comme des pseudoréférences pour évaluer les simulations des autres modèles en climat historique. Ceci permettra de quantifier la robustesse du schéma de pondération par rapport à la variabilité interne, une source d'incertitude intrinsèque au système climatique.

5.3.1 Choix de la fonction de pondération

Tel qu'indiqué en introduction, différentes méthodes existent dans la littérature pour obtenir des poids à partir de mesures de performance des modèles. Puisque l'utilisation d'une méthode plutôt qu'une autre est un choix subjectif, nous analysons dans cette section la sensibilité des poids obtenus à partir du choix entre deux fonctions de pondération connues dans la littérature, soit les fonctions INV et EXP décrites à la section 4.3 (Éq. 1 et 2,

respectivement).

La figure 16 montre les poids calculés pour chaque simulation individuelle, mais rassemblés pour chaque modèle (voir Tab. 2 en annexe C). Les deux panneaux montrent les poids obtenus en utilisant les fonctions INV et EXP, pour la moyenne climatologique de la température en hiver en utilisant tous les membres de ClimEx comme des pseudoréférences (donc pour un modèle n’ayant qu’un seul membre, 50 valeurs de poids sont affichées). La répartition des poids est similaire entre les deux fonctions du point de vue du rang des simulations selon leurs performances relatives, c’est-à-dire : plus le biais d’une simulation par rapport à la référence est petit, plus son poids est grand. Cependant, la distribution entre les poids des simulations est très différente selon la méthode utilisée. Pour la fonction INV, nous remarquons un écart plus grand entre les poids d’un nombre restreint de modèles qui dominant et le reste de l’ensemble. Pour EXP, un plus grand nombre de modèles montrent des simulations avec des poids dominants. Les poids sont donc mieux répartis entre les modèles pour la fonction EXP en comparaison avec la fonction INV.

Un point important à noter est la plage des valeurs de poids couverte par les membres de ClimEx. Puisque les membres de ClimEx sont utilisés à tour de rôle comme des pseudoréférences, il serait raisonnable de s’attendre à ce que les membres de ce modèle obtiennent des poids généralement élevés par rapport aux simulations des autres modèles. Or, le fait que certains membres de ClimEx aient des poids élevés dans certains cas, et des valeurs pratiquement nulles dans d’autres cas, peut paraître contre-intuitif. Comme nous allons le voir dans les prochaines sections, un ajustement du paramètre de la fonction de poids (σ_D) peut permettre de réduire la dispersion des poids associés aux 50 membres du modèle de référence, cette dispersion étant due à l’effet de la variabilité interne sur la référence. Contrairement à INV, la fonction EXP permet de tenir compte de cette variabilité à l’aide du rayon de similarité σ_D , son paramètre ajustable, permettant de borner les variations de poids lorsque la différence modèle-référence varie à l’intérieur de ce rayon. De plus, la fonction INV fait tendre le poids vers l’infini lorsque le biais tend vers zéro (voir Fig. 3). Le manque d’un paramètre d’ajustement pour tenir compte de la variabilité interne ainsi que la discontinuité de la fonction INV sont les raisons qui motivent l’utilisation de la fonction EXP dans la suite de ce rapport.

5.3.2 Ajustement de la fonction de pondération

Comme mentionné à la section 4.3, la quantité σ_D dans la fonction EXP est un paramètre libre. Le rôle de ce paramètre est primordial car il permet d’ajuster le niveau d’intensité associé à la pondération. Tel que montré sur la figure 3, une petite valeur de σ_D correspond à restreindre l’attribution de poids élevés seulement aux modèles très proches de la référence. À l’inverse, une grande valeur de σ_D mène à des poids de plus en plus équivalents entre les modèles, et à un ensemble (trop) démocratique ne tenant pas compte des différences entre les performances de modèles. Déterminer la valeur de σ_D , et donc le niveau d’intensité de la fonction de pondération, repose sur un choix subjectif, par exemple en choisissant de préserver 80% des modèles dans l’intervalle de confiance (Lorenz *et al.*, 2018). Nous pouvons également guider le choix de σ_D à partir de deux hypothèses sur la sélection des métriques, soit 1) les métriques d’évaluation des modèles sont robustes (c.-à-d. peu influencées par la variabilité interne), et 2) les métriques sont fiables pour contraindre l’incertitude des projections futures. Dans les prochaines sections, nous nous concentrons sur la première hypothèse de sélection des métriques, tandis que la seconde sera abordée à la section 5.3.6.

La figure 17 montre les poids calculés en utilisant la fonction EXP pour la moyenne climatologique de la température en hiver en utilisant les membres de ClimEx comme des pseudoréférences pour différentes valeurs de σ_D , soit σ_{D^*} (déjà présenté à la figure 16, panneau inférieur), $2\sigma_{D^*}$, $5\sigma_{D^*}$ et $10\sigma_{D^*}$, où σ_{D^*} est la valeur de référence correspondant à l’effet de la variabilité interne sur la différence entre deux membres de ClimEx (voir Sect. 4.3). Pour $\sigma_D = \sigma_{D^*}$, les poids obtenus sont pratiquement nuls pour environ la moitié des modèles, tandis que les poids se situent entre 0 et 0.07 pour les autres modèles, montrant une grande variabilité selon le membre ClimEx choisi comme référence. En comparant avec les trois autres panneaux, on voit que l’écart diminue entre les poids des différents modèles, ainsi qu’entre les poids des membres d’un même modèle. Pour $10\sigma_{D^*}$, certaines variations sont toujours visibles entre les poids des différents modèles même si les valeurs se situent généralement en dessous de 0.01.

Les figures 18 et 19 sont similaires à la figure 17, mais pour la variabilité interannuelle et les tendances historiques. Dans les deux cas, la convergence vers un ensemble démocratique (c.-à-d. avec des poids pratiquement égaux pour tous les modèles) est plus rapide que pour la moyenne climatologique (Fig. 17). Il est aussi intéressant de remarquer que la convergence vers l’ensemble démocratique est un peu plus rapide pour les tendances que pour la variabilité interannuelle. Ce résultat valable pour d’autres saisons et va-

riables. Sur la figure 20, nous montrons les poids calculés pour la température en hiver (gauche) et en été (droite), pour la moyenne climatologique (haut), la variabilité interannuelle (centre) et la tendance (bas), en utilisant une valeur commune de $\sigma_D = 3\sigma_{D^*}$. Pour les deux saisons, nous voyons que la variabilité des poids entre les différents modèles est la plus grande pour la moyenne climatologique, et la plus petite pour la tendance. Pour la tendance, les différents modèles ont un poids à peu près égal en moyenne. Ce résultat est également reproduit pour la précipitation en été et hiver (Fig. 21). Une hypothèse pour expliquer ce résultat est que les poids des modèles deviennent équivalents (en moyenne) lorsque la variabilité interne est comparable à l’incertitude inter-modèle. Pour vérifier cette hypothèse, nous avons calculé le rapport entre ces deux sources d’incertitude :

$$\Gamma = \frac{\sigma_{IV}}{\sigma_{IM}} \quad (7)$$

où σ_{IV} représente l’écart-type entre les valeurs obtenues (moyennes climatologiques, variabilité interannuelle ou tendances) pour les membres de ClimEx et σ_{IM} l’écart-type inter-modèle (en utilisant le premier membre de chaque modèle). Ces valeurs sont calculées pour la moyenne climatologique, la variabilité interannuelle et la tendance sur chaque panneau des figures 20 et 21. En comparant les trois valeurs de Γ pour chaque variable et les quatre saisons (voir Tab. 1), le rapport est en effet toujours le plus petit pour la moyenne climatologique et le plus grand pour la tendance. En d’autres termes, plus la variabilité interne représente une grande proportion de l’incertitude entre les modèles, plus les modèles tendront à obtenir des poids égaux en moyenne. Il en résulte que σ_D doit être choisi en fonction de la variabilité interne et de l’incertitude inter-modèle. Considérer les deux sources d’incertitude pour guider le choix de σ_D permet d’évaluer l’hypothèse de robustesse dans la sélection des métriques, permettant ainsi de pouvoir discerner un cas non robuste où l’incertitude est dominée par de la variabilité interne (comme pour les tendances) de celui d’une grande incertitude par rapport à la variabilité interne (moyennes climatologiques) où il est possible de discriminer les performances entre les modèles.

TABLE 1 – Valeurs de Γ (Éq. 7) représentant le rapport entre l'écart-type de l'ensemble ClimEx et l'écart-type multi-modèle (un membres par modèle, voir Tab. 2) pour deux variables (Tmoy et Pr), trois statistiques climatiques (moyenne climatologique (climato.), variabilité interannuelle (VIA) et tendance) et quatre saisons.

Variable	Statistique	Γ_{DJF}	Γ_{MAM}	Γ_{JJA}	Γ_{SON}
Tmoy	climato.	0.19	0.17	0.1	0.12
Tmoy	VIA	0.72	0.88	0.73	0.77
Tmoy	tendance	1.43	1.34	0.88	1.12
Pr	climato.	0.24	0.24	0.21	0.22
Pr	VIA	0.68	0.82	0.92	0.86
Pr	tendance	1.22	0.83	1.12	0.95

À la lumière de ces résultats, le choix subjectif d'attribuer une valeur à σ_D peut avoir un impact significatif sur les résultats. Plus sa valeur est petite, plus on tend à attribuer une surconfiance dans un petit nombre de modèles. Dans ce cas, le résultat de l'ensemble pondéré peut être hautement variable tout dépendant des modèles retenus. À l'opposé, une grande valeur de σ_D rend l'ensemble pondéré moins sensible à la référence et à la variabilité interne, mais à la limite rend également la pondération inefficace puisqu'elle revient alors au statu quo de la démocratie des modèles, c'est-à-dire : aucune pondération.

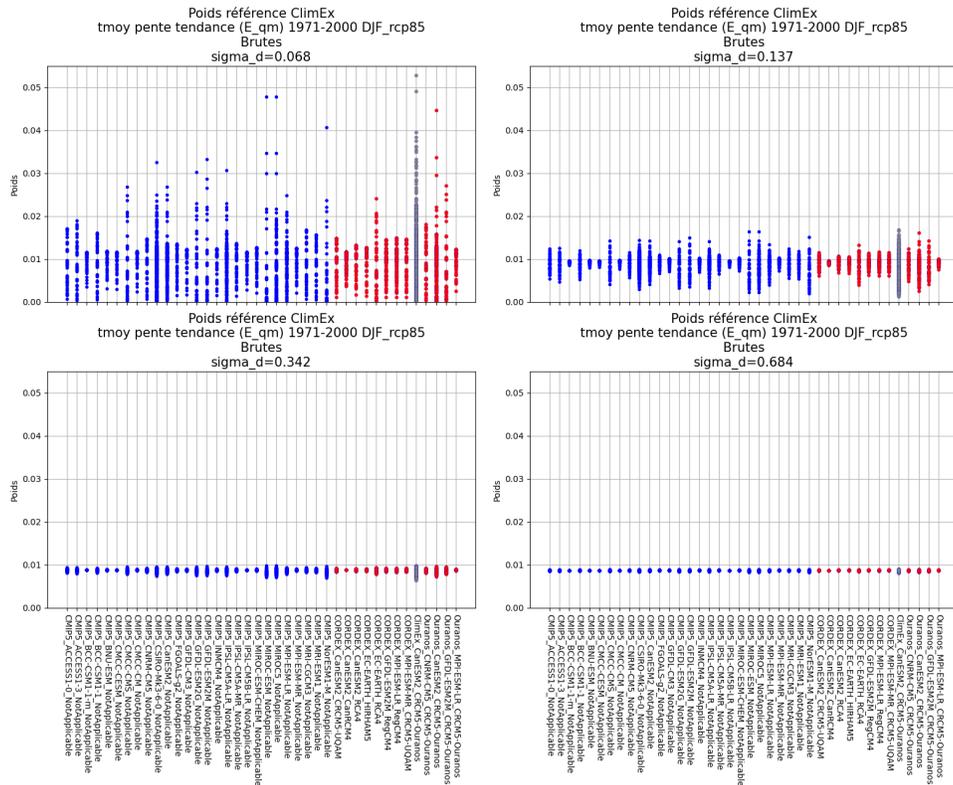


FIGURE 19 – Poids calculés pour la tendance historique de la température moyenne en hiver (DJF) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est variée selon $1\sigma_{D^*}$, $2\sigma_{D^*}$, $5\sigma_{D^*}$ et $10\sigma_{D^*}$.

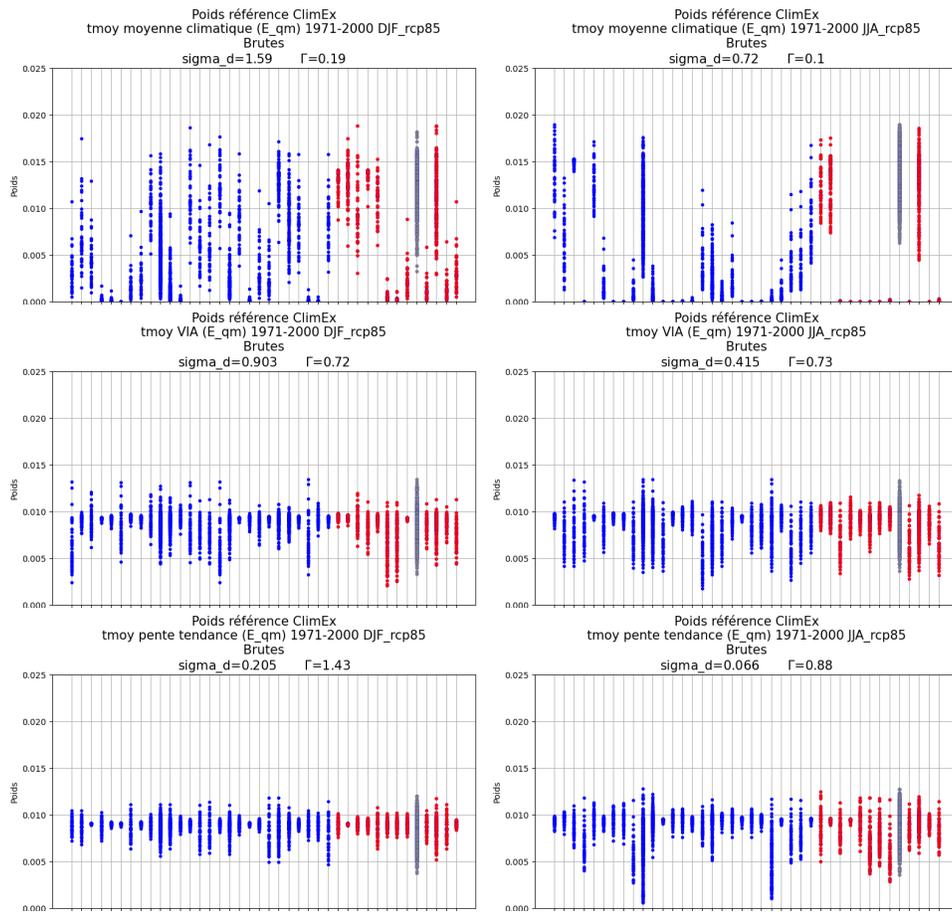


FIGURE 20 – Poids calculés pour la moyenne climatologique (haut), la variabilité interannuelle (centre) et la tendance historique (bas) de la température moyenne en hiver (DJF, gauche) et été (JJA, droite) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est fixée à $3\sigma_{D^*}$. Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2.

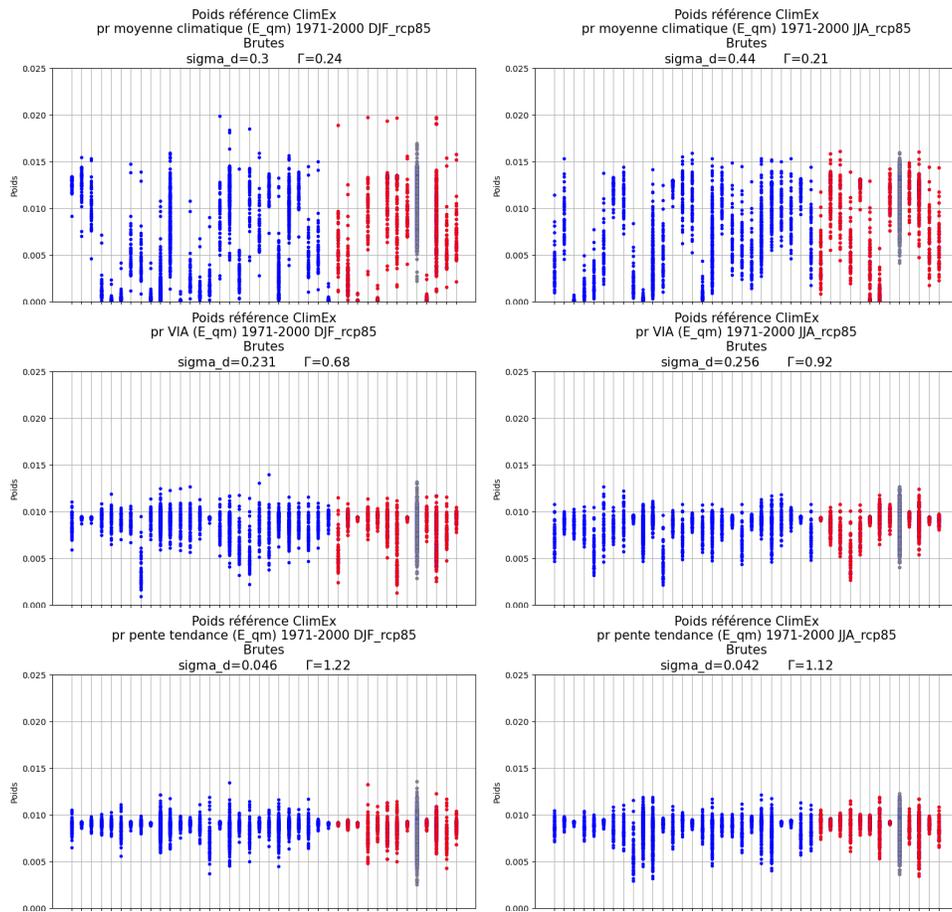


FIGURE 21 – Poids calculés pour la moyenne climatologique (haut), la variabilité interannuelle (centre) et la tendance historique (bas) des précipitations en hiver (DJF, gauche) et été (JJA, droite) en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. La valeur de σ_D est fixée à $3\sigma_{D^*}$. Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2.

5.3.3 Sélection des métriques de performance

Comme nous l’avons vu précédemment, la sélection de plusieurs métriques d’évaluation des modèles pour utilisation dans la formule de pondération (Éq. 4) est, tout comme le choix de la fonction de pondération elle-même et l’ajustement de son rayon de similarité, un choix subjectif. D’un côté, il peut être justifié d’utiliser un grand nombre de métriques afin d’évaluer le comportement des modèles selon plusieurs perspectives, et ainsi tenter d’identifier les “meilleurs modèles” de manière générale. D’un autre côté, il peut être tout autant justifié d’utiliser un nombre plus restreint de métriques, mais qui sont directement reliées au contexte d’application de la pondération. Dans le présent rapport, le nombre de métriques analysées correspond à :

$$N_{métriques} = N_{variables} \times N_{saisons} \times N_{stats} = 2 \times 4 \times 3 = 24 \quad (8)$$

où le nombre de métriques est égal au produit du nombre de variables (2, température de surface et précipitations totales), de saisons (4), et de statistiques climatiques (3, soit la moyenne climatologique, la variabilité internannuelle et la tendance historique). Nous allons maintenant évaluer l’effet de combiner plusieurs métriques d’évaluation sur les poids résultants. Pour ce faire, nous utiliserons seulement les métriques basées sur les moyennes climatologiques pour les deux variables et les quatre saisons, puisque ces métriques sont généralement peu influencées par la variabilité interne, et donc plus robustes comparativement à la variabilité interannuelle et aux tendances historiques.

5.3.4 Combinaison des métriques

Tel qu’indiqué précédemment, nous retenons huit métriques de performance pour analyser l’impact de la combinaison de plusieurs métriques dans le calcul des poids, soit les biais dans les moyennes climatologiques saisonnières pour la température et la précipitation. Il est à noter que ces huit indicateurs de performance pourraient être combinés de différentes manières selon le type d’application, un aspect qui n’est pas évalué dans le cadre de ce rapport.

La figure 22 montre la sensibilité des poids lorsqu’on augmente le nombre de métriques de performance en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Sur les huit panneaux de la figure, le calcul des poids est effectué d’abord par rapport à la température en incluant les saisons une à une, c’est à dire hiver seulement (une métrique), hiver et printemps (deux métriques), hiver-printemps-été (trois métriques),

les quatre saisons (quatre métriques), et ainsi de suite en ajoutant les quatre saisons pour la précipitation. À noter que pour chaque panneau, les métriques utilisées ont la même importance relative dans le calcul des poids, et ces derniers sont renormalisés après combinaison avec l'équation (5). À chaque étape, nous pouvons remarquer que les poids de certains modèles diminuent tandis que d'autres augmentent. Donc, pour un modèle se montrant très différent de la référence dans plusieurs métriques, son poids global diminuera rapidement au fur et à mesure que les métriques sont combinées. Nous pouvons aussi remarquer que certains modèles tendent à préserver un poids élevé au fil des combinaisons. Pour quatre métriques, il est intéressant de noter que les quatre modèles régionaux se démarquant par des poids plus élevés sont (de gauche à droite, en rouge) : CRCM5-UQAM, CanRCM4, RCA4 et CRCM5-Ouranos, tous pilotés par CanESM2. À huit métriques, seuls CRCM5-UQAM et CRCM5-Ouranos survivent à la combinaison des métriques. Notons aussi que le pic associé à CRCM5-Ouranos est le plus près de celui de ClimEx, et correspond également à la configuration la plus proche (seuls les membres de CanESM2 et la taille du domaine régional différent). Dans sa configuration actuelle ($\psi_D = 1$, voir éq. (6)), cette méthode permet d'identifier les modèles globalement les plus près de la référence, et de voir cette conclusion se renforcer au fur et à mesure que davantage de métriques sont combinées.

Les figures 22, 23 et 24 montrent exactement le même calcul, mais avec des valeurs de $\psi_D = 1, 2$ et 4 , respectivement. Pour $\psi_D = 2$ et huit métriques, une quinzaine de modèles obtiennent des poids significatifs. À partir de $\psi_D = 2$, la convergence est aussi beaucoup plus lente. Il est intéressant de noter que pour huit métriques, à $\psi_D = 4$, l'écart entre les poids des membres de ClimEx est d'environ $\pm 10\%$, tandis que ces variations sont de $\pm 30\%$ pour $\psi_D = 2$, et d'environ $\pm 100\%$ pour $\psi_D = 1$. Ainsi, une grande valeur de σ_D permet de contraindre les poids des membres d'un même modèle à l'intérieur d'un intervalle restreint (comme attendu puisqu'ils sont équivalents), mais ajoute également du poids à des modèles plus éloignés de la référence.

La figure 25 résume ces résultats en montrant les poids cumulés des modèles de l'ensemble (un membre par modèle) en ordre décroissant de valeur de poids. Les poids cumulés sont montrés en fonction de la combinaison des huit métriques de moyennes climatologiques (couleur des courbes) et pour différentes valeurs de ψ_D (les quatre panneaux). Chaque courbe peut être interprétée comme le nombre de meilleurs modèles qui dominent le poids total de l'ensemble. Les courbes pleines correspondent à la moyenne des courbes obtenues en utilisant les membres de ClimEx comme des pseudoréférences, tandis que les courbes pointillées représentent le même calcul en utilisant le

jeu de données sur grille de NRCan comme référence.

Pour $\psi_D = 0.5$, les résultats obtenus selon l'approche du modèle parfait (courbes pleines) montrent que l'utilisation d'une seule métrique (climatologie de la température en hiver) implique qu'environ vingt modèles permettent de représenter le poids total de l'ensemble, tandis que pour trois ou quatre métriques, environ dix modèles sont nécessaires. Pour huit métriques, le modèle au poids le plus élevé compte pour environ 95% du poids total. Lorsqu'on augmente ψ_D à 1, 2 et 4, les courbes se rapprochent progressivement de la ligne noire pointillée, qui représente le cas idéalisé d'un ensemble parfaitement démocratique (c.-à-d. où tous les modèles ont des poids égaux). Sur la même figure, les courbes pointillées montrent le même calcul, mais en utilisant les observations comme référence plutôt que la moyenne des courbes obtenues par l'approche du modèle parfait. Nous pouvons remarquer que les courbes associées aux observations suivent un comportement similaire à l'approche du modèle parfait, à la différence que, pour un nombre donné de modèles, ceux-ci représentent généralement une plus grande proportion du poids de l'ensemble lorsque les observations sont utilisées comme référence. Par exemple, pour $\psi_D = 0.5$ et une métrique (courbe rouge), l'approche du modèle parfait montre que les dix meilleurs modèles obtiennent 80% du poids total de l'ensemble, en comparaison avec 94% du poids total lorsque les observations sont utilisées. Autrement dit, le groupe des meilleurs modèles pour reproduire les observations est en général plus restreint que le groupe des meilleurs modèles pour reproduire les résultats de ClimEx. Il faut cependant noter que les courbes individuelles obtenues par la méthode du modèle parfait (non montrées) montrent une incertitude considérable, et donc la variabilité interne peut également expliquer une partie de ce résultat.

De manière générale, le fait d'augmenter le nombre de métriques d'évaluation mène à un résultat semblable à une diminution du rayon de similarité (ici via ψ_D), c'est-à-dire que cela augmente le poids donné à un nombre de plus en plus restreint de modèles situés près de la référence.

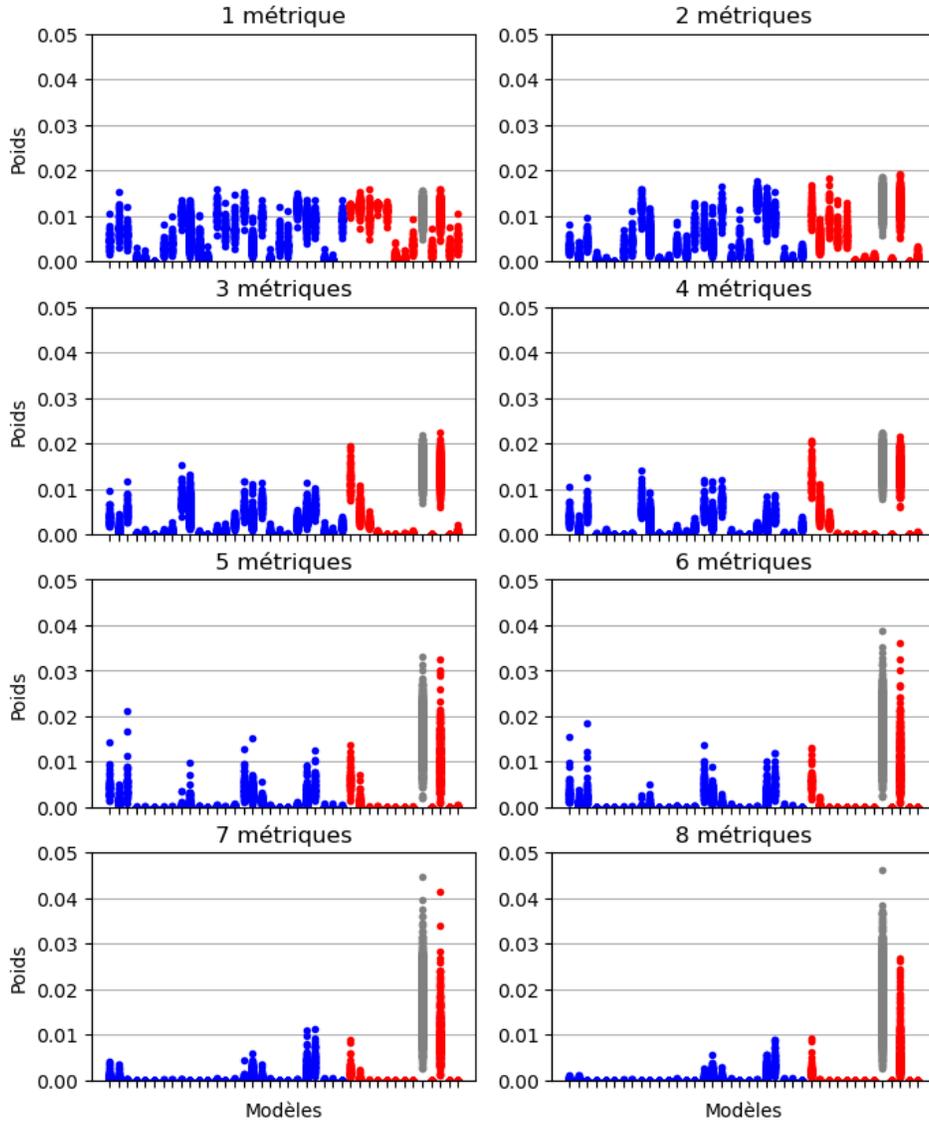


FIGURE 22 – Poids calculés à partir de combinaisons de huit métriques de moyennes climatologiques pour la période 1971-2000 en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Dans l'ordre sont combinés par multiplication les poids de hiver à automne pour la température moyenne (une à quatre métriques), en ajoutant ensuite les quatre saisons pour la précipitation (cinq à huit métriques). Données brutes pour CMIP5 (bleu), CORDEX (rouge) et ClimEX (gris). Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2. $\psi_D = 1$.

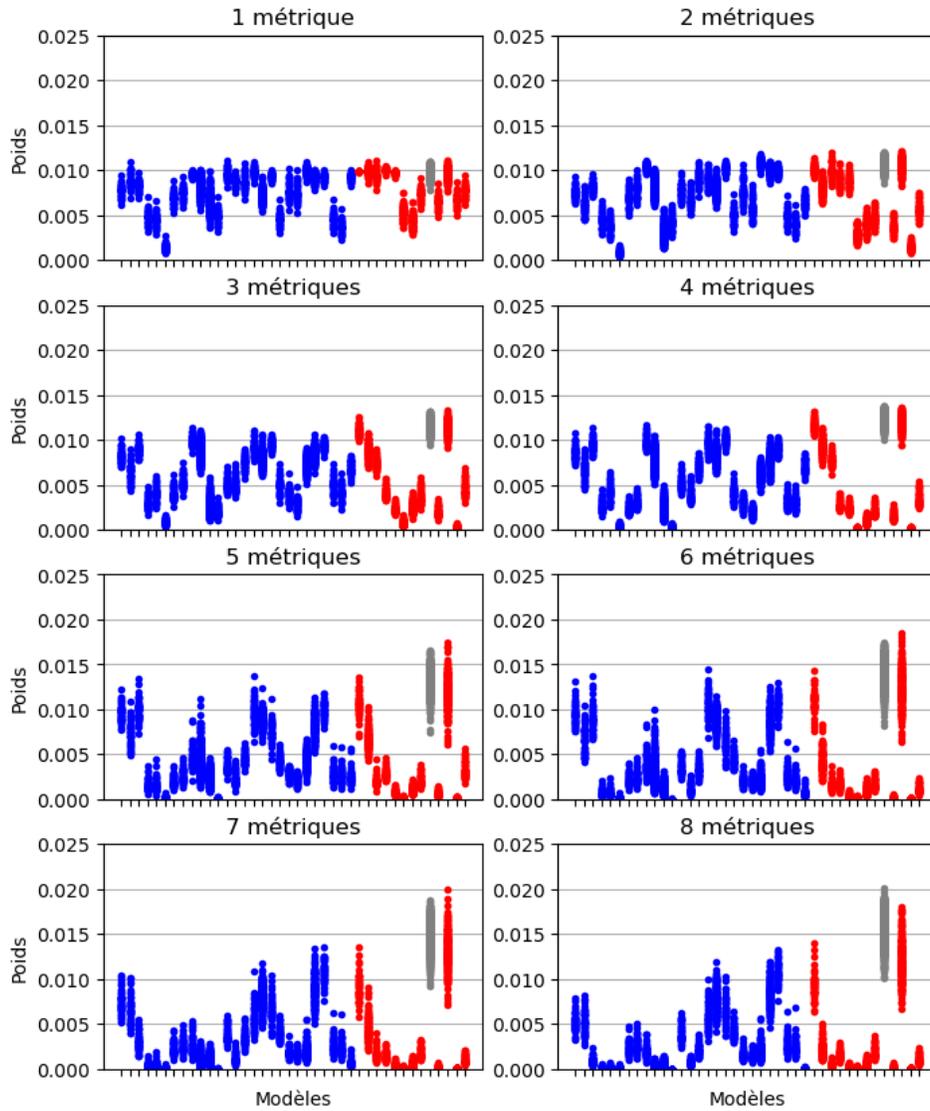


FIGURE 23 – Poids calculés à partir de combinaisons de huit métriques de moyennes climatologiques pour la période 1971-2000 en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Dans l'ordre sont combinés par multiplication les poids de hiver à automne pour la température moyenne (une à quatre métriques), en ajoutant ensuite les quatre saisons pour la précipitation (cinq à huit métriques). Données brutes pour CMIP5 (bleu), CORDEX (rouge) et ClimEX (gris). Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2. $\psi_D = 2$.

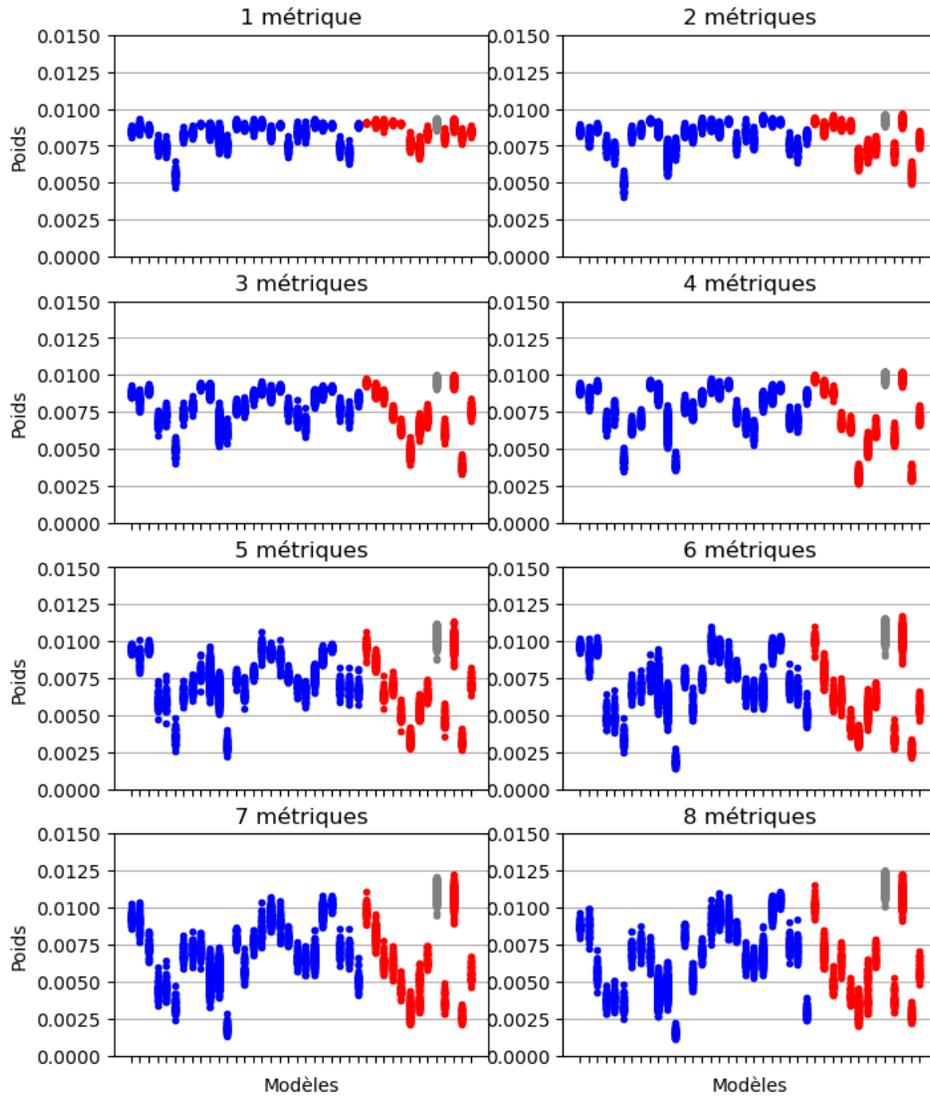


FIGURE 24 – Poids calculés à partir de combinaisons de huit métriques de moyennes climatologiques pour la période 1971-2000 en utilisant à tour de rôle les 50 membres de ClimEx comme des pseudoréférences. Dans l'ordre sont combinés par multiplication les poids de hiver à automne pour la température moyenne (une à quatre métriques), en ajoutant ensuite les quatre saisons pour la précipitation (cinq à huit métriques). Données brutes pour CMIP5 (bleu), CORDEX (rouge) et ClimEX (gris). Les modèles distribués sur l'axe horizontal apparaissent dans le même ordre qu'à la table 2. $\psi_D = 4$.

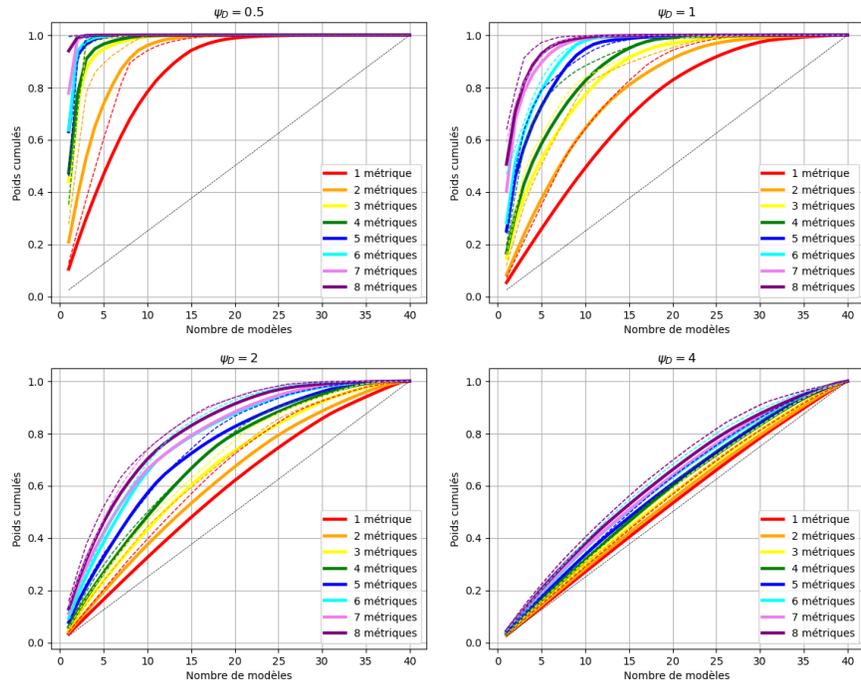


FIGURE 25 – Poids cumulés des modèles ordonnés en ordre décroissant de leur performance, en fonction du nombre de métriques combinées de 1 à 8 (de hiver à automne pour température, et ensuite en ajoutant les quatre saisons pour la précipitation), pour des valeurs de $\psi_D = 0.5, 1, 2$ et 4 . Le premier membre de chacun des 40 modèles de la table 2 est utilisé. Chaque courbe pleine correspond à la moyenne des courbes obtenues en utilisant les 49 membres résiduels de ClimEx comme des pseudoréférences. Les courbes pointillées représentent le même calcul obtenu en utilisant les observations de NRCan comme référence. La droite pointillée (noire) correspond aux poids cumulés pour des modèles à poids égaux.

5.3.5 Impact de la pondération a posteriori

Les résultats des sections précédentes ont montré que les poids attribués aux modèles peuvent dépendre fortement de plusieurs paramètres libres. Nous allons ici quantifier l’impact de ces paramètres lorsque les poids sont utilisés pour pondérer des deltas des changements climatiques en utilisant les observations comme référence. Pour ce faire, la figure 26 reprend les diagrammes de dispersion de la figure 15 pour les deltas de température et précipitation (2071-2100 par rapport à 1981-2010) pour les quatre saisons, mais en utilisant certaines des configurations de la fonction EXP comparées précédemment.

Analysons d’abord les deltas obtenus pour la saison d’hiver sur la figure 26. Le cercle vert représente la moyenne “un modèle un vote binaire” (premier membre utilisé pour chaque modèle de la table 2) que nous utiliserons comme point de repère pour comparer avec les moyennes pondérées de la figure 15. Les triangles rouges montrent l’évolution de la moyenne pondérée à un membre par modèle selon la métrique de la moyenne climatologique de la variable et saison étudiée (ici, la température et la précipitation en hiver). Les différents symboles (dans l’ordre de rotation horaire) montrent l’évolution de la moyenne d’ensemble lorsqu’on augmente le rayon de similarité de $\psi_D = 0.5$ à 4. En augmentant le rayon de similarité, la moyenne d’ensemble converge progressivement vers la moyenne démocratique de un modèle un vote. Les symboles mauves représentent le même calcul, mais en utilisant les huit métriques climatologiques. Le comportement de convergence vers la moyenne démocratique est similaire à l’utilisation d’une seule métrique, mais les variations du signal sont plus grandes avec l’utilisation de huit métriques. Tel que montré à la figure 25, l’augmentation du nombre de métriques tend vers une surconfiance dans un petit nombre de modèles, et donc la moyenne pondérée devient plus instable puisque basée sur un petit nombre effectif de modèles. Ces résultats sont essentiellement reproduits pour les autres saisons.

Il est intéressant de comparer ces résultats à ceux de la figure 15, où des pondérations a priori avaient été analysées. En général, les variations du rayon de similarité et du nombre de métriques ont un impact similaire sur les moyennes pondérées en comparaison avec les différentes techniques de sélection a priori. Cependant, dans certains cas, les méthodes a priori et a posteriori sont complémentaires en termes de couverture de l’incertitude des moyennes pondérées. Par exemple, au printemps, les pondérations a posteriori étendent la plage de valeurs couvertes par le delta en température par environ 1°C vers des valeurs plus froides, et par environ un demi-degré pour l’été et l’automne. L’incertitude des méthodes a posteriori permet également

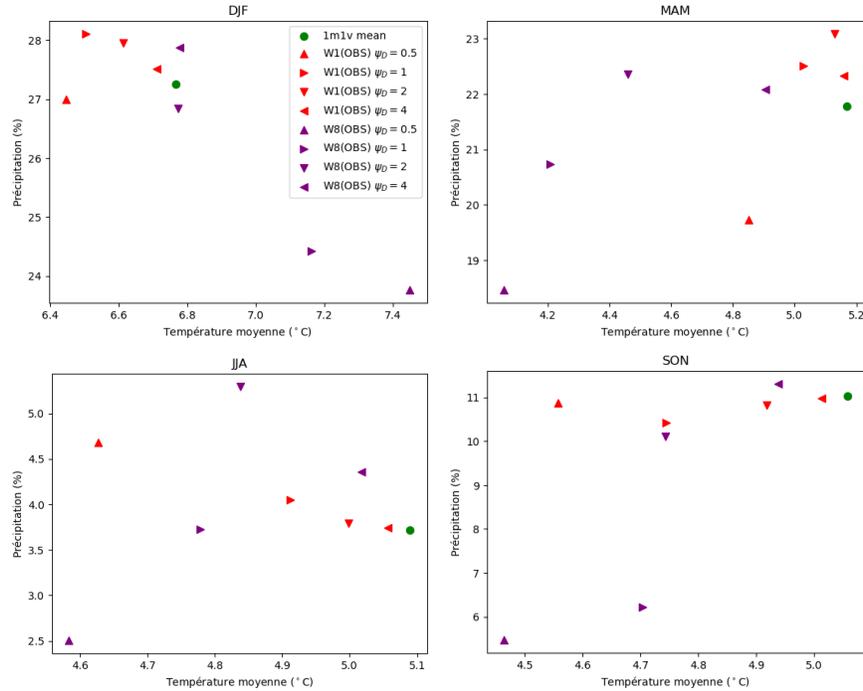


FIGURE 26 – Delta saisonnier en 2071-2100 par rapport à 1981-2010 pour différentes méthodes de pondération a posteriori en utilisant les observations comme référence, pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), selon RCP 8.5, données brutes. Une simulation par modèle est utilisée pour chacun des 40 modèles de la table 2.

d’augmenter la plage couverte par les deltas de précipitations, mais par des valeurs relativement faibles.

5.3.6 Contraindre les projections régionales à l’aide de métriques globales

Les sections précédentes ont permis d’évaluer la sensibilité de la pondération en fonction de la sélection des métriques d’évaluation des modèles. Un choix implicite n’ayant pas été pris en compte est celui de l’échelle spatiale associée aux métriques d’évaluation, qui se concentrait jusqu’ici exclusivement sur la région du Québec méridional, région d’intérêt pour Info-Crue. Or, il serait tout à fait acceptable d’utiliser des métriques d’évaluation à des échelles spatiales plus grandes que la région d’intérêt, pourvu que des liens

physiques puissent être établis entre les deux échelles. Dans le cadre du dernier rapport du GIEC (IPCC, 2021), il a été reconnu que certains modèles de l'ensemble CMIP6 ont des sensibilités climatiques trop élevées par rapport à ce que suggèrent plusieurs évidences observationnelles et théoriques (Sherwood *et al.*, 2020). En appliquant des contraintes sur la sensibilité climatique des modèles, il devient alors possible d'identifier un sous-ensemble de modèles dont les projections climatiques sont considérées plus probables. Cette approche a été appliquée dans le cadre du dernier rapport du GIEC pour contraindre des variables à l'échelle globale, mais il n'est pas clair si elle peut aussi permettre de contraindre l'incertitude des projections climatiques à l'échelle régionale.

La sensibilité climatique est souvent quantifiée à l'aide de deux métriques globales. D'abord, le TCR (pour "Transient Climate Sensitivity" en anglais), représente le changement dans la température globale résultant d'un doublement de la concentration de CO₂ dans l'atmosphère, en suivant un scénario d'augmentation de cette concentration à 1% par année (le doublement du CO₂ s'effectue après 70 ans). Le TCR se distingue de l'ECS ("Equilibrium Climate Sensitivity"), qui représente le changement dans la température globale du système climatique à l'équilibre après un doublement instantané du CO₂.

Afin d'évaluer le potentiel d'utiliser ces métriques globales pour contraindre les tendances futures de changement climatique sur le Québec méridional, la figure 32 montre le diagramme de dispersion du delta futur en température et précipitation sur le Québec méridional en fonction des métriques globales TCR et ECS. Ces figures montrent une relation de proportionnalité entre les deux métriques globales et les tendances régionales de température pour toutes les saisons. Les relations sont toutefois relativement faibles, tel qu'indiqué par les valeurs du coefficient de détermination se situant entre 0.2 et 0.61. Il est aussi intéressant de noter que les relations linéaires pour le changement régional de température sont légèrement plus fortes pour l'ECS que le TCR, ce qui peut paraître contre-intuitif étant donné les échelles de temps impliquées. Ce résultat a également aussi été mis en évidence par Grose *et al.* (2018). Il est toutefois utile de noter que les valeurs de ECS et TCR sont proportionnelles entre elles malgré les différences dans leurs définitions (voir Fig. 31 de l'annexe B). Pour le changement régional dans la précipitation, il ne montre aucune relation de proportionnalité avec l'ECS et le TCR.

Cette analyse est un bon exemple pour la seconde hypothèse de sélection des métriques afin de guider l'ajustement de la fonction de pondération (voir Sect. 5.3.2), soit celui de la fiabilité des métriques pour contraindre les projections futures. En effet, le fait que les tendances régionales de température

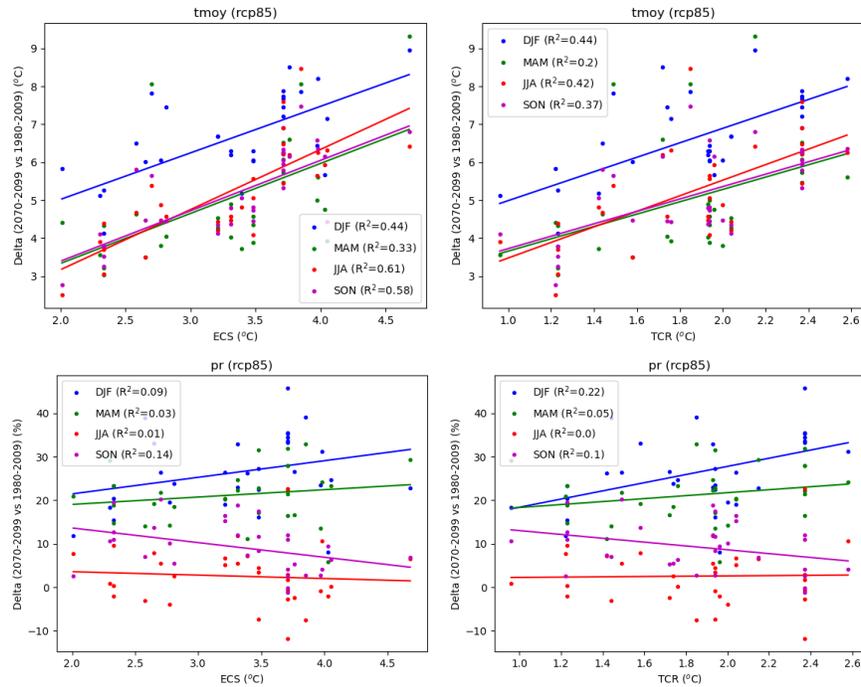


FIGURE 27 – Delta de température (haut) et précipitation (bas) entre 1980-2009 et 2070-2099 pour les quatre saisons en fonction des valeurs de ECS (gauche) et de TCR (droite), pour les données brutes des simulations suivant le RCP8.5, avec une simulation par modèle (31) de la table 2 dont les valeurs d’ECS et de TCR du MCG étaient disponibles (Flynn et Mauritsen, 2020).

montrent une certaine corrélation avec la sensibilité climatique globale, et ce malgré le bruit de la variabilité interne, suggère que nous pouvons nous fier dans une certaine mesure sur les métriques d’ECS et TCR pour contraindre les projections régionales de température.

Chapitre 6

Discussion

Dans le cadre de ce projet, nous avons comparé plusieurs approches pour pondérer les projections climatiques de l'ensemble de simulations utilisé dans le cadre du projet Info-Crue. L'ensemble de 179 simulations climatiques était disponible sous forme de sorties brutes de modèles ainsi que de scénarios post-traités utilisés pour piloter des modèles hydrologiques. Bien que les résultats du présent rapport se concentrent exclusivement sur des simulations climatiques extraites sur le domaine du Québec méridional, et que certaines des conclusions pourraient différer pour d'autres régions et applications, ce projet a permis de mieux comprendre les potentiels et limites concernant l'application de schémas de pondération à des simulations climatiques à l'échelle régionale.

Dans un premier temps, les données brutes et les scénarios (données post-traitées) des variables de température moyenne et de précipitation ont été évaluées par rapport à leurs caractéristiques de base en climat historique : moyenne climatologique, variabilité interannuelle et tendance de changement climatique. Les performances des différents modèles en climat historique ont été quantifiées en comparant chaque simulation avec le jeu d'observations sur grille de NRCan. Par construction, les scénarios sont très similaires à la référence pour les moyennes climatologiques, tandis que la variabilité interannuelle et les tendances sont peu améliorées par rapport aux sorties brutes des modèles. À partir des sorties brutes, il a été constaté que la variabilité interannuelle et les tendances sont fortement influencées par la variabilité interne du climat telle qu'obtenue à l'aide du grand ensemble ClimEx. À l'opposé, la variabilité interne a peu d'impact sur l'évaluation des moyennes climatologiques, suggérant que celles-ci puissent être utilisées comme métriques de performances des modèles à l'intérieur d'un schéma de pondération. Or,

bien que les métriques d'évaluation associées aux moyennes climatologiques soient robustes par rapport à la variabilité interne, il est attendu que celles-ci contiennent peu d'information à propos de la capacité des modèles à prévoir le climat futur. D'un autre côté, il a été démontré que les tendances historiques sont fortement influencées par la variabilité interne à l'échelle du Québec, et bien que celles-ci pourraient contenir de l'information très utile à propos des tendances climatiques dans le futur, le fort niveau de variabilité interne rend cette information difficilement utilisable dans le cadre d'un schéma de pondération.

Dans la suite de l'analyse, nous avons évalué l'impact de différentes approches possibles pour pondérer un ensemble de projections climatiques en utilisant les simulations brutes. Nous avons d'abord comparé différentes méthodes dites *a priori*, qui consistent en des stratégies de sélection des simulations à partir de caractéristiques connues des modèles, comme le type de modèle ou d'expérience, ou encore la résolution spatiale. Les résultats montrent une faible sensibilité de la moyenne d'ensemble des changements climatiques simulés, et ce malgré de fortes hypothèses posées lors de l'étape de sélection, ce qui est dû en partie au fait qu'un grand nombre de modèles et de simulations sont disponibles. Ces hypothèses de sélection doivent cependant être considérées avec grande attention lorsque des ensembles de plus petite taille sont utilisés. Par exemple, pour une application qui nécessiterait que la variabilité spatiale en lien avec l'hétérogénéité de la surface soit basée sur une modélisation des processus physiques comme pour les lacs et la cryosphère, ou encore pour représenter les aléas météorologiques aux échelles de quelques kilomètres, une sélection priorisant les modèles MRC pourrait être toute indiquée, bien qu'il existe tout de même à ces échelles une valeur potentielle ajoutée par les nouvelles méthodes de post-traitement multivariées des simulations de MGC (p. ex. Cannon 2016).

Nous avons ensuite évalué la sensibilité d'une pondération *a posteriori* par rapport à ses paramètres libres, c'est-à-dire en variant les aspects basés sur des choix d'experts. Après une brève comparaison entre deux méthodes de pondération, des expériences de sensibilité ont été appliquées à une équation de pondération découlant des travaux de Sanderson *et al.* (2015), laquelle a été utilisée à de nombreuses reprises dans la littérature récente (e.g. Knutti *et al.* 2017; Sanderson *et al.* 2017; Lorenz *et al.* 2018; Dalelane *et al.* 2018). La méthode inspirée de Sanderson *et al.* (2015) a été sélectionnée, offrant plus de flexibilité pour intégrer le concept de la variabilité interne dans la pondération que des méthodes plus anciennes. Des expériences de sensibilité ont ensuite été réalisées en utilisant l'approche du modèle parfait, soit en utilisant les 50 membres de l'ensemble ClimEx comme des pseudoréférences

interchangeables, permettant ainsi d'évaluer l'impact de la variabilité interne sur le calcul des poids.

Nous avons évalué la sensibilité de la fonction de pondération à la valeur du rayon de similarité. Nous avons d'abord utilisé la variabilité interne comme mesure de base de ce paramètre. En utilisant différents multiples de la valeur de base, nous avons démontré comment une augmentation du rayon de similarité faisait tendre les modèles vers des poids égaux, et donc vers le paradigme de la démocratie des modèles connu sous le nom de *un modèle un vote*. Ensuite, nous avons comparé les poids obtenus lorsque différentes statistiques climatiques sont utilisées comme métriques d'évaluation, soit la moyenne climatologique, la variabilité interannuelle et les tendances de changements climatiques. Cette comparaison a permis de démontrer que les poids des modèles tendent à être égaux lorsque la variabilité interne devient comparable à l'incertitude entre les modèles. Dans une telle situation, par exemple pour les tendances climatiques régionales qui sont particulièrement influencées par la variabilité interne, l'approche du modèle parfait a permis de démontrer que des poids différents n'étaient pas nécessairement dus à des différences détectables de performances. Ce résultat suggère que la variabilité interne devrait également, en plus de l'incertitude inter-modèle, être prise en compte dans la définition du rayon de similarité. Il est à noter que la littérature traitant du sujet de la pondération ne tient pas explicitement compte de la variabilité interne pour définir le rayon de similarité (Merrifield *et al.*, 2020). Comme dernière expérience de sensibilité, nous avons utilisé les métriques les plus robustes (soit celles associées aux moyennes climatologiques saisonnières de la température et la précipitation) pour évaluer l'impact d'une combinaison de plusieurs métriques pour obtenir un poids par modèle. L'impact général de combiner des métriques est généralement de converger progressivement vers une surconfiance dans un petit nombre de modèles (voire un seul) qui sont les plus proches de la référence. Combiner un grand nombre de métriques (p. ex. pour les huit moyennes climatologiques) a donc un effet similaire à diminuer le rayon de similarité. À l'opposé, un petit nombre de métriques et un grand rayon de similarité font tendre la distribution des poids vers le paradigme de la démocratie des modèles. Il est à noter que d'autres méthodes pourraient être utilisées pour combiner les métriques, par exemple par addition plutôt que par multiplication (Éq. 5), cet aspect n'ayant pas été évalué dans le cadre de ce rapport.

En comparant l'impact des méthodes *a priori* et *a posteriori* sur les deltas de changement climatique pour la fin du 21^e siècle, nous avons constaté que les deux types d'approches étaient complémentaires en termes de contributions à l'incertitude de la moyenne pondérée de l'ensemble. La complé-

mentarité de ces approches est aussi pertinente pour intégrer deux aspects importants d’une pondération, soit la performance et la dépendance (ou unicité) des modèles. En effet, la méthode de Sanderson *et al.* (2015) et Knutti *et al.* (2017) intègre déjà ces deux composantes dans une même équation, tandis que seulement le terme relié à la performance a été évalué quantitativement dans le cadre du présent rapport. L’aspect de la dépendance des modèles était ici plutôt évalué qualitativement à partir des méthodes de sélection a priori, par exemple la méthode “un centre un vote” (Leduc *et al.*, 2016), qui reste un compromis défendable face à une utilisation naïve de tous les modèles disponibles (Abramowitz *et al.*, 2019).

La dernière partie de l’analyse s’est penchée sur l’utilisation de métriques globales comme l’ECS et le TCR pour contraindre les projections climatiques à l’échelle régionale. Bien qu’il existe une relation entre les sensibilités climatiques globale et régionale, il est nécessaire de d’abord contraindre les valeurs d’ECS ou TCR afin d’appliquer ensuite la même pondération à l’échelle régionale. Selon Sherwood *et al.* (2020), l’intervalle de confiance entre les 5-95^e percentiles se situe entre 2.3-4.7°C, délimité par une plage entre 2-5.7°C. Dans le dernier rapport du GIEC (IPCC, 2021), l’intervalle de l’ECS considéré comme probable est de 2.5 à 4°C (Hausfather *et al.*, 2022). Les modèles utilisés pour produire la figure 32 indiquent des valeurs d’ECS qui varient entre 2.01°C et 4.68°C (Flynn et Mauritsen, 2020), et donc utiliser l’intervalle probable de Hausfather (2.5-4°C) offre un potentiel assez limité pour réduire l’incertitude dans les projections régionales de température sur le Québec méridional. Cependant, cette analyse montre un potentiel accru pour réduire les incertitudes des projections régionales dans le contexte de CMIP6, où plus du quart des modèles ont un ECS plus grand que 4.7°C (Hausfather *et al.*, 2022). De telles valeurs semblent en désaccord avec les évidences théoriques et observationnelles (Hausfather *et al.*, 2022; Sherwood *et al.*, 2020). Dans le cadre d’une analyse effectuée en marge de ce rapport (annexe D), l’ensemble CMIP6 semble effectivement montrer un plus grand potentiel pour contraindre avec l’ECS les projections climatiques sur le Québec méridional.

Chapitre 7

Conclusion

Ce projet a permis de procéder à une évaluation détaillée de l'intérêt de pondérer des projections climatiques sur le Québec méridional. Nous avons démontré que le paradigme de la démocratie des modèles est difficile à discréditer pour des simulations à l'échelle régionale. De manière générale, utiliser un petit nombre de métriques d'évaluation et un grand rayon de similarité tendent à maintenir le statu quo du traitement des ensembles, dont la communauté en science du climat est assujettie depuis plusieurs années (p. ex. Whetton *et al.* 2007). D'un autre côté, utiliser plusieurs métriques d'évaluation et un petit rayon de similarité mène à une surconfiance dans un petit groupe de modèles, ce qui peut être un pari risqué dans un contexte d'adaptation aux changements climatiques. Nous recommandons donc un compromis, soit un ensemble minimal de métriques soigneusement sélectionnées en fonction de l'application, et surtout qui soient robustes par rapport à la variabilité interne et fiables pour contraindre les projections futures. Un constat important de la présente étude est que l'utilisation de grands ensembles (p. ex. ClimEx) comme des pseudoréférences s'avère pratiquement indispensable pour estimer l'impact de la variabilité interne, souvent déterminante, dans le calcul des poids.

Au niveau de la recherche future, les méthodes de pondérations basées sur les performances devront tenir compte explicitement de la variabilité interne afin de moduler l'intensité de la pondération. Au cours des prochaines années, il sera également important de réaliser des travaux afin d'identifier les métriques les plus adéquates pour contraindre les projections climatiques dans différents contextes, plutôt que de tenter d'identifier une seule et même pondération applicable de manière générique. En guise d'exemple, l'utilisation de l'ECS ou du TCR pour contraindre les projections régionales de

température semble prometteuse, mais l’application de ces mêmes métriques pour contraindre les projections de précipitations est discutable lorsqu’aucun lien n’est établi entre cette variable et la métrique globale. Par exemple, l’élimination des “modèles chauds” de CMIP6 pour contraindre les projections régionales de précipitation peut mener à une détérioration de la diversité des projections de l’ensemble et une diminution injustifiée de l’incertitude.

Dans le contexte d’Info-Crue, il est particulièrement utile de se demander si des contraintes globales comme l’ECS et le TCR peuvent engendrer des bénéfices de réduction des incertitudes pour des projections hydrologiques. Par exemple, les liens physiques entre l’ECS, la tendance régionale future de température en hiver, et la fonte du couvert de neige seraient une suite de processus physiques pertinente à étudier pour évaluer le potentiel pour améliorer les projections hydroclimatiques associées aux crues printanières. Dans ce contexte, la moyenne climatologique (état initial en climat historique) de la quantité de neige au sol pourrait également aider à contraindre les projections. Dans des travaux de recherche futurs, il pourrait être utile de travailler à établir des ponts permettant de suivre les contraintes au travers de l’évolution des processus physiques entre les échelles globale, régionale et hydrologique. D’autres aspects à explorer pour établir de telles chaînes de processus entre les échelles pourraient se concentrer sur des phénomènes globaux ou régionaux reliés à la cryosphère, notamment la neige et la glace de mer, qui peuvent potentiellement influencer certaines tendances régionales (Senftleben *et al.*, 2020; Borodina *et al.*, 2017a). D’autres pistes de recherche seraient également d’intérêt pour l’hydrologie. Par exemple, l’amélioration de la variabilité spatiale dans les MGC post-traités et une comparaison avec des MRC.

Finalement, une application directe des résultats de ce rapport en hydrologie pourrait consister à identifier un sous-ensemble de modèles MGC et MRC pour lesquels les biais climatologiques sont minimisés. Bien que la portée d’utiliser les climatologies soit limitée pour influencer les projections futures, ce type de sélection pourrait permettre de piloter des modèles hydrologiques à l’aide de simulations brutes, et donc d’éviter certains défis du post-traitement (contexte multivarié, choix des données de référence). Les sorties climatiques associées à un plus grand nombre de variables météorologiques, radiation solaire, humidité ou vents, pourraient alors être utilisées tout en s’assurant de leur cohérence physique, qui est l’une des grandes forces des modèles de climat.

Chapitre 8

Bibliographie

- ABDALLA, S., ISAKSEN, L., JANSSEN, P. et WEDI, N. (2013). Effective spectral resolution of ecmwf atmospheric forecast models.
- ABRAMOWITZ, G., HERGER, N., GUTMANN, E., HAMMERLING, D., KNUTTI, R., LEDUC, M., LORENZ, R., PINCUS, R. et SCHMIDT, G. A. (2019). Esd reviews : Model dependence in multi-model climate ensembles : Weighting, sub-selection and out-of-sample testing. *Earth System Dynamics*, 10(1): 91–105.
- BISHOP, C. H. et ABRAMOWITZ, G. (2013). Climate Model Dependence and the Replicate Earth Paradigm. *Climate Dynamics*, 41(3-4):885–900.
- BOÉ, J. (2018). Interdependency in multimodel climate projections : Component replication and result similarity. *Geophysical Research Letters*, 45(6):2771–2779.
- BORODINA, A., FISCHER, E. M. et KNUTTI, R. (2017a). Emergent Constraints in Climate Projections : A Case Study of Changes in High-Latitude Temperature Variability. *Journal of Climate*, 30(10):3655–3670.
- BORODINA, A., FISCHER, E. M. et KNUTTI, R. (2017b). Potential To Constrain Projections of Hot Temperature Extremes. *Journal of Climate*, pages JCLI-D-16-0848.1.
- BRUNNER, L., PENDERGRASS, A. G., LEHNER, F., MERRIFIELD, A. L., LORENZ, R. et KNUTTI, R. (2020). Reduced global warming from cmip6 projections when weighting models by performance and independence. *Earth System Dynamics*, 11(4):995–1012.

- CANNON, A. J. (2016). Multivariate Bias Correction of Climate Model Output : Matching Marginal Distributions and Intervariable Dependence Structure. *Journal of Climate*, 29(19):7045–7064.
- CHEN, J., BRISSETTE, F. P., LIU, P. et XIA, J. (2017). Using Raw Regional Climate Model Outputs for Quantifying Climate Change Impacts on Hydrology. *Hydrological Processes*, 31(24):4398–4413.
- CHRISTENSEN, J. H. et BOBERG, F. (2012). Temperature dependent climate projection deficiencies in cmip5 models. *Geophysical Research Letters*, 39(24).
- CHRISTENSEN, J. H., KJELLSTRÖM, E., GIORGI, F., LENDERINK, G. et RUMMUKAINEN, M. (2010). Weight Assignment in Regional Climate models. *Climate Research*, 44(2-3):179–194. 10.3354/cr00916.
- DALELANE, C., FRÜH, B., STEGER, C. et WALTER, A. (2018). A pragmatic approach to build a reduced regional climate projection ensemble for germany using the euro-cordex 8.5 ensemble. *Journal of Applied Meteorology and Climatology*, 57(3):477–491.
- DÉQUÉ, M., ROWELL, D., LÜTHI, D., GIORGI, F., CHRISTENSEN, J., ROCKEL, B., JACOB, D., KJELLSTRÖM, E., de CASTRO, M. et van den HURK, B. (2007). An Intercomparison of Regional Climate Simulations for Europe : Assessing Uncertainties in Model Projections. *Climatic Change*, 81(0):53–70.
- EUM, H.-I., GACHON, P. et LAPRISE, R. (2016). Impacts of model bias on the climate change signal and effects of weighted ensembles of regional climate model simulations : A case study over southern québec, canada. *Advances in Meteorology*, 2016(nil):1–17.
- FLYNN, C. M. et MAURITSEN, T. (2020). On the climate sensitivity and historical warming evolution in recent coupled model ensembles. *Atmospheric Chemistry and Physics*, 20(13):7829–7842.
- FORSTER, P., STORELMO, T., ARMOUR, K., COLLINS, W., DUFRESNE, J.-L., FRAME, D., LUNT, D., MAURITSEN, T., PALMER, M., WATANABE, M., WILD, M. et ZHANG, H. (2021). *The Earth’s Energy Budget, Climate Feedbacks, and Climate Sensitivity*. In *Climate Change 2021 : The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N.*

Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnay, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

- FOWLER, H. J., BLENKINSOP, S. et TEBALDI, C. (2007). Linking Climate Change Modelling To Impacts Studies : Recent Advances in Downscaling Techniques for Hydrological Modelling. *International Journal of Climatology*, 27(12):1547–1578.
- GIORGI, F. et MEARNS, L. O. (2002). Calculation of Average, Uncertainty Range, and Reliability of Regional Climate Changes From AOGCM Simulations Via the "Reliability Ensemble Averaging" (REA) Method. *Journal of Climate*, 15(10):1141–1158.
- GIUNTOLI, I., VILLARINI, G., PRUDHOMME, C. et HANNAH, D. M. (2018). Uncertainties in projected runoff over the conterminous united states. *Climatic Change*, 150(3):149–162.
- GLECKLER, P. J., TAYLOR, K. E. et DOUTRIAUX, C. (2008). Performance Metrics for Climate Models. *J. Geophys. Res.*, 113(D6):1984–2012.
- GROSE, M. R., GREGORY, J., COLMAN, R. et ANDREWS, T. (2018). What climate sensitivity index is most useful for projections? *Geophysical Research Letters*, 45(3):1559–1566.
- HALL, A., COX, P., HUNTINGFORD, C. et KLEIN, S. (2019). Progressing emergent constraints on future climate change. *Nature Climate Change*, 9(4):269–278.
- HAUSFATHER, Z., MARVEL, K., SCHMIDT, G. A., NIELSEN-GAMMON, J. W. et ZELINKA, M. (2022). Climate simulations : Recognize the 'hot model' problem.
- HAWKINS, E. et SUTTON, R. (2009). The Potential To Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1107.
- HAWKINS, E. et SUTTON, R. (2011). The Potential To Narrow Uncertainty in Projections of Regional Precipitation Change. *Climate Dynamics*, 37(1-2):407–418.
- HUTCHINSON, M. F., MCKENNEY, D. W., LAWRENCE, K., PEDLAR, J. H., HOPKINSON, R. F., MILEWSKA, E. et PAPADOPOULOS, P. (2009). Development and Testing of Canada-Wide Interpolated Spatial Models of Daily

Minimum-Maximum Temperature and Precipitation for 1961-2003. *Journal of Applied Meteorology and Climatology*, 48(4):725 – 741.

IPCC (2007). *Climate Change 2007 : The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

IPCC (2013). *Climate Change 2013 : The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2021). *Climate Change 2021 : The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]*. Cambridge University Press. In Press.

KLAVER, R., HAARSMA, R., VIDALE, P. L. et HAZELEGER, W. (2020). Effective resolution in high resolution global atmospheric models for climate studies. *Atmospheric Science Letters*, 21(4):e952.

KNUTTI, R., FURRER, R., TEBALDI, C., CERMAK, J. et MEEHL, G. A. (2010). Challenges in Combining Projections From Multiple Climate Models. *Journal of Climate*, 23(10):2739–2758.

KNUTTI, R., MASSON, D. et GETTELMAN, A. (2013). Climate Model Genealogy : Generation CMIP5 and How We Got There. *Geophysical Research Letters*, 40(6):1194–1199.

KNUTTI, R., SEDLÁČEK, J., SANDERSON, B. M., LORENZ, R., FISCHER, E. M. et EYRING, V. (2017). A Climate Model Projection Weighting Scheme Accounting for Performance and Interdependence. *Geophysical Research Letters*.

LAVOIE, J., BOURGAULT, P., LOGAN, T., CARON, L.-P., GAMMON, S., SMITH, T. J., BINER, S. et LEDUC, M. (2023). ESPO-G6-E5L : Ensemble

de scénarios polyvalents d'Ouranos - Modèles Globaux CMIP6 - ERA5-Land / Ouranos Multipurpose Climate Scenarios - Global Models CMIP6 - ERA5-Land.

LEDUC, M., LAPRISE, R., de ELÍA, R. et SEPAROVIC, L. (2016). Is Institutional Democracy a Good Proxy for Model Independence? *Journal of Climate*.

LEDUC, M., MAILHOT, A., FRIGON, A., MARTEL, J.-L., LUDWIG, R., BRIETZKE, G. B., GIGUÈRE, M., BRISSETTE, F., TURCOTTE, R., BRAUN, M. et SCINOCCA, J. (2019). The Climex Project : A 50-member Ensemble of Climate Change Projections At 12-km Resolution Over Europe and Northeastern North America With the Canadian Regional Climate Model (CRCM5). *Journal of Applied Meteorology and Climatology*, 58(4):663–693.

LEE, J.-Y., MAROTZKE, J., BALA, G., CAO, L., CORTI, S., DUNNE, J., ENGELBRECHT, F., FISCHER, E., FYFE, J., JONES, C., MAYCOCK, A., MUTEMI, J., NDIAYE, O., PANICKAL, S. et ZHOU, T. (2021). *Future Global Climate : Scenario-Based Projections and Near-Term Information*. In *Climate Change 2021 : The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]*. Cambridge University Press. In Press.

LORENZ, R., HERGER, N., SEDLÁČEK, J., EYRING, V., FISCHER, E. M. et KNUTTI, R. (2018). Prospects and caveats of weighting climate models for summer maximum temperature projections over north america. *Journal of Geophysical Research : Atmospheres*, 123(9):4509–4526.

LUCAS-PICHER, P., ARSENAULT, R., POULIN, A., RICARD, S., LACHANCE-CLOUTIER, S. et TURCOTTE, R. (2020). Application of a high-resolution distributed hydrological model on a u.s.-canada transboundary basin : Simulation of the multiyear mean annualhydrograph and 2011 flood of therichelieu river basin. *Journal of Advances in Modeling Earth Systems*, 12(4):e2019MS001709. e2019MS001709 2019MS001709.

LUCAS-PICHER, P., LACHANCE-CLOUTIER, S., ARSENAULT, R., POULIN, A., RICARD, S., TURCOTTE, R. et BRISSETTE, F. (2021). Will evolving

climate conditions increase the risk of floods of the large u.s.-canada trans-boundary richelieu river basin? *JAWRA Journal of the American Water Resources Association*, 57(1):32–56.

MASSON, D. et KNUTTI, R. (2011). Climate Model Genealogy. *Geophys. Res. Lett.*, 38(8):L08703.

MCKENNEY, D. W., HUTCHINSON, M. F., PAPADOPOL, P., LAWRENCE, K., PEDLAR, J., CAMPBELL, K., MILEWSKA, E., HOPKINSON, R. F., PRICE, D. et OWEN, T. (2011). Customized Spatial Climate Models for North America. *Bulletin of the American Meteorological Society*, 92(12):1611–1622. 1611.

MEINSHAUSEN, M., SMITH, S. J., CALVIN, K., DANIEL, J. S., KAINUMA, M. L. T., LAMARQUE, J. F., MATSUMOTO, K., MONTZKA, S. A., RAPER, S. C. B., RIAHI, K., THOMSON, A., VELDEERS, G. J. M. et VUUREN, D. P. P. (2011). The Rcp Greenhouse Gas Concentrations and Their Extensions From 1765 To 2300. *Climatic Change*, 109(1-2):213–241.

MERRIFIELD, A. L., BRUNNER, L., LORENZ, R., MEDHAUG, I. et KNUTTI, R. (2020). An investigation of weighting schemes suitable for incorporating large ensembles into multi-model ensembles. *Earth System Dynamics*, 11(3):807–834.

MUERTH, M. J., GAUVIN ST-DENIS, B., RICARD, S., VELÁZQUEZ, J. A., SCHMID, J., MINVILLE, M., CAYA, D., CHAUMONT, D., LUDWIG, R. et TURCOTTE, R. (2013). On the Need for Bias Correction in Regional Climate Scenarios To Assess Climate Change Impacts on River Runoff. *Hydrology and Earth System Sciences*, 17(3):1189–1204.

NOTZ, D. et COMMUNITY, S. (2020). Arctic sea ice in cmip6. *Geophysical Research Letters*, 47(10):e2019GL086749. e2019GL086749 10.1029/2019GL086749.

OLSSON, J., ARHEIMER, B., BORRIS, M., DONNELLY, C., FOSTER, K., NIKULIN, G., PERSSON, M., PERTTU, A.-M., UVO, C., VIKLANDER, M. et YANG, W. (2016). Hydrological climate change impact assessment at small and large scales : Key messages from recent progress in sweden. *Climate*, 4(3):39.

O’NEILL, B. C., TEBALDI, C., van VUUREN, D., EYRING, V., FRIEDLINGSTEIN, P., HURTT, G., KNUTTI, R., KRIEGLER, E., LAMARQUE, J.-F., LOWE, J., MEEHL, J., MOSS, R., RIAHI, K. et SANDERSON, B. M. (2016).

The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development Discussions*, 2016:1–35.

PIRTLE, Z., MEYER, R. et HAMILTON, A. (2010). What Does It Mean When Climate Models Agree? A Case for Assessing Independence Among General Circulation models. *Environmental Science & Policy*, 13(5):351–361.

RONDEAU-GENESSE, G. et BRAUN, M. (2020). Production des scénarios climatiques pour les projets : Impact des changements climatiques sur les débits au Québec (cQ2) et la thématique Évolution du climat du projet de Soutien à INFO-Crue. Rapport présenté par Ouranos. Montréal. 14p. + annexes. Rapport technique.

RONDEAU-GENESSE, G., BRAUN, M. et CARON, L. (2021). Pondération des scénarios climatiques d’INFO-Crue sur la base de l’unicité et de la performance des simulations hydrologiques. Rapport technique présenté à la Direction de l’expertise hydrique du ministère de l’Environnement et de la Lutte contre les changements climatiques. Ouranos, Montréal, Canada. 51 p. + annexes. Rapport technique.

ROY, P., RONDEAU-GENESSE, G., JALBERT, J. et FOURNIER, É. (2023). Climate scenarios of extreme precipitation using a combination of parametric and non-parametric bias correction methods in the province of québec. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, nil(nil):1–17.

SANDERSON, B. M., KNUTTI, R. et CALDWELL, P. (2015). A Representative Democracy To Reduce Interdependency in a Multimodel Ensemble. *Journal of Climate*, 28(13):5171–5194.

SANDERSON, B. M., WEHNER, M. et KNUTTI, R. (2017). Skill and independence weighting for multi-model assessments. *Geoscientific Model Development*, 10(6):2379–2395.

SENFLEBEN, D., LAUER, A. et KARPECHKO, A. (2020). Constraining uncertainties in cmip5 projections of september arctic sea ice extent with observations. *Journal of Climate*, 33(4):1487 – 1503.

SHERWOOD, S. C., WEBB, M. J., ANNAN, J. D., ARMOUR, K. C., FORSTER, P. M., HARGREAVES, J. C., HEGERL, G., KLEIN, S. A., MARVEL, K. D., ROHLING, E. J., WATANABE, M., ANDREWS, T., BRACONNOT, P., BRETHERTON, C. S., FOSTER, G. L., HAUSFATHER, Z., von der HEYDT, A. S., KNUTTI, R., MAURITSEN, T., NORRIS, J. R., PROISTOESCU, C.,

- RUGENSTEIN, M., SCHMIDT, G. A., TOKARSKA, K. B. et ZELINKA, M. D. (2020). An assessment of earth’s climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, 58(4):e2019RG000678. e2019RG000678 2019RG000678.
- SKAMAROCK, W. C. (2004). Evaluating mesoscale nwp models using kinetic energy spectra. *Monthly Weather Review*, 132(12):3019 – 3032.
- TAYLOR, K. E., STOUFFER, R. J. et MEEHL, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, 93(4):485–498.
- TEBALDI, C. et KNUTTI, R. (2007). The Use of the Multi-Model Ensemble in Probabilistic Climate projections. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 365(1857): 2053–2075.
- TEBALDI, C., SMITH, R. L., NYCHKA, D. et MEARN, L. O. (2005). Quantifying uncertainty in projections of regional climate change : A bayesian approach to the analysis of multimodel ensembles. *Journal of Climate*, 18(10):1524–1540.
- TOKARSKA, K. B., STOLPE, M. B., SIPPEL, S., FISCHER, E. M., SMITH, C. J., LEHNER, F. et KNUTTI, R. (2020). Past warming trend constrains future warming in cmip6 models. *Science Advances*, 6(12).
- van der LINDEN, P. et MITCHELL, J. (2009). Report. ENSEMBLES : Climate Change and its Impacts : Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK. 160pp (2009).
- von TRENTINI, F., AALBERS, E. E., FISCHER, E. M. et LUDWIG, R. (2020). Comparing interannual variability in three regional single-model initial-condition large ensembles (smiles) over europe. *Earth System Dynamics*, 11(4):1013–1031.
- WHETTON, P., MACADAM, I., BATHOLS, J. et O’GRADY, J. (2007). Assessment of the Use of Current Climate Patterns To Evaluate Regional Enhanced Greenhouse Response Patterns of Climate Models. *Geophys. Res. Lett.*, 34.
- ZHAO, J., HE, S., WANG, H. et LI, F. (2022). Constraining CMIP6 Projections of an Ice-Free Arctic Using a Weighting Scheme. *Earth’s Future*, 10(10):e2022EF002708. e2022EF002708 2022EF002708.

Annexe A

Estimation de la variance d'une différence entre deux modèles

Soit X_i et Y_j les moyennes climatologiques sur 30 ans associées aux membres i et j des modèles X et Y , pour des ensembles de tailles I et J respectivement. La différence entre la paire de membres (i, j) se définit comme suit :

$$D_{ij} = X_i - Y_j \quad (9)$$

La variance de cette différence peut s'écrire comme :

$$Var(D) = \frac{1}{IJ} \sum_{ij}^{IJ} (D_{ij} - \bar{D}_{oo})^2 \quad (10)$$

où \bar{D}_{oo} correspond à la différence moyenne calculée sur toutes les paires de membres des deux ensembles :

$$\bar{D}_{oo} = \frac{1}{IJ} \sum_{ij}^{IJ} D_{ij} = \bar{X}_o - \bar{Y}_o \quad (11)$$

Après remplacement des termes, l'équation (10) devient :

$$Var(D) = \frac{1}{IJ} \sum_{ij}^{IJ} (X_i - Y_j - \bar{X}_o + \bar{Y}_o)^2 \quad (12)$$

$$= \frac{1}{IJ} \sum_{ij}^{IJ} \left((X_i - \bar{X}_o) - (Y_j - \bar{Y}_o) \right)^2 \quad (13)$$

qui devient après distribution des termes :

$$\begin{aligned} \text{Var}(D) = \frac{1}{I} \sum_i^I (X_i - \bar{X}_o)^2 + \frac{1}{J} \sum_j^J (Y_j - \bar{Y}_o)^2 \\ - \frac{2}{IJ} \sum_i^I (X_i - \bar{X}_o) \sum_j^J (Y_j - \bar{Y}_o) \end{aligned} \quad (14)$$

où le dernier terme à droite est nul par construction. Nous obtenons donc la décomposition de variance suivante :

$$\text{Var}(D) = \text{Var}(X) + \text{Var}(Y) \quad (15)$$

Supposons maintenant que Y représente des observations (O). Nous ne pouvons pas calculer $\text{Var}(O)$ directement car O n'a qu'un seul membre. Sous l'hypothèse que $\text{Var}(O) \approx \text{Var}(X)$, on obtient :

$$\text{Var}(D) \approx 2 \times \text{Var}(X) \quad (16)$$

Et donc, l'écart-type de la différence est :

$$\text{Std}(D) \approx \sqrt{2} \times \text{Std}(X) \quad (17)$$

Annexe B

Figures supplémentaires

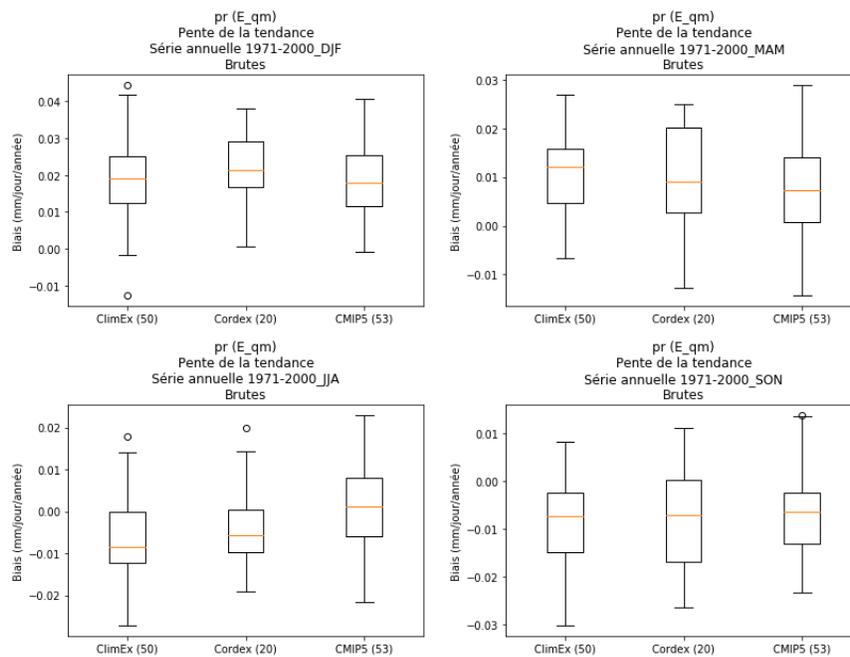


FIGURE 28 – Biais par rapport aux observations dans la tendance de changement climatique de la précipitation sur la période 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes.

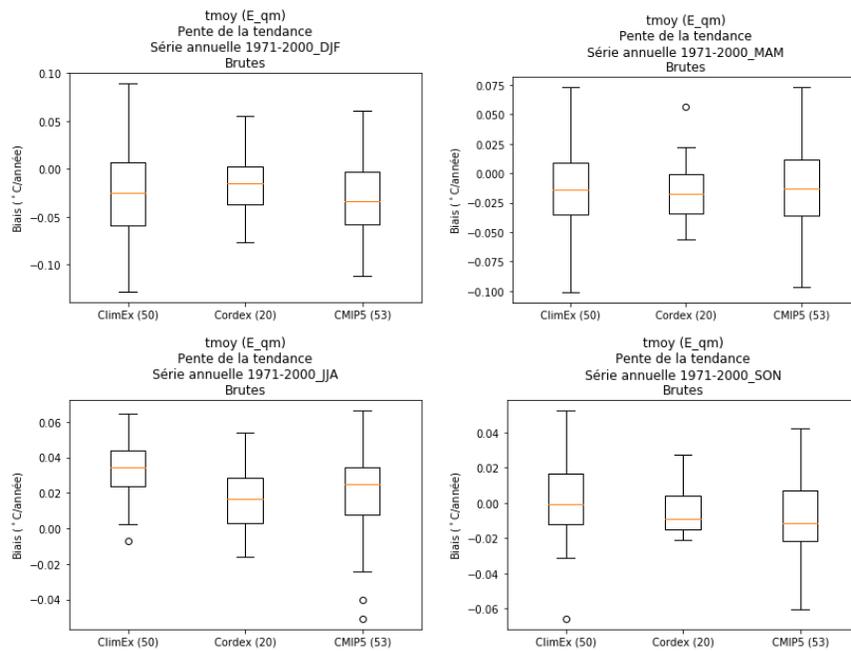


FIGURE 29 – Biais par rapport aux observations dans la tendance de changement climatique de la température moyenne sur la période 1971-2000 pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), données brutes.

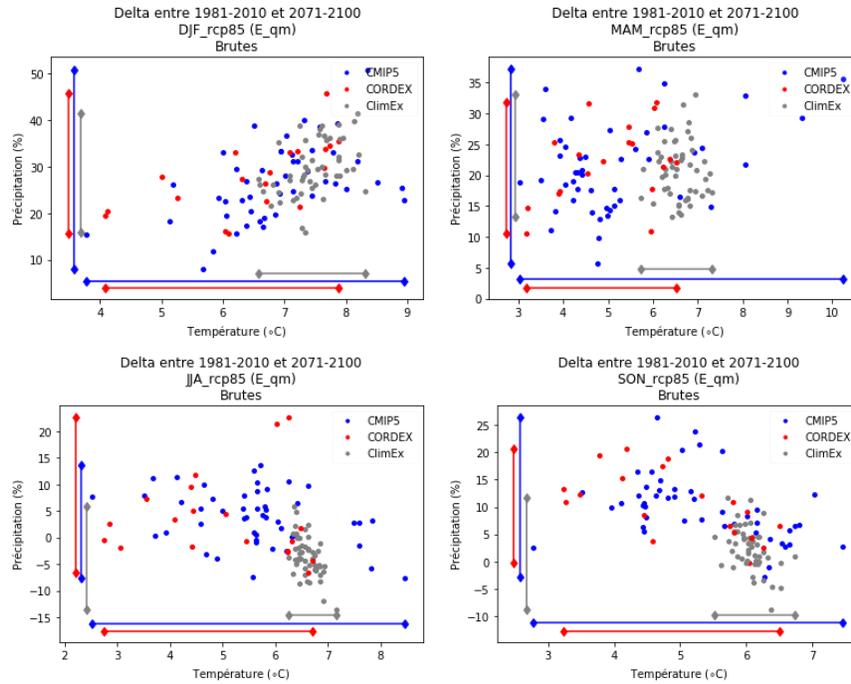


FIGURE 30 – Delta saisonnier 2071-2100 différents ensembles pour les saisons hiver (DJF), printemps (MAM), été (JJA) et automne (SON), selon RCP 8.5, données brutes.

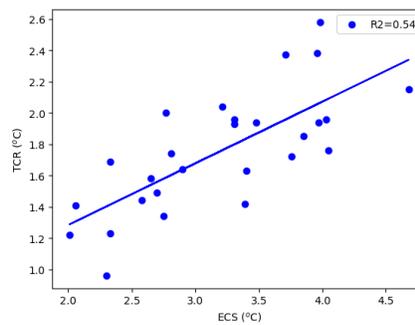


FIGURE 31 – Relation entre ECS et TCR pour 27 modèles globaux de CMIP5. Les données proviennent de Flynn et Mauritsen (2020).

Annexe C

Liste des modèles utilisant le scénario RCP8.5

TABLE 2 – Liste des modèles et du nombre de membres disponibles pour le scénario RCP8.5.

	Nom du modèle	N _{mem}
1	CMIP5_ACCESS1-0	1
2	CMIP5_ACCESS1-3	1
3	CMIP5_BCC-CSM1-1-m	1
4	CMIP5_BCC-CSM1-1	1
5	CMIP5_BNU-ESM	1
6	CMIP5_CMCC-CESM	1
7	CMIP5_CMCC-CMS	1
8	CMIP5_CMCC-CM	1
9	CMIP5_CNRM-CM5	1
10	CMIP5_CSIRO-Mk3-6-0	10
11	CMIP5_CanESM2	5
12	CMIP5_FGOALS-g2	1
13	CMIP5_GFDL-CM3	1
14	CMIP5_GFDL-ESM2G	1
15	CMIP5_GFDL-ESM2M	1
16	CMIP5_INMCM4	1
17	CMIP5_IPSL-CM5A-LR	4
18	CMIP5_IPSL-CM5A-MR	1
19	CMIP5_IPSL-CM5B-LR	1
20	CMIP5_MIROC-ESM-CHEM	1
21	CMIP5_MIROC-ESM	1
22	CMIP5_MIROC5	3
23	CMIP5_MPI-ESM-LR	3
24	CMIP5_MPI-ESM-MR	1
25	CMIP5_MRI-CGCM3	1
26	CMIP5_MRI-ESM1	1
27	CMIP5_NorESM1-M	1
28	CORDEX_CanESM2_CRCM5-UQAM	1
29	CORDEX_CanESM2_CanRCM4	2
30	CORDEX_CanESM2_RCA4	1
31	CORDEX_EC-EARTH_HIRHAM5	1
32	CORDEX_EC-EARTH_RCA4	1
33	CORDEX_GFDL-ESM2M_RegCM4	2
34	CORDEX_MPI-ESM-LR_RegCM4	2
35	CORDEX_MPI-ESM-MR_CRCM5-UQAM	1
36	ClimEx_CanESM2_CRCM5-Ouranos	50
37	Ouranos_CNRM-CM5_CRCM5-Ouranos	1
38	Ouranos_CanESM2_CRCM5-Ouranos	5
39	Ouranos_GFDL-ESM2M_CRCM5-Ouranos	1
40	Ouranos_MPI-ESM-LR_CRCM5-Ouranos	1

Annexe D

Relation entre l'ECS et les projections régionales sur le Québec à partir de l'ensemble CMIP6

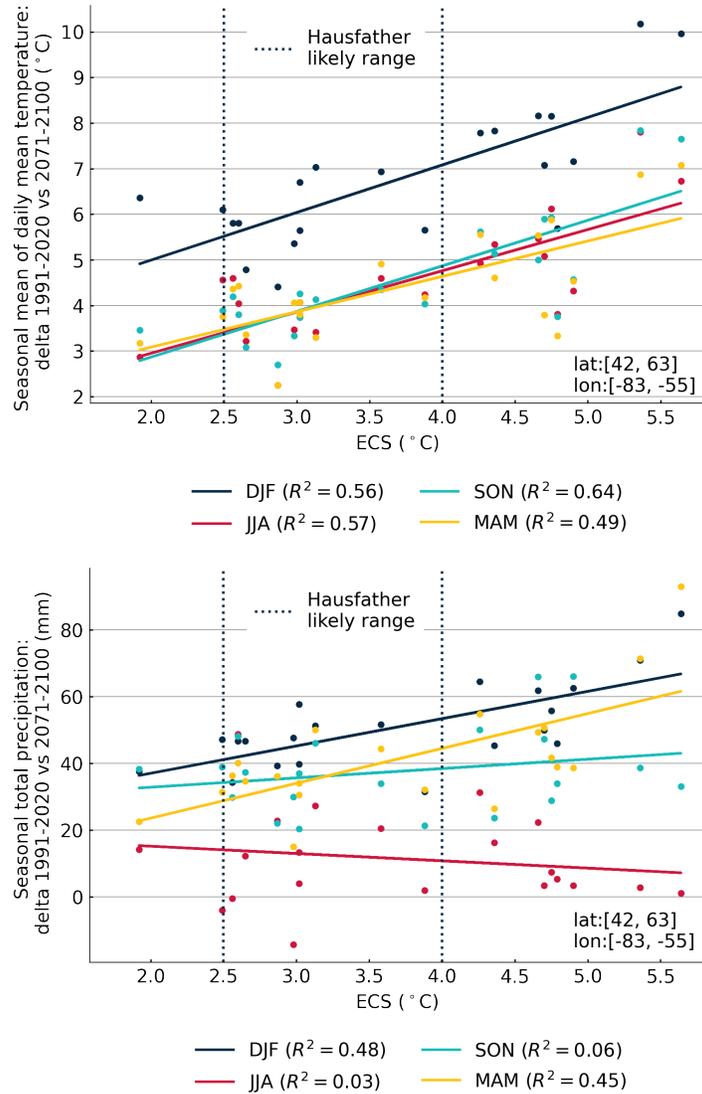


FIGURE 32 – Delta de température (haut) et précipitation (bas) sur la province de Québec entre 1991-2020 et 2071-2100 pour les quatre saisons en fonction des valeurs d’ECS, en utilisant 21 modèles de CMIP6 pour le scénario SSP370 (ensemble ESPO-G6-E5L d’Ouranos, Lavoie *et al.* (2023)). Les valeurs d’ECS proviennent de Hausfather *et al.* (2022). Crédits : Juliette Lavoie, Ouranos.



550 SHERBROOKE OUEST, TOUR OUEST, 19^e ÉTAGE, MONTRÉAL, QC, CANADA, H3A 1B9 | TÉLÉPHONE 514 282.6464 | TÉLÉCOPIEUR 514 282.7131 | WWW.OURANOS.CA