Évaluation de l'influence de la variabilité naturelle du climat et des changements climatiques anthropiques sur les extrêmes hydrométéorologiques

par

Jean-Luc MARTEL

THÈSE PAR ARTICLES PRÉSENTÉE À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE COMME EXIGENCE PARTIELLE À L'OBTENTION DU DOCTORAT EN GÉNIE
Ph. D.

MONTRÉAL, LE 8 AVRIL 2019

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

PRÉSENTATION DU JURY

CETTE THÈSE A ÉTÉ ÉVALUÉE

PAR UN JURY COMPOSÉ DE :

M. François Brissette, directeur de thèse
Département de génie de la construction à l'École de technologie supérieure

M. Alain Mailhot, codirecteur de thèse
Centre Eau Terre Environnement de l'Institut national de la recherche scientifique

M. Louis Lamarche, président du jury
Département de génie mécanique à l'École de technologie supérieure

Mme Annie Poulin, membre du jury
Département de génie de la construction à l'École de technologie supérieure

M. Jean-Philippe Vidal, examinateur externe
Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture (IRSTEA, France)

ELLE A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 8 FÉVRIER 2019

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je veux commencer en remerciant mon directeur de thèse François Brissette et mon codirecteur Alain Mailhot qui ont rendu l'ensemble de ce travail possible. François, j'ai adoré apprendre et collaborer avec toi durant mes deux stages de recherche, ma maîtrise et maintenant mon doctorat. Je me souviens encore de notre entretien dans ton bureau lorsque je t'ai demandé de me convaincre de faire une maîtrise. On peut dire que tu as vraiment réussi à me convaincre! Merci pour ta confiance et pour toute l'autonomie que tu m'as donnée durant toutes ces années. J'ai été choyé d'avoir un directeur aussi agréable et compréhensif que toi. Alain, je te remercie de m'avoir fait sortir de ma zone de confort afin de m'adapter à ton incroyable rigueur. Sache que, même si je semblais être découragé par moment, je suis reconnaissant de t'avoir eu comme codirecteur. Grâce à toi, je crois avoir acquis une maturité professionnelle qui fera de moi un bien meilleur enseignant et ingénieur.

Merci à tous mes collègues du DRAME (j'ai hésité à écrire les filles du DRAME) pour les nombreuses discussions non productives et les multiples pauses café qui m'ont permis de décrocher un peu! Je tiens à remercier Daniel Caya et de Magali Troin pour leurs multiples contributions scientifiques et philosophiques dans l'élaboration de ma thèse. Ç'a été un plaisir de pouvoir travailler avec vous deux! Aussi, un merci particulier à Richard Arsenault, que je considère comme un mentor, une source inépuisable de connaissances, mais surtout, un ami.

Un grand merci à mes parents, Léon et Nicole, qui m'ont constamment soutenu au mieux de leur capacité et encouragé à poursuivre mes études universitaires. Je vous suis infiniment reconnaissant de votre support inconditionnel.

Finalement, un énorme merci à ma conjointe Lucie, qui m'a accompagné tout au long de ce parcours et qui m'a offert un support moral inestimable. Merci d'avoir cru en moi et d'avoir été là durant les moments les plus difficiles! Je te suis éternellement reconnaissant d'avoir été une partenaire aussi fantastique!

Évaluation de l'influence de la variabilité naturelle du climat et des changements climatiques anthropiques sur les extrêmes hydrométéorologiques

Jean-Luc MARTEL

RÉSUMÉ

La capacité maximale des infrastructures d'eaux est conçue à partir des séries historiques d'extrêmes hydrométéorologiques et sur l'hypothèse que le climat est stationnaire. Toutefois, le consensus scientifique pointe vers un réchauffement planétaire causé par l'influence humaine, ayant des répercussions à long terme sur les précipitations et les crues extrêmes. De plus, plusieurs travaux récents indiquent que la variabilité naturelle du climat a le potentiel de masquer les effets de ces changements climatiques anthropiques, donnant une illusion de stationnarité du climat. Ainsi, l'objectif de cette thèse a pour but d'améliorer la compréhension des impacts de cette variabilité naturelle et des changements climatiques sur les extrêmes hydrométéorologiques.

Dans un premier lieu, la variabilité naturelle a été explorée à travers six indices climatiques importants (p. ex. : El Niño) influençant le climat en Amérique du Nord. Bien que ceux-ci aient démontré un certain potentiel de prédiction au niveau de la variabilité saisonnière des moyennes de températures et de précipitations, ils se sont trouvés beaucoup moins prometteurs au niveau de la variabilité des extrêmes hydrométéorologiques. La combinaison de ces indices climatiques n'a résulté qu'en une faible valeur prédictive de la variabilité des crues et précipitations extrêmes.

Dans un second lieu, la variabilité naturelle des précipitations moyennes et extrêmes a été examinée à l'aide de grands ensembles de simulations climatiques. Ces travaux ont permis de dresser un portait plus clair de son influence sur la détection du signal des changements climatiques. À l'échelle locale (p. ex. : une station météorologique), la variabilité naturelle dominera vraisemblablement le signal des changements climatiques des précipitations extrêmes jusqu'à la fin du 21^{e} siècle. Toutefois, à l'échelle régionale (p. ex. : plusieurs stations météorologiques), la détection du signal des changements climatiques serait plus rapide et robuste. Globalement, la variabilité naturelle a la capacité d'entraver la détection des changements climatiques sur les précipitations moyennes et extrêmes jusqu'à la moitié, voire même la fin du siècle, pour plusieurs régions de la planète.

Les grands ensembles de simulations climatiques ont aussi été utilisés pour évaluer l'impact des changements climatiques sur la probabilité de récurrence des événements hydrométéorologiques extrêmes. D'abord, les changements projetés pour la pluie 100 ans d'une durée d'une heure jusqu'à cinq jours entre les périodes de 1980-1999 et 2080-2099 ont été étudiés à l'aide de deux ensembles à l'échelle globale et un ensemble à l'échelle régionale couvrant le nord-est de l'Amérique du Nord et l'Europe. Les résultats des trois ensembles suggèrent que les événements de précipitation extrêmes, correspondant à la période de retour 100 ans de la période de référence, deviendront environ de quatre à cinq (deux à quatre) fois plus fréquents en moyenne pour le nord-est de l'Amérique du Nord (l'Europe). De plus, les

résultats suggèrent qu'en général, une période de retour plus élevée et/ou une durée plus courte entraineront des augmentations relatives plus importantes.

Ensuite, les changements projetés dans les crues extrêmes ont été investigués plus spécifiquement pour 3 567 bassins versants de grande taille (> 500 km^2) en Amérique du Nord. Les résultats suggèrent des patrons spatiaux très distincts en termes d'augmentation et de diminution de la crue 100 ans. Les changements les plus importants se résument en une diminution des crues générées par la fonte de la neige dans les bassins versants situés en haute latitude et/ou haute altitude et une augmentation pour les bassins versants situés dans le sud-est des États-Unis et sur la côte ouest, où les précipitations sont la principale cause des crues.

Finalement, des pistes de stratégies d'adaptation face aux changements climatiques ont été discutées à la suite de ces travaux. Ces dernières, de concert avec les conclusions de cette thèse, pourraient aider davantage les ingénieurs et les preneurs de décision à justifier l'implémentation de mesures d'adaptation permettant de mieux protéger le milieu bâti et les populations vulnérables.

Mots-clés : variabilité naturelle, changements climatiques, modélisation hydrologique, précipitation extrême, crue extrême.

Evaluation of the Influence of Natural Climate Variability and Anthropogenic Climate Change on Hydrometeorological Extremes

Jean-Luc MARTEL

ABSTRACT

The maximum capacity of water infrastructures is designed based on historical series of hydrological extremes and the hypothesis that the climate is stationary. However, the scientific consensus points towards a man-made global warming, with gradual repercussions on extreme precipitation and streamflow. In addition, several recent works indicate that the natural climate variability has the potential to mask the effects of these anthropogenic climate changes, giving the illusion of climate stationarity. Therefore, the aim of this thesis is to improve the understanding of the impacts of this natural variability and climate change on hydrometeorological extremes.

In the first place, natural variability was explored through six major climate indices (e.g., El Niño) influencing climate in North America. Although these showed some potential in predicting the seasonal variability of mean temperatures and precipitation, they were much less promising in the prediction of hydrological extreme events variability. The combination of these six climate indices only resulted in a low predictive value of extreme precipitation and streamflow variability.

In a second place, the natural variability of mean and extreme precipitation was examined using large ensembles of climate simulations. This work led to a clearer picture of its influence on the detection of the climate change signal. At the local scale (e.g., a single weather station), natural variability will likely dominate the climate change signal from extreme precipitation until the end of the 21st century. However, at the regional scale (e.g., multiple weather stations), the detection of the climate change signal would be faster and more robust. Globally, natural variability has the potential to impede the detection of climate change on mean and extreme precipitation up to the middle or even the end of the century for many regions of the world.

Large ensembles of climate simulations have also been used to assess the impact of climate change on the probability of recurrence of hydrological extreme events. To begin, the projected changes for the 100-year rainfall with durations ranging from one hour to five days between the 1980-1999 and 2080-2099 periods were studied using two ensembles at the global scale and one at the regional scale covering northeastern North America and Europe. The results from these three ensembles suggest that extreme precipitation events, corresponding to the 100-year return period of the reference period, will become about four to five (two to four) times more frequent in average for the northeastern North America (Europe). In addition, the results suggest that, in general, a larger return period and/or a shorter duration will result in more important relative increases.

Next, projected changes in extreme floods were investigated more specifically for 3 567 large (> 500 km^2) catchments in North America. The results suggest very distinct spatial patterns in terms of increases and decreases in the 100-year flood. The most significant changes were found to be reductions in flood generated from snowmelt in high latitude and/or high-altitude catchments and increases for catchments in the southeastern United States and the West Coast, where rainfall is the leading cause of flooding.

Finally, climate change adaptation strategies were discussed following this work. These, together with the conclusions from this thesis, could further help engineers and decision makers to justify the implementation of adaptation measures to better protect the built environment and vulnerable populations.

Keywords : natural variability, climate change, hydrological modelling, extreme precipitation, extreme streamflow

TABLE DES MATIÈRES

Page

LISTE DES TABLEAUX

Page

LISTE DES FIGURES

Page

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

90%DD	90% Detection Decade
ACC	Association canadienne de la construction
ACC	Anthropogenic climate change
ACTP	Association canadienne des travaux publics
AGCM	Atmospheric Global Circulation Model
AIC	Akaike Information Criterion
AMO	Atlantic Multi-Decadal Oscillation
AMP	Annual maximum precipitation
AMPS	Annual maximum precipitation series
AMS	Annual maxima series
AO	Arctic Oscillation
AOGCM	Atmospheric-Oceanic Global Circulation Model
BIC	Bayesian Information Criterion
BRIC	Bulletin de rendement des infrastructures canadiennes
CanESM2	Canadian Earth System Model version 2
CANOPEX	Canadian Model Parameter Experiment
CCCma	Canadian Centre for Climate Modelling and Analysis
CCSM3	Community Climate System Model version 3
CDF	Cumulative density function
CESM1	Community Earth System Model version 1
CLIMDEX	Climate Extreme Indices
ClimEx	Climate change and hydrological Extremes

CMIP	Coupled Model Intercomparison Project
CMP	Crue maximale probable
CNRC	Conseil national de recherches Canada
CORDEX	Coordinated Regional Climate Downscaling Experiment
CPM	Convection Permitting Model
CRCM5	Canadian Regional Climate Model version 5
CSA	Canadian Standards Association
DBC	Daily bias correction
DJF	December-January-February
DMIP	Distributed Model Intercomparison Project
DT	Daily translation
ECCC	Environnement et Changement climatique Canada
ENSO	El Niño-Southern Oscillation
ESM	Earth System Model
ESRL	Earth System Research Laboratory
ET	Evapotranspiration
ETCCDI	Expert Team on Climate Change Detection and Indices
EU	Europe
FCM	Fédération canadienne des municipalités
GCM	Global Circulation Model or Global Climate Model
GES	Gaz à effet de serre
GEV	Generalized Extreme Value distribution
GIEC	Groupe d'experts intergouvernemental sur l'évolution du climat

GR4J	Modèle du Génie rural à 4 paramètres journalier
ICV	Internal climate variability
IDF	Intensity-Duration-Frequency curves
IPCC	Intergovernmental Panel on Climate Change
JJA	June-July-August
KGE	Kling-Gupta Efficiency
LE	Large ensemble
LENS	Large Ensemble Community Project
LGP	Land grid points
LOCI	Local intensity scaling
MCG	Modèle de circulation générale ou Modèle climatique global
NA	North America
NAO	North Atlantic Oscillation
NARCCAP	North American Regional Climate Change Assessment Program
NCAR	National Center of Atmospheric Research
NCV	Natural Climate Variability
NRCAN	National Resource Canada
NSE	Nash-Sutcliffe Efficiency
NNA	Northeastern North America
NOAA	National Oceanic and Atmospheric Administration
OIQ	Ordre des ingénieurs du Québec
PDO	Pacific Decadal Oscillation
PDF	Probability Density Function

XXX

PMP	Pluie maximale probable
PNA	Pacific North American pattern
PRCPTOT	Total wet-day precipitation
OGCM	Oceanic General Circulation Model
QM	Mean daily streamflow
QX1day	Maximum 1-day streamflow
RCM	Regional Climate Model
RCP	Representative Concentration Pathway
RX1day	Max 1-day precipitation amount
RX5day	Max cumulative 5-day precipitation amount
S/N	Signal-to-noise ratio
SCE-UA	Shuffled Complex Evolution – University of Arizona
SCGC	Société canadienne de génie civil
SLP	Sea-level pressure
SMA	Séries des maxima annuels
SST	Sea surface temperature
SWE	Snow-water equivalent
TAS	Mean daily temperature
TOE	Time of Emergence
UCAR	University Corporation for Atmospheric Research
UQAM	Université du Québec à Montréal
USGS	United States Geological Survey
WMO	World Meteorological Organization

INTRODUCTION

0.1 Problématique de la thèse

Les événements hydrométéorologiques extrêmes sont d'une grande importance lors de la conception des divers ouvrages hydrauliques afin d'assurer la sécurité du public et de minimiser les dommages associés à de tels aléas. Par exemple, les événements de précipitations extrêmes peuvent servir de critère de conception pour différentes infrastructures urbaines (p. ex. : les systèmes de collecte des eaux usées), tandis que les débits extrêmes de rivière peuvent être utilisés pour le dimensionnement d'importantes structures hydrauliques (p. ex. : un pont ou un barrage) ou encore pour la délinéation des plaines inondables. En général, ces critères de conception sont basés sur une estimation de la probabilité de récurrence de ces événements (p. ex. : un orage 20 ans ou une inondation 100 ans) via l'utilisation d'une analyse fréquentielle. Puisque les séries chronologiques d'observations sont relativement courtes, une loi de probabilité basée sur la théorie des valeurs extrêmes est habituellement ajustée aux observations historiques afin de permettre l'extrapolation vers un événement plus rare (Meylan, Favre, & Musy, 2008).

En ce qui concerne les ouvrages hydrauliques de grande importance, tels qu'un barrage ou un pont, les critères utilisés sont encore plus sévères en raison de la gravité des conséquences liées à leur rupture potentielle. Par exemple, la Loi sur la sécurité des barrages du Québec adoptée en 2002 impose des normes minimales pour la récurrence de l'événement extrême à utiliser puisqu'un niveau de conséquences plus important lié à une rupture potentielle requiert l'utilisation d'un événement plus rare. Dans les cas où le niveau de conséquence est extrêmement important, la PMP (pluie maximale probable) ou la CMP (crue maximale probable), correspondant au pire cas physiquement possible, peuvent être utilisées, voire même imposées.

Pour les ouvrages plus communs, un certain compromis est réalisé entre les coûts de construction de l'ouvrage en question et les coûts qui seraient associés aux dommages

engendrés par un aléa hydrométéorologique qui dépasserait le critère de conception employé. En utilisant comme exemple un système de collecte des eaux pluviales, ce dernier serait dimensionné pour être en mesure de capter la majorité des événements de pluies, mais les événements plus rares engendreraient quand même des débordements, et possiblement des bris matériels. Cependant, le raisonnement logique derrière cette conception est qu'il serait moins dispendieux de dédommager les citoyens et de réparer les bris anticipés par un tel événement lorsque celui-ci survient, que de surdimensionner la capacité des ouvrages hydrauliques, engendrant des coûts de construction plus importants. La Figure 0.1 présente une vue idéalisée de la zone de conception optimale qui est généralement visée par les ingénieurs en fonction de ce compromis, correspondant à la région la plus basse de la courbe des coûts totaux (courbe noire).

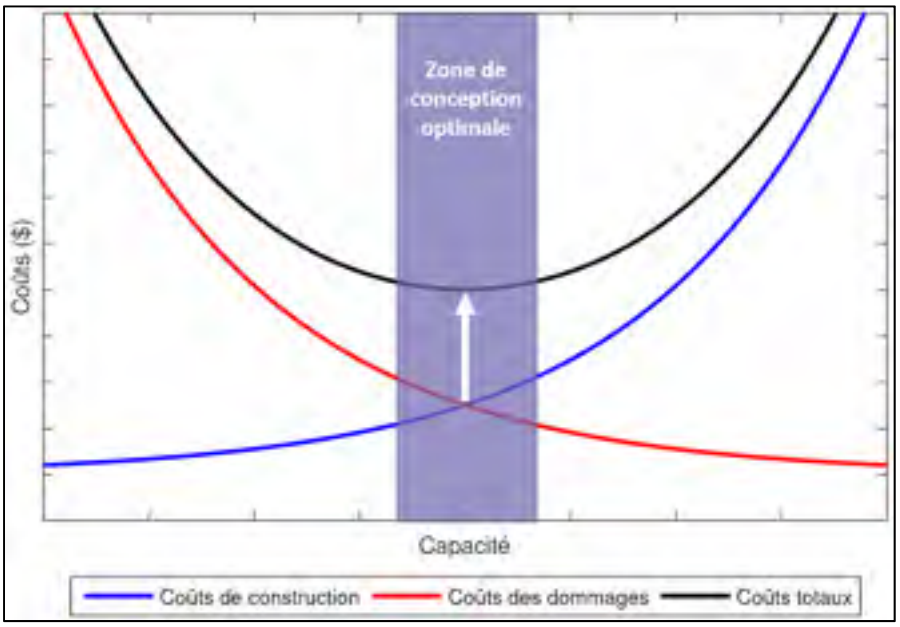

Figure 0.1 Idéalisation du compromis entre les coûts de construction (courbe bleue) et les coûts associés aux dommages (courbe rouge) représenté par les coûts totaux (courbe noire). La zone de conception optimale est représentée par la flèche blanche et le rectangle mauve

Les analyses fréquentielles utilisées pour estimer l'intensité de ces événements dépendent de plusieurs prémices, notamment celle de la stationnarité des séries temporelles d'observations utilisées (c.-à-d. que les propriétés statistiques, telles que la moyenne, la variance et les

extrêmes, ne varient pas en fonction du temps; Meylan et al., 2008). Cependant, le consensus des études scientifiques indique que le réchauffement planétaire est directement lié à l'influence humaine à travers les émissions de gaz à effet de serre (IPCC, 2013). De plus, il est attendu que cette augmentation globale de la température pourrait mener vers une augmentation de la fréquence et de l'intensité des événements de précipitations extrêmes, partiellement en raison d'une atmosphère plus chaude pouvant entreposer davantage d'humidité (IPCC, 2013; Lenderink & Fowler, 2017; Trenberth, 1999; Trenberth, Dai, Rasmussen, & Parsons, 2003). Par conséquent, dans l'éventualité où les événements hydrométéorologiques extrêmes deviennent plus ou moins fréquents, l'hypothèse de stationnarité traditionnellement utilisée dans les analyses fréquentielles doit être revisitée. La Figure 0.2 illustre comment la zone de conception optimale introduite dans la Figure 0.1 pourrait en fait sous-estimer les coûts associés aux dommages dans un climat futur où les extrêmes seraient plus fréquents, provoquant une augmentation des coûts totaux.

Il devient évident qu'une compréhension accrue de l'impact des changements climatiques sur la probabilité de récurrence des événements hydrométéorologiques extrêmes est primordiale. En effet, l'augmentation en fréquence et en intensité de ceux-ci a des implications directes au niveau de la sécurité du public et aussi sur les plans économiques et environnementaux.

Une considération importante doit aussi être attribuée à la nature chaotique du système climatique (c.-à-d. la variabilité naturelle du climat, aussi appelée variabilité interne et ci-après nommée variabilité naturelle). Il y a beaucoup d'indices qui tendent à démontrer que la variabilité naturelle puisse masquer localement et régionalement les effets des changements climatiques (c.-à-d. la tendance ou le signal), notamment au niveau des précipitations (p.ex. Deser, Knutti, Solomon, & Phillips, 2012a; Fischer & Knutti, 2014). Malgré que des tendances significatives soient actuellement détectables pour les données de température moyenne à l'échelle planétaire (IPCC, 2013), les changements au niveau des précipitations moyennes, et surtout extrêmes, sont beaucoup plus difficiles à détecter de façon significative à cause d'une plus grande variabilité naturelle (Westra, Alexander, & Zwiers, 2013).

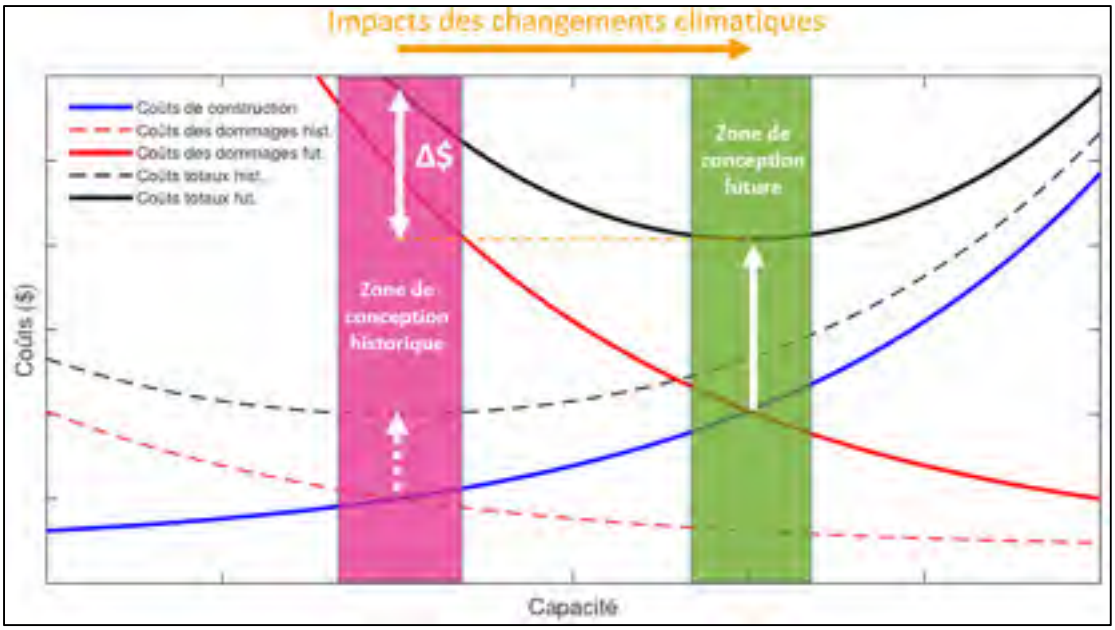

Figure 0.2 Impact des changements climatiques sur le compromis entre les coûts de construction (courbe bleue) et les coûts associés aux dommages (courbes rouges) représenté par les coûts totaux (courbes noires). Les courbes pointillées représentent la période historique et les courbes pleines la période future suite aux impacts des changements climatiques. La zone de conception historique (rectangle rose) présente des coûts plus élevés en climat futur, tandis que la zone de conception future (rectangle vert) présente la nouvelle zone de conception optimale

Ainsi, cette variabilité naturelle pourrait compromettre la capacité de détection du signal des changements climatiques pour certains endroits, donnant une illusion de stationnarité des séries d'observations temporelles (p. ex. : les séries d'une ou plusieurs stations météorologiques). Dans un tel cas, la décision de ne pas prendre en considération les impacts des changements climatiques pourrait être prise à tort. Il est donc essentiel de bien comprendre l'influence potentielle de la variabilité naturelle sur les séries chronologiques d'observations afin de mieux renseigner les décideurs et les ingénieurs sur l'importance d'adapter les infrastructures à un climat futur.

0.2 État de la situation au Canada

Le deuxième Bulletin de rendement des infrastructures canadiennes (BRIC, 2016) a été réalisé en partenariat avec quatre organismes : l'Association canadienne de la construction

(ACC), l'Association canadienne des travaux publics (ACTP), la Société canadienne de génie civil (SCGC) et la Fédération canadienne des municipalités (FCM). Cette deuxième édition dresse un portrait détaillé de l'état des infrastructures municipales dans l'ensemble du Canada en se basant sur les réponses obtenues à un sondage auprès de 120 municipalités. Deux messages clés d'intérêt pour cette thèse en sont ressortis :

1. « Le tiers de nos infrastructures municipales est en état passable, mauvais ou très mauvais, ce qui accroît le risque de perturbations de service (Figure 0.3) »
2. « Environ 19 % des municipalités répondantes ont indiqué qu'elles utilisent des mécanismes officiels (p. ex. : politiques municipales ou pratiques écrites) afin d'inclure des stratégies d'adaptation aux changements climatiques dans leur processus décisionnel. »

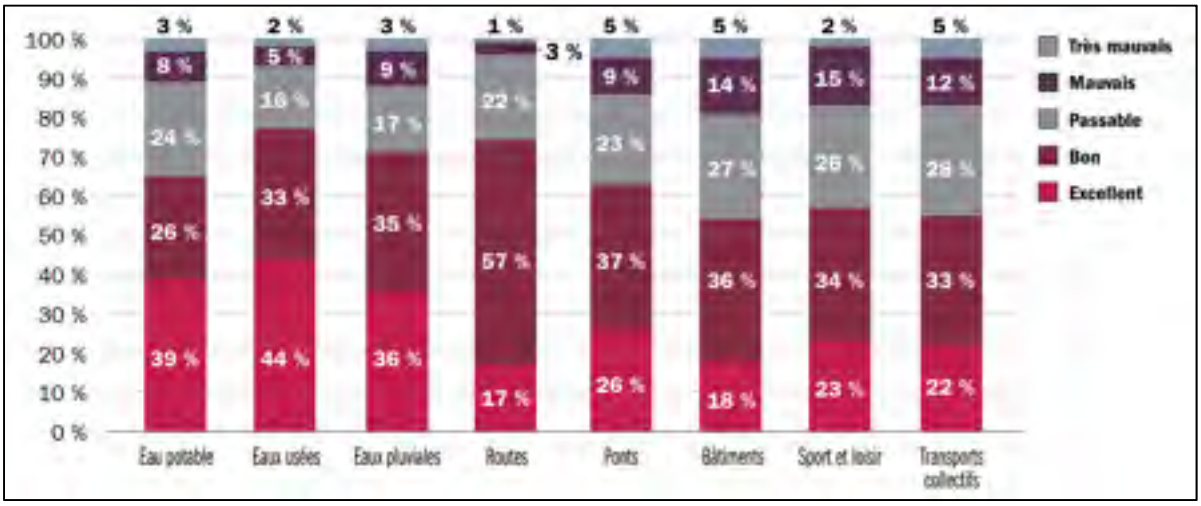

Figure 0.3 Résumé de l'état physique moyen des infrastructures municipales canadiennes. Tirée intégralement du BRIC (2016)

Un sondage d'Ingénieurs Canada a été réalisé par le Groupe CSA (*Canadian Standards Association*; CSA, 2012) à l'échelle nationale auprès des ingénieurs en infrastructures afin d'identifier leur niveau de connaissance et de sensibilisation aux impacts des changements climatiques. Parmi les conclusions basées sur les 3 362 réponses reçues, il a été noté que :

- Environ 3 ingénieurs sur 5 utilisent un outil ou une technique d'adaptation face aux changements climatiques dans leur pratique, tels que des facteurs de sécurité plus importants ;

- Environ 3 ingénieurs sur 4 estiment avoir besoin d'une plus grande quantité d'informations pour mieux aborder la question des changements climatiques dans leur pratique ;
- Les directives générales, les guides de meilleures pratiques et les ressources en ligne figurent parmi les méthodes de prédilection des ingénieurs pour obtenir de l'information à ce sujet ;
- Un des obstacles principaux qui empêchent les ingénieurs d'adresser les impacts des changements climatiques dans leur conception est le manque d'exigences dans les codes, les normes ou les politiques.

La situation actuelle des infrastructures municipales révèle donc un besoin urgent en termes de réhabilitation et reconstruction et il est évident que plusieurs de celles-ci devront être adaptées aux changements climatiques. Près du cinquième des municipalités prennent des mesures à ce sujet et devraient servir d'exemple pour les autres (BRIC, 2016). Les ingénieurs seront amenés à jouer un rôle de premier plan en informant leurs clients sur les stratégies d'adaptations aux changements climatiques. Cependant, tel que souligné par le sondage d'Ingénieurs Canada (CSA, 2012), il y a encore un manque à combler en termes de transfert de connaissance et d'outils décisionnels (p. ex. : par le biais d'exigences dans les codes, normes et politiques) leur permettant de bien répondre à ce rôle.

C'est dans cette optique que des directives générales ont été adoptées par Ingénieurs Canada en 2014 (*Engineers Canada*, 2014) afin de fournir des conseils aux organismes de réglementation du génie, tels que l'Ordre des ingénieurs du Québec (OIQ). Neuf principes y sont décrits pour encadrer la pratique professionnelle des ingénieurs au niveau de l'intégration des mesures d'adaptation face aux changements climatiques. Un constat de ce document qui mérite d'être souligné, est que les ingénieurs ont le devoir d'informer leurs clients et/ou employeurs sur les questions liées à l'adaptation face aux changements climatiques pouvant affecter les activités professionnelles dont ils sont responsables.

Il est aussi intéressant de mentionner que le Conseil national de recherches Canada (CNRC) et Infrastructure Canada ont annoncé qu'ils mettraient à jour les codes et les directives générales pour tenir compte des impacts des changements climatiques d'ici 2020.

0.3 Objectifs de la thèse

Dans ce contexte, cette thèse a pour but d'améliorer la compréhension des impacts de la variabilité naturelle et des changements climatiques sur les extrêmes hydrométéorologiques. L'approfondissement de ces connaissances permettra notamment d'établir des pistes de solutions en termes de stratégies d'adaptation face aux changements climatiques et d'en orienter des futurs travaux dans cette optique.

Pour y arriver, deux axes de recherche principaux seront investigués ayant pour but de :

1. déterminer l'influence de la variabilité naturelle sur la détection du signal des changements climatiques ;
2. évaluer l'influence des changements climatiques sur la récurrence des événements hydrométéorologiques extrêmes.

Cette thèse est composée d'un total de huit chapitres, excluant cette introduction, et de huit annexes. Le chapitre 1 expose d'abord une revue approfondie de la littérature scientifique pertinente. Le chapitre 2 introduit la démarche générale et l'organisation de la thèse par articles. Les chapitres 3 à 6 présentent chacun l'un des quatre articles scientifiques qui composent cette thèse en répondant aux deux axes de recherche principaux. Une discussion générale des conclusions principales obtenues est présentée au chapitre 7, suivie des conclusions au chapitre 8. En annexes figurent quatre autres articles scientifiques rédigés en collaboration (un publié, un accepté pour publication et deux soumis) qui ont découlé ou contribué à cette thèse, d'autres résultats contribuant aux objectifs de la thèse, ainsi qu'une synthèse des travaux et présentations (orales et par affiches) réalisés durant la réalisation de cette thèse.

CHAPITRE 1

REVUE DE LA LITTÉRATURE

Dans ce chapitre, une revue approfondie de la littérature y est exposée, couvrant la majorité des thématiques investiguées dans cette thèse. D'abord, une définition et un survol de travaux réalisés sur les impacts observés des changements climatiques d'origines anthropiques sur les extrêmes hydrométéorologiques sont établis. Une définition de la variabilité naturelle est présentée, ainsi que les différents indices utilisés pour la mesurer. Les différentes méthodes permettant de détecter le signal des changements climatiques à travers la variabilité naturelle des séries chronologiques sont ensuite abordées. Finalement, un recensement des études sur les impacts futurs des changements climatiques anthropiques sur les extrêmes hydrométéorologiques par le biais de la modélisation climatique et hydrologique est présenté.

1.1 Les changements climatiques d'origines anthropiques et leurs impacts

Il y a un consensus qui a été établi au travers des différentes études scientifiques supportant un réchauffement planétaire causé par l'action humaine via l'augmentation des émissions de gaz à effet de serre (GES). Un important recensement des travaux de recherche les plus récents est réalisé par les rapports du Groupe d'experts intergouvernemental sur l'évolution du climat (*Intergovernmental Panel on Climate Change* – IPCC – dont l'acronyme anglais sera utilisé dans cette thèse), permettant de dresser un portrait plus clair de la situation pour les preneurs de décisions (IPCC, 2013). Une des conclusions principales du cinquième rapport de l'IPCC (2013) en lien avec le réchauffement planétaire est que les événements de précipitations extrêmes vont très vraisemblablement devenir plus fréquents et plus intenses pour la majorité des régions situées dans les latitudes moyennes. En effet, il y a un grand nombre d'études indiquant qu'une atmosphère globalement plus chaude pourrait être en mesure d'emmagasiner une plus grande quantité d'humidité, entraînant par le fait même cette probable augmentation des précipitations extrêmes (Lenderink & Fowler, 2017; Trenberth, 1999; Trenberth et al., 2003). Par contre, la relation entre l'augmentation de la température et celle des extrêmes de précipitation, particulièrement à l'échelle sous-journalière, s'est

démontrée jusqu'à présent comme étant très complexe (Lenderink & Fowler, 2017; Lenderink, Mok, Lee, & van Oldenborgh, 2011; Westra et al., 2014).

Il est donc attendu que ces changements climatiques d'origine anthropique auront un impact important sur les événements extrêmes de précipitation, et conséquemment sur les extrêmes de débits en rivière. Cependant, tel que discuté précédemment, l'estimation des périodes de retour est encore traditionnellement faite sous l'hypothèse de la stationnarité du climat, ignorant donc toutes ces indications pointant notamment vers l'augmentation de la fréquence des extrêmes hydrométéorologiques (Katz, 2013; Mailhot & Duchesne, 2010; Milly et al., 2008b).

1.1.1 Augmentations observées sur les extrêmes de précipitations

Jusqu'à présent, un grand nombre d'études a démontré que près de deux tiers des zones terrestres couvertes par des données observées indiquent des tendances à la hausse dans les séries des maxima annuels de précipitation journalière (Alexander et al., 2006; Donat et al., 2013b; Min, Zhang, Zwiers, & Hegerl, 2011; Westra et al., 2013). Ces tendances sont également supportées par d'autres études scientifiques utilisant différentes bases de données et méthodes, dont les résultats indiquent aussi des augmentations importantes au niveau des extrêmes journaliers de précipitation (Easterling et al., 2000; Groisman et al., 2005; Madsen, Lawrence, Lang, Martinkova, & Kjeldsen, 2014; Seneviratne et al., 2012; Trenberth et al., 2007). Plus précisément, des augmentations régionalement répandues et partiellement significatives ont été observées à travers l'Amérique du Nord (Donat et al., 2013b; Easterling et al., 2000; Groisman et al., 2005).

Il y a aussi plusieurs indications que ces augmentations toucheront de façon encore plus importante les précipitations extrêmes sous-journalières, ayant notamment une grande importance au niveau des petits bassins versants et en milieux urbains (Madsen et al., 2014; Westra et al., 2014). Par exemple, un certain nombre d'études ont démontré que pour l'Amérique du Nord, une augmentation de la fréquence des événements extrêmes allant de

quelques heures à plusieurs jours est observée (Brommer, Cerveny, & Balling, 2007; Burn, Mansour, Zhang, & Whitfield, 2011; Kunkel et al., 2013; Muschinski & Katz, 2013).

1.1.2 Augmentations observées sur les extrêmes des débits en rivière

Au niveau des extrêmes des débits en rivière, la réponse face aux changements climatiques d'origine anthropique n'est pas aussi claire qu'elle l'est pour les extrêmes de précipitation. En effet, au niveau des crues, il n'y a pas de consensus à l'échelle globale de l'ampleur des changements observés en termes de fréquence et d'intensité, mais également au niveau du signe du changement (Kundzewicz et al., 2014; Kundzewicz et al., 2007; Trenberth et al., 2007). Ceci n'est pas une surprise, puisqu'il est attendu que le réchauffement planétaire aura une influence directe sur plusieurs composantes du cycle hydrologique (Bates, Kundzewicz, & Wu, 2008; Kundzewicz et al., 2014; Madsen et al., 2014). Mise à part la température, les plus importants changements seront ressentis sur les précipitations et ses diverses caractéristiques, notamment l'intensité, la durée et la phase (pluie ou neige). Il est également attendu que les changements dans les différentes variables climatiques auront des impacts directs sur les conditions menant à une crue extrême, comme l'humidité du sol précédent l'épisode important de précipitation, l'évapotranspiration et le couvert de neige.

À ce jour, les changements observés les plus évidents en termes de crues extrêmes concernent les bassins versants dominés par la neige. En effet, l'augmentation de température observée sur ces bassins a vraisemblablement déclenché une fonte de la neige plus hâtive et sur une plus longue période, générant ainsi des pointes de débits moins importantes (Groisman, Knight, & Karl, 2001; Li, Wrzesien, Durand, Adam, & Lettenmaier, 2017; Madsen et al., 2014; Seneviratne et al., 2012; Stewart, Cayan, & Dettinger, 2005; Zhang, Harvey, Hogg, & Yuzyk, 2001). Concernant les bassins versants pour lesquels les inondations sont principalement causées par d'importants événements de précipitation, la tendance observée se trouve à être vers des conditions plus humides avec des patrons mixtes d'augmentations et de diminutions des maxima annuels des débits en rivière et/ou des périodes de retour importantes (Burn & Whitfield, 2016; Groisman et al., 2001; Lins &

Slack, 1999; Madsen et al., 2014; McCabe & Wolock, 2002; Rice, Emanuel, Vose, & Nelson, 2015).

1.2 La variabilité naturelle du climat

Une attention particulière doit aussi être apportée à la nature chaotique du système climatique (c.-à-d. la variabilité naturelle). En effet, les résultats d'un grand nombre d'études indiquent que cette variabilité naturelle pourrait être en mesure de masquer le signal des changements climatiques anthropiques au niveau des précipitations à l'échelle locale et régionale (Deser et al., 2012a; Deser, Phillips, Bourdette, & Teng, 2012b; Deser, Phillips, Alexander, & Smoliak, 2014; Fischer, Beyerle, & Knutti, 2013; Fischer & Knutti, 2014; Fischer, Sedláček, Hawkins, & Knutti, 2014; Giorgi & Bi, 2009; Hawkins & Sutton, 2011, 2012; King et al., 2015; Maraun, 2013b; Mora et al., 2013; Sanderson, Oleson, Strand, Lehner, & O'Neill, 2018; Thompson, Barnes, Deser, Foust, & Phillips, 2015).

1.2.1 Définition de la variabilité naturelle

Le climat terrestre a toujours varié de façon significative sur plusieurs échelles temporelles, allant de seulement quelques années, à plusieurs décennies, jusqu'aux périodes glaciaires (Baede, Ahlonsou, Ding, & Schimel, 2001; Peixoto & Oort, 1992). Cette variabilité naturelle se manifeste en réponse aux différents changements dans les forçages externes et avec les interactions entre les composantes internes du système climatique (c.-à-d. : l'atmosphère, l'hydrosphère, la cryosphère, la lithosphère et la biosphère). Mis à part les changements d'origines anthropiques, les forçages externes sont essentiellement composés des variations astronomiques et terrestres (Baede et al., 2001; Peixoto & Oort, 1992).

Parmi les différents facteurs astronomiques, on retrouve principalement les cycles de Milankovitch et les variations des activités solaires (McGuffie & Henderson-Sellers, 2005; Peixoto & Oort, 1992). Les cycles de Milankovitch se résument par les changements dans l'excentricité de l'orbite terrestre, de l'obliquité terrestre et de la précession terrestre qui varient sur une échelle temporelle allant jusqu'à des dizaines de milliers d'années. L'orbite

terrestre oscille entre une orbite elliptique, puis circulaire, sur des cycles s'étendant à près de 110 000 années. L'obliquité ou l'inclinaison terrestre est l'angle qui se trouve entre l'axe de rotation et l'axe perpendiculaire sur le plan de son orbite. Celle-ci varie entre 22° et 24,5° suivant des cycles d'une durée d'environ 41 000 années. Finalement, à cause de l'attraction gravitationnelle des autres planètes, la précession terrestre apporte des variations dans la synchronisation des équinoxes. Les deux cycles principaux reliés à la précession terrestre sont d'environ 23 000 années. Quant à elles, les variations des activités solaires suivent des cycles variant de 80 à 100 années se manifestant entre autres par les taches solaires.

Au niveau des variations terrestres, on retrouve les changements de l'utilisation des terres (p. ex. : la déforestation et la désertification), les variations tectoniques causant le déplacement des continents ou même la création des montagnes et les variations des gaz composant l'atmosphère via les éruptions volcaniques et notamment l'activité humaine (Baede et al., 2001; Peixoto & Oort, 1992).

Cependant, même si tous les forçages externes du climat étaient constants (c.-à-d. : sans variations), la variabilité naturelle serait quand même présente. En effet, les différentes composantes du système climatique interagissent et tentent constamment de trouver l'équilibre, mais sans jamais y arriver à cause de leur temps de réaction très différent (Baede et al., 2001; Peixoto & Oort, 1992). Une rétroaction se produit suite à une interaction entre deux ou plusieurs composantes (p. ex. : l'atmosphère et l'océan) intégrant une portion du résultat dans l'interaction suivante (McGuffie & Henderson-Sellers, 2005; Peixoto & Oort, 1992). Ainsi, des rétroactions positives ou négatives sont résultantes de ces interactions complexes non linéaires. Par exemple, suite à une augmentation de la température à l'échelle planétaire, il y aura une diminution du couvert neigeux et de glace dans la cryosphère. La conséquence (c.-à-d. : la rétroaction) de cette diminution sera un albédo global plus faible, engendrant une moindre réflexion de la radiation solaire vers l'espace, augmentant davantage la température à l'échelle planétaire. Cependant, il y aura d'autres interactions avec les différentes composantes climatiques qui empêcheront que les rétroactions soient entièrement négatives ou positives. Il est très difficile d'établir des relations de proportionnalité entre les

différentes rétroactions, car la plupart des interactions entre les composantes du système climatique sont non linéaires. Par conséquent, le climat est considéré comme étant de nature chaotique et complexe le rendant imprédictible (Baede et al., 2001).

En lien avec les phénomènes décrits précédemment, la variabilité naturelle peut donc générer des changements climatiques pouvant s'étendre jusqu'aux ères glaciaires. Cependant, la période d'intérêt pour l'homme se restreint essentiellement de l'échelle annuelle à centennale. Dans cette optique, les variations atmosphériques et océaniques représentent les composantes principales résultant en la variabilité naturelle observée pour l'échelle temporelle d'intérêt. Plusieurs des processus externes affectant la variabilité naturelle, comme les cycles de Milankovitch, peuvent donc être fixés sans affecter les résultats.

1.2.2 Mesure de la variabilité naturelle

La variabilité naturelle aux échelles temporelles d'intérêt pour l'homme peut être explorée à travers des cycles de changements dans les états de l'atmosphère et des océans à différentes échelles temporelles allant de l'interannuelle à l'interdécennale. En général, des variations de pression ou de température entre deux endroits ou sur des régions particulières sont utilisées pour mesurer ces cycles océan-atmosphériques.

Plus spécifiquement pour l'Amérique du Nord, six indices climatiques principaux sont reconnus pour avoir une influence sur la variabilité hydroclimatique (dont les noms anglophones seront utilisés dans cette thèse) :

1. *El Niño-Southern Oscillation* (ENSO; Cayan, Redmond, & Riddle, 1999)
2. *Pacific North American pattern* (PNA; Rogers & Coleman, 2003)
3. *Arctic Oscillation* (AO; Déry & Wood, 2004)
4. *North Atlantic Oscillation* (NAO; Barlow, Nigam, & Berbery, 2001)
5. *Atlantic Multi-Decadal Oscillation* (AMO; Enfield, Mestas-Nuñez, & Trimble, 2001)
6. *Pacific Decadal Oscillation* (PDO; Stewart et al., 2005)

Chacun de ces indices climatiques représente un cycle se manifestant à une différente échelle temporelle, variant d'interannuelle (ENSO et PNA), à décennale (AO et NAO) jusqu'à interdécennale (AMO et PDO), affectant la variabilité climatique de l'Amérique du Nord de différentes façons (Rossi, Massei, & Laignel, 2011). Une description détaillée de chacun de ces indices est présentée dans l'article du chapitre 3.

Il est possible d'utiliser ces indices climatiques comme mesures de la variabilité naturelle pour expliquer une portion de la variabilité hydroclimatique (comme la température, les précipitations ou encore les débits des rivières.). Un grand nombre d'études ont notamment utilisé ces différents indices dans ce but pour différentes régions du Canada (Anctil & Coulibaly, 2004; Brabets & Walvoord, 2009; Burn, 2008; Gobena & Gan, 2009; Gobena, Weber, & Fleming, 2013; Kiffney, Bull, & Feller, 2002; Peters, Atkinson, Monk, Tenenbaum, & Baird, 2013; Spence, 2002; Thorne & Woo, 2011), des États-Unis (Hamlet & Lettenmaier, 1999; Hidalgo & Dracup, 2003; Massei et al., 2011; Maurer, Gibbard, & Duffy, 2006; Rogers & Coleman, 2003; Twine, Kucharik, & Foley, 2005) ou encore de l'Amérique du Nord (Bonsal & Shabbar, 2008; Déry & Wood, 2005; Fleming, Whitfield, Moore, & Quilty, 2007; Fu, James, & Wachowiak, 2012; Gobena & Gan, 2006; Nalley, Adamowski, Khalil, & Biswas, 2016; Rood, Samuelson, Weber, & Wywrot, 2005; Tootle, Piechota, & Singh, 2005; Wang, Whitfield, & Cannon, 2006; Whan & Zwiers, 2017; Woo & Thorne, 2008). Ces études utilisent une grande variété de méthodes statistiques (de très simples à très complexes) afin de tenter d'évaluer la relation entre les différents indices climatiques et la variabilité hydroclimatique sur plusieurs échelles temporelles (p.ex. saisonnière ou interannuelle).

En général, la majorité des travaux cités ci-dessus examinent principalement des variables climatiques dites moyennes, comme la température moyenne ou le cumul de précipitation à l'échelle saisonnière ou annuelle. En ce qui concerne la variabilité climatique des événements extrêmes, un nombre plus restreint d'études ont tenté d'en comprendre la relation avec ces mêmes indices climatiques (Fleming, Moore, & Clarke, 2006; Gershunov & Barnett, 1998; Gershunov & Cayan, 2003; Goly & Teegavarapu, 2014; Mo, Schemm, &

Yoo, 2009; St. George, 2007; Tan, Gan, & Shao, 2016; Wang et al., 2006; Whan & Zwiers, 2017).

Il est important de souligner que l'utilisation de différentes méthodes, qui elles sont employées sur une grande variété de sites d'étude, rendent difficile la capacité de dresser un portait clair de la relation entre la variabilité naturelle et ces différents indices. De plus, il est très probable que ces différents indices climatiques s'affectent mutuellement, rendant leur évaluation encore plus complexe de ce que la littérature peut parfois laisser sous-entendre (Emerton et al., 2017; Levine, McPhaden, & Frierson, 2017). Malgré tout, ces indices climatiques se montrent comme étant la meilleure option pour mesurer et comprendre la variabilité naturelle.

1.3 Détection du signal des changements climatiques

Pour mieux convaincre les preneurs de décisions de mettre en place des mesures d'adaptation face aux changements climatiques d'origine anthropique, il est essentiel de détecter la tendance de ces derniers au travers d'une série chronologique affectée par la variabilité naturelle. Il existe plusieurs méthodes plus ou moins complexes permettant d'y arriver, et en général elles peuvent être séparées en deux catégories : les tests non paramétriques et les tests paramétriques.

1.3.1 Tests non paramétriques

Malgré qu'il existe plusieurs tests non paramétriques permettant de détecter une tendance ou un point de rupture dans une série chronologique (p. ex. : test de Pettitt, test de Buishand, etc.), le test le plus répandu est sans aucun doute celui de Mann-Kendall (Kendall, 1975). Ce dernier a été utilisé dans de nombreuses études de changements climatiques, comme : Donat et al. (2013b); Lins et Slack (1999); Westra et al. (2013). Le test de Mann-Kendall permet de détecter une tendance significative dans une série temporelle locale (p. ex. : une station météorologique) avec un niveau de confiance choisit par l'utilisateur et est calculé de la façon suivante :

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{m} sign(x_i - x_j) \quad (1.1)$$

où x est la valeur de la variable étudiée au temps i et j, avec *sign( )* étant égal à +1 si x_i est plus grand que x_j et -1 si x_i est plus petit que x_j. S représente le nombre de fois que x_i est plus grand que x_j moins le nombre de fois que x_i est plus petit que x_j. Le signe de S indique aussi le signe de la tendance. Selon l'hypothèse nulle, la moyenne S du test est égale à 0. Pour de petits échantillons ($n < 10$), les valeurs du S tendent vers une distribution bêta, tandis que les distributions bêta et normale fournissent une bonne approximation pour les grandes valeurs de n (Hamed, 2009). Des corrections peuvent être aussi apportées au test de Mann-Kendall pour tenir compte des autocorrélations, dont l'existence pourrait faussement augmenter la probabilité de détection d'une tendance significative (Hamed & Rao, 1998).

Un test comme celui de Mann-Kendall est couramment utilisé à l'échelle locale plutôt que régionale. Tel que discuté dans la section précédente, il est attendu que la variabilité naturelle aura un impact plus important à l'échelle locale que régionale, augmentant donc la probabilité de masquer le signal des changements climatiques (Deser et al., 2012a; Fischer & Knutti, 2014).

Une alternative est d'effectuer un test de détection de la tendance à l'échelle régionale par rééchantillonnage des séries chronologiques (ou *field significance resampling* en anglais). Ce type d'analyse permet d'établir la significativité statistique d'une tendance à l'échelle régionale en tenant compte de la corrélation spatiale des stations météorologiques ou des points voisins d'un jeu de données sur grille. Parmi les méthodes employées pour ce type d'analyse, celles du *false discovery rate* (Renard et al., 2008; Wilks, 2006) et du rééchantillonnage par *bootstrap* (Douglas, Vogel, & Kroll, 2000; Kiktev, Sexton, Alexander, & Folland, 2003; Westra et al., 2013) se sont montrées adéquates et robustes. La méthode du rééchantillonnage par *bootstrap*, qui est décrite en détail dans l'article de Douglas et al. (2000) et est survolée ci-dessous.

Dans le cas d'un jeu de données sur grille couvrant la période de 1950 à 2010, une région peut être définie en utilisant un certain nombre de points (p. ex. : 3 × 3 = 9 points au total) en associant le résultat de l'analyse au point central de chaque région de 9 points de grilles. La moyenne régionale du *S* de Mann-Kendall ($\bar{S}_m$) est calculée comme étant la moyenne de la valeur du *S* pour chaque point de grille dans une région sélectionnée :

$$\bar{S}_m = \frac{1}{m}\sum_{k=1}^{m} S_k \tag{1.2}$$

où S_k est le *S* de Mann-Kendall (voir équation 1.1) pour le point de grille *k* dans une région de *m* points (*m* = 9 dans cet exemple).

Ensuite, pour déterminer si la tendance régionale est significative ou non, une approche de rééchantillonnage par *bootstrap* peut être réalisée tel que proposée par Douglas et al. (2000). Pour chaque échantillon du *bootstrap*, un échantillon de 60 années tiré aléatoirement avec replacement entre 1950 et 2010 est réalisé. Le même échantillon d'années tiré est réutilisé pour chaque point de grille de la région pour calculer le *S* de Mann-Kendall, permettant de conserver la corrélation spatiale entre les points de grille. La moyenne régionale de Mann-Kendall $\bar{S}_m$ est ensuite calculée avec l'équation 1.2. Cette procédure est réalisée un grand nombre de fois (p. ex. : 1 000 fois), en classant en ordre croissant les valeurs de *S* assignée à une probabilité de non-dépassement basé sur une formule de position empirique, telle que celle de Weibull :

$$P = \frac{r}{B+1} \tag{1.3}$$

où *r* est le rang de chaque échantillon et *B* est le nombre total d'échantillons du *bootstrap*. Un niveau de confiance (p. ex. : 95%) peut être obtenu à l'aide de la fonction de répartition empirique en prenant le 25^{e} rang (α = 2.5% ; tendance significative négative) et le 975^{e} rang (α = 97.5% ; tendance significative positive) avec *B* = 1 000. Afin de tenir compte de l'autocorrélation temporelle, une procédure de *bootstrap* par bloc-mobile peut aussi être

réalisée. La méthode est décrite en détail dans Wilks (1997) et Wilks (2011). Toutefois, les séries des maxima annuelles de précipitation semblent être généralement absente d'une telle autocorrélation temporelle (Westra et al., 2013).

1.3.2 Tests paramétriques

Différentes lois statistiques peuvent être aussi utilisées pour modéliser les séries chronologiques des extrêmes, permettant aussi d'extrapoler vers des événements plus rares avec une période de retour excédant la durée de la série. Une des lois statistiques les plus utilisées au niveau des événements extrêmes est la loi des extrêmes généralisés (*Generalized Extreme Value distribution* – GEV en anglais) définie par :

$$GEV(x; \mu, \sigma, \xi) = exp\left\{-\left[1 + \frac{\xi(x-\mu)}{\sigma}\right]^{-\frac{1}{\xi}}\right\}, 1 + \frac{\xi(x-\mu)}{\sigma} > 0 \qquad (1.4)$$

où trois paramètres sont ajustables : le paramètre de position (μ), le paramètre d'échelle (σ) et le paramètre de forme (ξ) (Coles 2001, Katz 2013). Afin d'ajuster les paramètres de cette loi en fonction des valeurs de la série chronologique d'intérêt (x), le maximum de vraisemblance peut être utilisé :

$$L(\theta;\ x_1, \ldots, x_n,) = f(x_1, x_2, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \qquad (1.5)$$

où L est la vraisemblance, θ la probabilité, x_i les données de la série chronologique et n le nombre de données. En maximisant la vraisemblance, le meilleur ajustement possible de la loi par rapport aux observations est obtenu.

Il est aussi possible de détecter une tendance significative au travers de la série chronologique utilisée en ajoutant un ou plusieurs paramètres covariables dans la loi statistique (Katz 2013, Westra et al. 2013). Par exemple, une tendance significative pourrait

faire en sorte que le paramètre de position (μ) évolue de façon linéaire dans le temps (t) à cause des changements climatiques, causant un déplacement dans la distribution des extrêmes, tel que décrit à l'équation 1.6 :

$$GEV(\mu_t, \sigma, \xi) \text{ avec } \mu_t = \beta_0 + \beta_1 x_t \tag{1.6}$$

où β_0 se retrouve à être le paramètre de position à la première année (p. ex. : 1950) et le β_1 est une tendance linéaire en mm/année qui est ajouté à la valeur initiale de β_0 selon l'année qui est modélisée (Katz 2013). En ajoutant une tendance temporelle linéaire à ce paramètre (induisant la non-stationnarité du modèle), il est possible de déterminer si l'ajustement du modèle avec ce paramètre supplémentaire est significativement meilleur que la variante stationnaire du modèle. Pour y arriver, différents critères, comme le *Akaike Information Criterion* (*AIC*; équation 1.7), le *Bayesian Information Criterion* (*BIC*; équation 1.8) ou encore le *Likelihood-ratio test* (équation 1.9) peuvent être utilisé.

$$AIC = 2k - 2\ln(L) \tag{1.7}$$

$$BIC = -2\ln(L) + k\ln(n) \tag{1.8}$$

$$D = -2\ln\left(\frac{L \text{ for null model}}{L \text{ for alternative model}}\right) \tag{1.9}$$

Ces différentes méthodes sont toutes basées sur le logarithme naturel du maximum de la vraisemblance (L) décrit à l'équation 1.5 et du nombre de paramètres (k) utilisé dans la loi statistique (Katz 2013). Dans le cas du *Likelihood-ratio test*, la statistique du test (D) et le nombre de degrés de liberté entre les deux modèles sont comparés à une distribution khi carré (χ^2). De cette façon, il est possible de comparer plusieurs variantes non stationnaires de la GEV en induisant une tendance temporelle sur un autre ou plusieurs de ses paramètres, comme les paramètres de position (μ) et d'échelle (σ). Dans un tel exemple, il y aurait un déplacement dans la distribution de même qu'un élargissement ou rétrécissement. Une autre alternative pourrait être d'utiliser des paramètres covariables qui varient en fonction de la

température moyenne de la région ou même de la planète. Ainsi, l'effet du réchauffement planétaire serait directement pris en considération via ces paramètres.

Il est également possible d'évaluer une tendance régionale plutôt que locale en incluant un plus grand nombre de séries chronologiques dans l'ajustement des paramètres de la GEV. Par exemple, encore dans le cas d'un jeu de données sur grille couvrant la période de 1950 à 2010, une région de 9 points de grille (3 × 3) peut aussi être utilisée pour définir une région. Lors de l'ajustement des paramètres de la GEV, les 9 séries chronologiques sont prises en considération, tenant compte de la corrélation spatiale dans la région sélectionnée.

1.4 Impacts futurs des changements climatiques

Les séries chronologiques d'observations historiques peuvent être difficilement utilisées pour évaluer les changements projetés sur les extrêmes hydrométéorologiques. La meilleure alternative disponible est l'approche du monde virtuel, dans lequel un modèle climatique est utilisé comme substitut du monde réel. Ce type de modèle permet de créer une réplique virtuelle du monde réel extrêmement riche en données climatiques. Dans le monde virtuel des modèles climatiques, une importante quantité de temps de calcul nécessite l'utilisation de superordinateurs pour générer des simulations climatiques dans un délai raisonnable. Ces simulations permettent ensuite de créer des séries chronologiques avec une échelle spatiale et temporelle beaucoup plus importante que celle des séries observations du monde réel, limitées aux stations météorologiques. Ce monde virtuel a aussi l'avantage d'être basé sur les lois fondamentales de conservation de la masse, de l'énergie et du moment, rendant les données physiquement cohérentes (Caya et Laprise, 1999; Music et Caya, 2007). En utilisant différents scénarios d'émissions de gaz à effet de serre, il est possible de projeter le climat dans un futur rapproché ou même éloigné afin d'évaluer les impacts de l'influence humaine.

1.4.1 Types de modèles climatiques

L'un des types de modèles climatiques le plus couramment utilisés est le modèle de circulation générale (MCG; ou souvent référé comme un modèle climatique global – ou

l'acronyme anglais GCM référant à *Global Circulation Model* ou *Global Climate Model*, qui sera utilisé ci-après) qu'on retrouve sous trois formes principales :

- Modèle de circulation générale atmosphérique ; AGCM
- Modèle de circulation générale océanique ; OGCM
- Modèle de circulation générale océan-atmosphère (couplés) ; AOGCM

De nos jours, les GCMs font souvent référence aux modèles océan-atmosphère couplés en incluant des modèles continentaux de glace et de surface. Les GCMs permettent donc de simuler l'ensemble des processus physiques du climat, comme la température, la précipitation, le vent, les nuages, etc., et leurs interactions. Avec la complexité grandissante de ces modèles, le modèle de système terrestre (ou *Earth System Model;* ESM en anglais) représente la nouvelle génération de modèle couplé qui inclut aussi des modèles de biogéochimie marine et de chimie atmosphérique. L'avantage d'un ESM est la possibilité de simuler l'évolution des gaz à effet de serre, plutôt que d'utiliser une composition prédéterminée comme il était le cas pour le GCM.

Les modèles climatiques étant de plus en plus utilisés auprès de la communauté scientifique, il existe de grands ensembles de modèles climatiques couplés (GCM ou ESM) couvrant la planète, tels que le CMIP3 (*Coupled Model Intercomparison Project* ; Meehl et al., 2007) et le CMIP5 (Taylor, Stouffer, & Meehl, 2011), rendant accessibles les simulations de près d'une trentaine de modèles. Une grande partie des derniers travaux de l'IPCC est entre autres basée sur ces deux ensembles (IPCC, 2013). Malgré les nombreux avantages des modèles climatiques, la résolution spatiale de ceux-ci est souvent trop grossière pour analyser certaines variables climatiques telles que les précipitations, notamment pour les extrêmes. Pour contrer ce manque, différentes méthodes de descente d'échelle peuvent être utilisées afin de transférer l'information obtenue par les GCMs/ESMs d'une résolution spatiale grossière vers une résolution spatiale plus fine. Ces méthodes sont regroupées en deux catégories : les méthodes de descente d'échelle statistique et dynamique.

1.4.2 Méthodes de descente d'échelle statistique

Il existe plusieurs méthodes statistiques de descente d'échelle. Le sujet est maintenant suffisamment vaste qu'au moins trois livres s'y consacrent (Benestad, Hanssen-Bauer, & Chen, 2008; Lee & Singh, 2018; Maraun & Widmann, 2018). Ces méthodes peuvent être regroupées en trois catégories : le pronostic parfait (*perfect prognosis;* PP), les statistiques des sorties de modèles (*model output statistics*; MOS) et les générateurs de météo (weather generators; WG) (Hertig et al., 2019; Maraun et al., 2010a).

Les PP sont calibrés uniquement sur les données historiques d'observations à l'aide de méthodes comme celles des régressions linéaires multiples et des analogues. Elles établissent des relations statistiques entre les prédicteurs (typiquement des variables atmosphériques à grande échelle comme les champs de pressions) et les prédictants (des variables locales et régionales comme la précipitation et la température). Ces dernières sont basées sur le principe que des valeurs similaires pour des variables à grande échelle (p. ex. les champs de pressions) conduisent aussi à des valeurs similaires pour des variables locales comme la précipitation (Hertig et al., 2019; Vidal, Hingray, Magand, Sauquet, & Ducharne, 2016). Les MOS couvrent principalement les méthodes de corrections de biais des modèles climatiques, faites à partir des séries chronologiques d'observations (Hertig et al., 2019). Ce type de méthode est couvert plus en détail à la section 1.5.3. Finalement, les WG sont des modèles stochastiques calibrés sur les données historiques d'observations permettant de générer des séries de variables météorologiques, comme la précipitation, avec des caractéristiques statistiques similaires aux observations (Fowler, Blenkinsop, & Tebaldi, 2007; Maraun et al., 2010a).

Bien que les méthodes de réduction d'échelle statistique aient l'avantage d'être plus faciles à appliquer et moins coûteuses en temps de calcul que les méthodes dynamiques, celles-ci ont des faiblesses qui doivent toutefois être considérées. Entre autres, les méthodes statistiques sont dépendantes de la longueur et de la qualité des séries chronologiques d'observations, elles ont tendance à sous-estimer la variance et à mal représenter les événements extrêmes.

De plus, toute relation établie entre les prédicteurs et les prédictants est par défaut considérée comme étant stationnaire (Fowler et al., 2007; Hertig et al., 2019; Maraun et al., 2010a).

1.4.3 Méthodes de descente d'échelle dynamique

La descente d'échelle dynamique est réalisée par le biais d'un modèle régional du climat (en anglais *Regional Climate Model*; RCM, ci-après utilisé). Ce type de modèle permet de créer un certain domaine d'étude (p. ex. : l'Amérique du Nord ou l'Europe) à une résolution plus fine et en étant piloté à ses frontières par des données climatiques à une résolution plus grossière (Fatichi, Rimkus, Burlando, & Bordoy, 2014; Fowler et al., 2007; Schmidli, Frei, & Vidale, 2006). Le RCM est habituellement piloté à ses frontières par les données issues d'un GCM/ESM (qui peut aussi avoir une structure différente du RCM) ou même d'une réanalyse météorologique (p. ex. : ERA-Interim; Dee et al., 2011) gardant une certaine cohérence avec les observations. Un certain nombre de grands ensembles de RCMs ont aussi été produits sur diverses zones d'études d'intérêts. Par exemple, NA-CORDEX (Giorgi, Jones, & Asrar, 2009) et NARCCAP (Mearns et al., 2012) couvrent l'Amérique du Nord et EURO-CORDEX (Jacob et al., 2014) et ENSEMBLES (van der Linden & Mitchell, 2009) couvrent l'Europe.

La descente à l'échelle dynamique avec l'utilisation des RCMs à haute-résolution a apporté des améliorations significatives par rapport aux GCMs/ESMs, notamment au niveau des précipitations moyennes et extrêmes aux échelles locales et régionales (Maraun et al., 2010a; Prein et al., 2013; Tripathi & Dominguez, 2013). Cependant, la résolution spatio-temporelle utilisée dans les RCMs est encore largement dépendante de la puissance informatique disponible, entraînant une résolution spatiale qui demeure relativement grossière ($\approx$ 12 km de résolution horizontale pour les nouvelles simulations couvrant l'ensemble du 21^{e} siècle). Certaines études suggèrent que les événements météorologiques synoptiques (généralement attribuables aux extrêmes de précipitations journalières et de plus longue durée) sont relativement bien simulés à la résolution spatiale des RCMs et potentiellement celles des GCMs/ESMs (Ban, Schmidli, & Schär, 2014; Chan et al., 2014; Kendon et al., 2017; Prein et al., 2015).

Cependant, même à une résolution d'environ ≈ 12 km, les processus de convection ne peuvent pas être complètement résolus et doivent plutôt être paramétrés. Lorsque les orages estivaux directement liés à la convection sont d'intérêt (p. ex. : pour les extrêmes sous-journaliers de précipitations), une résolution beaucoup plus fine (< 1 km) est nécessaire pour résoudre la convection (Chan et al., 2014; Kendon et al., 2017; Prein et al., 2015; Prein et al., 2017; Westra et al., 2014). Les modèles permettant la convection (en anglais *Convection Permitting Model*; CPM) exploités à des résolutions avoisinantes celle du kilomètre sont utilisés pour contourner cette problématique (Prein et al., 2015; Prein et al., 2017). Par contre, la puissance de calcul nécessaire pour utiliser de tel modèle est extrêmement élevée et il devient actuellement impensable d'obtenir de longues simulations, couvrant par exemple l'ensemble du 21^{e} siècle.

Il est notable que la majorité des études d'impact des changements climatiques qui ont été réalisées, notamment au niveau des extrêmes hydrométéorologiques, aient été principalement limitées à l'utilisation des GCMs/ESMs et des RCMs.

1.4.4 Les grands ensembles de simulations climatiques

Les grands ensembles de simulations climatiques d'un même modèle climatique sont une autre alternative aux grands ensembles de GCMs/ESMs à l'échelle globale (p. ex. : CMIP3 et CMIP5) et de RCMs à l'échelle régionale (p. ex. : NA-CORDEX et EURO-CORDEX). Ce type d'ensemble est obtenu en introduisant de faibles perturbations aléatoires dans les conditions initiales d'un même modèle climatique (GCM, ESM ou RCM), puis en conduisant les simulations avec les mêmes forçages externes (Deser et al., 2012a). La météo étant de nature chaotique, les simulations vont diverger rapidement, entraînant un ensemble de simulations climatiques équiprobables couvrant la même période (p. ex. : 1950 à 2100). Dans l'exemple illustré à la Figure 1.1, 5 simulations ont été perturbées 10 fois chacune, pour générer un ensemble de 50 simulations. Il est possible de constater qu'après seulement deux semaines (tel que représenté par une coupure dans le graphique pour y mettre l'emphase) que les simulations ont déjà complètement divergé les unes des autres.

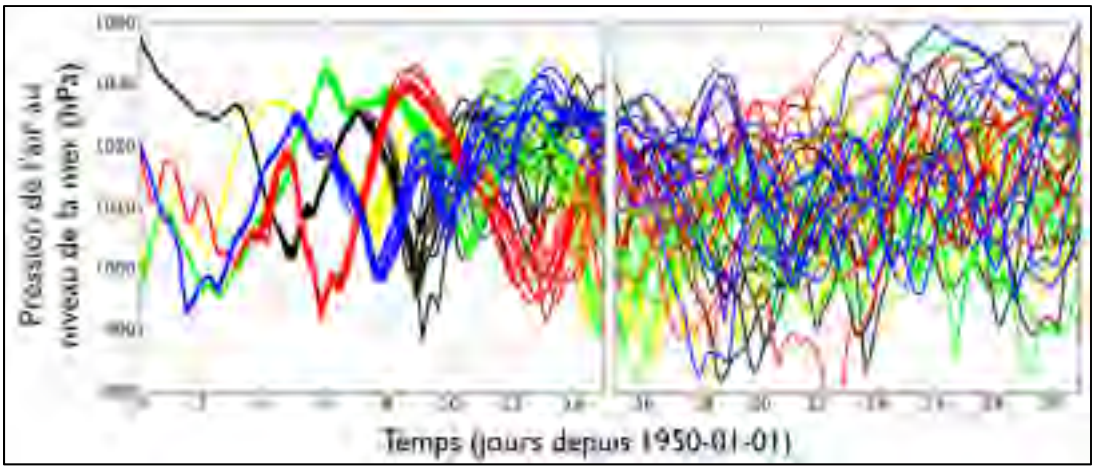

Figure 1.1 Période de démarrage du grand ensemble du CRCM5 représenté par la pression au niveau de la mer près de Munich pour le mois de janvier 1950 (jours 0 à 15 sur le panneau de gauche et jours 15 à 31 sur le panneau de droite).
Tirée de Leduc et al. (2016a)

Les grands ensembles de simulations sont de plus en plus utilisés dans les études d'impacts des changements climatiques (p. ex. : Aalbers, Lenderink, van Meijgaard, & van den Hurk, 2017; Deser et al., 2012a; Deser et al., 2012b; Deser et al., 2014; Fischer et al., 2014; Kay et al., 2015; Sanderson et al., 2018; Thompson et al., 2015).

1.4.5 Changements projetés sur les précipitations extrêmes

Une bonne quantité d'études ont eu recours aux GCMs/ESMs et aux RCMs pour évaluer les impacts futurs des changements climatiques sur les précipitations extrêmes à différentes échelles spatio-temporelles. En général, la majorité des études ont examiné des métriques d'extrêmes relativement fréquents, telles que les maxima annuels de précipitations journalières (RX1day) ou cumulatives sur 5 jours (RX5day) proposés par l'ETCCDI (*Expert Team on Climate Change Detection and Indices*) (Zhang et al., 2011). Un nombre beaucoup moins élevé de travaux a considéré les événements extrêmes plus rares, comme les périodes de retour 20 ans ou 100 ans.

À l'échelle globale, Kharin, Zwiers, Zhang, et Hegerl (2007) et Kharin, Zwiers, Zhang, et Wehner (2013) ont conduit des études à l'échelle globale, avec l'ensemble CMIP3 et CMIP5 respectivement, sur les changements projetés (horizon 2100) de la période de retour 20 ans estimés à partir des maxima annuels de précipitations journalières. Les résultats de ces deux études indiquent une augmentation globale de la fréquence des extrêmes de précipitations (correspondant à une réduction de la période de retour de référence), à l'exception de quelques régions tropicales et sous-tropicales. De plus, les conclusions sont demeurées les mêmes à travers trois scénarios d'émissions de GES (RCP2.6, RCP4.5 et RCP8.5) pour lesquels les valeurs médianes globales de la période de retour future correspondante à l'intensité de la période de retour 20 ans de référence étaient respectivement de 14, 11 et 6 ans.

À l'échelle régionale, les différentes études réalisées pointent principalement vers des augmentations importantes, notamment pour certaines régions de l'Amérique du Nord (Mailhot & Duchesne, 2010; Mailhot, Duchesne, Caya, & Talbot, 2007; Mladjic et al., 2011; Wehner, 2013; Wuebbles et al., 2013; Zhu, 2013). Il est aussi notable que les augmentations futures projetées obtenues dans ces différentes études demeurent en accord avec les augmentations observées discutées précédemment à la section 1.1.1.

1.4.6 Changements projetés sur les débits en rivière extrêmes

Malgré le fait qu'un grand nombre d'études ait été conduit sur l'impact des changements climatiques sur les débits en rivière, un nombre encore limité a examiné les projections au niveau des extrêmes (Seneviratne et al., 2012). En ce qui concerne les grandes périodes de retour (20 ans ou 100 ans), les projections suggèrent de grandes divergences au niveau des changements en termes d'augmentations et de diminutions sur l'ensemble des sites étudiés (Dankers & Feyen, 2008; Dankers, Feyen, & Christensen, 2009; Rojas, Feyen, Bianchi, & Dosio, 2012). Ces conclusions sont généralement cohérentes à travers les différents scénarios d'émission de GES (Dankers & Feyen, 2008), les différentes résolutions spatiales testées (Dankers et al., 2009) et au travers des différents modèles climatiques employés (Rojas et al.,

2012). Il est notable que la vaste majorité de ces études a employé une méthode de post-traitement pour corriger les biais pour les champs de précipitation et de température avant d'en alimenter le modèle hydrologique utilisé.

1.4.7 Impact du scénario d'émission de GES

Un nombre relativement important de scénarios d'émission de GES a été utilisé à travers les différentes études d'impact des changements climatiques, comme le RCP2.6, RCP4.5 et RCP8.5, dont le dernier se trouve à être le scénario le plus pessimiste dans cette liste (Meinshausen et al., 2011). La majorité des travaux sur les extrêmes hydrométéorologiques ayant été réalisés en utilisant plus d'un scénario indique que les conclusions générales demeurent les mêmes, avec des changements plus importants attribuables aux scénarios de plus grandes émissions (Dankers & Feyen, 2008; Kharin et al., 2007; Kharin et al., 2013; Rajczak & Schär, 2017).

1.5 Utilisation de la modélisation hydrologique

Malgré que les modèles climatiques simulent le cycle de l'eau et son ruissellement, le niveau de détail dans la représentation des processus est généralement insuffisant et requiert l'utilisation d'un modèle hydrologique pour la génération des débits en rivière afin d'en faire l'étude (Graham, Hagemann, Jaun, & Beniston, 2007; Rojas et al., 2012). Les études décrites dans la section 1.4.4 ont toutes eu recours à un modèle hydrologique pour obtenir les séries temporelles de débits en rivière dans un climat futur. Cependant, l'utilisation d'un modèle hydrologique pour la projection des débits en rivière génère plusieurs défis en soi, notamment : le choix du modèle, la calibration et la correction des biais des données d'entrées.

1.5.1 Choix du modèle hydrologique

Les deux principales catégories de modèles hydrologiques sont les modèles globaux (simple d'application et généralement facile à employer) et distribués (plus complexe à mettre en

place) (Pechlivanidis, Jackson, McIntyre, & Wheater, 2011). Essentiellement, le modèle de type global considère le bassin versant comme étant une entité unique, et les différentes variables d'états et météorologiques sont moyennées sur l'ensemble du bassin. En ce qui concerne le modèle distribué, celui-ci représente le bassin versant à l'aide de plusieurs sous-bassins (c.-à-d. : modèle semi-distribué) ou encore avec un maillage régulier ou irrégulier. Généralement, les processus hydrologiques sont simulés sur chaque maille (ou sous-bassin) et le routage est par la suite fait vers l'exutoire du bassin versant.

Une étude d'intercomparaison des deux types de modèles, le *Distributed Model Intercomparison Project* (DMIP) et une étude de suivi (DMIP2) ont été réalisés, entre autres, pour évaluer la valeur ajoutée d'un modèle distribué. À travers ces deux études, il a été démontré que les modèles globaux performent généralement aussi bien que les modèles distribués pour la simulation des débits à l'exutoire du bassin versant (Reed et al., 2004; Smith et al., 2012).

1.5.2 Calibration du modèle hydrologique

À ce jour, plusieurs processus physiques décrivant le cycle hydrologique ne sont pas entièrement compris et il y a aussi à un manque important de données pour l'utilisation des équations les plus complètes (Andreassian, Hall, Chahinian, & Schaake, 2006; Jakeman & Hornberger, 1993). Afin de contourner cette problématique, les modèles hydrologiques ont recours à une simplification des processus à travers différentes simplifications et un certain nombre de paramètres ajustables. Il est donc essentiel de performer une calibration du modèle hydrologique afin d'en obtenir les paramètres optimaux pour le bassin versant à l'étude.

La calibration du modèle hydrologique se fait avec une fonction objectif permettant d'obtenir un hydrogramme simulé le plus similaire possible aux observations (Arsenault, Poulin, Côté, & Brissette, 2014; Gupta et al., 2014). La fonction objectif de Nash et Sutcliffe (*NSE*; Nash & Sutcliffe, 1970) est la plus répandue dans le domaine de l'hydrologie (Jain & Sudheer,

2008). Toutefois, cette dernière est reconnue pour attribuer une trop grande importance sur les pointes de crues (McCuen, Knight, & Cutter, 2006). Le *NSE* est calculé à l'aide de l'équation suivante :

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{o,i} - Q_{s,i})^2}{\sum_{i=1}^{n}(Q_{o,i} - \overline{Q_{o,i}})^2} \quad (1.10)$$

où le *NSE* représente la fonction objectif de Nash et Sutcliffe, Q_o les débits observés, Q_s les débits simulés et *i* l'indice représente le jour de la simulation.

Une autre fonction objectif étant de plus en plus utilisée en hydrologie est celle de Kling et de Gupta (*KGE* ; Gupta, Kling, Yilmaz, & Martinez, 2009) et est définie comme suit :

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad (1.11)$$

où *KGE* est la fonction objectif de Kling et de Gupta, *r* est le coefficient de corrélation, α est une mesure de la variabilité et β est le biais entre les débits observés et simulés. Au moyen de sa nature multiobjective, visant à optimiser les erreurs liées à la corrélation, la variabilité et le biais, le *KGE* a été démontré comme étant une meilleure alternative à la fonction objectif du *NSE* (Pokhrel & Gupta, 2011), tout en y étant similaire.

Une fois la fonction objectif sélectionnée par l'utilisateur, la calibration du modèle hydrologique peut être réalisée en utilisant divers algorithmes d'optimisation. Un exemple répandu en hydrologie est l'algorithme du *Shuffled Compex Evolution* de l'Université de l'Arizona (SCE-UA; Duan, Sorooshian, & Gupta, 1992). Il existe des travaux ayant comparé la performance de différents algorithmes d'optimisation en fonction d'un budget d'évaluations fixe afin de fournir des directives sur quel algorithme sélectionner en fonction de la complexité du modèle hydrologique choisi (Arsenault et al., 2014). Par exemple, Arsenault et al. (2014) ont démontré que pour un modèle avec peu de paramètres, l'algorithme du SCE-UA avec un budget de 10 000 évaluations se trouve à être un choix optimal.

Plusieurs méthodes de calibration sont utilisées afin de valider la robustesse des paramètres obtenus. La méthode la plus classique demeure la validation de l'échantillon partagé, séparant la série temporelle d'observation en deux échantillons égaux, effectuant la calibration sur la première moitié et la validation sur la deuxième moitié (Klemeš, 1986). Un avantage d'un tel type de validation est la réduction de temps de calcul nécessaire à la calibration, puisque seulement la moitié de la série temporelle est utilisée lors de la simulation des débits. Une autre méthode très répandue est celle de la calibration des années paires/impaires (calibration sur les années paires et validation sur les années impaires), permettant de prendre en considération la tendance émergente des changements climatiques dans la série temporelle dans l'ajustement des paramètres (Arsenault, Essou, & Brissette, 2017; Essou, Arsenault, & Brissette, 2016; Gowda, Mulla, Desmond, Ward, & Moriasi, 2012).

Dans un contexte d'études des impacts des changements climatiques où la modélisation hydrologique est utilisée pour obtenir les débits à partie des variables de sortie des modèles climatiques, le modèle hydrologique doit être tout de même calé. Étant donné qu'il n'y a pas de séries temporelles de débits en rivière disponibles dans les modèles climatiques pour faire la calibration, le modèle hydrologique est généralement calé à l'aide des données observées (Chen, Brissette, & Leconte, 2011a; Chen, Brissette, Poulin, & Leconte, 2011b; Minville, Brissette, & Leconte, 2008).

1.5.3 Correction du biais des modèles climatiques

Malgré que de grandes améliorations soient constamment apportées dans la structure des GCMs/ESMs et RCMs, ces données sont souvent considérées comme étant trop biaisées pour être directement utilisées dans une étude d'impact des changements climatiques (Addor & Fischer, 2015; Chen, Brissette, Chaumont, & Braun, 2013; Chen, Brissette, & Lucas-Picher, 2015; Knutti, Masson, & Gettelman, 2013; Maraun et al., 2010a). Afin de répondre à cette problématique, plusieurs méthodes de post-traitement avec différents niveaux de complexités

ont été développées pour faire la correction de ces biais (Chen et al., 2013; Maraun, 2016; Maraun et al., 2010a; Teutschbein & Seibert, 2012).

Un nombre important d'articles (Addor & Seibert, 2014; Chen et al., 2013; Chen et al., 2015; Ehret, Zehe, Wulfmeyer, Warrach-Sagi, & Liebert, 2012; Maraun, 2016) ont souligné les limitations de ces différentes méthodes de correction qui doivent être considérées. Parmi celles-ci, il y en a notamment cinq qui méritent d'être soulignées :

- L'importance de la qualité des données observées utilisées pour faire la correction des biais (Addor & Fischer, 2015; Sunyer et al., 2013) ;
- L'hypothèse de stationnarité des biais (Buser, Künsch, Lüthi, Wild, & Schär, 2009; Chen et al., 2015; Maraun, 2012, 2013a; Nahar, Johnson, & Sharma, 2017; Piani, Haerter, & Coppola, 2010) ;
- La difficulté à corriger les événements extrêmes (Chen et al., 2013; Chen et al., 2011a; Maraun, 2013a; Maraun et al., 2010b; te Linde, Aerts, Bakker, & Kwadijk, 2010) ;
- Le manque de cohérence inter-variables (Chen et al., 2018) ;
- Le manque de cohérence spatiale (Maraun et al., 2017; Vrac & Friederichs, 2015).

Malgré les limitations des différentes méthodes de correction de biais, il en demeure que ces méthodes apportent des améliorations considérables pour la simulation des débits en rivière (Addor & Seibert, 2014; Muerth et al., 2012b; Rojas, Feyen, Dosio, & Bavera, 2011). Une des principales raisons est qu'un biais important dans une ou plusieurs variables utilisées comme données d'entrées du modèle hydrologie (p. ex. : précipitation et température) se retrouve à modifier considérablement le cycle hydrologique.

CHAPITRE 2

DÉMARCHE DU TRAVAIL ET ORGANISATION DU DOCUMENT

2.1 Démarche du travail

Tel que présenté en introduction, les objectifs de cette thèse se résument en deux principaux axes de recherche qui visent à :

1. Déterminer l'influence de la variabilité naturelle sur la détection du signal des changements climatiques ;
2. Évaluer l'influence des changements climatiques sur la probabilité de récurrence des événements hydrométéorologiques extrêmes.

Suivant la revue de la littérature du chapitre 1, un survol rapide des bases de données, des outils et de la démarche méthodologique est introduit dans les prochaines sections. Quatre articles ont été réalisés à partir de ces éléments afin de répondre aux deux axes de recherche et sont présentés dans les prochains chapitres.

2.1.1 Les grands ensembles de simulations climatiques

Comparativement aux ensembles de modèles climatiques (p. ex. CMIP5 regroupant plusieurs GCMs/ESMs), les ensembles de simulations climatiques d'un même modèle permettent d'investiguer les axes de recherche de cette thèse de façon plus robuste. Par exemple, ce type d'ensemble permet de séparer la variabilité intermodèle de la variabilité naturelle (Fischer et al., 2013; Kay et al., 2015), permettant entre autres de mieux cibler l'incertitude associée à la variabilité naturelle et de mieux comprendre comment celle-ci peut masquer le signal des changements climatiques. De plus, ayant accès à un grand nombre de séries temporelles couvrant la même période (p. ex. : 50 membres × 20 années = 1 000 années au total pour les ensembles du CanESM2 et du CRCM5), il est possible d'étudier des événements d'une période de retour plus grande (p. ex. : un orage de récurrence 1 dans 100 ans) tout en réduisant l'incertitude statistique (Schulz & Bernhardt, 2016).

C'est principalement pour ces raisons que les grands ensembles de simulations climatiques ont un rôle méthodologique prédominant dans les travaux de cette thèse. Trois grands ensembles de simulations climatiques sont utilisés dans cette thèse :

1. Les 50 membres du *Canadian Earth System Model version 2* (CanESM2) à l'échelle globale et couvrant la période de 1950 à 2100;
2. Les 40 membres du *Community Earth System Model version 1* (CESM1) à l'échelle globale et couvrant la période de 1920 à 2100;
3. Les 50 membres du *Canadian Regional Climate Model version 5* (CRCM5) à l'échelle régionale piloté à ses frontières par le grand ensemble du CanESM2 et couvrant la période de 1950 à 2100.

Un résumé de la résolution spatiale et de la couverture temporelle de ces trois grands ensembles est présenté à la Figure 2.1. Les trois ensembles sont décrits avec plus de détails dans les trois sous-sections suivantes.

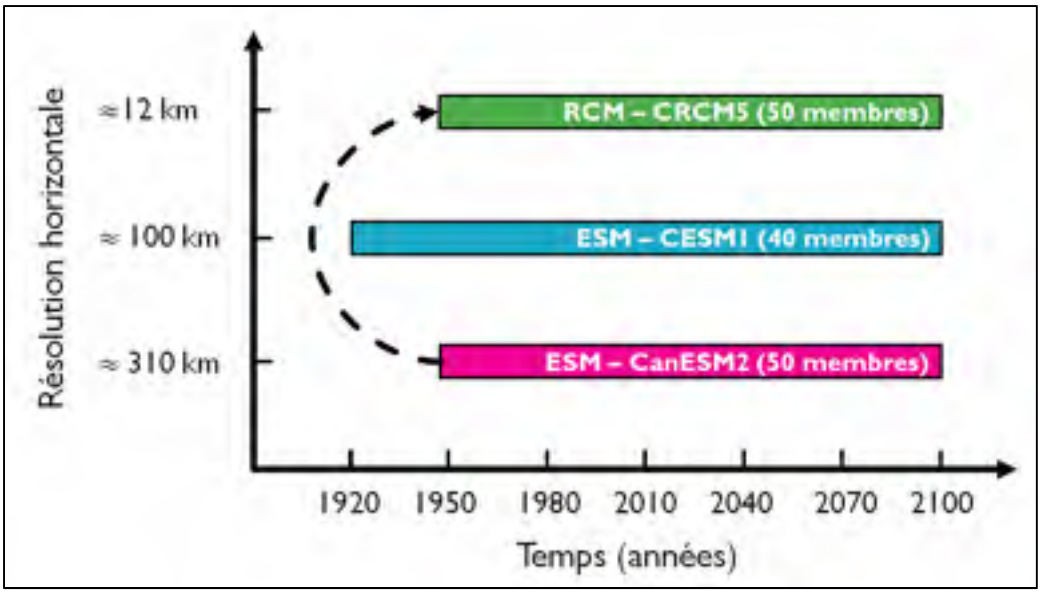

Figure 2.1 Résolution spatiale et couverture temporelle des trois grands ensembles de simulations climatiques disponibles pour les travaux de cette thèse

2.1.1.1 Grand ensemble global du CanESM2

Le grand ensemble de 50 membres du CanESM2 (Arora et al., 2011; Sigmond & Fyfe, 2016; von Salzen et al., 2013) du Centre canadien de la modélisation et de l'analyse climatique (en anglais : *Canadian Centre for Climate Modelling and Analysis*; CCCma) d'Environnement Canada est actuellement disponible depuis peu pour téléchargement sur le site web d'Environnement et Changement climatique Canada (ECCC). Cet ensemble couvre la période de 1950 à 2100 et est opéré à une résolution horizontale de 2.8° latitude et 2.8° longitude ou approximativement 310 km × 310 km à l'échelle globale. La Figure 2.2 explique la structure de ce grand ensemble. D'abord, il y a cinq simulations préindustrielles commençant en 1850 qui ont été arrêtées en 1950 permettant d'obtenir cinq états d'océan différents. À partir de ces cinq simulations, les conditions initiales de 1950 ont été perturbées aléatoirement 10 fois (générant 5 familles de 10 membres). Les 50 membres ont été simulés jusqu'en 2005 avec les forçages historiques, puis de 2006 à 2100 avec le scénario de forçage RCP8.5.

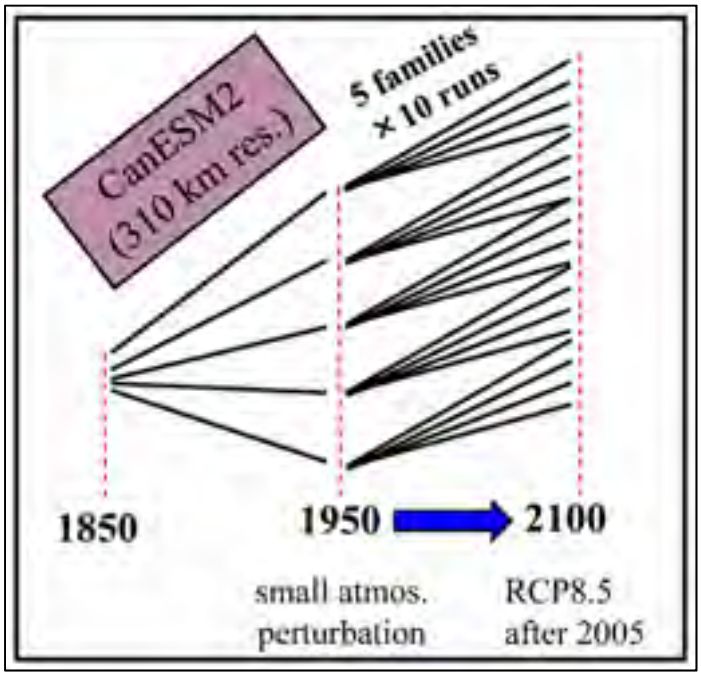

Figure 2.2 Structure du grand ensemble du CanESM2
Tirée de Leduc et al. (2016a)

2.1.1.2 Grand ensemble global du CESM1

Le grand ensemble de 40 membres du CESM1 (Kay et al., 2015) est produit par le *National Center of Atmospheric Research* (NCAR) et de l'*University Corporation for Atmospheric Research* (UCAR) et est disponible pour téléchargement sur le site web de NCAR. Une simulation à l'échelle globale utilisant les forçages préindustriels a été lancée de 1850 jusqu'en 1920, avant d'être perturbée 40 fois pour générer le grand ensemble. Tout comme l'ensemble de CanESM2, les forçages historiques sont utilisés de 1920 à 2005, puis le scénario d'émissions RCP8.5 de 2006 à 2100. La différence principale comparativement à CanESM2 est sa résolution spatiale qui est plus fine (1° × 1° ou ≈ 110 km × 110 km).

2.1.1.3 Grand ensemble régional du CRCM5

Le CRCM5 (Martynov et al., 2013; Šeparović et al., 2013) est développé par le centre ESCER de l'Université du Québec à Montréal (UQAM) en collaboration avec ECCC. Le grand ensemble du CRCM5 (Leduc et al., 2016a; Leduc et al., 2019) a été généré dans le cadre du projet ClimEx (*Climate change and hydrological Extremes*), dernière phase d'une collaboration de longue haleine entre le Québec et la Bavière. Il s'agit d'un ensemble de 50 membres avec une résolution spatiale de 0.11° × 0.11° (≈ 12 km × 12 km) couvrant deux domaines régionaux : le nord-est de l'Amérique du Nord et l'Europe (voir Figure 2.3). Les frontières des deux domaines ont été pilotées pour les 50 membres par les champs atmosphériques (à chaque 6 heures) et océaniques (journalier) sortant du grand ensemble du CanESM2 pour la période de 1950 à 2100. Comme les grands ensembles du CanESM2 et du CESM1, les forçages historiques sont employés de 1950 à 2005, puis le scénario d'émission du RCP8.5 de 2006 à 2100. Plus de détails sont disponibles sur le site web de ClimEx (http://www.climex-project.org/) et le grand ensemble sera aussi disponible pour téléchargement dans un futur rapproché. Une description détaillée et une validation de ce grand ensemble sont présentées dans l'article de l'Annexe I.

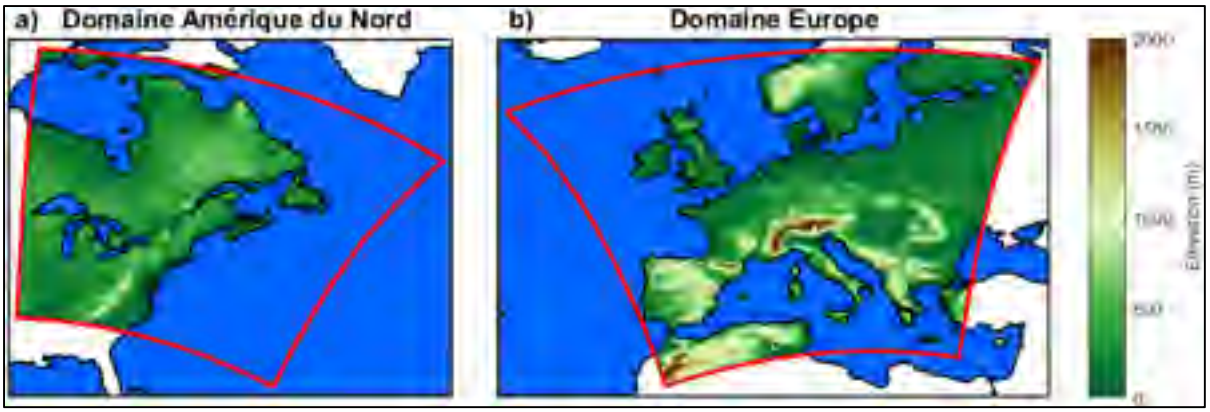

Figure 2.3 Modèle numérique d'altitude utilisé pour les domaines de a) l'Amérique du Nord et de b) l'Europe (délimité en rouge) couverts par le grand ensemble de simulations du CRCM5

2.1.2 La modélisation hydrologique

Une base de données de 5 797 bassins versants couvrant les États-Unis et le Canada est employée dans cette thèse pour étudier les débits en rivière. Les métadonnées et les débits des 5 265 bassins versants américains sont obtenus via l'*United States Geological Survery* (USGS). Malgré qu'une plus grande quantité de bassins versants était disponible, seulement ceux qui ont fourni une performance minimale (*NSE* > 0) avec le modèle hydrologique HSAMI (modèle global à 23 paramètres; Arsenault et Brissette (2014); Minville et al. (2008)) ont été conservés. Ce critère de sélection a vraisemblablement permis d'éliminer les bassins versants régulés, pour lesquels la modélisation hydrologique est plus complexe. Les données de précipitation et de température couvrant la période de 1950 à 2010 ont été extraites et spatialement moyennées à partir des données sur grilles de l'Université de Santa Clara (Maurer, Wood, Adam, Lettenmaier, & Nijssen, 2002). Concernant les 698 bassins versants canadiens, ceux-ci ont été extraits de la base de données CANOPEX (Arsenault, Bazile, Ouellet Dallaire, & Brissette, 2016a) et les données de précipitation et de température à partir des données sur grilles de Ressources naturelles Canada (RNCan) (Hutchinson et al., 2009). La base de données CANOPEX ne contient que des bassins versants sont non-régulés.

Puisqu'une grande fraction des séries d'observations des débits en rivière ne couvre pas la totalité de la période d'étude (1950-2010), un recours à la modélisation hydrologique est jugé comme nécessaire pour uniformiser les données. En utilisant les séries de précipitation et de température couvrant l'entièreté de la période, des séries de débits sont simulées et par la suite utilisées comme substituts aux observations. Considérant le nombre important de bassins versants (5 797), il est impraticable d'utiliser un modèle distribué dans ce contexte. En se basant sur la littérature de la section 1.5.1, un modèle global a été démontré comme étant une alternative permettant de répondre à cette problématique, tout en étant vraisemblablement aussi performant qu'un modèle distribué pour la simulation des débits à l'exutoire des bassins versants qui sont à l'étude. L'option d'un modèle hydrologique global a donc été sélectionnée pour la suite.

Le modèle hydrologique GR4J (modèle du Génie rural à 4 paramètres journalier; Perrin, Michel, & Andréassian, 2003), présenté à la Figure 2.4, couplé avec le modèle de neige CemaNeige (Valéry, 2010) et la formule d'évapotranspiration de Oudin (Oudin et al., 2005) a été choisi pour l'étape de la modélisation hydrologique. Une variante de cette structure à 9 paramètres suggérée par Poissant, Arsenault, et Brissette (2017) a été utilisée afin d'obtenir de meilleurs résultats sur les bassins versants ayant un couvert de neige important. Le choix de la structure complète du modèle hydrologique est fondé sur un certain nombre d'études ayant obtenu une bonne performance en utilisant cette structure (Troin, Arsenault, & Brissette, 2015a; Troin, Arsenault, Martel, & Brissette, 2018; Velázquez, Troin, Caya, & Brissette, 2015b) ainsi que sur des travaux présentés à l'Annexe VI.

Les résultats de l'Annexe VI ont permis de valider le choix de la structure du modèle hydrologique. Des cinq modèles de neige et des quatre formules d'évapotranspiration testées, CemaNeige et Oudin sont respectivement les choix optimaux pour une majorité de bassins versants nord-américains testés. En ce qui concerne le modèle hydrologique, trois modèles (GR4J ; Perrin et al. (2003), HMETS ; Martel, Demeester, Brissette, Poulin, et Arsenault (2017) et MOHYSE ; Fortin et Turcotte (2007)) ont été décomposés selon leur structure d'écoulement vertical et horizontal et interchangés pour déterminer la combinaison optimale.

La composante verticale de HMETS et horizontale de GR4J se sont trouvées à être la meilleure structure pour le modèle hydrologique. Toutefois, ces travaux n'étant pas encore publiés, il a été décidé d'opter pour une structure intacte du modèle hydrologique le plus performant des trois pour les autres travaux de cette thèse – le modèle hydrologique GR4J combiné avec le modèle de neige CemaNeige et la formule d'évapotranspiration d'Oudin.

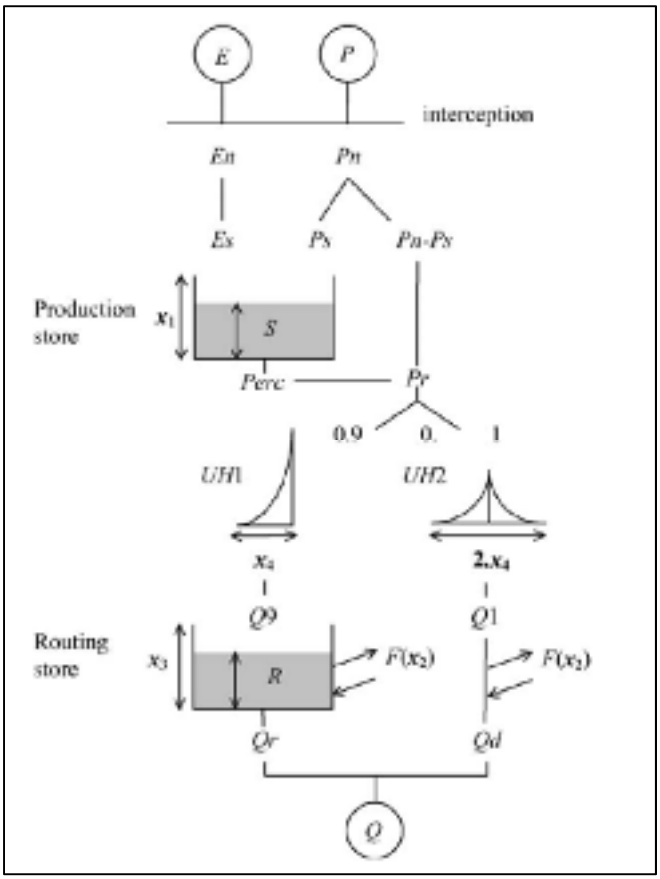

Figure 2.4 Structure du modèle hydrologique global
GR4J (Génie rural à 4 paramètres journalier)
Tirée de Perrin et al. (2003)

Pour la calibration du modèle hydrologique, la fonction objectif du *KGE* (équation 1.11) a été considérée comme étant la meilleure option basée sur les recommandations de la littérature (Pokhrel & Gupta, 2011). Un budget de 10 000 évaluations et l'algorithme d'optimisation du SCE-UA ont pour leur part été sélectionnés à partir des résultats des travaux de Arsenault et al. (2014).

La calibration a été réalisée sur chacun des 5 797 bassins versants de la base de données en utilisant l'entièreté des débits observés disponibles pour la période de 1950 à 2010. Les méthodes traditionnelles de calibration/validation décrites à la section 1.5.2 employée pour évaluer la robustesse du jeu de paramètres obtenus n'ont pas été utilisées suite aux résultats obtenus des travaux présentés dans l'Annexe III. Il a été démontré et recommandé dans cet article qu'il est plus optimal de réaliser la calibration sur la totalité des données disponibles, permettant d'avoir un jeu de paramètre plus robuste et couvrant un maximum d'information. Les autres travaux de cette thèse ont suivi ces recommandations. En revanche, cette méthode ne prend pas en considération la possibilité que le modèle puisse être calé avec des processus hydrologiques dominants différents d'un bassin versant à un autre. Par exemple, différentes phases de variabilité décennale ou multi-décennale pourraient être prises en considération lors de calibration si les observations sont disponibles sur différentes périodes.

Un premier triage a été réalisé parmi les 5 797 bassins versants disponibles afin de conserver les bassins avec une performance considérée adéquate. Seuls les bassins avec au moins 5 années complètes (c.-à-d. : une année complète est définie comme étant une année avec moins de 10% de données manquantes) de débits observés et un *KGE* supérieur à 0.4 ont été gardés. Ceci a donné lieu à une sélection de 4 536 bassins versants au total, illustrée à la Figure 2.5.

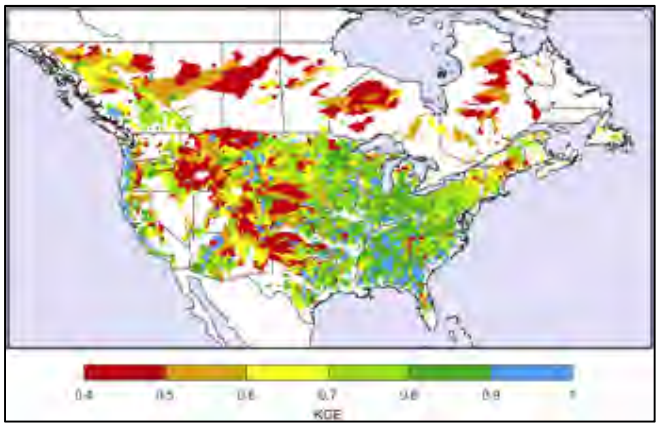

Figure 2.5 Carte des 4 536 bassins versants nord-américains sélectionnés lors du premier triage et la valeur du *KGE* obtenue lors de leur calibration

Un deuxième triage plus restrictif a été réalisé pour sélectionner les bassins versants les plus performants. Pour ce triage, seuls les bassins avec au moins 10 années complètes de débits observés et un *KGE* supérieur à 0.5 ont été gardés. En tout, 3 631 bassins versants ont été sélectionnés de cette façon.

2.1.3 Évaluation de l'influence de la variabilité naturelle

2.1.3.1 Compréhension de la variabilité hydroclimatique

Dans un premier lieu, les indices climatiques introduits dans la section 1.2.2 sont examinés afin d'améliorer la compréhension de la variabilité naturelle des différentes variables hydroclimatiques (précipitation, température et débits en rivière). Tel que discuté précédemment, il y a un manque à combler dans la littérature en termes de reproductibilité et de généralisation des résultats ayant été publiés jusqu'à présent. L'objectif ici est d'utiliser une méthode commune à l'échelle du bassin versant, tout en couvrant l'Amérique du Nord dans son ensemble, afin d'obtenir une portait plus détaillé pour améliorer la compréhension de la variabilité naturelle.

Les six indices climatiques reconnus pour leur influence à l'échelle nord-américaine présentés à la section 1.2.2 ont été sélectionnés. Afin de comprendre la relation entre ces indices climatiques et la variabilité hydroclimatique observée, une méthodologie basée sur le pourcentage de variance expliquée a été choisie. Ainsi, le pourcentage de la variance des différentes variables hydroclimatiques (température, précipitation et débits en rivière) à l'échelle saisonnière est évalué pour chacun des six indices climatiques et sur le premier ensemble des 4 536 bassins versants retenus. Le pourcentage de variance expliquée combinant les six indices climatiques a aussi été évalué pour comprendre l'influence du couplage des différents indices climatiques. Des travaux supplémentaires ont également été réalisés sur le couplage des différents indices climatiques et sont présentés à l'Annexe V et sont discutés brièvement au chapitre 7.

2.1.3.2 Détection du signal des changements climatiques

Pour mieux expliquer l'influence potentielle de la variabilité naturelle, les 50 représentations équiprobables d'un climat futur issues du grand ensemble du CanESM2 sont utilisées pour un point de grille spécifique. La Figure 2.6 présente le point de grille correspondant à la ville de Montréal pour la température moyenne annuelle (TAS – Figure 2.6-haut) et les maxima annuels de précipitations journalières (RX1day – Figure 2.6-bas) couvrant la période de 1950 à 2100. Il est important de noter que le scénario RCP8.5 d'émission de GES est utilisé dans ce grand ensemble, correspondant à un scénario pessimiste. En ce qui concerne les TAS, il est très évident qu'à ce jour (en 2018), le signal du changement climatique (tendance dans la courbe bleue) ait été détecté de façon significative à travers la variabilité naturelle (l'ensemble des courbes grises). En regardant les valeurs de RX1day, une augmentation de la moyenne et également de la variance sont perceptibles au fur et à mesure que la série temporelle progresse vers la fin du 21^{e} siècle. De plus, le signal du changement climatique se trouve à être beaucoup moins fort et la variabilité naturelle est plus importante qu'elle ne l'était pour les TAS. Ainsi, il peut être difficile de détecter une tendance significative en date d'aujourd'hui, et même encore dans plusieurs décennies. Le rôle de la variabilité naturelle dans ce contexte est examiné dans cette thèse.

Une meilleure compréhension du rôle de la variabilité naturelle sur la détection du signal des changements climatiques des précipitations moyennes et extrêmes est réalisée ici à l'aide des grands ensembles de simulations climatiques décrits précédemment. Afin de couvrir la planète au complet, les grands ensembles du CanESM2 et du CESM1 sont utilisés. Pour y arriver, une métrique définie comme la décennie pour laquelle une tendance significative du même signe a été détectée pour 90% des membres (90%DD) est employée. La significativité de la tendance est évaluée à l'échelle locale avec le test non paramétrique de Mann-Kendall et à l'échelle régionale un rééchantillonnage par bootstrap (section 1.3.1). La métrique du 90%DD se rapproche des métriques de « temps d'émergences » typiquement utilisés dans d'autres études d'impacts des changements climatiques (Giorgi & Bi, 2009; Hawkins & Sutton, 2012; IPCC, 2013; King et al., 2015; Maraun, 2013b). De plus, cette procédure est

répétée à l'échelle locale et régionale afin d'investiguer si la variabilité naturelle joue un rôle plus important localement que régionalement. Cette méthodologie est schématisée à la Figure 2.7 pour le point de grille correspondant à la ville de Montréal.

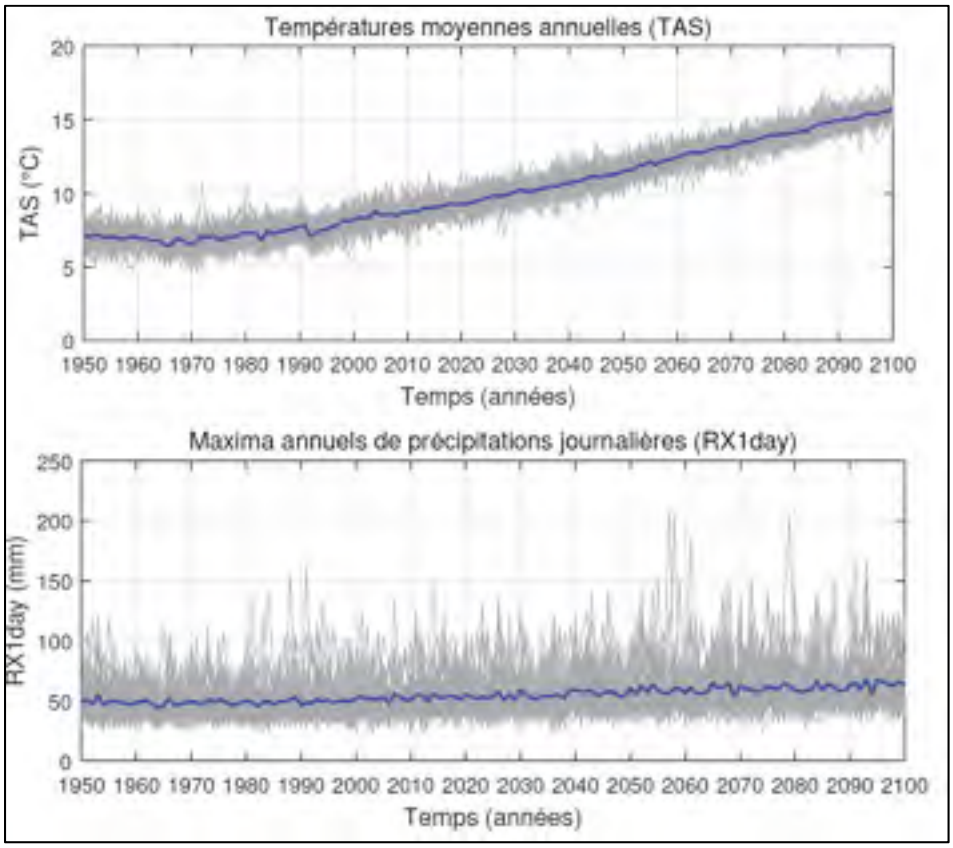

Figure 2.6 Séries temporelles pour le point de grille correspondant à la ville de Montréal de 50 réalisations équiprobables d'un climat futur de l'ensemble CanESM2 couvrant la période de 1950 à 2100 pour les températures moyennes annuelles (TAS – panneau du haut) et les maximas annuels de précipitations journalières (RX1day – panneau du bas). Les courbes grises représentent l'ensemble des 50 séries temporelles, la courbe bleue la moyenne d'ensemble

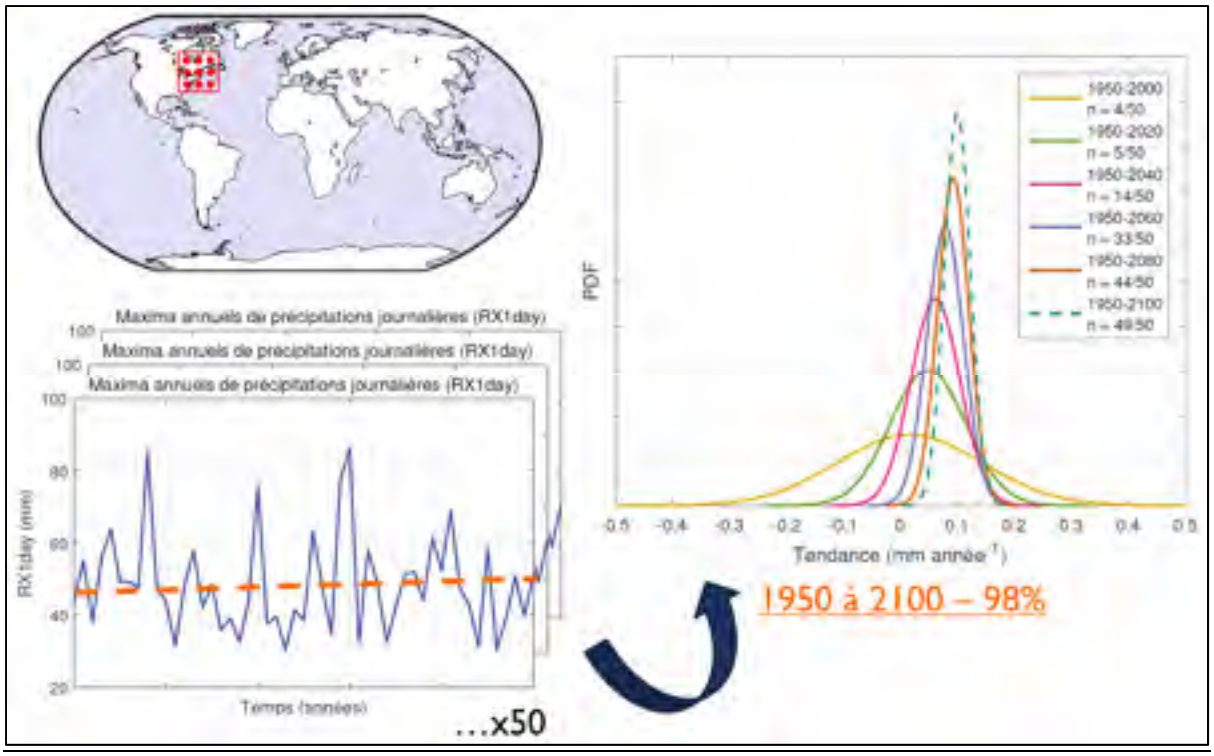

Figure 2.7. Résumé de la méthodologie permettant d'obtenir la métrique du 90%DD sur le point de grille du grand ensemble de CanESM2 correspondant à la ville de Montréal

2.1.4 Évaluation de l'influence des changements climatiques

2.1.4.1 Influence sur les précipitations extrêmes

Un autre avantage d'avoir recouru à de grands ensembles de simulations climatiques est également illustré dans la Figure 2.6 en ce qui concerne l'étude des événements extrêmes. En effet, tel que montré pour l'indice RX1day, des extrêmes moins fréquents sont générés dans un tel grand ensemble (les pointes des courbes grises les plus hautes). De cette façon, il est possible d'étudier les périodes de retour beaucoup plus grandes, tout en minimisant l'incertitude statistique typiquement liée à l'extrapolation de valeurs pour ces périodes de retour à partir d'une série d'observation trop courte (Schulz & Bernhardt, 2016).

Les trois grands ensembles décrits à la section 2.1.1 sont donc utilisés pour étudier l'évolution des périodes de retour des précipitations extrêmes. Différentes échelles

temporelles sont étudiées (1 jour et 5 jours pour les deux ESMs et de 1 heure à 5 jours pour le RCM). De plus, en utilisant le grand ensemble du CRCM5, il est possible d'étudier les deux domaines du nord-est de l'Amérique du Nord et l'Europe avec une résolution spatiale plus fine que celle des ESMs. Le survol méthodologique qui suit est reproduit sur la totalité des points de grille des trois grands ensembles pour chaque durée d'événements considérée.

D'abord, les séries des maxima annuels (SMA) de précipitation sont extraites d'un point de grille pour les périodes de référence (1980-1999) et future (2080-2099). Basés sur l'hypothèse de stationnarité sur les deux périodes de 20 années, les SMA des différents membres sont ensuite combinés, générant de nouvelles SMA de 1 000 années pour les ensembles du CanESM2 et du CRCM5 (20 années × 50 membres) et 800 années pour celui du CESM1 (20 années × 40 membres). Les SMA combinées sont alors classées en ordre croissant et les quantiles empiriques des périodes de retour d'intérêt (p. ex. : 2 ans, 20 ans ou même 100 ans) sont extraits de la SMA de référence. Finalement, les périodes de retour futures correspondant aux valeurs associées aux périodes de retour de référence sont déterminées. Un résumé de la méthodologie est illustré à l'aide de la Figure 2.8.

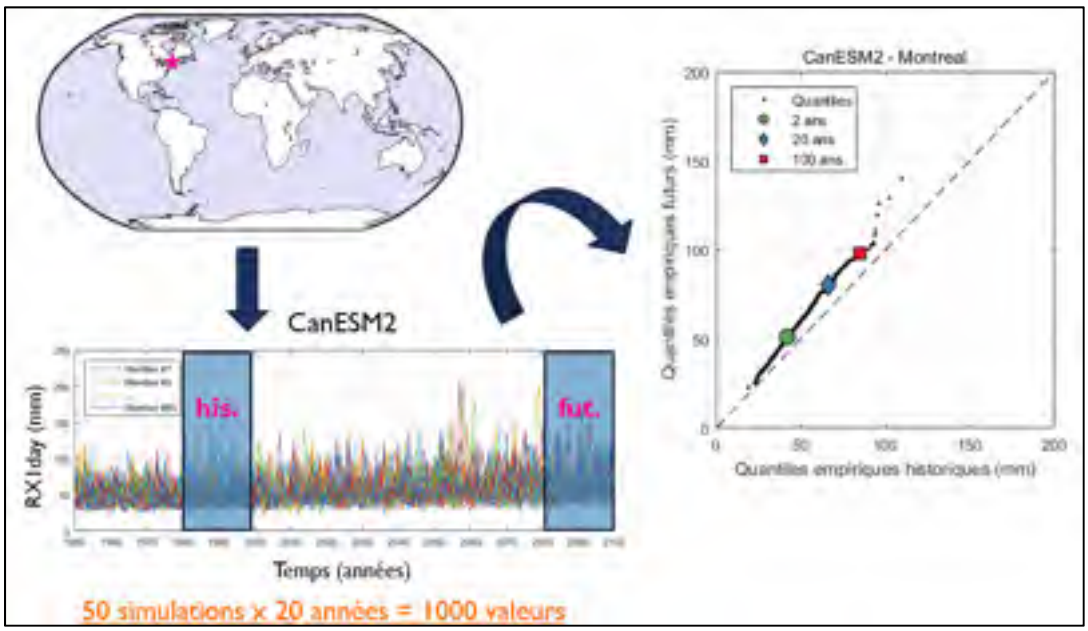

Figure 2.8 Résumé de la méthodologie utilisée basée sur le point de grille du grand ensemble de CanESM2 correspondant à la ville de Montréal

Tel que discuté à la section 1.5.3, il y a plusieurs limitations au niveau de la correction des biais des données sortantes des modèles climatiques, notamment au niveau des valeurs extrêmes. Pour cette raison, il a été considéré comme étant une meilleure alternative d'évaluer directement les données brutes sortantes des modèles climatiques, et de plutôt en analyser les changements relatifs au sein des modèles. En évitant de faire un post-traitement des sorties de modèle, une source d'incertitude additionnelle est évitée, rendant l'interprétation des résultats plus simple.

2.1.4.2 Influence sur les débits en rivière extrêmes

Dans un deuxième temps, les deux grands ensembles issus des deux ESMs (c.-à-d. : CanESM2 et CESM1) sont utilisés pour étudier l'évolution de la période de retour 100 ans des débits en rivière. La raison pour laquelle le grand ensemble du CRCM5 a été laissé de côté pour cette étude est que le domaine du nord-est de l'Amérique du Nord est trop petit, réduisant significativement le nombre de bassins versants pouvant être analysés. D'un autre côté, un avantage du grand ensemble de CRCM5 est qu'il a une résolution spatio-temporelle plus fine, donnant lieu à un plus grand réalisme au niveau des extrêmes de précipitation. Cependant, les bassins versants étudiés ayant une superficie élevée (> 500 km^2), les extrêmes des débits en rivière se retrouvent à être générés par une combinaison d'éléments relativement bien simulés à l'échelle des ESMs employés.

Tel que discuté précédemment à la section 1.5.3, il y a un certain nombre de limitations non négligeable directement en lien avec l'utilisation d'une méthode de correction de biais. Cependant, lorsque vient le temps d'utiliser un modèle hydrologique alimenté par les données sortantes d'un modèle climatique, la correction des biais devient incontournable. En effet, puisque les débits agissent comme un intégrateur des différentes variables hydroclimatiques (p. ex. : précipitation et température), un biais dans une variable en particulier peut avoir un grand impact sur la modélisation du cycle de l'eau. Par exemple, un important biais positif de température sur un bassin versant dominé par la neige viendra générer une moins grande quantité de neige et une fonte beaucoup plus rapide que la réalité. Ainsi, la crue printanière,

générant habituellement le maximum annuel du débit dans la rivière, ne produira pas la réponse hydrologique attendue. En corrigeant les biais en fonction des données observées, ceci permet de fournir une réponse hydrologique plus réaliste, et d'en évaluer par la suite les changements dans un climat futur.

La méthode de correction de biais employée ici est celle de la correction des biais journaliers *(*DBC; *Daily Bias Correction*), basée sur une approche par distribution des quantiles (Chen et al., 2013). D'abord, une mise à l'échelle de l'intensité locale (LOCI; *Local Intensity Scaling*) est employée afin de corriger la probabilité d'occurrence d'un événement de précipitation (Schmidli et al., 2006). Ensuite, la méthode de translation journalière (DT; *Daily Translation*) est utilisée pour corriger la distribution de fréquence pour la précipitation et la température (Mpelasoka & Chiew, 2009). Une fois les variables hydroclimatiques post-traitées avec la méthode DBC, la modélisation hydrologique est réalisée sur la sélection de 3 631 bassins versants décrite précédemment avec la structure du modèle GR4J/CemaNeige/Oudin décrit à la section 2.1.2.

Pour la suite, la méthodologie employée se trouve à être très similaire à celle décrite ci-haut à l'aide de la Figure 2.7. Cependant, au lieu d'utiliser les précipitations à différentes échelles temporelles, les séries de débits journaliers obtenues suite à la modélisation hydrologique sont employées.

2.2 Organisation du document

Le survol méthodologique présenté dans ce chapitre a servi d'introduction pour le restant de la thèse dont l'organisation se fait comme suit :

Le chapitre 3 introduit le premier article intitulé « *Impacts of the dominant modes of natural climate variability on hydroclimatic variability over North American catchments* » et soumis au journal « *Water Resources Research* ». Cet article a pour objectif d'évaluer la relation entre les grands indices climatiques, servant de mesure de la variabilité naturelle, et la

variabilité hydroclimatique à l'échelle des bassins versants de l'Amérique du Nord. Les résultats permettent de clarifier si une partie de la variabilité naturelle des extrêmes hydrométéorologiques, pouvant entraver le signal des changements climatiques, peut être expliquée à l'aide des indices climatiques.

Le chapitre 4 présente le deuxième article intitulé : « *Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales* » et publié dans la revue « *Journal of Climate* ». L'objectif principal de cet article est de déterminer comment la variabilité naturelle peut contrevenir à la détection du signal des changements climatiques des événements de précipitation extrêmes. Formant l'une des pierres angulaires de cette thèse, les résultats de ces travaux aident à mieux comprendre et interpréter une absence potentielle de tendance liée aux changements climatiques.

Le chapitre 5 présente le troisième article qui lui est intitulé : « *Global and regional projected changes in 100-year sub-daily, daily and multi-day precipitation extremes estimated from three large ensembles of climate simulations* » et soumis à la revue « *Journal of Climate* ». Dans ces travaux, l'évolution des événements extrêmes de précipitations est évaluée dans un contexte de changements climatiques limitant l'incertitude statistique. Les conclusions découlant de ces résultats permettent de mieux comprendre l'ampleur potentielle de l'impact des changements climatiques sur les extrêmes de précipitations et à quel point l'hypothèse de stationnarité des séries d'observations doit être revisitée.

Le chapitre 6 présente le quatrième et dernier article de cette thèse : « *Impacts of climate change on daily extreme streamflow estimated from two large climate simulation ensembles over 3 567 North American catchments* » qui sera soumis à la revue « *Nature Climate Change* ». Cet article introduit une méthode similaire à celle utilisée dans le troisième article, mais pour des bassins versants de grande superficie couvrant l'Amérique du Nord. Les résultats permettent de mieux comprendre l'impact des changements climatiques sur les crues extrêmes.

En annexes, quatre articles qui ont été rédigés en collaboration au courant du doctorat, ainsi que d'autres résultats sont présentés :

1. Un premier article intitulé : *« 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)* » et accepté pour publication avec révisions majeures dans la revue *« Journal of Applied Meteorology and Climatology »*. Ce premier article en collaboration introduit et présente une validation du grand ensemble de simulations climatiques du CRCM5 à l'échelle régionale du nord-est de l'Amérique du Nord et de l'Europe utilisé dans l'article du chapitre 5.
2. Un deuxième article intitulé : *« Uncertainty of hydrological model components in climate change studies over two Nordic Quebec catchments »* et publié dans la revue « *Journal of Hydrometeorology* ». Cet article présente une étude sur l'incertitude des composantes de la modélisation hydrologiques, notamment en comparaison avec l'incertitude associée à la variabilité naturelle.
3. Un troisième article intitulé : *« The hazards of split-sample validation in hydrological model calibration »* et publié dans la revue *« Journal of Hydrology »*. Cet article a permis d'établir la stratégie optimale de calibration d'un modèle hydrologique en démontrant qu'une validation traditionnelle diminue la performance de la calibration. Les conclusions de ces travaux ont été appliquées dans les articles des chapitres 3 et 6, ayant recours à la modélisation hydrologique.
4. Un quatrième article intitulé : *« Relative importance of internal climate variability versus anthropogenic climate change in global climate change* » et soumis à la revue « *Journal of Climate* ». Cet article a aussi investigué l'influence de la variabilité naturelle sur la détection des changements climatiques. Malgré que la méthodologie employée dans cet article se trouve à être différente de celle de l'article du chapitre 4, des résultats similaires ont été obtenus, renforçant les conclusions présentées dans la thèse.
5. Les résultats d'un article en préparation intitulé : *« Individual and coupled effects of the dominant modes of natural climate variability on seasonal precipitation and temperature over North America »*. Ces travaux sont la suite de l'article présenté au chapitre 3, ayant

pour but de démontrer l'importance du couplage des différents indices climatiques pour expliquer la variabilité hydroclimatique en fonction de leur phase négative et positive.

6. Les résultats d'un article en préparation intitulé : « *Uncertainty of hydrological model components over North American watersheds* ». Ces travaux sont la suite de l'article présenté à l'Annexe II et ont pour but de mieux comprendre l'incertitude associée aux différentes composantes de la structure du modèle hydrologique. Ces travaux ont aussi contribué au choix du modèle hydrologique utilisé dans les articles des chapitres 3 et 6 pour la simulation des débits en rivière.
7. Les résultats d'une présentation orale faite à *«* l'*European Geosciences Union – General Assembly »* en 2016. Ces résultats investiguent l'influence de l'incertitude statistique dans l'estimation des crues 1 dans 20 ans dans un contexte de changement climatique. Ceux-ci n'ont actuellement pas abouti sur la rédaction d'un article scientifique, mais apportent une certaine contribution à la discussion générale et aux recommandations de cette thèse.
8. Finalement, la huitième et dernière annexe présente une liste des articles publiés, soumis et en préparation, des présentations orales et par affiches, comme premier auteur ou en collaboration. L'ensemble de ces publications regroupent donc les travaux qui ont découlé de cette thèse.

CHAPITRE 3

IMPACTS OF THE DOMINANT MODES OF NATURAL CLIMATE VARIABILITY ON HYDROCLIMATIC VARIABILITY OVER NORTH AMERICAN CATCHMENTS

Jean-Luc Martel[a], Magali Troin[b], François Brissette[c]

[a, b, c] Département de Génie de la construction, École de technologie supérieure

Article soumis à la revue « Water Resources Research », avril 2018

Key Points

1. The impact of six climate indices on five hydroclimatic variables is analyzed for 4536 North American catchments over the 1950-2010 period
2. Temperature variability over the study catchments is influenced by some climate indices, with strong seasonal and regional variations.
3. Seasonal variability in precipitation and streamflows are moderately correlated to the climate indices for the investigated catchments

Abstract

The impacts of the dominant North American modes of natural climate variability (NCV) on observed hydroclimatic variability are investigated for 4 536 catchments. The NCV is explored on the basis of six climate indices from interannual − El Niño-Southern Oscillation (ENSO) and Pacific North American pattern (PNA), decadal − Arctic Oscillation (AO) and North Atlantic Oscillation (NAO), to interdecadal − Atlantic Multidecadal Oscillation (AMO) and Pacific Decadal Oscillation (PDO) − time scales, by considering individual and all-combined impacts of these indices on five hydroclimatic variables over the 1950-2010 period. While correlations with some individual indices are statistically significant for several regions, their effects on the hydroclimatic variables are stronger and more spatially coherent for winter than summer. The climate indices explain a larger part of the seasonal temperature variance compared to the seasonal precipitation and streamflow variances. The PNA tends to be associated with winter temperature variability over western Canada, while the PDO

combined to AO and NAO explains much of the winter temperature variance over the southeastern United States. The analysis underlines the difficulty to identify a single index as a predictor for the ensemble of hydroclimatic variables across North America, because of the large regional effects. By considering the all-combined impacts of the indices on seasonal hydroclimatic variables, the relationships are intensified both in amplitude and spatial extend for a large number of catchments. Of the assessment of the linkages between extreme hydroclimatic variables and climate indices, some patterns vaguely emerge when the combination of all indices are considered.

3.1 Introduction

There is a consensus among scientists that human activities, mainly through greenhouse-gas emissions, are likely the dominant cause of the observed global warming since the middle of the 20th century (e.g., Cook et al., 2013; Cook et al., 2016; Qin et al., 2014). However, superimposed on this long-term change in the mean state or variability of the climate properties (commonly referred to as climate change) is the natural variability of climate. The natural climate variability (NCV) is a natural consequence to external forcings or to the non-linear internal interactions between the components of the climate system, such as the atmosphere and the hydrosphere (Baede et al., 2001). Because these climate components have different response times to these interactions, they never reach equilibrium. It is expected that the NCV, notably through the atmosphere-ocean interactions, will delay (or accelerate) the detection of the climate change signal on the hydrologic cycle at the local and regional scale (Barnett et al., 1999; Deser et al., 2012a; Fischer & Knutti, 2014; Fischer et al., 2014; Hegerl et al., 2015; Hegerl et al., 1996; Martel, Mailhot, Brissette, & Caya, 2018; National Research Council, 1998; Stott et al., 2006; Swanson, Sugihara, & Tsonis, 2009). To provide reliable future projections of the hydrologic cycle, an estimation of the contribution from the NCV on the long-term changes on key hydroclimatic variables is thus required.

NCV is generally explored according to fluctuations in the mean state of the atmosphere. One manner to describe fluctuations is by using climate indices, which represent the

variations of the mean state of the atmosphere associated, or not, with an oceanic coupling (Rossi et al., 2011). It became apparent that climate fluctuations are not randomly distributed in time and space, but tend towards relatively coherent spatial patterns through time (National Research Council, 1998). Thus, consistent regional spatio-temporal patterns that identify the modes of NCV emerge. The indices are an illustration of the various oceanic-atmospheric modes of variability from interannual to multi-decadal time scales that can influence regional hydroclimatic variations. The most thoroughly investigated indice that appears to affect interannual hydroclimatic variability and to be related to the frequency of extreme events, such as hurricanes and floods in many regions of the world, is associated with the El Niño-Southern Oscillation (ENSO) phenomenon (Bell & Chelliah, 2006; Cayan et al., 1999; Goldenberg, Landsea, Mestas-Nuñez, & Gray, 2001; Ward, Beets, Bouwer, Aerts, & Renssen, 2010; Ward, Eisner, Flörke, Dettinger, & Kummu, 2014a; Ward et al., 2014b; Ward, Kummu, & Lall, 2016; among others). Other indices also influence the regional climate and hydrological variability. For instance, Massei et al. (2010) showed that it was possible to associate the mode of variability of 17 years observed in the Seine River flow (France) with some components of the North-Atlantic Oscillation (NAO); while Hidalgo et Dracup (2003) present the evidence of the Pacific Decadal Oscillation (PDO) signature on the long-term hydroclimatic variations of the Upper Colorado River Basin (United States). Because of the attributes of the indices, it is hoped that they yield benefits for decadal hydroclimatic predictability (Chikamoto, Timmermann, Stevenson, DiNezio, & Langford, 2015; Mehta, Wang, Mendoza, & Rosenberg, 2014; Wanders & Wada, 2015) similar to those obtained for seasonal-to-interannual predictions through ENSO (Chandimala & Zubair, 2007; Córdoba-Machado, Palomino-Lemus, Gámiz-Fortis, Castro-Díez, & Esteban-Parra, 2016; Souza Filho & Lall, 2003). Climate indices thus provide a suitable avenue by which the search for a predictable hydroclimatic signal should be pursued.

Over North America (NA), the major climate indices affecting hydroclimatic variability include El Niño-Southern Oscillation (ENSO; Cayan et al., 1999), the Pacific North American pattern (PNA; Rogers & Coleman, 2003), the North Atlantic Oscillation (NAO; Barlow et al., 2001), which is closely related to the Arctic Oscillation (AO; Déry & Wood,

2004), the Pacific Decadal Oscillation (PDO; Stewart et al., 2005), and the Atlantic Multi-Decadal Oscillation (AMO; Enfield et al., 2001). These indices involve different temporal variability, from interannual (ENSO and PNA) to decadal (NAO and AO) towards interdecadal (PDO and AMO) time scales, which can affect NA hydrosystems in various manners (Rossi et al., 2011).

In recent years, many studies have explored the relationships of some of these indices with hydroclimatic variables for several NA regions (Bonsal & Shabbar, 2008; Déry & Wood, 2005; Fleming et al., 2007; Fu et al., 2012; Gobena & Gan, 2006; Nalley et al., 2016; Rood et al., 2005; Tootle et al., 2005; Wang et al., 2006; Whan & Zwiers, 2017; Woo & Thorne, 2008) and specific catchments in Canada (Anctil & Coulibaly, 2004; Brabets & Walvoord, 2009; Burn, 2008; Gobena & Gan, 2009; Gobena et al., 2013; Kiffney et al., 2002; Peters et al., 2013; Spence, 2002; Thorne & Woo, 2011) and in the United States (US) (Hamlet & Lettenmaier, 1999; Hidalgo & Dracup, 2003; Massei et al., 2011; Maurer et al., 2006; Rogers & Coleman, 2003; Twine et al., 2005). A range of hydrological variables were used, including annual streamflows (Assani, Landais, Mesfioui, & Matteau, 2010; Coulibaly & Burn, 2004; Déry, Hernández-Henríquez, Owens, Parkes, & Petticrew, 2012; McCabe, 1995), seasonal streamflows (Barlow et al., 2001; Coulibaly & Burn, 2005; Schmidt, Lipp, Rose, & Luther, 2001; Stewart et al., 2005), as well as low flow and high flow extremes (Andrews, Antweiler, Neiman, & Ralph, 2004; Biron, Assani, Frenette, & Massicotte, 2014; Cayan et al., 1999; Khaliq, Ouarda, Gachon, & Sushama, 2008). Strong correlations between the climate indices and streamflows have identified notable changes in the long-term hydroclimatic behavior, such as those observed in the 1970s over the Mid-Atlantic region of the US and the Canadian-wide region (Coulibaly & Burn, 2004; Schulte, Najjar, & Li, 2016).

Various combinations of a large number of statistical techniques have been applied for exploring the relationships between the climate indices and hydroclimatic variables. The trends in hydroclimatic time series are usually detected and characterized by using parametric and/or non-parametric statistical techniques. This includes, for instance, linear regression analysis, Bayesian and non-Bayesian change-detection algorithms, signal-to-noise ratios,

principal component analysis, as well as various forms of low pass filtering and spectral analysis, such as Fourier transform and wavelet transform techniques (Barlow et al., 2001; Detzel & Mine, 2014; Fleming & Weber, 2012; Fu et al., 2012; Hidalgo & Dracup, 2003; Stewart et al., 2005). The use of the wavelet transform techniques to explore trends in hydroclimatic time series is relatively recent; many works have demonstrated their useful applications in the analysis and extraction of non-stationary characteristics of hydroclimatic variables (e.g., Coulibaly & Burn, 2004; Gobena & Gan, 2009; Massei et al., 2011; Nalley et al., 2016; Nourani, Hosseini Baghanam, Adamowski, & Kisi, 2014; Rathinasamy et al., 2014). The significance of trends in hydroclimatic time series is evaluated based on statistical tests, such as the Mann-Kendall test or variance analysis (Assani, Landry, & Laurencelle, 2012; Burn & Hag Elnur, 2002; Déry et al., 2012). The relationships between the climate indices and hydroclimatic variables are then analyzed through quantitative comparisons of trends – similarities between trends and patterns detected in climate indices such as ENSO and hydroclimatic variables related to precipitation and streamflows. For example, this is commonly made using the one-step or two-step correlation methods (Biron et al., 2014; Khaliq et al., 2008). The significance of the coefficients of correlation can also be tested by the way of resampling (i.e., Monte-Carlo, Chi-squared and bootstrapping tests, permutation procedure) methods (Biron et al., 2014; Burn & Hag Elnur, 2002; Enfield et al., 2001; Khaliq et al., 2008). Some recent studies have inter-compared several of the previous mentioned statistical methods and have provided a critical analysis associated with the detection of trends in hydroclimatic time series over catchments (e.g., Detzel & Mine, 2014; Fleming & Weber, 2012; Khaliq et al., 2008).

The literature is replete with descriptions of the major climate indices affecting NA hydroclimatic variability; however, it is challenging to find a synthesis of the impacts of the ensemble of these indices on hydroclimatic variability over all NA catchments. Most studies conducted to far only focus on small regions or samples of catchments, which greatly differ from a study to another, constraining direct inter-comparison amongst studies. Although some general patterns emerge, with trends varying by catchments, the effects of the climate indices on the seasonal and annual hydroclimatic variability are not clear at the regional and

continental scales of NA. Moreover, few works explore the relationships of these indices with variability both in climate (precipitation and temperature) and streamflows. However, these relationships are required to understand the behavior of hydrologic trends and to provide a comprehensive response of a catchment to the NCV. Besides the selection of the study area and hydroclimatic variables, the choice of the climate indices and definition of the corresponding time intervals is appreciated through different ways according to the methodology employed. The combinations of statistical techniques used for the analyses also vary significantly between studies, rendering difficult the extrapolation of conclusions in the regional and continental NA.

The aim of the research presented here is to improve the understanding of how the dominant NA climate indices, both individually and all-combined, influence the hydroclimatic variability of the catchments. This work will tackle the limitations from previous studies, by using a large database of NA catchments and by considering climatic (precipitation and temperature) and hydrologic (streamflow) variables at the seasonal scales. The potential relationships between hydroclimatic variables and individual or all-combined climate indices (AMO, PDO, AO, NAO, PNA and ENSO) are investigated using a combination of statistical techniques. To attain the research aim, it was assessed if a single index or a combination of indices emerges for explaining the observed NA hydroclimatic variability over the 1950-2010 period. In the following, Section 2 presents the experimental design of the study. Section 3 discusses the relevant results of the impacts of the climate indices on hydroclimatic variability over the NA catchments. Concluding comments are provided in Section 4.

3.2 Experimental design

To explore the relationships between climate indices and observed hydroclimatic variability over the NA catchments, two types of data sets are used: three hydroclimatic variables (temperature, precipitation and streamflow) and oceanic-atmospheric data for the Atlantic and Pacific Oceans, corresponding to the six major climate indices (AMO, PDO, AO, NAO, PNA and ENSO) affecting NA hydroclimate variability.

3.2.1 Study area and hydroclimatic datasets

The United States Geological Survey (USGS) database is used to extract the metadata (boundary and drainage area) of 5 265 catchments over the contiguous US, as well as daily streamflow data over the 1950-2010 period. Daily precipitation and temperature data are extracted from the University of Santa-Clara gridded dataset (Maurer et al., 2002) over each catchment for the 1950-2010 period. This dataset is based on the National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer meteorological stations, with an average density of one station per 700 km^2. The network of meteorological stations is interpolated on a 0.125° × 0.125° (≈ 14 km) grid using the Shepard's synergraphic mapping system algorithm (Shepard, 1984), as implemented by Widmann et Bretherton (2000). An arithmetic average using all available grid points within the catchment's boundaries is then performed to obtain average daily precipitation and temperature data for each catchment. Because of the fine and constant grid size of the University of Santa Clara gridded dataset, the use of a more complex method (e.g., Thiessen's polygons) is not considered necessary.

Similarly, the Canadian Model Parameter Experiment (CANOPEX) database (Arsenault et al., 2016a) is used to extract metadata (boundary and drainage area) and daily streamflow data for 698 catchments in Canada over the 1950-2010 period. Average daily precipitation and temperature data are obtained using the same methodology as that employed for the US catchments, but by using the Natural Resource Canada (NRCAN) gridded dataset (Hutchinson et al., 2009). This dataset is based on the Environment Canada weather stations, where the number of stations varies roughly between 2000 and 3000 for precipitation data and between 1500 and 3000 for temperature data over the 1950-2010 period. The meteorological network is interpolated on a 0.083° × 0.083° (≈ 9 km) grid using a thin plate-smooting splines (ANUSPLIN) algorithm (Hutchinson, 1995; Hutchinson & Xu, 2004).

While the University of Santa-Clara and NRCAN gridded datasets provide continuous daily time series of precipitation and temperature data over the 1950-2010 period, the streamflow observational records are incomplete for many of the selected catchments. A first screening is

done by removing all catchments with less than five complete years of observations over the 1950-2010 period. A complete year is defined as a year with less than 10% of missing data.

To overcome the issue of the non-existence and/or shortcomings of the hydrological time series, a hydrological model is used to simulate streamflow over the complete 1950-2010 period. The ensemble of average daily climatic variables (temperature and precipitation) is used to calibrate a hydrological model to provide a fully coherent dataset of daily streamflow data for all catchments over the study period. The hydrological model used in this study is the GR4J model (modèle du Génie Rural à 4 paramètres Journaliers - daily rural engineering model with four parameters; Perrin et al., 2003) coupled with the CEMANEIGE degree-day snow model (Valéry, 2010). GR4J is a four-parameter lumped rainfall-runoff model that operates at the daily time step. GR4J is divided into two stores: a production store and a routing store. Inputs for the GR4J-CEMANEIGE model for a given day are precipitation, temperature and potential evapotranspiration. Potential evapotranspiration is calculated using the Oudin formula (Oudin et al., 2005), based on the daily extraterrestrial radiation. The GR4J-CEMANEIGE-OUDIN model structure was used in many previous studies and has proven to perform well in simulating daily streamflows on most of the NA catchments (Troin et al., 2015a; Troin et al., 2018; Velázquez et al., 2015b). A more detailed description of the GR4J-CEMANEIGE-OUDIN model structure can be found in Troin et al. (2018). In this study, the total number of parameters are fixed at nine (five for GR4J and four for CEMANEIGE) based on a previous study on the optimal number of parameters for this model (Poissant et al., 2017). The GR4J-CEMANEIGE-OUDIN model is calibrated with the available daily streamflow data over all catchments over the 1950-2010 period based on the Kling-Gupta Efficiency (*KGE*; Gupta et al., 2009) defined as,

$$KGE = 1 - \sqrt{(r-1)^2+(\alpha-1)^2+(\beta-1)^2} \tag{3.1}$$

where r is the correlation coefficient; α is a measure of variability; and β is the bias between observed and simulated streamflow data. *KGE* is a multi-objective metric allowing the correlation error, variability error and bias error to be optimized. *KGE* has been recently

demonstrated to be more appropriate than the accepted Nash-Sutcliffe Efficiency benchmark for hydrological model calibration (Pokhrel & Gupta, 2011). Readers can refer to Gupta et al. (2009) for more details in the calculation of this metric. The Shuffled Complex Evolution-University of Arizona (SCE-UA; Duan et al., 1992) optimization method with a total of 10 000 model evaluations is used to calibrate the model, as suggested by Arsenault et al. (2014).

A second screening is then performed to remove catchments with unsatisfactory calibration results. In the following, only the catchments with *KGE* values above 0.4 during the calibration are selected for further analyses. This leads to 4 153 catchments in the US and 384 catchments in Canada, with a total of 4 536 catchments across NA. Figure 3.1 shows the results of the calibration (with *KGE* above 0.4) illustrating, at the same time, the selected catchments in this work. The simulated daily streamflow data are then considered as pseudo-observations for the purpose of this study.

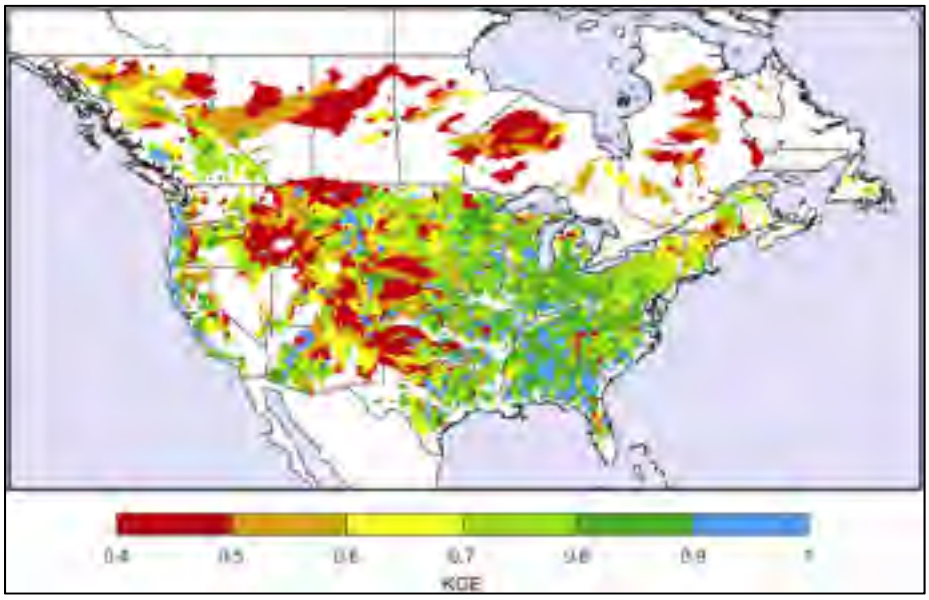

Figure 3.1 Map of the calibration results for the 4 536 catchments using the Kling-Gupta Efficiency (*KGE*) criterion. Catchments with the smallest area are overlapped on the catchments with the largest area

3.2.2 Climate indices

The standardized monthly values of the six climate indices analyzed in this study over the 1950-2010 period are shown in Figure 3.2. The indices are standardized by subtracting the mean and dividing by the standard deviation. Additional details regarding the selected climate indices are provided in Table 3.1.

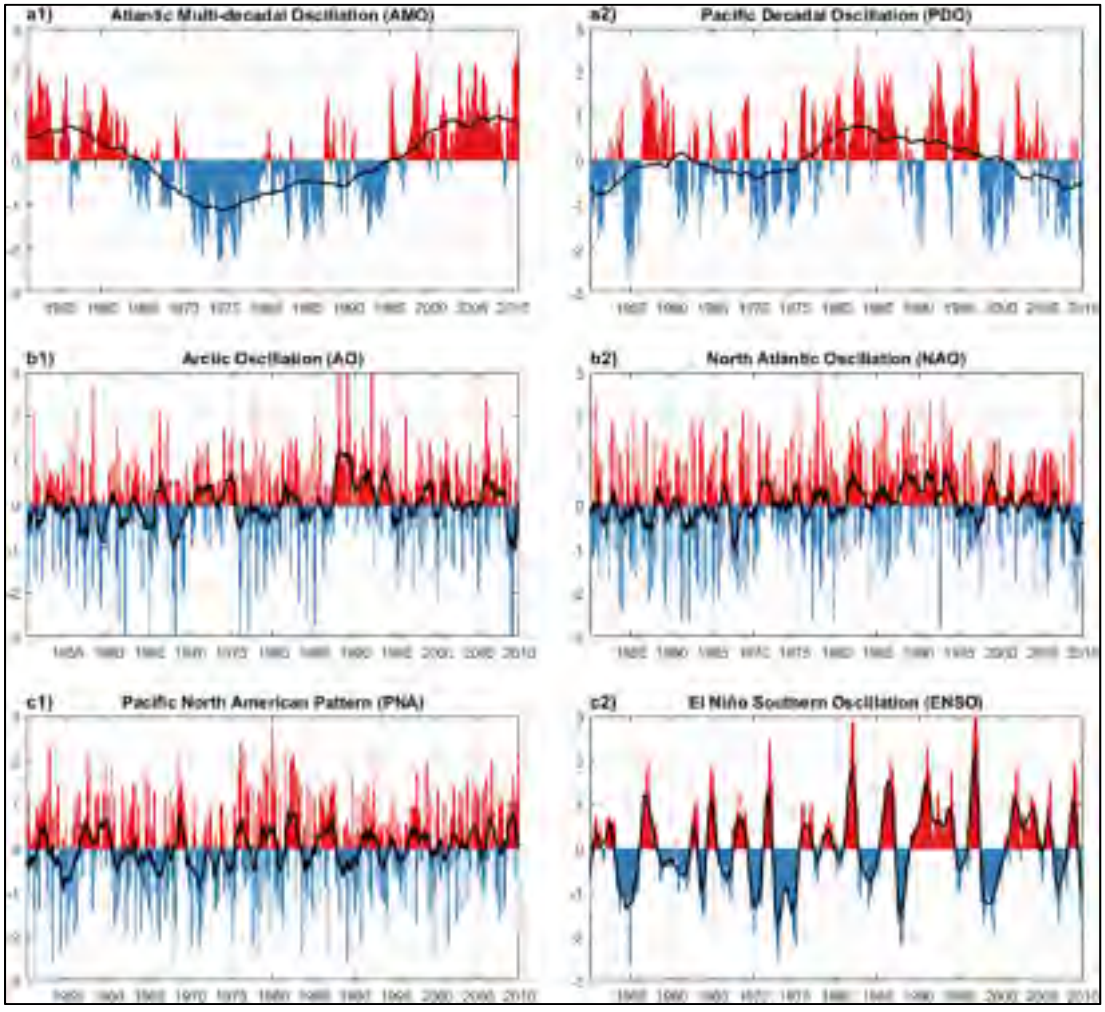

Figure 3.2 Standardized monthly values of the six climate indices over the 1950-2010 period. The top row represents the interdecadal climate indices – a1) AMO and a2) PDO. The middle row shows the decadal climate indices – b1) AO and b2) NAO. The bottom row illustrates the interannual climate indices – c1) PNA and c2) ENSO. Solid black lines are the 10-year running means for AMO and PDO, and the 1-year running means for the other four climate indices; the running means are plotted to show the different cycles between the climate indices

Table 3.1 Description of the climate indices used in this study. SST = sea surface temperature; SLP = sea level pressure.

Index	Type of phenomenon	Periodicity of the signal	Description	Location of the phenomenon
Atlantic Multidecadal Oscillation (AMO)	Oceanic	Interdecadal	Mean SST in the Atlantic, north of the Equator between 0°N -60°N and 75°W-7.5°W	North Atlantic Ocean
Pacific Decadal Oscillation (PDO)	Oceanic	Interdecadal	The leading principal component of monthly SST anomalies in the North Pacific Ocean, poleward of 20°N	North Pacific Ocean
Arctic Oscillation (AO)	Atmospheric	Decadal	Difference in SLP between the North Pole and the 45°N parallel	Extratropical Northern Hemisphere
North Atlantic Oscillation (NAO)	Atmospheric	Decadal	Difference in SLP between the north (Reykjavik, Iceland) and the south (Ponta Deldaga, Spain) of the North Atlantic Basin	Extratropical North Atlantic zone
Pacific North American (PNA) pattern	Atmospheric	Interannual	Quadrupole atmospheric model of anomalies in the geopotential height fields (low pressure)	From the subtropical west Pacific to the east coast of North America
El Niño and Southern Oscillation (ENSO)	Oceanic-Atmospheric	Interannual	Difference in SLP anomalies between Tahiti and Darwin (known as the Southern Oscillation Index) and SST in the equatorial Pacific	Tropical South Pacific

Interdecadal oceanic data include the Atlantic Multidecadal Oscillation (AMO) and the Pacific Decadal Oscillation (PDO). The AMO index consists of mean sea surface temperature (SST) anomalies in the North Atlantic Ocean region; and the PDO index corresponds to the leading principal component of monthly SST anomalies in the North Pacific Ocean (Table 3.1). Using a 10-year running mean for the period 1950 to 2010, the warm phases (from 1950 to 1963 and from 1995 to 2010) of the AMO index are a positive numerical value of the standardized monthly index values, while the cold phase (from 1964 to 1994) is a negative numerical value (Figure 3.2-a1). Similarly the PDO index exhibits cold phases from 1950 to 1976 and from 1998 to 2010 and a warm phase from 1977 to 1997 (Figure 3.2-a2).

Decadal atmospheric data include the Arctic Oscillation (AO) and the North Atlantic Oscillation (NAO). The AO index represents the difference in sea level pressure (SLP) between the North Pole and the 45°N parallel over the extratropical Northern Hemisphere region; while the NAO index is defined as the difference in normalized mean winter

(December to March) SLP anomalies between Iceland and Spain of the North Atlantic Basin (Table 3.1). Since 1950, AO and NAO display both interannual variability and long-term fluctuations (Figure 3.2-b1 and Figure 3.2-b2). For the two indices, this results in a cold (negative) phase from 1950 to 1970. Positive/negative fluctuations are shown between 1970 and 2010, not always synchronous between both indices (Figure 3.2-b1 and Figure 3.2-b2).

Interannual oceanic-atmospheric data include the Pacific North American (PNA) pattern and the El Niño and Southern Oscillation (ENSO). The PNA index is depicted by a quadrupole atmospheric model of anomalies in the geopotential height from the subtropical west Pacific to the east coast of North America (Table 3.1). ENSO represents the difference in SLP anomalies between Tahiti and Darwin (known as the Southern Oscillation Index) and SST in the equatorial Pacific (Table 3.1). In this study, the ENSO is represented using the SST anomalies over the Niño 3.4 SST region located along the equatorial Pacific Ocean (5°S–5°N, 170°–120°W). PNA and ENSO display strong interannual variability over the 1950-2010 period (Figure 3.2-c1 and Figure 3.2-c2).
All the six climate indices values are available from the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL) (https://www.esrl.noaa.gov/psd/data/climateindices/).

3.2.3 Methods

3.2.3.1 Hydroclimatic variables

Three hydroclimatic variables are analyzed in this study: temperature, precipitation and streamflow. By considering these three variables, the comprehensive response of the catchments to the NCV can be explored. In the following, the focus is set on the mean daily values of temperature (TAS) and streamflow (QM), and the total wet-day precipitation (PRCPTOT). Extreme values are also investigated by considering the total maximum 1-day precipitation (RX1day) and streamflow (QX1day). The analysis is conducted over the winter (DJF; December-January-February) and summer (JJA; June-July-August) seasons.

3.2.3.2 Relationships between hydroclimatic variables and climate indices

First, the individual impacts of the climate indices on the hydroclimatic variables are investigated. The relationship between the climate indices and the hydroclimatic variables is calculated using the single correlation method. This method is chosen because it was used in many studies of hydroclimatic trend detection, making the direct inter-comparison of results possible. Each hydroclimatic variable (X) is compared to each climate index (Y) based on the correlation coefficient (ρ) as

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{3.2}$$

where cov is the covariance between X and Y; σ_X and σ_Y are the standard deviations of X and Y, respectively. The square of the correlation coefficient is a measure of how well the variance of the dependent variable (i.e., the hydroclimatic variable, such as TAS) is predictable by the independent variable (i.e., the climate index, such as ENSO), and corresponds to the percentage of the variation explained (R^2) by a best-fit regression line between the two variables. In the following, only significant ρ values ($\alpha = 0.05$) are considered, resulting in a minimum R^2 value of 6.5%. This methodology is repeated for the 4 536 selected catchments over the 1950-2010 period. Note that the raw data is used for computing R^2 between a climate index and a hydroclimatic variable.

Next, the all-combined impacts of the climate indices on the hydroclimatic variables are explored. The relationship between all-combined climate indices and the hydroclimatic variables is calculated using the stepwise linear regression method. The stepwise linear regression method uses forward and backward stepwise regressions, where, at each step, some terms (i.e., climate indices) are added or removed from the regression, based on the adjusted value of the percentage of variance explained (R^2_{adj}) as

$$R^2_{adj} = 1 - (1 - R^2)\frac{(n-1)}{(n-p-1)} \tag{3.3}$$

where n is the number of observations and p is the number of regression coefficients. The advantage of using the R^2_{adj} criterion compared to the regular R^2 criterion is that the increase in R^2 value due to the inclusion of a new explanatory variable must be statistically significant instead of showing spurious increases. The R^2_{adj} value obtained from the regression is used to express the total percentage of explained variance predictable by the combination of one or more climate indices. The R^2_{adj} value is, by default, equal or higher than the highest individual R^2 value by any of the individual climate index. This methodology is repeated for the 4 536 catchments over the 1950-2010 period.

In the following, the intensity of the relationships between the hydroclimatic variables and the climate indices are categorized as insignificant with R^2 and R^2_{adj} values below 6.5%, weak between 6.5% and 30%, moderate between 30% and 55%, strong between 55% and 80%. Note that no value of R^2 and R^2_{adj} is obtained above 77%.

3.3 Results and discussion

The results of the individual and all-combined impacts of the climate indices on the hydroclimatic variables are presented in Figures 3.3 to 3.9. Additional details concerning the number of catchments significantly influenced by the climate indices are provided in Table 3.2. For all hydroclimatic variables and catchments, only significant R^2 and R^2_{adj} values (above 6.5%) are considered for further analysis corresponding to the 95% confidence level. An analysis of the R^2_{adj} values for the five hydroclimatic variables is presented in Table 3.3.

3.3.1 Relationships between temperature and climate indices

Figure 3.3 presents the results of the R^2 values for winter mean temperature (DJF TAS).

Table 3.2 Number of catchments (%) with significant relationships between the hydroclimatic variables and the climate indices. The analysis is presented for the three hydroclimatic variables (TAS, PRCPTOT and QM) over the two seasons (DJF and JJA) when considering individual and all-combined climate indices. Note that the relationships between the hydroclimatic variables and the climate indices are considered as significant with a value of R^2 or R^2_{adj} above 6.5%

	AMO	PDO	AO	NAO	ENSO	PNA	Combined	Best index
DJF TAS	14%	60%	60%	68%	29%	57%	98%	95%
JJA TAS	56%	7%	37%	36%	17%	5%	92%	83%
DJF PRCPTOT	6%	33%	15%	11%	34%	41%	77%	69%
JJA PRCPTOT	3%	12%	20%	19%	10%	3%	57%	45%
DJF QM	8%	23%	14%	22%	27%	18%	75%	64%
JJA QM	14%	11%	8%	16%	5%	6%	56%	45%

Table 3.3 Number of catchments (%) by class of R^2_{adj} values. The analysis is presented for the five hydroclimatic variables (TAS, PRCPTOT, QM, RX1day and QX1day) over the two seasons (DDJ and JJA) when considering all-combined climate indices

	Insignificant (< 6.5%)	Weak (6.5% - 30%)	Moderate (30% - 55%	Strong (55% - 80%)
DJF TAS	2%	35%	49%	14%
JJA TAS	8%	57%	34%	<1%
DJF PRCPTOT	23%	67%	10%	<1%
JJA PRCPTOT	43%	54%	2%	0%
DJF QM	25%	70%	5%	<1%
JJA QM	44%	53%	4%	<1%
DJF RX1day	38%	58%	4%	<1%
JJA RX1day	54%	45%	2%	0%
DJF QX1day	37%	59%	4%	<1%
JJA QX1day	44%	50%	5%	<1%

AMO winter temperature responses are sporadic and variable across NA, with only 14% of the catchments showing a potential link to this climate index (Figure 3.3-a1 and Table 3.2). The intensity of the DJF TAS response to AMO is weak or moderate, with R^2 values ranging between 6.5% and 40%. A loose tendency to increased strength of the relationship between DJF TAS and AMO is observed for the catchments located in eastern (E) Canada.

As for the other five climate indices, indications of the related effects on winter temperature are more geographically widespread, with an influence encompassing up to 68% of the NA catchments (Table 3.2). Concerning the impacts of PDO, ENSO and PNA on DJF TAS, two distinct regions across NA are identified: the southeast (SE) and the northwest (NW) regions (Figure 3.3-a2-a5-a6). For the SE region, the intensity of the relationships between winter temperature and the three climate indices increases as it moves towards the state of Louisiana.

The strongest SE regional link is observed for PDO, where it explains up to 50% of the observed winter temperature variance (Figure 3.3-a2). This is particularly interesting, since PDO is a pattern of North Pacific climate variability, whose effects are not expected to directly affect temperature of SE NA. Mantua et Hare (2002) have also reported the potential impacts of PDO on the SE region of the US, corroborating the PDO links shown in this study. This climate index is well-known for its prominent effects on the wintertime temperatures of the NW NA (Bonsal, Shabbar, & Higuchi, 2001; Mantua, Hare, Zhang, Wallace, & Francis, 1997; Papineau, 2001; Shabbar & Bonsal, 2004; Stahl, Moore, & McKendry, 2006), that confirms the PDO - DJF TAS linkages observed on Figure 3.3-a2.

Additionally, the dependence of PNA on winter temperature is clearly seen for the catchments located in the SE and NW regions, with strong R^2 values reaching up to 70% of the observed DJF TAS variance over western (W) Canada and moderate R^2 values for the extreme W US (Figure 3.3-a6). This is in agreement with Yu, Lin, Wu, et Merryfield (2016), which demonstrated that NA winter temperature variability is largely controlled by PNA. Given the documented association between PNA and both winter climates (Leathers, Yarnal, & Palecki, 1991) and synoptic weather patterns (Wallace & Gutzler, 1981), this provides indication that the strong winter temperature responses over the NW-SE regions might be linked to an enhanced PNA pattern (Liu et al., 2015).

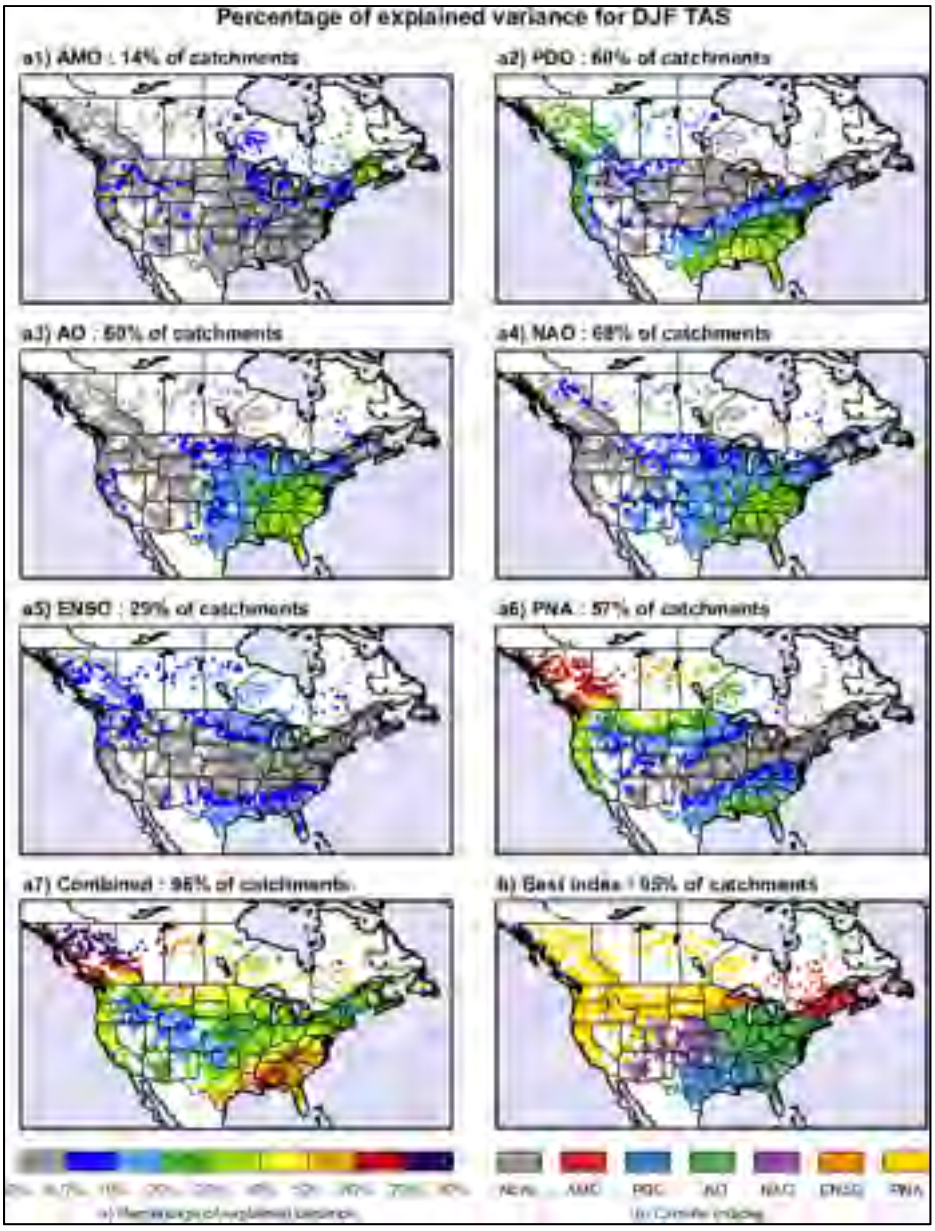

Figure 3.3 Percent explained variance (R^2) map for winter (DJF) mean temperature (TAS). The first three rows show the maps for each climate index: a1) AMO, a2) PDO, a3) AO, a4) NAO, a5) ENSO and a6) PNA. The bottom left panel (a7) represents the R^2_{adj} values when considering all-combined climate indices. The bottom right panel (b) shows the climate index for which the highest value of R^2 is obtained. For all panels, the colored points correspond to the catchments with significant R^2 or R^2_{adj} values (above 6.5%); insignificant R^2 or R^2_{adj} values (below 6.5%) are represented by gray color points and are plotted behind not to mask significant R^2 and R^2_{adj} values. The number of catchments (%) with significant R^2 values is also indicated for each climate index

Results indicate an overall weak response of DJF TAS to ENSO, with a maximal R^2 intensity of 20% (Figure 3.3-a5). This encompasses 29% of the catchments located in the two regions (Table 3.2). The identified regional effects of ENSO on winter temperature are consistent with the literature with, however, a lower amplitude. Many studies pointed to the fact that large regions of NA experience significant modulation of winter temperature by ENSO. Shabbar et Khandekar (1996) showed strong relationships between ENSO and winter temperature over west-central (W-C) regions of Canada, mainly attributed to changes in mid-tropospheric circulation patterns associated with the PNA. Bonsal et al. (2001) also evidenced regional influences of ENSO on winter temperature variability across Canada, and underlined significant modulating effect of winter PDO pattern on ENSO related temperature responses. A number of studies have described the effects of ENSO on winter temperature variability over the SE US (e.g., Redmond & Cayan, 1994; Ropelewski & Halpert, 1986, 1996).

As for NAO and AO, these indices show a similar band of significant DJF TAS correlations mainly located in the SE region of the US, with moderate R^2 values up to 50% over the Appalachian region (Figure 3.3-a3-a4). Strong NAO and AO related winter temperature responses also occur for several catchments confined to northeastern (NE) and E Canada, somewhat similar to the AMO regional influence. NAO is the dominant index affecting winter temperature, with up to 68% of the catchments impacted across NA (Table 3.2). This is consistent with previous observational studies of NA winter temperature (Bonsal et al., 2001; Higgins, Leetmaa, & Kousky, 2002; Hurrell, 1996; Lim & Schubert, 2011).

When the influence of the six climate indices is combined, the intensity of the relationship on DJF TAS significantly increases, with a more geographically widespread impact encompassing 98% of the catchments (Figure 3.3-a7 and Table 3.2); among these catchments, 14% show strong R^2_{adj} values (Table 3.3). The largest amplitudes are found for catchments in W Canada and in the extreme W US, mainly driven by PNA (Figure 3.3-a6). Strong impacts are also observed for the SE US and the NE Canada, where the PDO, PNA, AO, NAO and the AMO, AO, NAO influences, respectively, are detected (Figure 3.3-a1 to

a4). The ENSO contribution is somewhat small with regional impacts confined at the west of the Great Lakes and in the SE US.

For many NA regions, the results suggest that moderate to strong R^2 values are obtained when using individual or all-combined climate indices. This provides an indication of the potential of some climate indices to be used as predictors of NA winter temperature. However, the choice of the climate index depends on the investigated region (Figure 3.3-b). For instance, one climate index (PNA) might allow a reasonable DJF TAS prediction to be realized for the catchments in NW NA, while a combination of two indices (PDO and AO) appears more appropriate for the catchments in SE NA.

Figure 3.4 presents the results of the R^2 values for summer mean temperature (JJA TAS).

Compared to winter temperature, the influence of the climate indices on summer temperature is not only weaker but is also more geographically restricted across NA. Among the six climate indices, the indications of the related effects on JJA TAS are more apparent for AMO, encompassing 56% of the catchments mainly situated in the SE-E NA (Figure 3.4-a1 and Table 3.2). More precisely, AMO has significant effects on summer temperature that extend over the SE US, with R^2 values up to 40%; however, moderate R^2 values are also sporadically widespread throughout E Canada.

While AO and NAO are two physically close indices, they lead to contrasted influences on summer temperature. The obtained NAO-JJA TAS pattern closely resembles the AMO-JJA TAS pattern, however, with somewhat less geographically amplitude in E NA (Figure 3.4-a4 and Table 3.2). AO also has an impact on summer temperature for the 37% of catchments located over N-E-SE NA, with R^2 values ranging between 6.5% and 30% (Figure 3.4-a3 and Table 3.2).

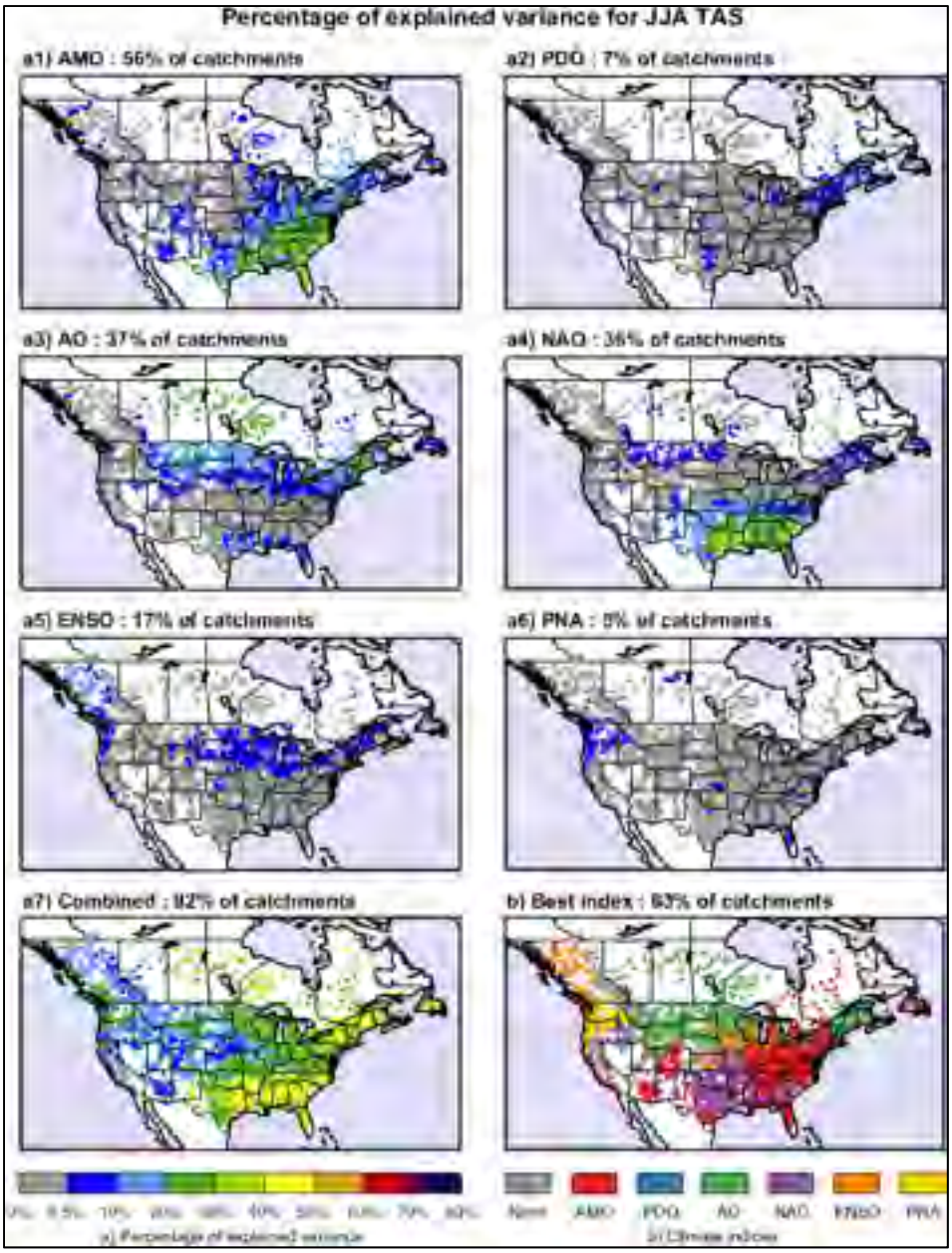

Figure 3.4 Same as Figure 3.3, but for summer (JJA) TAS

As for ENSO, PDO and PNA, the influence of these indices on summer temperature is relatively weak, with maximum R^2 values of 20% (Figure 3.4-a2-a5-a6). The spatial influence of ENSO is slightly larger than PDO and PNA, including 17% of the catchment over NW-C-E NA. It is noted that the related influences of PDO and PNA on JJA TAS strongly differ than those on DJF TAS, where PDO and PNA, associated with NAO, are

considered as the prominent indices affecting winter temperature variability. This is in agreement with findings from previous studies, where researchers showed that these climate indices reach small amplitude and spatial extent during summer, with, consequently, low expected effects on NA temperature variability for this season (e.g., Fleming & Whitfield, 2010; Halpert & Ropelewski, 1992; Hurrell, Kushnir, Ottersen, & Visbeck, 2003).

When considering the all-combined influences of the climate indices on JJA TAS, it is evident that the zone of influence of the indices increases impacting 92% of the study catchments across NA (Figure 3.4-a7 and Table 3.2). As expected, higher R^2_{adj} values between JJA TAS and the climate indices are observed, with a maximum value of 60% over the SE US. Nevertheless, the intensity of the relationship between all-combined climate indices and JJA TAS is lower compared to that obtained for DJF TAS; the majority of the catchments are situated in the weak category (Table 3.3).

An indication of the climate indices potentially used for predicting NA summer temperature is shown on Figure 3.4-b. As for DJF TAS, one or more climate indices can be associated to a catchment characterizing a specific region. The influence zone of AO defines the N NA, while the influence zone of AMO-NAO and ENSO-PNA is the SE and NW NA, respectively.

3.3.2 Relationships between precipitation and climate indices

Figure 3.5 presents the results of the R^2 values for winter (DJF) total wet-day precipitation (PRCPTOT).

Variations of NA winter precipitation are modulated only moderately by the climate indices in comparison with seasonal temperature. It is identified that AMO, AO and NAO cannot explain the winter precipitation variability in a significant manner, since their influence is relatively weak with R^2 values ranging between 6.5% and 20% and it encompasses less than 20% of the NA catchments (Figure 3.5-a1-a3-a4 and Table 3.2). It is noted that the AMO

influence is confined to western NA while both the AO and NAO influences cover the Appalachian region, the peninsula of Florida and sporadically the NW and C regions of NA.

However, more of winter precipitation variance of the NA catchments can be explained by PDO, ENSO and PNA with R^2 values up to 40% (Figure 3.5-a2-a5-a6); these indices reflect the conditions in the Pacific Ocean (Table 3.1). The PDO-DJF PRCPTOT regional pattern tightly resembles that of PNA-DJF PRCPTOT, with, respectively, 33% and 41%, of the NA catchments impacted; this primarily includes the Appalachian region, the peninsula of Florida and the regions along the NW-S diagonal (Table 3.2). The ENSO related influence on DJF PRCPTOT is noticeable for 34% of the catchments over the southern Great Lakes region and the W-S-SE regions of NA. The influence of these three indices on the temporal variability of winter precipitation over NW-W NA has been recognized by several authors (e.g., Fleming & Whitfield, 2010; Hu & Huang, 2009; Ropelewski & Halpert, 1986; Stahl et al., 2006).

The findings of the present study are in partial agreement with the other studies that showed that ENSO can serve as a useful predictor of winter precipitation variability for the catchments on the entire region of S NA (e.g., Enfield et al., 2001; Goly & Teegavarapu, 2014; Hidalgo & Dracup, 2003; Kalra & Ahmad, 2012; Kim, Valdés, Nijssen, & Roncayolo, 2006). For the other regions, a combination of two or more climate indices seems to be required for the predictions of winter precipitation; as for instance, when considering the all-combined impacts of the climate indices on DJF PRCPTOT, this leads to enhanced relationships both in amplitude (R^2_{adj} values up to 70%) and in spatial extend over and around the Appalachian region (Figure 3.5-a7). For that region, the best predictability of winter precipitation variability might be obtained using a combination of four indices: PDO, PNA, AO and NAO (Figure 3.5-b).

Figure 3.6 presents the results of the R^2 values for summer (JJA) total wet-day precipitation (PRCPTOT).

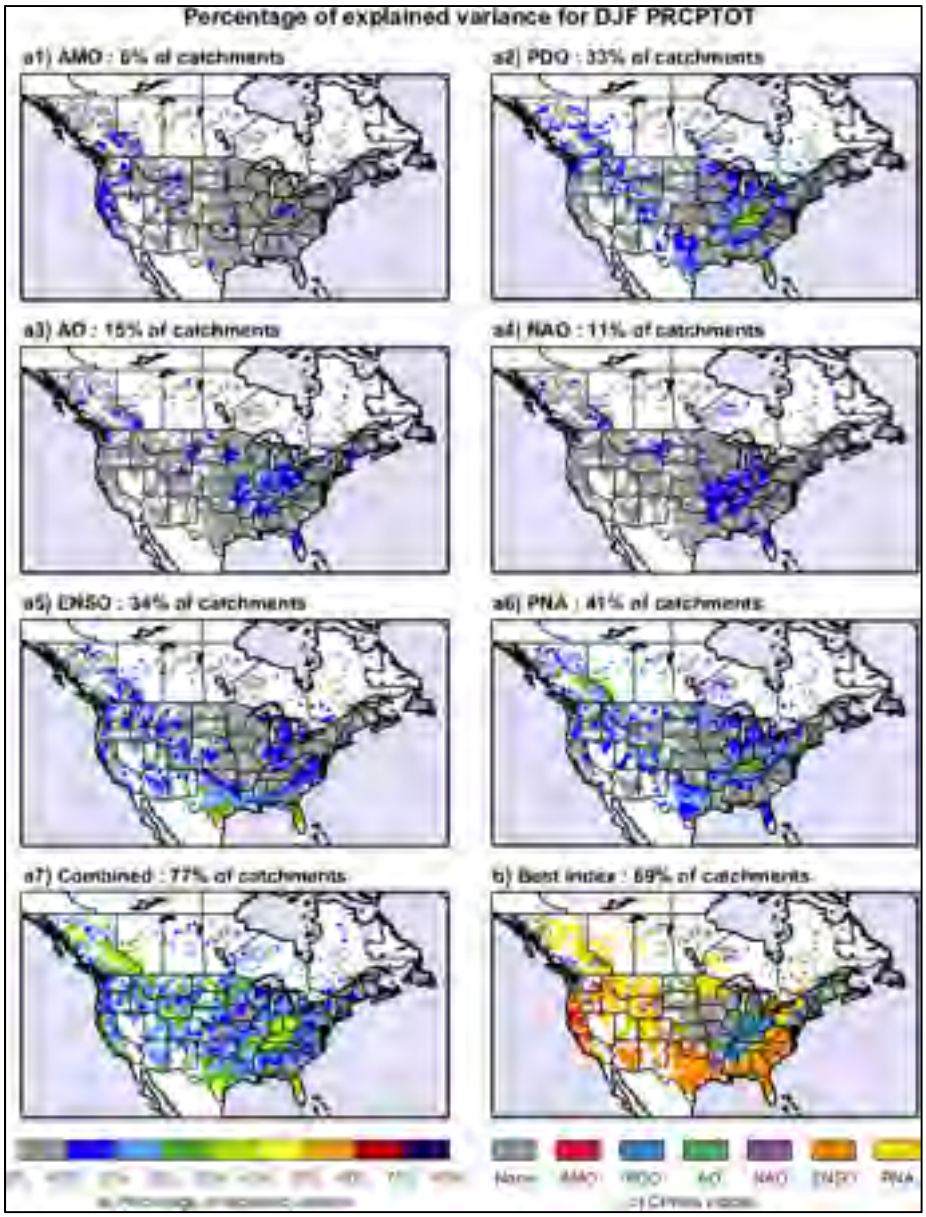

Figure 3.5 Same as Figure 3.3, but for winter (DJF) PRCPTOT

Compared to winter precipitation, variability of NA summer precipitation is even less impacted by the climate indices. The intensity of the relationships between JJA PRCPTOT and both AMO and PNA are weak, with R^2 values between 6.5% and 20%. No clear pattern is identified, as 97% of the catchments are insensitive to AMO and PNA with regards to summer precipitation (Figure 3.6-a1-a6 and Table 3.2).

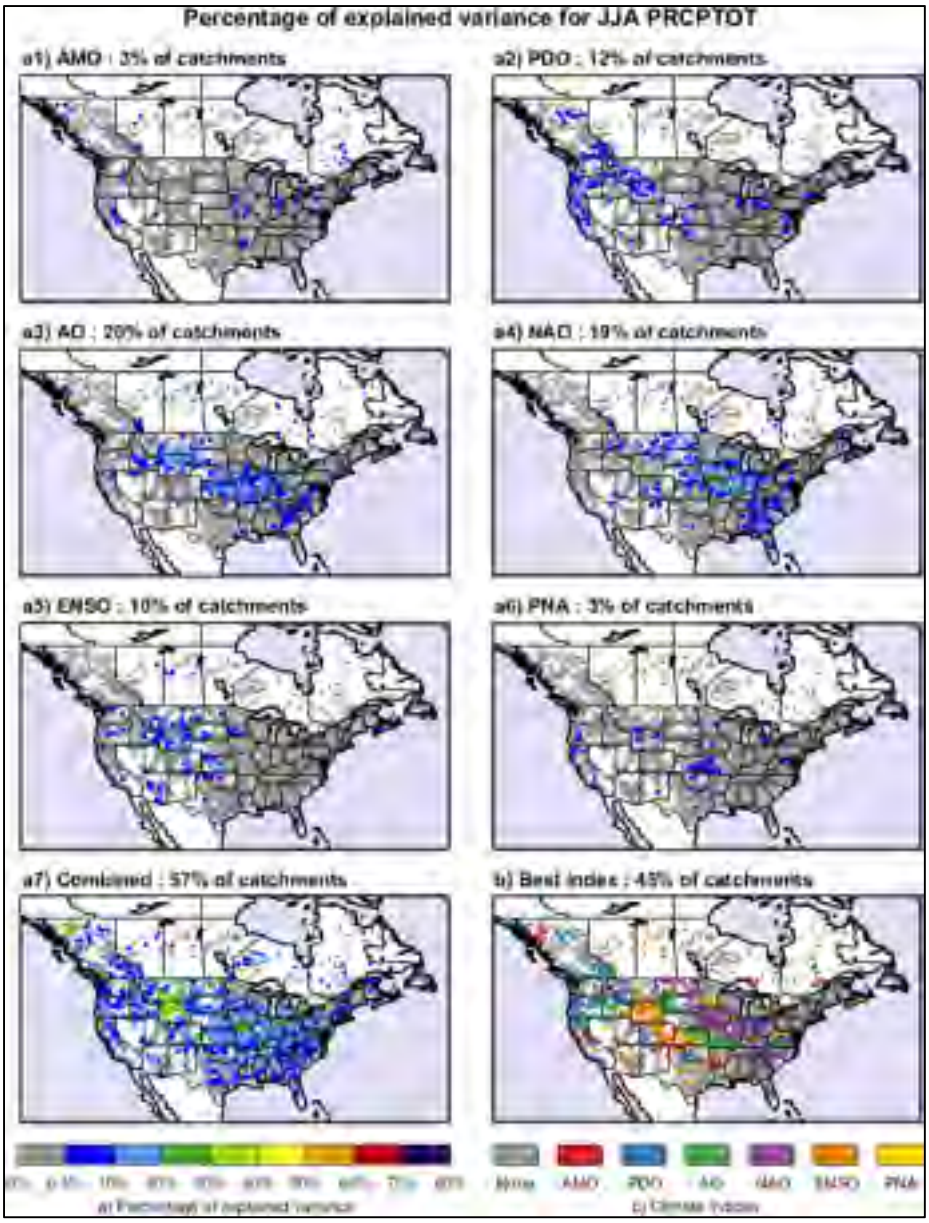

Figure 3.6 Same as Figure 3.3, but for summer (JJA) PRCPTOT

As for PDO and ENSO, a slight increase impact on JJA PRCPTOT is observed, where the R^2 values reach up to 30% over W-C regions, respectively (Figure 3.6-a2-a5). Nevertheless, only 10-12% of the NA catchments seem to be influenced by ENSO and PDO (Table 3.2).

Indications of the AO and NAO related effects on JJA PRCPTOT are somewhat more geographically widespread, with an influence encompassing until 20% of the NA

catchments, principally located over C-SE regions of NA (Figure 3.6-a3-a4 and Table 3.2). However, the R^2 values remain weak. When all the climate indices are combined, moderated impacts on JJA PRCPTOT are notable in the middle of the US, mostly in areas with relatively low annual precipitation compared to other areas (Figure 3.6-a7). Besides that, the results suggest, to some extent, insignificant influences of the combination of all-indices on JJA PRCPTOT, since 43% of the catchments are insensitive to their effects (Table 3.3). Thus, a single climate index (or a combination of more than one) does provide a limited predictability of summer precipitation variability for only 45% (57%) of the NA catchments (Figure 3.6-a7-b).

3.3.3 Relationships between streamflow and climate indices

Figure 3.7 presents the results of the R^2 values for winter (DJF) mean streamflow (QM).

The climate indices affect very moderately the variability of winter mean streamflow across NA. Compared to winter precipitation (Figure 3.5), there is a general tendency to a decreasing intensity of the relationships between DJF QM and the most prominent indices (PDO, ENSO and PNA), both in the amplitude (maximum R^2 values of 30%) and in the spatial extent (maximum 27% of catchments impacted). However, noteworthy for these three indices is that the overall patterns obtained for winter mean streamflow are quite similar to those for winter precipitation (Figure 3.7-a2-a5-a6). Again, the PDO-DJF QM pattern has some matching sides with the PNA-DJF QM pattern, with an influence confined over the Appalachian region, the peninsula of Florida and the regions along the NW-S diagonal. The ENSO related influence on DJF QM affects the catchments over the western Great Lakes region and the W-S-SE regions of NA. The review of several studies, focusing on PDO, PNA and ENSO, reveals a clear linkage between these indices and the seasonal streamflow variability over the NW-W regions of NA (e.g., Gobena & Gan, 2006, 2009; Gobena et al., 2013; Hamlet & Lettenmaier, 1999; Kahya & Dracup, 1993), as well as for the S NA (Hidalgo & Dracup, 2003; Kahya & Dracup, 1992; Maurer, Lettenmaier, & Mantua, 2004) and the peninsula of Florida (Tootle & Piechota, 2005).

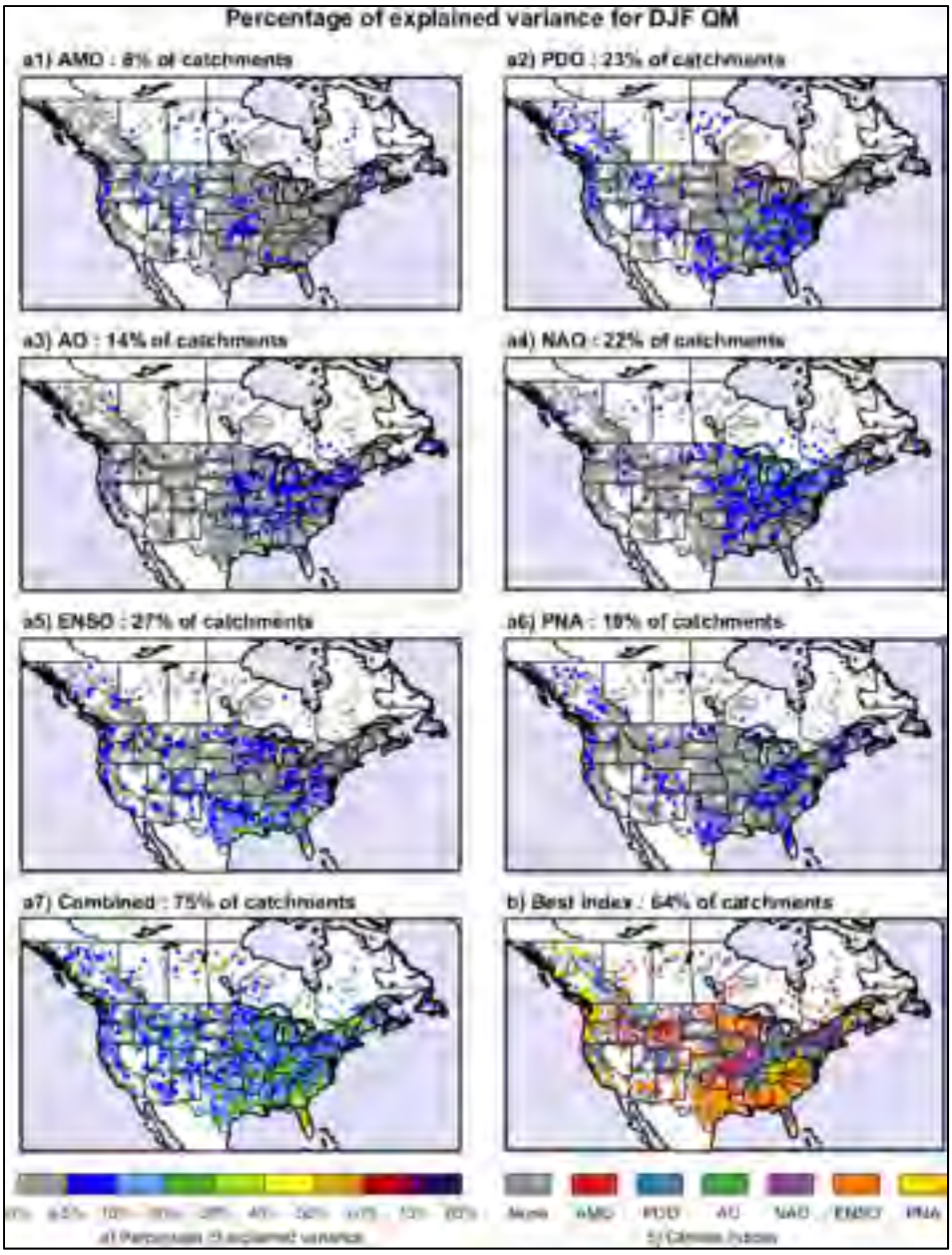

Figure 3.7 Same as Figure 3.3, but for winter (DJF) QM

For the other three indices (AMO, NAO and AO), the similarities with winter precipitation patterns are less perceptible and, curiously, their related influence on DJF QM slightly increases in terms of spatial extent compared to DJF PRCPTOT; though, the R^2 values remain weak and below 20% (Figure 3.7-a1-a3-a4 and Table 3.2). Again, AMO is the index that affects the least of catchments without a well-defined influence zone across NA (8%;

Figure 3.7-a1). The impact of AO and NAO dominates the Appalachian region and the surrounding areas, as well as some catchments located in the extreme E-N Canada and in C US in a more ad hoc manner (Figure 3.7-a3-a4).

As expected, the patterns of all-combined indices-DJF QM and all-combined indices-DJF PRCPTOT are very similar in terms of spatial extent (75% and 77% of catchments impacted, respectively) but the amplitude of relationships is reduced for DJF QM (maximum R^2_{adj} values of 60%; Figure 3.7-a7 and Table 2); 70% of the catchments show weak R^2_{adj} values and only 5% of catchments presents moderated R^2_{adj} values (Table 3.3). For this hydroclimatic variable, it is not surprising that ENSO is presented as the best index as regards with the predictability of DJF QM in the S-SE US (Figure 3.7-b). This coincides with the findings of many studies (e.g., Kahya & Dracup, 1992; Massei et al., 2011; Maurer et al., 2004; Tootle & Piechota, 2005), which showed the well-established ENSO signal on the streamflow variability over these regions. Nevertheless, for the other regions, it is evident that the prediction of DJF QM goes through a combination of several indices, though it appears difficult to conclude which combinations of indices will dominate the DJF QM prediction.

Figure 3.8 presents the results of the R^2 values for summer (JJA) mean streamflow (QM).

Compared to winter mean streamflow, variability of NA summer mean streamflow is less impacted by the climate indices, with the exception of AMO. AMO is the single index for which the intensity of the relationships on JJA QM increases compared to DJF QM, both in the amplitude (R^2 values up to 40%) and in spatial extent (Figure 3.8-a1 and Table 3.2), until becoming the best index for predicting JJA QM for the catchments situated over W-N-NE NA (Figure 3.8-b). AMO is recognized to have a substantial impact on seasonal streamflow variability over these regions (e.g., Assani et al., 2012; Maurer et al., 2004; Mazouz, Assani, Quessy, & Légaré, 2012; Rogers & Coleman, 2003; Tootle & Piechota, 2005). The results are, however, hard to understand since AMO appears to have little effect on JJA TAS and JJA PRCPTOT (Figure 3.4-a1 and Figure 3.6-a1). While streamflow characteristics should

reflect changes in the characteristics of the precipitation regime, the nonlinear precipitation–streamflow patterns may lead to a different spatial pattern of climate-related signals in streamflow from that of precipitation (Kahya & Dracup, 1993). Piechota, Dracup, et Fovell (1997) argues that, in contrast to precipitation, streamflow is less prone to contamination by noise in time and space as it is a naturally filtered product of precipitation.

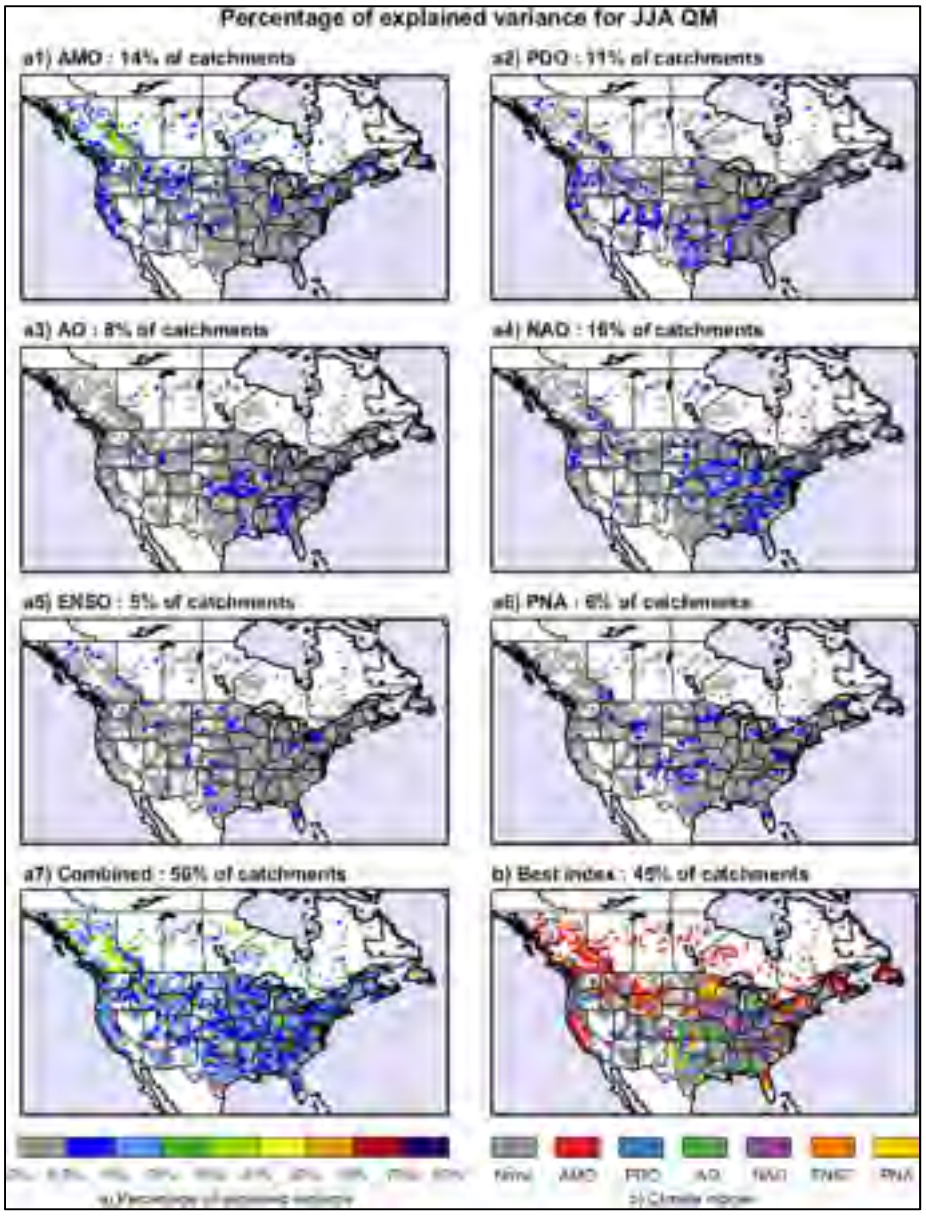

Figure 3.8 Same as Figure 3.3, but for summer (JJA) QM

As for the other five indices, their related influence on JJA QM is relatively weak with R^2 values ranging between 6.5% and 20%. The effects of ENSO and PNA on JJA QM are feeble, since they affect less than 6% of the NA catchments without an obvious pattern (Figure 3.8-a5-a6 and Table 3.2). Links between JJA QM and PDO-AO-NAO are somewhat more perceptible, with a larger spatial widespread encompassing between 8% and 16% of the NA catchments (Figure 3.8-a2-a3-a4 and Table 3.2). The PDO influence is still confined over the regions along the NW-S diagonal and part of the Appalachian region; while the AO-NAO influences cover the SE and mid-E of US and punctually the catchments over W-N NA.

By considering the all-combined impacts of the indices on JJA QM, the amplitude of the relationships slightly intensifies over northwestern NA with R^2_{adj} values up to 60% for 4% of catchments (Figure 3.8-a7 and Table 3.3). Nevertheless, the impacts of the combined indices on JJA QM remain weak, since only 56% of NA catchments are impacted (Figure 3.8-a7). As previously mentioned, except for the W-N-NE regions of NA where AMO is presented as the only index potentially adapted for the JJA QM prediction, the findings are more nuanced for the other regions, where either none of the study indices seems to be suitable to or a combination of more two indices is minimally required to predicting summer mean streamflows (Figure 3.8-b).

3.3.4 Relationships between extreme hydroclimatic variables and climate indices

An evaluation of the influence of the climate indices on hydroclimatic extremes is performed by examining the potential linkages of the indices with precipitation (RX1day) and streamflow (QX1day) extremes for the winter (DJF) and summer (JJA) seasons. The analysis is restricted to the all-combined impact of the indices on hydroclimatic extremes.

Of Figure 3.9 and compared with the mean hydroclimatic variables' results (Figures 5 to 8), it can be inferred that the all-combined influence of the indices on hydroclimatic extremes is less intense in terms of magnitude (R^2_{adj} values) spatial extent (number of catchments

impacted). Similarly, the impact of the all-combined indices on hydroclimatic extremes is more noticeable for the winter than the summer season (Figure 3.9 and Table 3.3).

The all-combined indices signal influencing winter precipitation extreme is spatially non-uniform throughout NA. Although the relationships between DJF RX1day and the all-combined indices is relatively weak for many catchments, more marked effects emerge over W NA and the Appalachian region, with moderate R^2_{adj} values (Figure 3.9-a1). As for JJA RX1day, the majority of the catchments show insignificant to weak R^2_{adj} values, with some moderate values sporadically distributed over N NA; though, no clear pattern is identified in the variance of JJA RX1day (Figure 3.9-a2).

Considering streamflow extremes, the DJF QX1day pattern is somewhat similar to the DJF RX1day pattern with a close value of impacted catchments (Figure 3.9-a1-a3). Similarly, the SW NA and the Appalachian region show moderate R^2_{adj} values. However, the DJF QX1day pattern exhibits moderate values over N NA and Florida which are not identified on the DJF RX1day pattern. In contrast, no clear similarities between the JJA RX1day and JJA QX1day patterns are observed (Figure 3.9-a2-a4); and stronger all-combined influence of the indices is displayed for JJA QX1day both in terms of amplitude (R^2_{adj} values up to 70%) and spatial extent (56% versus 46% of catchments impacted). The leading influence zones encompass the NW-N-E regions of NA (Figure 3.9-a4).

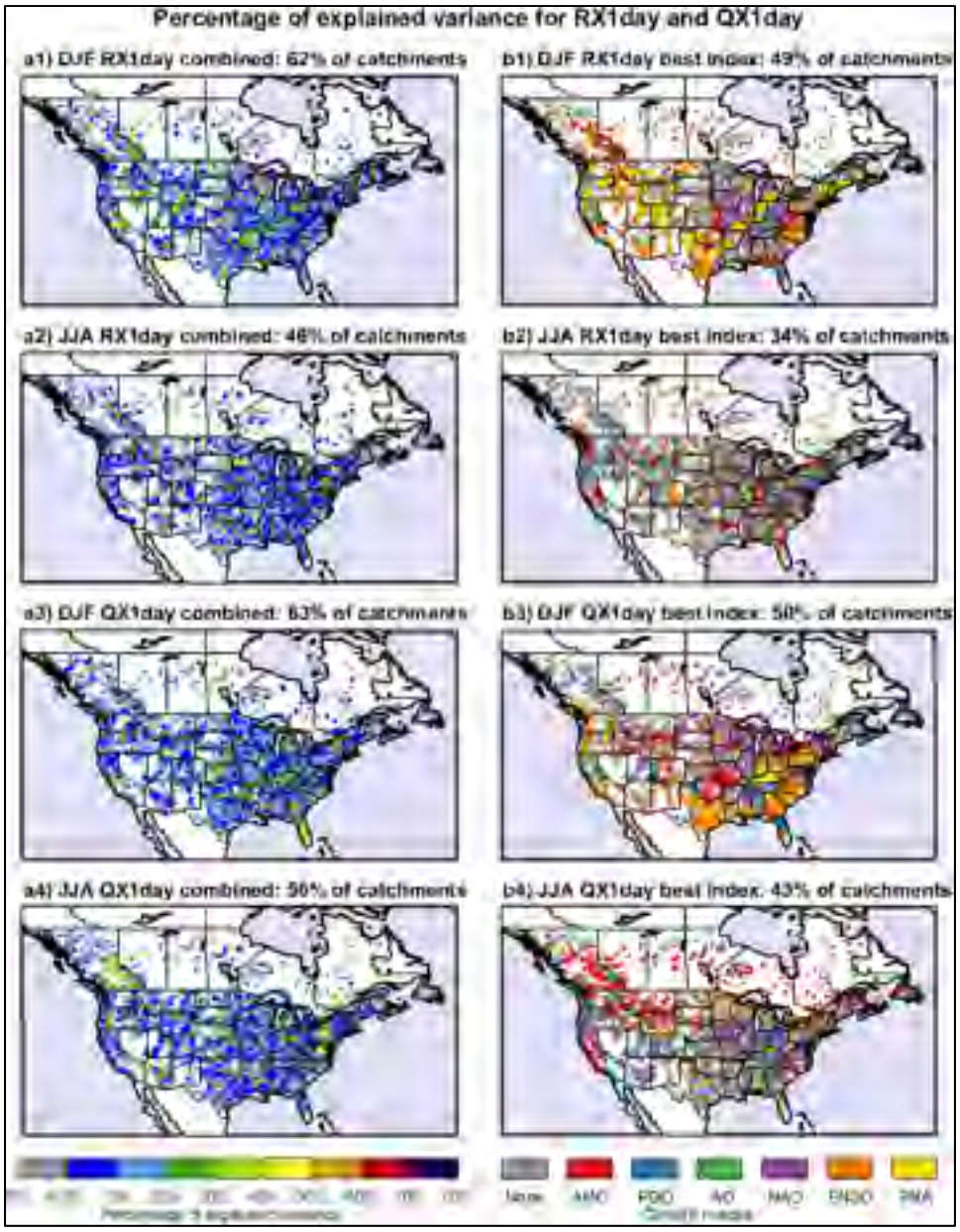

Figure 3.9 Percent explained variance (R^2_{ajd}) map for winter (DJF) and summer (JJA) RX1day and QX1day. The left-hand side panels (a) represent the R^2_{adj} value when considering all-combined climate indices. The right-hand side panels (b) show the climate index for which the highest value of R^2 (like previous figures) is obtained. For all panels, the colored points correspond to the catchments with significant R^2_{adj} values (above 6.5%); insignificant R^2_{adj} values (below 6.5%) are represented by gray color points and are plotted behind not to mask significant R^2_{adj} values. The number of catchments (%) with significant R^2_{adj} values is also indicated for each climate index

Of the analysis of Figure 3.9-b1 to b4, it can be deduced that ENSO is the primary index of importance in the DJF RX1day and DJF QX1day predictability for the S-SE NA. The prediction of JJA QX1day for the catchments confined over the W-NW-N-NE regions of NA is dominated by AMO. A combination of two indices (PNA and NAO) emerges as the best predictors for DJF RX1day, DJF QX1day and, to some extent, JJA QX1day over the region of Appalachian and southern Great Lakes. For the other regions, a combination of two or more indices seems to be required for predicting to some extent DJF RX1day, DJF QX1day and JJA QX1day. As for JJA RX1day, there is less spatial coherence between indices, and no single index or combination of indices can be clearly identified for the prediction of this extreme over the NA catchments.

The results provide evidence for the potential linkages of both the AMO and ENSO on hydroclimatic extremes across NA. This is partially confirmed by the literature. Goly et Teegavarapu (2014) indicated a strong influence of AMO and ENSO on seasonal extreme precipitation over Florida. Whan et Zwiers (2017) reported significant influence from ENSO in SW-S-SE on extreme winter precipitation. Over the contiguous US, Mo et al. (2009) explored the influence of ENSO and AMO on droughts, while Gershunov et Barnett (1998); Gershunov et Cayan (2003) investigated the ENSO influence on seasonal extreme precipitation and temperature. Several investigations have been also undertaken over Canada in view of exploring the relationships between some indices and hydroclimatic extremes (Fleming et al., 2006; St. George, 2007; Tan et al., 2016; Wang et al., 2006; Whan & Zwiers, 2017); the authors have reported significant influences of PDO, AO/NAO, ENSO and PNA on extreme precipitation and streamflow over various regions of Canada. However, these influences are not shown in our results. This is probably due to the fact that the hydroclimatic extreme events are not analyzed in a similar manner (e.g., indices, methodology). In addition, to our knowledge, no study has been carried out between the hydroclimatic extremes and the PNA over the mid-E of the US; future investigations should be expanded in that direction.

3.4 Summary and conclusions

The present study provides several new insights into the influences of the primary modes of NCV on recent hydroclimatic variability over 4536 NA catchments. The NCV is explored on the basis of six climate indices, depicting atmospheric-oceanic oscillations from interannual to interdecadal time scales. The analysis is conducted on five seasonal hydroclimatic variables including both mean characteristics and extremes over the 1950-2010 period. While the influence of the indices on hydroclimatic variables at large spatial scales is useful to be understood (Tootle et al., 2005), the analysis at a regional level through the NA continent is equally essential for evaluating the spatio-temporal patterns of the indices and trying to connect influences of one or more indices for explaining the observed hydroclimatic variability. The key results of the study are summarized as follows:

1) The effects of the individual climate indices on the five hydroclimatic variables are stronger and more spatially coherent for the winter than the summer season.
2) The influence of regional climatology restricts the spatial influence of the climate indices on temperature, precipitation and, by extension, on streamflow. The pattern of the influence zones varies spatially according to the temporal frame (i.e., season). This is noticeable when the climate indices are analyzed individually or all-combined.
3) The evaluation of the influence of the climate indices on winter (summer) temperatures indicates substantial linkages with PDO, AO/NAO and PNA (AMO and AO/NAO). The effects are particularly marked for the catchments confined over the SE and NW regions of NA.
4) Compared to temperature, precipitation variability in both seasons is less connected with the climate indices. The winter (summer) precipitation variability is dominated by PDO, ENSO and PNA (NAO and AO), although the connections show less spatial variability than those of temperature. The influence zones are restricted to the catchments located over the Appalachian region, the peninsula of Florida and along the NW-S diagonal.
5) Seasonal streamflow variability is weakly synchronized with the climate indices. A comparison of the water balance components of the catchments reveals that there is an

agreement between the variability of precipitation and streamflow, and some climate indices. Winter (summer) streamflow variability exhibits linkages with ENSO, NAO and PDO (AMO and NAO) for the catchments situated over the Appalachian region, the peninsula of Florida and along the NW-S diagonal (W-N-NE-SE regions).

6) The influence of the all-combined climate indices on precipitation and streamflow extremes, in comparison with the mean characteristics, is clearly less spatially uniform across NA. Despite that, potential influences of AMO, ENSO, PNA and NAO on hydroclimatic extremes emerge for some NA regions. As expected, the strongest linkages of precipitation and streamflow extremes with the all-combined climate indices occur in winter.
7) Combining altogether the climate indices leads to more consistent relationships with the hydroclimatic variables in both seasons. The analysis also underlines the difficulty to identify a single index as a predictor for the ensemble of hydroclimatic variables across NA, because of the large regional effects. This provides indications that the NA hydroclimatic variability is conditioned by several climate indices. The coupled influence of some indices on seasonal hydroclimatic variability for some particular regions, as identified in this work and other studies (e.g., Levine et al., 2017), deserves further investigation. Thus, as it has been communicated by Emerton et al. (2017) for ENSO, the influence of climate indices on hydroclimatic variability is much more complex than the general impression conveyed by the literature.

This work establishes an interesting unseen picture of the potential relationships between the leading climate indices affecting NA variability and the hydroclimatic variables for a large number of catchments over the contiguous NA continent. This will have further implications for a better understanding and enhanced prediction of seasonal hydroclimatic variability that is critical for future efficient management of water resources throughout the Canada and the US.

3.5 Acknowledgments

This work was partially supported by École de technologie supérieure and the Natural Sciences and Engineering Research Council of Canada (NSERC). The data used in this study are listed in the experimental design section and references.

CHAPITRE 4

ROLE OF NATURAL CLIMATE VARIABILITY IN THE DETECTION OF ANTHROPOGENIC CLIMATE CHANGE SIGNAL FOR MEAN AND EXTREME PRECIPITATION AT LOCAL AND REGIONAL SCALES

Jean-Luc Martel[a], Alain Mailhot[b], François Brissette[c], Daniel Caya[d]

[a, c, d] Département de Génie de la construction, École de technologie supérieure
[b] Institut national de la recherche scientifique – Eau, Terre et Environnement

Article publié dans « Journal of Climate », mai 2018

Abstract

Climate change will impact both mean and extreme precipitation, having potentially significant consequences on water resources. The implementation of efficient adaptation measures must rely on the development of reliable projections of future precipitation and on the assessment of their related uncertainty. Natural climate variability is a key uncertainty component, which can result in apparent decadal trends that may be greater or lower than the long-term underlying anthropogenic climate change trend. The goal of the present study is to assess how natural climate variability affects the ability to detect the climate change signal for mean and extreme precipitation. Annual and seasonal total precipitation are used as indicators of the mean, whereas annual and seasonal maximum daily precipitation are used as indicators of extremes. This is done using the CanESM2 50-member and CESM1 40-member large ensembles of simulations over the 1950-2100 period. At the local scale, results indicate that natural climate variability will dominate the uncertainty for annual and seasonal extreme precipitation going up to the end of the century in many parts of the world. The climate change signal can however be reliably detected much earlier at the regional scale for extreme precipitation. In the case of annual and seasonal total precipitation, the climate change signal can be reliably detected at the local scale without resorting to a regional analysis. Nonetheless, natural climate variability can impede the detection of the anthropogenic climate change signal until the mid to late century in many parts of the world for mean and extreme precipitation.

4.1 Introduction

Research conducted in the past decades has emphasized human influence on the climate system through anthropogenic emissions of greenhouse gases (IPCC, 2013). It is also expected that global climate warming will induce significant changes in many parts of the world in the distribution of extremes such as extreme precipitation events, droughts and floods. To ensure public safety, the most important infrastructures are typically designed based on an estimate of the recurrence likelihood of a specific extreme precipitation event (e.g. the 100-year storm). This estimate is itself usually based on available historical annual daily maxima data. Since such infrastructures often have typical lifespans exceeding 75 years, the potential impact of the anthropogenic climate change signal (referred to as *climate change signal* hereafter) on extreme precipitation events has important implications for design practice and public safety.

While the climate change signal needs to be accounted for in design practice, consideration also needs to be given to the inherent chaotic nature of the climate system (i.e., the unforced variability which naturally appears in the climate system, and which will be hereafter referred as *natural variability*). There are many indications that natural variability may mask the climate change signal for short- and long-term precipitation at both the local and regional scales (Deser et al., 2012a; Deser et al., 2012b; Deser et al., 2014; Fischer et al., 2013; Fischer & Knutti, 2014; Fischer et al., 2014; Giorgi & Bi, 2009; Hawkins & Sutton, 2011, 2012; King et al., 2015; Maraun, 2013b; Mora et al., 2013; Sanderson et al., 2018; Thompson et al., 2015). A good example of how natural variability can conceal the climate change signal at the decadal scale is the hiatus in the rise of the global mean surface temperature observed between 1998 and 2012 (Hawkins, Edwards, & McNeall, 2014; IPCC, 2013).

To convince policy makers of the importance of adapting infrastructures to climate change, it is crucial to better understand and explain the influence of natural variability on the climate system. However, the ability to assess natural variability is strongly hampered by the short length of available historical records for key weather variables. An alternative approach is to

study it through simulations of a General Circulation Model (GCM) or of an Earth System Model (ESM). Most published studies use many GCMs and/or ESMs (e.g., models from the Coupled Model Intercomparison Project version 5 - CMIP5; Taylor et al., 2011) to gather a large enough ensemble of models to perform such analyses (Fischer et al., 2014; Giorgi & Bi, 2009; Hawkins & Sutton, 2012; IPCC, 2013; King et al., 2015; Maraun, 2013b; Mora et al., 2013). In many such studies, the concept of Time of Emergence (TOE) is defined to assess the moment when the climate change signal emerges from natural variability (Giorgi & Bi, 2009; Hawkins & Sutton, 2012; IPCC, 2013; King et al., 2015; Maraun, 2013b). Generally, it is defined through a signal-to-noise (S/N) ratio based on a measure of the anthropogenic climate change signal (S) and some measure of natural variability (i.e. noise; N). The TOE is then estimated for each simulation (either from an individual model or from different models), and then some measure of the TOE distribution over all simulations (e.g., mean or median TOE) is used.

Most of these studies look at mean climate variables, and few analyze precipitation extremes under such a framework (Fischer et al., 2014; King et al., 2015; Maraun, 2013b). Most of them, though, share a common limitation in their ability to separate natural variability from inter-model variability (uncertainties) since they combine simulations from various models. To correctly assess the sole impact of natural variability, one must first disentangle the inter-model uncertainties from natural variability (Fischer et al., 2013; Kay et al., 2015).

This can be done using a large ensemble of climate simulations from a single GCM or ESM to assess the simulated natural variability (Fischer et al., 2013; Kay et al., 2015). To date, quite a few studies of this kind using large ensembles have been conducted on mean precipitation (as well as other mean climate variables, Deser et al., 2012a; Deser et al., 2012b; Deser et al., 2014; Fischer et al., 2014; Kay et al., 2015; Sanderson et al., 2018; Thompson et al., 2015). These studies showed that natural variability has a substantial influence over mean precipitation trends at the local and regional scales.

A relatively limited number of studies have been conducted on the influence of natural variability on the detection of climate change signals for precipitation extremes, based on large ensembles of climate simulations from a single model (Fischer et al., 2013; Fischer & Knutti, 2014; Fischer et al., 2014). One of the key findings in these studies is that the signal for precipitation extremes is more robust than that for mean precipitation, indicating a potential earlier emergence of the climate change signal from natural variability in many regions. However, the impact of natural variability on the probability of detecting a climate change signal at the local and regional scales remains a complex problem.

Accordingly, the main objective of the present study is to look at how natural variability could impair the detection of the climate change signal for both precipitation means and extremes at the local and regional scales. This is addressed using two large ensembles of 150-year climate simulations. The models and methods used are developed in Section 4.2. A comparison of model data against observations as well as results for both mean and extreme precipitation are presented in Section 4.3 and discussed in Section 4.4. Concluding remarks are presented in Section 4.5.

4.2 Datasets and methods

4.2.1 The CanESM2 and CESM1 large ensembles

The first large ensemble used in the present study is composed of fifty climate simulations with a 2.8° resolution, derived from the Canadian Centre for Climate Modelling and Analysis (CCCma) second-generation Canadian Earth System Model (CanESM2; Arora et al., 2011; Sigmond & Fyfe, 2016). Five simulations covering the 1850-1950 historical period were performed to generate five different states of the ocean in 1950. Then, ten coupled ocean-atmospheric simulations were run from each of these five historical simulations using randomly perturbed initial conditions (in 1950), for a total of fifty 150-year simulations over the 1950-2100 period. Because of the chaotic nature of the climate system, small perturbations in the initial 1950 conditions quickly resulted in different atmospheric states after a few days following the perturbation (Deser et al., 2012a; IPCC, 2013). The

simulations were conducted from 1950 to 2006 using historical greenhouse gas concentrations data. From 2006 on, the RCP8.5 (Representative Concentration Pathway) scenario resulting in a 8.5 Wm^{-2} increase in the atmospheric radiative forcing in 2100 was used (IPCC, 2013).

The second large ensemble is made up of forty climate simulations with a 1° resolution, derived from the Community Earth System Model version 1 (CESM1) coupled with CAM5.2 for the atmospheric component (Kay et al., 2015). The covered period ranges from 1920 to 2100, but only the 1950-2100 period was analyzed in this study to allow a direct comparison with the CanESM2 large ensemble. The same RCP8.5 scenario was considered from 2006 until the end of the simulation period. Aside from the model structure, the main differences between the two ensemble simulations lie in the spatial resolutions (2.8° for CanESM2 vs 1° for CESM1) and the initial ocean conditions (five different ocean states for CanESM2 vs a single ocean state for CESM1).

4.2.2 Precipitation indices

Two precipitation indices were used in this study: the total wet-day precipitation (PRCPTOT) from days ≥ 1 mm and the max 1-day precipitation amount (RX1day). Both indices were analyzed at the annual and seasonal scales for winter (December-January-February; DJF) and summer (June-July-August).

These two indices are recommended by the Expert Team on Climate Change Detection and Indices (ETCCDI; Klein Tank, Zwiers, & Zhang, 2009; Sillmann, Kharin, Zhang, Zwiers, & Bronaugh, 2013a; Sillmann, Kharin, Zwiers, Zhang, & Bronaugh, 2013b; Zhang et al., 2011). Using the same indices allows a comparison and further discussion of the results obtained here with observed datasets (Donat et al., 2013a; Donat et al., 2013b) and with other climate change studies (e.g. Fischer et al., 2013; Fischer & Knutti, 2014; Fischer et al., 2014; IPCC, 2013). Having both mean and extreme indices furthers our understanding of the role of natural variability in the climate change signal.

4.2.3 Probability of detecting the climate change signal at the local scale

Eleven periods (1950-2000, 1950-2010, 1950-2020, …, 1950-2100) were considered to investigate annual and seasonal time series of PRCPTOT and RX1day indices at each grid point of both ensembles. The non-parametric Theil-Sen estimator (Sen, 1968), which corresponds to the median of the slopes over all pairs of sample points, was used to estimate the slope of a linear trend over each period for all 50 members. This estimator was mainly used to compare observed trends with the simulated trends of both ensembles (see Section 4.3.1). The local trend significance of each grid point was estimated using the non-parametric Mann-Kendall test (Kendall, 1975) at a 95% confidence level:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{m} sign(x_i - x_j) \tag{4.1}$$

where x is the index value (i.e. PRCPTOT or RX1day) at time i and j, with $sign(\ )$ being equal to +1 if x_i is greater than x_j and -1 if x_i is smaller than x_j. S represents the number of times x_i is greater than x_j minus the number of times x_i is smaller than x_j. The sign of S also indicates the sign of the trend.

The Mann-Kendall test was used to characterize the climate change signal at the local scale (i.e. over a given grid point without considering regional spatial correlations) over the corresponding periods. The probability of detecting the climate change signal for a given period was then defined by the percentage of members with a significant trend of a given sign (positive or negative) at the 95% confidence level. The eleven pre-defined periods allowed the investigation of the probability of locally detecting the climate change signal over the eleven periods.

An advantage of using these two tests is that they do not make assumptions about the distribution of the analyzed variable and they can be applied to both observed and simulated series. When dealing with recorded series, the Theil-Sen estimator and Mann-Kendall test are

often used to detect the non-stationarity associated with the climate change signal (Donat et al., 2013b; Lins & Slack, 1999; Westra et al., 2013).

The 90% Detection Decade (90%DD) was defined as the decade ending the first period (e.g., decade 2060-2070 of the 1950-2070 period), where at least 45 out of 50 members for CanESM2 or 36 out of 40 members for CESM1 (therefore, a 90% probability of detecting the trend among the various simulations) had a significant trend (95% confidence level) of the same sign (either positive or negative) over that period and over all subsequent periods up to the 1950-2100 period (in our example the trend must remain over the 1950-2080, 1950-2090 and 1950-2100 periods). The 45 members of the CanESM2 (36 members for CESM1) thresholds were chosen such that the probability of having 5 members (4 members) with a non-significant trend due to type II errors (false negatives) was less than 5%. The 90%DD was estimated using the annual index series and the seasonal index series. The 90%DD is, to some extent, related to the Time of Emergence (TOE) used in previous studies (Giorgi & Bi, 2009; Hawkins & Sutton, 2012; IPCC, 2013; King et al., 2015; Maraun, 2013b). Results shown hereafter, based on the local trends analysis, are referred to as the "local scale."

An example of an estimated 90%DD is shown in Figure 4.1 for the land grid point containing the city of Toronto, Canada, for the RX1day index. In this example, the 90%DD is the 2090-2100 decade. Two main features can be observed in Figure 4.1 for both ensembles (a- CanESM2 and b- CESM1) as the length of data increases: 1) the distribution becomes narrower, and 2) there is a shift in the central value of the distribution. This suggests that, when using a smaller number of decades, natural variability has a greater influence on the detected trend resulting in a wider distribution. However, when a greater number of decades is used, the distribution becomes narrower as the signal increases and the influence of natural variability on the trends decreases. Moreover, as the climate change signal becomes stronger, the central value of the distribution shifts to the right.

There was a possibility of inaccurate results being obtained when the estimated 90% probability of detecting the climate change signal was reached near the end of the 1950-2100

period, since it could theoretically have fallen below the 90% threshold in the decades after 2100. This situation was investigated by looking at the probability of a grid point that had reached the 90% probability threshold before 2100 dropping back below the threshold of 45/50 members for CanESM2 or 36/40 members for CESM1 in any subsequent periods. The probability of occurrence of such cases was estimated to average 0.0103 for PRCPTOT and 0.0039 for RX1day over all land grid points (for both ensembles and for annual and seasonal scales). It would therefore be very unlikely that grid points with a reported 90%DD before 2100 would be changed beyond 2100.

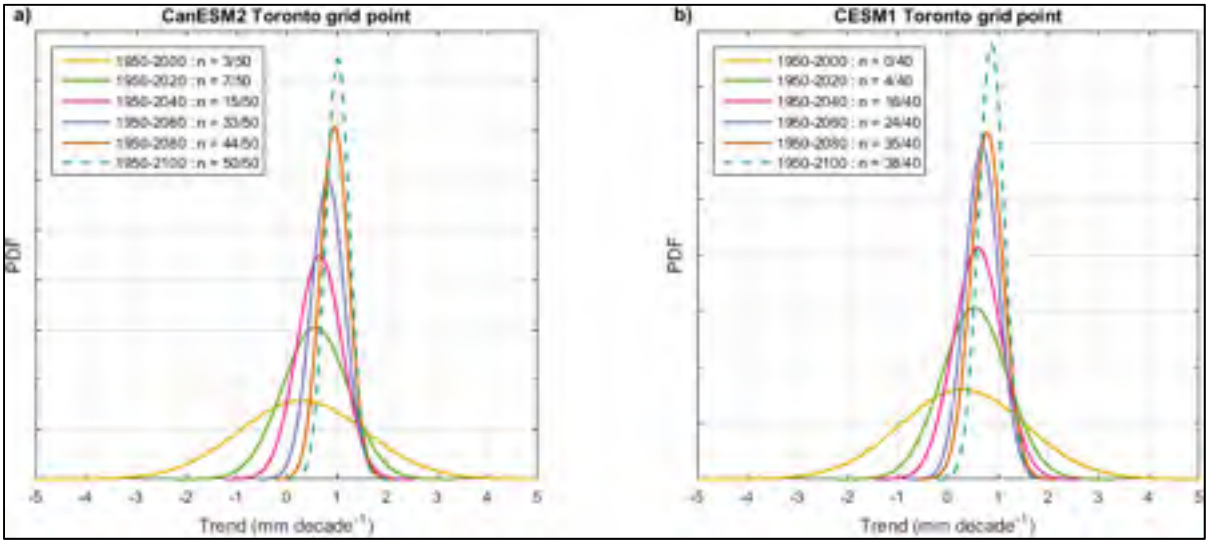

Figure 4.1 Gaussian distribution of the non-parametric Theil-Sen trend estimators for the grid point corresponding to the city of Toronto, Canada, for the RX1day index estimated from a) 50-member CanESM2 ensemble and b) 40-member CESM1 ensemble. Continuous lines correspond to the distributions for periods with fewer than 45 members with a 95% significant trend and dashed lines correspond to periods with 45 or more members with a 95% significant trend of the same sign. The number of members with a significant trend (*n*) is shown in the legend for each period

4.2.4 Probability of detecting the climate change signal at the regional scale

The methodology described in Section 4.2.3 does not take into consideration a possible spatial correlation between neighboring grid points. It is expected that if grid point values are spatially correlated, this could result in earlier 90%DD than expected at the local scale.

To investigate regional trends, a field significance test combined with a resampling approach by bootstrap is performed over each grid point of both ensembles. The method proposed for assessing the regional trend significance is also described in Douglas et al. (2000); Kiktev et al. (2003); Westra et al. (2013). Figure 4.2 describes the method through an example using the grid point containing the city of Toronto, Canada for the RX1day index with the CanESM2 ensemble.

Regions were defined in CanESM2 by using the nearest neighboring grid points for each grid point (3 × 3 = 9 total grid points). In CESM1, a relatively similar surface area was selected to allow a fair comparison with CanESM2 results. 81 grid points (9 × 9 = 81 grid points) were taken for each region. The result of the test was associated to the middle grid point of each region.

The regional average Mann-Kendall's S ($\bar{S}_m$) is then computed as the average of the local trend values from each grid point within the region:

$$\bar{S}_m = \frac{1}{m}\sum_{k=1}^{m} S_k \tag{4.2}$$

where S_k is the Mann-Kendall S (see equation (4.14.1) for the k^{th} grid point in a region of m grid point (m = 9 for CanESM2 and m = 81 for CESM1).

To determine whether or not the regional trend is significant, a bootstrap resampling approach was performed (Douglas et al., 2000). For each bootstrap sample, a sample of years with replacement corresponding to the period analyzed (i.e. 1950-2000, 1950-2010, … 1950-2100) was randomly generated (Figure 4.2). The same sample of years was then used for each grid points of the region to compute the Mann-Kendall's S metric (equation 4.1). Using the same years allows us to keep track of the spatial correlation between neighboring grid points. The regional average Mann-Kendall $\bar{S}_m$ is then computed using Equation 4.2. This

procedure is repeated 1 000 times and sorted in ascending order of S assigning a non-exceedance probability based on the Weibull plotting position formula:

$$P = \frac{r}{B+1} \tag{4.3}$$

where r is the rank of each sample and B = 1 000 (1 000 samples). The 95% confidence level of the empirical CDF obtained is then defined as the Mann-Kendall $\bar{S}_m$ associated with the 25th rank (α = 0.025; negative significant trend) and the 975th rank (α = 0.975; positive significant trend).

This methodology is then repeated using all available members of both ensembles for each of the 11 periods. As for the local trend analysis described in Section 4.2.3, the 90%DD is defined as the decade ending the first period where at least 45 out of 50 members for CanESM2 (36 out of 40 members for CESM1) had a significant trend at the 95% confidence level of the same sign over that period and over all subsequent periods. Finally, the methodology is reproduced over all grid points using the same sample of years for the bootstrap. Results shown hereafter, based on the regional trends described in this section, are referred to as the "regional scale."

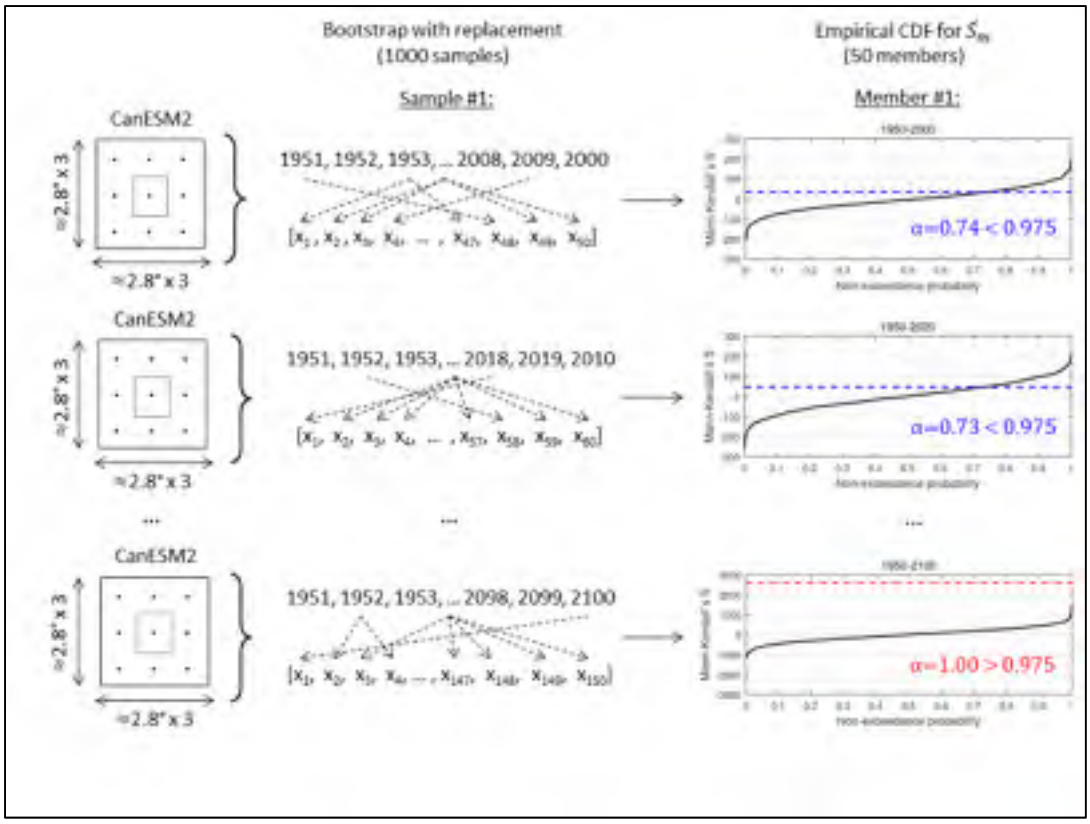

Figure 4.2 Summary of the methodology used to assess regional trend (see Section 4.2.4). The grid point containing the city of Toronto, Canada in CanESM2 is used in this example

The proposed regional trend analysis is based on the hypothesis that PRCPTOT and RX1day annual series are temporally uncorrelated. The median value of the lag-1 autocorrelation coefficient across all land grid points over the 1950-2100 period for annual values (similar values for DJF and JJA) was equal to 0.011 for the PRCPTOT index and -0.010 for the RX1day index in the CanESM2 ensemble and 0.046 and -0.004 for the CESM1 ensemble. Autocorrelations were computed on the residuals from a linear regression. These small values suggest that, on average, the hypothesis of temporal independence is valid for both indices. Nonetheless, the field significance resampling approach was also performed using a moving-block bootstrap method to account for autocorrelations (Wilks, 1997, 2011). A moving-block of 2 years was used in the bootstrapping (which was above the median obtained for both indices and both ensembles). The results were consistent with that obtained under the hypothesis of temporal independence (not shown for conciseness).

4.2.5 Global region analysis

The analyses described in the previous two sections were performed globally and then using the 21 geographical regions listed in Table 4.1 and shown in Figure 4.3. These 21 geographical regions were also used by Giorgi et Francisco (2000); Sanderson et al. (2018); Sillmann et al. (2013a); Sillmann et al. (2013b). An analysis of the combined land grid points from these 21 regions (LGP, excluding Antarctica) is also shown.

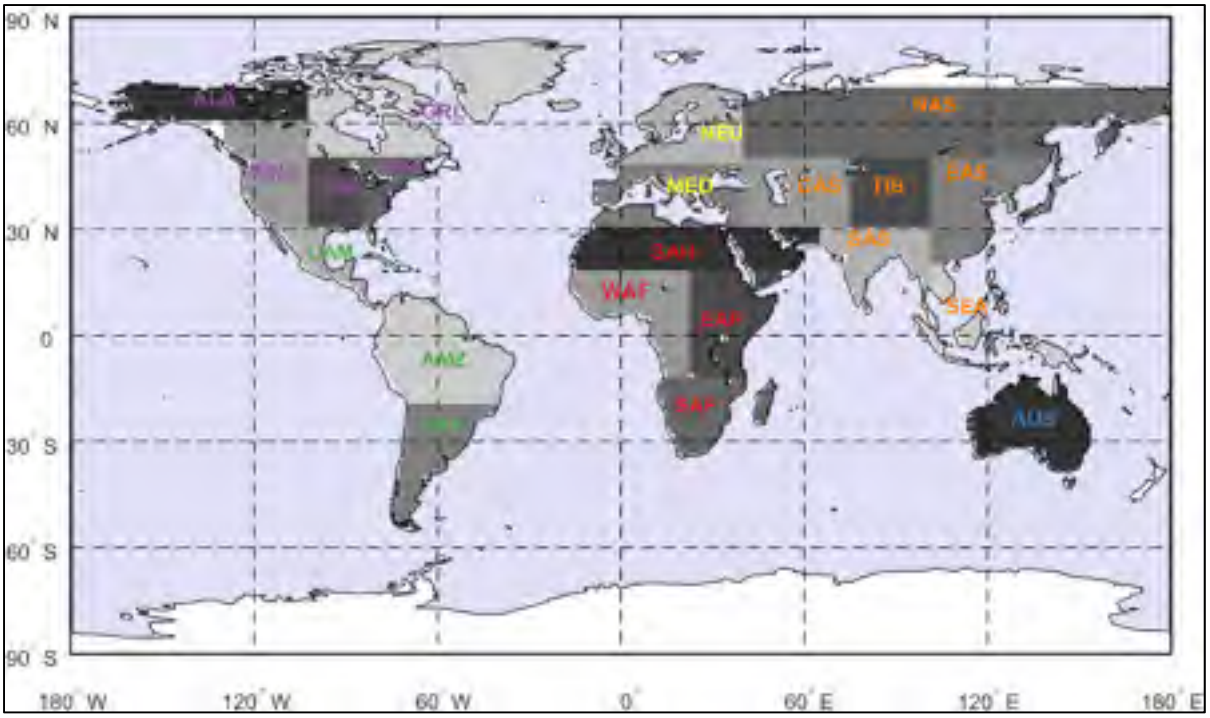

Figure 4.3 Map of the 21 geographical regions used in this study. Label colors represent the respective continents: Oceania (blue), South America (green), North America (purple), Europe (yellow), Africa (red) and Asia (orange). Only land grid points were considered in the regional analysis (see Table 4.1) (Giorgi & Francisco, 2000; Sillmann et al., 2013a; Sillmann et al., 2013b). Grid points correspond to CESM1 grid points

Table 4.1 List of the 21 regions used in this study. The number of land grid points available in CanESM2, CESM1, HadEX2 and GHCNDEX gridded datasets for the PRCPTOT (PRCP.) and RX1day indices over the 1950-2010 period are also shown and will be further discussed in Section 4.3. Only HadEX2 and GHCNDEX grid points with at least 40 years of data available over the 1950-2010 period were considered

Name	Acronym	Latitude (°)	Longitude (°)	Number of land grid points					
				CanESM2	CESM1	HadEX2		GHCNDEX	
				Both indices	Both indices	PRCP.	RX1day	PRCP.	RX1day
Australia	AUS	45S-11S	110E-155E	89	587	42	18	108	86
Amazon Basin	AMZ	20S-122N	82W-34W	127	856	32	9	39	15
Southern South America	SSA	56S-20S	76W-40W	70	462	50	23	36	17
Central America	CAM	10N-30N	116W-83W	32	183	29	17	35	14
Western North America	WNA	30N-60N	30W-103W	89	553	75	59	109	84
Central North America	CNA	30N-50N	103W-85W	41	301	43	42	68	85
Eastern North America	ENA	25N-50N	85W-60W	34	215	29	27	52	47
Alaska	ALA	60N-72N	170W-103W	68	503	71	7	80	10
Greenland	GRL	50N-85N	103W-10W	186	1272	88	14	84	15
Mediterranean Basin	MED	30N-48N	10W-40E	64	460	55	35	76	45
Northern Europe	NEU	48N-75N	10W-40E	85	569	67	67	110	103
Western Africa	WAF	12S-18N	20W-22E	71	551	10	1	0	0
Eastern Africa	EAF	12S-18N	22E-52E	82	571	5	1	0	0
Southern Africa	SAF	35S-12S	10W-52E	66	417	35	19	47	22
Sahara	SAH	18N-30N	20W-65E	127	743	11	1	7	1
Southeast Asia	SEA	11S-20N	95E-155E	41	293	22	13	21	8
East Asia	EAS	20N-50N	100E-145E	103	693	95	67	140	90
South Asia	SAS	5N-30N	65E-100E	58	391	46	25	21	6
Central Asia	CAS	30N-50N	40E-75E	77	557	78	18	95	26
Tibet	TIB	30N-50N	75E-100E	63	438	62	22	81	12
North Asia	NAS	50N-70N	40E-180E	296	1995	277	119	420	203
Land grid points	LGP	-	-	1869	12610	1222	604	1629	889

4.3 Results

4.3.1 Representation of natural variability in CanESM2 and CESM1

Since the representation of natural variability in CanESM2 (resolution of 2.8° latitude × 2.8° longitude) and CESM1 (resolution of 1° latitude × 1° longitude) is a key

element of the present study, variability in trends in both ensembles is compared to corresponding values in the observed HadEX2 (Donat et al., 2013b; resolution of 2.5° latitude × 3.75° longitude) and GHCNDEX (Donat et al., 2013a; resolution of 2.5° latitude × 2.5° longitude) gridded datasets. These two datasets have different spatial and temporal coverage due to the different data sources used and quality control performed (Dittus, Karoly, Lewis, & Alexander, 2015). There is also a larger number of grid points available for the PRCPTOT index compared to the RX1day index in both HadEX2 and GHCNDEX datasets due to the interpolation technique used to create these datasets (Donat et al., 2013a; Donat et al., 2013b).

Only grid points with at least 40 (out of 60) years over the 1950-2010 period were considered for the observed datasets (resulting in a total of 1222 grid points and 1629 grid points for the PRCPTOT index and 604 grid points and 889 grid points for the RX1day index for HadEX2 and GHCNDEX respectively). The number of land grid points within each of the 21 analyzed regions analyzed is shown in Table 4.1.

The performance of the CanESM2 and CESM1 ensembles is first assessed through the comparison of the 60-year annual mean and annual standard deviation (1950-2010) of the PRCPTOT and RX1day indices with the HadEX2 and GHCNDEX datasets (Figure 4.4 and Figure 4.5). For both ensembles, the median of the distribution of annual mean and annual standard deviation values (i.e. one value for each member over the 1950-2010 period) at each grid point was considered.

As shown in maps on the left-hand side of Figure 4.4, the spatial distribution of the annual mean PRCPTOT values is globally well reproduced by both ensembles when compared to the HadEX2 and GHCNDEX datasets. Similarly, both ensembles capture relatively well the observed spatial pattern of annual standard deviation as shown by the maps on the right-hand side of Figure 4.4.

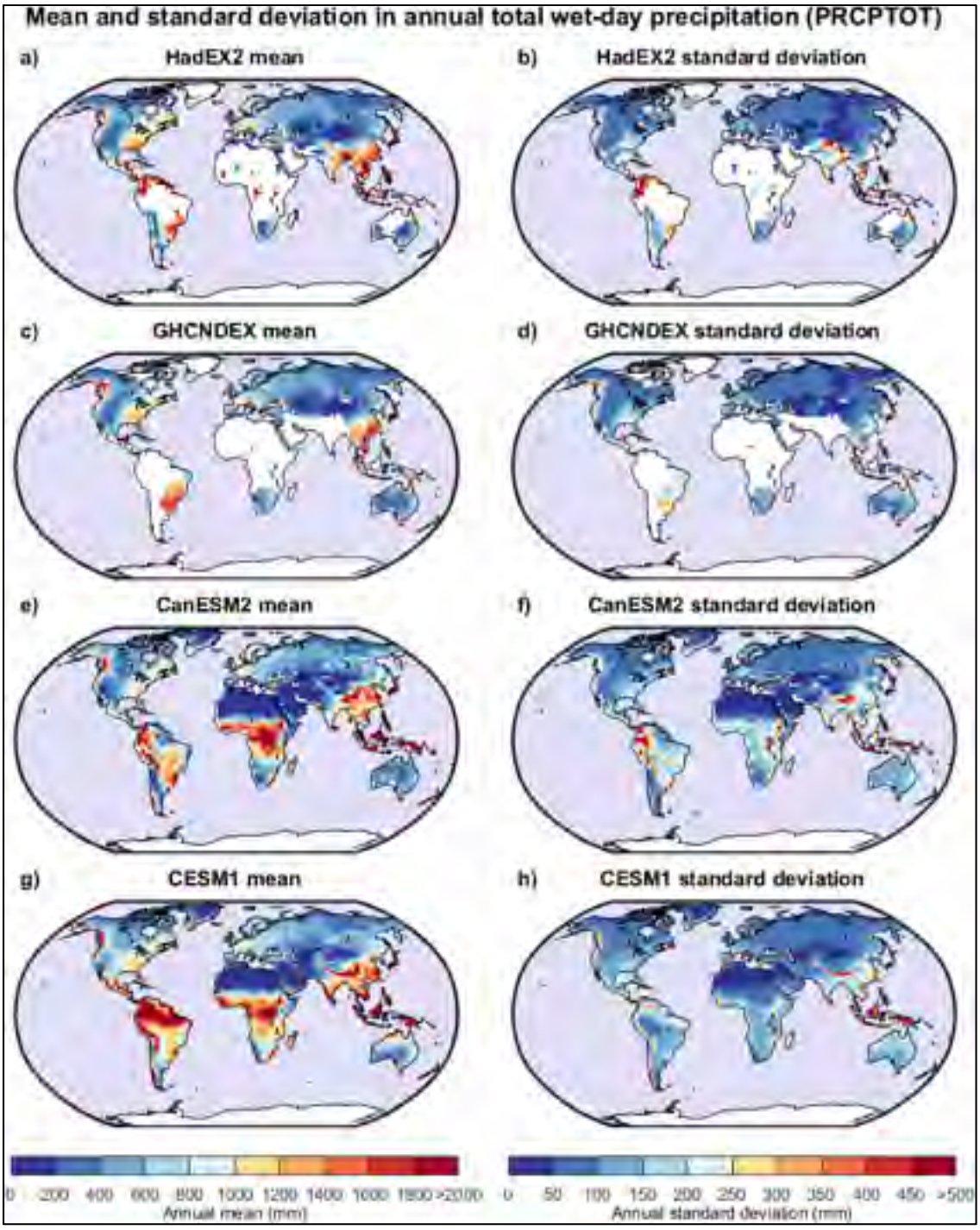

Figure 4.4 Comparison of the mean and standard deviation of annual total wet-day precipitation (PRCPTOT) over the 1950-2010 period. Shown are the observed (a),(b) HadEX2 and (c),(d) GHCNDEX datasets, and the median value over the (e),(f) CanESM2 50-member ensemble and (g),(h) CESM1 40-member ensemble

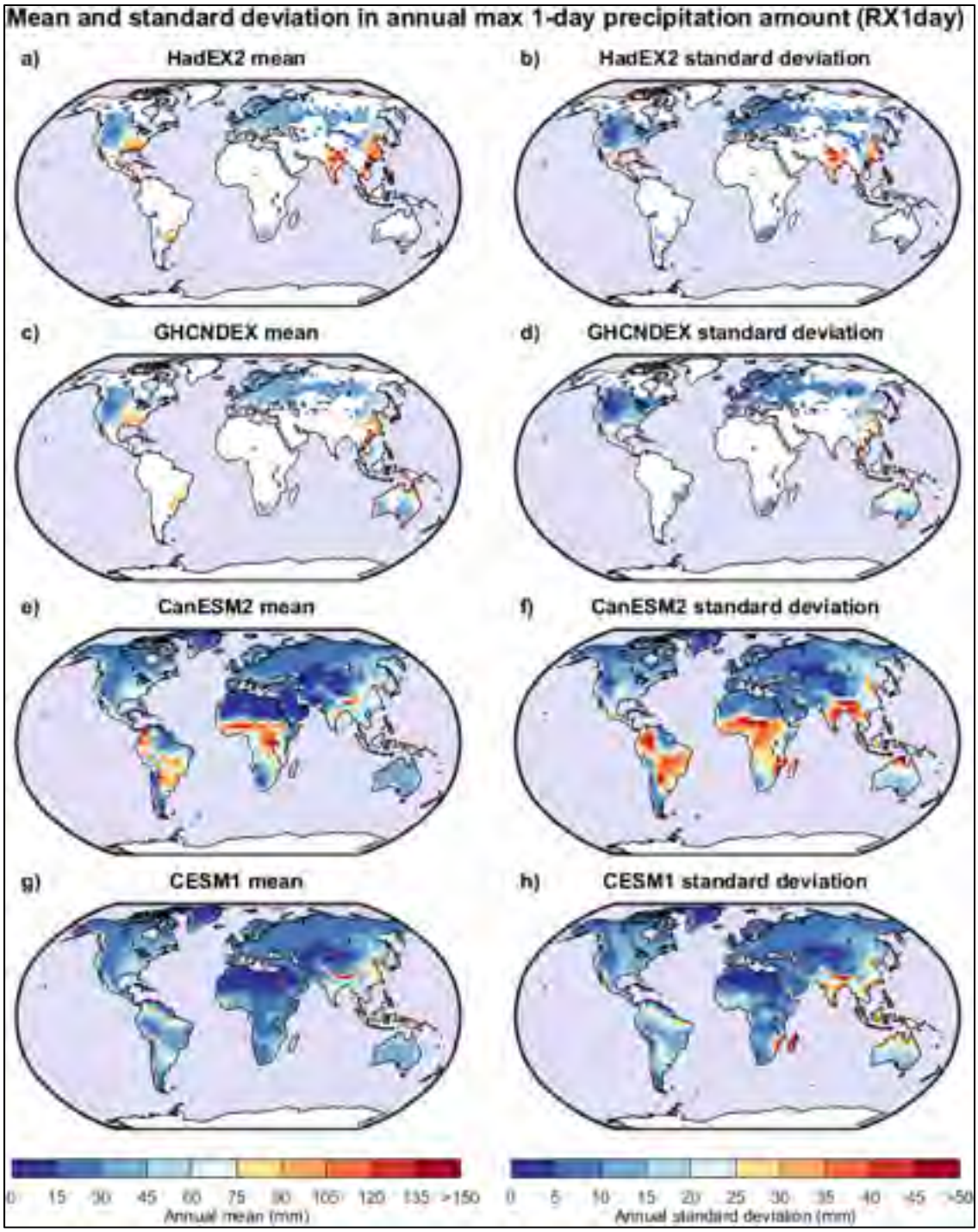

Figure 4.5 Same as Figure 4.4, but for RX1day

Mean annual values of the RX1day index (maps on the left-hand side of Figure 4.5) are generally underestimated by both ensembles when compared to the observed datasets. Such results were however expected because of the spatial mismatch between the ensembles resolution and the smoothed grid point estimates constructed in the HadEX2 and GHCNDEX datasets (Sillmann et al., 2013a). However, the inter-annual variability as estimated by the annual standard deviation is well captured by both CanESM2 and CESM1 ensembles (maps on the right-hand side of Figure 4.5).

Trends estimated by the Theil-Sen estimator and the Mann-Kendall test for the PRCPTOT and RX1day annual time series over the 1950-2010 period were also compared. Figure 4.6 (PRCPTOT) and Figure 4.7 (RX1day) show maps of land grid points comparing local linear trend values from HadEX2 and GHCNDEX datasets to the member with the smallest, median, and largest global trend (defined as the median of the distribution of trends over all grid points) for both ensembles.

As seen for the PRCPTOT index (Figure 4.6) and RX1day index (Figure 4.7), a larger number of grid points displayed a significant trend for PRCPTOT than for RX1day (44.5% versus 16.3% for HadEX2 and 45.8% versus 14.8% for GHCNDEX). A similar behaviour was observed for individual members of both ensembles. However, as shown in Figure 4.6 and Figure 4.7, there is a much smaller fraction of grid points with a significant trend in the different members of both ensembles as compared with observations when comparing the same areas. These results outline the stronger influence of natural variability at the local scale for the RX1day index and the ability of the two ensembles to reproduce this behaviour. The selected individual members also highlight the large range of possible local trends (individual grid points). This range is due to the uncertainty related to natural variability, which can even span negative and positive trends at a given grid-point for various members.

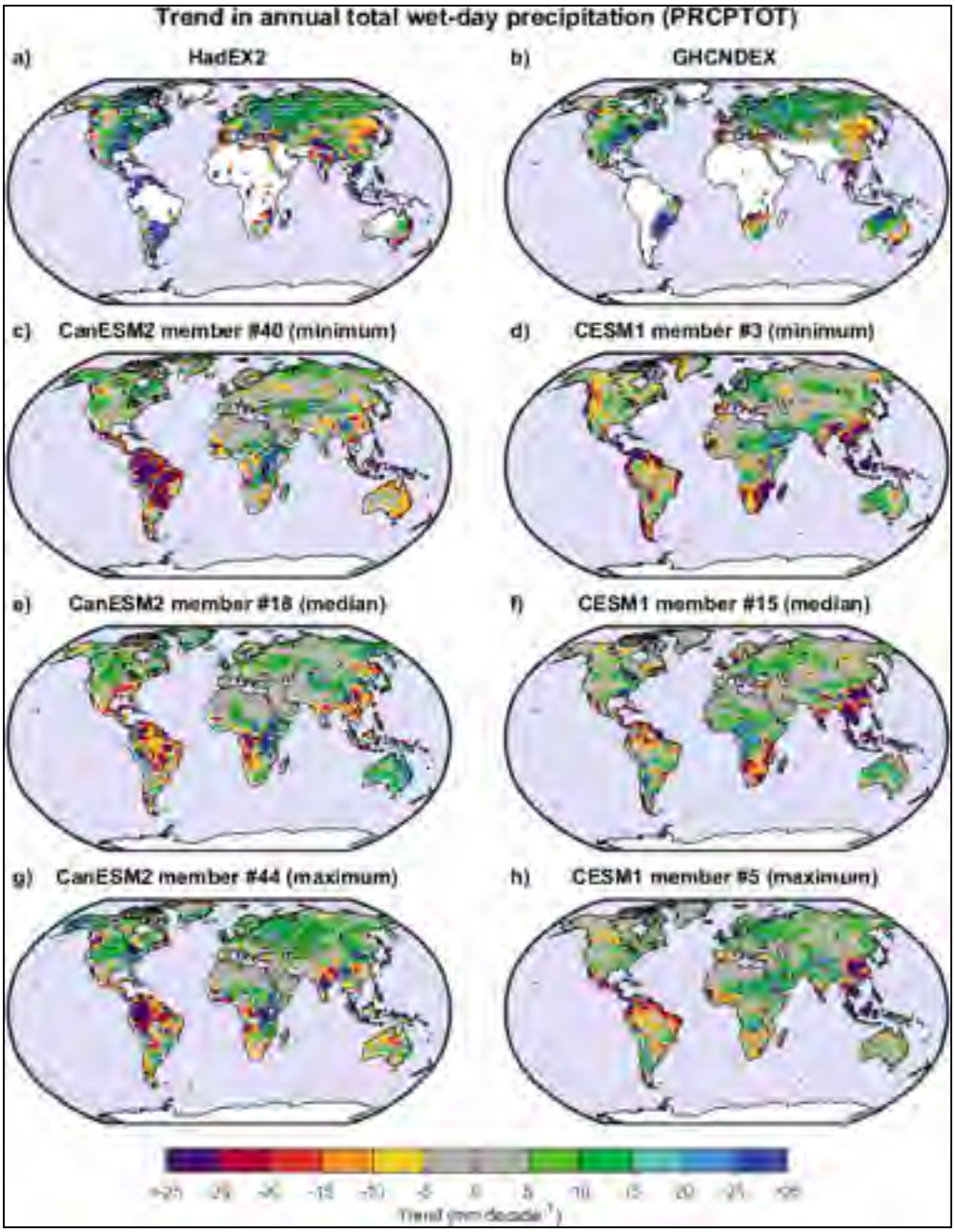

Figure 4.6 Linear trends in annual total wet-day precipitation (PRCPTOT) over the 1950-2010 period as defined by the Theil-Sen estimator. The first row represents observed (a) HadEX2 and (b) GHCNDEX dataset decadal trends. Grid points where fewer than 40 out of 60 years of data were available over the 1950-2010 period are shown in white. The remaining rows represent decadal trend for two individual members corresponding to the (c),(d) smallest, (e),(f) median, and (g),(h) largest median global trend value in the CanESM2 50-member ensemble and CESM1 40-member ensemble, respectively

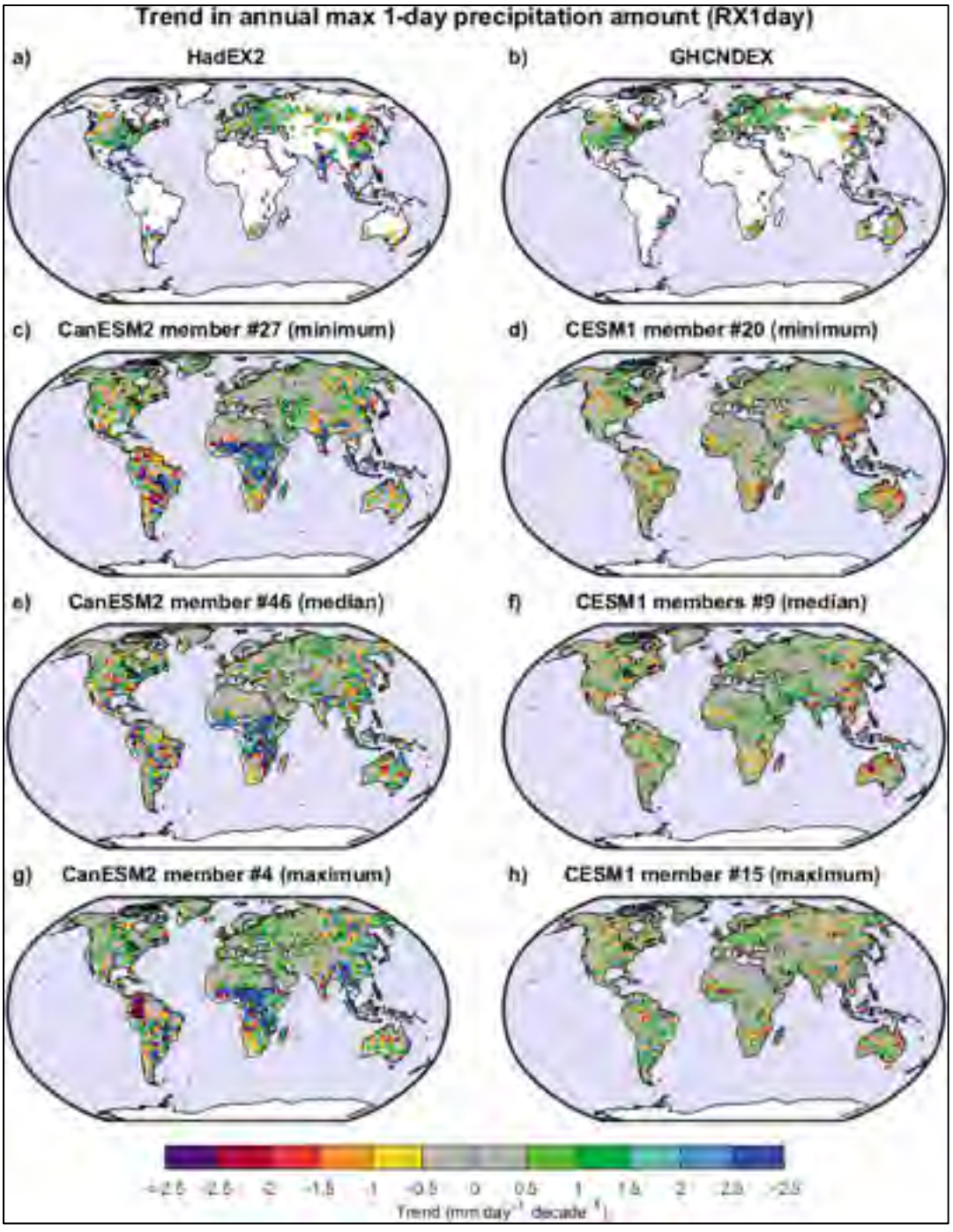

Figure 4.7 Same as Figure 4.6, but for RX1day

4.3.2 PRCPTOT index

An analysis of the 90%DD for annual and seasonal total precipitation (using the PRCPTOT index) allows an overview of how natural variability affects the detection of the climate change signal in both the CanESM2 and CESM1 ensembles. Figure 4.8 (local scale) and Figure 4.9 (regional scale) show maps of the decade in which the PRCPTOT index reaches 90%DD. Figure 4.10 (local scale) and Figure 4.11 (regional scale) give a more detailed analysis of these results over the 21 geographical regions listed in Table 4.1.

A global comparison between Figure 4.8 and Figure 4.9 suggests that there is a relatively good agreement between both ensembles for both the annual and seasonal scales. Figure 4.8 indicate that the PRCPTOT 90%DD based on local trends occurs before the end of the century over large fractions of ocean and land surface areas, especially at higher latitudes and over the tropics. The seasonal analysis of 90%DD for DJF and JJA shows a later detection than in the annual case. These results show that the likelihood of detecting a significant signal (stippled regions) is greater at the annual scale than at the seasonal scale for most regions. Figure 4.9 shows very similar results for the regional trends analysis based on field significance resampling approach. Overall, the 90%DD is reached somewhat earlier (slightly darker colors) and there is less noise in the maps as compared to the results obtained at the local scale.

The spatial patterns of average trend signs tend to be similar over both the annual scale and DJF, but differ in JJA. For instance, average trends are of different sign over most parts of Europe and North America, where more negative trends are observed for JJA as compared to the annual scale and DJF (Figure 4.8 and Figure 4.9 panels e and f). Overall, for CanESM2 (CESM1), there are 75.8% (76.4%) of all grid points with a positive trend, 70.8% (74.2%) for DJF and 68.5% (68.7%) for JJA.

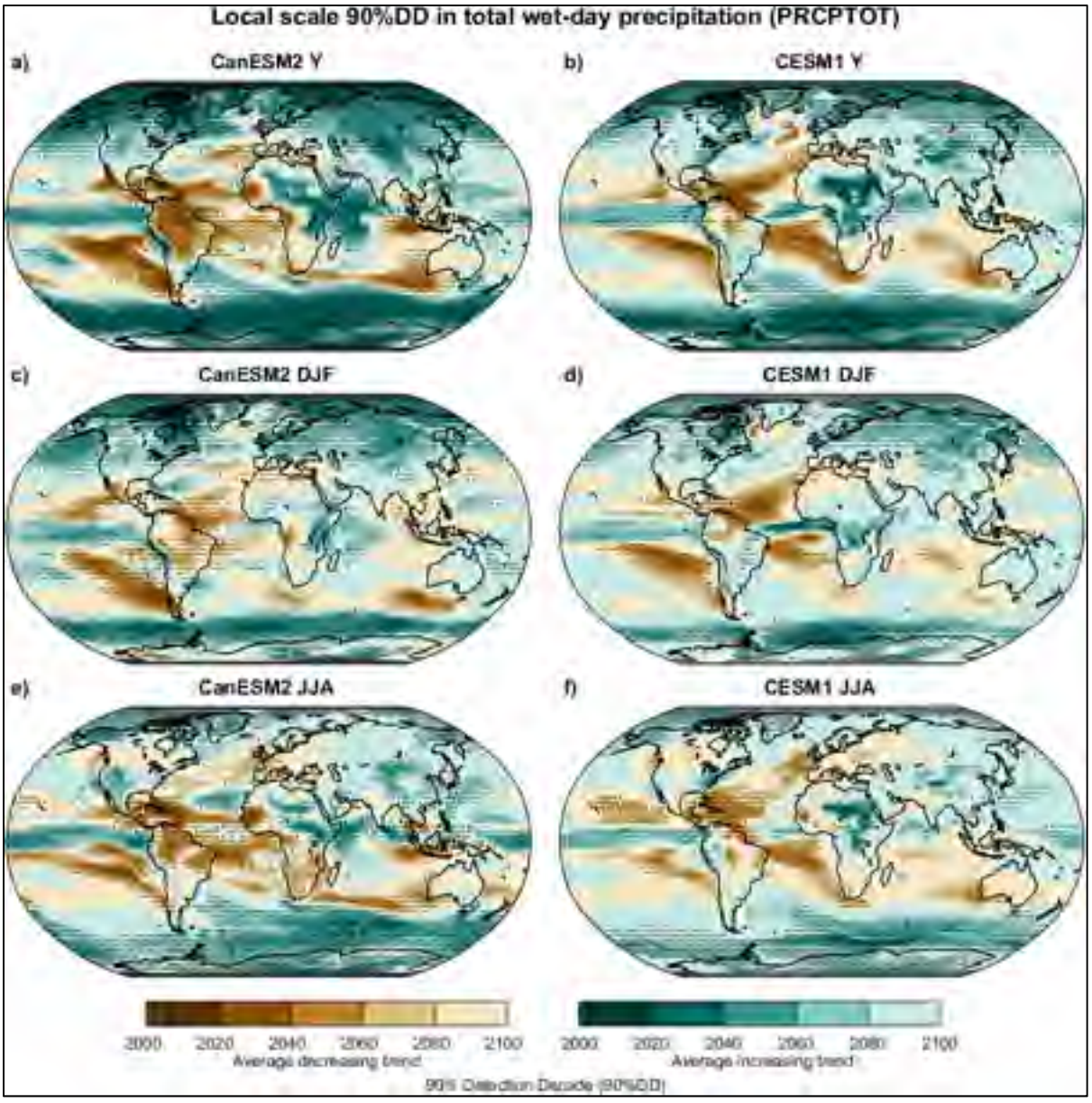

Figure 4.8 Global local trends analysis (i.e. corresponding to each grid point) of the estimated 90%DD for the PRCPTOT index, showing results based on (left) the CanESM2 ensemble and (right) the CESM1 ensemble for the (top)-(bottom) annual (Y), DJF and JJA scales. The brown colors represent an average decreasing trend while the blue-teal colors indicate an average increasing trend. The stippled patterns identify grid points where the estimated 90%DD occurred before 2100

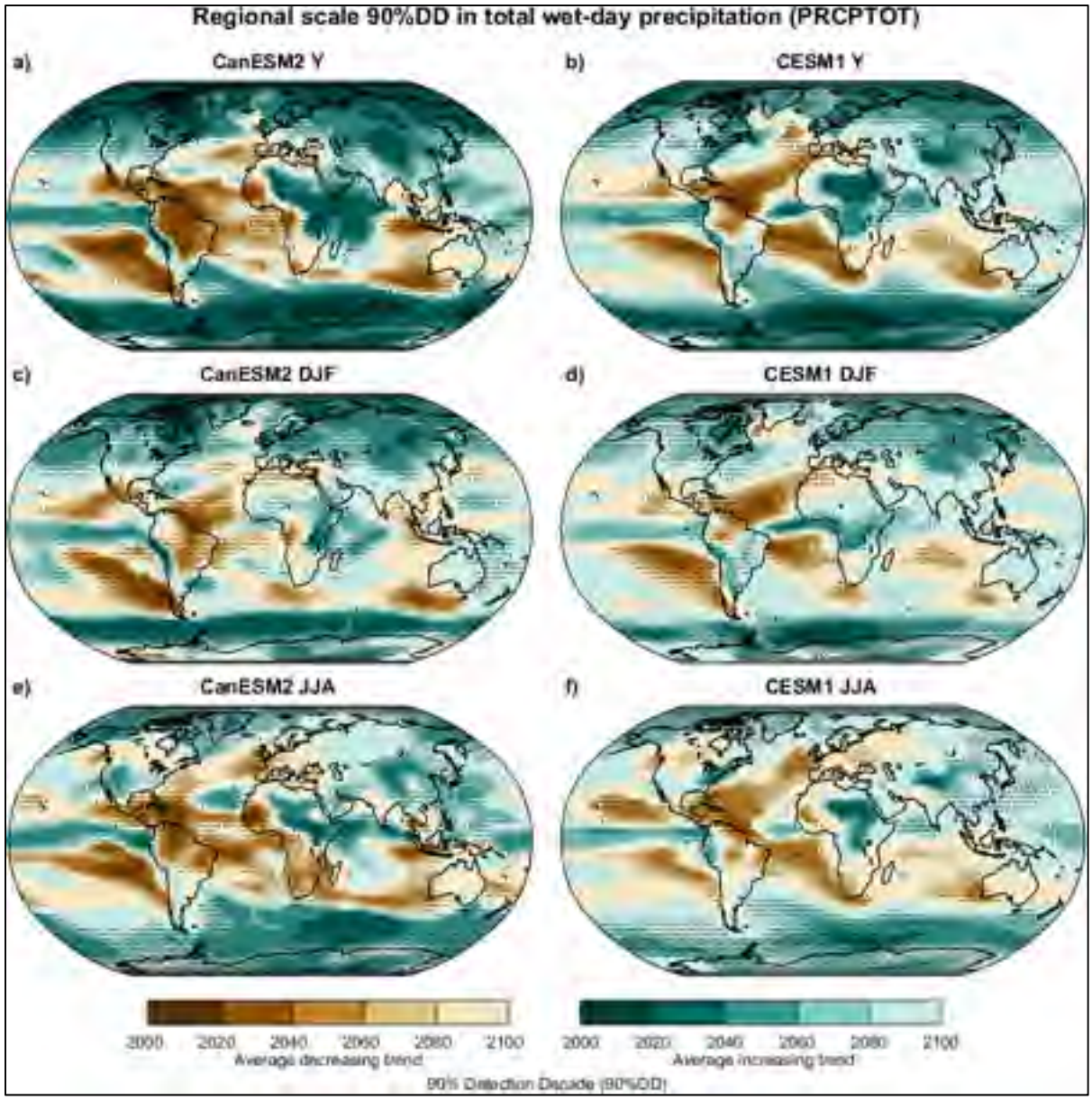

Figure 4.9 Global regional trends analysis (i.e. computed with the field significance resampling approach) of the estimated 90%DD for the PRCPTOT index, showing results based on (left) the CanESM2 ensemble (region size of 3 × 3 grid points) and (right) the CESM1 ensemble (region size of 9 × 9 grid points), for the (top)-(bottom) annual (Y), DJF, and JJA scales. The brown colors represent an average decreasing trend while the blue-teal colors indicate an average increasing trend

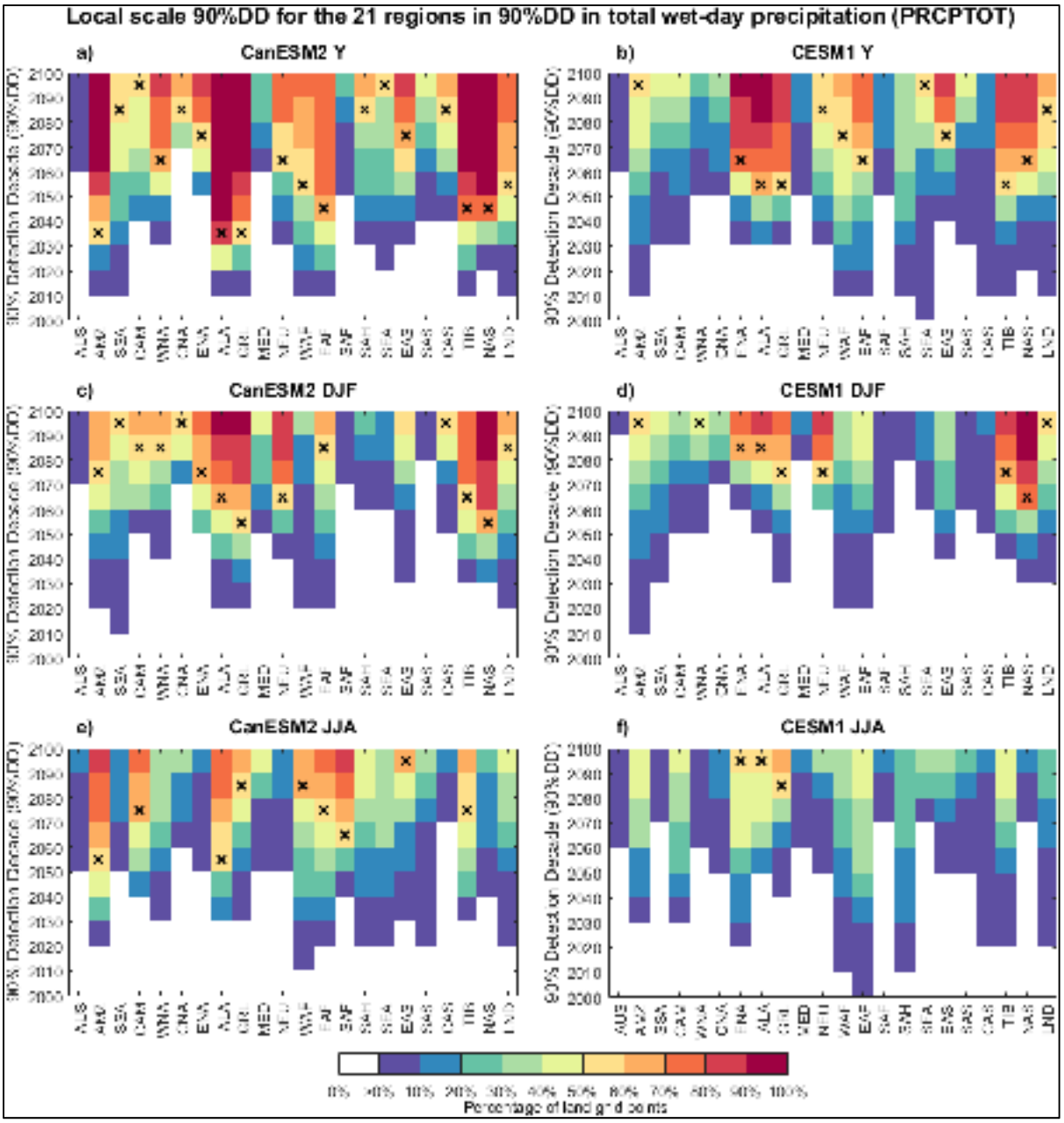

Figure 4.10 Cumulative percentage of land grid points within each of the 21 regions (listed in Table 4.1) with the local trend analysis (i.e. corresponding to each grid point) estimated 90%DD occurring in each decade for the PRCPTOT index, showing results based on (left) the CanESM2 ensemble (region size of 3 × 3 grid points) and (right) the CESM1 ensemble (region size of 9 × 9 grid points), for the (top)-(bottom) annual (Y), DJF and JJA scales. The black "x" indicates the decade when more than 50% of the region's land grid points reached the 90% probability of detecting the climate change signal. The white boxes correspond to regions that had no 90%DD at any grid point in (and prior to) that decade

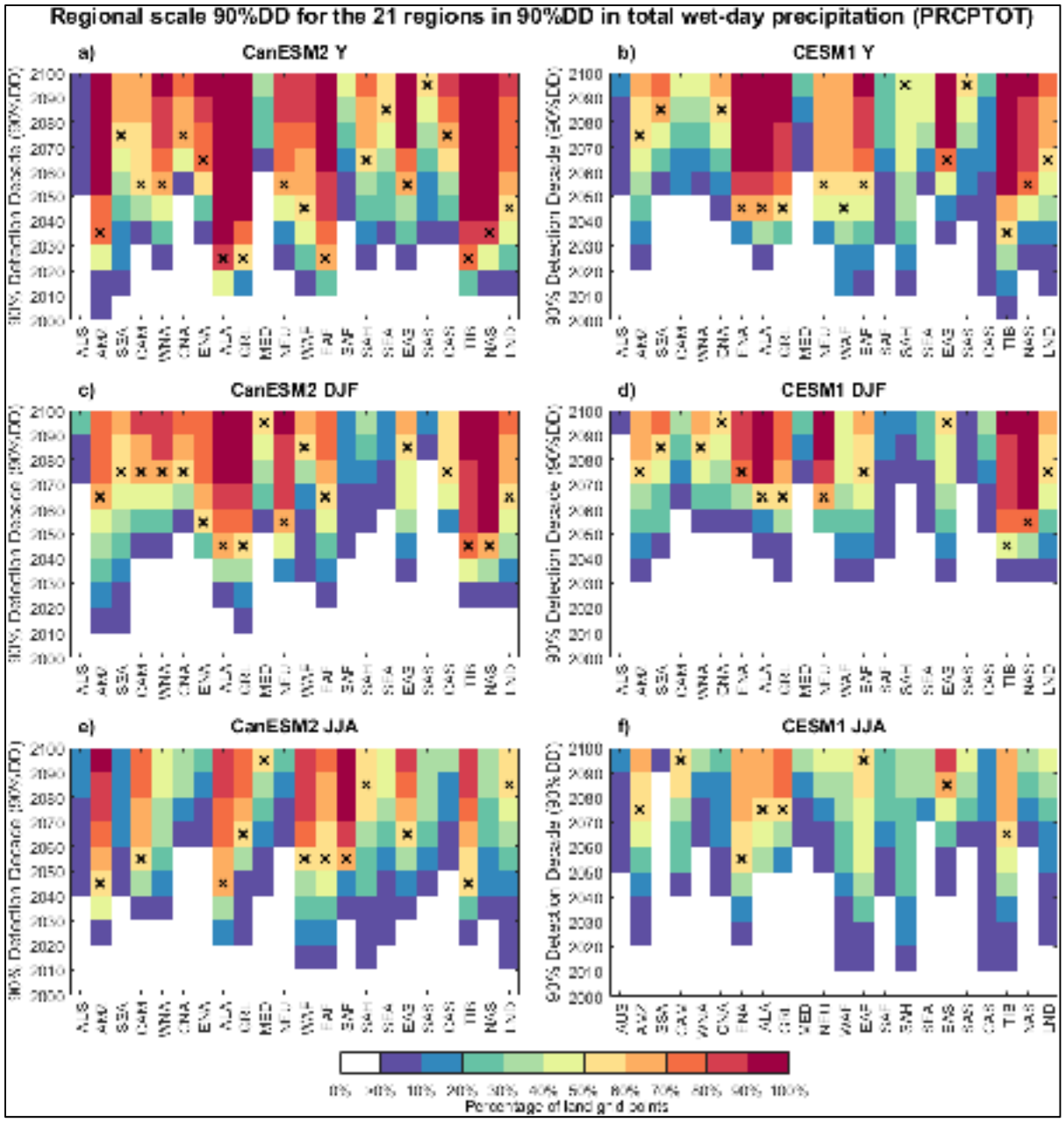

Figure 4.11 Cumulative percentage of land grid points within each of the 21 regions (listed in Table 4.1) with the regional trend analysis (i.e. computed with the field significance resampling approach) estimated 90%DD occurring in each decade for the PRCPTOT index, showing results based on (left) the CanESM2 ensemble (region size of 3 × 3 grid points) and (right) the CESM1 ensemble (region size of 9 × 9 grid points), for the (top)-(bottom) annual (Y), DJF and JJA scales. The black "x" indicates the decade when more than 50% of the region's land grid points reached the 90% probability of detecting the climate change signal. The white boxes correspond to regions that had no 90%DD at any grid point in (and prior to) that decade

As shown in the left-hand side panels of Figure 4.10 and Figure 4.11 for CanESM2, 17 (18) regions out of 21 have 50% of their land grid points with 90%DD occurring prior to 2100 at the annual scale, 13 (16) regions for DJF and 9 (11) regions for JJA based on the local (regional) scale. Not a single region crosses this 50% of land grid points threshold before 2040 at the local scale and 2030 at the regional scale (and for most regions this will only occur a few decades later) at both the annual and seasonal scales.

For CESM1, the 90%DD is reached later than for CanESM2 as shown in the right-hand side panels of Figure 4.10 and Figure 4.11. A total of 12 (15) regions out of 21 have 50% of their land grid points reach their 90%DD prior to 2100 at the annual scale, 9 (13) regions for DJF and 3 (8) regions for JJA at the local (regional) spatial scale. Not a single region crosses the threshold before 2060 for the local and 2050 for regional trends (except for TIB) at both the annual and seasonal scales (two decades later than for CanESM2). On average, the threshold where 50% of the regions' land grid points reach their 90%DD in CESM1 is 1.6 decades later than for CanESM2 for annual, 1.6 for DJF and 0.9 for JJA at the regional scale.

Despite CESM1 having a later 90%DD than CanESM2, as well as some differences in their spatial patterns (see Figure 4.8 and Figure 4.9), both ensembles agree in many respects. The 90%DD is reached earlier at the regional scale for all 21 geographical regions and at the global land scale. The regions with the earliest 90%DD are TIB, the tropical zones (Amazon Basin – AMZ (except for CESM1), Eastern Africa – EAF and Western Africa – WAF) and high latitude zones above the 50th parallel (Alaska – ALA, Greenland – GRL, and North Asia – NAS). Eastern North America (ENA) is also one of the region with the earliest 90%DD with CESM1, but this is not as clear for CanESM2. At the annual scale, a clear climate change signal emerges worldwide for the PRCPTOT index, except for Australia (AUS), the Mediterranean Basin (MED), South Africa (SAF) and the South Asia (SAS) regions. When looking at DJF and JJA, the climate change signal emerges later. By the end of the century, the climate change signal will most likely be detected in many regions of the world at the local or regional scales for this index.

4.3.3 RX1day index

A 90%DD analysis was also realized for precipitation extremes (using the RX1day index). Figure 4.12 (local scale) and Figure 4.13 (regional scale) show maps of the 90%DD, while Figure 4.14 (local scale) and Figure 4.15 (regional scale) show the results for the 21 geographical regions.

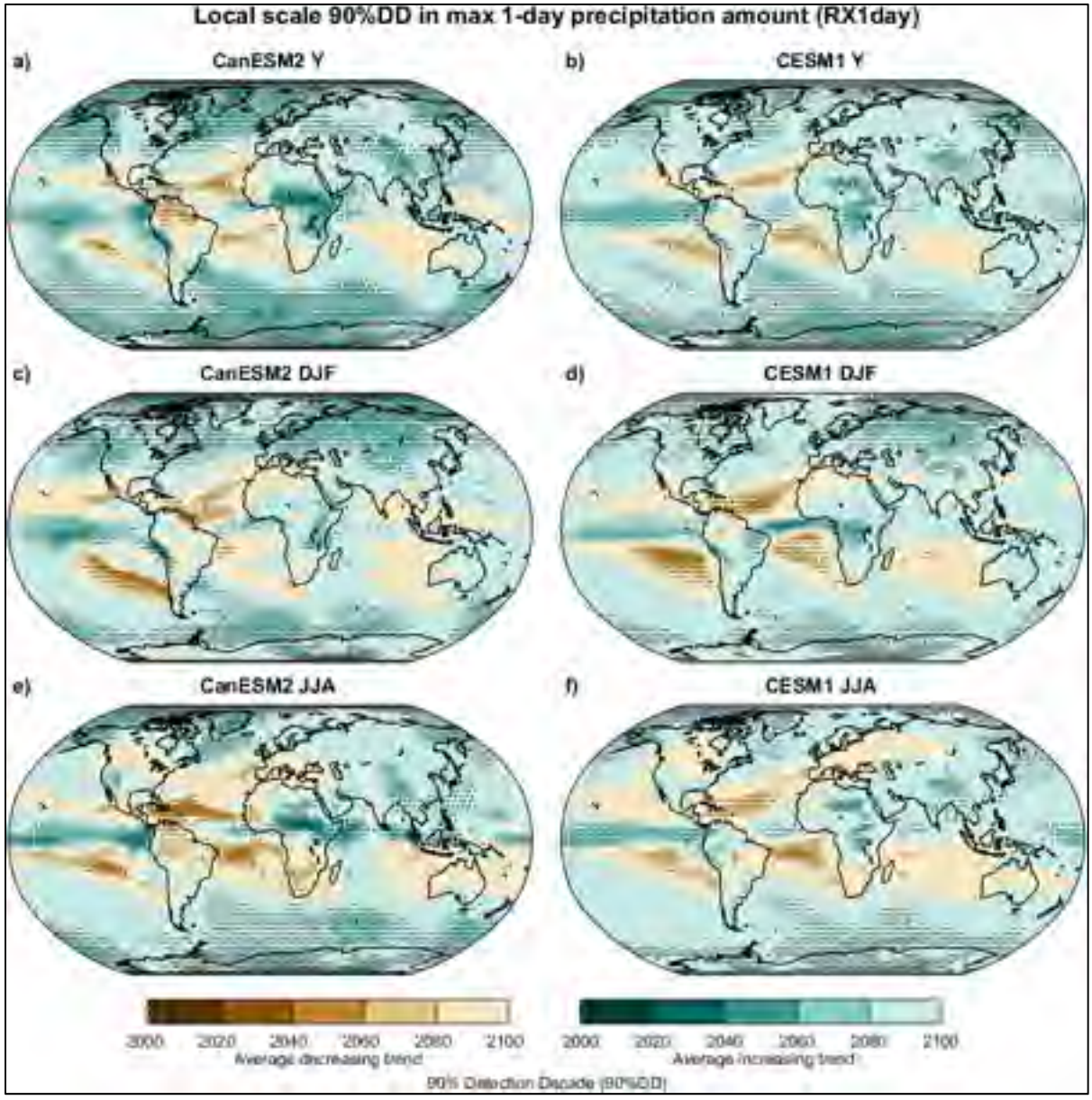

Figure 4.12 Same as Figure 4.8, for RX1day index

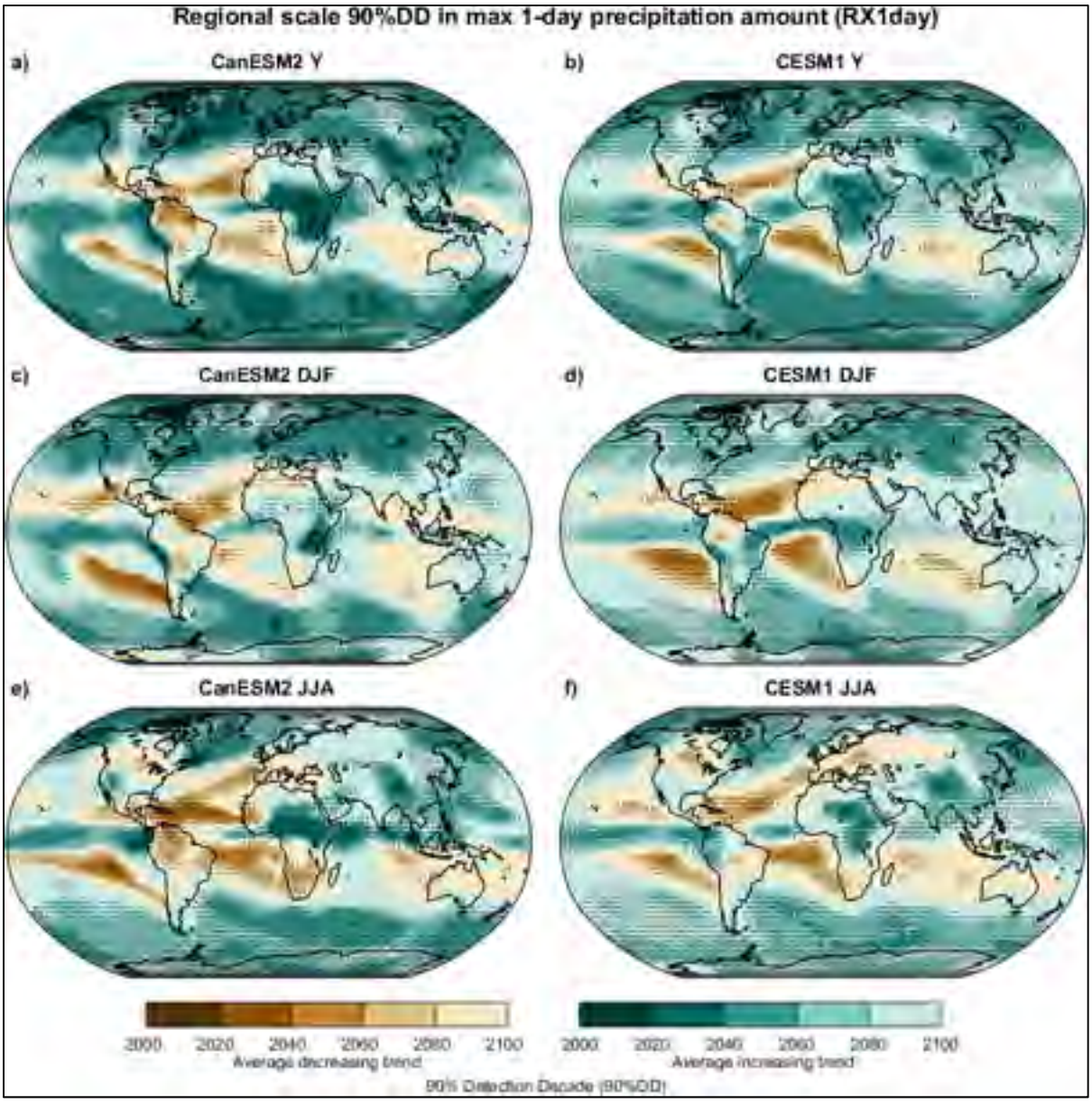

Figure 4.13 Same as Figure 4.9, for RX1day index

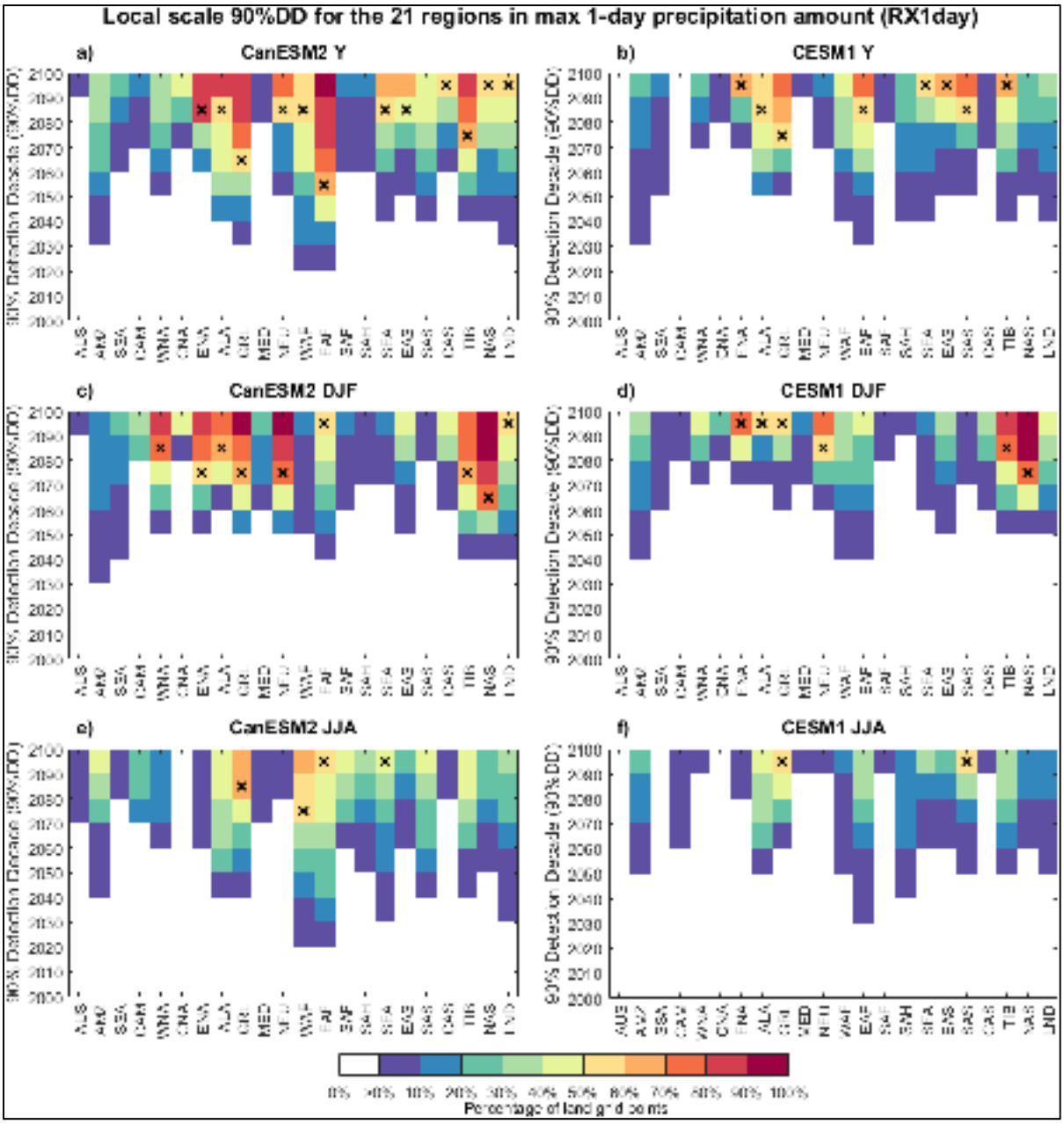

Figure 4.14 Same as Figure 4.10, for RX1day index

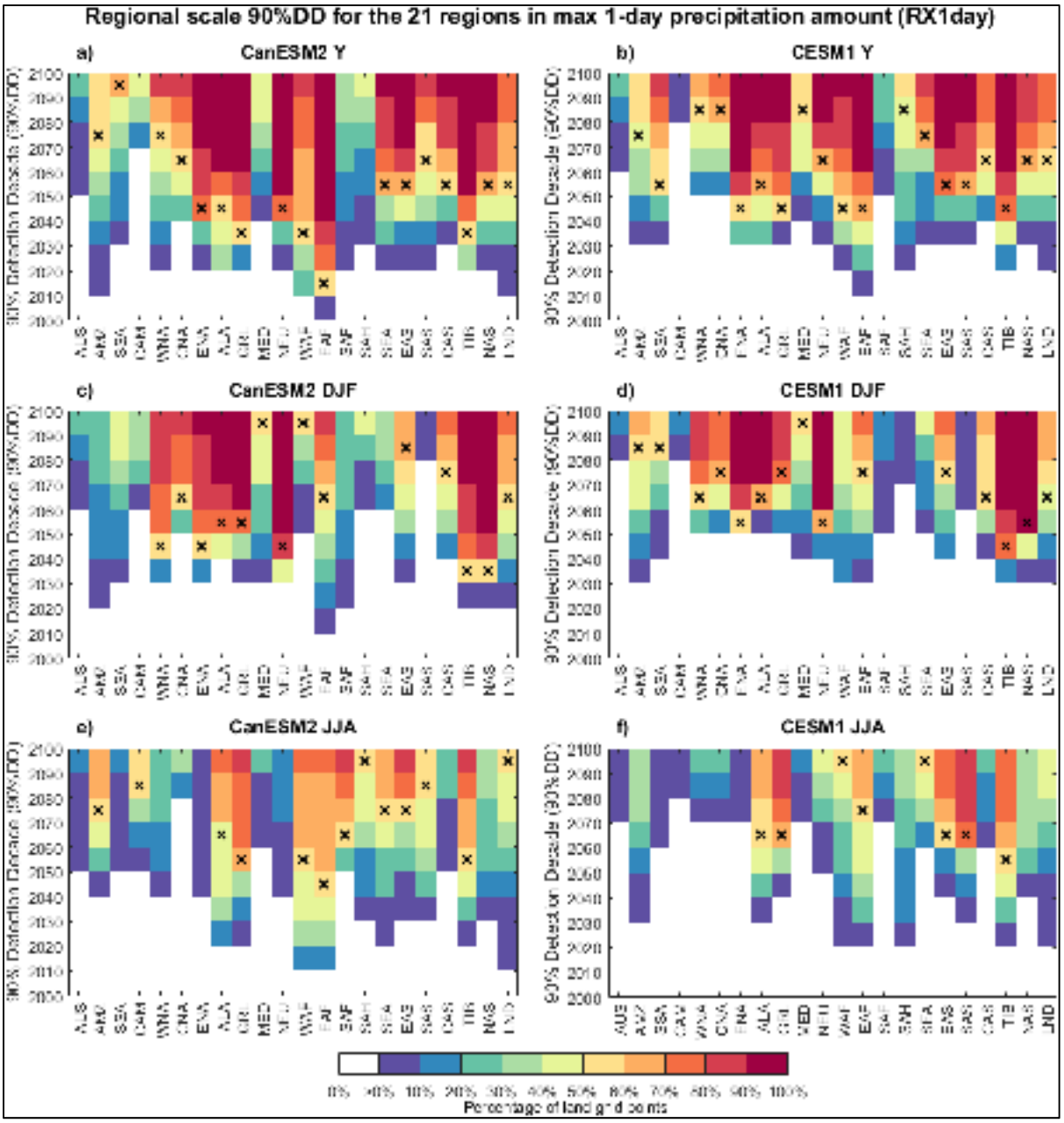

Figure 4.15 Same as Figure 4.11, for RX1day index

The comparison between Figure 4.12 and Figure 4.13 indicates that differences between both ensembles is much smaller than for the PRCPTOT index. Figure 4.12 shows that the local scale results have a much larger fraction of both oceans and land surface areas which do not reach the 90%DD by the end of the simulation (non-stippled areas) for both annual and

seasonal scales. However, Figure 4.13 shows that the 90%DD occur earlier at the regional scale.

One distinctive feature here is that a larger number of average positive trends is observed for RX1day than for PRCPTOT. The percentage of all grid points showing a positive trend for CanESM2 (CESM1) is 86.4% (90.3%) at the annual scale, 80.5% (83.3%) for DJF and 77.1% (77.6%) for JJA. As was the case for PRCPTOT, the spatial patterns are similar for the annual scale and DJF, but notable differences are seen for JJA. Negative trends are observed across large parts of Europe and North America at the JJA scale for RX1day (Figure 4.12 and Figure 4.13 panels e and f).

As shown in the left-hand side panels of Figure 4.14 and Figure 4.15 for CanESM2, 11 (21) regions out of 21 have 50% of their land grid points with a 90% probability of detecting the climate change signal before the end of the simulation at the annual scale, 8 (13) regions for DJF, and 4 (12) for JJA at the local (regional) scale. The threshold of 50% of land grid points was not achieved for any of the 21 geographical regions before 2050 (and most regions beyond that decade) at the local scale, and 2030 at the regional scale (with the exception for EAF at the annual scale, where it reached as early as 2010).

For the CESM1 ensemble, the 90%DD was also reached slightly later than for CanESM2, as shown in the right-hand side panels of Figure 4.14 and Figure 4.15. Overall, 8 (18) regions out of 21 reached the same threshold at the annual scale, 6 (14) regions for DJF and 2 (8) regions for JJA. For this ensemble, the regions that have 50% of their land grid points reaching the 90%DD the earliest were two high-latitude regions beyond 50°N: GRL (2070 for annual) and NAS (2070 for DJF) at the local scale. For the regional scale, the earliest was 2040 for 5 regions at the annual scale (ENA, GRL, WAF, EAF and TIB), for TIB only for DJF and 2050 in the TIB region also for JJA. On average, the threshold where 50% of the land grid points reached their 90%DD in CESM1 is 0.4 decades later than for CanESM2 at the annual scale, 0.3 for DJF and 0.8 for JJA at the regional scale.

Similarly to the PRCPTOT index, the geographical regions with the earliest 90%DD are also consistent for both ensembles when looking at the regional scale. This is also reflected in the combined land grid points (LGP), where we see a similar percentage of grid points reaching 90%DD globally. The regions with the earliest 90%DD for the RX1day index are the tropical zones (EAF and WAF), high latitude zones above the 50th parallel (ALA, GRL, and NAS), regions affected by monsoons (SAS, East Asia – EAS, and TIB) and ENA, which is affected by hurricanes. These regions share in common the fact that an increase in warming will likely result in a robust climate change signal for RX1day. A later 90%DD is expected at the seasonal scale.

4.4 Discussion

4.4.1 Validation of both ensembles

The comparison with observations suggest that the spatial pattern of inter-annual variability and mean PRCPTOT index values, and to a lesser extent, of the RX1day index, as simulated by both CanESM2 and CESM1 ensembles are globally in agreement with corresponding patterns of the observed HadEX2 and GHCNDEX datasets. Differences can be partly explained by natural variability, as the distribution of annual mean and standard deviation over the various members can be quite dispersed, especially for grid points displaying large inter-annual variability (see Figure S4.16 for PRCPTOT and Figure S4.17 for RX1day). These discrepancies can also be due to biases in both ensembles and also from sampling errors and uncertainties in HadEX2 and GHCNDEX datasets.

Furthermore, the comparison of trends between models and observations for both indices suggests that it is difficult to directly compare the global spatial distribution of trends obtained by the different members of each ensemble to observed trends. When comparing one realization (the observed recent past) against a probabilistic distribution (ensemble members), the best possible outcome is to frame this realization within the possible predicted range according to the expected statistical frequency. However, the large variability of trends extracted for each ensemble members demonstrate the challenge of detecting the climate

change signal at the local scale. Comparison of observed and simulated trends was only achieved at the local scales.

A comparison for each region listed in Table 4.1 could also have been performed, but a qualitative analysis of all members for each ensemble (not shown due to lack of space) clearly outlined a very large inter-member variability at the scale of the regions and would not have changed the above conclusion. Other difficulties when dealing with local and regional comparisons arise from the different sources of uncertainty in observation datasets, such as short observational records, homogeneity problems and missing data (Hegerl et al., 2015). Furthermore, since gridded observed datasets are typically constructed by interpolated point values (e.g. station), various upscaling/downscaling problems are always present (Avila et al., 2015; Chen & Knutson, 2008; Herold, Behrangi, & Alexander, 2017; Sillmann et al., 2013b).

The 90%DD is shown to be conservative estimate as it corresponds to the decade where the climate change signal is detected in most of simulations from a large ensemble of simulations, as compared to a "single realization" of the climate system when dealing with the real world. This is well illustrated with Supplementary materials (Figure S4.18 to Figure S4.21), showing the probability of detecting a significant trend during a given decade at both the local and regional scales. While the probability increases overall as we move further into the 21st century, very high probabilities are only reached after the mid-century, and even later for many grid points. For the 1950-2010 period, this probability remains relatively low for most regions.

A limited qualitative comparison of CanESM2 and CESM1 against the CMIP5 multi-model mean signal was made to frame the general behavior of both climate models against other GCMs/ESMs. Globally, spatial patterns of increasing and decreasing trends match the multi-model average changes obtained by Sillmann et al. (2013a) for both annual total and extreme precipitation indices. Furthermore, both the sign of the change and robustness of the climate change signal (characterized in this study by an early 90%DD) match the signal obtained by

(Fischer et al., 2014) remarkably well (especially for RX1day) in regions where at least 12 out of 15 CMIP5 models agreed on the direction of change. The regions with the most robust climate change signal for precipitation extremes obtained by Scoccimarro, Gualdi, Bellucci, Zampieri, et Navarra (2013) are consistent with the regions with the earliest 90%DD for the majority of land grid points for both ensembles. There are no reasons to assume that the conclusions drawn from both ensembles would be markedly different when using another GCM/ESM.

4.4.2 Impact of natural variability at the local and regional scales

Local, as well as regional trend based analyses were performed to determine how the spatial correlation affects results for the PRCPTOT and RX1day indices. In general, the field significance resampling approach showed that a more robust climate change signal can be detected from natural variability at the regional scale as compared to the local scale. Supplementary materials (Figures S4.18 to S4.21) showed that the increasing probability of detecting a significant trend is initially larger and grows faster at the regional scale.

For the PRCPTOT index, results from both the local and regional scales are quite similar for the annual and seasonal scales. This suggests that mean precipitation trends can likely be detected at the local scale. However, for RX1day, spatial dependence was shown to have a great influence, as the results for the regional scale were markedly different from those at the local scale. Figures 4.14 and 4.15 clearly show the difference between the local and regional scales for RX1day.

These results show that when investigating extreme precipitation at the local scale, it is likely that natural variability will strongly impede the detection of a statistically significant climate change signal over a long period. Overall, this is also in agreement with Fischer et al. (2013), who concluded that it is not possible to provide stakeholders with reliable information for changes in extreme precipitation when investigating at the local scale.

Westra et al. (2013) investigated trends on the HadEX2 dataset for the RX1day index using a field significance resampling approach. The areas that showed the most significant trends were the United States, Europe, South Africa and some parts of India and South-East Asia. With the exception of South Africa, the results obtained here for these areas (Figure 4.12 and Figure 4.13) also showed a relatively early 90%DD, corresponding to areas with a robust climate change signal.

Further comparisons were made using the CESM1 ensemble to investigate the effect of using an increasing region size in the field significance resampling approach. Regions made of 1 (1 × 1), 9 (3 × 3), 25 (5 × 5), 49 (7 × 7) and 81 (9 × 9) grid points were used for this analysis. Results can be seen on Supplementary materials – Figure S4.22 (PRCPTOT) and Figure S4.23 (RX1day). Overall, for RX1day, the results indicate a convergence around the 5 × 5 domain, with minor changes seen as we move to a larger domain. As for the PRCPTOT index, there was no significant difference at any of the sizes tested, which is consistent with the previously discussed results. It is expected that using a larger region would eventually lead to an overlap of wetter and dryer regions which could impair the ability to detect trends at the regional scale.

4.4.3 Impact of natural variability on PRCPTOT and RX1day indices

The discussion from the previous section clearly outlines one of the main differences between both indices, which is the strong influence of natural variability at the local scale for RX1day, and its much smaller influence for the PRCPTOT index.

Fischer et Knutti (2014); and Fischer et al. (2014) show that there is a greater expectation of extreme precipitation to emerge from natural variability than mean precipitation. They argue that natural variability is indeed greater in the case of extreme precipitation. However, this difference is likely due because precipitation extremes respond more strongly to global warming than does mean precipitation (Fischer & Knutti, 2014; Fischer et al., 2014). Results in Table 4.2 show a comparison between both indices of the percentage of grid points that

have reached their 90%DD before the end of the century. When looking at the local scale, we see that the RX1day index has fewer grid points reaching their 90%DD as compared to PRCPTOT (e.g., the CESM1 annual scale LGP percentage for PRCPTOT is 57.3% against 38.9% for RX1day). However, at the regional scale, the RX1day ends up with a larger number of grid points reaching their 90%DD than the PRCPTOT index (e.g., the CESM1 annual scale LGP percentage for PRCPTOT is 67.1% against 81.5% for RX1day). These results indicate that the climate change signal for RX1day is indeed more robust than for PRCPTOT at the global scale, which is in agreement with previous studies.

Table 4.2 Comparison of the percentage of grid points (either all grid points or land grid points (LGP)) with 90%DD before the end of the simulations in 2100. Results for CanESM2 and CESM1 ensembles at the annual (Y) and seasonal (DJF and JJA) scales are shown for both indices at both the local and regional scales

Model	Scale	All grid points				Land grid points (LGP)			
		Local scale		Regional scale		Local scale		Regional scale	
		PRCPTOT	RX1day	PRCPTOT	RX1day	PRCPTOT	RX1day	PRCPTOT	RX1day
CanESM2	Y	60.5%	59.6%	64.1%	74.1%	66.0%	53.2%	71.2%	78.2%
	DJF	46.4%	43.8%	52.6%	59.7%	56.0%	55.0%	64.1%	69.3%
	JJA	44.0%	41.5%	51.7%	59.1%	30.0%	26.8%	38.4%	42.7%
CESM1	Y	54.6%	52.2%	60.3%	78.8%	57.3%	38.9%	67.1%	81.5%
	DJF	42.1%	37.0%	49.2%	61.1%	50.7%	47.1%	60.2%	68.8%
	JJA	38.7%	34.7%	47.3%	60.9%	23.8%	21.1%	33.6%	43.2%

For PRCPTOT, many regions will experience an increase in precipitation (especially at high latitudes), while a considerable number of regions will also see a decrease in precipitation (see Figure 4.8 and Figure 4.9). However, for RX1day (see Figure 4.12 and Figure 4.13), nearly all land grid points show an increasing trend due to climate change. Globally, the RX1day index shows more increasing trends than the PRCPTOT index, both for the annual and seasonal scales (with the smallest percentage at the JJA scale). Thus, there will be regions that will see a decrease in annual total precipitation, but an increase in RX1day. While the RX1day index increases globally at the annual scale, many regions will see a decrease at the JJA (e.g., AUS, Central North America – CNA, MED, Northern Europe – NEU and SAF). The AMZ and Central America (CAM) seem to be the only regions where

decreases are observed year round. Overall, these spatial patterns of average increasing or decreasing trends agree with the general behavior of the expected climate change signal described by the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) and other published studies (Fischer et al., 2013; Hegerl et al., 2015; IPCC, 2013; King et al., 2015; Maraun, 2013b).

The results indicate that for both indices the climate change signal will be affected by natural variability until past the mid-century for most of land grid points at the local and regional scales. At the global scale, Fischer et Knutti (2014) showed that a significant fraction of grid points will experience increases. It is also likely that this influence will be stronger during summer (JJA) than during winter (DJF) or at the annual scales. When looking at the 21 geographical regions, high latitude (e.g. GRL, ALA and NAS) and tropical (e.g. AMZ, WAF and EAF) climate change signals will be detected much earlier than in other regions for both ensembles; other regions will see their 90%DD reached later in the century.

Overall, natural variability represents a considerable source of uncertainty and it can mask or amplify the climate change signal at both the local and regional scales. This conclusion agrees with those from previous studies (Deser et al., 2012a; Deser et al., 2012b; Fischer & Knutti, 2014; Kay et al., 2015; Sanderson et al., 2018; Thompson et al., 2015).

4.5 Discussion of Limitations

The following issues need to be discussed as their outcome may impact the conclusions of this study:

4.5.1 Coarse resolution of the ESM

There are indications that even with their coarse spatial resolutions, both GCMs and ESMs do a reasonably good job capturing the large-scale events usually associated with synoptic weather patterns (IPCC, 2013; Sillmann et al., 2013b). However, smaller scales weather events in GCMs or ESMs are not directly simulated but considered through convection

parameterization schemes (Chan et al., 2014; Jones & Randall, 2011; Kendon et al., 2017; Kendon, Roberts, Senior, & Roberts, 2012; Prein et al., 2015; Prein et al., 2017). A spatial resolution of the order of the kilometer would be required to adequately simulate deep convection which plays a significant role in the generation of extreme rainfall in some regions at the daily scale (Prein et al., 2015). Thus, the impact of spatial resolution and deep convection parameterization needs to be investigated using a large ensemble of simulations at very high resolutions (~ few kilometers). The only available simulations are still limited to small regions (Prein et al., 2015; Prein et al., 2017).

4.5.2 Representative concentration pathway

There is evidence to suggest that the rate of increase in extreme precipitation does not depend specifically on the emission scenario (as it does for mean precipitation), but rather on the total amount of warming (Pendergrass, Lehner, Sanderson, & Xu, 2015). The RCP8.5 used in this study represents the scenario with the largest increase in greenhouse gas concentrations typically used in climate change studies (IPCC, 2013). It is reasonable to think that under less significant anthropic forcing, natural variability could be expected to hide the anthropogenic climate change signal over longer time periods since forcing is weaker. This hypothesis could only be validated by comparing two large ensembles of simulations from the same model with different forcing scenarios.

A study by Sanderson et al. (2018) used two large ensembles from the Community Earth System Model with identical settings (30 members using RCP8.5 and 15 members using RCP4.5) to explore the role played by greenhouse gas concentration trajectories. Their results suggest a considerable overlap in possible outcomes for both ensembles even in the 2080 decade. Some significant changes between both scenarios started appearing, albeit with considerable overlap after 2040 at the regional scale in Northern Europe, while no difference was observed at the local scale.

By extending these conclusions to this work, under the weaker RCP2.6 or RCP4.5 scenarios, lower probabilities of detecting the climate change signal could be expected resulting in later 90%DD than those obtained for the RCP8.5 at the regional scale, but with little differences at the local scale.

4.5.3 Simulation period

Trend analyses were performed on sub-periods of the 1950-2100 simulations. Extending this the pre-1950s period, and ultimately to the 19th century, when anthropogenic forcing began, could possibly have an impact on trends detection in the climate change signal. This is because trends detection will very likely be impacted when using longer time series, which could in turn have an impact on the estimated trend detection probability during forthcoming periods. Such work could only be performed if both large ensembles had simulations using extended periods prior to 1950.

This was tested to see the impact of using the longer simulation period available for CESM1 (from 1920 to 2100) and shown in the Supplementary materials, Figure S4.24 and Figure S4.25 for PRCPTOT, and Figure S4.26 and Figure S4.27 for RX1day. Using an extended period early in the 20th century did not provide different conclusions from those obtained using the simulations starting in 1950. Thus, it is reasonable to assume that this limitation should not have a significant impact on the results and conclusions obtained in the present paper.

4.6 Concluding remarks

For precipitation extremes, natural variability is likely to dominate the climate change signal at the local scale until the next century in many parts of the world. To properly estimate trends in extreme precipitation it is essential to take into account spatial dependence. This is less critical for annual and seasonal total precipitation, which is comparably less affected by natural variability at the local scale. When accounting for spatial dependence, trend detection

for precipitation extremes is expected to occur for a larger number of grid points than for annual and seasonal total precipitation.

In some instances, natural variability may undermine our ability to detect the climate change signal at the local and regional scales. This should not prevent us from implementing adaptation measures, especially when dealing with precipitation extremes. In other words, the uncertainty linked to natural variability should not detract decision makers from underlying anthropogenic changes. Nonetheless, results from this study clearly show that natural variability can impede the detection of the anthropogenic signal for a few to several decades over many parts of the world, and this should be considered when implementing adaptation strategies.

4.7 Acknowledgements

The authors acknowledge the contribution of Environment and Climate Change Canada's Canadian Centre for Climate Modelling and Analysis in executing and making available the CanESM2 Large Ensemble simulations used in this study, and thank the Canadian Sea Ice and Snow Evolution Network for proposing the simulations. The authors would also like to thank the Ouranos Consortium for helping with data transfer. The CESM1 ensemble was downloaded from the Large Ensemble Community Project (LENS) website (http://www.cesm.ucar.edu/projects/community-projects/LENS/). The HadEX2 and GHCNDEX gridded datasets were downloaded from the Climate Extreme indices (CLIMDEX) website (www.climdex.org). Finally, the authors would like to thank the anonymous referees that helped to significantly improve the quality and relevance of this paper.

4.8 Supplementary materials

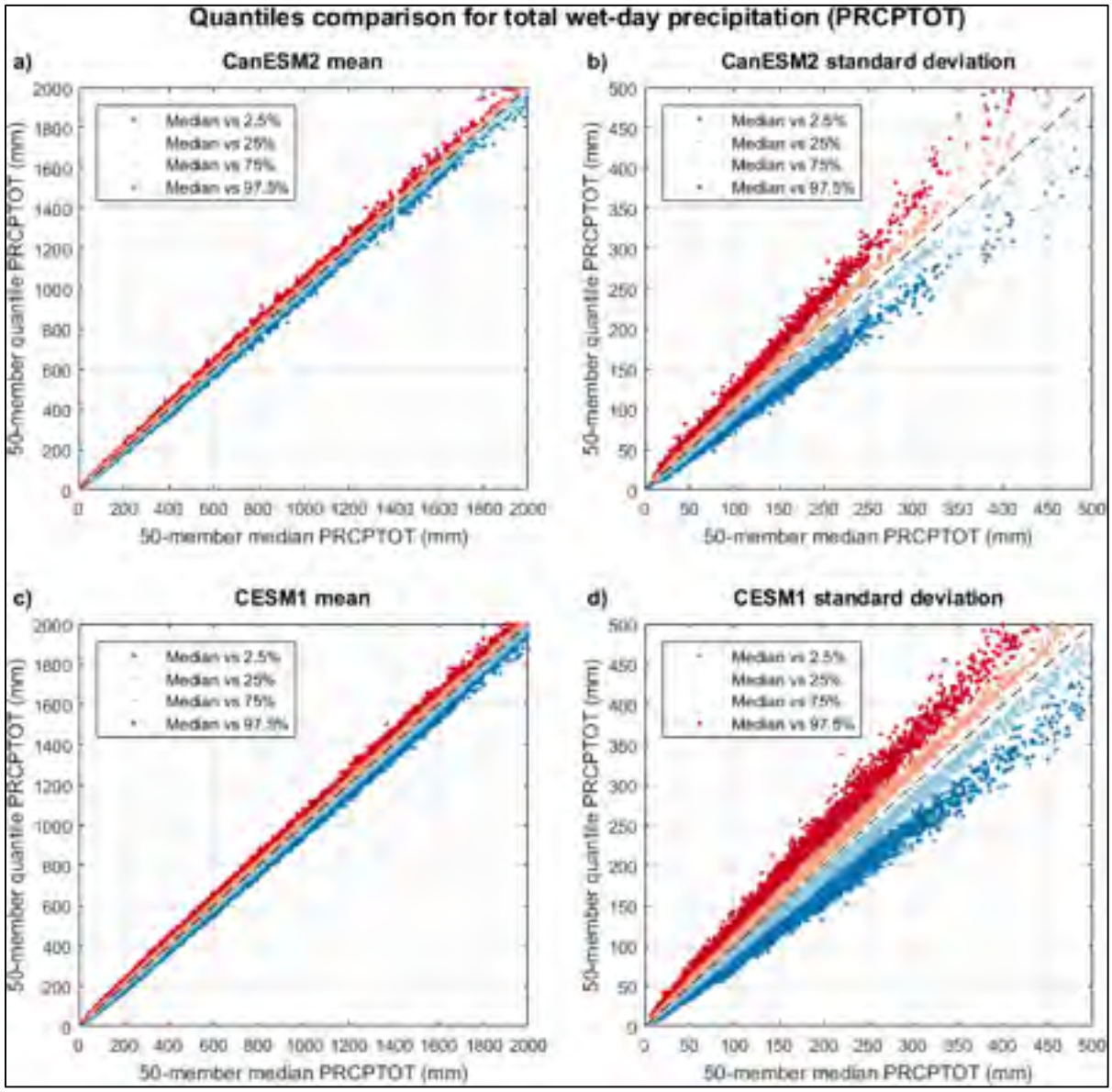

Figure S4.16 Scatter plots of various quantiles (2.5%; 25%; 75% and 97.5%) of the distribution of the mean (left) and standard deviations (right) of the annual PRCPTOT series (1950-2010) over the 50-member CanESM2 (top) and 40-member CESM1 (bottom) simulations at a given grid point as a function of the corresponding median value of the 50-member CanESM2 distribution. Some outliers are not displayed for ease of interpretation

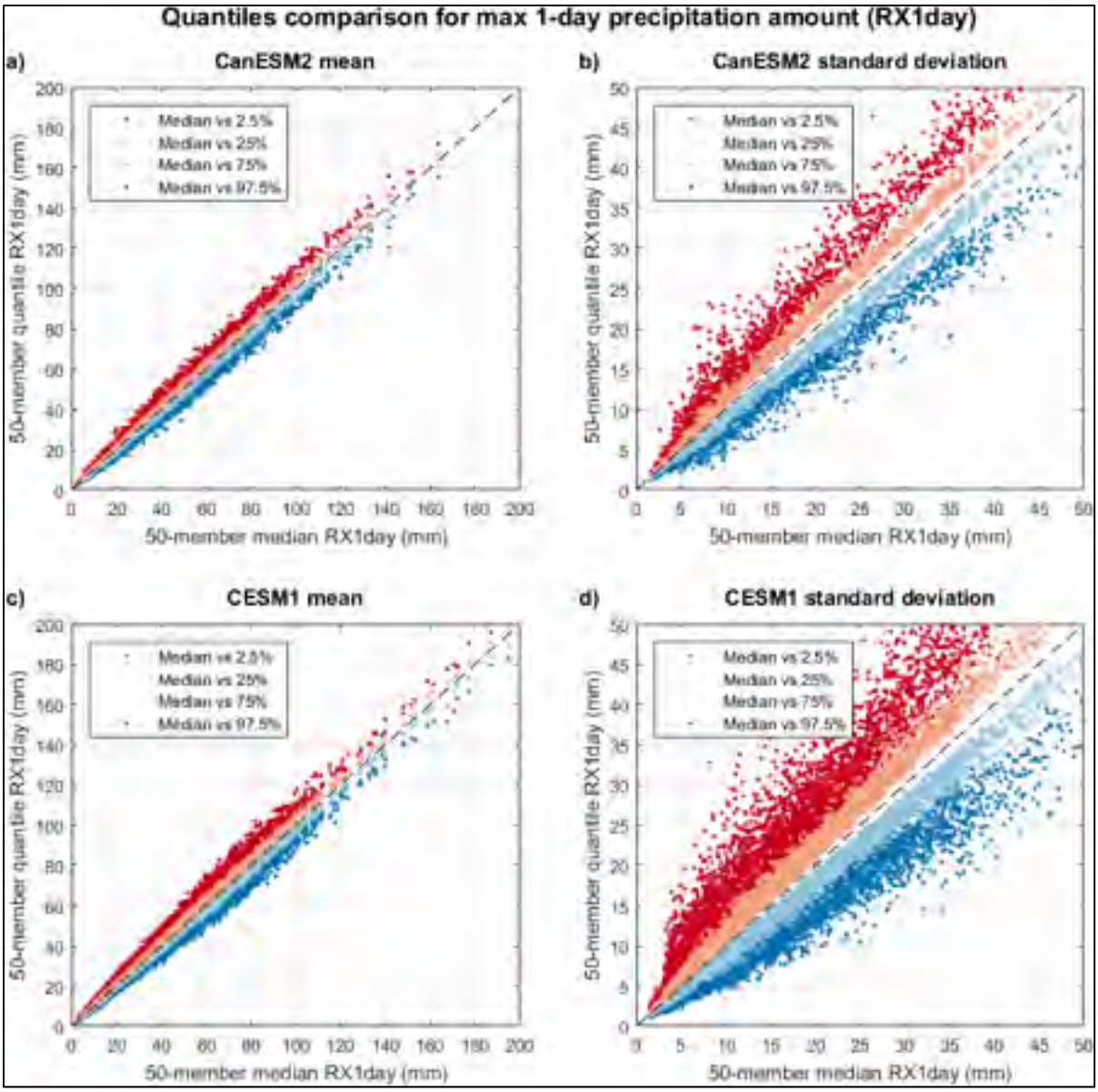

Figure S4.17 Same as Figure S4.16, for the RX1day index

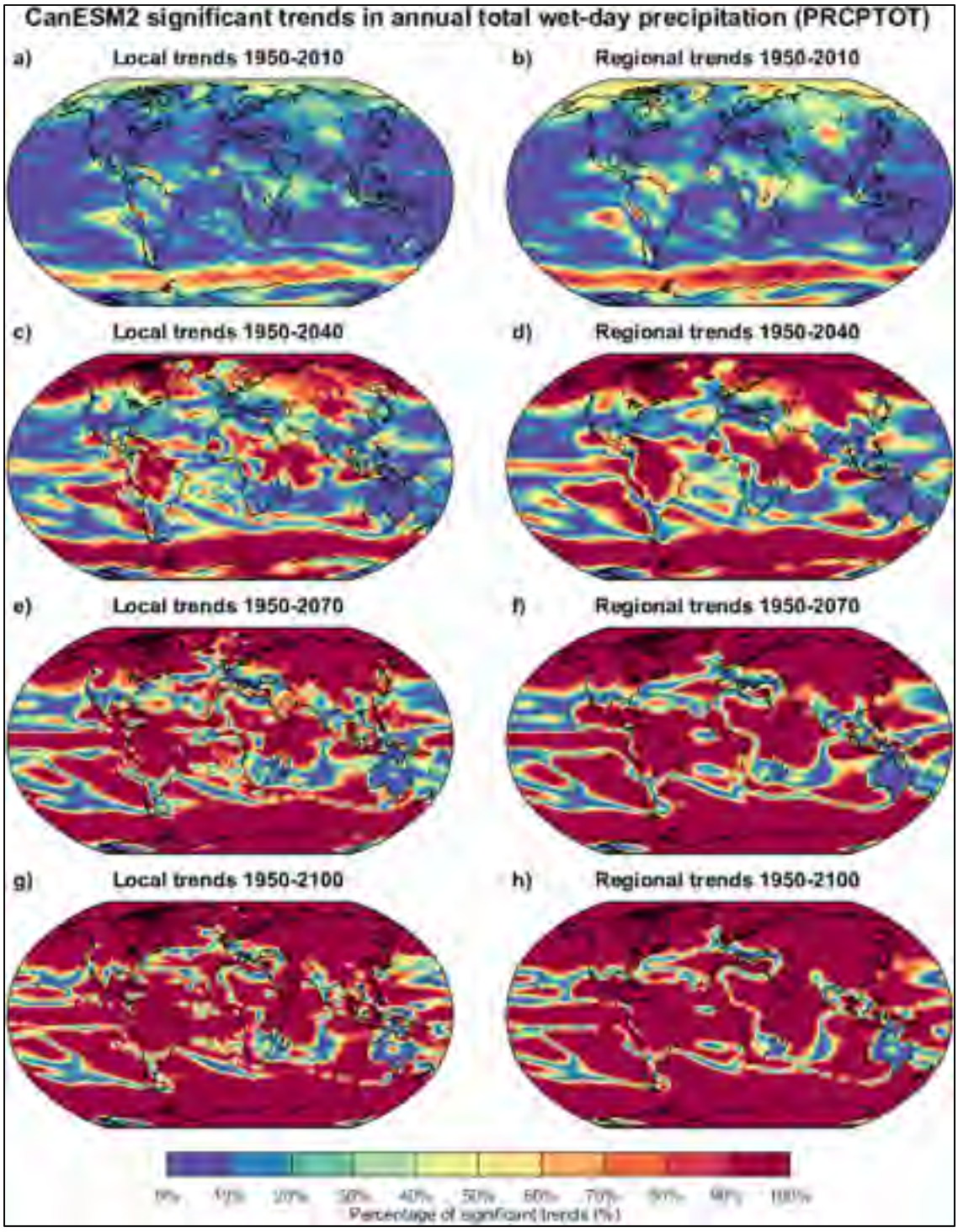

Figure S4.18 Maps of the percentage of significant trends of the same sign for each CanESM2 grid points for the PRCPTOT index, showing results based on (left) the local scale and (right) the regional scale for the (top)-(bottom) 1950-2010, 1950-2040, 1950-2070 and 1950-2100 periods

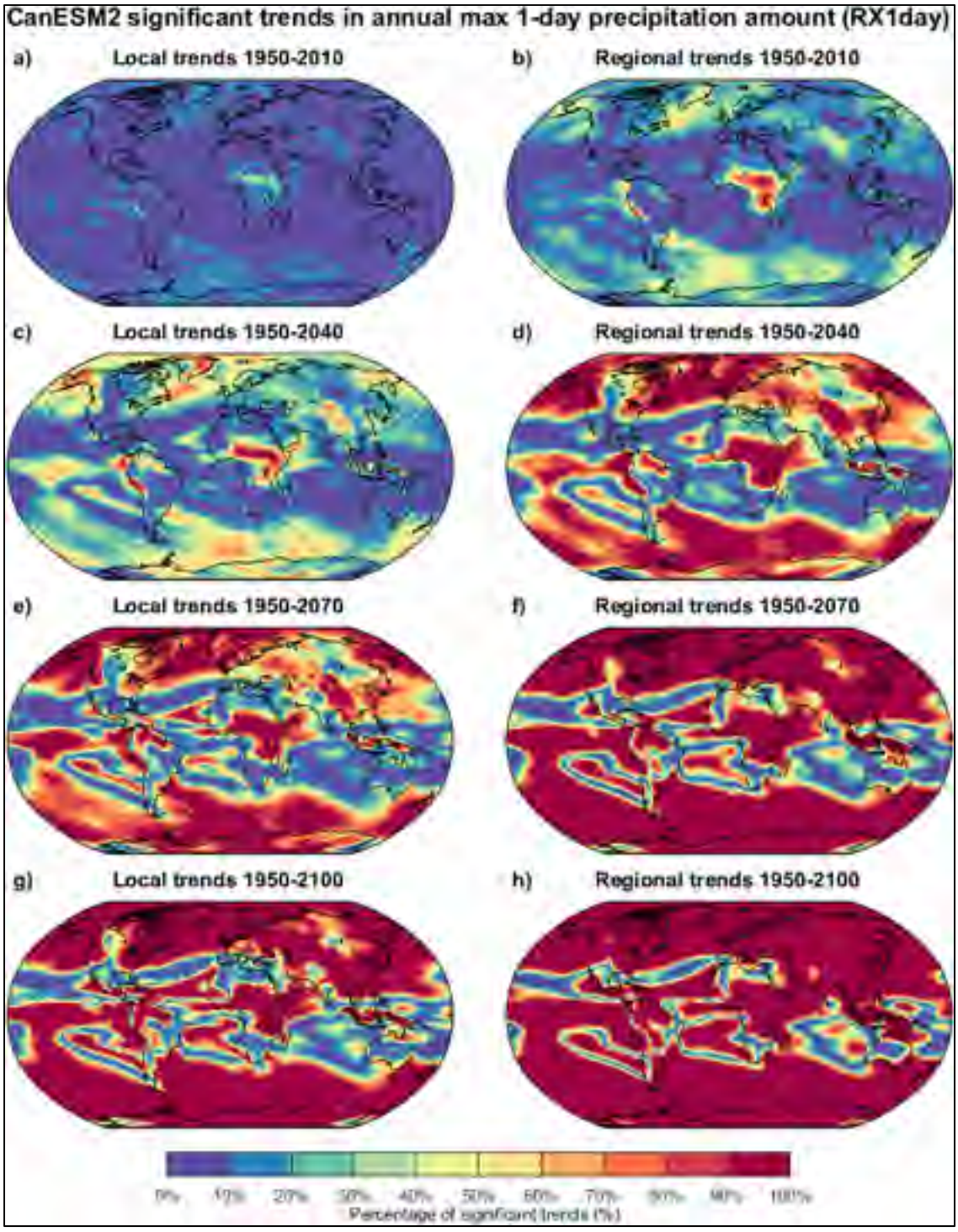

Figure S4.19 Same as Figure S4.18, but for the CESM1 model

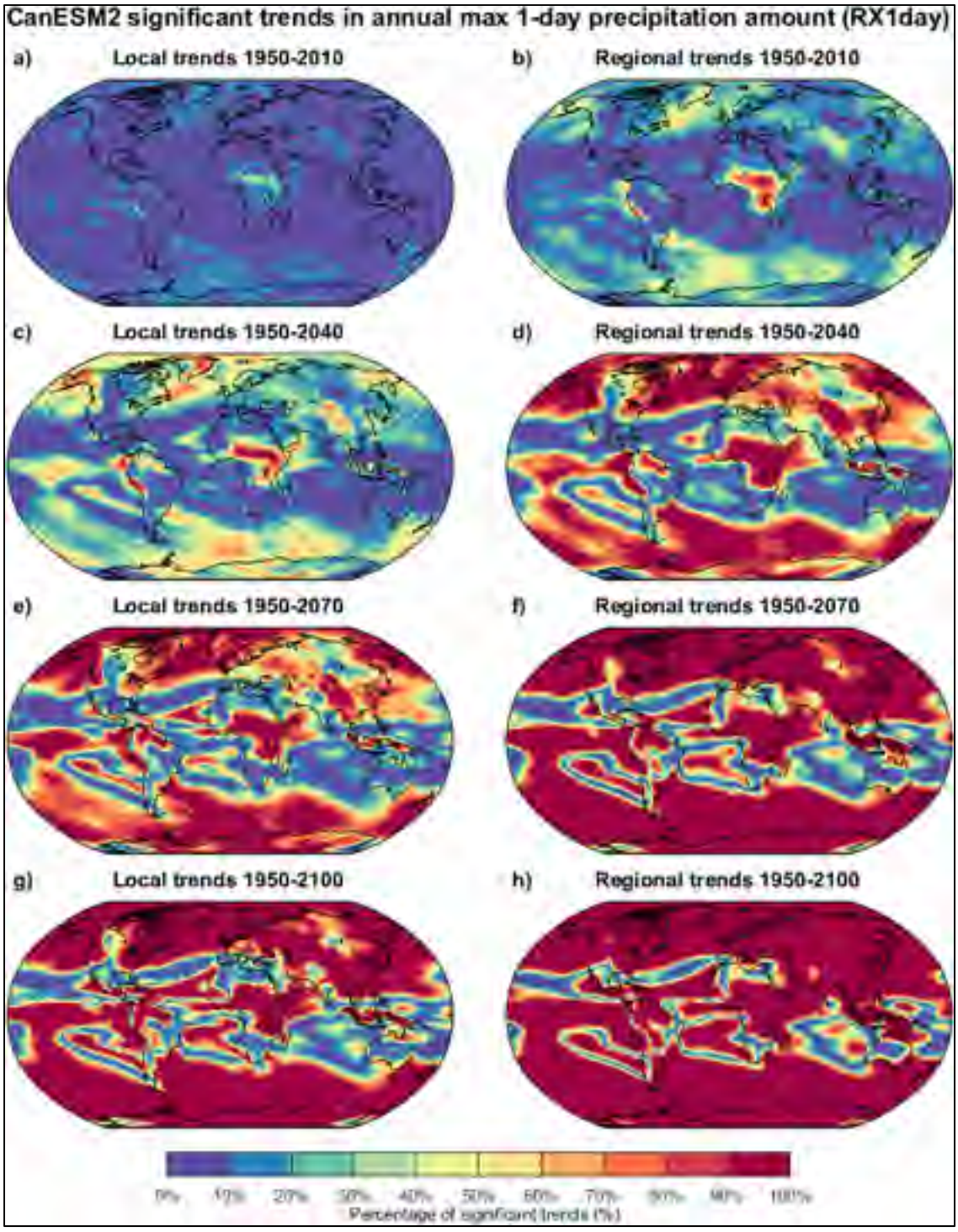

Figure S4.20 Same as Figure S4.18 for RX1day index and CanESM2

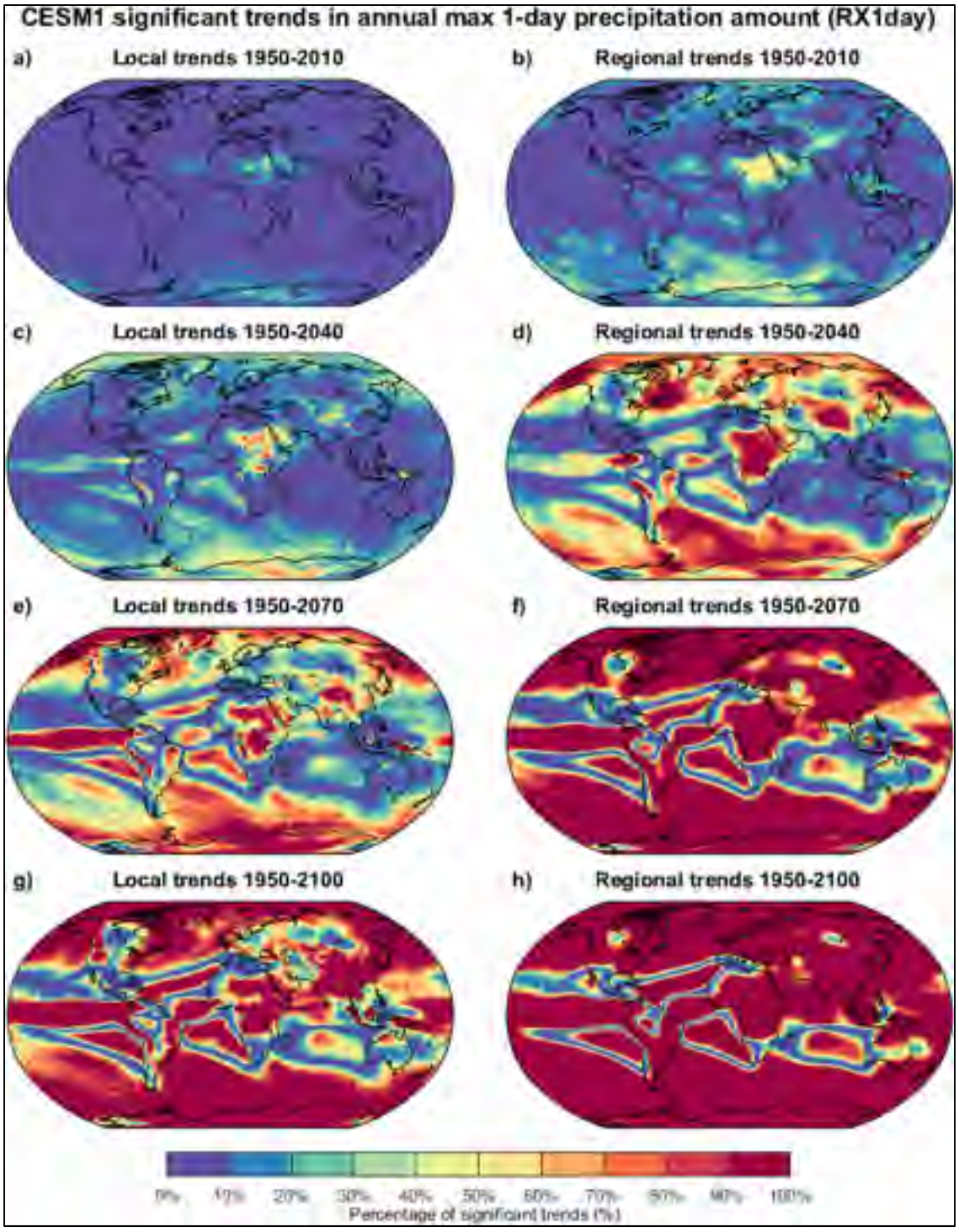

Figure S4.21 Same as Figure S4.18 for RX1day index and CESM1

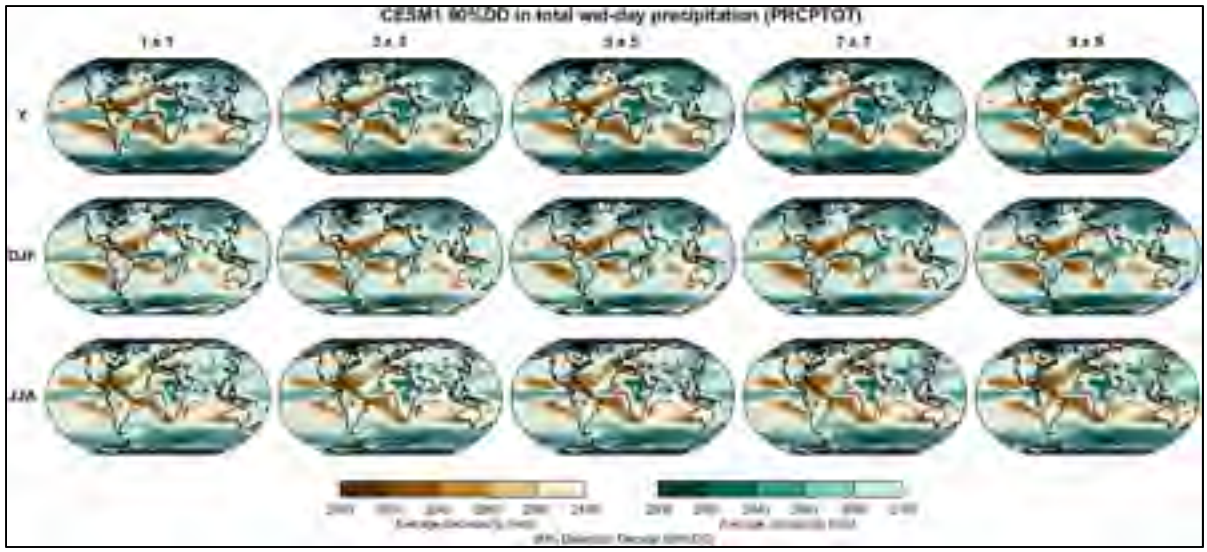

Figure S4.22 Comparison of the global analysis of the estimated 90%DD for the PRCPTOT index and CESM1 ensemble over the 1950-2100 period using different region size, showing results based on (left)-(right) the local scale using 1 (1 × 1) grid point and the regional scale using 9 (3 × 3), 25 (5 × 5), 49 (7 × 7) and 81 (9 × 9) grid points for the (top)-(bottom) annual (Y), DJF and JJA time scales. The brown colors represent an average decreasing trend while the blue-teal colors indicate an average increasing trend. The stippled patterns identify grid points where the estimated 90%DD occurred before 2100

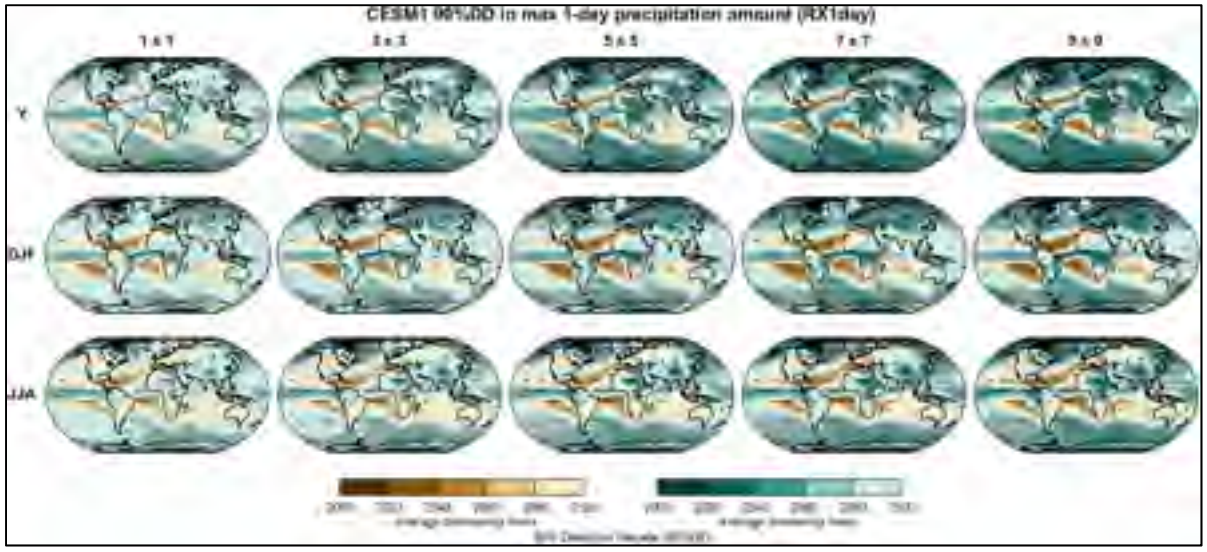

Figure S4.23 Same as Figure S4.22 for RX1day index

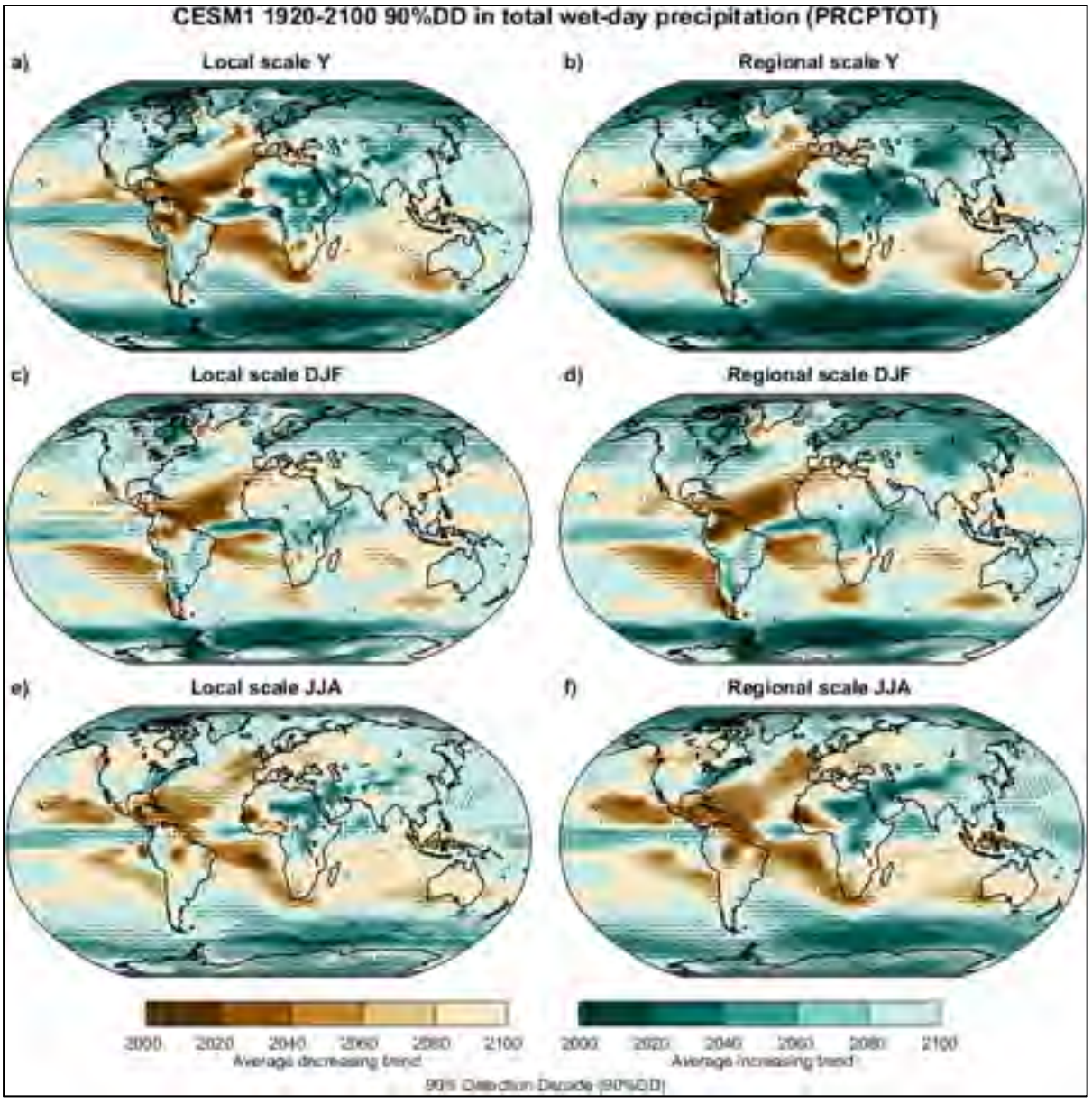

Figure S4.24 Global analysis of the estimated 90%DD for the PRCPTOT index and CESM1 ensemble over the 1920-2100 period, showing results based on (left) the local trends (i.e. corresponding to each grid point) and (right) the regional trends computed with the field significance resampling approach for the (top)-(bottom) annual (Y), DJF and JJA time scales. The brown colors represent an average decreasing trend while the blue-teal colors indicate an average increasing trend. The stippled patterns identify grid points where the estimated 90%DD occurred before 2100

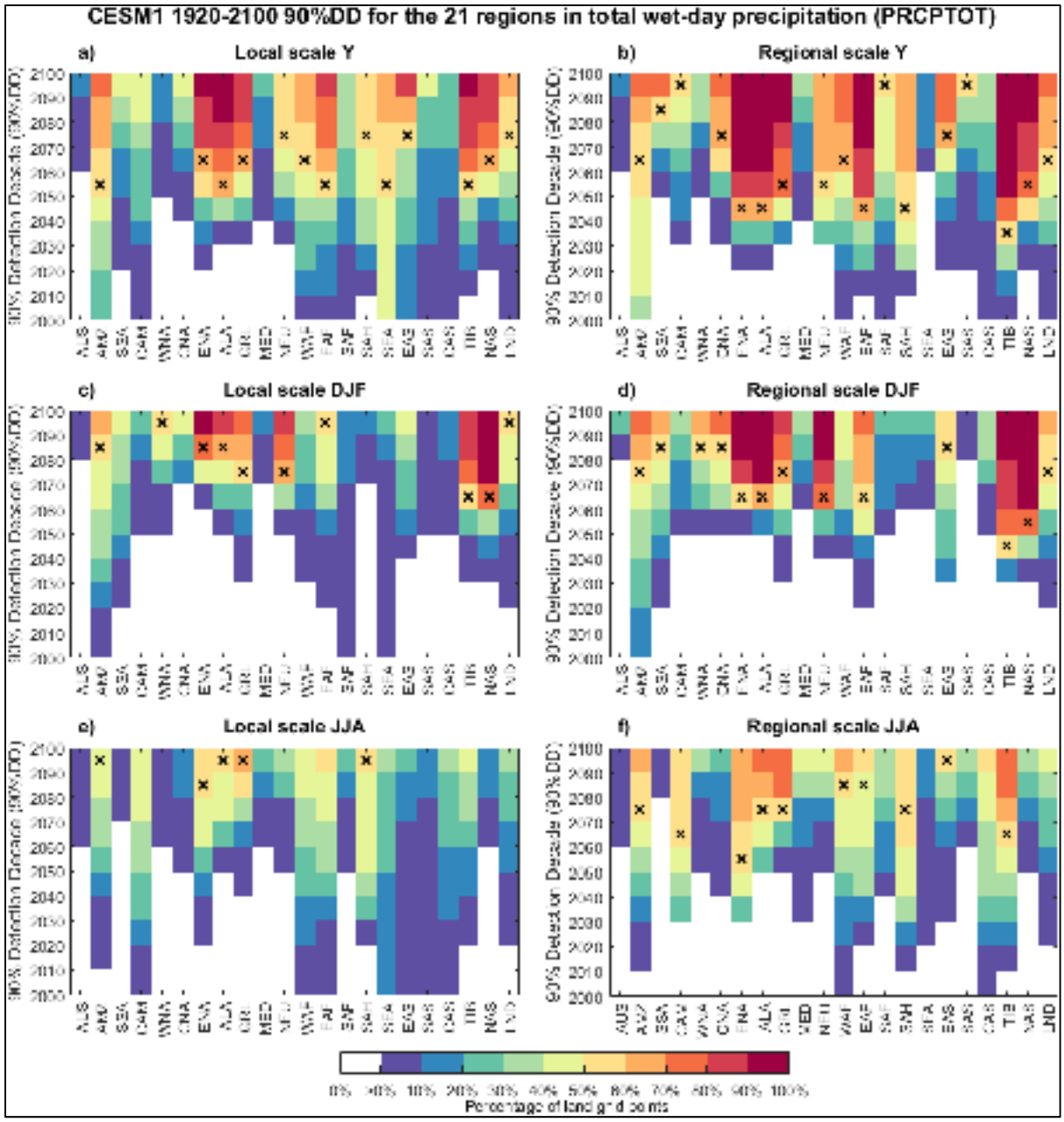

Figure S4.25 Cumulative percentage of land grid points within each of the 21 regions (listed in Table 4.1) with estimated 90%DD occurring in each decade for the PRCPTOT index and CESM1 ensemble over the 1920-2100 period, showing results based on (left) the local trends (i.e. corresponding to each grid point) and (right) the regional trends computed with the field significance resampling approach for the (top)-(bottom) annual (Y), DJF and JJA time scales. The black "x" indicates the decade when more than 50% of the region's land grid points have reached the 90% probability of detecting the climate change signal. The white boxes correspond to regions without 90%DD at any grid points in (and prior to) that decade

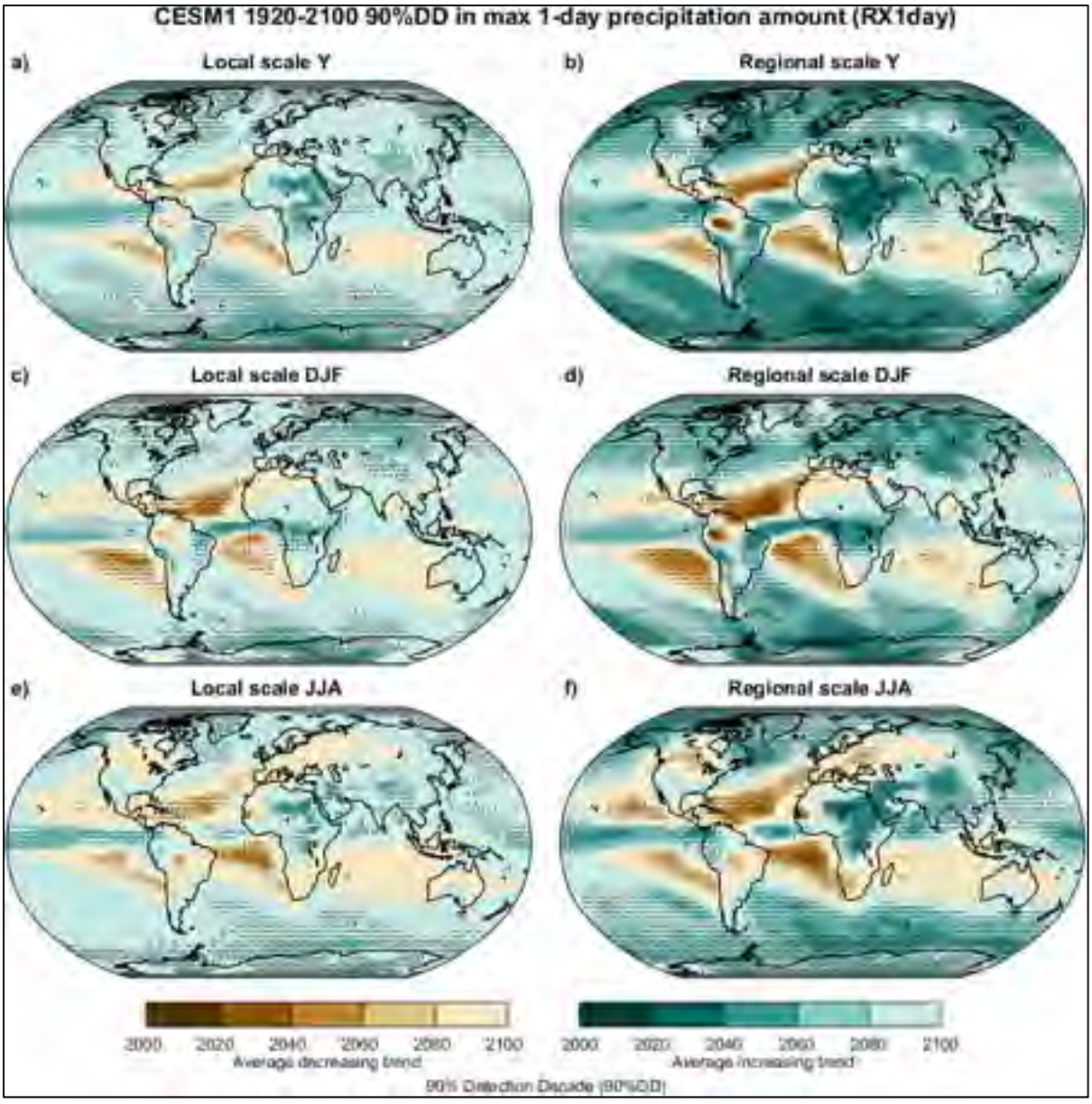

Figure S4.26 Same as Figure S4.24 for RX1day index

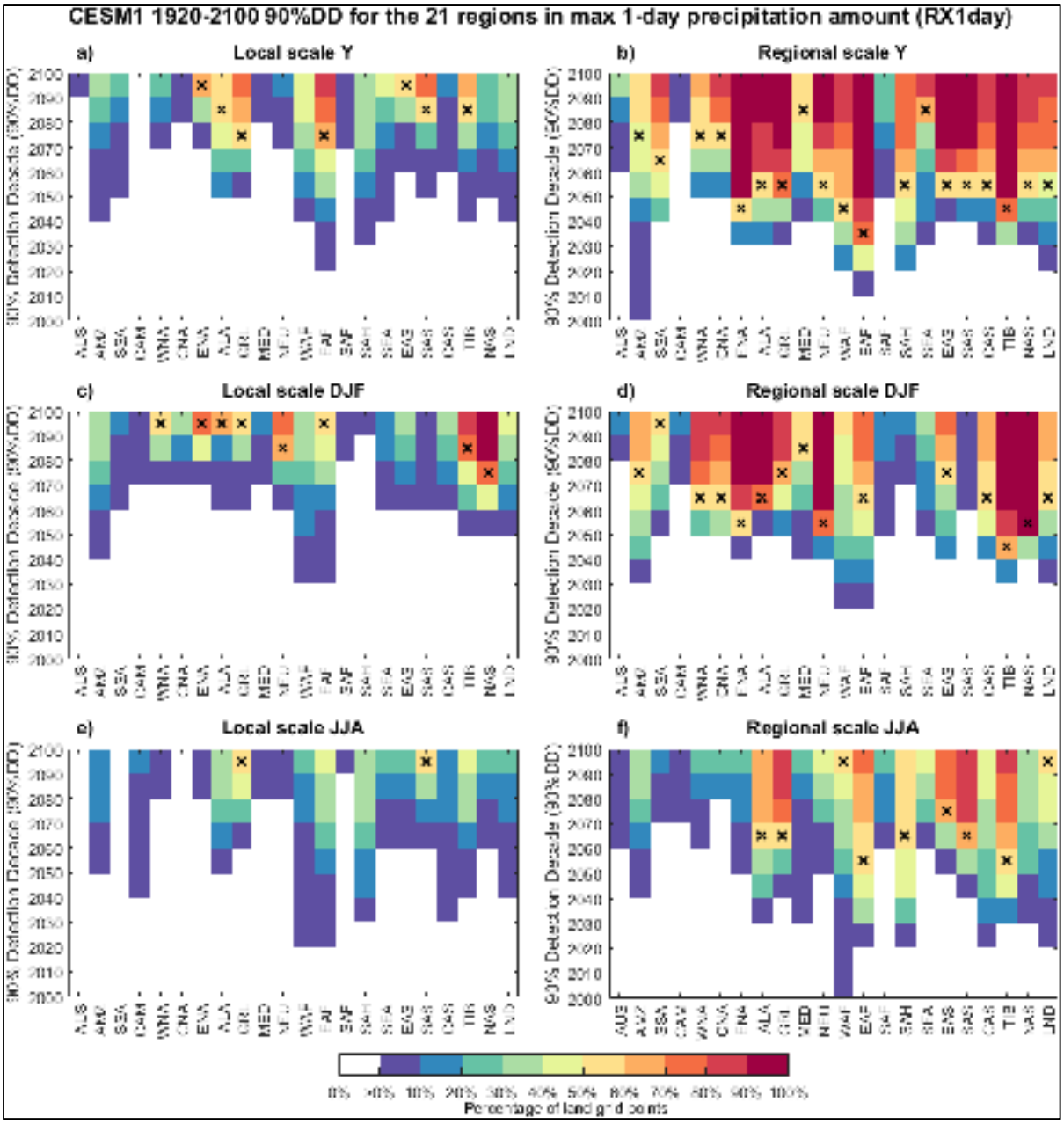

Figure S4.27 Same as Figure S4.25 for RX1day index

CHAPITRE 5

GLOBAL AND REGIONAL PROJECTED CHANGES IN 100-YEAR SUB-DAILY, DAILY AND MULTI-DAY PRECIPITATION EXTREMES ESTIMATED FROM THREE LARGE ENSEMBLES OF CLIMATE SIMULATIONS

Jean-Luc Martel[a], Alain Mailhot[b], François Brissette[c]

[a, c] Département de Génie de la construction, École de technologie supérieure
[b] Institut national de la recherche scientifique – Eau, Terre et Environnement

Article accepté avec révisions majeures à la revue « Journal of Climate », février 2019

Abstract

Many studies have reported projected increases in the frequency and intensity of extreme precipitation events in a warmer future climate. These results challenge the assumption of climate stationarity, a standard hypothesis in the estimation of extreme precipitation quantiles (e.g., 100-year return period) often used as key design criteria for many infrastructures. In this work, changes in hourly to 5-day precipitation extremes occurring between the 1980-1999 and 2080-2099 periods are investigated using three large ensembles (LE) of climate simulations. The first two are the global CanESM2 50-member ensemble at a 2.8° resolution and the global CESM1 40-member ensemble at a 1° resolution. The third is the regional CRCM5 50-member ensemble at a 0.11° resolution, driven at its boundaries by the 50-member CanESM2 ensemble over the Northeastern North America (NNA) and Europe (EU) domains. Results indicate increases in the frequency of future extreme events, and, accordingly, a reduction of the return period of current extreme events for all tested spatial resolutions and temporal scales. Agreement between the three ensembles suggests that extreme precipitations, corresponding to the 100-year return period over the reference period, become around four to five (two to four) times more frequent on average for the NNA (EU) domain for daily and 5-day annual maximum precipitation. Projections by CRCM5-LE show even larger increases for sub-daily precipitation extremes. Considering the lifespan of many public infrastructures, these changes may have important implications on service levels and

design of many water infrastructures and for public safety, and should therefore be taken into consideration in establishing design criteria.

5.1 Introduction

Extreme precipitation events are of high importance when considering the design of public water infrastructures with long life expectancy. Both daily and sub-daily precipitation events are of interest, especially when dealing with small watersheds. Due to the relatively short observational records and rarity of such events, methods based on the statistical theory of extreme value have traditionally been applied to estimate high precipitation quantiles usually used as design criteria (e.g., the 100-year return period; Katz (2013); Schulz et Bernhardt (2016)). However, actual design criteria generally assume that the climate is stationary, which ignores the scientific evidence pointing to human-induced global warming (Katz, 2013; Mailhot & Duchesne, 2010; Milly et al., 2008a). It is expected that projected increases in global temperature will lead to an increase in the intensity of extreme rainfall events, partly because a warmer atmosphere can hold more moisture (IPCC, 2013; Lenderink & Fowler, 2017; Trenberth, 1999; Trenberth et al., 2003). However, the relationship between the increase in temperature and the projected change in extreme precipitation, especially at the sub-daily scale, has been shown to be rather complex (Lenderink & Fowler, 2017; Lenderink et al., 2011; Westra et al., 2014).

Some studies, using various trend detection methods (Min et al., 2011; Westra et al., 2013) and datasets (Alexander et al., 2006; Donat et al., 2013b), have found that about two-thirds of the data-covered global land areas exhibit positive trends in annual daily precipitation extremes for the latter half of the 20th century. Significant increases using different indices of precipitation extremes have also been reported in various studies (Easterling et al., 2000; Groisman et al., 2005; Madsen et al., 2014; Seneviratne et al., 2012; Trenberth et al., 2007). More specifically, widespread increasing trends in daily precipitation extremes were observed for the eastern half of North America (Donat et al., 2013b; Easterling et al., 2000; Groisman et al., 2005). Similar increases over many regions of Europe, notably over eastern

and northern Europe, were also reported, while decreases were observed in southern Europe and around the Mediterranean basin (Donat et al., 2013b; Easterling et al., 2000; Groisman et al., 2005; Madsen et al., 2014; van den Besselaar, Klein Tank, & Buishand, 2013; Zolina, 2012).

There is also evidence, as highlighted in reviews by Westra et al. (2014) and Madsen et al. (2014), indicating that anthropogenic climate change is also likely leading to an increase in sub-daily precipitation extremes. For instance, studies are pointing to general increases in extremes for durations ranging from multiple hours to days for North America (Brommer et al., 2007; Burn et al., 2011; Kunkel et al., 2013; Muschinski & Katz, 2013) and for Europe and the Mediterranean basin, for which the increases depend on the region, season and duration of the event (Arnone, Pumo, Viola, Noto, & La Loggia, 2013; Leahy & Kiely, 2011; Madsen, Arnbjerg-Nielsen, & Mikkelsen, 2009; Ntegeka & Willems, 2008; Wang et al., 2011).

The Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5; IPCC, 2013) concluded with high confidence that by the end of the 21st century, the frequency of daily precipitation extremes will likely increase for mid-latitude land masses and wet tropical regions due to global warming. This claim is supported by many studies using the General Circulation Models (GCMs) from the Coupled Model Intercomparison Project version 3 (CMIP3; Meehl et al., 2007) and version 5 (CMIP5; Taylor et al., 2011) at the global scale (Kharin et al., 2007; Kharin et al., 2013; Sillmann et al., 2013a; Sillmann et al., 2013b; Wuebbles et al., 2013). Furthermore, in a study using the CMIP3 ensemble (Kharin et al., 2007) and a follow-up study using the CMIP5 ensemble (Kharin et al., 2013), the projected change in the 20-year return period of annual maximum daily precipitation was estimated. Both studies found worldwide increase in precipitation extremes, in terms of frequency of occurrence, except for some subtropical and tropical regions. Similar conclusions were found for the United States in a study by Wuebbles et al. (2013).

Dynamical downscaling of GCMs through the use of high-resolution a Regional Climate Model (RCM) brought significant improvements in the representation in both daily and sub-daily precipitation extremes (Prein et al., 2013; Tripathi & Dominguez, 2013) at the local and regional scales (Maraun et al., 2010a). RCMs with a spatial resolution of ~10 km using parameterized convection have been shown to adequately capture the intensity of daily precipitation extreme events (Ban et al., 2014). However, a resolution of the order of a few kilometers is required to adequately resolve the convective processes directly linked to sub-daily summer extreme rainfall, which occurs at very small scales (Chan et al., 2014; Kendon et al., 2017; Prein et al., 2015; Westra et al., 2014). Indeed, convective permitting scales have demonstrated clear added value for reproducing precipitation and circulation associated with orography (Rasmussen et al., 2011). Convective precipitation has also shown to be largely sensitive to the selection of the model physics (Mooney, Broderick, Bruyère, Mulligan, & Prein, 2017).

Multi-model ensembles of RCM simulations, combining various RCMs (and consequently, GCMs as drivers of the RCMs), were also produced to assess the impact of climate change over specific regional domains. Examples of such ensembles are the NA-CORDEX (Giorgi et al., 2009) and the NARCCAP (Mearns et al., 2012) ensembles for North America and the ENSEMBLES (van der Linden & Mitchell, 2009) and the EURO-CORDEX (Jacob et al., 2014) ensembles for Europe. The use of various combinations of RCMs/GCMs in several studies allowed investigating future changes in precipitation extremes over more specific regions, along with an assessment of uncertainties related to model structure. Regarding the more extreme precipitation events, a number of studies have investigated changes in different regions of North America (e.g., Mailhot, Beauregard, Talbot, Caya, & Biner, 2012; Mailhot et al., 2007; Mladjic et al., 2011; Wehner, 2013; Zhu, 2013) and Europe (e.g., Aalbers et al., 2017; Hosseinzadehtalaei, Tabari, & Willems, 2018; Rajczak & Schär, 2017). In general, large projected changes (mostly increases) have been reported by these studies, in agreement with the sign of observed trends over both North America and Europe.

An important factor relevant to most of previously cited studies is the impact from the spatio-temporal resolution. Indeed, a coarser spatial resolution is expected to have both an averaging effect and inferior representation of various physical processes (Chen & Knutson, 2008; Volosciuk, Maraun, Semenov, & Park, 2015). When it comes to temporal resolution, it is expected that there will be a stronger intensification for extreme rainfalls of shorter duration (Mailhot et al., 2012; Westra et al., 2014). These spatio-temporal scaling issues have been previously explored with both observational datasets (Gehne, Hamill, Kiladis, & Trenberth, 2016; Westra et al., 2014; Zolina et al., 2014) as well as climate model experiments (Chan et al., 2014; Chen & Knutson, 2008; Mailhot et al., 2012; Volosciuk et al., 2015), but remain a limitation for most studies.

Based on the consensus highlighted in the previous studies, it becomes clear that anthropogenic climate change played, and will play, a significant role in the evolution of both daily and sub-daily precipitation extreme events. However, since these studies generally combine different models, they lack the ability to properly disentangle inter-model variability from natural climate variability (Fischer et al., 2013; Kay et al., 2015; Martel et al., 2018). Furthermore, most of the reported studies could not assess possible changes in very rare events (such as the 100-year return period event) without large statistical uncertainties, due to the relatively short simulated series or the small number of ensemble members (Schulz & Bernhardt, 2016).

Large Ensemble (LE) from the same climate model can be used to partly overcome these limitations. Such ensembles are generated by running a climate model many times with slightly different initial conditions, resulting in different simulations for the same time period and forcing scenario after merely a few weeks (Deser et al., 2012a; Martel et al., 2018).

These ensembles have gained in popularity in the past years, and a growing number of climate change studies are using them (e.g., Aalbers et al., 2017; Deser et al., 2012a; Deser et al., 2012b; Deser et al., 2014; Fischer et al., 2014; Kay et al., 2015; Martel et al., 2018; Sanderson et al., 2018; Thompson et al., 2015). However, due to their high computational

costs, these ensembles are still relatively rare, especially when it comes to high-resolution RCMs (Aalbers et al., 2017; Mizuta et al., 2017).

The objectives of this study are two-fold: 1) assess the projected future changes in the frequency of large extreme precipitation events (up to 100-year return period events) using three LE of climate simulations (two global ensembles – CanESM2-LE and CESM1-LE – and one regional ensemble – CRCM5-LE); 2) check the global consistency of the projected changes in extreme precipitations and evaluate the impact of both the spatial resolution and temporal scale of LE models on the projected future changes. The three LE of climate simulations and the methods used in their production are presented in Section 5.2, while the analysis methods are described in Section 5.3. Results and discussion are shown in Section 5.4. Finally, concluding remarks are presented in Section 5.5.

5.2 Datasets and Methods

Three LE were considered in this study. Two of these ensembles are global, and use two different ESMs, while the third one is produced by an RCM driven by a GCM over two distinct domains.

5.2.1 CanESM2 and CESM1 Large Ensembles

The first ensemble is the Canadian Earth System Model version 2 Large Ensemble (CanESM2-LE) from the Canadian Centre for Climate Modelling and Analysis (CCCma), with a 2.8° spatial resolution (Arora et al., 2011; Sigmond & Fyfe, 2016; von Salzen et al., 2013). First, five simulations over the 1850-1950 period were launched to obtain five different states of the oceans, and these were then used as the initial conditions for five simulation families. This step allowed accounting for some of the oceanic variability, which is characterized by a much longer response time than the atmosphere. These five initial conditions were each randomly perturbed ten times, and simulations run until 2006 using historical forcing, and then using the RCP8.5 forcing scenario (Representative Concentration Pathway; Meinshausen et al., 2011) until 2100. Fifty simulations (called members) covering

the 1950 to 2100 period, for a total of 7500 (50 simulations x 150 years) simulated years were therefore produced.

The Community Earth System Model version 1 large ensemble CESM1-LE consists of 40 climate simulations (members) at a 1° resolution, covering the 1920 to 2100 period (Kay et al., 2015). First, a multi-century control simulation starting in 1850, and using constant preindustrial forcing was run until 1920. In contrast to CanESM2-LE, all forty simulations were produced using the same initial ocean state. From there, the air temperature fields were randomly perturbed 40 times at the round-off error level. The RCP8.5 forcing scenario was also introduced from 2006 to the end of the century.

Data from both CanESM2-LE and CESM1-LE are available for public download respectively from the Environment and Climate Change Canada (ECCC) and the National Center for Atmospheric Research (NCAR) websites.

The representation of inter-annual mean and variability within CanESM2-LE and CESM1-LE was investigated for both the total wet-day precipitation (PRCPTOT; total precipitation on days with precipitation ≥ 1 mm) and the annual maximum 1-day precipitation (RX1day) in a study by Martel et al. (2018). Both precipitation indices were compared against the observed HadEX2 (Donat et al., 2013b; resolution of 2.5° latitude x 3.75° longitude) and GHCNDEX (Donat et al., 2013a; resolution of 2.5° latitude x 2.5° longitude) gridded datasets. The spatial patterns of inter-annual mean and variability were found to be globally in good agreement with those of HadEX2 and GHCNDEX datasets for both PRCPTOT and RX1day, albeit to a lesser extent for the latter.

5.2.2 CRCM5 Large Ensemble

The Canadian Regional Climate Model version 5 (CRCM5; Martynov et al., 2013; Šeparović et al., 2013) was developed by the ESCER Centre at UQAM (Université du Québec à Montréal) in collaboration with Environment and Climate Change Canada (ECCC). The

CRCM5 was run at a resolution of 0.11° over two different regional domains: Northeastern North American (NNA) and European (EU) (Figure 1). The CRCM5 50-member ensemble (CRCM5-LE; Leduc et al., 2019) was produced within the ClimEx (Climate change and hydrological Extremes) project, part of a long-term collaboration between Bavaria and Québec (http://www.climex-project.org/). The 50 members were run using the 6-hour atmospheric and daily oceanic field outputs from CanESM2-LE at the boundaries of both domains covering the 1950 to 2100 period. ClimEx data will be available for public download in the near future. Six selected cities (three per domain) that will be further analyzed in this study are also shown in Figure 5.1. These were selected as they represent distinct change patterns in the results obtained.

A general climatological evaluation of CRCM5-LE output was performed using different observed gridded datasets for both domains (Leduc et al., 2019). Regarding mean daily precipitation, a wet bias was observed throughout the year for both the Northeastern North America (NNA) and Europe (EU) domains, with a strong dominant component in winter for both domains. A dry bias was also observed in summer for south-western NNA and Eastern Europe. A comparison with a CRCM5 run driven by the ERA-Interim reanalysis (ERA-CRCM5) showed that a significant portion of the wet bias can be attributed to CanESM2-LE (Leduc et al., 2019). The CRCM5 performance, notably in terms of extreme precipitation quantiles, annual and daily cycles, has been evaluated in a study by Innocenti et al. (under review). A comparison of the ERA-CRCM5 native resolution against observational records of weather stations displayed a good agreement over short durations extreme precipitation quantiles and for the 2-, 10- and 25-year return periods, but over-estimated for the daily and longer durations for many stations. The ERA-CRCM5 run has also shown a good representation of both the annual and diurnal cycles.

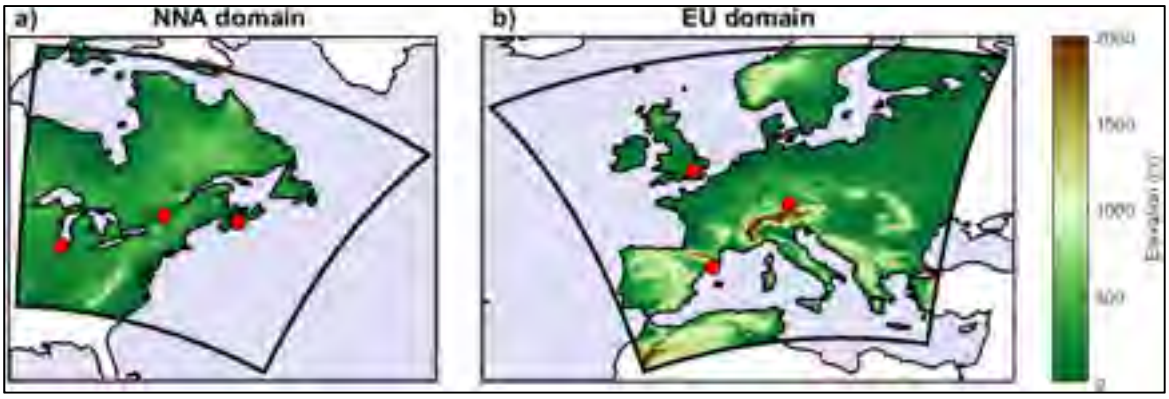

Figure 5.1 Digital Elevation Model (DEM) used in CRCM5-LE for the (a) Northeastern North America (NNA) and (b) Europe (EU) domains. Red dots appearing in the NNA domain represent, from left to right, Chicago, Montreal and Halifax. Red dots appearing in the EU domain correspond, from top to bottom, to London, Munich and Barcelona

5.3 Methods

Cumulative annual maximum precipitation (AMP) series (AMPS) for various temporal scales (from 1 hour to 5-day) were extracted and used to estimate long return periods (i.e. 100-year return period). Daily and 5-day annual precipitation extremes were first extracted for all grid points of all CanESM2 and CESM1 members for both the 20-year reference (1980-1999) and future (2080-2099) periods. Hourly outputs being only available for CRCM5-LE, moving windows were instead used to extract the AMP (e.g., a 24-hour moving window was used instead of the daily value to create the 24-hour AMPS). CRCM5-LE 1-hour to 120-hour (5-day) AMPS were thus similarly constructed for all grid points over both domains. AMPS from the different members were then pooled for each period, leading to a 1000-year AMPS for CanESM2-LE and CRCM5-LE (20 years × 50 members) and an 800-year AMPS for CESM1 (20 years × 40 members). The framework used in this study is built on the assumption of an ergodic process (Nikiéma, Laprise, & Dugas, 2017), given that the human-induced climate change signal (external forcing) will be dominated by natural variability when investing short time window (i.e. the time series are stationary over this time window).

Pooling the AMPS from the different members is based on the hypothesis that these series can be considered stationary over the given 20-year period. The existence of local trends at

each grid point for the three LE was therefore assessed using the non-parametric Mann-Kendall test (Kendall 1975) at a 95% confidence level. The distribution of the number of grid points with a given number of members with significant trends was compared to the binomial distributions, and differences between these distributions were assessed through the Kolmogorov-Smirnov test (95% confidence level). The binomial distribution was considered as it represents the distribution of the number of members that would randomly display a significant trend (5% probability to randomly generate significant trends) over 50 trials (members). The null hypothesis was not rejected in all cases for both the 1980-1999 and 2080-2099 periods. Thus, the reference and future 20-year time series can be considered stationary. A similar analysis was conducted for 30-year time series (1980-2009 and 2070-2099) and many grid points were not stationary.

The pooled time series were then sorted and used to estimate the empirical quantiles based on the Cunnane plotting position (Cunnane, 1978; Meylan et al., 2008). Empirical estimates were used, considering the length of available series (1000 years for CanESM2-LE and CRCM5-LE and 800 years for CESM1-LE). The 100-year return period AMP was first estimated for the reference period at each grid point. The 2080-2099 return period corresponding to the 1980-1999 100-year precipitation intensity was then estimated, therefore providing the projected change in the frequency of this extreme event (hereafter called future 1-day-T and 5-day-T for CanESM2-LE and CESM1-LE and future 1-hr-T, 24-hr-T and 120-hr-T for CRCM5-LE, depending on the temporal scale). A reference 100-year 24-hour rainfall AMP of 100 mm increasing to 130 mm over the future period, and where a 100 mm rainfall now corresponds to a 20-year AMP can be used as an example. In this case, the future changes in the reference 100-year AMP could be expressed as a 30% relative increase, or by becoming a 20-year AMP (an event 5 times as frequent). Expressing the future change in terms of changes in future return period of the reference 100-year AMP was preferred in this paper as it gives potential users a sense of the actual rarity of the event they experienced in a past climate (and for which they may have an idea of the impact).

5.4 Results and discussion

5.4.1 Projected changes at the global scale

Projected future changes for 1-day and 5-day AMPs at the global scale using CanESM2-LE and CESM1-LE were first analyzed. Figure 5.2 presents the projected return period over the 2080-2099 period of 1-day and 5-day AMPs (1-day-T and 5-day-T) having the same intensity as the 100-year event over the 1980-1999 reference period. Cumulative distributions of land grid point projected return periods of the reference 100-year AMP are shown in Figure 5.3.

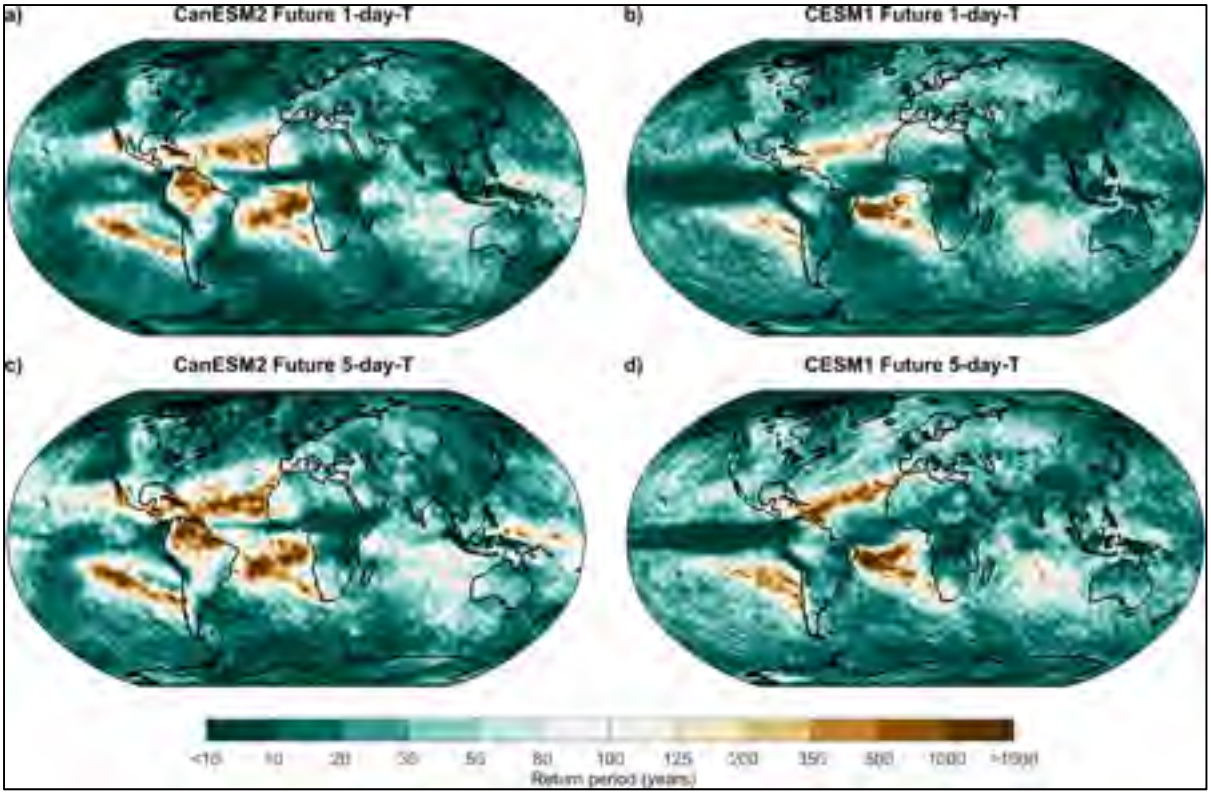

Figure 5.2 Maps of the projected return period over the 2080-2099 period of the 1-day (a,b) and 5-day (c,d) AMP with same intensity as the 100-year AMP over the 1980-1999 period for CanESM2-LE (a,c) and CESM1-LE (b,d)

Figure 5.2 shows that the frequency of future 1-day-T and 5-day-T will increase for almost all grid points, with the exception of some subtropical and tropical regions, notably for CanESM2-LE. Previous studies based on the CMIP5 ensemble have shown that these regions have strong natural variability and a relatively low inter-model agreement in the projected changes in precipitation extremes (IPCC, 2013; Kharin et al., 2007; Kharin et al., 2013). For instance, both ensembles display opposite signs in the projected changes over the Amazon basin, but CESM1-LE tends to be closer to the results from the CMIP5 inter-model mean obtained in previous studies. Aside from these discrepancies, there is global agreement between both large ensembles in the projected global change in precipitation extremes. Global patterns of changes for 1-day-T and 5-day-T for both LE are also very similar (Figure 5.2-a versus Figure 5.2-c and Figure 5.2-b versus Figure 5.2-d).

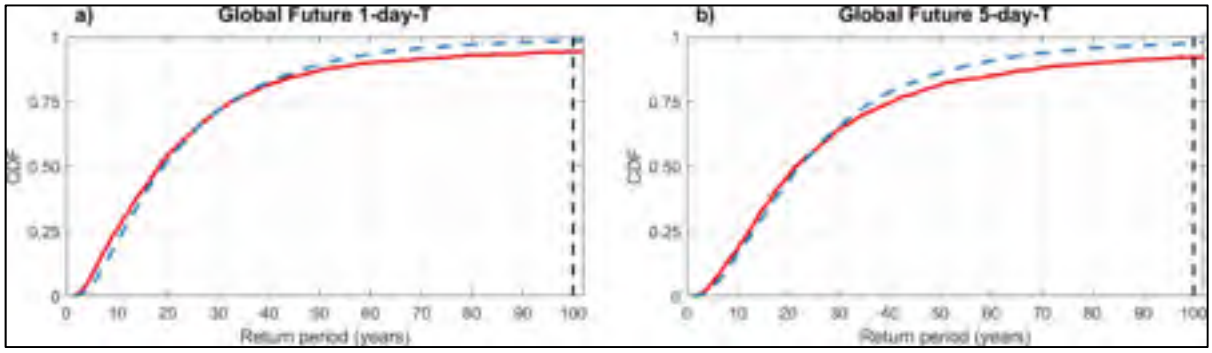

Figure 5.3 Cumulative distribution functions (CDF) of the projected return period over the 2080-2099 period of the 1-day (a) and 5-day (b) AMP with same intensity as the 100-year AMP over the 1980-1999 period. Distributions are shown for all land grid points from CanESM2-LE (red curve) and CESM1-LE (blue curve). The dashed vertical line corresponds to the 100-year return period. Values on x-axis longer than 100 years are not displayed.

Figure 5.3 shows that there is an overall strong agreement between the distributions of projected changes by CanESM2-LE and CESM1-LE for both 1-day-T and 5-day-T. Approximately 94% (98%) of land grid points for CanESM2-LE and 92% (98%) for CESM1-LE experience more frequent 1-day (5-day) precipitation extremes in the future. The median value of the future 1-day-T (5-day-T) is 18 (22) years for land grid points (without Antarctica) for CanESM2-LE and 19 (22) years for CESM1-LE, meaning that a 100-year

return period over the 1980-1999 reference period is about 4 to 5 times more frequent in 2080-2099 for half of the land grid points.

Similar results were obtained for the reference 20-year AMP both in terms of the sign and the magnitude of the projected changes (Supplemental Material Figure S5.9 and Figure S5.10). The median value over land grid points of the future 1-day-T (5-day-T) is 6 (6) years for CanESM2-LE and 6 (7) years CESM1-LE, or between 3 to 4 times more frequent. It should be noted that these results corroborate those obtained in the studies of Kharin et al. (2007); Kharin et al. (2013), where the multi-model average (29 CMIP5 models) global median value for all land grid points of the projected return period of the 20-year daily AMP by 2100 with the same intensity as the 20-year return period over 1986-2005 was also found to be 6 years.

5.4.2 Projected changes over the North American and European domains

The projected return period with the same intensity as the 100-year AMP over the 1980-1999 period simulated by CRCM5-LE over the NNA and EU domains are displayed in Figure 5.4. Cumulative distributions of land grid point values at various durations over CRCM5-LE are presented in Figure 5.5.

For the NNA domain, an overall increase in frequency of extreme precipitation events is projected (Figure 5.4-a-c and Figure 5.4-a). A west-to-east gradient towards a greater reduction in the future return period can be seen for all durations, especially for the 1-hr AMP. Some orographic effects can be observed for the 1-hr AMP as greater reductions in the return period are observed above the Appalachian Mountains (see the topography in Figure 5.1-a). The shortest projected return periods would be reached over the east coast, with return periods shorter than 10 years for the future 1-hr-T, corresponding to more than a 10-fold increase in frequency of the reference 100-year AMP events. Approximately 30% of NNA land grid points experience more than a 10-fold increase of the 1-hr AMP. This fraction quickly decreases as the duration increases since fewer than 5% of land grid points experience a 10-fold increases or greater for the 24-hr AMP. For the future 24-hr-T, the

lower projected return period values would range around 10- to 20-year, representing a 5- to 10-fold increase in frequency. As shown in Figure 5.5-a, the median future 1-hr-T is 14 years, while for the future 24-hr-T and 120-hr-T, median values over land grid points are 18 and 20 years, respectively.

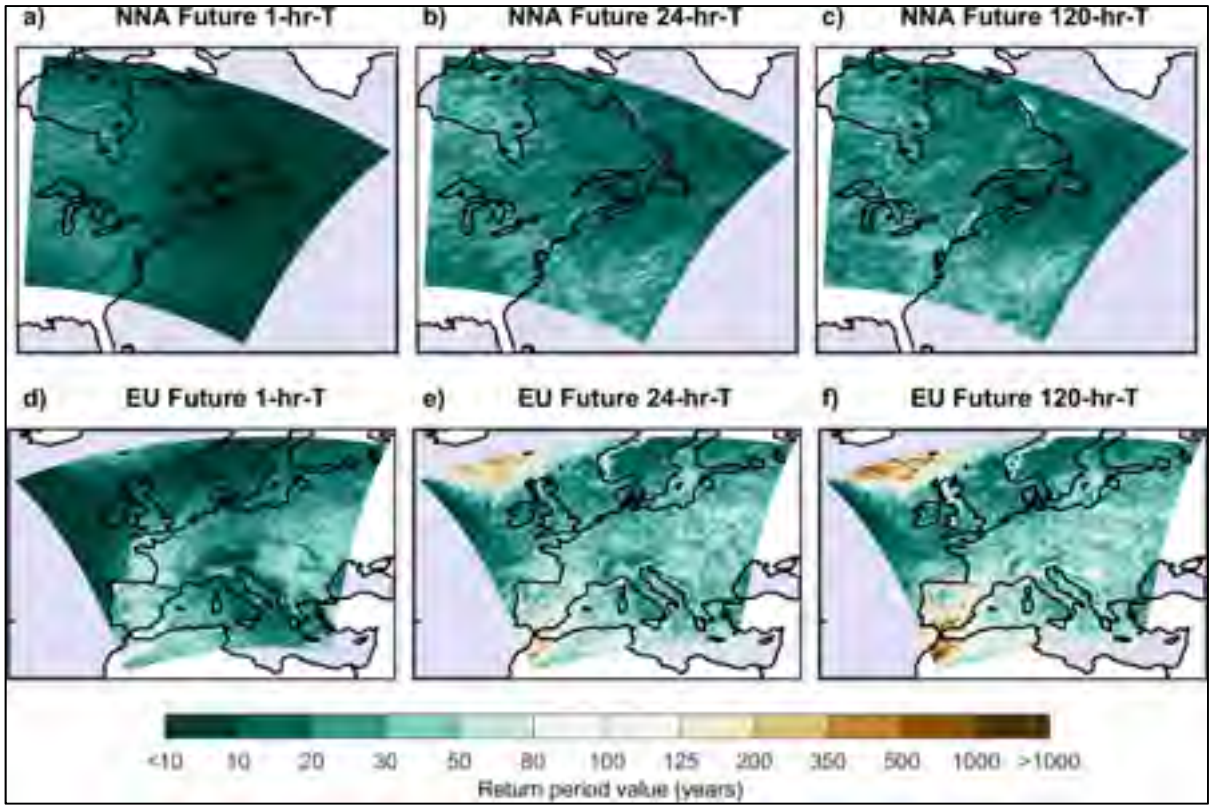

Figure 5.4 Projected return period over the 2080-2099 future period of the 1-hr (a,d), 24-hr (b,e) and 120-hr (c,f) AMP with same intensity as the 100-year AMP over the 1980-1999 period as simulated by CRCM5-LE for the NNA (a,b,c) and EU (d,e,f) domains

Significant increases in precipitation extremes over most of the EU domain were also observed although these increases are not as high as those seen over the NNA domain (Figure 5.4-d-f and Figure 5.5-b). For the 1-hr AMP (future 1-hr-T), the median future return period is 26 years, as compared to 35 and 42 years respectively for the 24-hr and 120-hr AMP. The greatest reductions in return period are observed in high-altitude regions (e.g., the European Alps; also see (Giorgi et al., 2016)), the Scandinavian countries and Ireland (with values ranging between 5-year and 10-year return periods for the 1-hr AMP). It should be

noted that the European Alps can easily be seen in Figure 5.4-d-f, especially for the 1-hr duration. Other high-altitude regions also stand out, such as the Pyrenees, the Dinaric Alps, the Balkan and the Carpathian Mountains (see Figure 5.4-d and Figure 5.1-b for high-altitude regions). Figure 5.5-c-d also shows that projected increases in extreme precipitation at the hourly and sub-daily scales are greater than for daily and multi-day extremes over both regions.

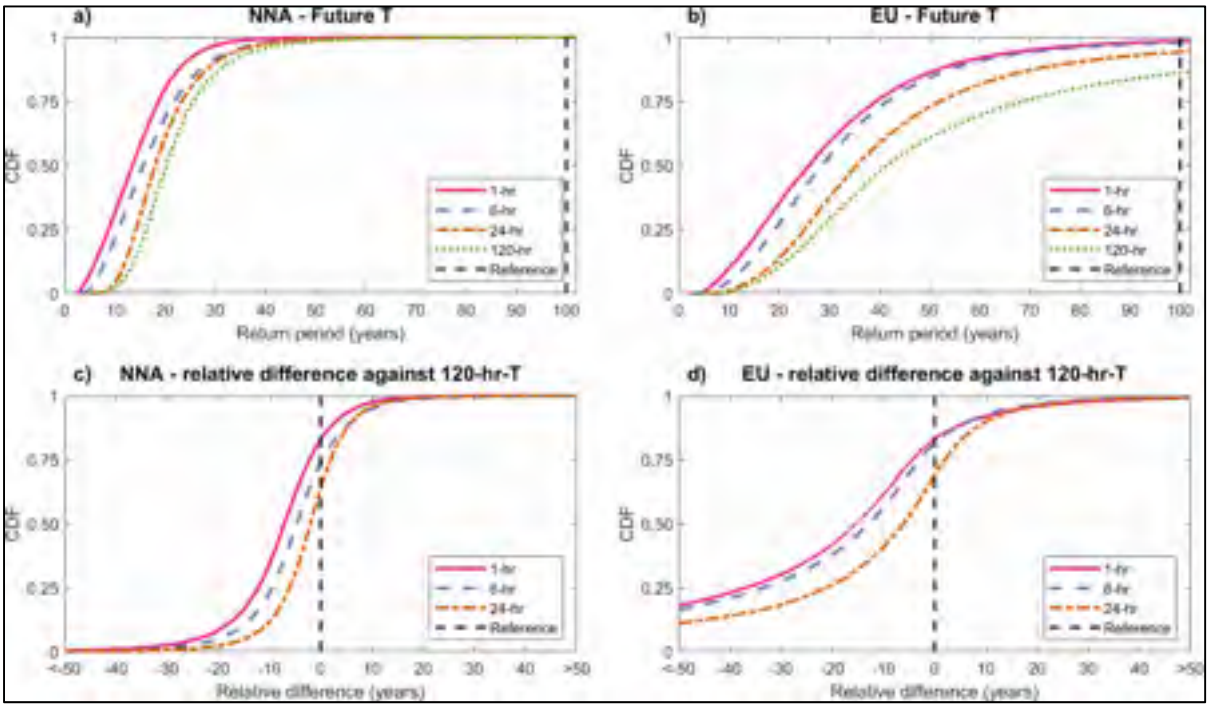

Figure 5.5 Cumulative distribution functions (CDF) of the projected return period over the 2080-2099 period of AMP with same intensity as the 100-year AMP over the 1980-1999 period for durations ranging from 1 hour to 120 hours (a,b) and the differences between grid point 120-hr-T values and corresponding 1-hr-T (pink), 6-hr-T (blue) and 24-hr-T (orange) (c,d) values. Only land grid points inside the NNA domain (a,c) and EU domain (b,d) were considered. The dashed vertical line corresponds to the 100-year return period. Values on x-axis longer than 100 years are not displayed on panels a and b and those smaller than 50-years are not displayed on panels c and d

Results from CRCM5-LE, CanESM2-LE, and CESM1-LE were then compared to assess the consistency of projected changes for 1-day and 5-day (24-hr and 120-hr for CRCM5-LE) AMP over the NNA and EU domains. A total of 81 (100) land grid points of CanESM2-LE

and 525 (662) for CESM1-LE were therefore considered for the NNA (EU) domains. The CDF of land grid point projected return periods over the 2080-2099 period for the three LE for both domains are presented in Figure 5.6. Maps of the projected changes are shown in Figure S5.11 and Figure S5.12 of Supplementary Materials.

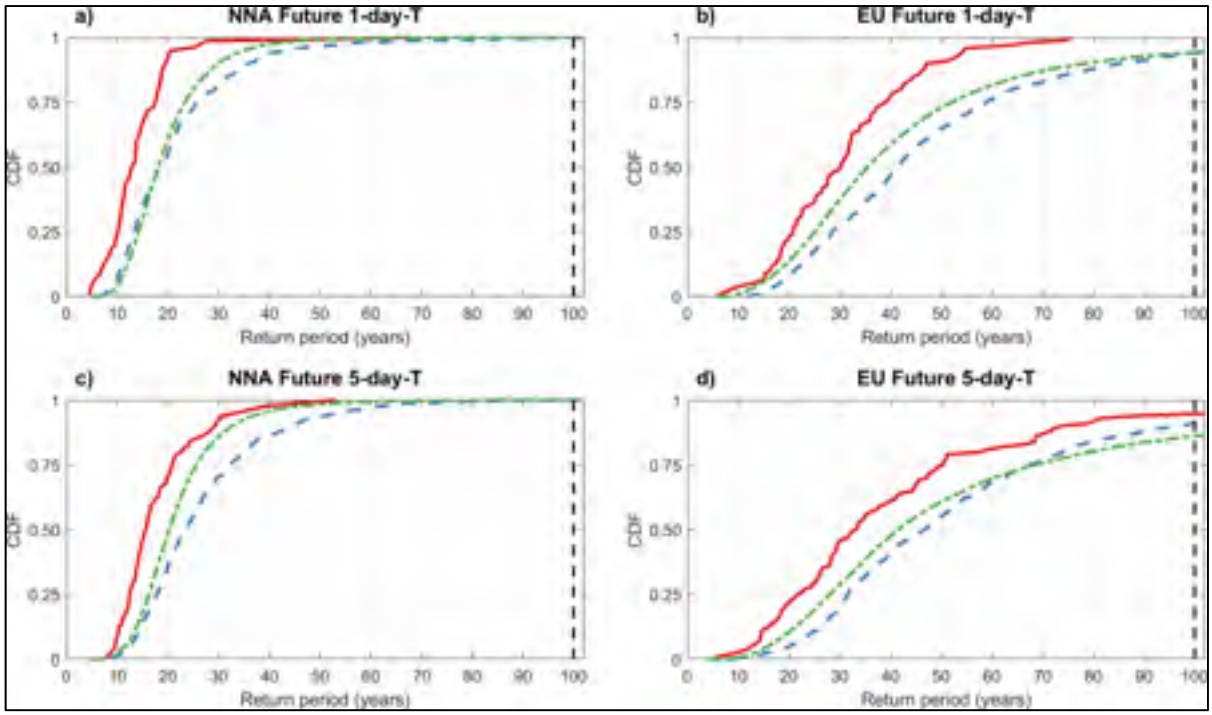

Figure 5.6 Cumulative distribution functions (CDF) of the projected return period over the 2080-2099 period for 1-day (24-hr) (a,b) and 5-day (120-hr) (c,d) AMP for CanESM2-LE and CESM1-LE (CRCM5-LE) with same intensity as the 100-year AMP over the 1980-1999 period. Only land grid points from CanESM2-LE (red), CESM1-LE (blue) and CRCM5-LE (green) over the NNA (a,c) and EU (b,d) domains were considered. The dashed vertical lines correspond to the 100-year return period. Values on x-axis longer than 100 years are not shown

Figure 5.6-a-c shows that all land grid points of the NNA domain experience increases in AMP (corresponding to decreases in the projected return period) according to the three LE for both 1-day (24-hr) and 5-day (120-hr) AMP. A very large fraction of land grid points similarly displays large increases in extreme precipitation for the EU domain (100%, 94%, 94% for 1-day (24-hr) AMP and 95%, 91%, 86% for 5-day (120-hr) AMP, according to CanESM2-LE, CESM1-LE, and CRCM5-LE, respectively). Also, for the NNA domain, all

ensembles point to large increases in AMP, with more than 4- to 5-fold increases in frequency for half of the land grid points. With respect to the EU domain, half of the land grid points display more than 2- to 3-fold increases in frequency of the reference 100-year AMP. In general, other studies based on different methodologies and on various extreme precipitation metrics have also reported significant increases over the NNA (Mailhot et al., 2012; Mailhot et al., 2007; Mladjic et al., 2011; Wehner, 2013; Wuebbles et al., 2013) and EU (Aalbers et al., 2017; Rajczak & Schär, 2017) domains, but in terms of relative changes for the different return periods investigated.

Figure 5.6 also shows that, on both domains, projected future return periods associated with 1-day (24-hr) and 5-day (120-hr) 100-year AMP in the reference climate as simulated by CanESM2-LE are shorter than corresponding values simulated by CESM1-LE and CRCM5-LE. It should be noted that despite large differences in terms of spatial resolution and model structures, projected changes are consistent across all three climate models on both domains for the 100-year return period AMP. It is also interesting to note that when projected changes are expressed in terms of changes in return period, the greatest changes in future 100-year return period AMP are estimated by CanESM2-LE, which has the coarsest spatial scale.

Figure 5.7 presents the CDF of the land grid point relative changes (1-day/24-hr and 5-day/120-hr) occurring between the 1980-1999 and 2080-2099 periods for 20- and 100-year AMP over both domains. This figure provides a complementary perspective to Figure 5.6. It shows the projected future increased in terms of AMP intensity of the reference 100-year AMP, while Figure 5.6 shows changes in terms of frequency. Complementary maps of relative changes in the 20- and 100-year AMP for Figure 5.7 are shown in the Supplementary Materials, Figure S5.13 to Figure S5.16.

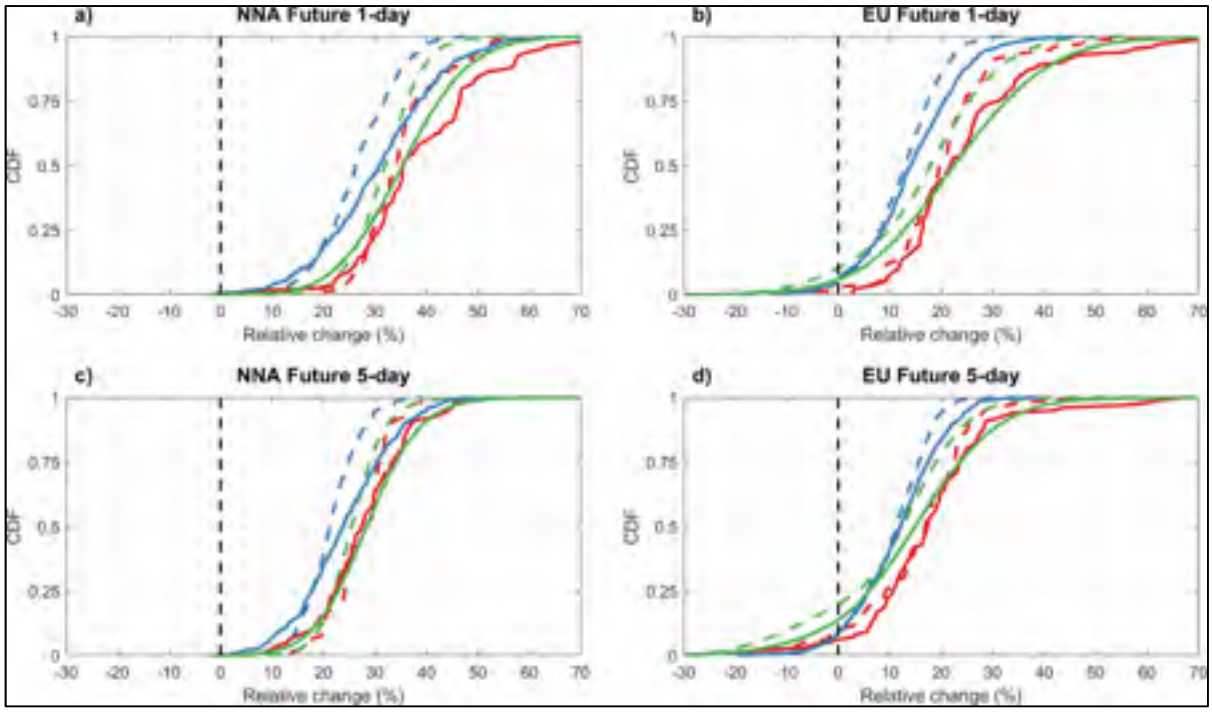

Figure 5.7 Cumulative distribution functions (CDF) of the projected relative changes (%) in the 20-year (dashed curves) and 100-year (continuous curves) AMP between the 1980-1999 and the 2080-2099 periods for 1-day (24-hr) (a,b) and 5-day (120-hr) (c,d) AMP for the CanESM2-LE and CESM1-LE (CRCM5-LE) over the NNA (a,c) and EU (b,d) domains. Distributions are shown for CanESM2-LE (red), CESM1-LE (blue) and CRCM5-LE (green) land grid points

Figure 5.7 shows consistent results over the NNA domain with all ensembles projecting similar increases for both 20- and 100-year, as well as 1-day (24-hr) and 5-day (120-hr) AMP. For instance, all three ensembles project that a larger proportion of land grid points experience greater increases for 100-year AMP than for 20-year AMP, and greater increases for 1-day (24-hr) AMP than for 5-day (120-hr) AMP. These are important results, as they suggest that climate change impact more extreme events (associated with long return periods) more severely, and have a more severe impact on daily AMP than on multi-day AMP (also supported by Pendergrass (2018)). Similar results can also be observed over the EU domain (Figure 5.7-b-d). Increases in that case are lower (with up to a quarter of the land grid points experiencing relative decreases), but still greater for longer return periods and daily AMP (see also Figure S5 and S6 for the 20-year AMP, S7 and S8 for the 100-year AMP in

Supplementary Materials). The picture over both domains is remarkably consistent for all three LE and for both durations (1-day/24-hr and 5-day/120-hr).

A comparison of grid point values from the three ensembles over the reference period (1980-1999; Figure S5.17 to Figure S5.21 in Supplementary Materials) shows that AMP values as simulated by CRCM5-LE are much greater than corresponding values estimated from CESM1-LE or CanESM2-LE for all durations, return periods and domains. Considering that the output from the different climate models are interpreted as spatial averages over each grid point, the spatial resolution can partly explain these differences in AMP values (Chen & Knutson, 2008; Sunyer et al., 2013), especially for 100-year return period AMP. Increasing the spatial resolution enables a better representation of local precipitation extremes, notably for summer convective storms (Maraun et al., 2010a; Prein et al., 2013; Tripathi & Dominguez, 2013). It is, however, interesting to note that despite these large differences in terms of simulated AMP quantiles in the reference climate, the three ensembles provide quite consistent changes in terms of daily and multi-day AMP increases over both domains.

5.4.3 Projected changes for some specific grid points

Considering the importance of assessing the impact of climate changes on sub-daily extreme precipitation in urban areas, grid points corresponding to six cities, three located in the NNA (Chicago, Halifax, Montreal) and three in the EU (Barcelona, London, Munich) domains were selected (Figure 5.1). Projected relative changes occurring between the 1980-1999 and 2080-2099 periods for AMP durations from 1 hour to 120 hours and return periods equal to 2, 20, and 100 years were assessed using CRCM5-LE (Figure 5.8). Bootstrapping with resampling (10000 samples) was used to estimate the 95% confidence interval.

Relative changes in the IDF curves shown in Figure 5.8 point to large increases for all cities, durations and frequencies (except for Barcelona, for the short return period and long duration AMP, and for London, where relative changes are almost identical for the three return periods), reaching up to a 71% increase for Halifax. However, while large relative increases

are observed, they differ strongly for each duration and frequency analyzed. In general, the relative increases tend to be larger as the return period increases or as the duration decreases. Furthermore, there are strong local variations going from one city to another.

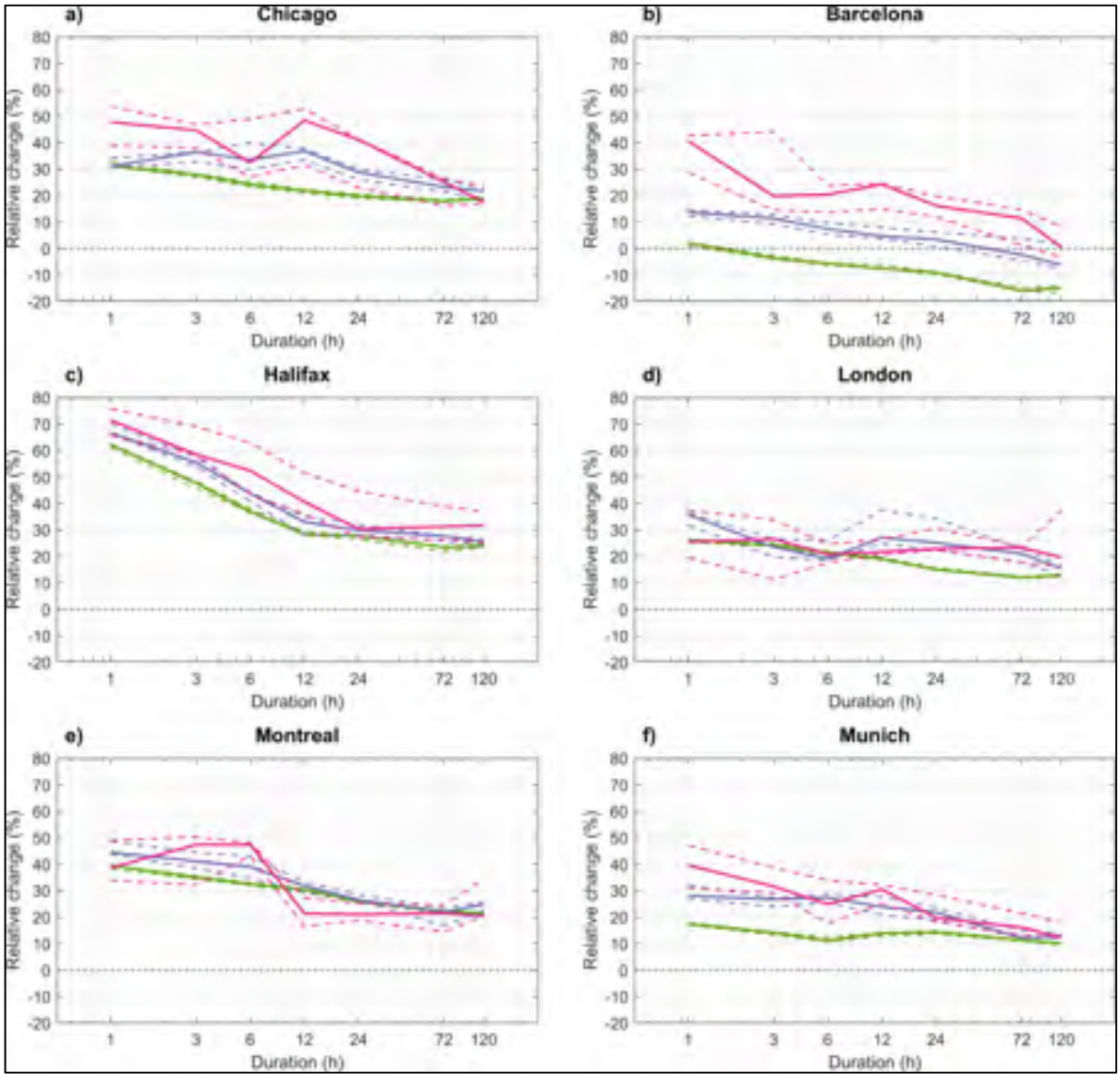

Figure 5.8 Relative change in the intensity-duration-frequency (IDF) curves between the 1980-1999 and 2080-2099 periods for grid points encompassing: (a) Chicago; (b) Barcelona; (c) Halifax; (d) London; (e) Montreal; (f) Munich as estimated from CRCM5-LE. Two-year (green), 20-year (blue) and 100-year (pink) return period AMP are shown. The dashed curves show the 95% empirical confidence interval

It should be noted that increases in AMP relative changes as durations decrease are not as smooth for the 100-year return period (e.g. for Chicago, relative increases for the 6-hr 100-year AMP are less than for 12- and 24-hr AMP increases). This may be due to larger uncertainties (sampling errors) in empirical AMP quantile estimates in this case. These can be significant, even when 1000-year series are used to empirically estimate the 100-year AMP.

Although these projected relative changes are based on a single land grid point series (some regional analyses could provide a more regional picture and reduce sampling uncertainties for the 100-year AMP), they show that large increases in AMP can be expected at local scales, and should be accounted for, even if these estimates remain uncertain.

5.5 Concluding remarkds

The three large ensembles (LE) of climate simulations, two global (the 40-member CESM1-LE and the 50-member CanESM2-LE), and one based on a regional model (the 50-member CRCM5-LE), were considered in this study. CRCM5-LE simulations are available over two domains, one covering Northeastern North America (NNA) and the other covering Europe (EU). The CRCM5-LE regional ensemble was generated by dynamically downscaling the CanESM2-LE over the 1950 to 2100 period. The RCP8.5 forcing scenario was considered for all three ensembles.

The three ensembles point to a significant reduction in the projected return period (corresponding to an increase in intensity) over the 2080-2099 period of the 100-year AMP over the 1980-2000 period at both global and regional scales. The 100-year annual maximum precipitation (AMP) is more frequent in future climates for most land grid points, with up to a 10-fold frequency increase for some grid points. Over a given state or country, this means that larger populations and more cities, towns, municipalities face larger extreme rainfall events.

Future return periods of 100-year 1-day and 5-day AMP estimated from CanESM2-LE and CESM1-LE are consistent over most of the land grid points, aside from some discrepancies over subtropical and tropical regions. At the regional level, on the Northeastern North America and Europe domains, CRCM5-LE also showed similar changes in future return periods for durations ranging from 1 hour to 120 hours.

More specifically, at the global scale, using both CanESM2-LE and CESM1-LE, 100-year 1-day and 5-day AMP over the reference period become 4 to 5 times more frequent by the end of this century for half of the world's land grid points. At the regional scale, all three LE project major increases in the 100-year 1-day and 5-day (24-hr and 120-hr for the CRCM5-LE) AMP, with more than half of land NNA (EU) grid points experiencing a 4- to 5-fold (2- to 3-fold) increase in frequency. Despite having different model structures and resolutions, all three ensembles point to a strong reduction in the projected return period of the reference 100-year AMP for the 1-day and 5-day durations.

Results suggest that complex orography plays an important role in the projected changes. Greater reductions in return periods for high-altitude regions are observed within the regional model (e.g., the Appalachian Mountains in the NNA domain and the Alps in the Europe domain), notably for the 1-hour AMP. This shows the "added value" of higher spatial resolution when assessing regional changes in extreme precipitation. The impact of complex topography on the projected changes in precipitation extremes should be further investigated in future work.

Looking more specifically at the impact of AMP durations on projected changes for the regional CRCM5-LE, it was shown that short-duration 100-year AMP experiences longer return period reductions. Therefore, the model results imply that the impact of climate changes are greater on short-duration AMP.

Looking at the relative changes in IDF curves at some specific sites, it was shown that higher relative increases can be expected for shorter durations and longer return periods with strong

local variations (from city to city). For example, projected increases in intensity/frequency of the 1-hour 100-year rainfall are greater than for the 6-hour 20-year rainfall. These results suggest that no unique safety factor (i.e. a ratio that expresses how much stronger infrastructures must be to cope with an intended load) can account for the impact of climate change on extreme precipitation over all duration, return periods and locations.

These large increases in the frequency of extreme precipitation events have important implications since the design of many types of infrastructure relies on estimations of extreme precipitation for various durations and return periods. Based on the results highlighted in this work (which corroborate those from other recent studies), climate change adaptation strategies and design criteria must be revised to account for these expected large reductions in the projected future return periods for precipitation extremes.

Several limitations of this study should be investigated in future work. A single regional model was used to explore the sub-daily time scale. There are, however, very few examples of such large ensembles run at the regional scale (Aalbers et al., 2017; Mizuta et al., 2017). Analyses combining many LE generated from various RCM/GCM combinations should help in assessing the impact of climate model structure and resolution on projected sub-daily extreme precipitation. Multi-model analysis is important since it is expected that inter-model variability is likely one of the main sources of uncertainties for long-term projections of extreme precipitation. Initiatives such as the HighResMIP (Haarsma et al., 2016) could provide the necessary datasets to further investigate this source of uncertainty.

Despite CRCM5's relatively high spatial resolution (0.11°), many processes, such as deep convection, occur at spatial scales too small to be resolved explicitly at the model grid scale, and are therefore parametrized. While synoptic weather patterns are generally well simulated at this resolution, a finer resolution is required to comprehensively simulate small scale convective events (Chan et al., 2014; Kendon et al., 2017; Prein et al., 2015; Prein et al., 2017). Convective-permitting models (CPM) are therefore needed to resolve sub-daily convective extreme precipitation events, and confirm the results obtained from convection-

parametrized regional models for sub-daily extreme precipitation (Prein et al., 2015). There are also new efforts to build on suggestions in Prein et al. (2015) and generate multi-model ensembles at these scales (see Coppola et al., 2018).

Only the RCP8.5 forcing scenario, which is a high-end emission scenario (IPCC, 2013; Meinshausen et al., 2011), was considered in this study. Clearly, other forcing scenarios (e.g., RCP2.6 or RCP4.5) should be considered in order to assess the sensitivity of the projected changes to anthropogenic forcing. This is even more important as emission scenarios play an important role for the more distant future periods, as considered in this work. Some authors suggest that the most likely future emission scenario is probably closer to the middle range of forcing scenarios currently proposed, such as RCP4.5 or RCP6.0 (Raftery, Zimmer, Frierson, Startz, & Liu, 2017).

5.6 Acknowledgements

This work was partly financed through the ClimEx project funded by the Bavarian State Ministry for the Environment and Consumer Protection. The authors acknowledge the contributions from the Canadian Centre for Climate Modelling and Analysis (Environment and Climate Change Canada, ECCC) for simulating and making available the CanESM2-LE used in this study, and the Canadian Sea Ice and Snow Evolution Network for proposing the simulations. The authors would also like to thank the Ouranos Consortium for helping with data transfer. The CanESM2-LE dataset is now available on the ECCC website (http://collaboration.beta.cmc.ec.gc.ca/cmc/cccma/CanSISE/output/CCCma/CanESM2/). The CESM1 ensemble was downloaded from the Large Ensemble Community Project (LENS) website (http://www.cesm.ucar.edu/projects/community-projects/LENS/).

The CRCM5 was developed by the ESCER centre at Université du Québec à Montréal (UQAM; www.escer.uqam.ca) in collaboration with ECCC. Computations with the CRCM5 for the ClimEx project were made on the SuperMUC supercomputer at the Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities. The

operation of this supercomputer is funded via the Gauss Centre for Supercomputing (GCS) by the German Federal Ministry of Education and Research and the Bavarian State Ministry of Education, Science and the Arts,

5.7 Supplementary materials

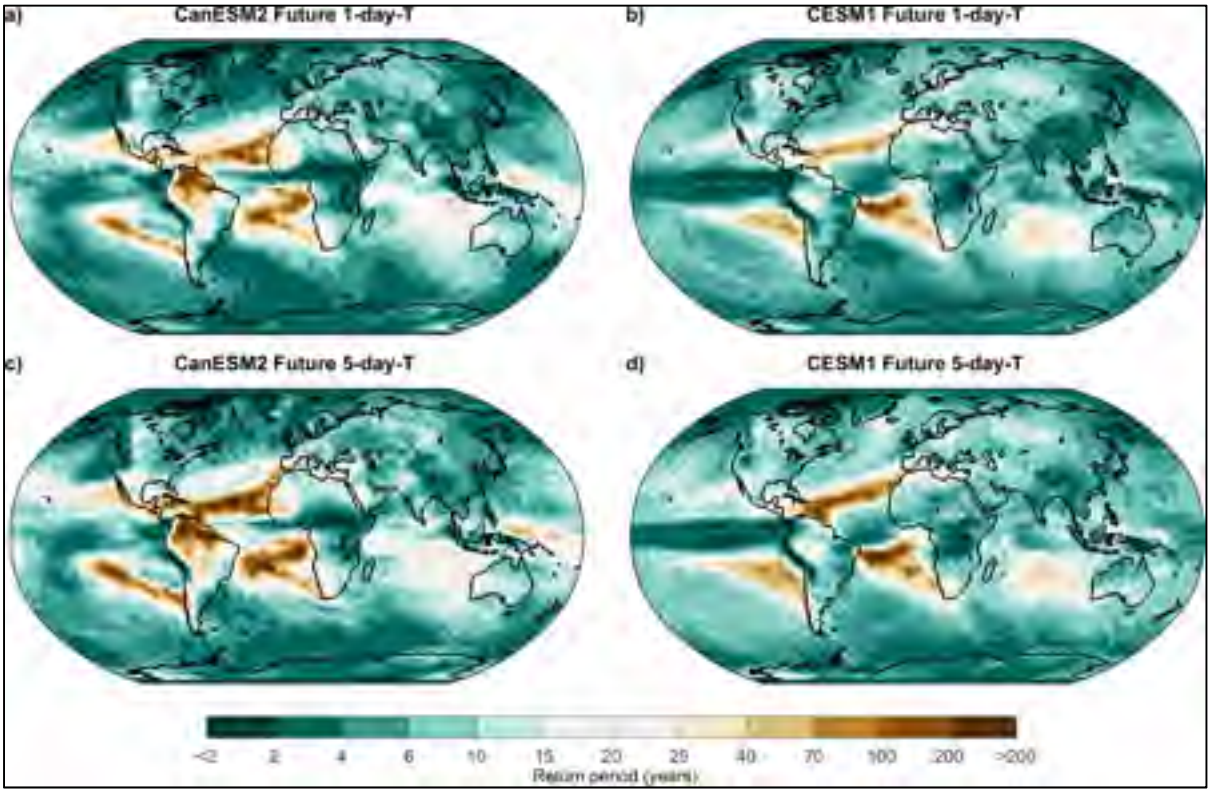

Figure S5.9 Projected return period over the 2080-2099 period for 1-day (a,b) and 5-day (c,d) AMP with same intensity as the 20-year AMP over the 1980-1999 period for CanESM2-LE (a,c) and CESM1-LE (b,d)

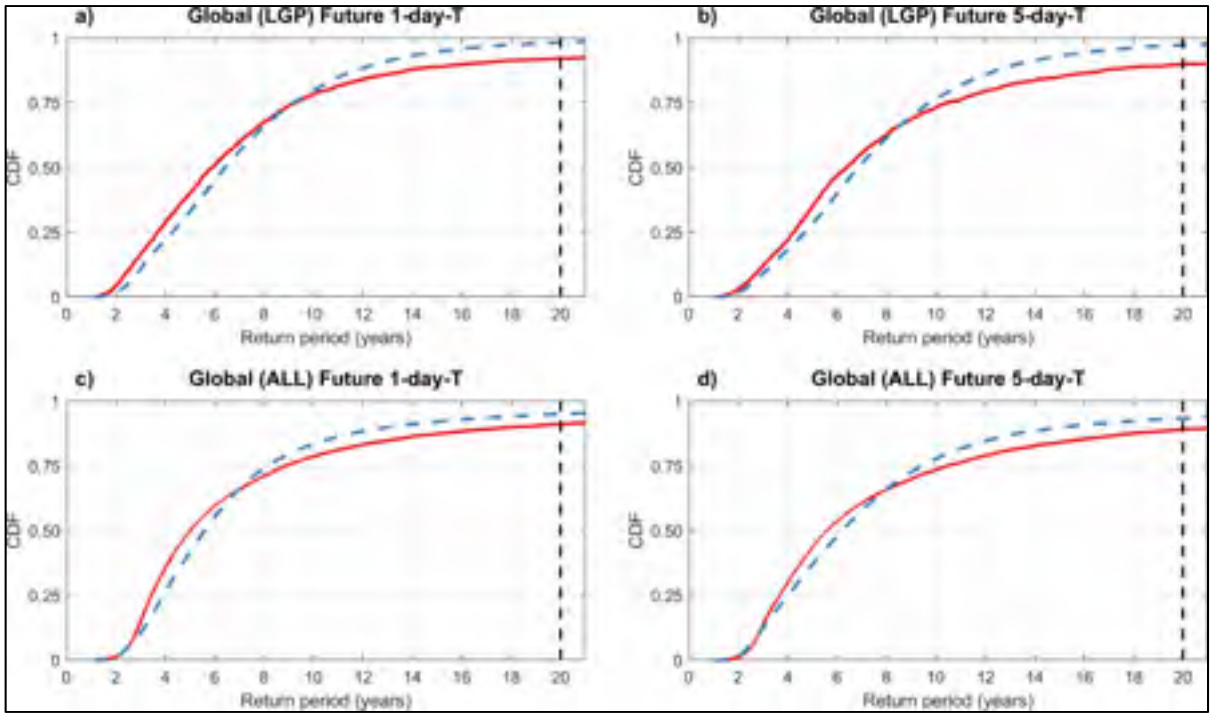

Figure S5.10 Cumulative distribution functions (CDF) of the projected return period over the 2080-2099 period of the 1-day (a,c) and 5-day (b,d) AMP with same intensity as the 20-year AMP over the 1980-1999 period. Distributions are shown for land grid points (LGP; a,b) and for all grid points (ALL; c,d) from CanESM2-LE (red curve) and CESM1-LE (blue curve) at the global scale. The dashed vertical line corresponds to the 20-year return period. Values on x-axis longer than 20 years are not displayed

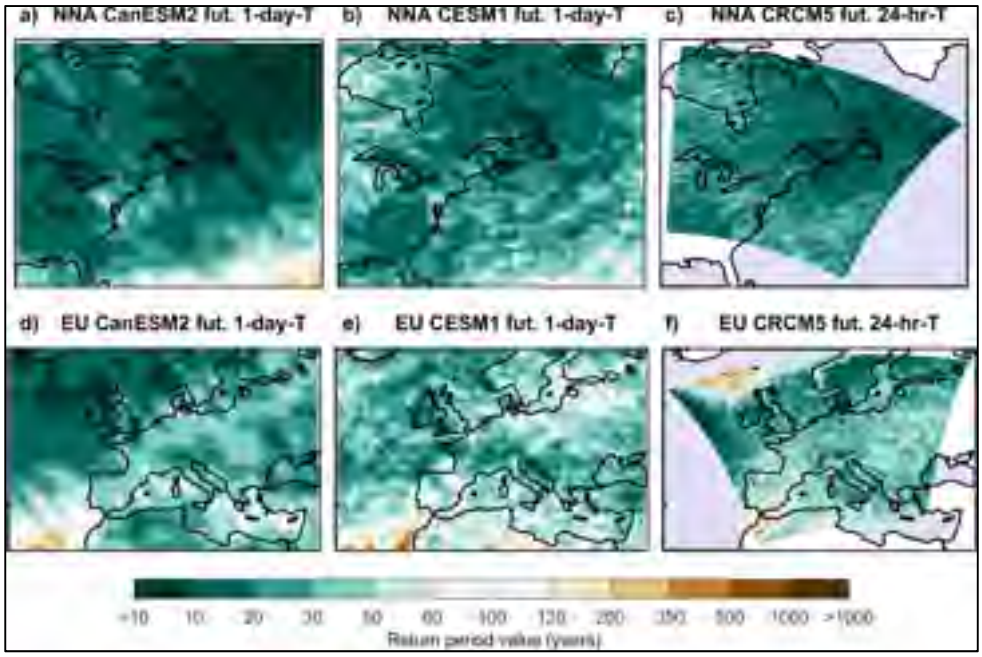

Figure S5.11 Projected return period over the 2080-2099 period for 1-day (24-hr) AMP with same intensity as the 100-year AMP over the 1980-1999 period as simulated by (a,d) CanESM2-LE, (b,e) CESM1-LE and (c,f) CRCM5-LE for the NNA (a,b,c) and EU (d,e,f) domains

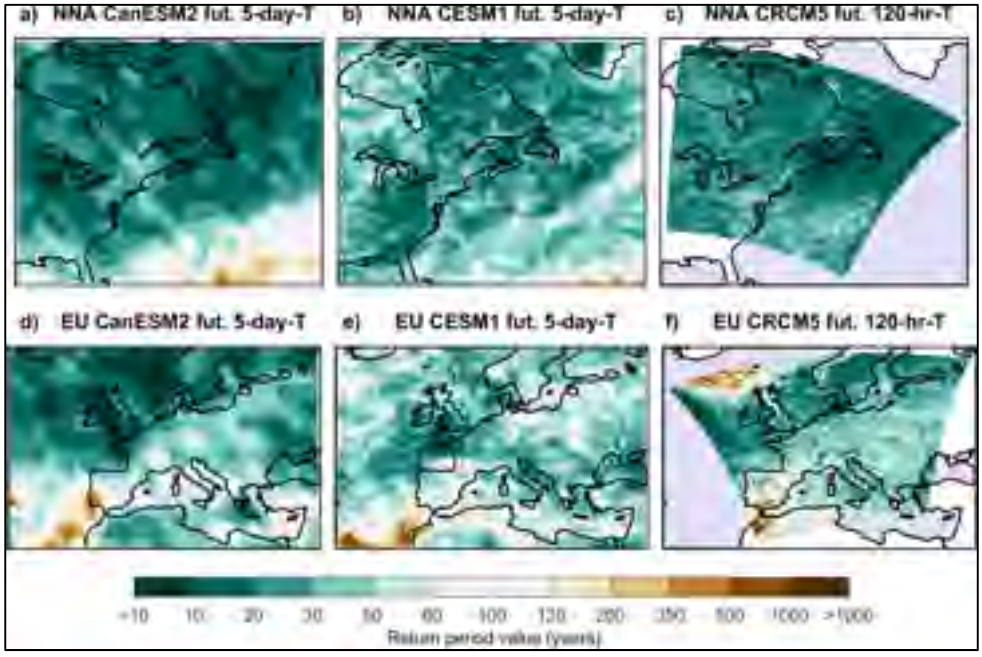

Figure S5.12 Same as Figure S5.11 for 120-hr AMP

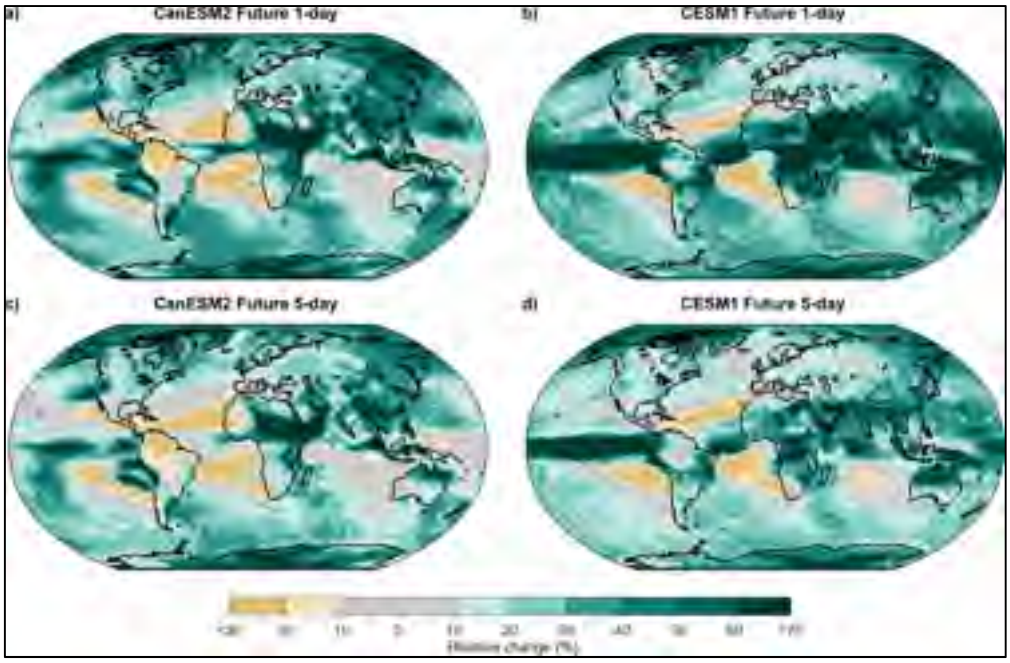

Figure S5.13 Relative changes (%) of the 20-year AMP between the 1980-1999 and the 2080-2099 period for 1-day (a,b) and 5-day (c,d) AMP using CanESM2-LE (a,c) and CESM1-LE (b,d)

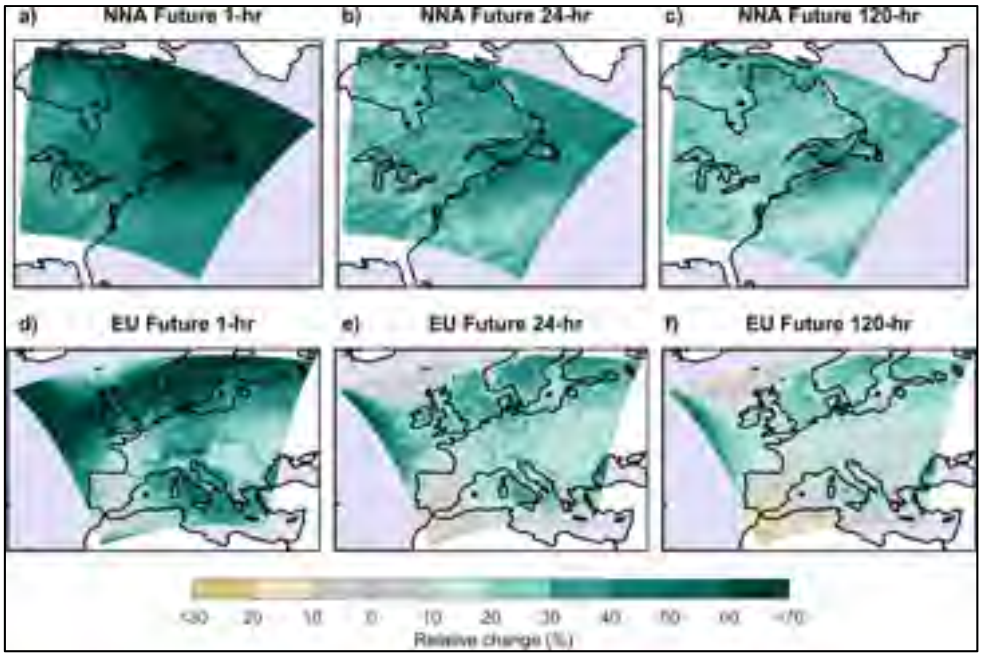

Figure S5.14 Relative changes (%) of the 20-year AMP between the 1980-1999 and the 2080-2099 period for 1-hr (a,d), 24-hr (b,e) and 120-hr (c,f) AMP over the NNA (a,b,c) and EU (d,e,f) domains

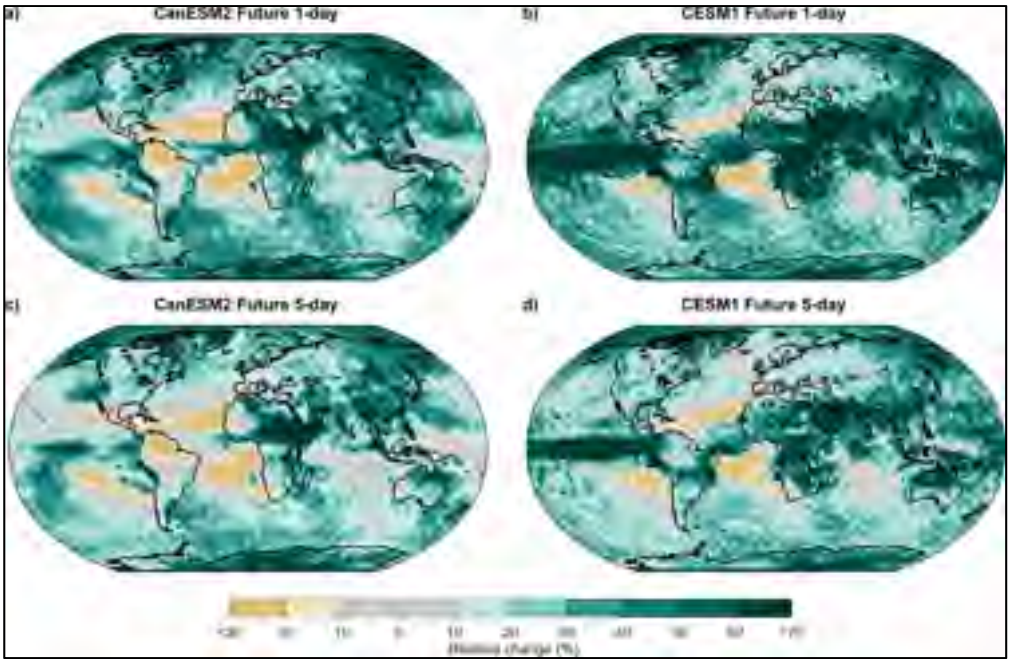

Figure S5.15 Same as Figure S5.13, but for the 100-year AMP

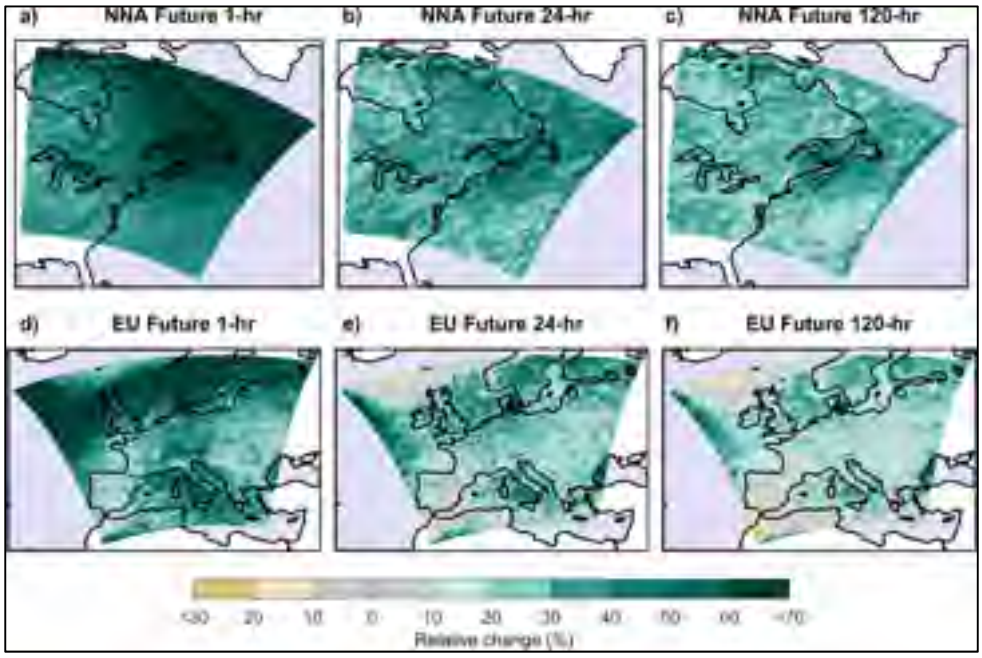

Figure S5.16 Same as Figure S5.14, but for the 100-year AMP

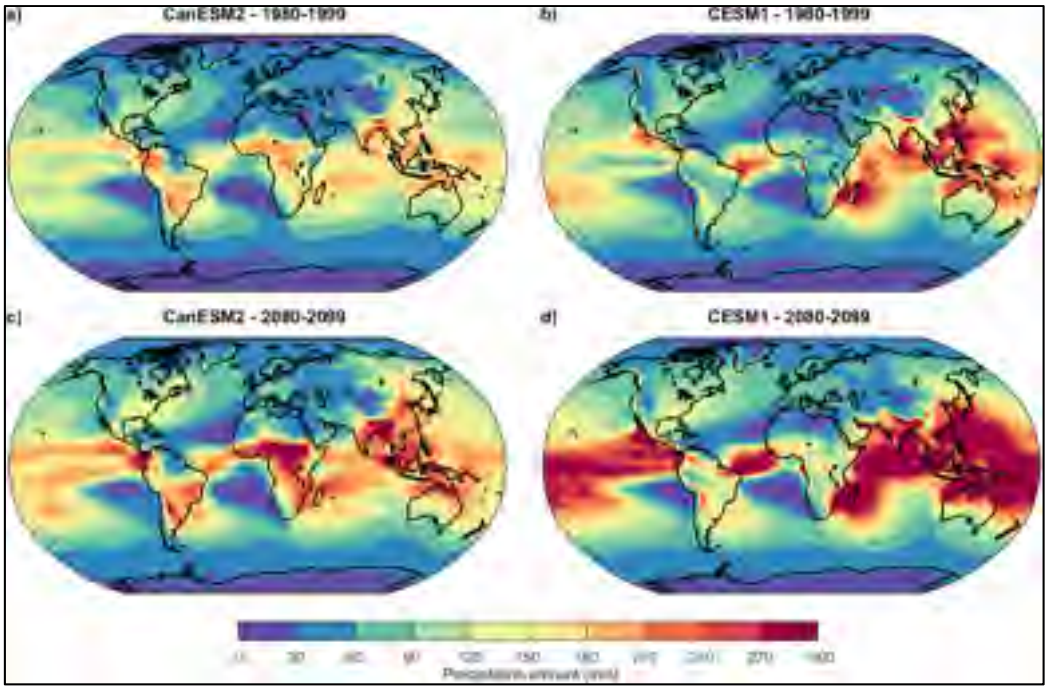

Figure S5.17 Precipitation amounts (mm) for the 100-year 1-day AMP over the 1980-1999 period (a,b) and 2080-2099 period (c,d) for CanESM2-LE (a,c) and CESM1-LE (b,d)

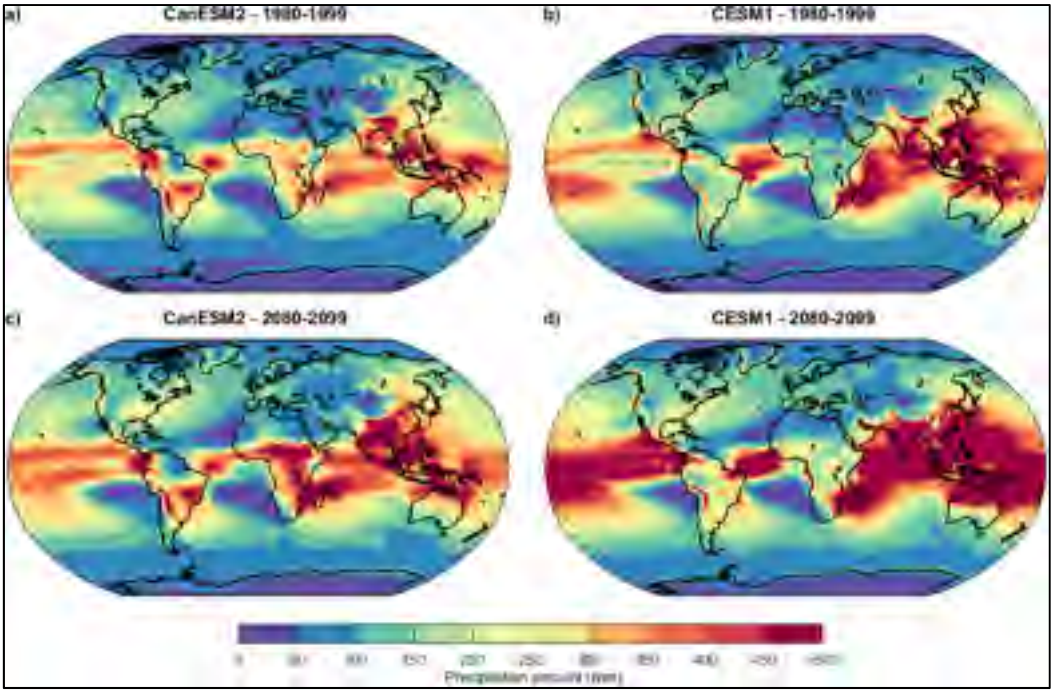

Figure S5.18 Same as Figure S5.17 for the 5-day AMP

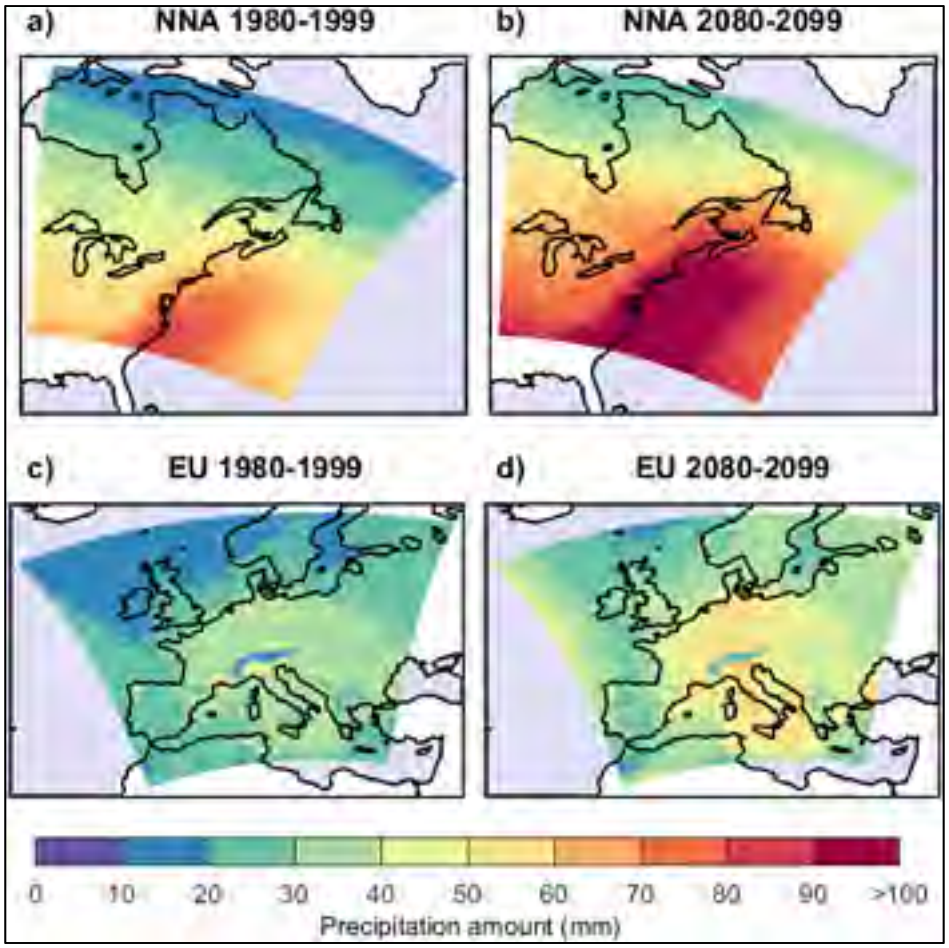

Figure S5.19 Precipitation amounts (mm) for the 100-year 1-hr AMP over the 1980-1999 period (a,c) and 2080-2099 period (b,d) for the NNA (a,b) and EU (c,d) domains as simulated by CRCM5-LE

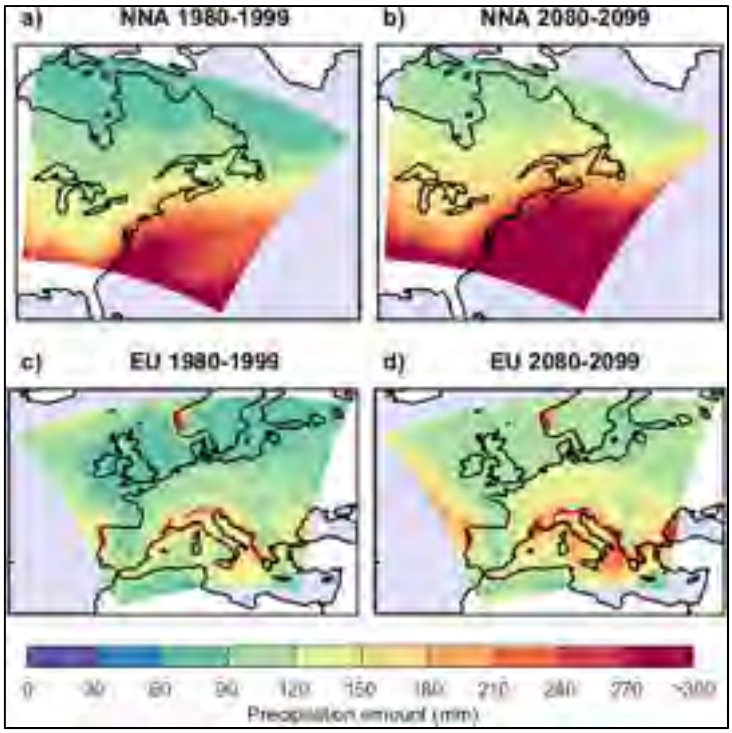

Figure S5.20 Same as Figure S5.19, but for the 24-hr AMP

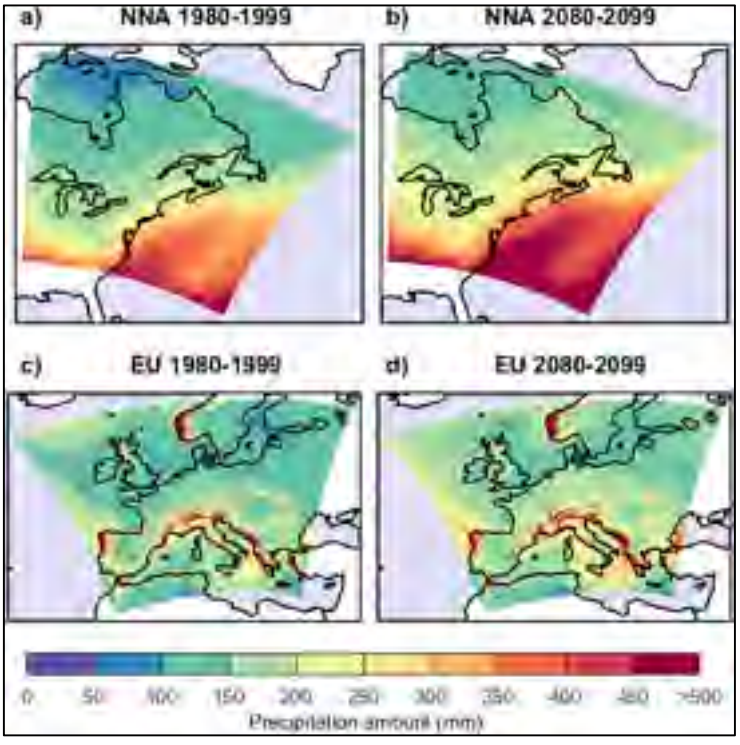

Figure S5.21 Same as Figure S5.19, but for the 120-hr AMP

CHAPITRE 6

IMPACTS OF CLIMATE CHANGE ON DAILY EXTREME STREAMFLOW ESTIMATED FROM TWO LARGE CLIMATE SIMULATION ENSEMBLES OVER 3567 NORTH AMERICAN CATCHMENTS

Jean-Luc Martel[a], François Brissette[b], Alain Mailhot[c]

[a, b] Département de Génie de la construction, École de technologie supérieure
[c] Institut national de la recherche scientifique – Eau, Terre et Environnement

Article en préparation pour soumission à la revue « Nature Climate Change »

Abstract

With the consensus emerging towards projected increases in rainfall extremes in response to human-induced global warming, there is a need to better assess the evolution of future extreme flood events. The impacts of climate change on extreme flooding was investigated using two large ensembles of climate simulations over 3 567 medium-to-large size North America catchments. The use of two large ensembles allowed for an accurate estimation of the 100-year flood in current and future climates. Results show very distinctive spatial patterns with both increasing and decreasing annual maximum of daily streamflow in all regions of North America by the end of the 21st century. The largest changes were found to be decreases in floods generated from the snowmelt in catchments located in high latitudes and/or altitudes and increases for catchments located within the Southeast region of the United States and the West Coast where rainfall is the leading cause of flooding.

The design of water infrastructures such as bridges and dams, as well as flood protection through the delineation of floodplains is largely based on statistical assessments of extreme floods. The intensity of such events (e.g., 100-year flood event) is typically estimated through a flood frequency analysis using available observational records. In most cases, these analyses assumed that recorded time series are stationary. Considering the scientific consensus regarding the human-induced global warming (IPCC, 2013), this hypothesis needs to be revisited (Milly et al., 2008b; Wagener et al., 2010).

A considerable number of recent studies using observational records is pointing towards an increase in the frequency and intensity of both daily (IPCC, 2013; Min et al., 2011; Westra et al., 2013) and sub-daily (Madsen et al., 2014; Westra et al., 2014) precipitation extremes due to anthropogenic forcing. This claim has been reinforced by several studies exploring the projected changes of extreme precipitation events using one or multiple climate models at the global (Kharin et al., 2007; Kharin et al., 2013) and regional scales (Mailhot et al., 2012; Mailhot et al., 2007; Wuebbles et al., 2013; Zhu, 2013). This consensus towards an increase in the frequency and intensity of extreme precipitation events also has evident implications for extreme flood events (Burn & Whitfield, 2016; Kundzewicz et al., 2014).

However, when it comes to streamflow, there is still a lack of evidence in the observed change in terms of intensity and frequency, and even in the direction of the change (Kundzewicz et al., 2014; Kundzewicz et al., 2007; Trenberth et al., 2007). This is not surprising since the expected response to human-induced climate change poses a bigger challenge due to the resulting changes in several components of the hydrological cycle (Bates et al., 2008; Kundzewicz et al., 2014; Madsen et al., 2014), for which streamflow acts as an integrator. Beside temperature, the most important component affected by climate change is precipitation through its various characteristics, such as intensity, duration and state (rain or snow). Changes in other variables, such as temperature, are expected to impact the conditions required to generate extreme events, with antecedent soil moisture, evapotranspiration and snow cover playing crucial roles at the regional scale.

Currently, the strongest evidence in observed changes in flood events are over catchments with important snow cover, where the increase in temperature is likely to trigger an earlier and longer snowmelt period resulting in lower peak floods (Groisman et al., 2001; Li et al., 2017; Madsen et al., 2014; Seneviratne et al., 2012; Stewart et al., 2005; Zhang et al., 2001). Over catchments where rainfall is the main cause of flooding, the tendency is towards wetter conditions with mixed patterns of increases and decreases in annual maximum streamflow (Burn & Whitfield, 2016; Groisman et al., 2001; Lins & Slack, 1999; Madsen et al., 2014; McCabe & Wolock, 2002; Rice et al., 2015). With respect to large return periods, projected future changes suggest similar conclusions, where large discrepancies in terms of increases and decreases are obtained over all study sites (Dankers & Feyen, 2008; Dankers et al., 2009; Rojas et al., 2012; Shrestha, Cannon, Schnorbus, & Zwiers, 2017).

The goal of this paper is to investigate the projected changes of rare extreme flood events (i.e. the 100-year return period events) over North America by the end of the 21st century. This return period level was selected since it is traditionally used for floodplain mapping and for the design of most major water infrastructures.

The projected return period of the current 100-year flood (1990-1999 period) was estimated over the 2080-2099 time horizon. A selection of 3 567 medium to large size catchments (minimum, median and maximum sizes being respectively 500, 1 797 and 1 074 127 km^2; Figure S6.7) over Canada and the continental United States (US) allowing a large coverage of North America's climatological regions. The lower limit was chosen to only include catchments with a response time commensurate with the daily time resolution of the climate ensembles. Consequently, results presented in this work should not be extended to smaller catchments, and particularly to urbanized catchments whose flooding behaviour is strongly dependent on sub-daily convective rainfall. Data from two large ensembles of climate simulations using the RCP8.5 emission scenario (van Vuuren et al., 2011) were used: the Canadian Earth System Model (CanESM2) 50-member large ensemble (CanESM2-LE) (Arora et al., 2011; Sigmond & Fyfe, 2016) at a 2.8° resolution and the Community Earth

System Model (CESM1) 40-member large ensemble (CESM1-LE) (Kay et al., 2015) at a 1° resolution. Using these large datasets (20 years × *n* members) allowed for the reduction of the sampling uncertainty and to reliably estimate the 100-year flood without resorting to the extreme value theory.

6.1 Projected changes in streamflow extremes

The projected changes of the 100-year flood are shown in Figure 6.1. An increase (decrease) in the frequency of the 1990-1999 flood event translates into a decrease (increase) of the corresponding 2080-2099 return period. Hence, a future return period of 25 years of the current 100-year flood, indicates a fourfold increase in its frequency.

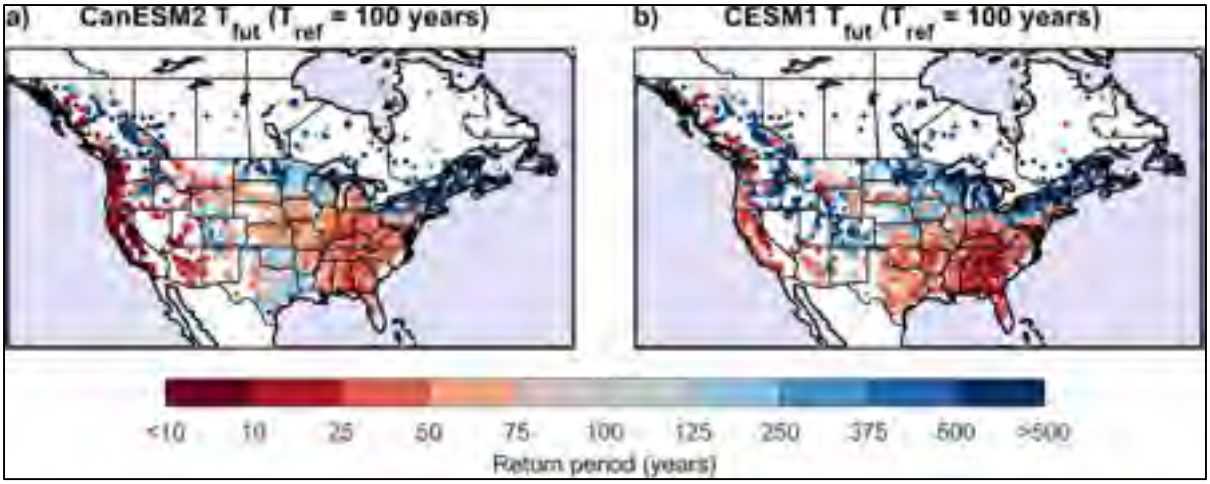

Figure 6.1 Projected future return periods (2080-2099) corresponding to the 100-year flood event over the reference period (1990-1999) estimated by the CanESM2-LE (a) and the CESM1-LE (b) for the 3 567 North American catchments. The dots are located at the centroid of each catchment

Results shown in Figure 6.1 indicate clear and distinct regional spatial patterns of both projected increases and decreases of the frequency of the corresponding reference extreme flood event. It can be noted that the main discrepancies between both ensembles are observed over the South of the US (notably over the state of Texas). For this region in particular, low

agreement was also previously obtained between different climate models with respect to projected changes in mean precipitation (IPCC, 2013). Analyses for the 20-year return period were conducted and showed similar patterns in terms of projected changes, but with more important increases for the 100-year return period (Figure S6.8). Overall, both ensembles are in strong agreement with respect to the spatial distribution of the projected future changes of the reference 100-year return period. Analyses of to the impact of the catchment drainage area were also performed but showed no significant differences in the spatial patterns of projected increases and decreases.

The largest decreases in frequency were obtained over high-latitude and/or high-altitude mountainous catchments (notably along the Rockies). For these catchments, it is expected that the change in the snowmelt regime influences both the timing and intensity of the annual maximum flood event. This would be consistent with both the projected increase in global temperature (IPCC, 2013), and the reported decreases of flooding over catchments with snowmelt-generated floods (Groisman et al., 2001; Li et al., 2017; Madsen et al., 2014; Seneviratne et al., 2012; Stewart et al., 2005; Zhang et al., 2001).

The largest increases in flood frequency are obtained all over the Southeast region of the US as well as the West Coast, and to some extent, over some of the mountain ranges such as the Sierra Nevada and Cascades. Since flooding events over these catchments are generally not triggered exclusively by snowmelt (except for the mountain ranges catchments), but rather by strong and persistent rainfall events and antecedent soil moisture, it is likely that increases in extreme rainfall events and wetter conditions explain this result. This is also supported by observed and projected increases in rainfall extremes (IPCC, 2013; Kharin et al., 2007; Kharin et al., 2013; Wuebbles et al., 2013) (see also Figure S6.9 for the projected change in non-bias corrected precipitation extremes in both ensembles) and probable wetter conditions for these regions (Burn & Whitfield, 2016; Lins & Slack, 1999; McCabe & Wolock, 2002; Rice et al., 2015). Regarding the catchments located in the mountain ranges (e.g., the Coastal Range and the Sierra Nevada), major changes in the hydrological cycle seem to be

responsible of the reduction in the future return periods. Overall, these results are further investigated through the analysis of changes in runoff volume (Figure 6.2) and snowpack (Figure 6.3).

6.2 Projected changes in runoff volume

Projected changes in annual runoff volume are first investigated to provide a better understanding of changes shown in Figure 6.1. Total annual runoff volume (ΔV) was directly computed based on the hydrological model daily streamflow outputs. The differences between the median reference and future values of the ΔV (which corresponds to the 2-year return period) are shown in Figure 6.2.

The total annual water balance can be roughly detailed as V = P – ET, where V is the runoff volume, P the precipitation and ET the evapotranspiration. Thus, a more detailed breakdown of the water balance based on differences for precipitation (ΔP) and evapotranspiration (ΔET) is also provided in Figure S6.10.

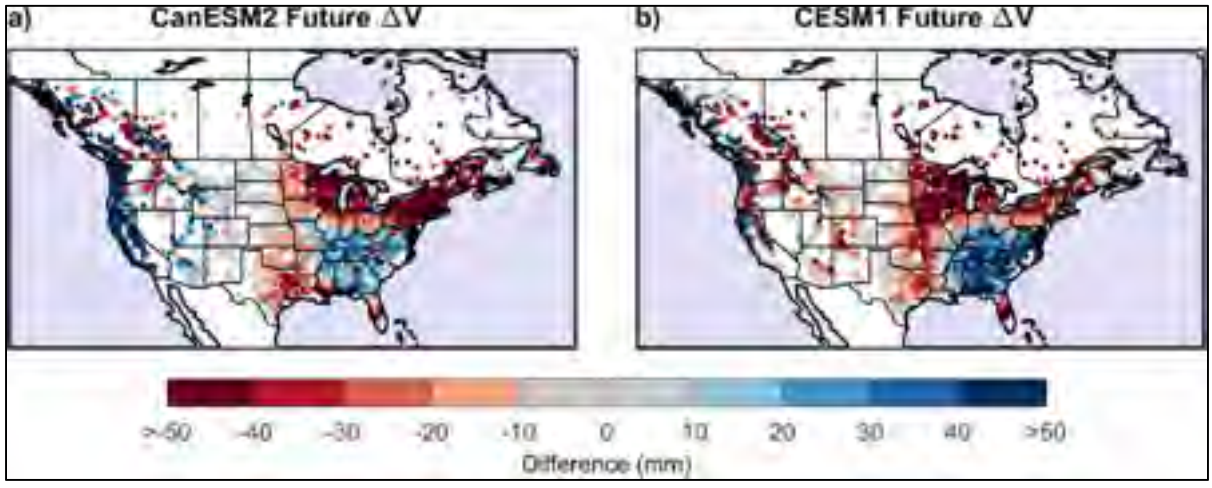

Figure 6.2 Same as Figure 6.1, but for the changes in the total annual runoff volume (ΔV)

The Southeast region of the US and the West Coast see a strong positive ΔV. This is very likely due to the superior increase in ΔP compared to ΔET for both ensembles (notably for CESM1-LE in the Southeast region of the US and CanESM2-LE for the West Coast – Figure S6.10). It is also notable that the frequency and intensity of extreme rainfall events are also expected to increase over these two regions (Figure S6.9). With increases in average total annual precipitation, there is a higher probability of reaching wetter antecedent soil moisture prior to an extreme rainfall event, which would likely contribute to the increase in frequency and intensity of extreme flood events. Overall, the regions with decrease in the projected return periods presented in Figure 6.1 show a strong correlation with this spatial feature the increase in ΔV.

At the other end of the spectrum, the Northeast region above the 40th parallel, high-altitude mountainous catchments (notably in Canada), and the states of Texas, Arizona and Florida are all experiencing strong negative ΔV. This can be explained by the larger increase in ΔET compared to ΔP for these regions (Figure S6.10). In general, this behaviour is also relatively well correlated with regions of increases in the projected return periods (Figure 6.1).

Overall, there is little discrepancy between the two ensembles with respect to the expected direction of the change. In terms of intensity, increases tend to be larger for CESM1-LE in the Southeast region of the US and for CanESM2-LE in the West Coast. There is a strong correlation between results presented in Figure 6.1 and Figure 6.2.

6.3 Projected changes in snow cover

In order to further investigate the projected changes over catchments where the annual maximum streamflow are dominated either by snowmelt or a mixed combination of rain over snow events, an evaluation of the projected change in the catchments' snowpack was carried out. Figure 6.3 presents the reduction in the annual maximum snow water equivalent median value (SWE_{max} – corresponding to a 2-year return period value).

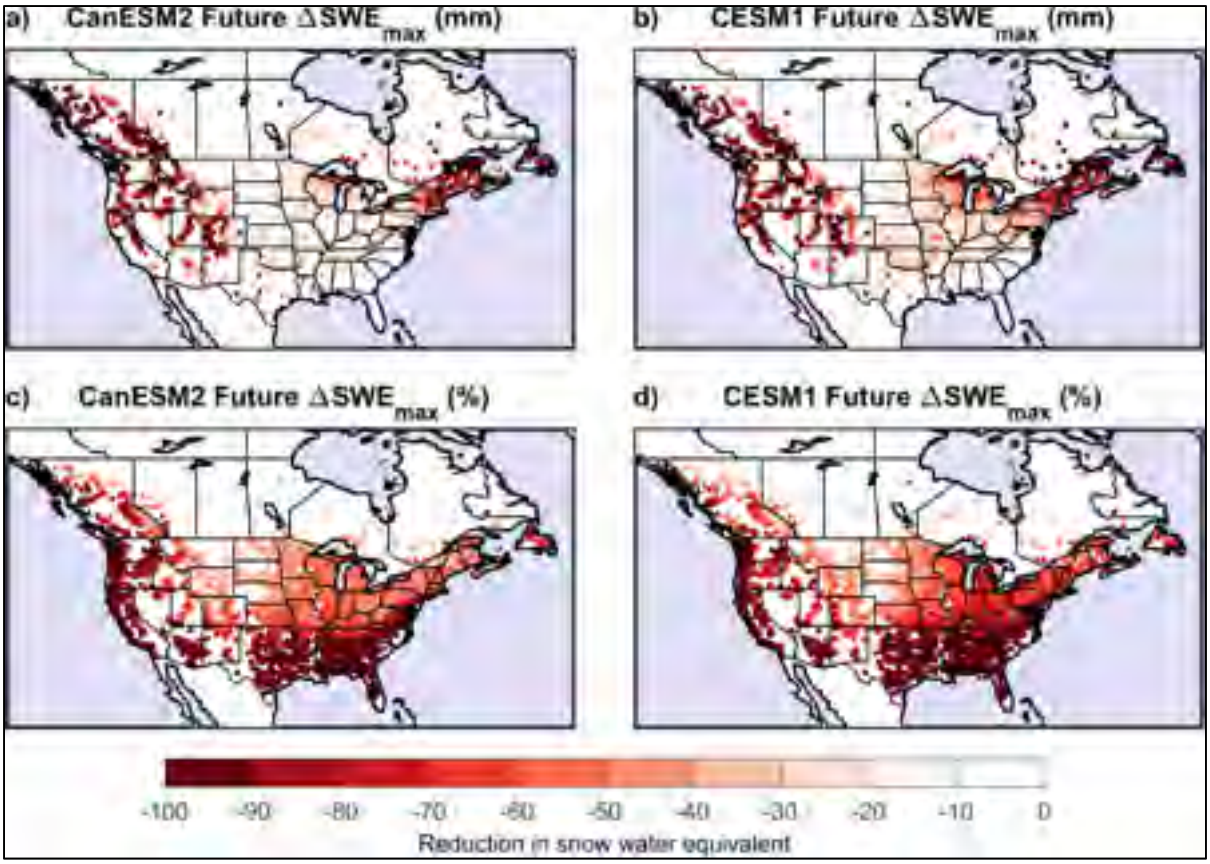

Figure 6.3 Same as Figure 6.1, but for the changes in the annual maximum snow water equivalent (SWE_{max}) in mm (top) and in % (bottom)

Results from Figure 6.3-a-b show that catchments with large reduction in the snowpack's annual maximum SWE (and consequently snowmelt volumes) will very likely experience decreases in the frequency and intensity of extreme flood events (Figure 6.1). These are generally high-latitude and/or high-altitude mountainous catchments mostly located within the Canadian Shield and the Rockies. These results corroborate conclusions from previous studies on observed data (Groisman et al., 2001; Li et al., 2017; Madsen et al., 2014; Seneviratne et al., 2012; Stewart et al., 2005; Zhang et al., 2001).

For catchments located in the southern mountainous regions of the West Coast (e.g. the Sierra Nevada and Cascades), a shift from a large to a quasi-inexistent snowpack is obtained

(Figure 6.3-c-d). These catchments experience important changes in the hydrological cycle, with flooding going from being generated by snowmelt to rainfall. These changes explain the large reduction in the return periods seen in Figure 6.1 for these specific catchments.

Some catchments located in both high-latitude and high-altitude (e.g. the states of Montana and Wyoming as well as in Western Canada) experience a smaller reduction in the snowpack (Figure 6.3-c-d). This feature, combined with an increase in ΔV (Figure 6.2), points toward in increases in the frequency and intensity of flood events.

The projected future change in the peak timing (i.e. change in the most probable day when the annual maximum streamflow value is obtained – Δdate Q_{peak}) is shown in Figure 6.4 for both climate ensembles. Figure 6.5 and Figure 6.6 show a comparison of these results against those from Figure 6.1 to show the main patterns of changes. Reference and future mean annual hydrographs of the main patterns of changes shown in Figure S6.12 can also be related with results previously shown, and in particular the annual maximum snowpack's SWE show in Figure 6.3.

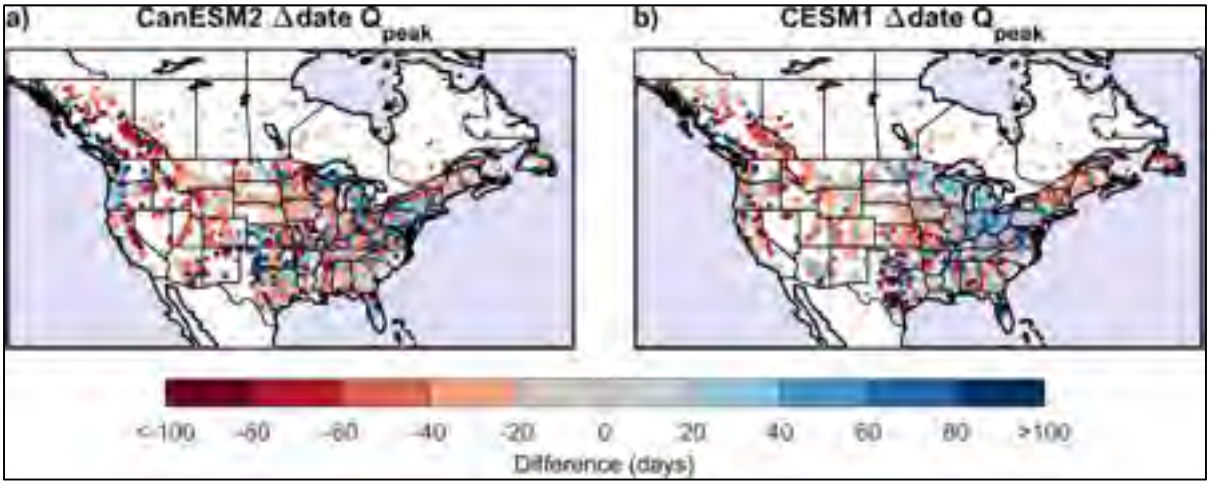

Figure 6.4 Same as Figure 6.1, but for the changes in the ordinal day of occurrence of daily annual maximum streamflow (Δdate Q_{peak}). A positive change means that future occurrence will occur in average later during the year

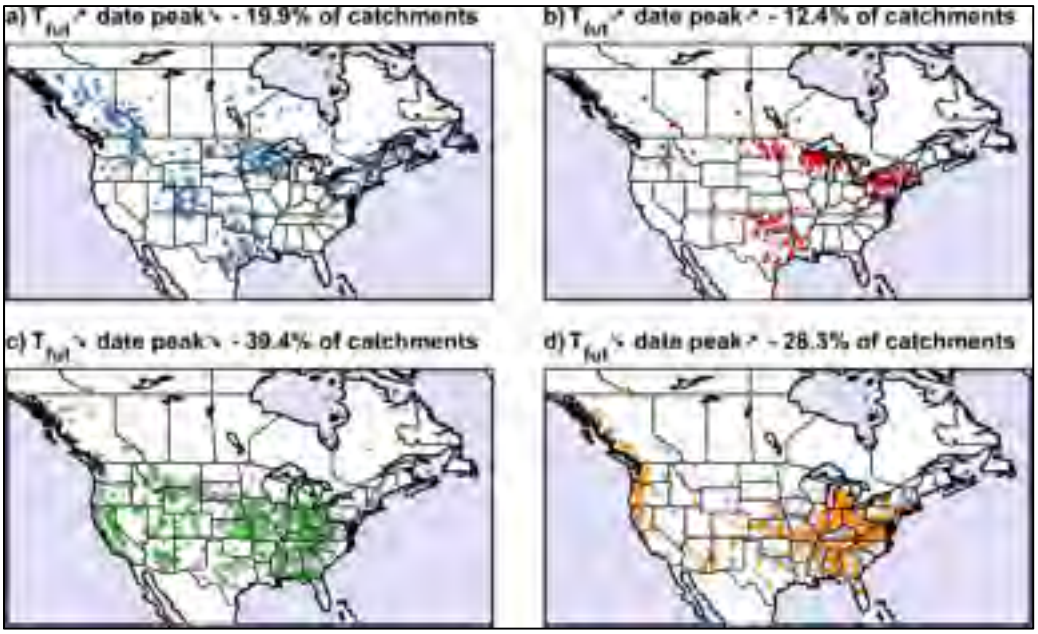

Figure 6.5 Direction of the projected future changes (2080-2099) in the 100-year flood event (T_{fut}) and in the most probable date for the annual maximum streamflow (date peak) for all 3 567 catchments using the CanESM2-LE. Gray dots correspond to those from Figure 6.1 and Figure 6.4

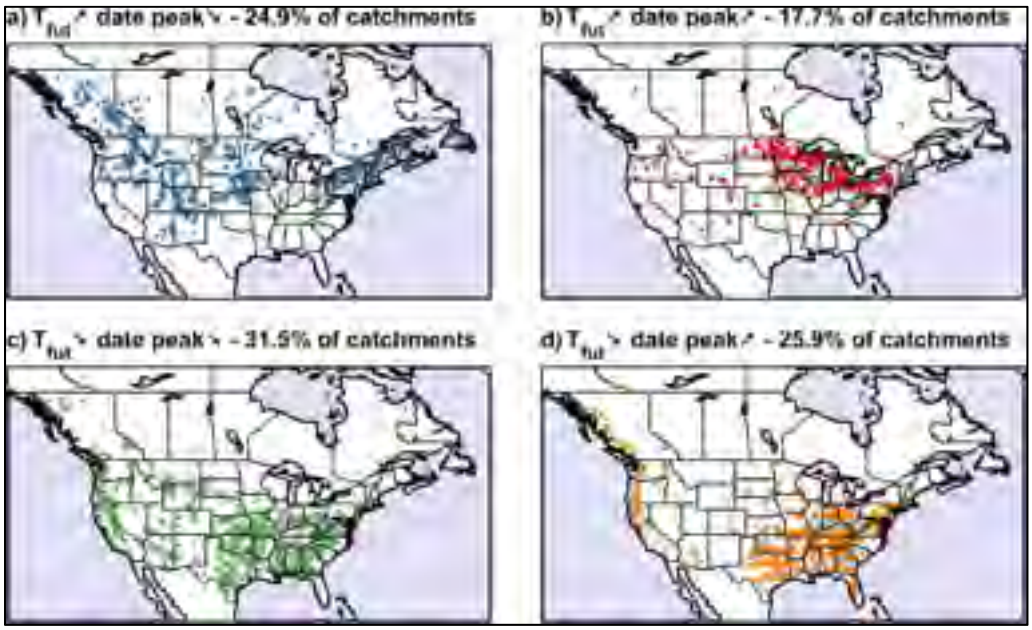

Figure 6.6 Same as Figure 6.5, but for CESM1-LE

Results displayed in Figure 6.4 suggest that earlier flood peaks are seen all over North America, except for Texas and catchments located in the Interior Lowlands regions (close to the Great Lakes). Smaller values are obtained over the Southeast region of the US and the West Coast mainly pointing towards a *status quo*.

High-latitude and/or high-altitude mountainous catchments generally have their flood peaks up to three months earlier, likely due to the increase in global temperature triggering a faster and longer snowmelt period (Pepin et al., 2015) (Figure 6.5-a and Figure 6.6-a). This tends to lead to both lower total runoff volume and intensity of the flood events. Some high-altitude catchments over the Canadian Rockies also switch from one to two snowmelt periods, further decreasing the intensity of flood extremes (Figure S6.12-d). However, as previously discussed, some catchments located in both high-latitude and high-altitude project increases in the intensity of flood events even if the snowmelt period is triggered faster (Figure 6.5-c and Figure 6.6-c).

As presented in Figure 6.5-b and Figure 6.6-b, catchments surrounding the Interior Lowlands region are showing later (around 2 months) and weaker flood events. This can be explained with the important decrease in the snowpack (Figure 6.3-c-d), switching from large flood events generated by snowmelt to smaller flooding generated by rainfall events. Regarding the Texas catchments for CanESM2, their behaviour can be explained due to the large increase in precipitation for the months of November and December, inducing a change in the timing of the annual maximum streamflow (from the beginning of the summer).

With respect to results shown in the bottom panels of Figure 6.5 and Figure 6.6, mixed patterns of increases and decreases in Δdate Q_{peak} combined with an increase in the intensity of flood events are shown over Southeast region of the US and the West Coast catchments.

6.4 The issue of spatial and temporal resolution

Most studies investigating extreme events using climate model simulations are generally hampered by a limited spatial and temporal resolution. To overcome this limitation, a minimal catchment size of 500 km^2 was imposed. For such catchments, the hydrological response time is proportionate with the daily temporal resolution of the climate models.

The large spatial resolution of climate models also brings some concerns. While a spatial resolution of the order of the kilometer is required to properly resolve convection, large weather systems have shown to be adequately simulated at the model resolutions used in this work (Kendon et al., 2017; Prein et al., 2015). When it comes to large catchments (> 500 km^2), mesoscale and synoptic scale weather systems are more likely to be involved in the generation of extreme flood events than intense convective storms. It can be expected that a higher resolution climate ensemble would have led to similar conclusions for the catchments analysed in this study.

The results presented in this work should not be extrapolated to smaller catchments. Small catchments have a sub-daily response time and even a sub-hourly response time for small urban catchments. For such catchments, extreme convective rainfall is the main driving force behind extreme floods, irrespective of the presence of snow or not. Both climate model ensembles used in this work do not allow for a top down study of floods on such catchments. Climate models with a finer temporal and spatial resolution would be needed to study their projected changes. However, outputs from global and regional climate models (Mailhot et al., 2012; Mailhot et al., 2007; Zhu, 2013) strongly suggest important increases in precipitation extremes, and especially for shorter durations and larger return periods. Accordingly, the flooding risk is expected to rise sharply on smaller catchments irrespective of the geographical location.

6.5 Conclusion and message to policy makers

While projected future in precipitation extremes are very likely to increase in response to global warming (IPCC, 2013; Kharin et al., 2007; Kharin et al., 2013), the picture is different when it comes to flood extremes. In this study, four main patterns of changes have been highlighted for North American catchments:

1) Important reduction in the annual maximum snowpack and earlier snowmelt period resulting in smaller volume of water and flood events. Most of catchments located over high-latitude and/or high-altitude mountainous catchments are impacted by these changes.
2) A shift in flood events that are currently generated by snowmelt towards flooding occurring due to strong rainfall. The affected catchments are mostly located within the Interior Lowlands around the Great Lakes.
3) More important and earlier snowmelt period generating larger volume of water and flood events. Some catchments located within both high-altitude and high-latitude mountainous catchments showed this behaviour (e.g. the states of Montana and Wyoming as well as in Western Canada).
4) Larger rainfall extremes with likely wetter antecedent soil conditions leading to an increase in the frequency and intensity of extreme flood events. Most of these catchments are located over the Southeast region of the US and the North American West Coast.

Overall, this study showed important climate change impacts on the frequency, intensity and seasonality of annual maximum streamflow. These findings have important implication when it comes to the design of resilient water infrastructures and flood adaptation measures directly affecting public safety and vulnerable populations. The results of this study are not applicable to small catchments where the flooding risk is expected to rise, and especially so for urban watersheds. Also, the impacts from changes ice jams are expected to play a significant role but were not considered in this work.

6.6 Methods

In the absence of proper routing of runoff outputs in climate models, a hydrological model was used to simulate streamflow at catchment outlets the reference (1990-1999) and future (2080-2099) periods. Daily streamflow were simulated using bias corrected time series of daily precipitation and temperature from two climate model large ensembles. The Daily Bias Correction (DBC) method used is a quantile mapping approach which combines a local intensity scaling (Schmidli et al., 2006) to correct for the probability of occurrence of precipitation, and a daily translation (Mpelasoka & Chiew, 2009) to correct the frequency distribution of both precipitation and temperature series (Chen et al., 2013).

The parsimonious GR4J lumped rainfall-runoff model (Perrin et al., 2003) coupled with the CemaNeige degree-day snow model (Valéry, 2010) was used to simulate daily streamflow from the bias corrected meteorological data. Potential evapotranspiration was computed using the Oudin formula (Oudin et al., 2005). Previous studies have used the same model structure which performed well in the simulation of continuous daily streamflow for a large number of the catchments analyzed in this study (Troin et al., 2015a; Troin et al., 2018; Velázquez et al., 2015b). Also a 9-parameter variant of the coupled model (5 for GR4J and 4 for CemaNeige) was adopted to improve the simulation over catchments with an important snow cover (Poissant et al., 2017).

A database of catchments' observations was elaborated over both Canada and the contiguous US to calibrate the hydrological model. Metadata (i.e. boundaries and drainage area) and daily streamflow were both extracted from Canadian Model Parameter Experiment (CANOPEX) database (Arsenault et al., 2016a) and from the United States Geological Survey (USGS). Only catchments with at least ten complete years of streamflow observation over the 1950-2010 period (defined as a year with less than 10% of missing data) were selected. Observed daily precipitation and temperature data were obtained through the Natural Resource Canada (NRCAN) (Hutchinson et al., 2009) and the University of Santa-

Clara (Maurer et al., 2002) gridded datasets for Canadian and American catchments respectively.

The hydrological model was calibrated over all catchments using the available daily observed streamflow observed data covering the 1950-2010 period. The Kling-Gupta Efficiency (*KGE*) (Gupta et al., 2009) multi-objective metric, giving weight to correlation, variability and bias errors, was used:

$$KGE = 1 - \sqrt{(r-1)^2+(\alpha-1)^2+(\beta-1)^2} \tag{6.1}$$

where r is the correlation coefficient; α is a measure of relative variability; and β is the bias between observed and simulated streamflow data. The calibration was performed using the Shuffled Complex Evolution-University of Arizona (SCE-UA) (Duan et al., 1992) optimization method with a total of 10,000 model evaluations (Arsenault et al., 2014). Finally, only catchments with a *KGE* value above 0.4 were kept.

Using the CanESM2-LE (Arora et al., 2011; Sigmond & Fyfe, 2016) and CESM1-LE (Kay et al., 2015) simulated daily streamflow time series, the corresponding annual maxima series were extracted for each catchment for both 1990-1999 reference and 2080-2099 future periods. The different members AMS are then pooled together, leading to a 1 000 (20 years × 50 members) and 800 (20 years × 40 members) datasets. Stationarity of the annual maxima series over the reference and future 20-year periods was validated using the Mann-Kendall (Kendall, 1975) test at a 95% confidence level.

These pooled datasets were then sorted in ascending order and the empirical probability of exceedance was based on the Cunnane plotting position (Cunnane, 1978; Meylan et al., 2008). Considering the 1 000 and 800-year annual maxima series provided by each ensemble, the empirical 100-year flood was used as a best estimate without resorting to using the Generalized Extreme Value theory. This was done for both the reference and future

periods. The large sample of annual maxima provided by the large-ensembles allows for a precise estimation of the 100-year flood.

6.7 Acknowledgements

This work is partly financed by the ClimEx project, which is funded by the Bavarian State Ministry for the Environment and Consumer Protection. The authors acknowledge Environment and Climate Change Canada's Canadian Centre for Climate Modeling and Analysis for executing and making available the CanESM2 Large Ensemble simulations used in this study, and the Canadian Sea Ice and Snow Evolution Network for proposing the simulations. The authors would also like to thank the Ouranos Consortium for helping with data transfer. The CanESM2 ensemble data is now available on the Environment and Climate Change Canada website (http://collaboration.beta.cmc.ec.gc.ca/cmc/cccma/CanSISE/output/CCCma/CanESM2/). The CESM1 ensemble was downloaded from the Large Ensemble Community Project (LENS) website (http://www.cesm.ucar.edu/projects/community-projects/LENS/).

6.8 Supplementary materials

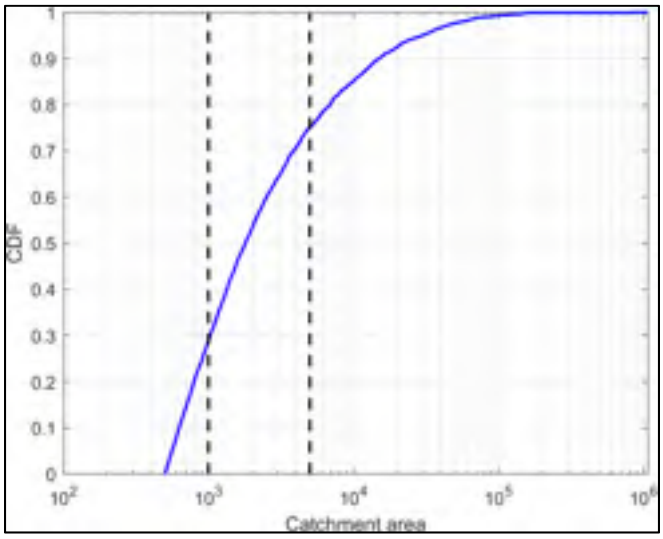

Figure S6.7 Cumulative distribution functions (CDF) of the 3 567 catchment areas

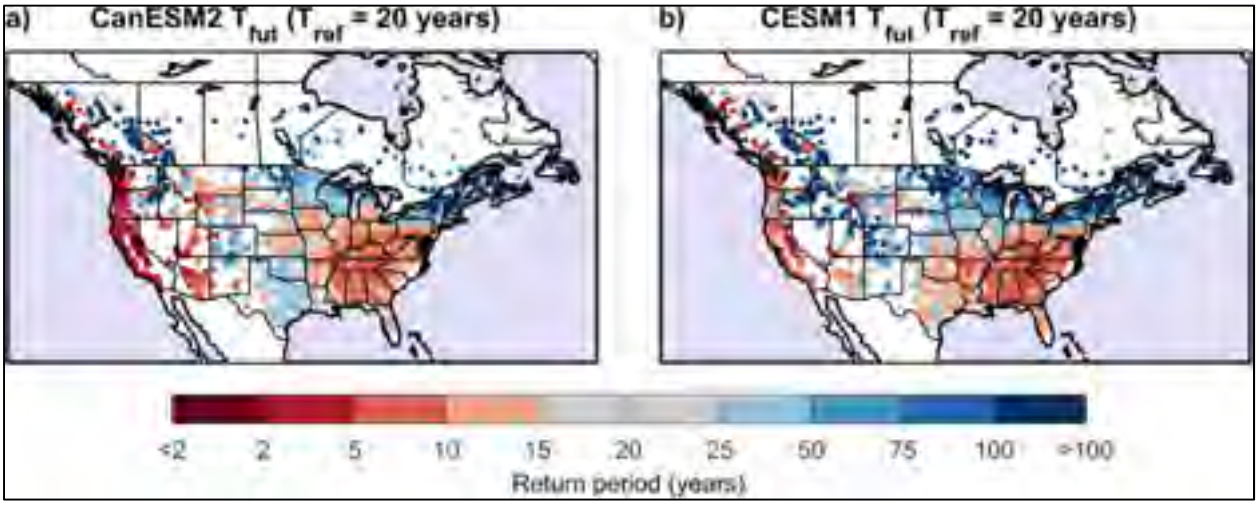

Figure S6.8 Projected future return periods (2080-2099) corresponding to the 20-year flood event over the reference period (1990-1999) estimated by the CanESM2-LE (a) and the CESM1-LE (b) for the 3 567 North American catchments

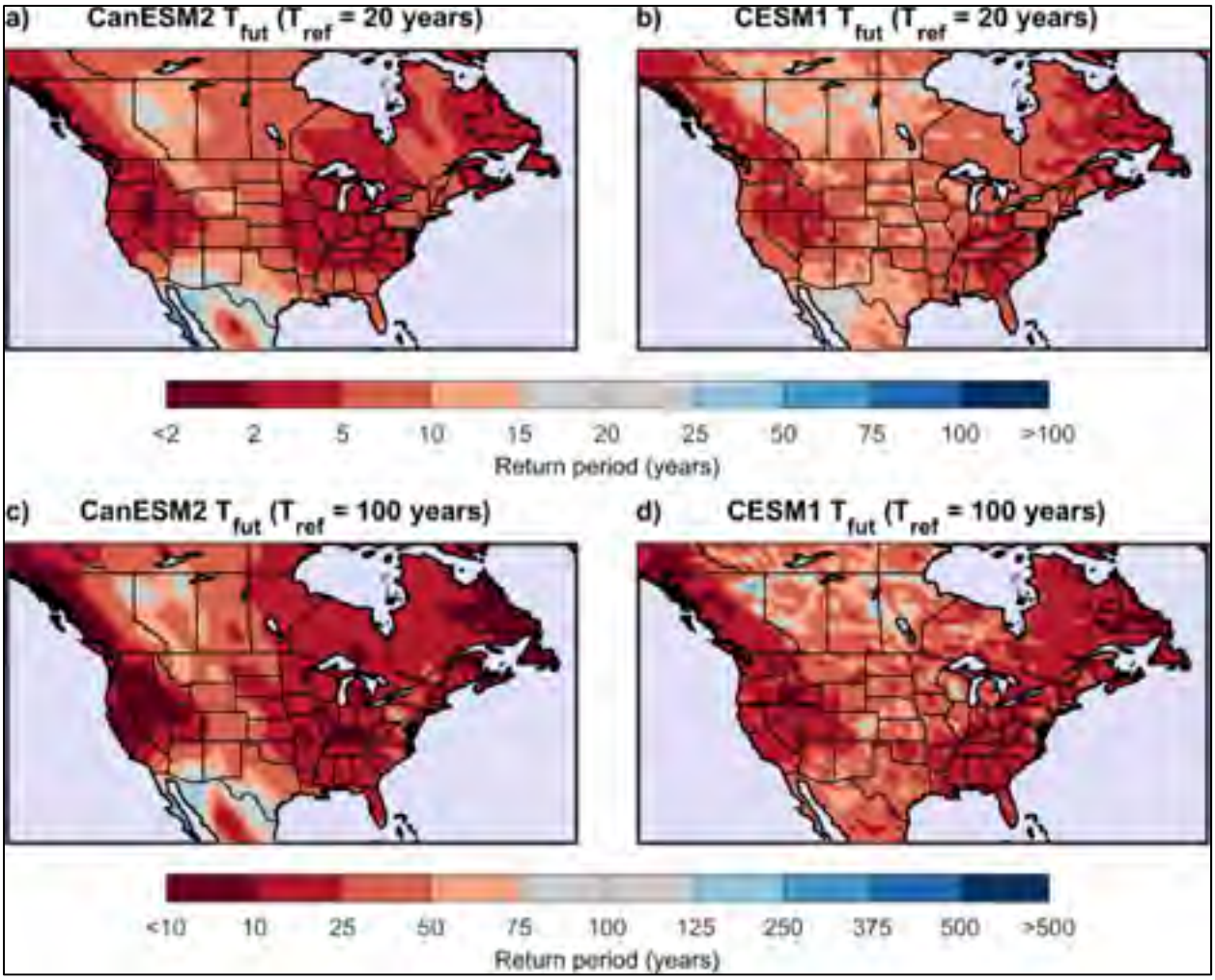

Figure S6.9 Same as Figure S6.8, but for the 20-year (top) and 100-year (bottom) daily precipitation event. Land raw data are presented instead of daily bias corrected data

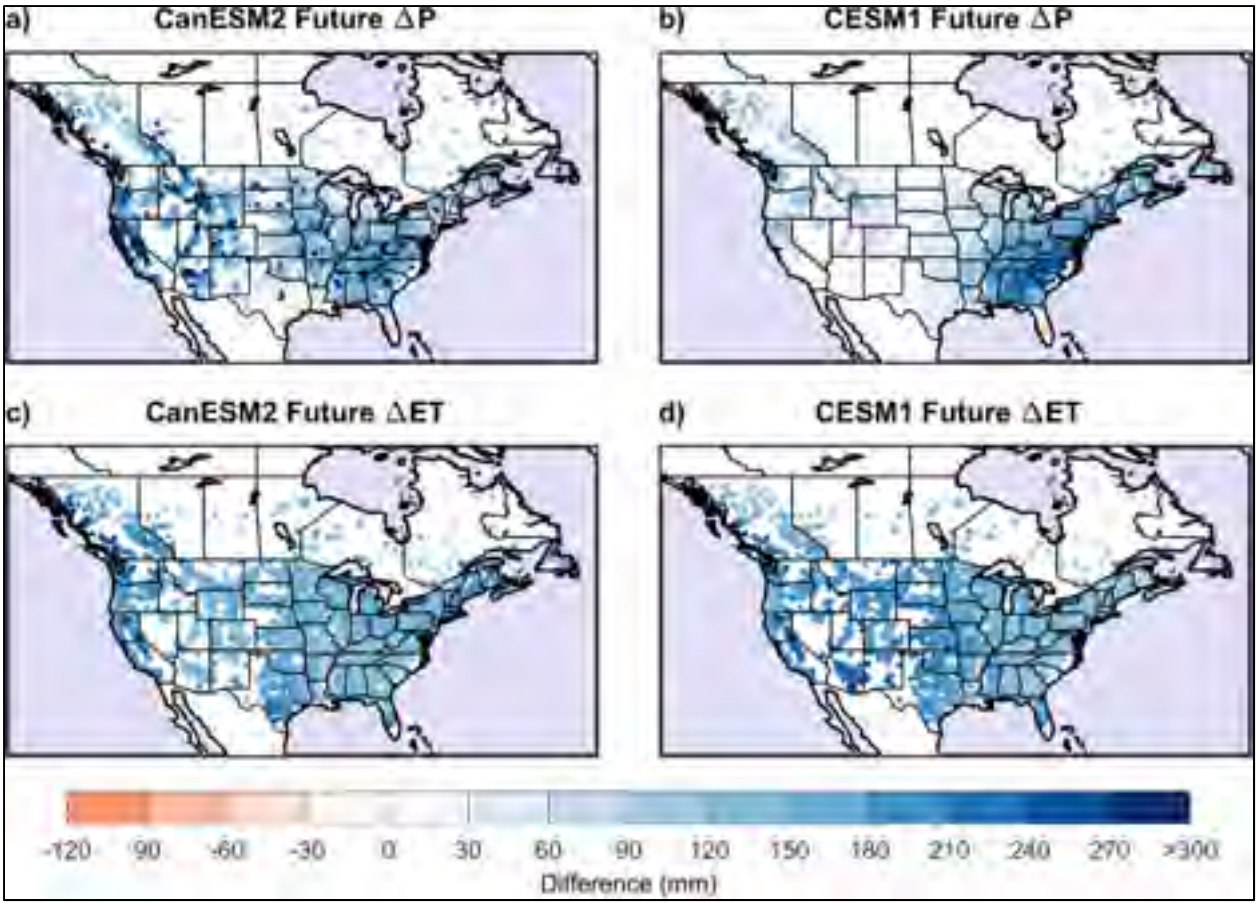

Figure S6.10 Same as Figure S6.8, but for the changes in the total annual precipitation (ΔP – top) and total annual evapotranspiration (ΔET – bottom)

Table S6.1 Contingency table between CanESM2-LE and CESM1-LE for the results shown in Figure 6.5 and Figure 6.6

		CESM1			
		Tfut↗ date peak↘	Tfut↗ date peak↗	Tfut↘ date peak↘	Tfut↘ date peak↗
CanESM2	Tfut↗ date peak↘	11.4%	3.8%	2.3%	1.8%
	Tfut↗ date peak↗	3.4%	5.8%	1.8%	1.9%
	Tfut↘ date peak↘	6.7%	4.3%	17.7%	11.7%
	Tfut↘ date peak↗	3.4%	3.8%	9.7%	10.5%

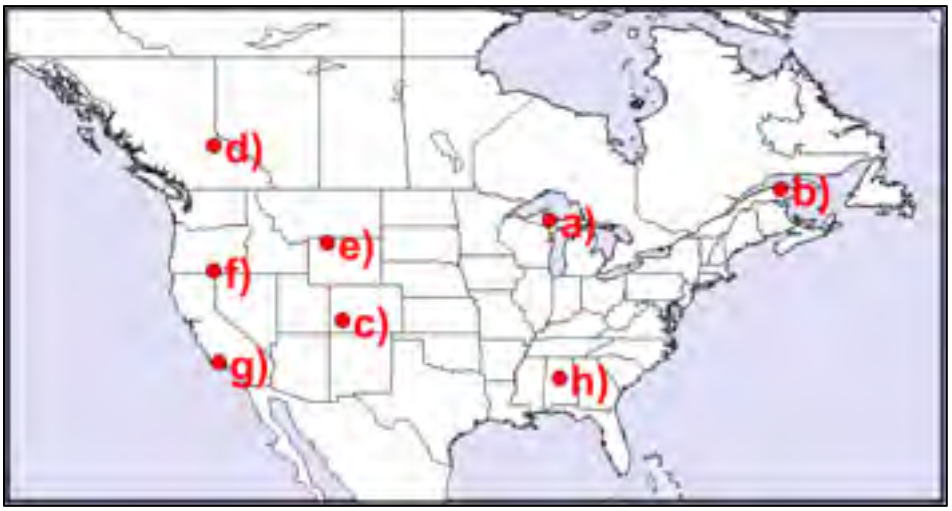

Figure S6.11 Map of the 8 selected catchments for the mean hydrograph analysis shown in Figure S6.12

Table S6.2 The eight selected catchments' in Figure S6.11 characteristics and mean annual hydrometeorological variables

	a)	b)	c)	d)	e)	f)	g)	h)
Catchment characteristics:								
Drainage area (km^2)	544	1200	686	904	7738	644	717	2659
Altitude centroid (m)	462	484	2744	1035	2016	1717	1930	147
Latitude centroid (degrees)	46.4	49.0	38.1	52.7	44.6	42.2	34.6	33.3
Longitude centroid (degrees)	-87.8	-65.6	-107.8	-120.1	-109.2	-120.1	-119.6	-86.9
Mean hydrometeorological variables:								
Annual total precipitation (mm)	838	1065	704	869	500	561	650	1417
Rainfall	598	589	385	380	290	293	642	1403
Snowfall	240	476	319	489	210	268	8	14
Annual daily temperature (°C)								
Minimum	-1.8	-2.8	-5.1	-5.2	-4.3	-1.9	4.1	10.1
Maximum	10.4	6.0	11.5	4.7	10.7	13.3	17.9	23.5
Mean	4.3	1.6	3.2	-0.3	3.2	5.7	11	16.8
Annual daily streamflow (m^3s^{-1})	5.4	28.4	5.6	45.2	32.2	4.1	2.6	43.9

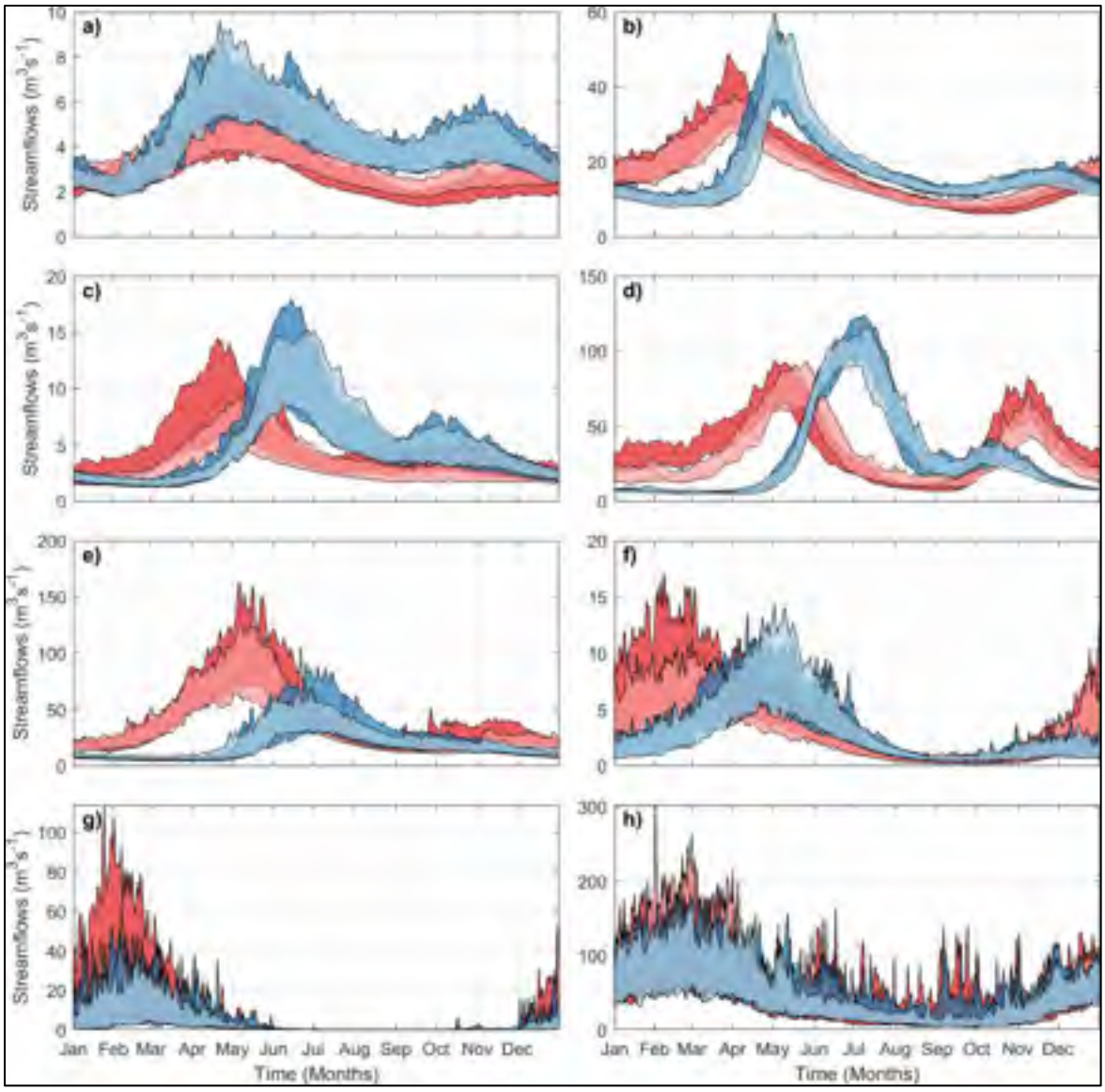

Figure S6.12 Envelop of the reference (blue) and future (red) mean annual hydrograph over the CanESM2-LE (dark colors) and CESM1-LE (light colors) for the selected 8 catchments presented in Figure S6.11

CHAPITRE 7

DISCUSSION GÉNÉRALE

Dans ce chapitre, les conclusions principales de l'ensemble des travaux de recherche réalisés sont discutées. Cette discussion permet de synthétiser les principaux résultats qui répondent à l'objectif principal de cette thèse, soit, l'amélioration de la compréhension des impacts de la variabilité naturelle et des changements climatiques sur les extrêmes hydrométéorologiques. De plus, cette discussion vise aussi à mettre en perspective l'importance de ces conclusions pour le Québec. Finalement, des pistes de stratégies d'adaptations ainsi que des recommandations pour la suite de ces travaux sont présentées.

7.1 Capacité de prédiction de la variabilité des extrêmes hydrométéorologiques

Les indices climatiques ont démontré un certain potentiel de prédiction au niveau de la variabilité saisonnière des températures et précipitations moyennes. Ces indices se trouvent à être moins prometteurs pour la variabilité saisonnière des extrêmes hydrométéorologiques. Les résultats du chapitre 3 présentent de faibles (voire non significatifs) pourcentages de variance expliquée à l'échelle des bassins versants de l'Amérique du Nord, et ce, même en combinant les six indices climatiques les plus pertinents.

Des analyses supplémentaires sur les précipitations et températures moyennes ont été réalisées afin de déterminer si les différents indices climatiques peuvent s'affecter mutuellement à travers leurs différentes phases (positives et négatives). Les résultats présentés à l'Annexe V ont examiné le couplage des phases de ces indices en comparant les cycles multidécennaux (AMO et PDO) avec les cycles décennaux (AO et NAO) et interannuels (ENSO et PNA). La conclusion principale qui est ressortie de ces analyses est que les indices climatiques sont probablement interdépendants et que l'interprétation des impacts individuels des indices pourrait être fortement biaisée par les phases des indices multidécennaux. Ceci suggère que la variabilité naturelle du climat est trop complexe pour être estimée en utilisant de simples indices climatiques. Ainsi, l'utilisation de ces indices

pour expliquer une partie de la variabilité naturelle des extrêmes hydrométéorologiques est probablement limitée et son interprétation devrait être faite avec précaution. Toutefois, des travaux comme ceux de (Renard & Lall, 2014), suggèrent qu'il y a un certain potentiel de prédiction, notamment pour les inondations dans cette étude, en utilisant les champs atmosphériques et océaniques plutôt que les simples indices climatiques.

7.2 Détection du signal des changements climatiques d'origine anthropique

Le deuxième article de cette thèse a employé les grands ensembles de simulations climatiques afin de dresser un portrait plus clair du rôle de la variabilité naturelle sur la détection du signal des changements climatiques à travers les séries chronologiques de précipitations moyennes et extrêmes. Des résultats similaires ont été obtenus via des analyses additionnelles et sont présentés dans l'article de l'Annexe IV, renforçant davantage les conclusions qui suivent.

En ce qui concerne le Québec, les résultats suggèrent qu'à l'échelle régionale, il est probable que la détection du signal des changements climatiques pour les séries des maxima annuels de précipitation soit retardée jusqu'au milieu du 21^{e} siècle à cause de la variabilité naturelle. Dans le cas de la détection à l'échelle locale, le rôle de la variabilité naturelle est encore plus important, retardant une détection potentielle jusqu'à la fin du siècle.

Les conclusions qui découlent de ces analyses indiquent que l'analyse d'une série d'observations issue d'une seule station météorologique (échelle locale) ou d'un ensemble de stations (échelle régionale) pourrait ne pas montrer de tendance significative (notamment à l'échelle locale). Ces conclusions apportent une perspective importante à deux niveaux :

1. Dans la mesure du possible, il est essentiel de réaliser des analyses régionales et non locales pour minimiser l'effet de la variabilité naturelle sur la détection du signal des changements climatiques ;

2. La décision de ne pas mettre en place des mesures d'adaptation contre les changements climatiques dans le cas d'un signal apparemment absent (possiblement à cause de la variabilité naturelle) pourrait avoir des répercussions importantes à moyen et long terme.

7.3 Projection de l'évolution des événements extrêmes hydrométéorologiques

7.3.1 Projection de l'évolution des événements extrêmes de précipitations

Les résultats obtenus dans le troisième article de cette thèse ont permis de clarifier l'impact des changements climatiques sur les événements de précipitations exceptionnels (p. ex. la période de retour 1 dans 100 ans) à l'échelle locale, régionale et même globale. Il est important de noter que les conclusions qui en ressortent corroborent le consensus de la littérature pointant vers des augmentations de la fréquence et de l'intensité des extrêmes de précipitation pour la majorité des régions de la planète, entre autres, le Québec.

La conclusion principale qui peut être tirée de ces résultats est que les changements projetés sont dépendants de la période de retour et de la durée des événements. Il est attendu qu'une période de retour plus élevée et/ou une durée plus courte entraîneront vraisemblablement des changements relatifs plus importants, pointant majoritairement vers des augmentations. Pour le Québec, des augmentations sont attendues pour les extrêmes de précipitations, qui pourraient avoir des répercussions importantes, surtout au niveau des zones urbanisées et des petits bassins versants avec une réponse hydrologique infra-journalière. Pour ce type de bassin, les pluies convectives d'été de courte durée entraînent généralement les débordements plus sévères. Toutefois, il mérite d'être rappelé que les modèles climatiques ont des limitations importantes, notamment au niveau de leur résolution spatiale, qui doivent être considérées lors de l'interprétation des changements projetés dans les événements de précipitation extrêmes à l'échelle infra-journalière.

À la lumière de ces conclusions, des événements de surverses en milieu urbain pourraient donc devenir beaucoup plus importants, rendant cruciale une bonne gestion des eaux urbaines. Ceci renforce l'importance de recourir à des pratiques de gestions optimales des eaux pluviales à la source afin de mieux gérer le ruissellement direct. De plus, le dimensionnement du drainage des réseaux mineur (conduites) et majeur (routes) devrait être fait en prenant en considération les impacts des changements climatiques à moyen et long terme. Pour des ouvrages ayant une durée de vie de plusieurs décennies, leur capacité utile sera constamment à la baisse au fur et à mesure que l'ouvrage vieillit et que les extrêmes de précipitations deviennent de plus en plus importants.

7.3.2 Projection de l'évolution des événements extrêmes de débits en rivière

Les changements projetés pour les extrêmes de débits en rivière ne sont pas aussi clairs que pour les précipitations. Les résultats présentés au chapitre 6 indiquent que pour les grands bassins versants (> 500 km^2 et caractérisé par une réponse hydrologique supérieure à une journée) de l'Amérique du Nord, ceux-ci peuvent se résumer via quatre patrons principaux :

1. Une réduction importante du couvert de neige maximal et une fonte plus hâtive, se traduisant en une diminution des volumes d'eau et des crues ;
2. Un déplacement des crues générées par la fonte de la neige vers des crues moins fortes se produisant par des pluies abondantes ;
3. Une augmentation de la fonte résultant en des volumes d'eau et des crues plus importants ;
4. Des précipitations plus abondantes conduisant vers une augmentation de l'intensité et de la fréquence des crues extrêmes.

En ce qui concerne la région du Québec, celle-ci semble être fortement affectée par les deux premiers patrons (le premier pour le sud et le deuxième pour le nord). En effet, la fonte de la neige combinée aux événements de pluies sur neige joue un rôle fondamental au niveau de l'intensité des crues printanières québécoises. L'impact des changements climatiques pour

les grands bassins versants du Québec pointe donc vers une diminution de la fréquence des événements extrêmes. Par exemple, ces résultats suggèrent que les inondations exceptionnelles ayant frappé le Québec au printemps 2017 ne semblent pas avoir une signature des changements climatiques. En effet, les résultats d'une étude réalisée par Teufel et al. (2018) indiquent que ces inondations sont le résultat d'une combinaison de précipitation nettement au-dessus de la moyenne sur un couvert de neige existant, suivi de deux fortes pluies consécutives. Malgré que ces fortes précipitations soient probablement plus fréquentes dans un climat futur, les effets de celles-ci sur la crue printanière seront contrebalancés par la réduction attendue dans le couvert de neige.

Il est important de noter que ces conclusions ne sont pas applicables pour des petits bassins versants ou des bassins fortement urbanisés. La ligne séparant les grands et les petits bassins, au-delà de laquelle des augmentations pourraient être observées demeure incertaine. Par conséquent, les plaines inondables devraient être gérées au cas par cas. De plus, le modèle hydrologique et les simulations climatiques sont des outils qui pourront apporter une aide considérable à ce niveau.

7.4 Incertitude statistique dans l'estimation des extrêmes hydrométéorologiques

Dans le contexte des travaux réalisés dans cette thèse, les grands ensembles de simulation climatiques ont permis d'analyser des changements projetés dans les événements hydrométéorologiques extrêmes pour des périodes de retour allant jusqu'à celle de 1 dans 100 ans. Cependant, en pratique, l'incertitude statistique des périodes de retour moins fréquentes est très importante et est trop souvent ignorée par les ingénieurs.

Cette problématique peut être illustrée en se basant sur une expérimentation de Schulz et Bernhardt (2016). En utilisant le grand ensemble de simulations du CRCM5, une série de 1 000 années est extraite du point de grille englobant la ville de Montréal pour la période de 1980 à 1999. À l'aide de cette longue série chronologique, des pluies de récurrence 1 dans

100 ans sont estimées avec une distribution de Gumbel en utilisant une fenêtre mobile de 30, 100 et 300 ans pour l'ensemble de la série. Ces valeurs sont ensuite comparées à la meilleure estimation de la récurrence 1 dans 100 ans basée sur les 1 000 valeurs de maxima annuels. L'hypothèse sous-jacente de cette analyse est que le climat est un système ergodique, permettant ainsi de combiner les séries de 20 années dans une très longue série de 1 000 années. Cette dernière est présentée à la Figure 7.1 permet de mieux visualiser l'impact de l'incertitude statistique sur l'estimation d'une telle période de retour, surtout si une série chronologique relativement courte est utilisée pour l'estimation.

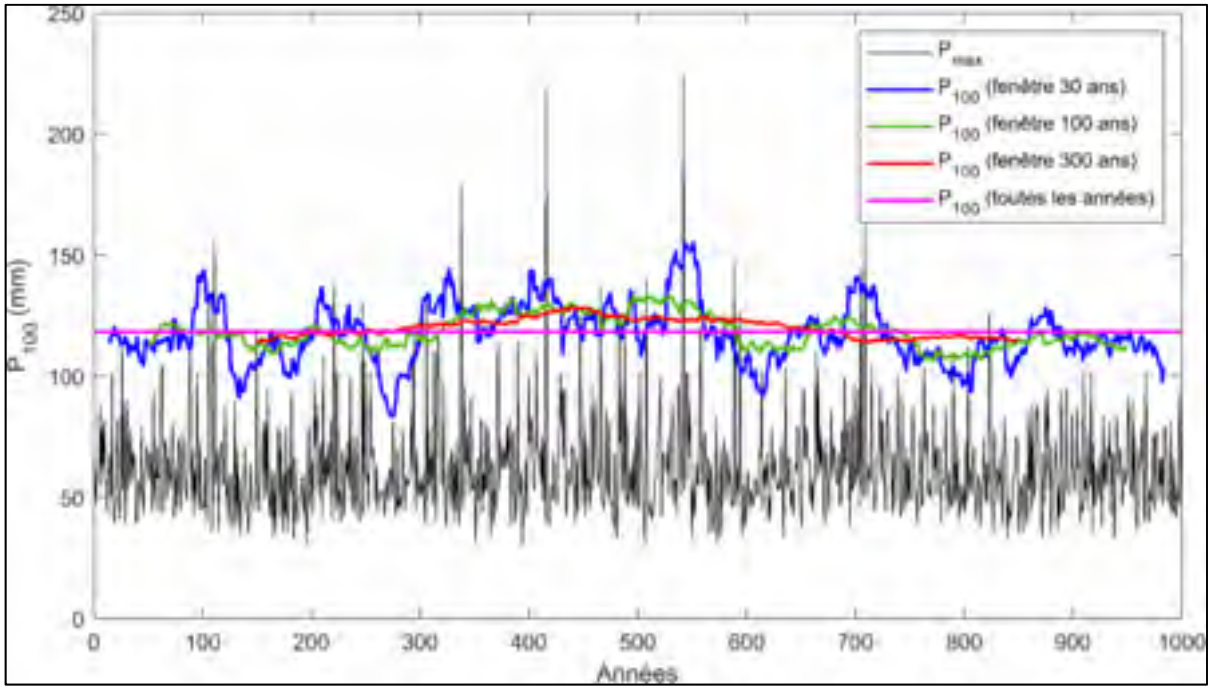

Figure 7.1 Série temporelle construite à partir des séries des maxima annuels de précipitation journalière tirée des 50 membres du CRCM5 pour la période de 1980 à 1999 centrée sur le point de grille de Montréal. Les valeurs de la P_{100} sont estimées à l'aide de la distribution Gumbel. La courbe bleue, verte et rouge représente les valeurs estimées de la P_{100} en utilisant une fenêtre mobile de 30, 100 et 300 ans respectivement. La ligne rose représente la valeur de la P_{100} en utilisant les 1 000 valeurs disponibles pour son estimation

7.5 Pistes de stratégies d'adaptation face aux changements climatiques

Depuis 2016, au Québec, un facteur de sécurité de 18% doit être ajouté par défaut pour les événements d'une récurrence supérieure à 2 ans, peu importe la période de retour et la durée (Ministère du Développement durable, 2017). À la lumière des conclusions de l'article du chapitre 5, cette approche pourrait s'avérer comme étant trop optimiste et non-adaptée pour faire face aux changements climatiques. Cependant, il est important de souligner qu'il s'agit d'un pas dans la bonne direction qui doit être encouragé et servir d'exemple pour les autres provinces et pays n'ayant pas encore mis en place de mesures d'adaptation. Par exemple, en reprenant les changements relatifs dans les courbes IDF du point de grille correspondant à la ville de Montréal présentés dans le chapitre 5, ces derniers pourraient être beaucoup plus élevés que le 18 % suggérés.

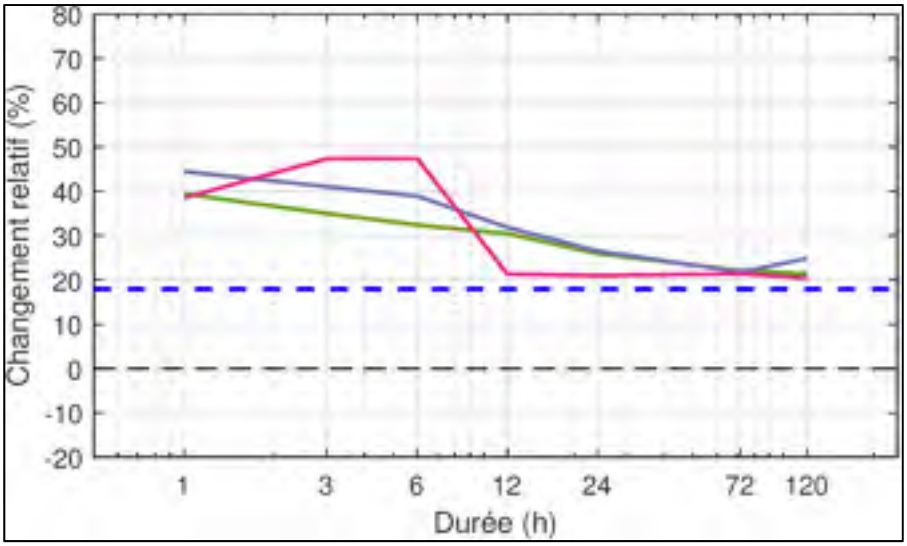

Figure 7.2 Changements relatifs dans les courbes intensité-durée-fréquence (IDF) entre les périodes de 1980-1999 et 2080-2099 pour le point de grille correspondant à la ville de Montréal estimés à partir du CRCM5-LE. Les périodes de retour 2 ans (vert), 20 ans (mauve) et 100 ans (rose) sont présentées. La ligne pointillée bleue présente le facteur de sécurité de 18 % qui doit être ajouté par défaut au Québec

7.5.1 Revue des stratégies actuellement employées

Il est intéressant de présenter certaines stratégies d'adaptations qui ont été mises en place par d'autres pays et qui pourraient servir d'alternative temporaire le temps que des guides et des normes plus peaufinés soient développés. Les travaux de Madsen et al. (2014) ont notamment recensé certaines stratégies d'adaptations de divers pays, notamment :

- Belgique : 30% par défaut pour tous les événements ;
- Danemark : 20%, 30% et 40% d'augmentations pour les périodes de retour de 2-, 10- et 100 ans respectivement ;
- Royaume-Uni : 10%, 20%, 30% d'augmentations pour les périodes futures de 2055, 2085 et 2115 respectivement ;
- Suède : 5% à 30% par défaut pour tous les événements dépendant de la région.

7.5.2 Utilisation d'une variante non stationnaire d'une loi statistique

L'utilisation d'une variante non-stationnaire d'une loi statistique, telle que la loi des extrêmes généralisés décrite à la section 1.3.2 peut s'avérer une alternative intéressante qui mérite d'être mentionnée. En variant un ou plusieurs des paramètres de la distribution en fonction d'une autre variable (p. ex. : le temps (en années) ou même de l'augmentation régionale ou même globale de la température moyenne), l'impact des changements climatiques serait considéré lors de l'extrapolation vers événements plus rares. Cependant, l'absence de détection du signal des changements climatiques limiterait l'efficacité d'une telle méthode. En effet, en l'absence d'une tendance significative locale ou régionale, un critère comme celui du BIC ou du *Likelihood-ratio test* décrit à la section 1.3.2, ne recommanderait probablement pas l'utilisation du ou des paramètres non stationnaires additionnels testés.

7.5.3 Tenir compte de l'incertitude statistique

Une autre suggestion serait de tenir compte au minimum de l'incertitude statistique de l'estimation de la période de retour désirée. Pour des périodes de retour plus élevées (p. ex. une récurrence 1 dans 20 ans ou 1 dans 100 ans), leur extrapolation à partir d'une série chronologique relativement courte entraine une importante incertitude statistique. La Figure 7.3 présente un exemple basé sur l'analyse de la Figure 7.1, démontrant de l'ampleur de l'incertitude statistique (intervalle de confiance 90% de la distribution Gumbel) de la période de retour 100 ans. En tenant compte de cette incertitude statistique (c.-à-d. en prenant la borne supérieure de l'intervalle de confiance), un critère de conception plus sévère peut être obtenu, aboutissant à des infrastructures plus résilientes.

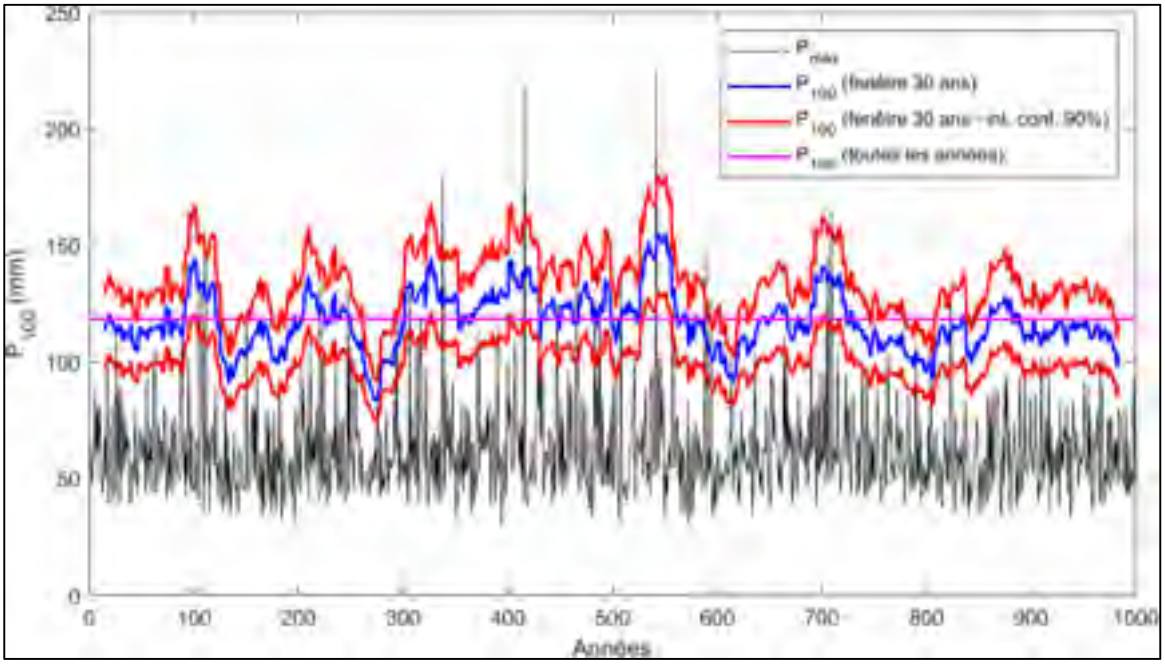

Figure 7.3 Comme la Figure 7.1, mais avec l'intervalle de confiance 90% (courbes rouges) de la fenêtre mobile de 30 ans (courbe bleue)

Par ailleurs, l'incertitude statistique est généralement fournie, comme avec les courbes IDF d'Environnement Canada ou d'Agrométéo. Par exemple, pour la station de l'aéroport

international Pierre-Elliot-Trudeau de Montréal, l'estimation d'une pluie 100 ans d'une durée d'une heure est de 53 mm avec un intervalle de confiance à 95% entre 46.6 mm (-13.7%) et 58.5 mm (9.4%).

L'annexe VII présente les résultats d'une communication orale réalisée en 2016 dans le cadre de cette thèse. La stratégie décrite dans ce paragraphe y avait été employée pour trois bassins versants québécois en utilisant la loi statistique de Gumbel avec un intervalle de confiance de 90% sur les séries des maximas annuels journaliers des débits en rivière. Le grand ensemble de 50 simulations du modèle CanESM2 avait été utilisé afin de comparer l'estimation de la période de retour 20 ans de chaque simulation entre elles. Les résultats suggèrent que l'intervalle de confiance à 90% englobait la quasi-totalité des 50 estimations obtenues en climat de référence, et également une grande portion de l'incertitude liée aux changements climatiques.

Ainsi, dans un contexte de changement climatique, tenir compte de l'incertitude statistique pourrait s'avérer être un pas vers des infrastructures plus résilientes, grâce à l'utilisation de critères de conception plus sévères que ceux actuellement utilisés.

7.5.4 Ajouter des facteurs de sécurité adaptés aux différentes conditions

Similairement à la majorité des stratégies actuellement employées, des facteurs de sécurité adaptés aux différentes conditions pourraient se trouver à être une solution relativement simple à mettre en place et efficace. Cependant, tel que discuté précédemment, il essentiel de considérer des facteurs variant en fonction de la durée et de la période de retour.

En se basant sur la méthodologie présentée dans les chapitres 5 et 6, les modèles climatiques, notamment les grands ensembles de simulations climatiques, pourraient être utilisés pour générer ces facteurs de sécurité. Par exemple, un tableau présentant les changements relatifs attendus dans les courbes IDF pour différentes villes, ou encore des cartes présentant les

changements en fonction d'une durée et d'une période de retour sélectionné pourraient être une méthode intéressante pour communiquer ces valeurs. Un exemple d'un tel tableau est présenté pour les villes de, Gaspé, Montréal et Saguenay à partir de l'ensemble du CRCM5 (Tableau 7.1). Tel que discuté précédemment, les changements attendus pour ces trois villes pointent vers des augmentations plus importantes pour les courtes durées et les périodes de retour moins fréquentes.

Tableau 7.1 Différences relatives (%) entre les valeurs d'intensité pour les périodes de 1980-1999 et 2080-2099 pour différentes durées et périodes de retour. Le point de grille du CRCM5 est centré sur les villes de Gaspé, Montréal et Saguenay respectivement

	Gaspé			**Montréal**			**Saguenay**		
Duration (h)	Période de retour (années)								
	2	20	100	2	20	100	2	20	100
1	63.8	75.2	81.3	39.4	44.5	38.6	53.6	52.1	63.7
3	57.0	71.0	67.1	35.0	41.1	47.4	47.7	58.0	61.3
6	44.5	56.1	64.6	32.4	38.9	47.4	42.5	54.0	61.9
12	36.3	42.5	47.3	30.4	31.9	21.3	38.6	47.2	55.6
24	33.1	31.3	34.6	26.0	26.5	21.0	31.9	36.6	43.1
72	27.0	32.0	23.1	22.1	21.8	21.4	26.7	29.8	34.6
120	27.4	28.0	28.9	21.4	24.8	20.4	22.2	28.9	39.9

Considérant les limitations des grands ensembles de simulations climatiques (p. ex. la résolution spatiale des modèles), il est évident que les valeurs obtenues auront une incertitude non négligeable. Cependant, celles-ci sont vraisemblablement les meilleures estimations qu'on puisse actuellement obtenir, et pourraient servir de base pour l'établissement de facteurs de sécurité. Toutefois, les différentes sources d'incertitudes devraient être quantifiées dans la mesure du possible, et les facteurs devraient être actualisés en fonction des avancées en modélisation et des connaissances en la matière.

7.6 Recommandations et travaux futurs

Il y a certaines questions et limitations qui ont été soulevées tout au long de cette thèse qui méritent d'être adressées dans de futurs travaux. Ces travaux supplémentaires permettraient d'approfondir davantage les connaissances acquises dans cette thèse, et de les étendre vers de meilleures stratégies d'adaptation face aux changements climatiques. Les points suivants présentent les recommandations principales en termes de travaux futurs :

- Évaluer l'incertitude des différentes composantes de la chaine de modélisation. Les trois principales sources d'incertitude peuvent se résumer à :
    - Les grands ensembles de simulations climatiques ;
    - Les scénarios d'émissions de gaz à effet de serre ;
    - Les modèles hydrologiques (pour la simulation des débits en rivière).

  Quantifier l'incertitude de ces diverses composantes pourrait potentiellement renforcer les conclusions principales de cette thèse.
- Inclure des bassins versants de petite taille (< 500 km^2) parmi l'échantillon analysé dans cette thèse. Avec une réponse hydrologique généralement infra-journalière, l'impact des changements climatiques sur leurs crues extrêmes serait vraisemblablement différent. En général, pour les bassins de petite taille ou les zones urbaines, les inondations se produisent principalement suite à des événements de précipitation intense d'une durée relativement courte. Ainsi, pour une région comme le Québec, où la crue maximale des grands bassins versants est fortement influencée par la fonte de la neige, les changements projetés pour les petits bassins se trouveraient probablement à l'opposé du spectre.
- Reproduire les analyses en utilisant un grand ensemble de simulations climatiques provenant d'un « *Convective Permitting Model* » lorsque ceux-ci seront enfin disponibles. La résolution spatiale plus élevée de ce type de modèle permet de résoudre la convection, ce qui entraînera vraisemblablement une meilleure représentation des événements extrêmes de précipitation plus locaux et de courte durée. Ces résultats fourniraient potentiellement un portrait plus clair de l'impact des

changements climatiques sur les précipitations à l'échelle horaire et sous-horaire, jouant aussi un rôle important sur les ruissellements extrêmes des bassins versants de petite taille et en milieux urbains.

- Développer des meilleures techniques de correction de biais pour les événements extrêmes. Par exemple, les changements projetés pour les précipitations extrêmes ont été formulés en termes de changements relatifs en utilisant les sorties brutes des modèles climatiques. Toutefois, une méthodologie qui permettrait de corriger de façon plus robuste les biais des extrêmes tirés des modèles par rapport aux observations, et d'en analyser par la suite les changements projetés, pourrait donner une plus grande confiance dans les conclusions obtenues.
- Conduire plus spécifiquement des travaux sur les bassins anthropisés (régulés) afin de déterminer l'impact des différentes méthodes de gestion des ressources hydriques sur les extrêmes hydrologiques futurs.
- Développer des méthodes permettant d'aider avec l'interprétation des différents biais dans les sorties des modèles climatiques. Par exemple, si un modèle présente un signal à la hausse de 20% sur les précipitations centennales, mais a un biais négatif de 20% en climat présent, est-ce que les résultats obtenus avec ce modèle devraient-être conservés ou écartés?
- Finalement, tester en profondeur les différentes pistes de solutions discutées précédemment afin de fournir aux ingénieurs et aux décideurs de meilleurs outils et stratégies d'adaptation potentielles contre les changements climatiques.

CHAPITRE 8

CONCLUSION

Traditionnellement, la conception de la majorité des infrastructures est réalisée en fonction de l'hypothèse que le climat est stationnaire. Considérant le vaste consensus (97% des études scientifiques) indiquant que le réchauffement planétaire est causé par l'influence humaine, il est évident que cette hypothèse n'est plus valide et que nos critères de conception doivent rapidement être adaptés. En effet, une des conclusions principales du cinquième rapport l'IPCC est qu'il y aura une augmentation de l'intensité et de la fréquence des précipitations extrêmes, notamment pour une région comme le Québec, en raison du réchauffement planétaire. Les résultats de cette thèse ont appuyé cette conclusion en dressant un portrait plus clair des augmentations attendues.

Une première partie des travaux de cette thèse présentée dans le chapitre 3 ont montré qu'il est très difficile de prédire la variabilité naturelle des extrêmes hydrométéorologiques à partir des indices climatiques connus comme celui d'El Niño. En plus de la difficulté à prédire cette variabilité naturelle, les résultats montrés au chapitre 4 indiquent qu'il est probable, qu'en raison de cette dernière, qu'il ne soit pas possible de détecter l'influence des changements climatiques sur les séries chronologiques d'observations de précipitation d'une, voire même plusieurs stations météorologiques. Ainsi, dans le cas d'un signal apparemment absent, et ce possiblement à cause de cette variabilité naturelle, il demeure impératif de mettre de l'avant les stratégies d'adaptation contre les changements climatiques.

Il demeure difficile pour l'ingénieur de déterminer comment adapter leur conception face aux changements climatiques. C'est dans cette optique que les résultats et les conclusions de cette thèse prennent tout leur sens, servant de fondation vers des travaux futurs qui permettront d'établir davantage de guides et de normes pour les ingénieurs. Les résultats présentés au chapitre 5 ont présenté l'ordre de grandeur des changements attendus principalement au niveau de la période de retour 1 dans 100 ans des extrêmes de précipitations pour différentes durées. Les courbes intensité-durée-fréquence (IDF), des outils constamment utilisés dans la

conception des infrastructures, ont été investiguées du même coup. Les résultats suggèrent qu'il y aura une importante augmentation dans la fréquence des événements extrêmes de précipitations de toutes les durées sur l'ensemble du Québec. Par exemple, l'intensité d'un orage d'une durée de 24 heures qu'on voyait historiquement (1980-1999) en moyenne une fois tous les 100 ans, pourrait devenir de quatre à six fois plus fréquent vers la fin du 21^{e} siècle (2080-2099). Étant donné que de nombreuses infrastructures publiques ont une durée de vie supérieure à 75 ans, ces augmentations ont des implications importantes et devraient être considérées par les décideurs et les ingénieurs.

Concernant les débits en rivière, la réponse face aux changements climatiques est très différente de celle des précipitations, telle que présentée au chapitre 6. Des réductions au niveau des événements extrêmes de faible récurrence sont attendues sur les bassins versants de grande taille dominés par la fonte de la neige, tels que ceux du Québec, en lien avec une fonte plus rapide du couvert neigeux. Toutefois, les petits bassins versants et les zones urbaines ayant une réponse hydrologique infra-journalière (ou même infra-horaire) ne réagiront vraisemblablement pas de la même façon et afficheraient vraisemblablement des augmentations importantes comme pour les précipitations. D'un autre côté, une grande partie des régions nord-américaines pour lesquelles la neige ne joue qu'un rôle secondaire (ou pour les bassins dont de temps de réaction est très rapide), des augmentations similaires à celle des précipitations extrêmes sont projetées.

Ainsi, il devient évident que les changements climatiques auront aussi des impacts directs sur la sécurité publique (p. ex. : blessures, pertes de vies, relocalisation des sinistrés, etc.) et sur plusieurs enjeux d'intérêt public (p. ex. : endommagement des infrastructures, impacts sur l'environnement, coûts élevés de réparation, etc.). La société devient également de plus en plus sensibilisée face à cette réalité, surtout à la suite des événements exceptionnels auxquels nous sommes de plus en plus exposés.

Afin de limiter les impacts qu'auront les changements climatiques via l'augmentation des extrêmes hydrométéorologiques, il est impératif d'agir le plus rapidement possible. Les ingénieurs seront amenés à jouer un rôle de premier plan à travers les choix de conception de nos infrastructures. Ainsi, les résultats et les conclusions de cette thèse, et des travaux qui en découleront pourront servir de références à ce niveau. Une liste d'idées de stratégies potentielles d'adaptation est aussi présentée et discutée brièvement au chapitre 7 de la thèse. Ces dernières pourraient servir de point de départ et d'outils supplémentaires aux ingénieurs permettant de justifier la nécessiter d'adapter leurs conceptions face aux changements climatiques.

Somme toute, nous avons un choix à faire en tant que société. Nous pouvons attendre et accepter la réalité que les événements exceptionnels du passé deviennent monnaie courante et de payer pour les dégâts qui en résulteront; ou nous pouvons plutôt mettre en place des stratégies d'adaptation dès maintenant afin de protéger l'ensemble de la population, et surtout, les plus vulnérables.

La balle est dans notre camp.

ANNEXE I

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)

Martin Leduc[a], Alain Mailhot[b], Anne Frigon[c], Jean-Luc Martel[d], Ralf Ludwig[e], Gilbert B. Brietzke[f], Michel Giguère[g], François Brissette[h], Richard Turcotte[i], Marco Braun[j], John Scinocca[k]

[a, c, g, j] Ouranos, Montréal (Québec), Canada
[b] INRS – Eau, Terre et Environnement, Québec (Québec), Canada
[d, f] École de technologie supérieure, Montréal (Québec), Canada
[e] Ludwig Maximilians University of Munich (Munich), Germany
[f] Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities, Garching, Germany
[i] Ministère du Développement durable, Environnement et Lutte contre les changements climatiques (MDDELCC), Québec (Québec), Canada
[k] Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria (British-Columbia), Canada

Article accepté à la revue « Journal of Applied Meteorology and Climatology », janvier 2019. La version présentée ici est la version acceptée avec révisions majeures.

Abstract:

In recent years, large initial-conditions ensembles using a single Global Climate Model (GCM) have become increasingly popular for studying the impacts of anthropogenic greenhouse-gas emissions on natural climate variability and extreme events. Due to high computational costs, however, these ensembles are rarely produced and generally feature coarse grid resolutions that render them inadequate for direct use in most regional studies of climate-change impacts. Large ensembles featuring a high-resolution Regional Climate Model (RCM) instead of a GCM are even less common because a GCM large ensemble needs to be purposely archived to serve as boundary conditions to an RCM. In the scope of the ClimEx project that investigates the implications of extreme hydrometeorological events on water management in Bavaria and Québec, a 12-km resolution RCM large ensemble was produced to drive hydrological models. This unprecedented dataset — the Canadian Regional

Climate Model (CRCM5) Large Ensemble (CRCM5-LE) — consists of a dynamically downscaled version of the coarse resolution Canadian Earth System Model 50-member initial-conditions ensemble (CanESM2-LE). Downscaling was performed over two domains, Europe and northeastern North America, and covers the period from 1950 to 2100 following the RCP8.5 greenhouse-emission scenario from 2006 onward. This paper presents a regional analysis of the CRCM5-LE dataset, including validation with observations, as well as its climate-change projections, inter-annual variability and extreme events.

I.1 Introduction

The ClimEx (Climate change and hydrological Extremes) project is the latest phase of a Bavaria and Québec's long-term collaboration to support the development of robust climate-change adaptation strategies based on assessments of the hydrological impacts of climate change, natural variability and extreme events. In practice, climate impacts are relevant at the local scale, where communities or natural ecosystems are exposed. However, the physical processes that drive important impacts involve a wide range of space and time scales. For instance, global mean sea-level rise has important impacts on coastal regions and small islands (Nicholls & Cazenave, 2010), but changes in storm intensity and frequency can also contribute to local flood risk (Rizzi et al., 2017). It is also generally accepted that natural climate variability already plays an important role in the occurrence of extreme events, but since climate change modulates several interlinked processes, assessing climate risk will therefore become increasingly complicated in the future (e.g. Milly et al., 2008b).

In hydrology, the key physical processes related to flooding may differ due to basin size (Musy, Hingray, & Picouet, 2014). For large basins, flooding generally results from intense precipitation over several days; in higher latitudes, melting snow also contributes to flooding. For smaller basins with quick reaction times, flooding often occurs due to short, intense precipitation events. Coarse spatio-temporal resolution climatological data may be used to drive hydrological models for large basins, but are inadequate for short-duration, high-

intensity precipitation that drives summer and fall floods. In such cases, spatio-temporal data must be at a much finer resolution to accurately represent rainfall-discharge processes (Ogden & Julien, 1993; Singh, 1997). In general, local hydrological impacts of climate change are studied using a variety of impact models, which use state-of-the-art climate model simulations for inputs. For instance, Global Climate Models (GCMs) (Earth System Models in their current generation) are commonly used to generate large scale climate-change projections over periods from decades to centuries (Collins et al., 2013). However, since GCMs are computationally expensive to run due to their high complexity, they typically use rather coarse spatial resolutions — ranging from 100 to 450 km in the Coupled Model Intercomparison Project Phase 5 (CMIP5) ensemble. These resolutions are often too coarse for hydrological applications (Fatichi et al., 2014; Fowler et al., 2007; Wigley, Jones, Briffa, & Smith, 1990). To fill the gap between GCMs and local scales, downscaling methods have been developed to refine GCM output before driving the hydrological model over a region of interest (Fowler et al., 2007; Xu, 1999).

Regional Climate Models (RCMs) offer a convenient approach to downscale GCM output at sufficiently high resolutions for impact modeling. RCMs represent an intermediate step that enables the concentration of computational power on a limited area (rather than on the entire globe as with a GCM) to obtain downscaled climate projections at spatial resolutions typically ranging from 12 to 50 km (Giorgi & Gutowski, 2015). RCMs are essentially built as GCMs in terms of dynamical core and parameterizations of sub-grid processes, but must be driven by either GCMs or reanalyses through their lateral and surface boundaries. With their higher resolution, RCMs provide a much better representation of the heterogeneity in surface forcings (e.g., land-sea contrasts, orography, distribution of lakes and rivers, canopy types from vegetation to urban surfaces and soil properties), and an extended range of resolved atmospheric spatio-temporal scales toward finer processes (Lucas-Picher, Laprise, & Winger, 2017). For all these reasons, RCMs are excellent candidates for driving hydrological models since, compared to coarse-resolution GCMs, they can better account for processes relevant to the scale of many hydrological applications.

Since they provide hydrologically relevant output variables such as precipitation, runoff and evapotranspiration, RCMs can already be used to assess some hydrological impacts from climate change without the need to run a hydrological model (e.g. Music et al., 2012). At the basin scale, however — where complex topography and heterogeneity in soil characteristics are important factors — applications using RCM-driven hydrological models are increasingly popular in the assessment of the hydrological impacts of climate change. It is generally recognized that RCM data may need to be bias-corrected to ensure that calibrated hydrological models are driven by realistic meteorological conditions (Muerth et al., 2012a). However, there is some debate as to whether an RCM output should be, or not, bias-corrected prior to drive a hydrological model, as bias correction may introduce further uncertainty into future hydrological simulations (Chen, Brissette, Liu, & Xia, 2017). Therefore, raw RCM outputs may be preferred to drive hydrological models for some applications, as when Lucas-Picher, Riboust, Somot, et Laprise (2015) reconstructed the Richelieu River flooding of spring 2011, one of the most important flood that occurred in Québec over the last years.

The use of a hydro-modelling chain including a GCM, an RCM and a hydrological model appears to be necessary for the proper assessment of hydrological impacts driven by climate change. This approach, however, requires the various sources of uncertainty that may affect climate-change projections be considered. The World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project (CMIP) multi-model datasets CMIP3 (Meehl et al., 2007), CMIP5 (Taylor et al., 2011) and the upcoming CMIP6 (O'Neill et al., 2016) are vast multi-model ensembles that allow to sample the three main sources of uncertainty: 1) future pathway (scenario) of greenhouse-gas and aerosol (GHGA) emissions; 2) climate sensitivity (structural uncertainty) to fixed GHGA emissions scenario; 3) natural climate variability. These uncertainties are sampled using an "ensemble of opportunity" framework: modelling centres around the world voluntarily generate simulations (based on their own resources and interests) using different GHGA-emission scenarios and GCM models. Some modelling centres also generate multiple realizations of the same experiment (i.e. a specific GCM model driven by a specific GHGA scenario), by adding slight

perturbations to the model's initial conditions to sample the effect of natural climate variability (Deser et al., 2014) — an approach that reflects the intrinsic chaotic nature of the climate system (Lorenz, 1963). Ensembles involving multiple RCMs are also increasingly common, as they are built upon CMIP-like ensembles of GCMs, such as the CORDEX-coordinated project (Giorgi & Gutowski, 2015), which consists of a multi-scenario, multi-GCM, multi-RCM ensemble.

Given the large amount of resources involved in the production of climate model simulations, the multi-model ensemble framework does not generally provide every possible combination of scenarios and models. In addition, models are often represented by a single realization, leading to a weak sampling of natural climate variability. In this sense, it is important to note that, for short-term climate projections, the contribution from natural climate variability to uncertainties is often more important than the contributions from the other factors (Hawkins & Sutton, 2009, 2011). Moreover, as extreme events are by definition rare, multiple realizations from one model are important to more robustly assess how climate change may affect their occurrence and intensity. For extremes floods, for instance, short-term data records translate into large uncertainties for 100-year return-level estimates (Schulz & Bernhardt, 2016).

To better understand the role of natural variability and extreme events in current climate projections, it has become increasingly popular in recent years to use the large-ensemble framework, consisting of using a single GCM to generate several realizations of a same experiment. Recent examples are the Community Earth System Model Large Ensemble (CESM-LE) (Kay et al., 2015), which now contains at least 40 members of transient climate-change projections under the RCP8.5 emissions scenario, or its 15-member RCP4.5 version (Sanderson et al., 2018). Similarly, the CanESM2 Large Ensemble (CanESM2-LE) (Fyfe et al., 2017) was produced by the Canadian Centre for Climate Modelling and Analysis (CCCma) at Environment and Climate Change Canada (ECCC), and consists of 50 members under the RCP8.5 scenario. Two 40-member ensembles use the CESM model driven by

historical radiative forcing, one using a dynamical ocean model, and the other one observed sea-surface temperatures (Mudryk, Kushner, & Derksen, 2014). The Dutch Challenge Project produced another ensemble, consisting of 62 members from the Community Climate System Model (CCSM1) driven by a business-as-usual scenario (Selten, Branstator, Dijkstra, & Kliphuis, 2004). Also worth noting is the "Essence" project (Sterl et al., 2008), a 17-member ensemble of climate-change simulations using the ECHAM5/MPI-OM climate model forced by the "Special report on Emissions Scenarios" (SRES) A1B pathway. All of these large ensemble projects use many initial-condition members to filter the effects of internal variability to better detect the climate-change signal related to a phenomenon of interest.

As natural climate variability can highly depend on the spatial scale under consideration (Giorgi, 2002), a better assessment of local climate-change impacts from natural variability and extreme events implies that the regional climate modelling community also began to follow the large-ensemble framework (Deser et al., 2014). An important recent example is "Database for Policy Decision-Making for Future Climate Change" (d4PDF) (Mizuta et al., 2017), which involved the dynamical downscaling of a GCM large ensemble at a spatial resolution of 20 km over Japan. Also, the Canadian Regional Climate Model version 4 (CanRCM4) was used to perform a 35-member ensemble over North America on a 50-km grid mesh (Fyfe et al., 2017). Another example is the 16-member ensemble performed over western Europe and the Alps using the Royal Netherlands Meteorological Institute's regional model KNMI-RACMO2 at 12-km resolution driven by the EC-EARTH global model (Aalbers et al., 2017).

In the scope of the ClimEx project, a 50-member ensemble of climate-change projections at 12-km resolution was produced to assess hydrological impacts from climate change in Bavaria and Québec. This paper presents initial results from this new dataset — the Canadian Regional Climate Model (CRCM5) Large Ensemble (CRCM5-LE; Ouranos 2017, unpublished data) — which is characterized by continuous simulations from 1950 to 2100 under the RCP8.5 GHGA emission scenario and was produced over two domains, Europe

and northeastern North America. CRCM5-LE consists of a dynamically downscaled version of CanESM2-LE, which was used to drive the CRCM5 through its boundary conditions.
This paper is organized as follows. Section I.2 describes the experimental framework of CRCM5-LE. Section I.3 presents an analysis of the results as follows: model validation with observations (Section I.3.1) and climate-change projections of mean climate (Section I.3.2), natural variability (Section I.3.3) and extremes (Section I.3.4). Section I.4 provides a discussion and conclusions.

I.2 The ClimEx experimental framework

Figure-A I-1 shows the general framework of the ClimEx experiment. The Canadian Earth System Model version 2 (CanESM2; Arora et al., 2011), developed at the CCCma, was used to generate a large initial-condition ensemble of climate-change projections at 2.8 resolution. This dataset, namely the CanESM2 Large Ensemble (CanESM2-LE; Fyfe et al., 2017), was generated by first using the CMIP5 piControl equilibrium simulation of 1,000 years under pre-industrial conditions (i.e. under constant 284.7 ppm atmospheric CO2 concentration) and then applying random atmospheric perturbations (in the cloud-overlap parameter) resulting in the five CMIP5 historical runs. These runs were forced with observed emissions (CO2 and non-CO2 GHGs, aerosols and land use) including observed explosive volcanoes and solar-cycle forcings up to year 2005. To obtain 50 members, cloud-overlap random perturbations were applied ten times to each of the five members on 1 January 1950, leading to 50 ensemble members that become independent after days to weeks, and thus being assumed as independent atmospheric-climate realizations over yearly to multi-decadal time scales. The set of 50 historical simulations was extended from 2006 to 2100 by using radiative forcing from the representative concentration pathway RCP8.5. Post 2005, simulations are forced by a repetition of roughly the last observed solar cycle (prior to 2006) without volcanic aerosol forcing.

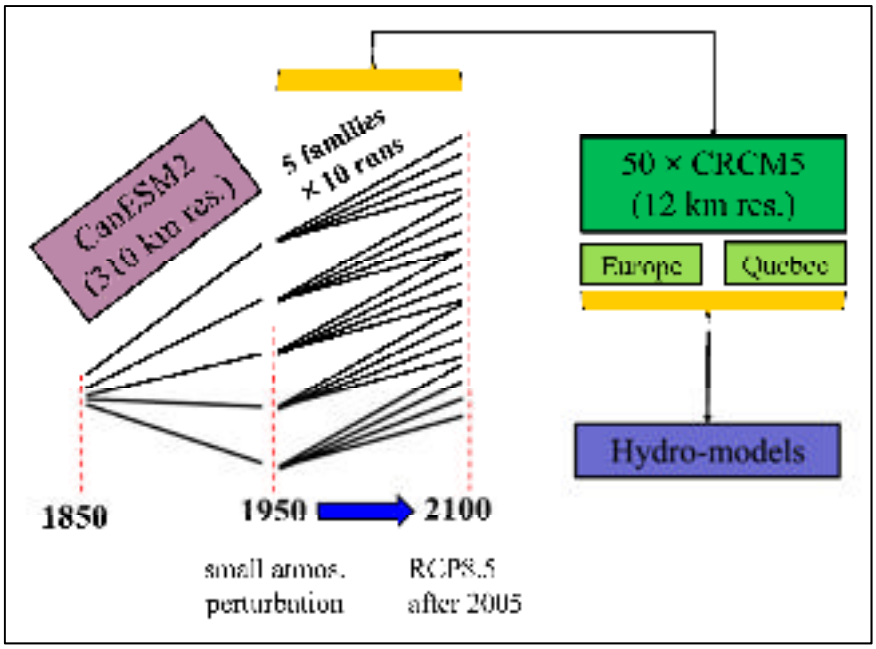

Figure-A I-1 Schematic representation of the ClimEx modelling chain where the CanESM2 members are used to drive the CRCM5 and hydrological models

The Canadian Regional Climate Model version 5 (CRCM5; Martynov et al., 2013; Šeparović et al., 2013), developed by the ESCER Centre (Centre pour l'Étude et la Simulation du Climat à l'Échelle Régionale) of l'Université du Québec à Montreal in collaboration with Environment and Climate Change Canada, was used by the Ouranos Consortium on Regional Climatology and Adaptation to Climate Change to dynamically downscale CanESM2-LE from 2.8° (≈310 km) to 0.11° (≈12 km) resolution over the 1950-2100 period. The downscaling experiment was performed for two domains — Europe (EU) and northeastern North America (NNA) — both using an integration domain of 380 x 380 grid points (Figure-A I-2). CRCM5 lateral boundary conditions are updated every six hours and linearly interpolated to the five-minute time step of the model. GCM output fields of temperature, surface pressure, specific humidity and horizontal wind components are used to drive the RCM with a one-way nesting procedure over a 10 grid points (≈120 km) surrounding blending zone (Davies, 1976). A smooth spectral nudging of large scales (Riette & Caya,

2002; Separovic, de Elía, & Laprise, 2012) was applied to the horizontal wind component within the RCM domain interior. The spectral nudging configuration consists of large-scale features being defined with a half-response wavelength of 3,113 km and a relaxation time of 13.34 hours. These large scales are imposed inside the RCM domain and vary along the vertical: the nudging strength is set to zero from the surface to a height of 500 hPa and increases linearly onward to the top of the model's simulated atmosphere (10 hPa). In order to validate the performance of the CRCM5 over the two domains, ERA-Interim driven runs were performed at the same resolution (12 km) and covering the period from 1979 to 2012. A smooth spectral nudging of large scales was applied as well to these simulations.

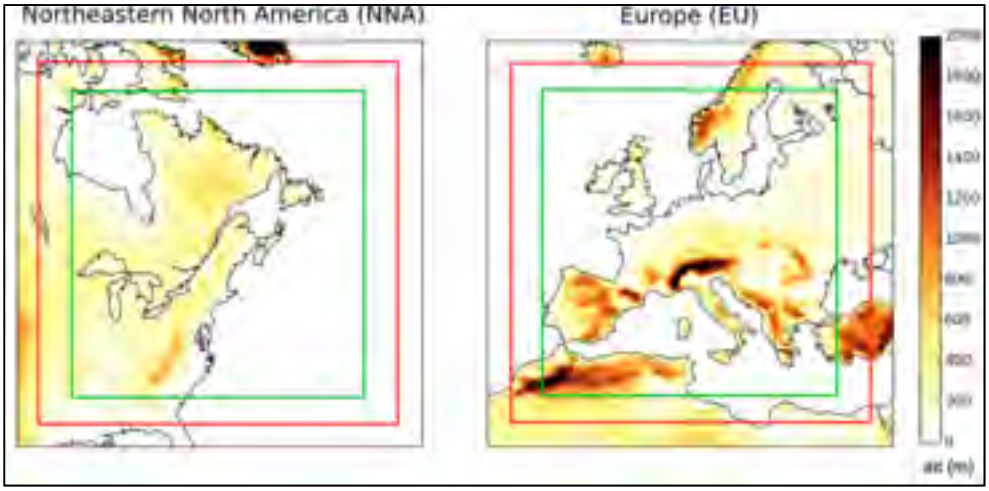

Figure-A I-2 Topography used by CRCM5 to produce the ClimEx large ensemble over northeastern North America (NNA: left panel) and Europe (EU: right panel). For each case is shown the integration domain (380x380 grid points, in black), the "free domain" (340x340 grid points, in red) where the model is technically free from direct imposition of lateral boundary conditions, and the "analysis domain" (280x280, in green), that is the region where are archived all the output fields

Removing both the 10 grid point wide Davies' blending zone and the 10 point halo (which provides upstream data in the semi-Lagrangian interpolation) included in the periphery of the integration domain results into a 340 x 340 "free domain", where the model is technically

free from direct imposition of lateral boundary conditions. However, RCM applications are known to suffer from boundary effects inside their free domain because small-scale features — which are absent from the lateral boundary conditions — need space (Leduc & Laprise, 2009; Leduc, Laprise, Moretti-Poisson, & Morin, 2011; Matte, Laprise, Thériault, & Lucas-Picher, 2017) and time (de Elía, Laprise, & Denis, 2002) to develop from the coarse-resolution boundary conditions. For this reason, an additional 30 grid point wide security zone was removed within the free domain to favour the development of fine-scale features over the region of interest, corresponding to a 280 x 280 grid points analysis domain (Figure 2) over which all CRCM5 outputs were archived. The lists of variables and archival time frequencies to be outputted in the scope of the CRCM5 Large Ensemble (CRCM5-LE) were designed according to a balance between project priorities, potential use for future projects and disk space limitations. The archival time period varies between one hour (for precipitation only), three hours (e.g., surface-air temperature) and daily (e.g., radiative fluxes). The complete list of available variables with corresponding archival frequencies is posted at www.climex-project.org.

I.3 Results

I.3.1 Validation of the historical climate

In this section, a general climatological evaluation of the CRCM5 model output is performed. Historical climates, as simulated by the model when driven by either CanESM2 or ERA-Interim, are compared with observational gridded datasets: E-OBS (0.22° resolution; Haylock et al., 2008) for the EU domain and the Climatic Research Unit time series (CRU TS; 0.5° resolution; Harris, Jones, Osborn, & Lister, 2014) for NNA. Data are linearly interpolated onto the CRCM5 grid for comparison. Here, only one member of the large ensemble (rather than the ensemble mean) is used to make a proper comparison with the single realization of the ERA-Interim run. Using a 32-year climate period for validation, the climates of the different members slightly differ due to internal variability, but the general

conclusions drawn from this validation hold across the ensemble. While the following discussion focusses on the differences between CRCM5 output and the observed climatology, the simulated climatology of the different variables and domains can be found in Supplementary Figures-A I-S19 to I-S23.

Beginning with the seasonal mean surface-air temperature averaged over the 1980-2012 period for the EU domain, Figure-A I-3 (left column) presents the difference between the CRCM5 climatology when driven by CanESM2 and the E-OBS dataset. It can first be seen that biases in the CRCM5 simulations depend on geographical location and season. A generalized cold bias (around -2 °C) occurs during fall and spring. Winter biases are generally small and mostly range between -1 °C to 1 °C, while warm biases larger than 2 °C arise during summer in the eastern part of the domain. By comparing with the CRCM5 topography in Figure-A I-2, systematic warm biases in winter appear over mountainous regions such as the Alps, Pyrenees, Balkans and the Carpathians.

Figure-A I-4 (left column) shows the difference between CRCM5 (CanESM2 driven) and the E-OBS dataset for the mean daily precipitation over the EU domain for the 1980-2012 period. Throughout the year, there is a wet bias over several parts of Europe. During winter, it exceeds 4 mm/day in western Europe and remains around 1 mm/day in eastern Europe. The wet biases decrease during spring in eastern regions, leading to dry biases (larger than -1 mm/day) in summer.

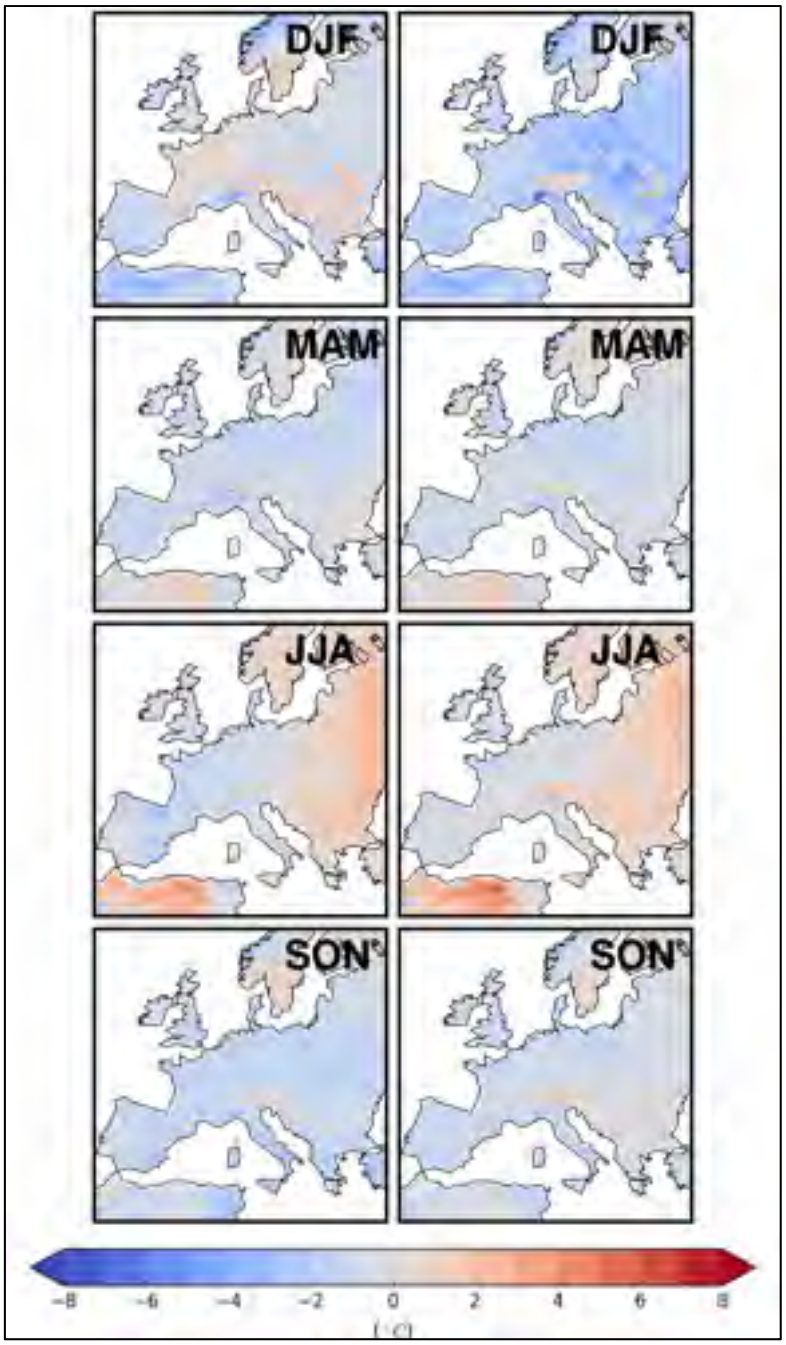

Figure-A I-3 Differences between mean seasonal surface-air temperature over the 1980-2012 period between E-OBS and the (first column) CRCM5 run driven by CanESM2 (first member) and the (second column) CRCM5 run driven by ERA-Interim for the EU domain. A positive difference corresponds to an overestimation of the observed values

Figure-A I-4 Differences between mean seasonal precipitation over the 1980-2012 period between E-OBS and the (first column) CRCM5 run driven by CanESM2 (first member) and the (second column) CRCM5 run driven by ERA-Interim for the EU domain. A positive difference corresponds to an overestimation of the observed values

Biases in the large regional ensemble can originate from both driving (here CanESM2) and driven (here CRCM5) models. To better discriminate between these bias sources, it is useful to compare the biases found in the previously analyzed CanESM2-driven run, due to both the CRCM5 and its lateral boundary conditions (CanESM2), with those of a reanalysis-driven run, which encompasses CRCM5 and reanalysis biases. For temperature, CRCM5 driven by the ERA-Interim reanalysis (Figure-A I-3, right column) shows a strong cold bias during winter, which was not observed when driven by CanESM2 (Figure-A I-3, left column). However, the ERA-Interim driven run slightly attenuates the warm biases in summer, as well as the cold biases in fall. For precipitation (Figure-A I-4), the wet bias during winter, spring and fall for the CanESM2 driven simulation is generally smaller for the ERA-Interim driven simulation. Notably, during winter, the bias changes from more than 5 mm/day to less than 2 mm/day. Similarly, the summer dry bias in the eastern part of the domain decreases by about 1 mm/day. For both temperature and precipitation, it is worth noting that the part of the bias of the CRCM5 when driven by the reanalysis is comparable to that of other RCMs, according to Kotlarski et al. (2014), where multiple RCMs were validated using the E-OBS dataset in the scope of the coordinated CORDEX experiment.

Figure-A I-5 (left column) shows the difference between CRCM5 (CanESM2 driven) and the CRU dataset for the seasonal mean surface-air temperature averaged from 1980 to 2012 over the NNA domain. Bias patterns have a high seasonal dependency for northern Québec, while a systematic warm bias is visible throughout the year over a large part of the domain, exceeding 5 °C in winter and summer. The cold bias in northern Québec (around -4 °C) appears in winter and progresses southward during spring, while warm bias generally appears during summer and fall. Despite the lack of validation data over the Great Lakes, it is also worth mentioning the contrasted enhancement of air temperature during winter as the lakes do not completely freeze, a distinctive climatic feature of this region (not shown).

Figure-A I-6 (left column) shows the difference between CRCM5 (CanESM2 driven) and CRU for precipitation over the NAA domain. A systematic wet bias ranging between 1-2

mm/day exists for most of the domain during winter and spring, and to a lesser extent during fall. During summer, a dry bias emerges in the western part of the domain and exceeds -2 mm/day. While this is not visible in these maps, precipitation enhancement can be observed downwind (eastward) of the Great Lakes, when water is warmer than surrounding land and remains partially ice free during winter. Other analyses using CRCM5 simulations (Lucas-Picher et al., 2017) and other RCM (Notaro, Zarrin, Vavrus, & Bennington, 2013) have shown that this corresponds to the well-known lake-effect snowbelt patterns, another distinctive feature of the Great Lakes region.

Similarly to the EU domain, the right column in Figures-A I-5 and I-6 shows, for temperature and precipitation respectively, the difference between the ERA-Interim driven CRCM5 simulation and CRU in order to evaluate the sensitivity of the CRCM5 bias to the nature of the lateral boundary conditions over the NAA domain. For temperature, the important warm bias that was detected over most parts of the domain and through the year for the CanESM2 driven run (Figure-A I-5, left column) is systematically attenuated (from more than 7 °C to less than 2 °C in several regions) for ERA-Interim driven run (Figure-A I-5, right column). The cold bias in northern Québec, however, persists independently of the lateral boundary conditions, which suggests that the bias may originate from the CRCM5 itself. For precipitation (Figure-A I-6), the large dry bias observed during summer in south-west part of the domain practically vanishes, while a systematic wet bias (around -1 mm/day) persists over the entire domain and throughout the year for the reanalysis-driven run. As a comparison with other state-of-the-art RCMs, Martynov et al. (2013) showed that the performance of the ERA-Interim reanalysis-driven CRCM5 was satisfactory and similar to other models in terms of reproducing the current climate over North America. Similarly, Diaconescu, Gachon, Laprise, et Scinocca (2016) have shown that the performance of the CRCM5, when driven by CanESM2, was comparable to other RCMs in terms of several precipitation indices over the Canadian Arctic.

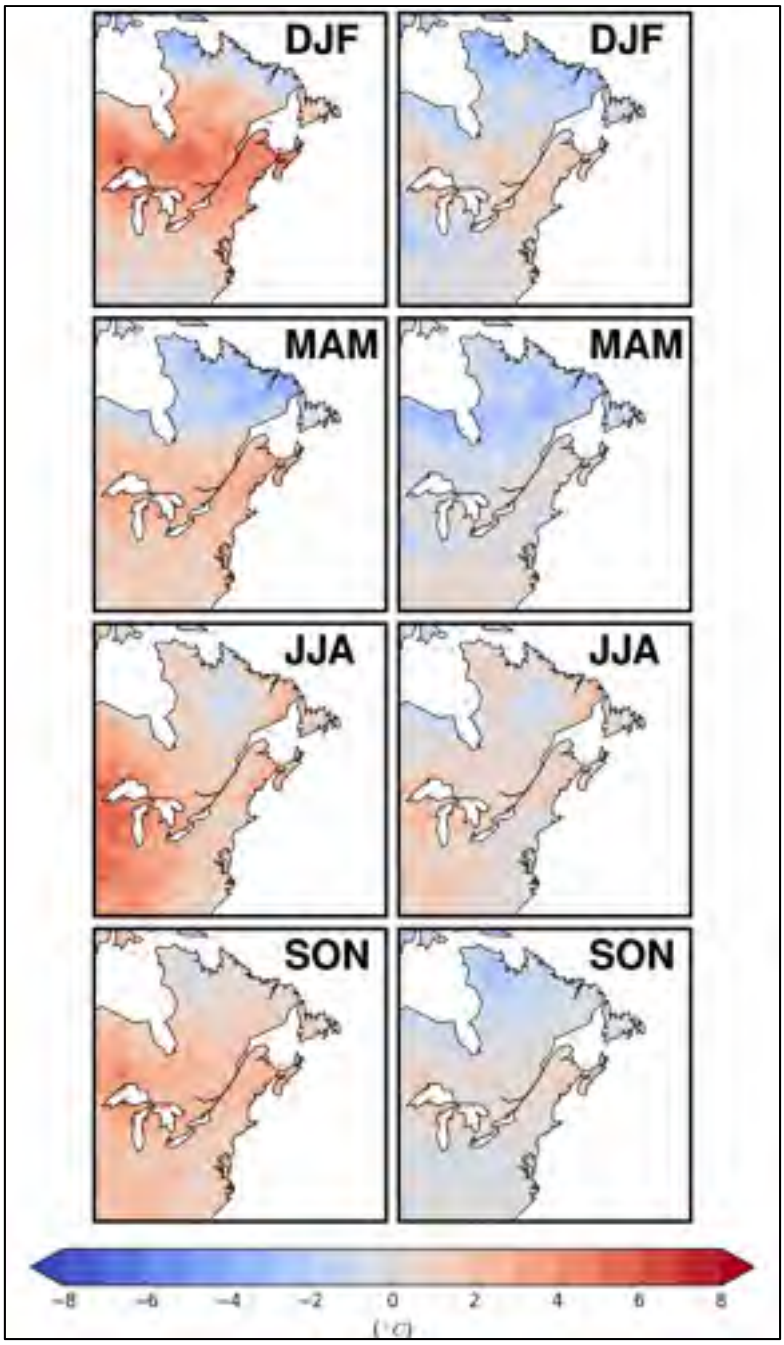

Figure-A I-5 Differences between mean seasonal surface-air temperature over the 1980-2012 period between CRU and the (first column) CRCM5 run driven by CanESM2 (first member) and the (second column) CRCM5 run driven by ERA-Interim for the NNA domain. A positive difference corresponds to an overestimation of the observed values

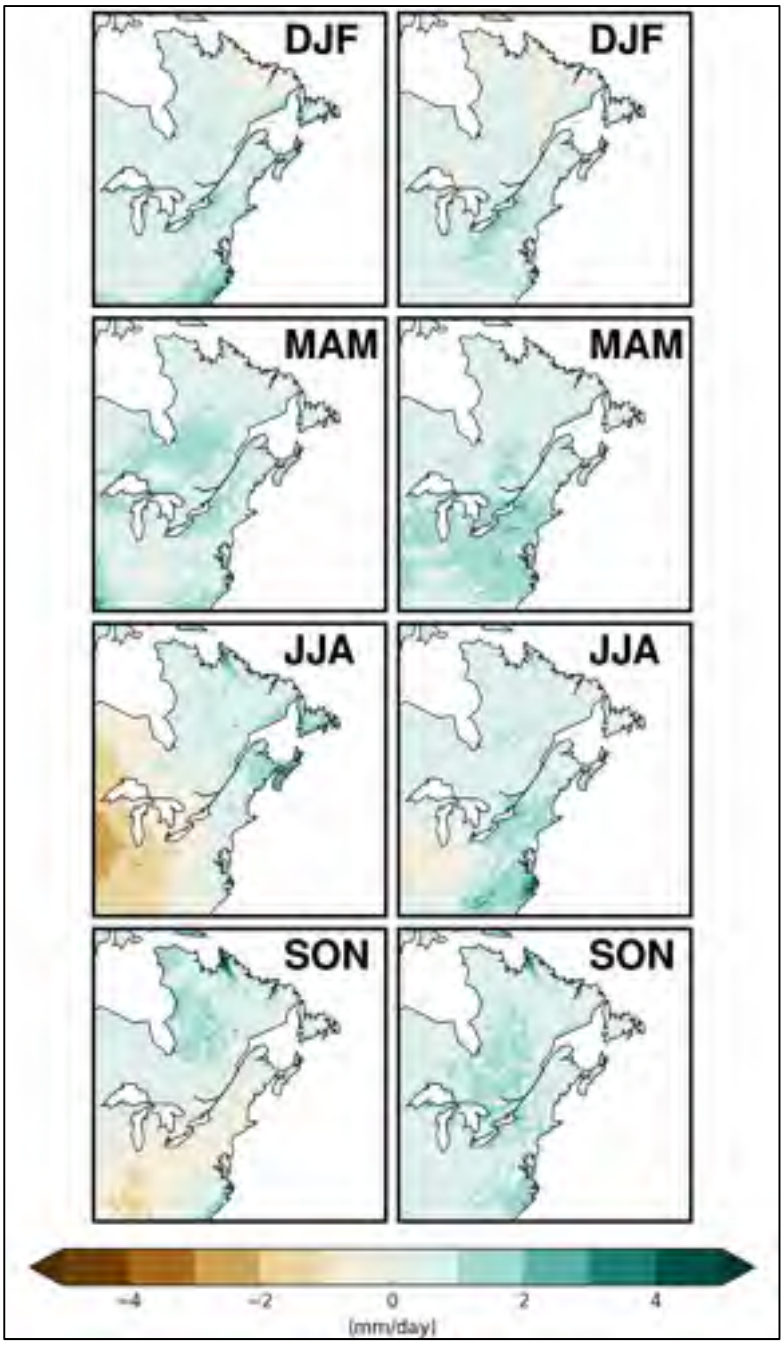

Figure-A I-6 Differences between mean seasonal precipitation over the 1980-2012 period between CRU and the (first column) CRCM5 run driven by CanESM2 (first member) and the (second column) CRCM5 run driven by ERA-Interim for the NNA domain. A positive difference corresponds to an overestimation of the observed values

I.3.2 Projected changes in climatological means

Figure-A I-7 presents the short-term projected changes (2020-2039 versus 2000-2019) in precipitation for December estimated from ensemble members 1 to 24 over both domains. Recalling that the ensemble members differ only by slight random perturbations in their atmospheric initial conditions, these results clearly show how natural variability can lead to very different projections, both in terms of intensity and spatial distribution, for this particular variable, and time horizon. Some regions with strong changes may even show opposite signs for different members (e.g., members 4 and 6 over both domains). This also demonstrates how the practical use of single-member ensembles of regional climate projections may lead to misleading recommendations for planning short-term adaptation strategies to climate change. To focus on climate-change features that are robust across the ensemble, the ensemble mean signal is analyzed in the following. The statistical significance of the signal will be quantified by applying a Student's t test on the difference between future and historical ensemble-mean climates, and the dependence of this measure to the time horizon and the ensemble size will be assessed.

Ensemble mean climate-change signal between the 2000-2019 and 2080-2099 periods for the monthly mean surface-air temperature over the EU domain is first analyzed (Figure-A I-8). The signal is stronger from June to September, with August showing temperature increases exceeding 8 °C in western and southeastern Europe. There is also an enhanced warming in the northeastern part of the domain during winter, partly attributable to the decreasing snow cover-albedo feedback (Fischer et al., 2010).

Figure-A I-9 shows the ensemble mean climate-change signal for monthly mean precipitation over the EU domain (2080-2099 versus 2000-2019). These simulations show that the climate in Europe will become dryer in summer and wetter in winter. Precipitation increase in December is as large as 2 mm/day on the west side of the Alps and along the west coast of the Balkan Peninsula. A large decrease of 2 mm/day in summer precipitation is detected

during July and August on both the north and south sides of the Alps. However, the projected changes in precipitation are not significant everywhere, even for such a far horizon, as can be seen from the hatched regions, where the signal is not statistically significant. Notably, precipitation changes in winter over the Mediterranean Sea and the Iberian Peninsula are too weak to emerge from the noise of natural climate variability.

In order to investigate the relative contribution of natural variability and climate-change signal, changes in temperature and precipitation over different future periods were estimated and compared for the ensemble mean of the 50 members and a much smaller sample based on the first five members of the ensemble. Figures 10-a, b and c show the 50-member ensemble mean temperature change (for December only) for three different time horizons; 2020-2039 (short term), 2040-2059 (mid-term) and 2080-2099 (long term; repeated from Figure-A I-8 for clarity) respectively. Similarly, Figure-A I-10-d, e, and f show the five-member ensemble mean temperature over the same three future periods. The five-member mean results are very similar to those of the full ensemble and the signal remains statistically significant everywhere in the domain for both mid-term (2040-2059) and long-term (2080-2099) projections. However, when considering short-term projections (2020-2039), the 50-member ensemble still shows statistically significant changes (Figure-A I-10-a), while the signal has not emerged from natural variability over most land areas for the five-member ensemble (Figure-A I-10-d). Similar conclusions hold for other months (see also Supplementary Figures-A I-S23, I-S27 and I-S28).

Comparing the 50- and five-member ensemble mean precipitation change for July (Figure-A I-10-g to l), the general features seen for the 50-member ensemble are still present for the five-member ensemble. Particularly, for long-term projections, the decrease in precipitation is statistically significant, although the intensity of the change is greater for this particular five-member ensemble. For short-term projections (2020-2039), the 50-member ensemble allows to detect small significant decreases in precipitation for western and southwestern Europe (Figure-A I-10-g), while the five-member ensemble mean displays practically no

region with statistical significance changes in the short term, and very few statistically significant areas in the mid-term (Figure-A I-10-j). It is interesting to note that larger part of the domain with statistically significant changes for the short-term period are reported for the 50-member ensemble than for the mid-term period for the five-member ensemble. These conclusions generally hold for the other months (see also Supplementary Figures-A I-S24, I-S29 and I-S30). In several cases, even the long-term projections show very low statistical significance for the five-member ensemble while the 50-member ensemble generally allows to detect a signal over an appreciable fraction of the domain.

Repeating the previous analysis for the NNA domain, the climate-change signal in 2080-2099 for the monthly mean temperature is shown in Figure-A I-11 based on the 50-member ensemble. A prominent maximum increase of temperature appears over the Hudson Bay. It exceeds 14 °C from January through March and attenuates in April. It is worth noting that this regional feature is mostly inherited from the CanESM2 driving model, because its sea-surface temperature and sea-ice values are prescribed to the CRCM5. The positive ice-albedo feedback occurs as Hudson Bay becomes partially covered, instead of completely covered, by sea ice during winter by the end of the 21st century in the CanESM2 simulations (not shown). The important temperature change in winter extends into northern Québec and is influenced by the feedback from Hudson Bay sea ice, and by snow-albedo feedback as snow cover decreases. The latter effect is particularly visible in northern Québec from April through June, when the transition zone (area where there is still snow in some members while not in the others) migrates northward through the future.

Figure-A I-12 shows the projected changes in precipitation over the NNA domain. From November through May, precipitation increases over land regions (exceeding 0.8 mm/day in northern Québec), Hudson Bay and Atlantic Ocean. In June, precipitation decreases by more than 0.4 mm/day over most land regions with the exception of northern Québec, and this drying pattern slowly decays until August, when only a small drying area remains over Ontario. Over the Atlantic, minimal change is observed during December, while precipitation

decreases slightly during April/May, to reach values exceeding -1.8 mm/day in July/August. The important decrease in summer precipitation occurs in the area of the North Atlantic storm track and might be related to the poleward shift of mid-latitude storm tracks (Woollings, Gregory, Pinto, Reyers, & Brayshaw, 2012), as well as to the weakening of the North Atlantic Meridional Overturning Circulation (Brayshaw, Woollings, & Vellinga, 2009) in CanESM2 simulations.

As for the EU domain, reducing the ensemble from 50 to 5 members does not significantly modify the patterns in temperature change (Figures-A I-13-a to f, results shown for December only). Short-term projections are also statistically significant for the 50-member ensemble (Figure-A I-13-a) while for the five-member ensemble (Figure-A I-13-d) the southern half of the domain shows practically no statistically significant change during winter. Similar conclusions are obtained for the other months, namely that statistically significant changes are observed everywhere with the exception of some regions in the short-term projection for the five-member ensemble (see also Supplementary Figures-A I-S25, I-S31 and I-S32).

Comparing the 50-member ensemble with a five-member ensemble for precipitation over the NNA domain (for July only), Figures-A I-13-j to l show that the fraction of the domain with statistically significant changes is very small for the five-member ensemble. For short-term projections, however, the 50-member ensemble (Figure-A I-13-g) already shows a significant, though small, decrease in precipitation in the western part of the domain, which progressively extends in size for the mid-term and long term projections. Similar results are obtained for the other months, that is, no statistically significant changes over the largest fraction of the domain for the five-member ensemble, even in long-term projections are observed, while the 50-member ensemble generally allows to detect such changes (see also Supplementary Figures-A I-S26, I-S33 and I-S34). But it is also important to note that precipitation change remains a challenging variable even with the full ensemble, as the signal is generally weak while the variability is high.

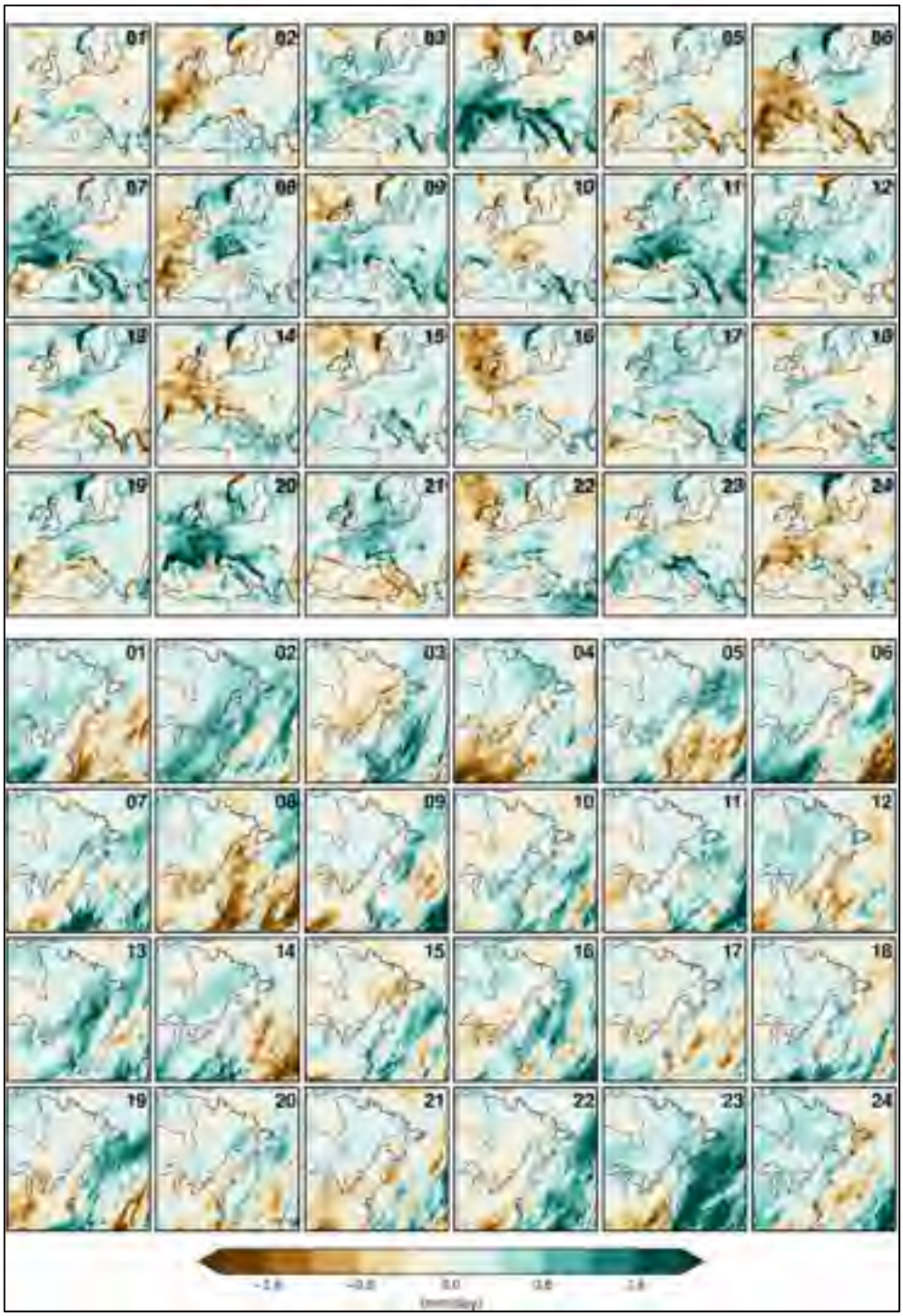

Figure-A I-7 Short-term climate-change projections (2020-2039 vs 2000-2019) for mean December precipitation from the ensemble members 1 to 24 over the EU (top rows) and NNA (bottom rows) domains

Figure-A I-8 The CRCM5 50-member ensemble mean climate-change signal for surface-air temperature computed as the difference between the 2080-2099 and 2000-2019 monthly climate means for the EU domain. All reported changes are statistically significant at the 99% confidence level (Student's t test with unequal variances)

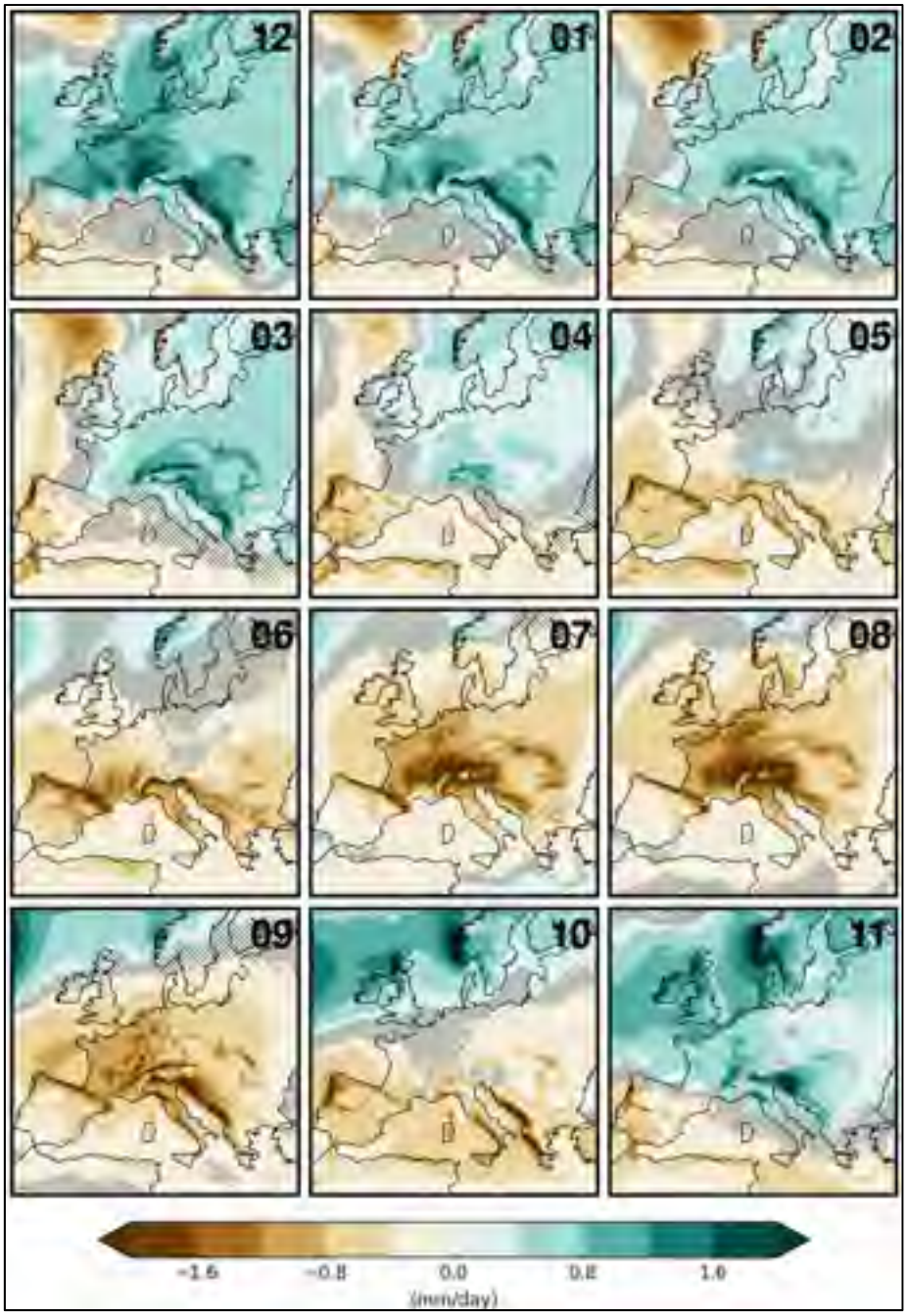

Figure-A I-9 Same as Figure-A I-8 for precipitation during the 2080-2099 period over the EU domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

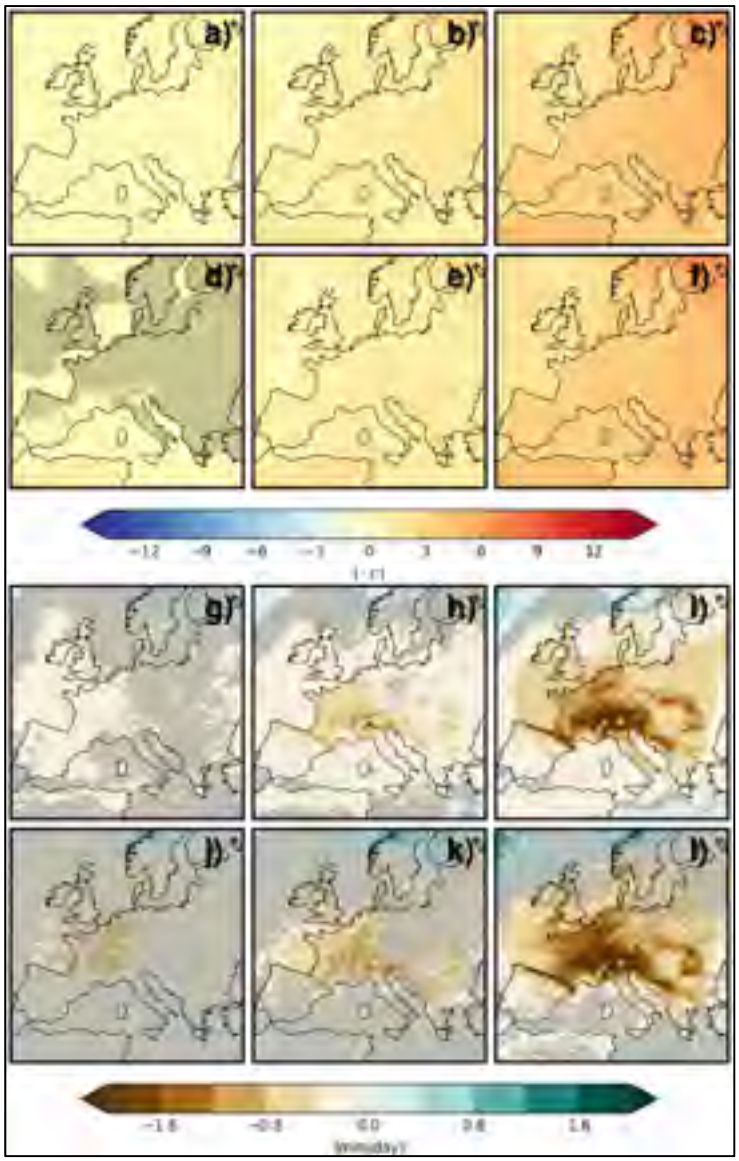

Figure-A I-10 (a) to (c): CRCM5 50-member ensemble mean climate-change signal for surface-air temperature during December over the EU domain computed for the (a) 2020-2039, (b) 2040-2059, and (c) 2080-2099 periods relative to 2000-2019; (d) to (f): Same as (a) to (c) for the first five members of the ensemble; (g) to (i) and (j) to (l): Same as (a) to (c) and (d) to (f)for precipitation during July. Panels (c) and (i) are reproduced from Figures-A I-8 and I-9 for clarity. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

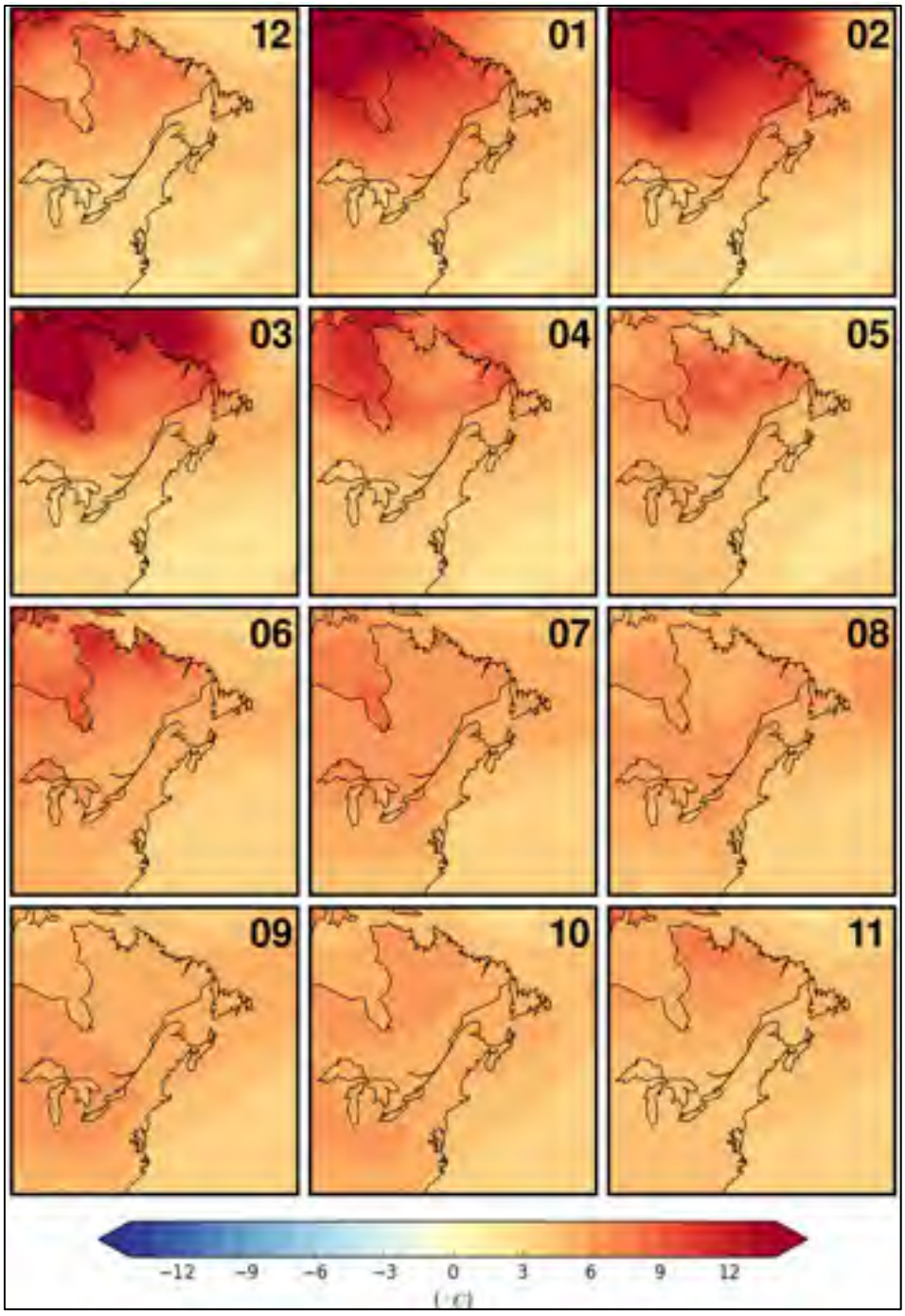

Figure-A I-11 Same as Figure-A I-8 for surface-air temperature during the 2080-2099 period over the NNA domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

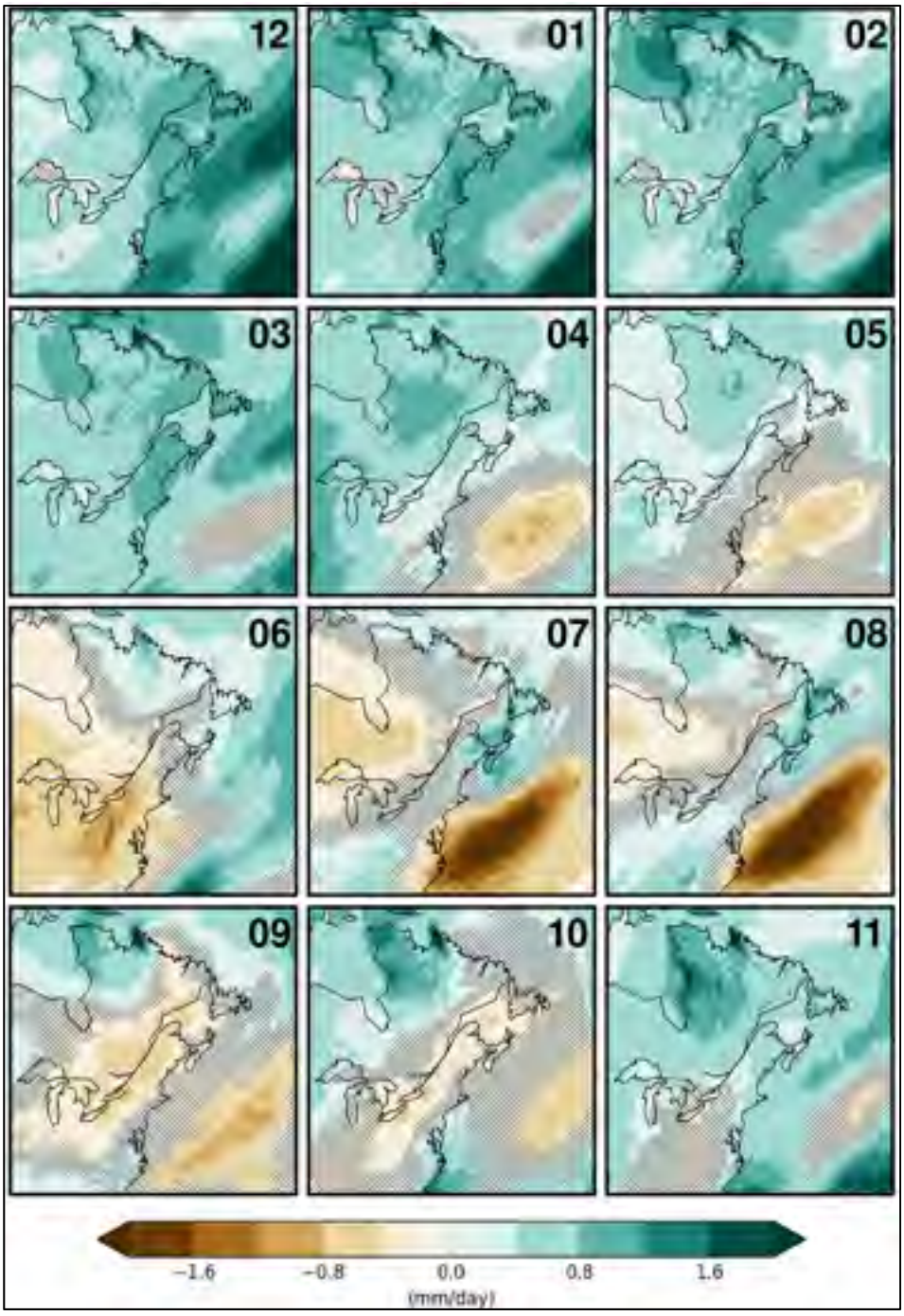

Figure-A I-12 Same as Figure-A I-8 for precipitation during the 2080-2099 period over the NNA domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

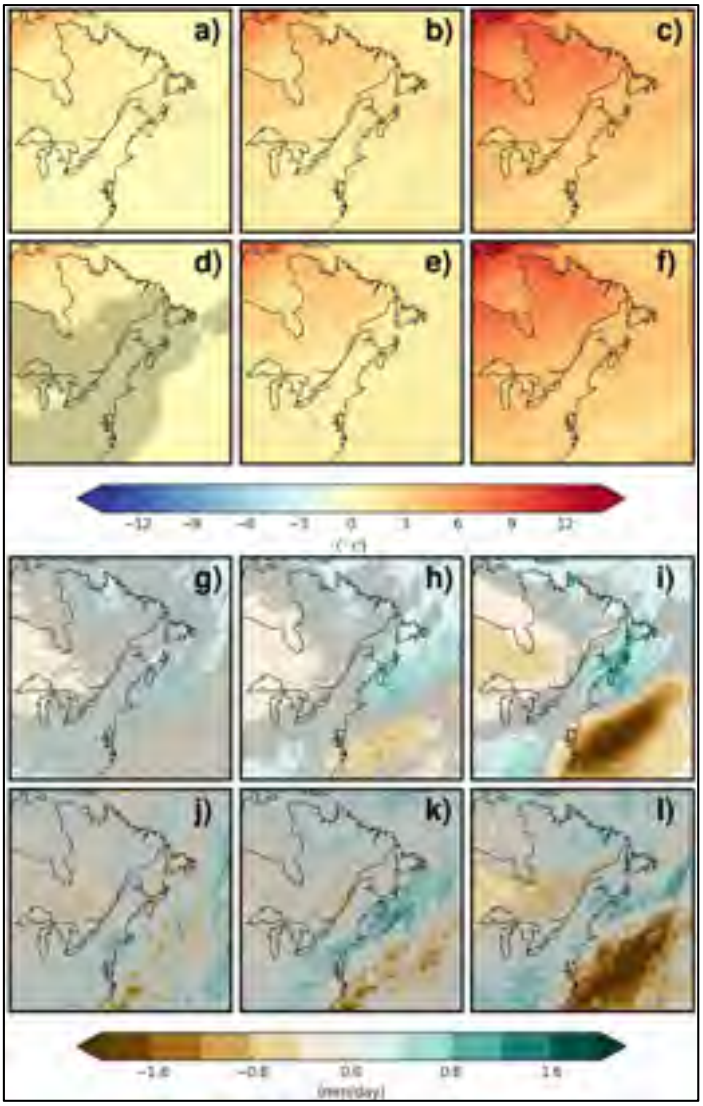

Figure-A I-13 (a) to (c): CRCM5 50-member ensemble mean climate-change signal for surface-air temperature during December over the NNA domain computed for the (a) 2020-2039,(b) 2040-2059, and (c) 2080-2099 periods relative to 2000-2019; (d) to (f): Same as (a) to (c) for the first five members of the ensemble; (g) to (i) and (j) to (l): Same as (a) to (c) and (d) to (f) but for precipitation during July. Panels (c) and (i) are reproduced from Figures-A I-11 and I-12 for clarity. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

I.3.3 Projected changes in temperature interannual variability

Here the large ensemble is used to assess the effect of climate change on temperature interannual variability, which can be defined as follows. Given a time window extending from year *a* to *b* inclusively, the overall variance calculated over this period of $P = b - a + 1$ years at a given grid-point can be written as

$$\sigma_{a,b}^2 = \frac{1}{P(N-1)} \sum_{t=a}^{b} \sum_{i=1}^{N} (X_{it} - \bar{X}_{ot})^2 \quad \text{(I.1)}$$

where N is the ensemble size ($N = 50$), X_{it} the monthly mean temperature over the given time period for member i and year t, and X_{ot} the ensemble mean (average over all members) at year t. Assuming ergodicity between temporal and inter-member variances (Nikiéma et al., 2017), $\sigma_{a,b}$ (i.e. the square root of equation I.1) can be interpreted as an estimate of the interannual variability for this specific time period. In the case of a climate system under transient forcing, the use of equation I.1 to assess temporal variability using the inter-member spread involves weaker assumptions than calculating the residual temporal variability from detrended time series. The latter approach is nevertheless popular when assessing natural variability using small ensembles (Hawkins & Sutton, 2009, 2011; Leduc, Laprise, de Elía, & Šeparović, 2016b; Leduc, Matthews, & de Elía, 2016c; Räisänen, 2002).

Figure-A I-14 shows the monthly patterns of interannual variability of surface-air temperature calculated over the 2000-2019 period for the EU domain. These patterns show a marked annual cycle reaching a maximum of around 4 °C during winter in the northern regions, while the variability generally remains below 2.5 °C for the rest of the year. The relative changes in interannual variability from 2000-2019 to 2080-2099 are presented in Figure-A I-15, where the statistical significance is assessed using the F-test with a 99% confidence level. A large increase in interannual variability occurs from May through September over most of western and central Europe, and extending into the Scandinavian

Peninsula. The maximum change is reached in August, when interannual variability increases by more than 70% (approximately 1 °C), compared to the 2000-2019 period for which interannual variability is around 1.5 °C (Figure-A I-14). In addition to the mean surface-air temperature increase of around 7 °C over this area and month in 2080-2099 (Figure-A I-8), this highlights the importance of considering the effect of climate change on both mean climate and interannual variability when investigating the effect of climate change on heat waves, for instance (Schär et al., 2004).

The important projected decrease in mean precipitation during summer (see Figure-A I-9) leads to a decrease in soil-moisture content (not shown) over a large part of Europe. The heat capacity of the land surface thus decreases, strengthening land-atmosphere coupling. As described in Seneviratne, Lüthi, Litschi, et Schär (2006), the enhancement of the land-atmosphere coupling over Europe is an important contributor to the projected increase in temperature interannual variability. For instance, the surface-air temperature becomes more strongly influenced by variations in incident solar radiation, which is converted into sensible rather than latent heat flux (Brown, Ming, Li, & Hill, 2017). This suggests that local temperature variability could highly depend on geophysical characteristics in this case. It is also worth noting that the increase in summer temperature interannual variability is known to relate to both land-atmosphere interactions and projected changes in global atmospheric circulation patterns (Meehl & Tebaldi, 2004).

For the rest of the year (i.e. October through April), Figure-A I-15 shows that interannual variability tends to decrease throughout the 21st century. Several physical mechanisms support this result. Sea-ice retreat in the North Atlantic plays a role as westerly circulation becomes less affected by sea-ice albedo variability, but also as the atmosphere is no more isolated from the ocean which has a much greater heat capacity (Stouffer & Wetherald, 2007). As another key physical mechanism that could explain this decreasing variability, it is known that sub-seasonal temperature variability is strongly affected by Arctic amplification. As shown by Screen (2014), rapid warming in the Arctic translates into a warming of cold air

advected by northerly winds, which decreases sub-seasonal variability of surface-air temperature.

Figure-A I-16 shows the annual cycle of interannual variability over the NNA domain for the period 2000-2019. Variability is much larger during the cold season in the northern part of the domain, which is in general agreement with observations (see Figure 1 in de Elía, Biner, & Frigon, 2013). From January through March, interannual variability exceeds 3 ° C for Hudson Bay and most of Québec. High values persist into April and May in a narrow region of maximum temperature variability that extends from the south shore of Hudson Bay and across Québec. It is worth noting that these regions are also characterized by a high level of interannual variability in snow-cover fraction (not shown). This corresponds with the transition zone separating permanent snow cover in the north and rare spring snow in the lower latitudes (Krasting, Broccoli, Dixon, & Lanzante, 2013). This link between high temperature variability and the edges of snow-covered regions is consistent with the results of Fischer, Lawrence, et Sanderson (2011).

Figure-A I-17 shows changes in monthly mean temperature interannual variability over the NNA domain from 2000-2019 to 2080-2099. There is a systematic decrease in interannual variability during winter over a dominant fraction of the domain and an increase during summer for the southern regions. This is in agreement with the relationship between temperature variability and thermal advection (Holmes, Woollings, Hawkins, & de Vries, 2016), based on the fact that land-sea temperature contrasts will tend to increase during summer and decrease during winter, while the temperature gradient from pole to equator decreases mostly during winter due to Arctic amplification.

The northernmost part of Québec experiences a 80% increase (corresponding to about 1 °C) in interannual temperature variability in May. This can be partly explained by the northward migration of the snow transition zone, which is located in the northernmost part of Québec in 2080-2090 while being around 10 further south in the reference period. In other words, the

snow cover in a specific year may completely disappear in May in the northernmost region for some ensemble members while persisting in others. So interannual variability increases in a region when persistent snow cover transforms into a new transition region (northernmost region of Québec), while inversely, a transition region that becomes permanently without snow will rather experience a decrease in interannual variability. This may also explains the narrow east-west band in northern Québec where variability decreases by 30% during May.

While a rich literature describes the physical mechanisms underlying changes in temperature variability, the patterns of these changes are often difficult to assess with a high degree of confidence when using smaller ensembles. Similarly to what was done in Section I-3.2, it can be shown that using only the first five members of the ensemble leads to much less regions where changes in temperature interannual variability are statistically significant. Nevertheless, it is worth noting that some general features can still be detected with the smaller ensemble, such as the general decrease in variability over the northern regions during winter, or the increasing variability that is specific to central Europe during summer. More details about these results can be found in Supplementary Figures-A I-S35 and I-S36.

Figure-A I-14 Interannual variability of monthly mean surface-air temperature over the EU domain calculated as the yearly inter-member spread averaged during the 2000-2019 period

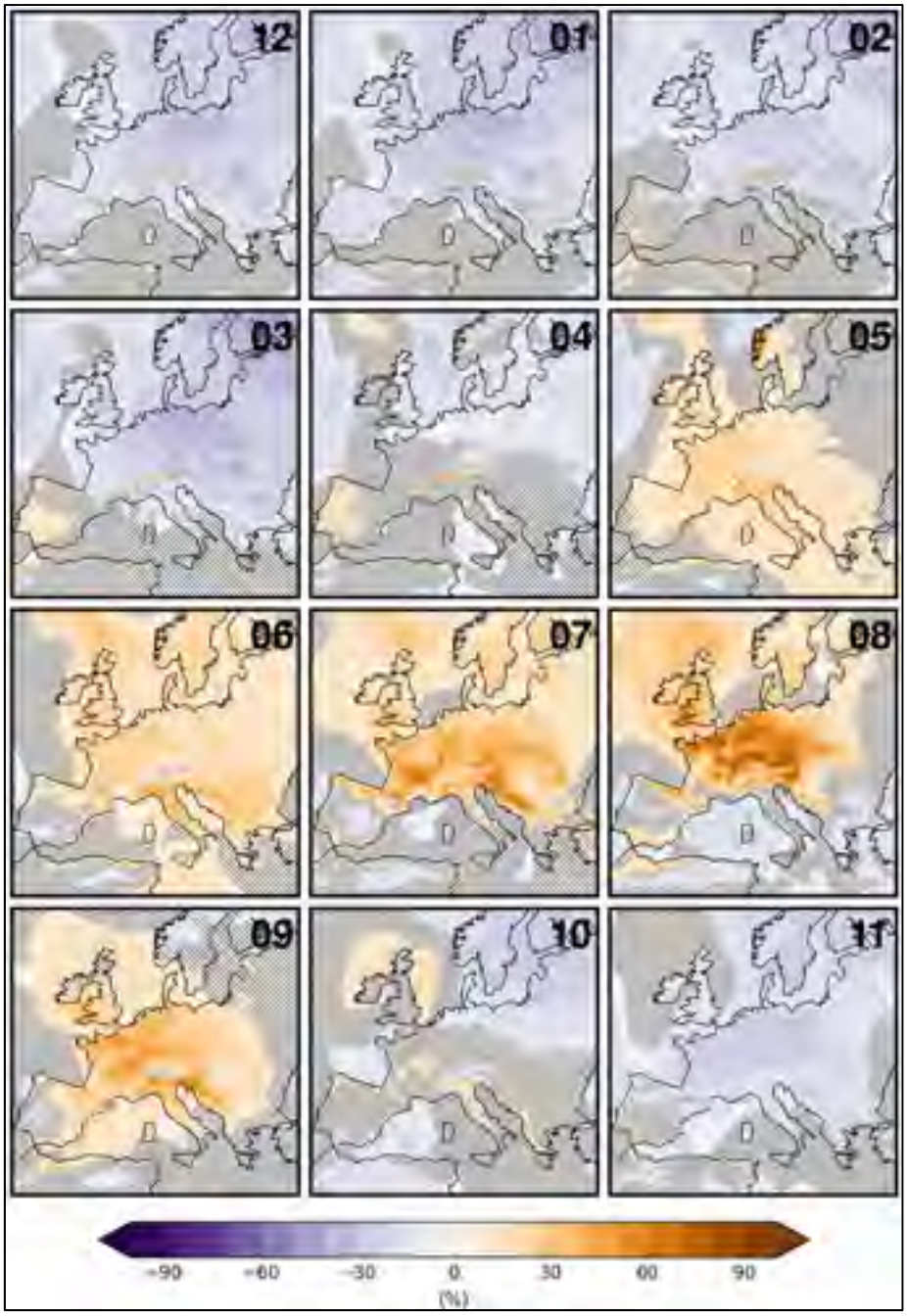

Figure-A I-15 Relative change in interannual variability for the monthly mean surface-air temperature during 2080-2099 relative to 2000-2019 over the EU domain. Hatched regions identify where the change in variability is not statistically significant at the 99% confidence level (F-test)

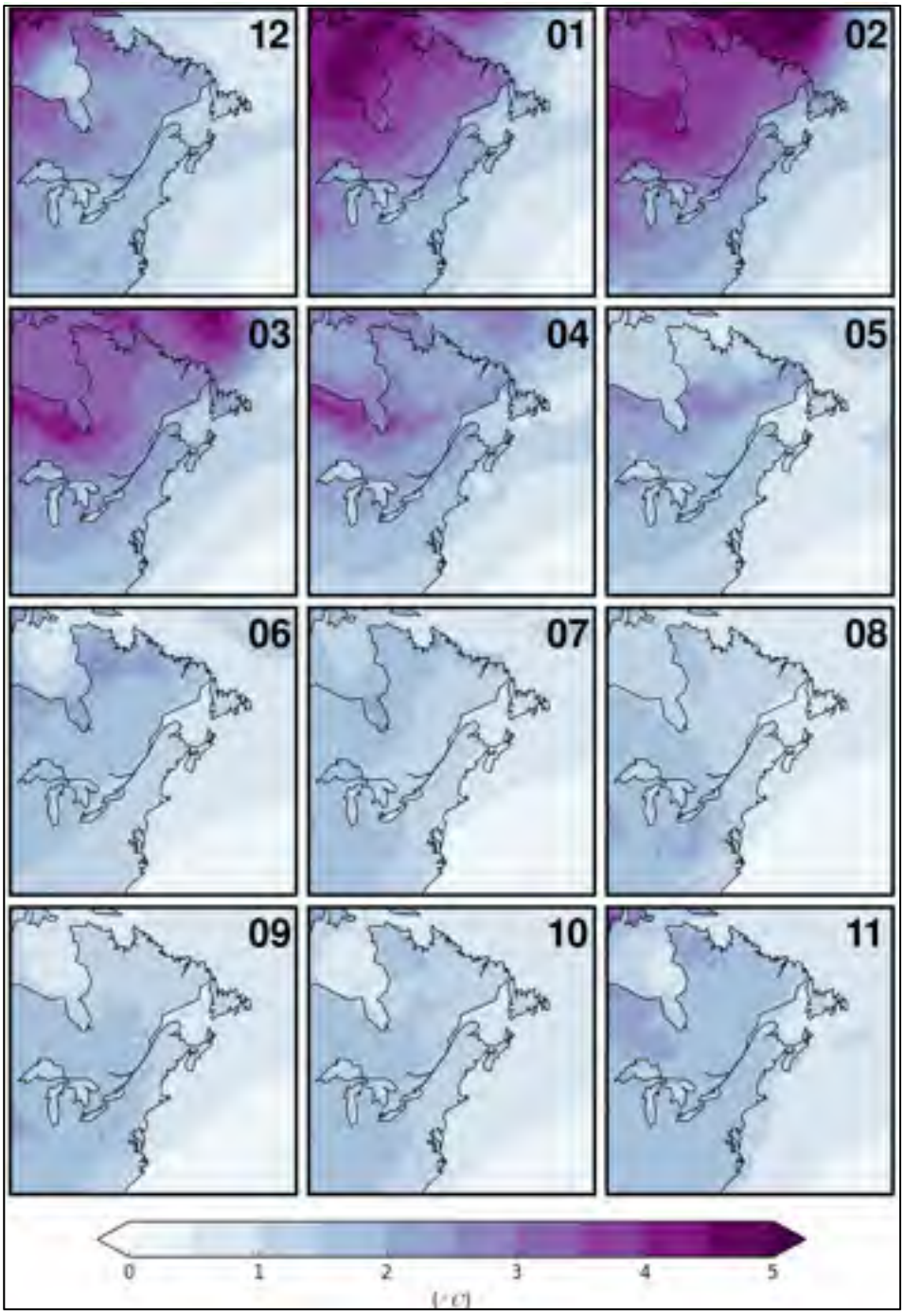

Figure-A I-16 Interannual variability of monthly mean surface-air temperature over the NNA domain calculated as the yearly inter-member spread averaged during the 2000-2019 period

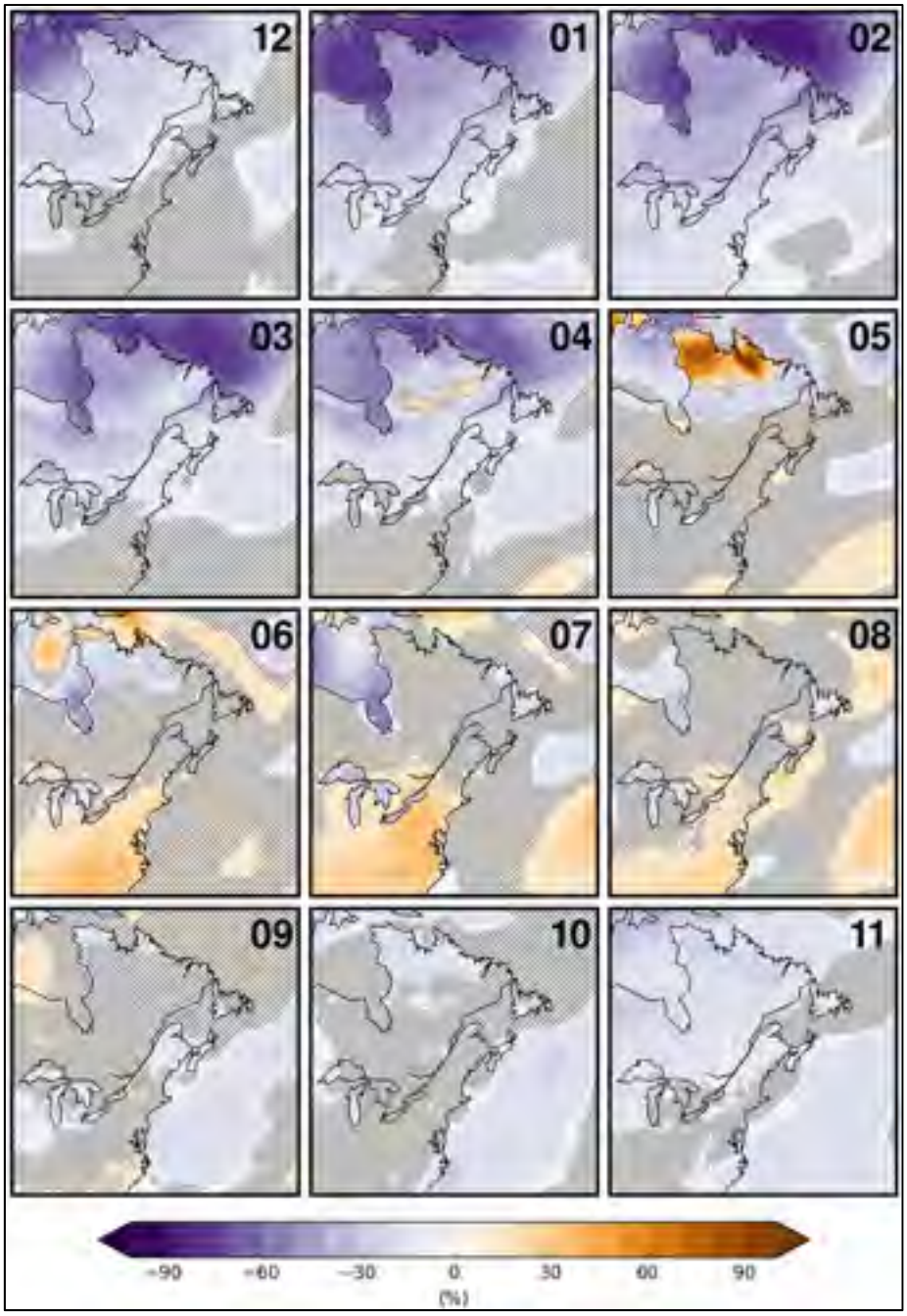

Figure-A I-17 Relative change in interannual variability for the monthly mean surface-air temperature during 2080-2099 relative to 2000-2019 over the NNA domain. Hatched regions identify where the change in variability is not statistically significant at the 99% confidence level (F-test)

I.3.4 Projected changes in precipitation extremes

The CRCM5-LE gives an opportunity to investigate the impact of climate change on the evolution of very extreme precipitation events. For instance, quantiles as high as the 100-year return level can be estimated empirically. For the reference period (2000-2019), 24-hour annual maxima precipitation series were constructed at each grid point (using all 50 members) and the 99^{th} percentiles precipitation (100-year return level) were estimated.

The 100-year return 24-hour annual maximum precipitation in the reference period is shown in Figure-A I-18-a for the EU domain. Extreme values exceeding 90 mm/24h are generally obtained for the most part of Europe, while values higher than 120 mm/24h are found over mountainous regions and the Mediterranean Sea. In particular, regions with high topography such as the Alps show very high localized values around 300 mm/24h. For the NNA domain (Figure-A I-18-b), values are larger than 60 mm/24h in the northern regions, while generally increasing southward to exceed 240 mm/24h near the southern boundary and over the east coast. Very large maximum values exceeding 300 mm/day are also found over the Atlantic.

Return period in future climate (2080-2099) of events with the same intensities as the 100-year return level in the reference period were estimated. These return periods are shown in Figure-A I-18-c and d for the EU and NNA domains, respectively. For the EU domain, the 100-year return period events often become approximately 50-year return events in the future period, while in northern regions and the northern slope of the Alps it may become a 20-year return period event or less. For the NNA domain, the future return period is in general smaller than 30 years. These preliminary results show that extreme precipitation (100-year event) will experience significant increase and that, for some sub-regions, these rare events could be much more frequent.

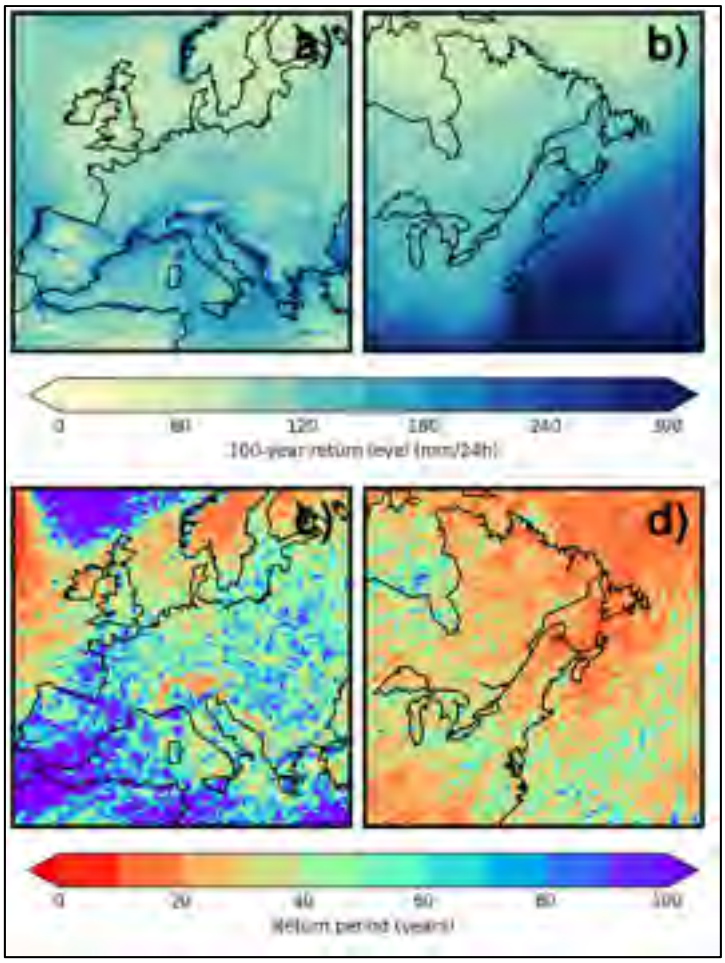

Figure-A I-18 The 100-year return period values of the 24-hour annual maximum precipitation for the reference period (2000-2019) for (a) EU, and (b) NNA domains, respectively. Corresponding return period of events with same intensities for the future period (2080-2099) are presented in (c) and (d)

I.4 Discussion and conclusions

The series of extreme flood events that occurred in Bavaria and Québec in recent decades has been of great concern to local governments, and has led to the ClimEx project, which builds on the longstanding collaboration between Bavaria and Québec. The main goal of ClimEx is to help decision makers to implement robust climate-change adaptation strategies regarding

flood risk, and more particularly, to better understand the role of natural climate variability and extreme meteorological events in the quantification of risk. The project is structured as a hydro-modelling chain: a Global Climate Model (GCM) large ensemble is dynamically downscaled with a Regional Climate Model (RCM), whose outputs will serve as input to hydrological model simulations over Bavaria and Québec. More specifically, the CRCM5 Large Ensemble (CRCM5-LE) consists in the dynamically downscaled version of the CanESM2 large ensemble from 2.8° (≈310 km) to 0.11° (≈12 km) resolution using the CRCM5. Simulations cover the 1950-2100 period using historical forcings from explosive volcanoes and solar cycle by the end of 2005, and following the RCP8.5 emission scenario thereafter. The 50 CanESM2 members that are used to drive the CRCM5 at its boundaries differ by small perturbations in their atmospheric initial conditions on 1 January 1950. RCM downscaling was applied over two regions of interest, namely Europe (EU) and northeastern North America (NNA).

This paper aimed to introduce the CRCM5 Large Ensemble (CRCM5-LE) to the scientific community as the complete dataset will become publicly available (see www.climex-project.org). The main climate features of the CRCM5-LE were analyzed over both the EU and NNA domains. The analysis started with a validation of the simulations over the historical period (2000-2019). Climate projections of monthly means were then analyzed for both surface-air temperature and precipitation, followed by an analysis of the projected changes in temperature interannual variability. The paper ended with an analysis of extreme daily precipitation.

In the first part of the paper, a climatological validation of monthly mean surface-air temperature and precipitation was performed based on the E-OBS and CRU observational gridded datasets. As for other climate models, CRCM5 reproduces the historical climate with some biases. RCM biases depend mainly on two sources, namely the RCM model itself (e.g., domain configuration, parameterization packages, land-surface scheme) and the type of boundary conditions (e.g., GCM or reanalysis). To separate these two sources, the biases of

the CanESM2-driven CRCM5 were compared with ERA-Interim driven simulations, the latter providing some indications of the biases that are intrinsic to the CRCM5, while the former combined CRCM5 bias and those introduced by GCM. For the analyzed variables, it was shown that CanESM2 generally contributed to increase the CRCM5 biases compared to the reanalysis-driven run, with the exception of a cold bias during winter in Europe. This validation should provide guidance to future users of the large ensemble to select the most suitable bias-correction methods when using the CRCM5-LE as an input for impact models (e.g. Muerth et al., 2012a).

Climate-change projections of the monthly mean climate were next analyzed. Since climate-change projections can be difficult to assess when natural climate variability is large relative to the anthropogenic-induced signal, a sensitivity analysis was performed by comparing the full 50-member ensemble with a five-member ensemble, as well as by considering three time horizons for the projections: short term (2020-2039), mid-term (2040-2059 and long term (2080-2099). The value of the large ensemble was highlighted by comparing these different cases through the spatial extent of the statistically significant climate-change signal. As expected, the highest statistical significance was obtained using the full ensemble, as well as for long-term projections for the temperature field. The value of the large ensemble was much more obvious for short-term climate projections of precipitation, but this was also true (although to a lesser extent) for temperature. These results are in agreement with Hawkins et Sutton (2009, 2011), who have shown that for short-term climate-change projections, natural climate variability is often the largest contributor to the total uncertainty in comparison with the choice of the emissions scenario and model climate sensitivity.

Following the analysis of projected changes for the monthly means, the projected changes in interannual variability of monthly mean surface-air temperature were investigated. While the projected patterns of change in natural variability generally agreed with other studies using GCMs (e.g. Holmes et al., 2016), this ensemble allowed to obtain a more detailed characterization of the regional variability compared to its predecessors given the higher

spatial resolution, but also due to the large number of members, thus allowing to detect small changes when the noise is important. This analysis also recalled the importance of considering both the climate trends in both mean and variability when analyzing the effect of climate change on extreme events such as heat waves.

Finally, an example of analysis of extreme events was performed based on the 100-year return level of 24-hour annual maximum precipitation. An advantage of using a large ensemble was implicitly shown as the 100-year return levels can be estimated empirically due to the large number of available values, hence bypassing the strong assumptions being made in the context of the extreme value theory framework. This analysis showed that for both domains, the 24-hour 100-year return level of precipitation intensity in the reference period (2000-2019) will correspond to a 50-year or even more frequent event. Particularly, it will correspond to less than a 20-year return period in some regions for the 2080-2099 period. These results are particularly important in the context of hydrological modelling as it is not clear how the 100-year return levels translate into local hydrological extremes, and future work within the scope of the ClimEx project will likely lead to new insights regarding this question. This is even more important as observational datasets cover a limited period, which leads to huge uncertainties in the estimation of the 100-year return levels (Schulz & Bernhardt, 2016).

This paper presented some of the advantages of using a large ensemble, such as the ability to better assess regional climate changes on a monthly basis regarding means, interannual variability and extremes. However, some caveats should be noted. First, it is worth recalling that this framework does not address the structural uncertainty of climate-change projections, since it uses only a single combination of global (CanESM2) and regional climate models (CRCM5), along with a single future pathway of GHGA emissions (RCP8.5). The CRCM5-LE thus samples the internal variability of the CanESM2 model, which was downscaled at the regional scale using the CRCM5 that also adds its own internal variability (although in general much smaller than that of the GCM). Despite not spanning the current range of

structural uncertainty, the natural climate variability associated with a high-resolution regional climate system (here defined as the CanESM2/CRCM5 combination) was assessed at a degree of detail never reached before.

Another important caveat, common to most RCM experiments, relates to the generation of small scales within the regional domain from the coarse-resolution boundaries. This intrinsic characteristic of the one-way nesting procedure is known as the "spatial spin-up" (Leduc & Laprise, 2009; Leduc et al., 2011; Matte et al., 2017) of small scales. In the current experiment, given the relatively large gap of spatial resolution between CanESM2 and CRCM5, the downscaling challenge was important and small-scale features may not be fully developed inside the domain (particularly near the western inflow lateral boundary). To address this issue, a security zone of 30 grid points surrounding the free domain was removed (in addition to the Davies' blending zone and the halo), leaving a validation area that is much less prone to undesirable boundary effects.

To facilitate future cross-disciplinary research collaborations based on this dataset, about 50 variables were archived, covering most facets of the regional climate. The time resolution of the archives was selected depending on the variable, and to balance disk-space constraints (more than 500 Terabytes of archived data), partners' priorities (e.g., hourly archived precipitation) and potential uses in future projects. Also, performing simulations over two domains (northeastern North America and Europe) will open new collaboration opportunities for researchers in the field of extreme impacts from climate change. Further applications of the CRCM5-LE include the characterization of extreme compound events such as heat waves, forest fires, floods and droughts. This new ensemble may also serve as a benchmark to test innovative techniques involving machine-learning algorithms that link meteorological patterns to high-impact events, among others.

Finally, it is important to note that the CanESM2-LE was the only candidate allowing the CRCM5 to be driven by 50 continuous climate simulations from 1950 to 2100, while the

CESM large ensemble (Kay et al., 2015) also provides output to drive an RCM but for a limited number of 10-year periods. Given the small number of existing GCM large ensembles, the concept of an RCM large ensemble is slowly emerging despite the need for significant computational resources and that the necessary driving fields from a GCM large ensemble must be available. The authors stress that more GCM large-ensemble experiments should archive the necessary fields to drive RCM models, allowing the production of other high-resolution large ensembles in the future, which will help to better cope with the uncertainty related to future GHGA emissions, climate sensitivity (i.e. structural uncertainty) and natural variability within a common framework, at a spatial scale that is suitable for climate-change impact research.

I.5 Acknowledgments

The authors would first like to thank Michel Giguère (Ouranos) and Michel Valin for their participation to the run-production phase; Gilbert Brietzke (LRZ) and Mourad Labassi (Ouranos) for their technical support with the SuperMUC and Ouranos computational infrastructures; René Laprise (UQAM) for scientific discussions; and James Anstey (Environment and Climate Change Canada) for reviewing the first draft of the paper. We would also like to thank Laxmi Sushama and Katja Winger (UQAM) for supporting the development of the CRCM5. The ClimEx project is a new achievement for the Quebec-Bavaria International Collaboration on Climate Change (QBic3), which would have been impossible without the contribution of several important actors, notably Dieter Kranzlmüller (LRZ), Diane Chaumont (Ouranos) and Simon Ricard (Quebec Environment Ministry).

The ClimEx project is funded by the Bavarian State Ministry for the Environment and Consumer Protection. The CRCM5 is developed by the ESCER centre of Université du Québec à Montréal (UQAM; www.escer.uqam.ca) in collaboration with Environment and Climate Change Canada. We acknowledge Environment and Climate Change Canada's Canadian Centre for Climate Modelling and Analysis for executing and making available the

CanESM2 Large Ensemble simulations used in this study, and the Canadian Sea Ice and Snow Evolution Network for proposing the simulations. Computations with the CRCM5 for the ClimEx project were made on the SuperMUC supercomputer at Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities. The operation of this supercomputer is funded via the Gauss Centre for Supercomputing (GCS) by the German Federal Ministry of Education and Research and the Bavarian State Ministry of Education, Science and the Arts.

I.6 Supplementary materials

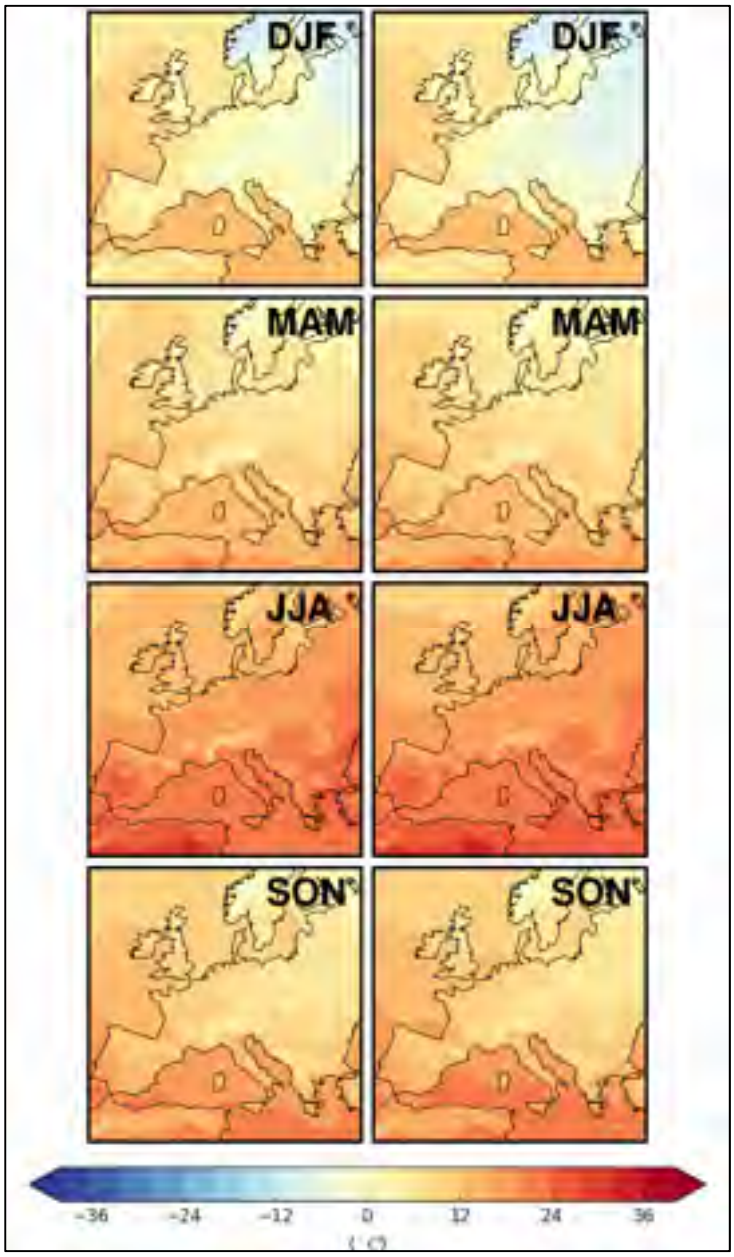

Figure-A I-S19 (left column) Seasonal mean surface-air temperature during the 1980-2012 period as simulated by the CRCM5 driven by CanESM2 (first member) over the EU domain; (right column) same for the ERA-Interim driven CRCM5 run

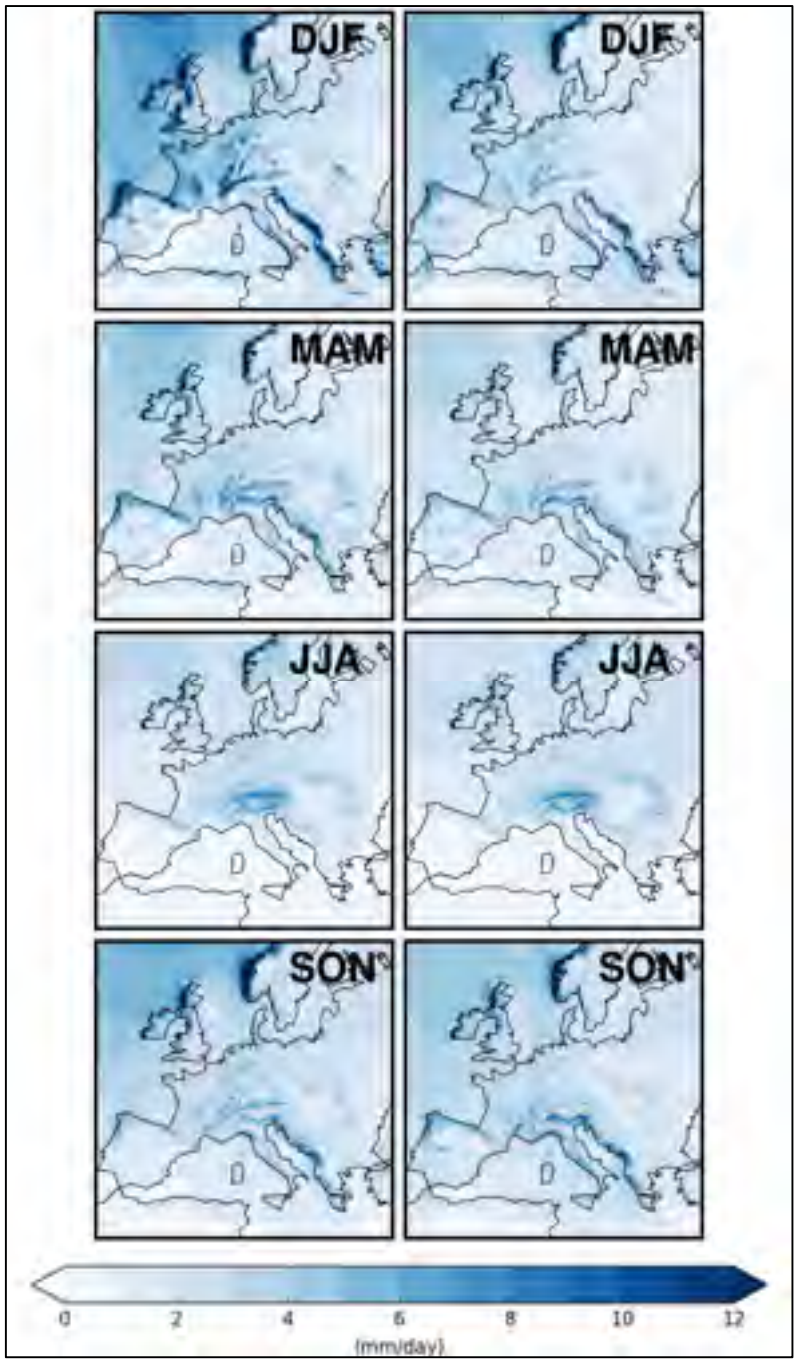

Figure-A I-S20 (left column) Seasonal mean precipitation during the 1980-2012 period as simulated by the CRCM5 driven by CanESM2 (first member) over the EU domain; (right column) same for the ERA-Interim driven CRCM5 run

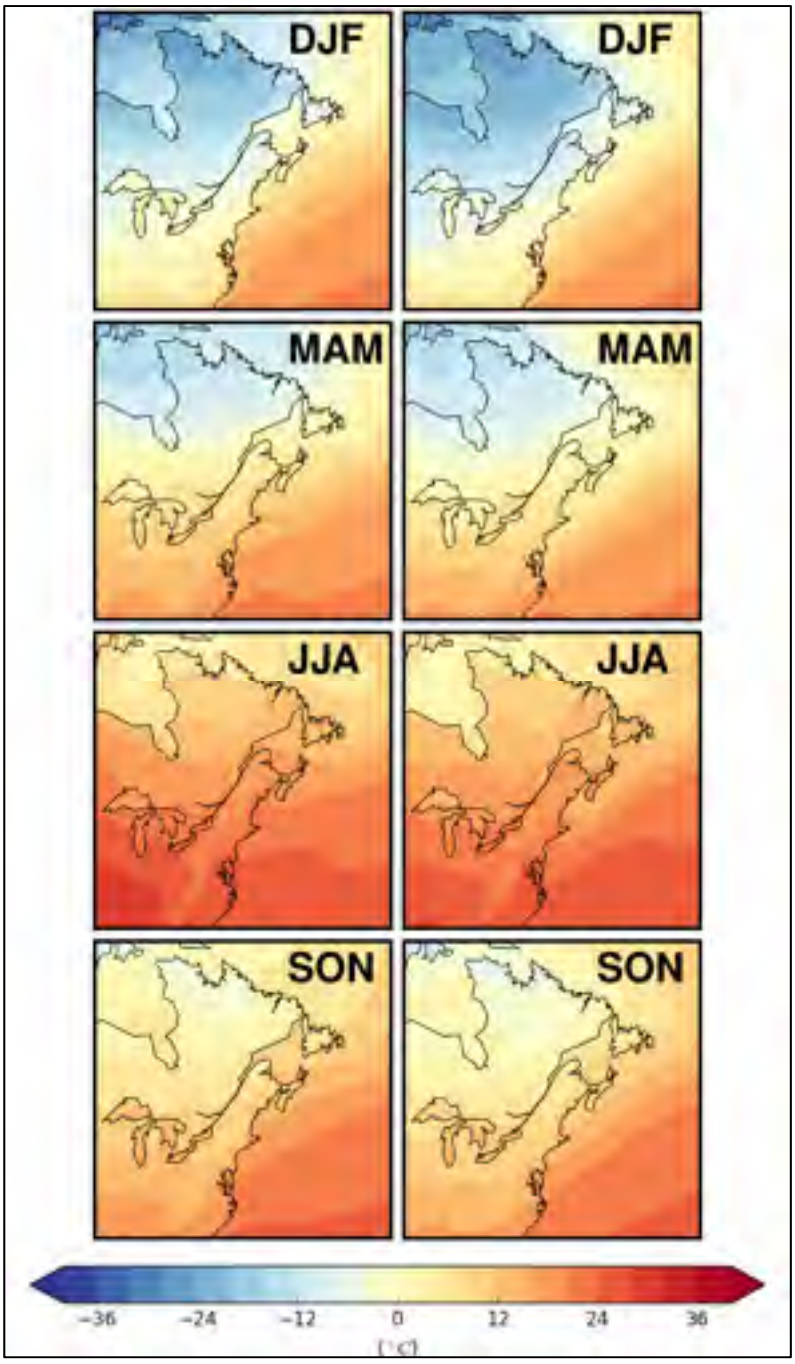

Figure-A I-S21 (left column) Seasonal mean surface-air temperature during the 1980-2012 period as simulated by the CRCM5 driven by CanESM2 (first member) over the NNA domain; (right column) same for the ERA-Interim driven CRCM5 run

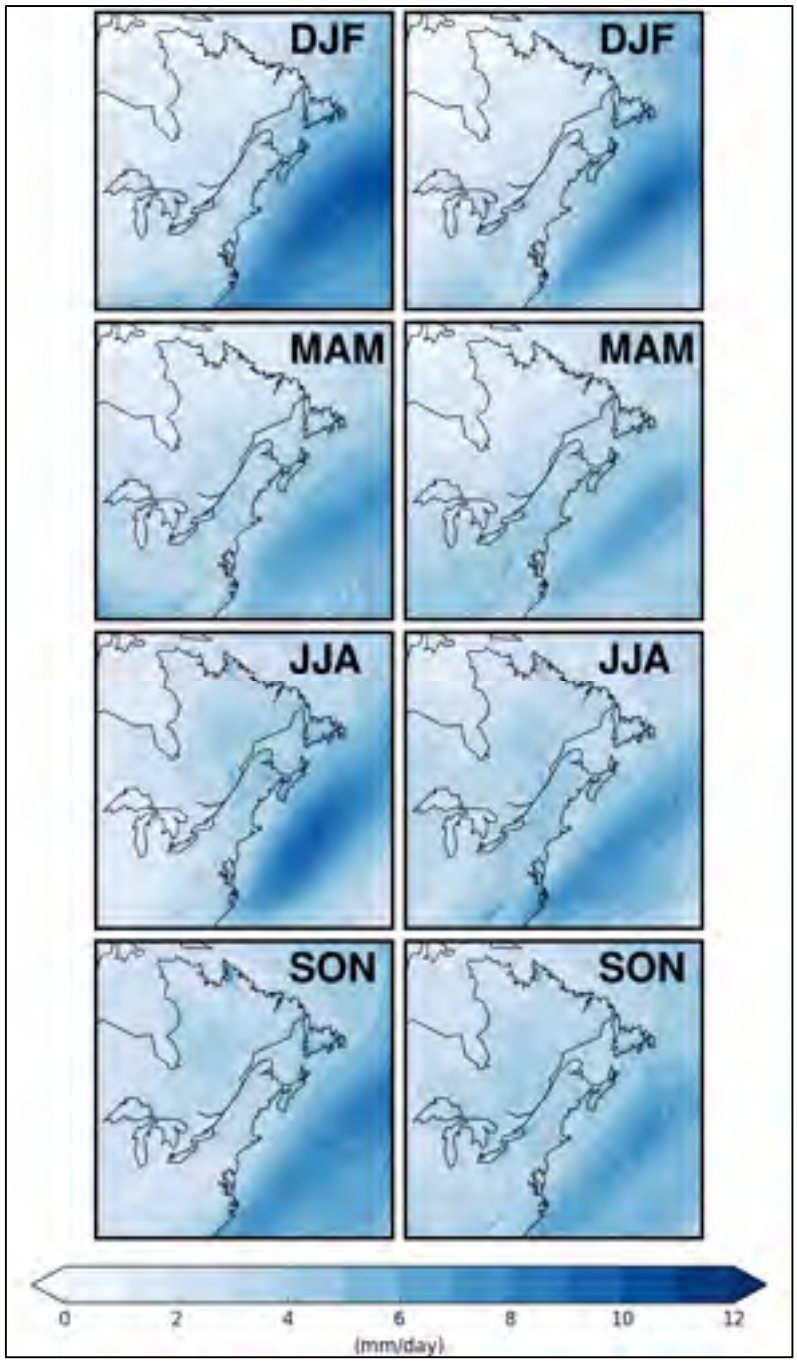

Figure-A I-S22 (left column) Seasonal mean precipitation during the 1980-2012 period as simulated by the CRCM5 driven by CanESM2 (first member) over the NNA domain; (right column) same for the ERA-Interim driven CRCM5 run

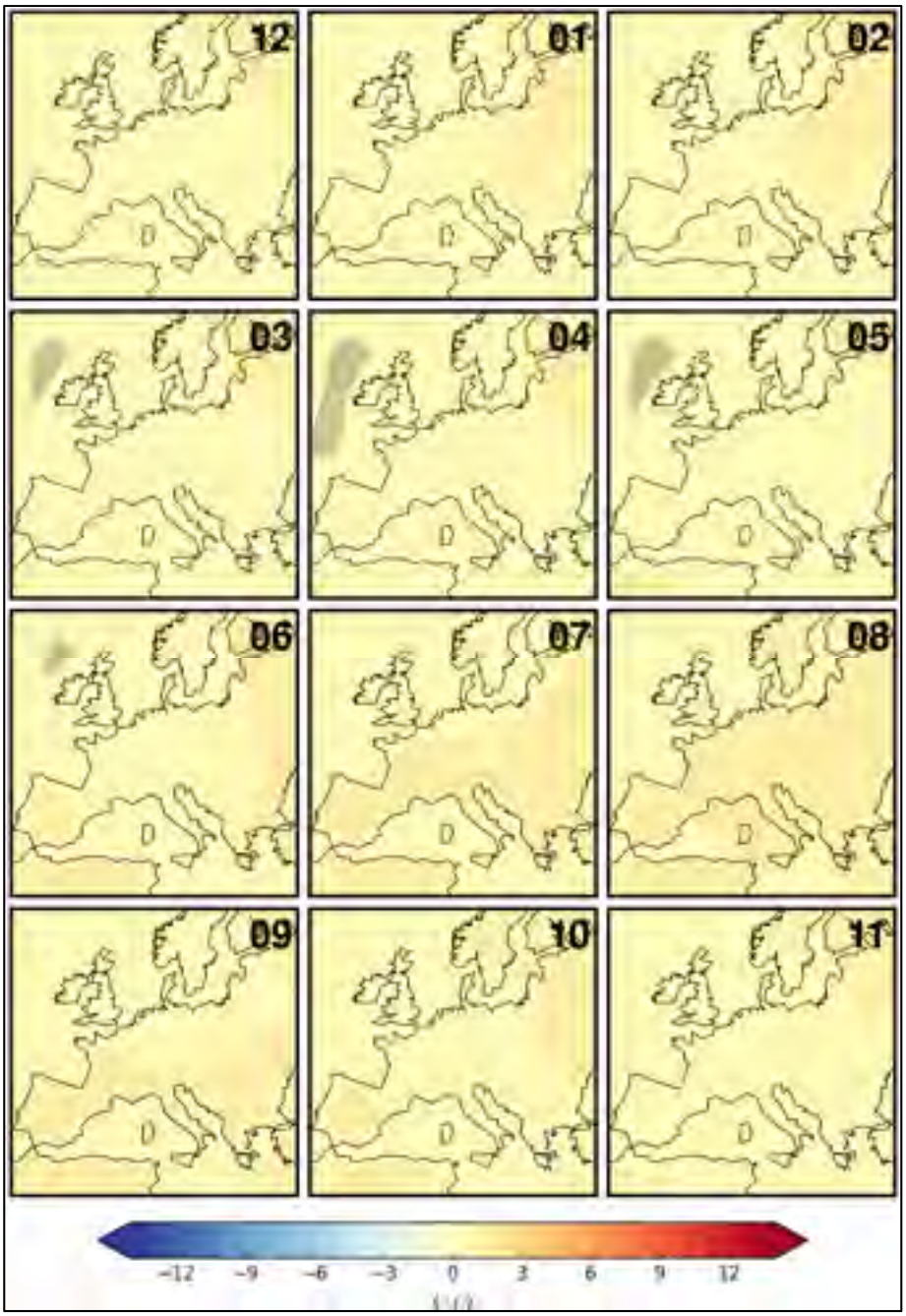

Figure-A I-S23 The CRCM5 50-member ensemble mean climate-change signal for surface-air temperature computed as the difference between the 2020-2039 and 2000-2019 monthly climate means for the EU domain. Hatched region identify regions where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

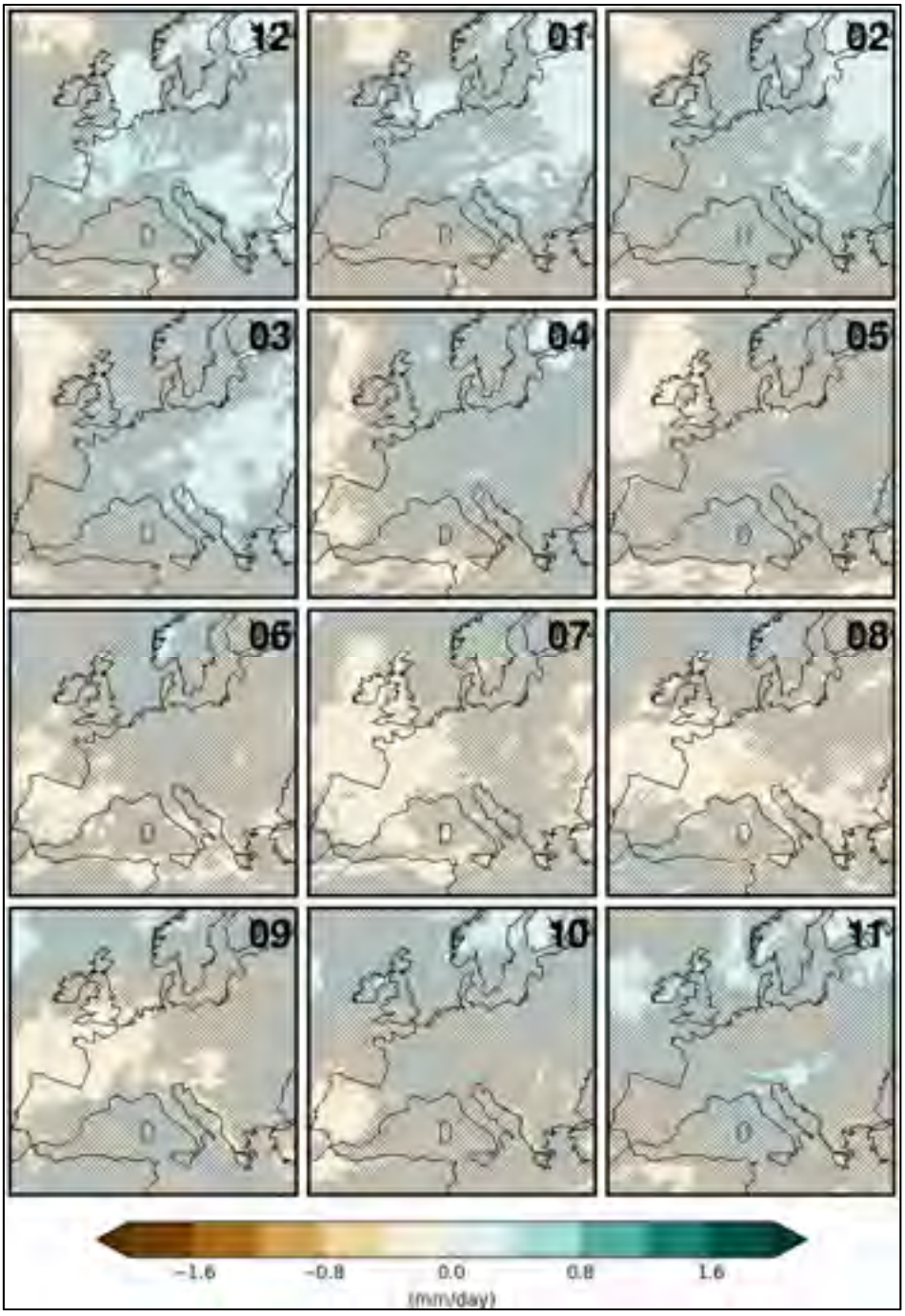

Figure-A I-S24 Same as Figure-A I-S23 for precipitation during 2020-2039 over the EU domain (50 members)

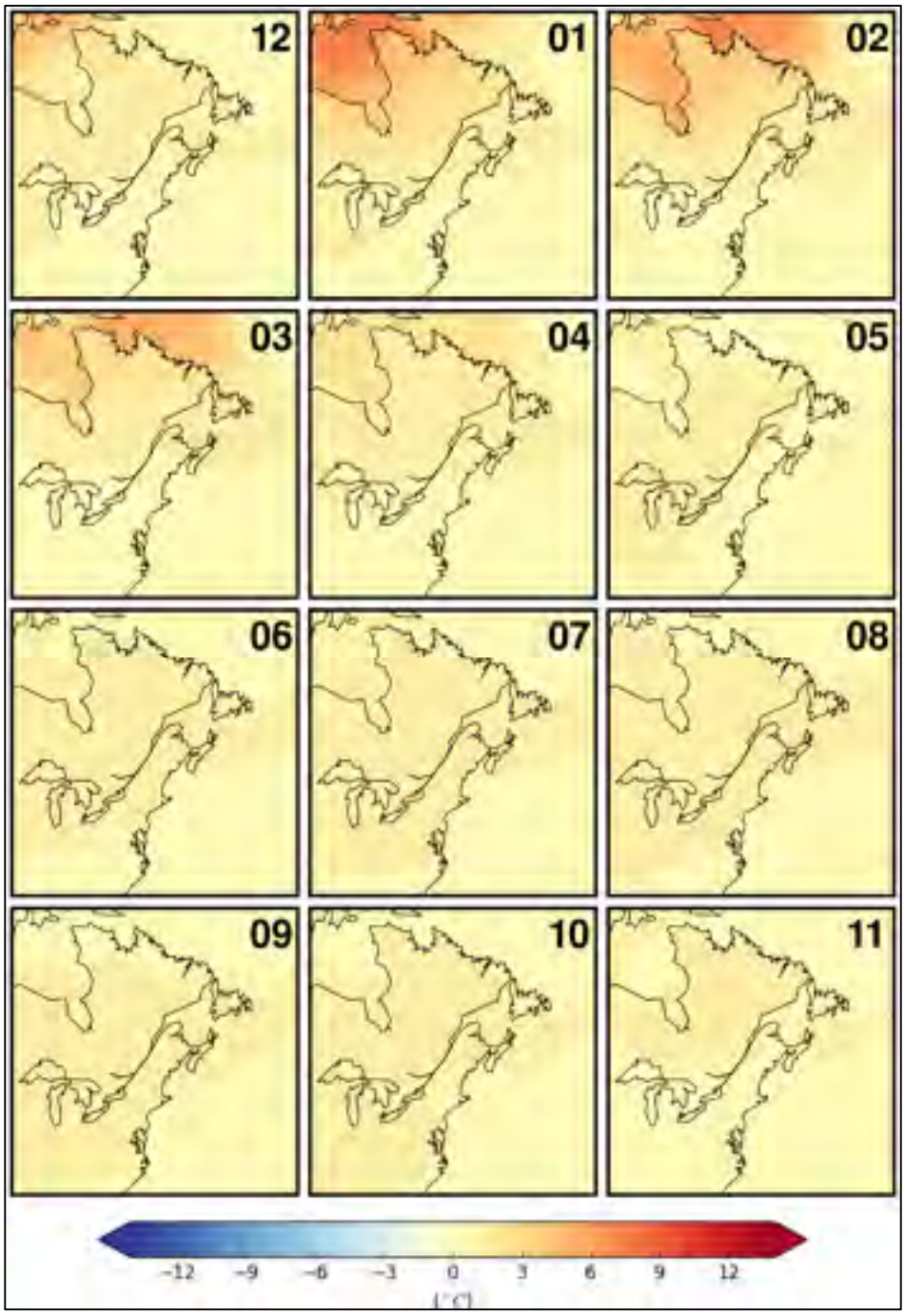

Figure-A I-S25 Same as Figure-A I-S23 for surface-air temperature during 2020-2039 over the NNA domain (50 members)

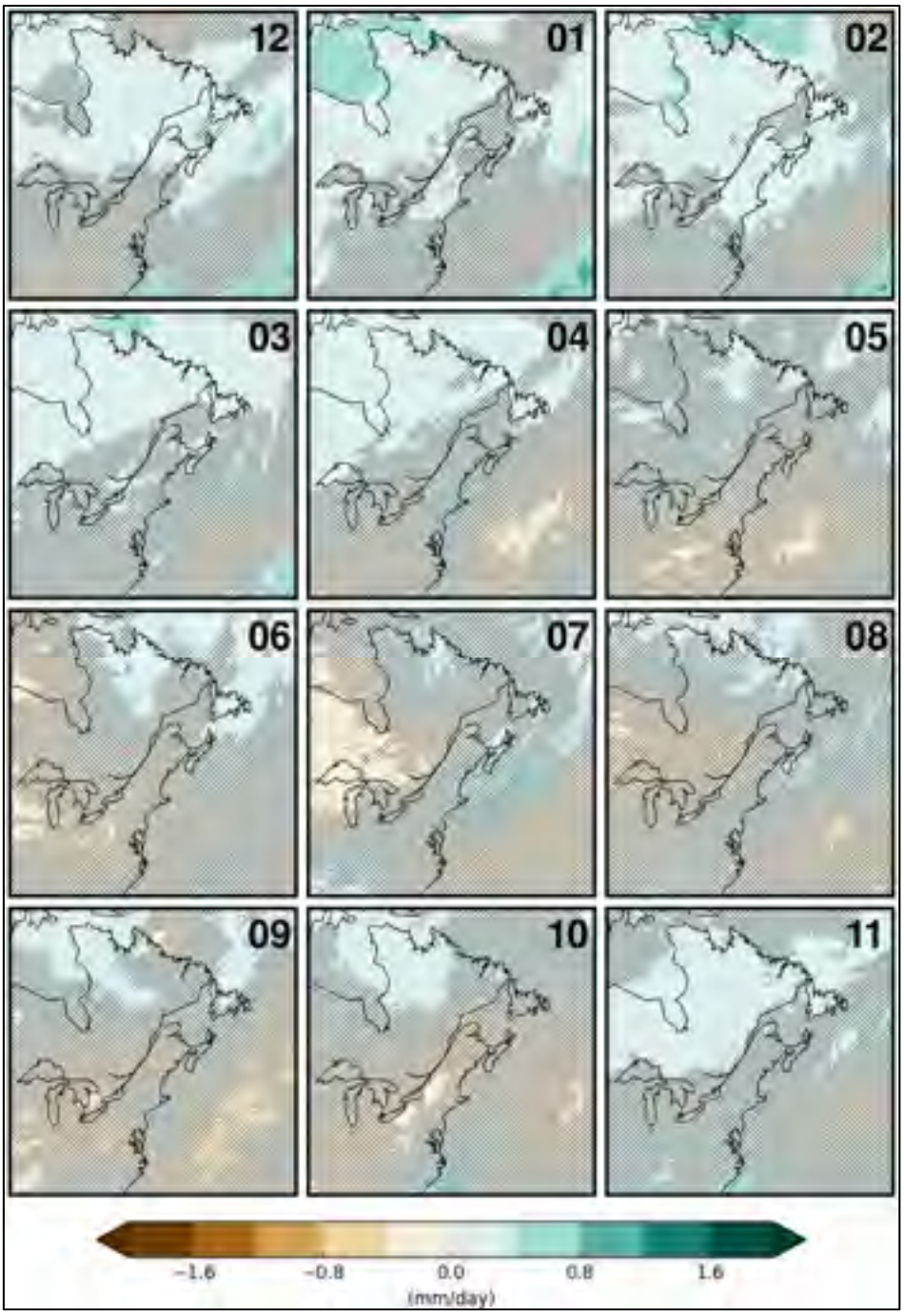

Figure-A I-S26 Same as Figure-A I-S23 for precipitation during 2020-2039 over the NNA domain (50 members)

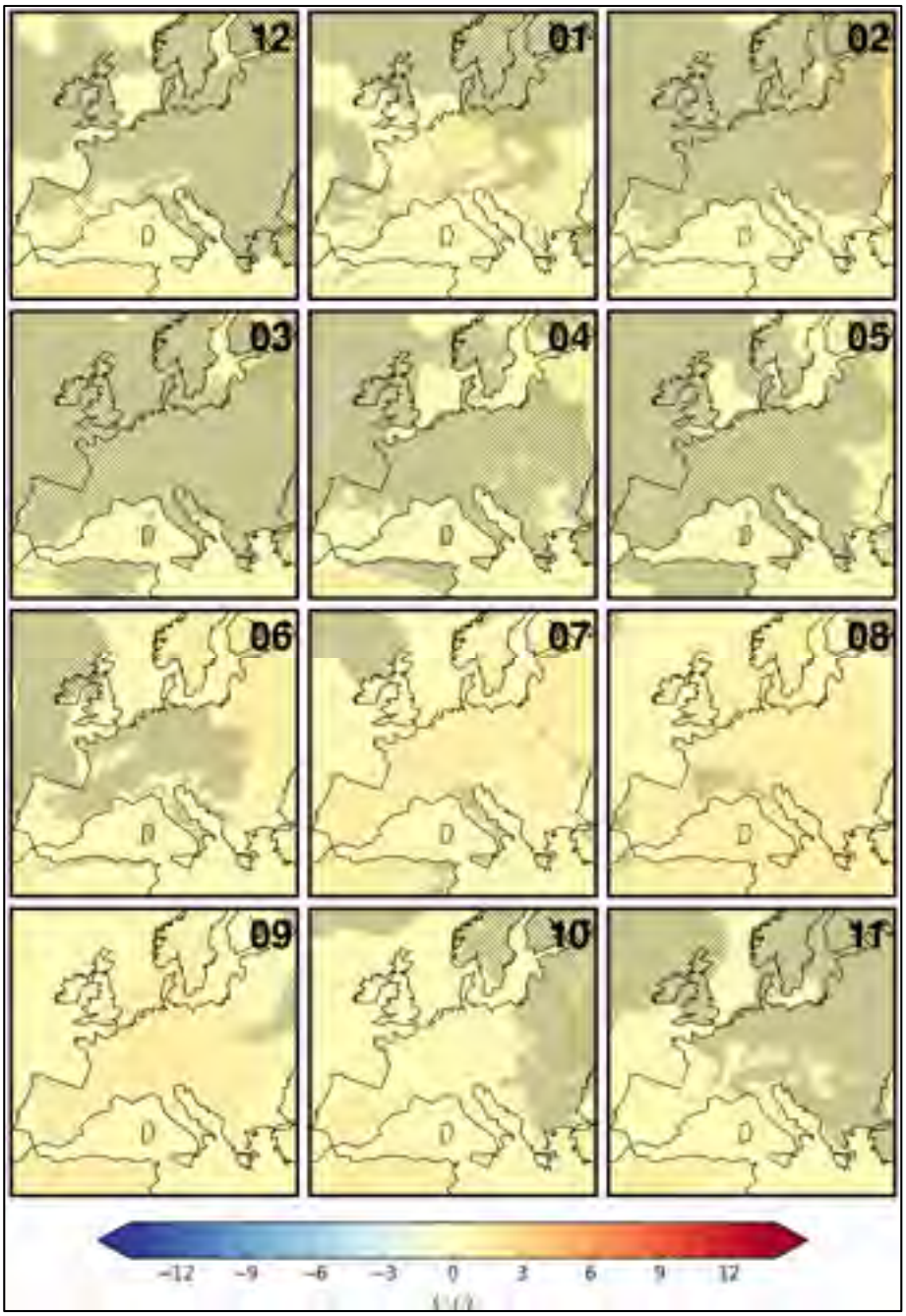

Figure-A I-S27 The CRCM5 five-member ensemble mean climate-change signal for surface-air temperature computed as the difference between the 2020- 2039 and 2000-2019 monthly climate means for the EU domain. Hatched regions identify where the signal is not statistically significant at the 99% confidence level (Student's t-test with unequal variances)

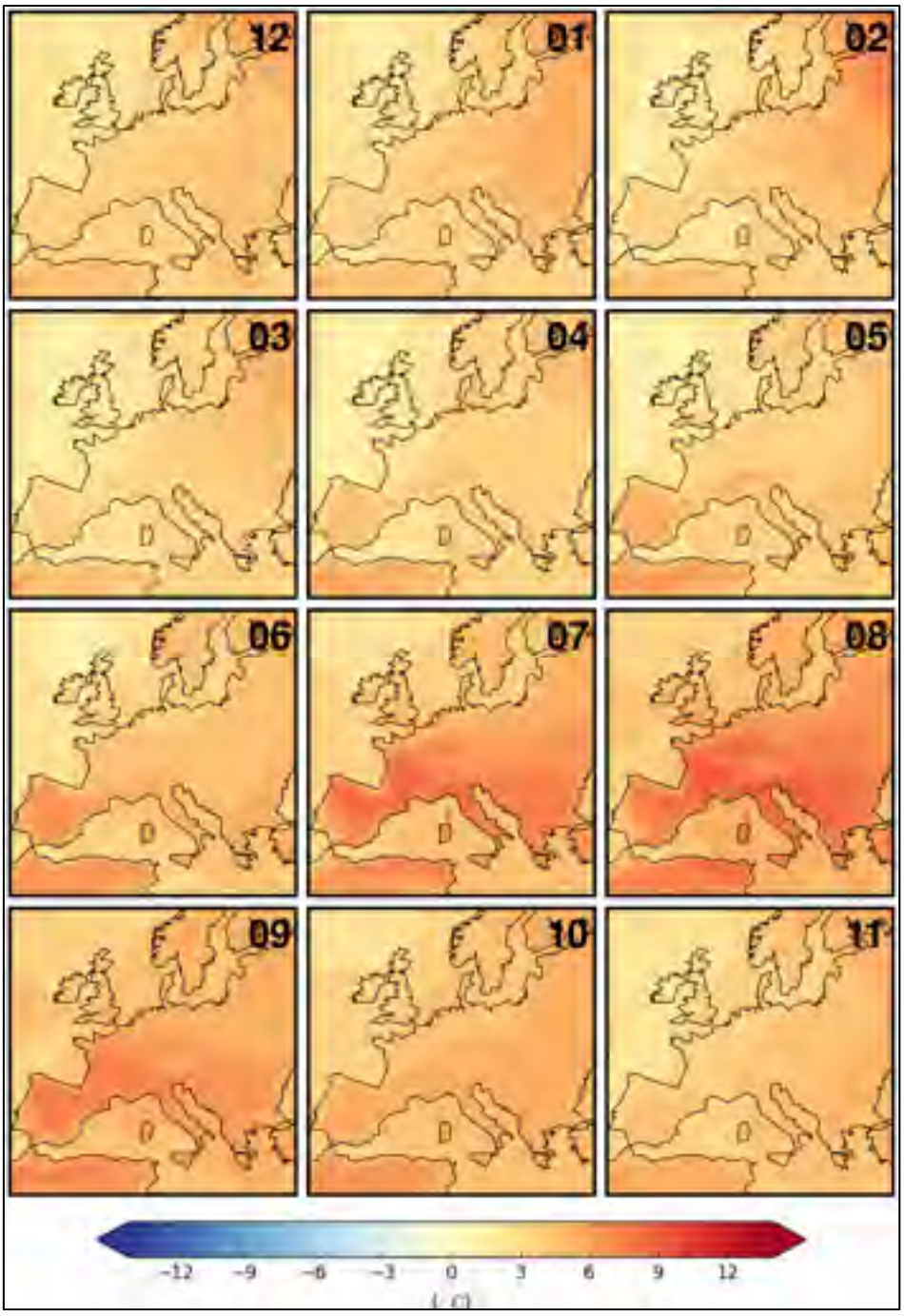

Figure-A I-S28 Same as Figure-A I-S27 for surface-air temperature during 2080-2099 over the EU domain (five members)

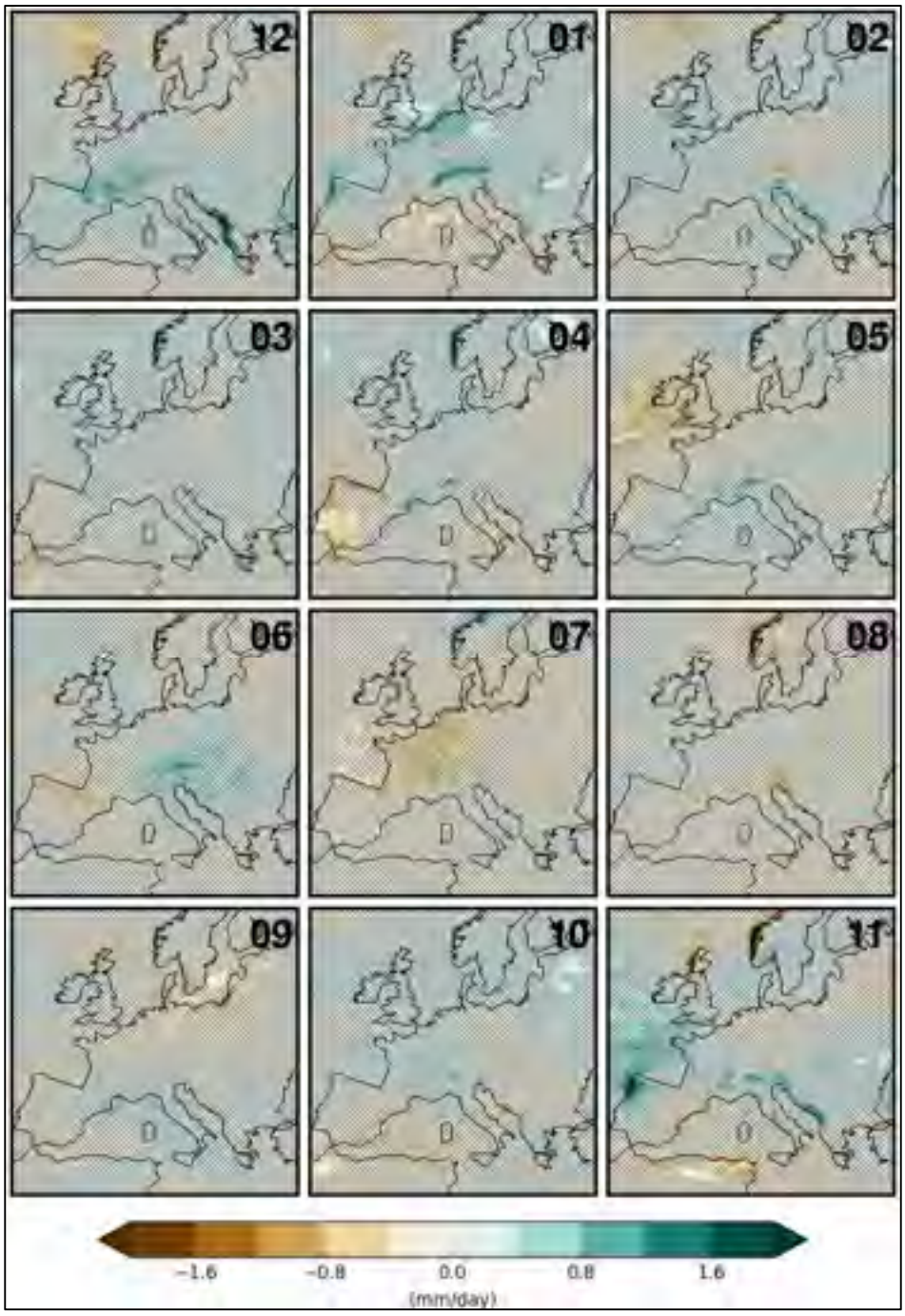

Figure-A I-S29 Same as Figure-A I-S27 for precipitation during 2020-2039 over the EU domain (five members)

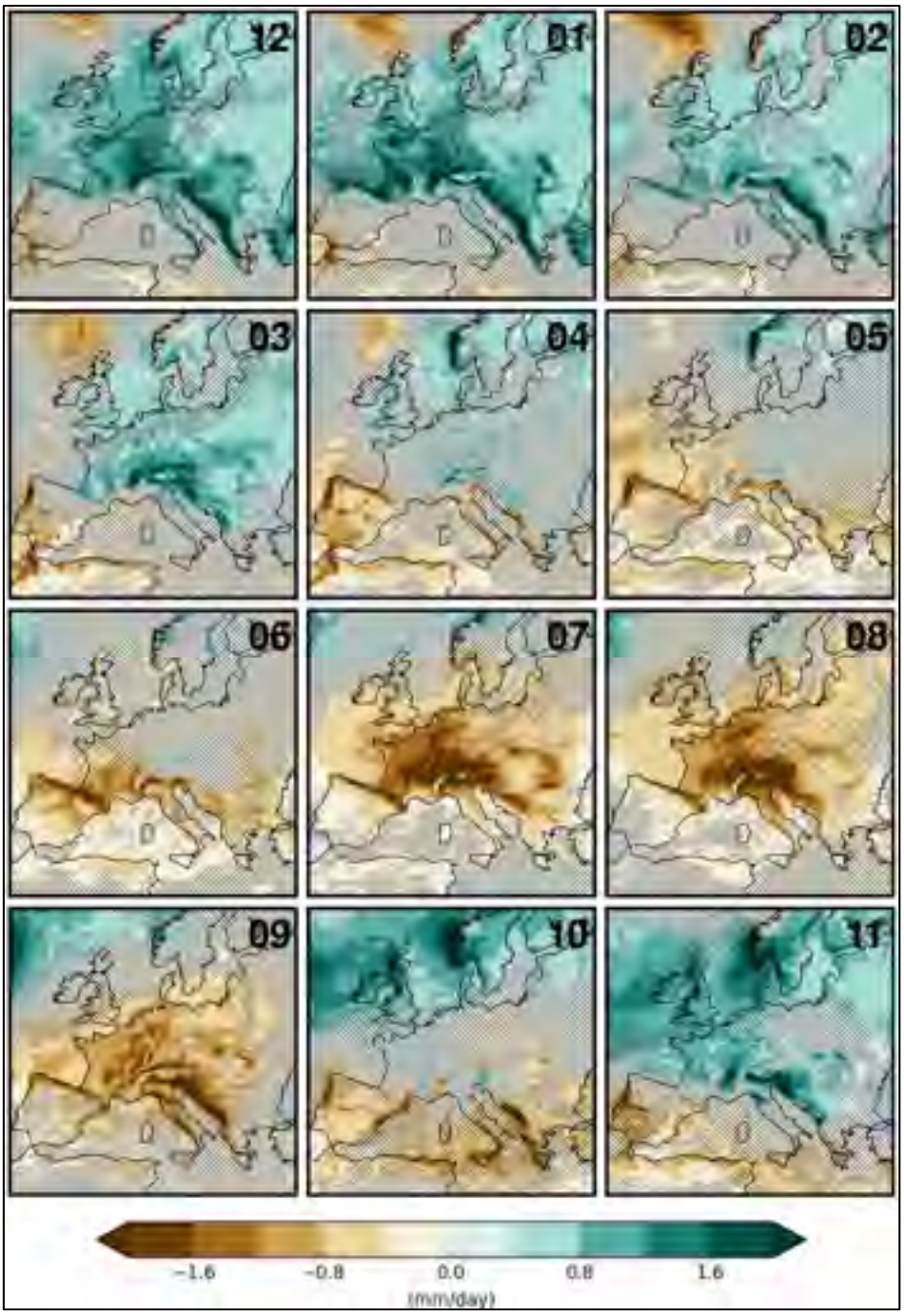

Figure-A I-S30 Same as Figure-A I-S27 for precipitation during 2080-2099 over the EU domain (five members)

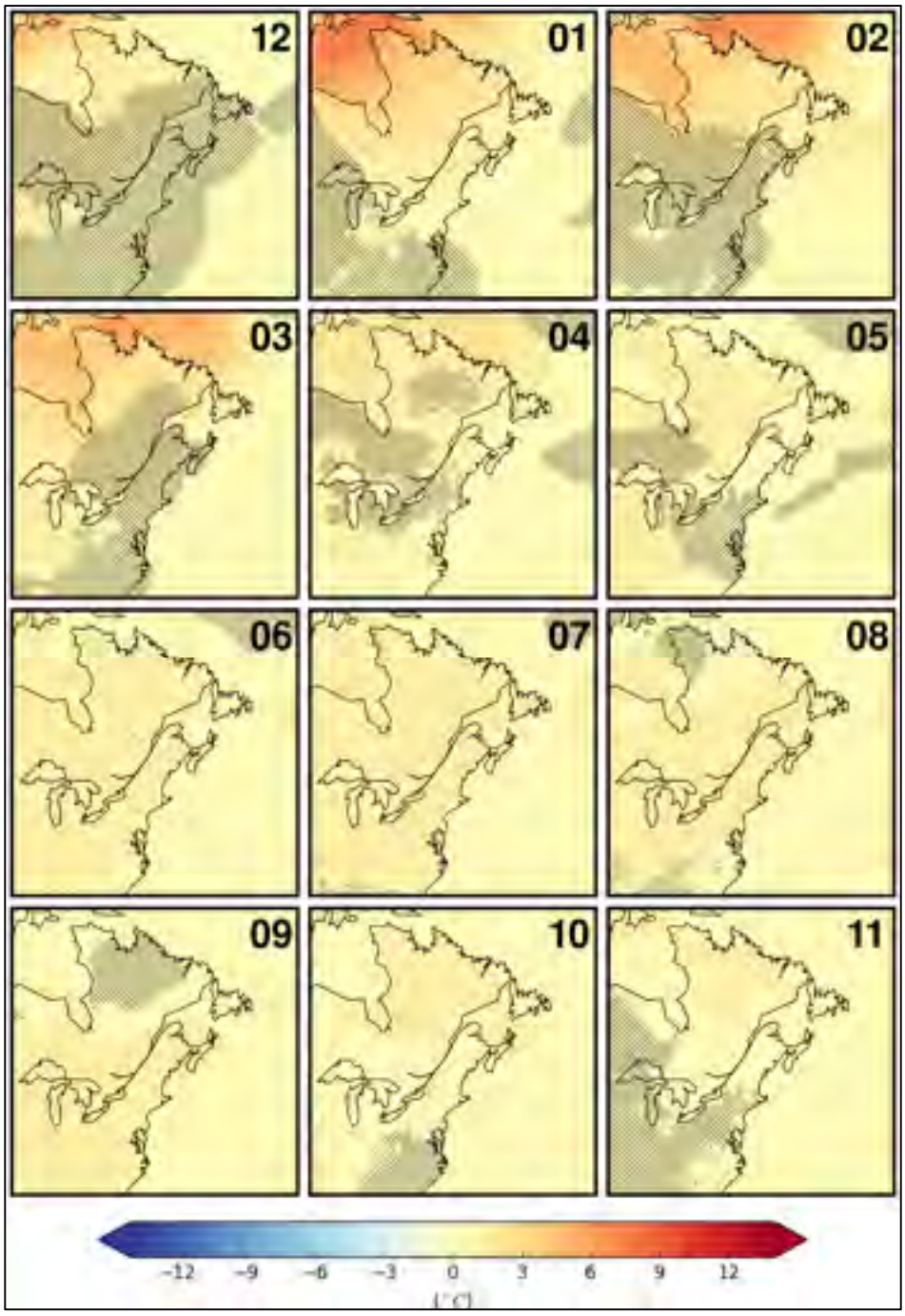

Figure-A I-S31 Same as Figure-A I-S27 for surface-air temperature during 2020-2039 over the NNA domain (five members)

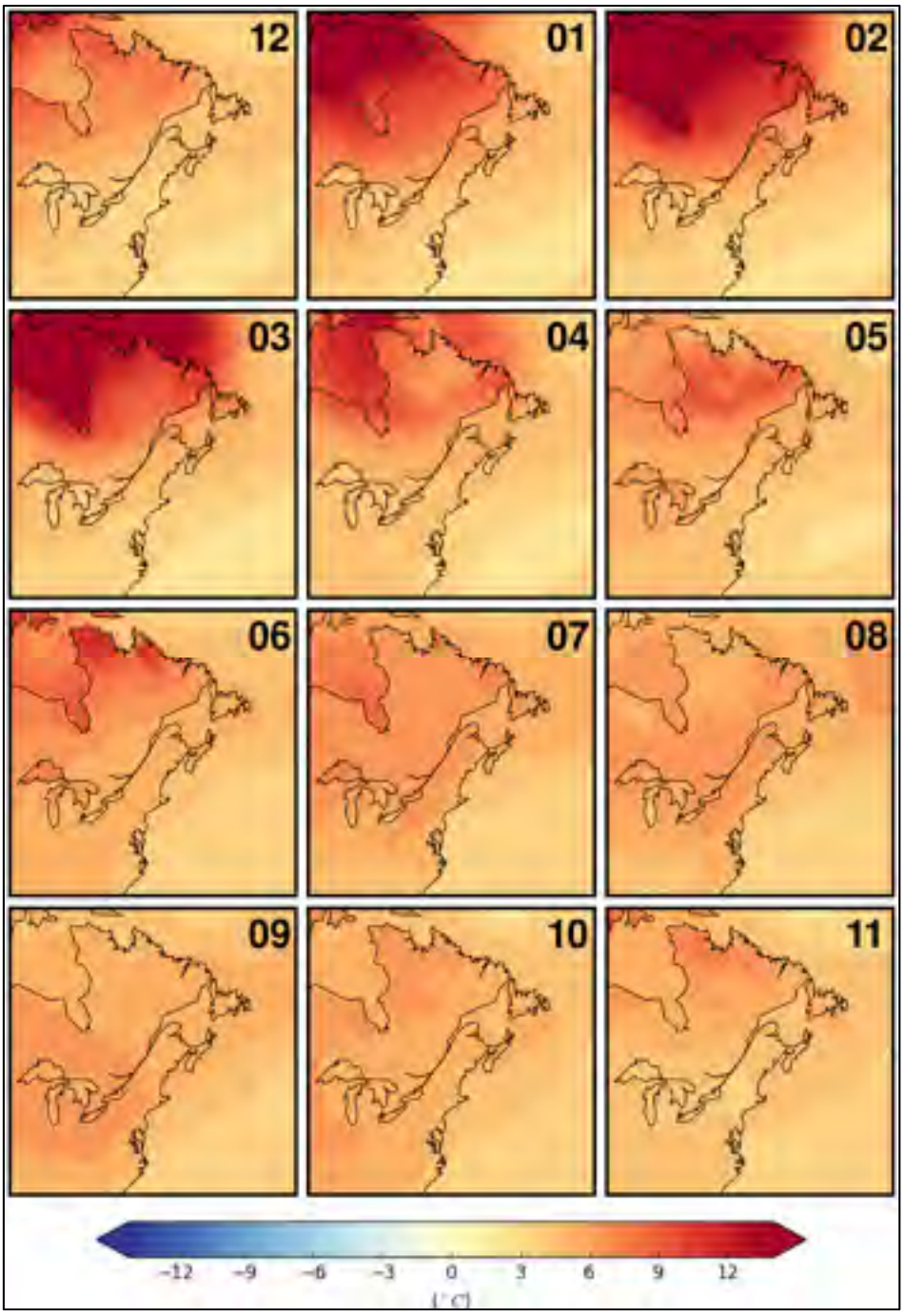

Figure-A I-S32 Same as Figure-A I-S27 for surface-air temperature during 2080-2099 over the NNA domain (five members)

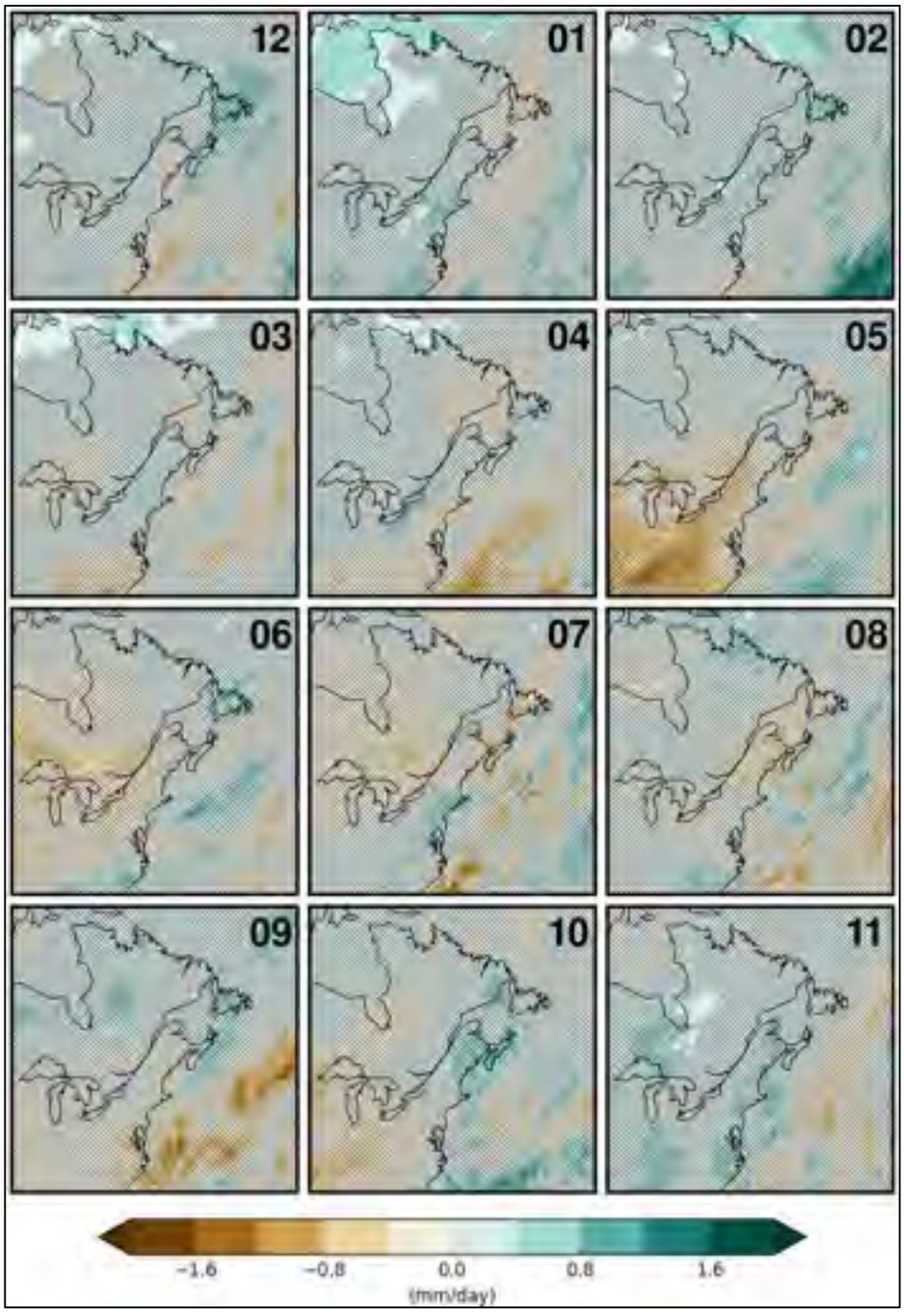

Figure-A I-S33 Same as Figure-A I-S27 for precipitation during 2020-2039 over the NNA domain (five members)

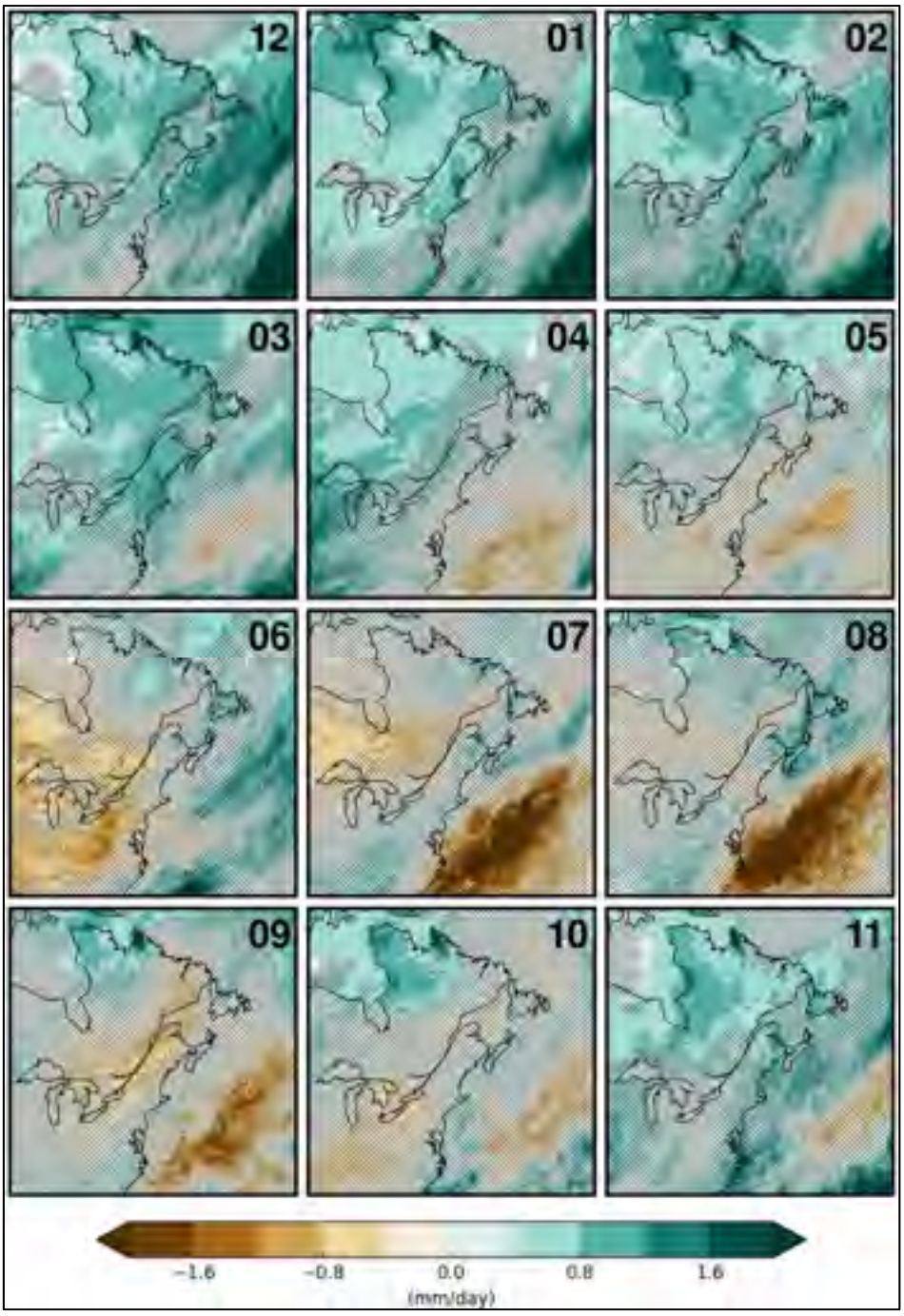

Figure-A I-S34 Same as Figure-A I-S27 for precipitation during 2080-2099 over the NNA domain (five members)

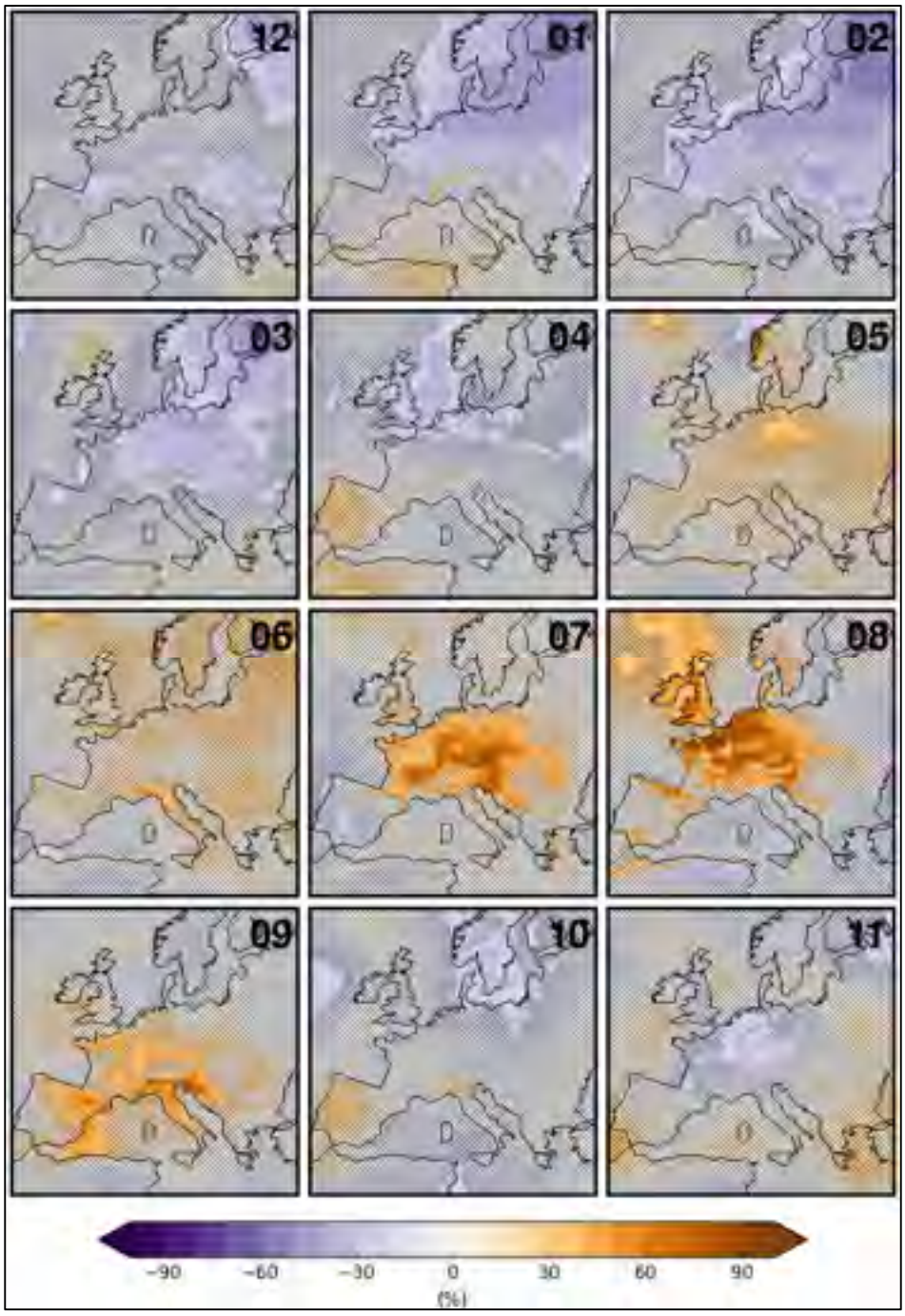

Figure-A I-S35 Relative change in interannual variability for the monthly mean surface-air temperature during 2080-2099 relative to 2000-2019 over the EU domain. Hatched regions identify where the change in variability is not statistically significant at the 99% confidence level (F-test) (five members)

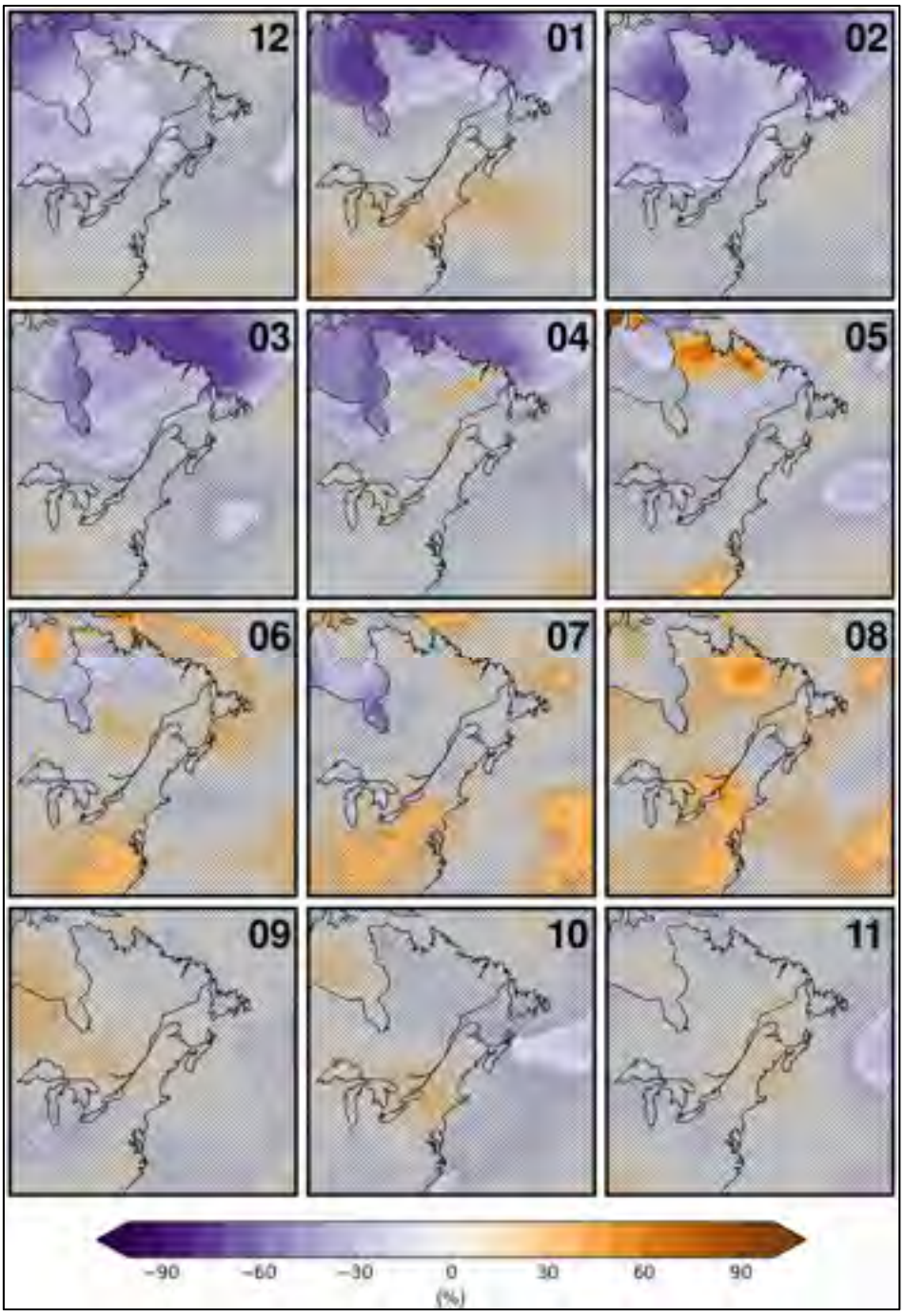

Figure-A I-S36 Same as Figure-A I-S35 over the NNA domain (five members)

ANNEXE II

UNCERTAINTY OF HYDROLOGICAL MODEL COMPONENTS IN CLIMATE CHANGE STUDIES OVER TWO NORDIC QUEBEC CATCHMENTS

Magali Troin[a], Richard Arsenault[b], Jean-Luc Martel[c], François Brissette[d]

[a, b, c, d] Département de Génie de la construction, École de technologie supérieure

Article publié dans « Journal of Hydrometeorology », octobre 2017

Capsule:

Various combinations of snow models (SM) – potential evapotranspiration (PET) methods – hydrological model (HM) structures are evaluated to simulate discharges for the current and future climates.

Abstract:

Projected climate change effects on hydrology are investigated for the 2041-2060 horizon under the A2 emission scenarios using a multi-model approach over two snowmelt-dominated catchments in Canada. An ensemble of 105 members was obtained by combining seven snow models (SM), five potential evapotranspiration (PET) and three hydrological model (HM) structures. The study was performed using high-resolution simulations from the Canadian Regional Climate Model (CRCM-15 km) driven by two members of the Canadian Global Climate Model (CGCM3). This study aims to compare various combinations of SM-PET-HM in terms of their ability to simulate streamflows under the current climate, and to evaluate how they affect the assessment of the climate change-induced hydrological impacts at the catchment scale. The variability of streamflow response caused by the use of different SMs (degree-day versus degree-day/energy balance), PET methods (temperature-based versus radiation-based methods) and HM structures is evaluated, as well as the uncertainty due to the natural climate variability (CRCM inter-member variability). The hydroclimatic simulations cover 1961-1990 in the present period and 2041-2060 in the future period. The ensemble spread of the climate change signal on streamflow is large, and varies with

catchments. Using the variance decomposition on three hydrologic indicators, the HM structure was found to make the most substantial contribution to uncertainty, followed by the choice of the PET methods or natural climate variability, depending on the hydrologic indicator and the catchment. Snow models played a minor, almost negligible role in the assessment of the climate change impacts on streamflow for the study catchments.

Keywords: Regional climate modeling; virtual world; snow models; PET methods; hydrological models; natural climate variability.

II.1 Introduction

There is an increasing interest in impact studies for future climate projections and the impacts on hydrological systems in order to identify suitable strategies for adaptation and mitigation. Considerating uncertainty is an essential step in climate change impact studies, since the value of a hydrologic projection can be restricted without an adequate evaluation of associated uncertainty (Georgakakos, Seo, Gupta, Schaake, & Butts, 2004; Wagener & Gupta, 2005). To provide a complete uncertainty analysis on hydrological projections, five sources of uncertainty need to be explored throughout the modeling process: (1) emission scenarios; (2) global climate model (GCM); (3) natural variability of climate; (4) downscaling method; and (5) hydrological model (Wilby & Harris, 2006). Uncertainties of emission scenario, climate model projection (e.g. Bae, Jung, & Lettenmaier, 2011; Giorgi & Mearns, 2002; Hawkins & Sutton, 2009, 2011; Jung & Chang, 2011; Tebaldi, Smith, Nychka, & Mearns, 2005), and downscaling method (e.g. Chen et al., 2011a; Fowler et al., 2007; Troin, Velázquez, Caya, & Brissette, 2015b) have been intensively analysed in recent studies; there are generally accepted as being the most important sources of uncertainty (Chen et al., 2011b; Kay, Davies, Bell, & Jones, 2009; Prudhomme & Davies, 2009a, 2009b). Natural climate variability (McGuffie & Henderson-Sellers, 2005; Peixoto & Oort, 1992) could also have a significant effect on the total uncertainty, especially at the catchment scale (e.g. de Elía et al., 2013; Deser et al., 2012a; Deser et al., 2012b; Fatichi et al., 2014;

Haque, Rahman, Hagare, Kibria, & Karim, 2015; Hawkins & Sutton, 2009, 2011; Seaby, Refsgaard, Sonnenborg, & Højberg, 2015; Thompson et al., 2015; Troin, Poulin, Baraer, & Brissette, 2016). The uncertainty associated with hydrological modeling can also be significant, and should be assessed when quantitative hydrologic projections of climate change impacts are expected – a robust characterization of uncertainty is also a critical and challenging issue for operational hydrologic forecasting (Arsenault & Brissette, 2016; Liu & Gupta, 2007; Sudheer, Lakshmi, & Chaubey, 2011; Van Griensven, Meixner, Srinivasan, & Grunwald, 2008).

To assess hydrological modeling uncertainty, two major sources need to be explored under stationary climate conditions: (1) model parameters describing the difficulty to attribute exact parameter values because of finite lengths and uncertainty in the calibration data, imperfect understanding of process, approximations and (2) model structure, mentioned to as "model uncertainty", due to lumped and simplified representation of hydrological processes (Lindenschmidt, Fleischbein, & Baborowski, 2007; Renard, Kavetski, Kuczera, Thyer, & Franks, 2010). Additional uncertainties can be considered under non-stationary climate conditions (changing climatic conditions), stemming from instability of parameters because of variations in both catchment characteristics and main hydrological processes (Brigode, Oudin, & Perrin, 2013). Wilby (2005) showed that the uncertainty of the non-uniqueness of parameters affects the representation of climate change signals on streamflows. Other authors reported the strong influences of the model structure, with substantial inter-model differences on projected streamflows (e.g. Mendoza et al., 2015; Vano, Kim, Rupp, & Mote, 2015; Velázquez, Troin, & Caya, 2015a). When both sources are compared to each other, the model structure uncertainty remains the greatest one under stationary and non-stationary climate conditions (Fu, James, & Yao, 2015; Najafi, Moradkhani, & Jung, 2011; Poulin, Brissette, Leconte, Arsenault, & Malo, 2011). The characterization and reduction of uncertainties are thus challenging in the hydrological modeling field (Clark et al., 2016). The understanding of these uncertainties is still incomplete, and dedicated researches are needed to substantially reduce model uncertainty.

Many components structure a hydrological model (HM); each component brings an uncertainty contributing to the total uncertainty of the model structure. The uncertainties of the most sensitive components to non-stationary climate conditions require to be explored in order to provide reliable streamflow projections. The most sensitive components to non-stationary climate conditions include the snow models (SM) and the potential evapotranspiration (PET) computation methods. In a recent work, Troin et al. (2016) made an attempt to estimate the uncertainty on the snow water equivalent projections related to the type of SM by using an ensemble of eight degree-day (DD) and mixed degree-day/energy balance (DD/EB) SM. The authors concluded that both types of SM lead to similar snow water equivalent projections at the basin scale. This study confirmed the results of Troin et al. (2015a), where, using the same ensemble of SM, a low level of SM uncertainty in simulating snowmelt flows is found under the current climate. Bae et al. (2011) explored the impacts of three HMs with PET-related computation methods on streamflow projections. The authors showed that the PET methods have contrasted sensitivities to climate projections, resulting in a large spread of projected runoff changes at the catchment scale. A key research effort towards the reduction of model structure uncertainty therefore aims to better characterize the uncertainty derived from each main HM component by using adequate modeling frameworks. Recent studies focusing on the analysis of model uncertainties have highlighted that the multi-model strategy is a skilful pathway to explore the associated uncertainties to hydrological projections resulting from models (e.g. Ajami, Duan, Gao, & Sorooshian, 2006; Bohn, Sonessa, & Lettenmaier, 2010; Duan, Ajami, Gao, & Sorooshian, 2007; Velázquez, Anctil, Ramos, & Perrin, 2011). Comparing SMs-PET methods-HM structures multi-combinations and showing how the combinations of these multi-model components can affect streamflow projections could improve the knowledge of model structure uncertainty.

The aim of this study is to evaluate different types of SM-PET-HM combinations, by taking into account uncertainties from the SMs (DD versus DD/EB models), the PET methods (temperature-based versus radiation-based methods) and the HM structures. The multi-model

approach is composed of 105 combinations of SM-PET-HM. The ensemble is implemented over two typical catchments in Quebec, Canada in terms of land cover types and Nordic climates. The Arnaud Basin is located in the Nunavik region, Northern Quebec and the Mistassini Basin is situated in the Saguenay-Lac-Saint-Jean region in Central Quebec. Figure-A II-1 shows the location of each basin and Table-A II-1 lists some general characteristics. For the two catchments, daily climate data came from the high-resolution simulations ensemble of the Canadian Regional Climate Model (CRCM-15 km) drived by the Canadian GCM (CGCM3). The CRCM simulations feed the multi-models to simulate the catchments' hydrology. The evaluation is made under current (1961-1990) and future (2041-2060) climates by evaluating the differences between the hydrological simulations and the CRCM pseudo-observations in the virtual world. More information is given in section II.2.2. The virtual world provides a complete database without missing data or space and time incoherences (Arsenault & Brissette, 2014, 2016; Maraun, 2012; Troin et al., 2016; Velázquez et al., 2015b). In addition to the decreasing of the uncertainties due to the non-existence and/or shortcomings of the meteorological time series in the real world over the study basins (e.g., solar radiation, snow albedo, surface pressure), this approach provides hydrologic projections under climate change conditions. The CRCM simulations used in this work allow the uncertainty of the natural variability of climate to be explored, and then compared to hydrological model uncertainty. Additional uncertainties related to the RCMs (e.g., scenario and inter-model variability) cannot be evaluated with this ensemble, since it rests on a single scenario and regional model.

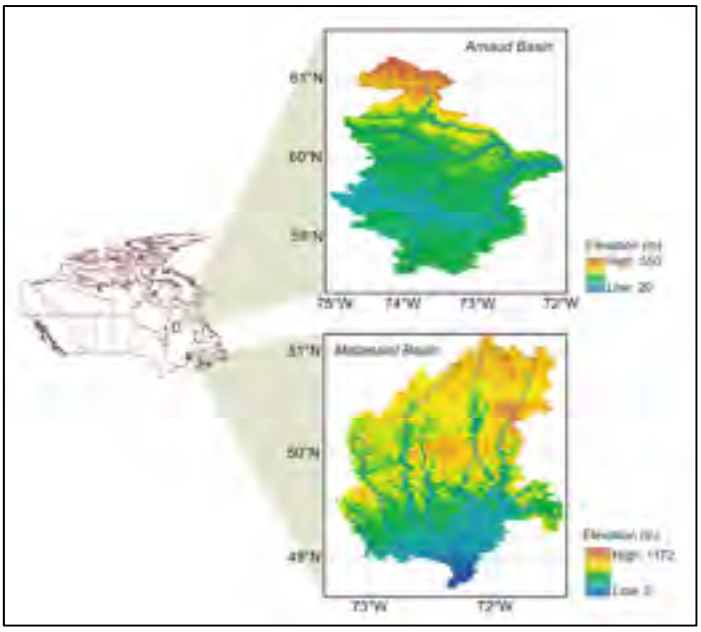

Figure-A II-1 Location map of the study basins

Table-A II-1 General characteristics of the study basins

	Basins	
	Mistassini	Arnaud
Drainage area (km^2)*	11250	26775
Elevation range (m)	120-700	22-547
Main land cover type	Forest (90%)	Tundra (90%)
Hydrology – Virtual world		
Mean daily discharge (m^3/s)	198	204
Annual maximum snow water equivalent (mm)	215	171
Annual average actual evaporation flux (mm)	395	168
Climate – Virtual world		
Annual average precipitation total (mm)	960	409
Annual daily temperature (°C)		
Min.	-51.5	-56.8
Max.	+34.0	+29.4
Mean	-2.5	-12.1
Number of grid points		
CRCM (15-km)	50	119

* The reported basin surface areas are based on the streamflow gauge position and not on the basin outlet.

In the following, the CRCM ensemble and the virtual world concept are presented in Section II.2. Section II.3 described the multi-model approach. In Section II.4, the results from the multi-model approach are analyzed for the historical and future periods. Section II.5 discusses the uncertainties of the model components and the natural climate variability on the streamflow projections. Concluding comments are provided in Section II.6.

II.2 Experimental setup

II.2.1 RCM simulations

In the present study, we used the CRCM version 4.2.4, with a 15-km high-resolution horizontal grid-point spacing (Caya & Laprise, 1999; Laprise, Caya, Frigon, & Paquin, 2003; Plummer et al., 2006). Its domain (225 × 225 grid points) is focused over the Quebec province, which dictates the selection of the basins to this geographical area. Energy and water exchanges between the surface and the atmosphere are simulated by the Canadian LAnd Surface Scheme (CLASS 2.7; Verseghy, McFarlane, & Lazare, 1993). Snow cover and vegetation canopy are considered in CLASS (Verseghy, 1991; Verseghy et al., 1993). An evaluation of the RCM-modeled water cycle can be found in Music et Caya (2007). Additional information about the simulated snow features by CLASS is provided in Langlois et al. (2014); Langlois, Royer, Fillol, Frigon, et Laprise (2004).

II.2.2 The virtual world principle

The analysis is performed in the virtual world based upon the CRCM simulations forced at the lateral boundary by a 2-member ensemble of the third-generation Canadian Centre for Climate Modelling and Analysis (CCCma) Coupled Global Climate Model (CGCM3.1v2; Flato and Boer 2001). The CRCM simulations are from an older experiment, where the driving CGCM was drived by the strongest A2 emissions scenario during the 21st century. Nevertheless, regarding the recent literature on comparisons of climate change signals in the third and fifth phase of the Coupled Model Intercomparison Project (CMIP3 and CMIP5),

several studies have reported quite similar trends in projections of future climate between both ensembles (e.g. Lutz, Immerzeel, Gobiet, Pellicciotti, & Bierkens, 2013; Supharatid, 2015; Woldemeskel, Sharma, Sivakumar, & Mehrotra, 2016). The virtual world rests on the principle that the CRCM simulations are considered as pseudo-observations for both the historical and future periods (Troin et al., 2016). The daily climate variables extracted from the CRCM simulations are maximum and minimum temperatures, precipitation, snow albedo, incoming shortwave radiation, wind vectors at 10-meter height, screen-specific humidity and surface pressure.

The pseudo-observed streamflow, on the other hand, is not directly available in the CRCM outputs. Indeed, the climate model simulates overland and below ground runoffs in the land-surface scheme for each grid cell, but excess water is simply removed from the water column at each time step rather than being routed to a river as streamflow. A method was devised by Arsenault et Brissette (2014) to transform distributed runoffs into outlet streamflows. The streamflow generation is carried out in the following three general steps. The first step is to separate the catchment into grids and classify them into zones of influence defined as the distance to the outlet of the basin. This ensures that the routing scheme takes transit time into account on larger basins. Larger basins are awarded more zones than smaller ones. The second step is to average the runoff values of all grid points inside each zone, to require a single averaged value of surface runoff and below ground runoff per zone. Finally, the unit hydrograph parameters are calibrated using the interannual observed and reconstructed hydrographs on the reference period. Even though the day to day representation of discharge from the model and the associated routing scheme is biased by the generation process, the approach allows for comparative analyses as is conducted in current conditions. The routing scheme is applied to generate the daily pseudo-observed flows in both the historical (1961-1990) and future (2041-2060) periods. Readers can refer to Arsenault and Brissette (2014) for a more detailed methodology.

The advantages of working within a virtual setting were put to use in various recent studies. This method was successfully applied in regionalization method analysis, where regionalization approaches limitations and process uncertainties were quantified (Arsenault & Brissette, 2016). It was also used to determine the optimal rain gauge observation network density in data-sparse regions (Arsenault & Brissette, 2014). Furthermore, Minville et al. (2014) calibrated the HM with observed data and conditioned the ET values to the CRCM-ET values. The skill of the HM was significantly improved. The extreme flooding event in the Richelieu River Basin on the USA/Canada border in 2011 have been remarkably well modelled by two RCMs driven by European Reanalysis data (Lucas-Picher et al., 2015). These studies indicate that the RCM simulations can convey important information where no measurements exist. Finally, Beauchamp, Leconte, Trudel, et Brissette (2013) used the virtual setting to calculate new maximum probable precipitation and flood values, which compare favorably with existing methods. There is thus a growing interest in using virtual laboratories to improve process simulation and reduce uncertainty.

II.3 The multi-model pathway

The procedure to assess the hydrologic uncertainty of model components in streamflow projections using different SM-PET-HM combinations is shown in Figure-A II-2. The multi-model pathway requires to couple seven snow models with five PET computation methods through three HMs in view of obtaining 105 discharge simulations per catchment over the historical and future periods. In the following, each model component is briefly presented.

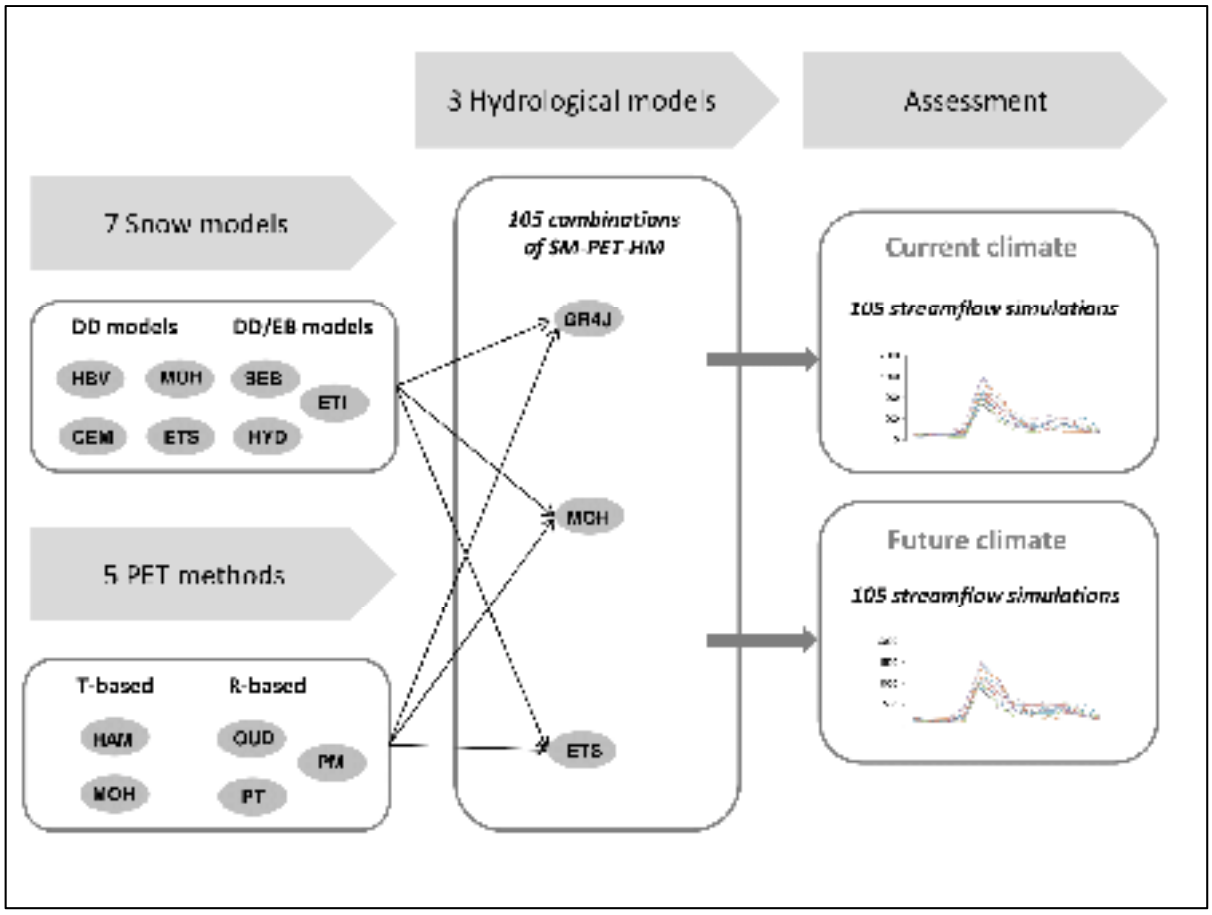

Figure-A II-2 Schematic description of the multi-model approach conducted per catchment in this study.

I.3.1 Snow models description

In the present study, we compare four SMs derived from the HBV, MOHYSE, HMETS and HYDROTEL HMs and three extra SMs (CEMANEIGE, SEB and ETI). The seven SMs are grouped in DD (HBV, MOHYSE, CEMANEIGE and HMETS) and DD/EB models (HYDROTEL, ETI and SEB). The discrepancies between the seven SMs are listed in Table-A II-2. The algorithms linked to the snowmelt routines are only described in Sections II.3.1.1 and II.3.1.2. A more detailed presentation of the SMs is found in Troin et al. (2016).

Table-A II-2 Conceptual differences between SM model components. The acronyms of models and methods used in this study are also specified; T = temperature, P = precipitation amount (liquid and/or solid), ISR = incoming shortwave radiation and α = snow albedo

Snow models	Model type	Input data	Snow parameters	References
MOHYSE (MOH)	DD	T, P	2	Fortin et Turcotte (2007)
HBV (HBV)	DD	T, P	3	Bergstrom (1976)
CEMANEIGE (CEM)	DD	T, P	2	Valéry (2010)
HMETS (ETS)	DD	T, P	10	Vehviläinen (1992b)
HYDROTEL (HYD)	DD/EB	T, P	5	Turcotte, Fortin, Fortin, Fortin, et Villeneuve (2007)
SEB (SEB)	DD/EB	T, P, ISR, α	2	Machguth, Paul, Hoelzle, et Haeberli (2006)
ETI (ETI)	DD/EB	T, P, ISR, α	2	Pellicciotti et al. (2005)

II.3.1.1 DD models

II.3.1.1.1 HBV

The SM derived from HBV (Hydrologiska Byråns Vattenbalansavdelning; Bergstrom, 1976) is a simple model based on a temperature index. If the mean daily air temperature (T) is above the threshold value (T_s), snowmelt begins following this equation:

$$SNOWMELT = (T - T_s) \times c_f \quad \text{(II.1)}$$

where SNOWMELT is the snowmelt (mm of water); c_f is the factor of melt (mm °C^{-1} d^{-1}); T is the mean air temperature (^{0}C); and T_s is a threshold temperature value (°C).

II.3.1.1.2 MOHYSE

The MOHYSE SM (MOdèle HYdrologique Simplifié à l'Extrême) is based upon a DD relation (Fortin & Turcotte, 2007). The snowpack's mass balance is calculated as:

$$SWE = SNOW + WE\_SNOWMELT \quad \text{(II.2)}$$

with $$WE\_SNOWMELT = \min(c_f \times \max(T - T_s, 0), SWE) \quad \text{(II.3)}$$

where *SWE* is the snowpack's snow water equivalent (mm of water); *SNOW* is the water equivalent of snow precipitation (mm of water); and *WE_SNOWMELT* is the snowmelt's water equivalent (mm of water).

II.3.1.1.3 HMETS

The DD-based HMETS SM (Hydrological Model of Ecole de technologie supérieure) comes from the study of Vehviläinen (1992b). The potential snowmelt (*pSNOWMELT*) is computed as:

$$pSNOWMELT = \max(0, ddf \times (T - T_{bm})) \quad \text{if } T > T_{bm} \quad \text{(II.4)}$$

with $$ddf = ddf_{min} \times (1 + k_{cum} \times CSM) \quad \text{(II.5)}$$

where *ddf* is a factor ranging between a minimum (ddf_{min}) and a maximal value which is function of the snowmelt cumulative amount (*CSM*) (mm °C^{-1}); T_{bm} is the base melting temperature (°C); k_{cum} is an empirical parameter (mm^{-1} of water); and *CSM* is the snowmelt cumulative amount (mm of water).

II.3.1.1.4 CEMANEIGE

The SM derived from CEMANEIGE is a complex version of the MOHYSE SM, where the simulation of snowpack dynamic is improved (Valéry, 2010). The mass balance for the snowpack is computed as follows:

$$SWE = SNOW - ASNOWMELT \quad \text{(II.6)}$$

$$ASNOWMELT = (0.1 + SNOWCOVER \times 0.9) \times PSNOWMELT \quad \text{(II.7)}$$

where *ASNOWMELT* is the actual snowmelt (mm of water); *SNOWCOVER* is the percentage of snow over the catchment; and *PSNOWMELT* is the potential snowmelt.

II.3.1.2 DD/EB models

II.3.1.2.1 SEB

For the Simplified Energy Balance (SEB; Machguth et al., 2006) model, the energy and mass balances within the snowpack are given by:

$$SWE = SNOW - SNOWMELT \quad \text{(II.8)}$$

with $$SNOWMELT = \frac{ISR \times (1 - \alpha) + C_0 + C_1 \times T}{\rho \times L_f} \times 1000 \quad \text{(II.9)}$$

where α is the snow albedo; *ISR* is the incoming shortwave radiation (W m^{-2}); C_0 and C_1 are factors accounting for the temperature-dependent energy fluxes (W m^{-2} and W m^{-2} °C^{-1}, respectively); ρ is the density of snow (kg m^{-3}); and L_f is the fusion latent heat (J kg^{-1}).

II.3.1.2.2 ETI

The Enhanced Temperature-Index (ETI; Pellicciotti et al., 2005) model estimates the energy and mass balances in the snowpack following these two equations:

$$SWE = SNOW - SNOWMELT \quad \text{(II.10)}$$

with $$SNOWMELT = SRf \times (1 - \alpha) \times ISR + Tf \times T \quad \text{if } T > T_s \tag{II.11}$$

where *SRf* and *Tf* are the factors for the shortwave radiation (m^2 mm^{-1} W^{-1} d^{-1}) and temperature (mm $°C^{-1}$ d^{-1}), respectively.

II.3.1.2.3 HYDROTEL

In the HYDROTEL SM (Turcotte et al., 2007), a mass balance is used to calculate the *SWE* as:

$$SWE = SPLIQ + SNOW - M + WATER \tag{II.12}$$

with $$M = \frac{-U}{C_f \times \rho_w} \tag{II.13}$$

where *WATER* is the snow water equivalent on the ground (mm of water); *M* is melting water (mm of water); *U* is the calorific deficit (J m^{-2}); C_f is the melting heat (J kg^{-1}); and ρ_w is the water density (kg m^{-3}).

II.3.2 Description of potential evapotranspiration methods

Five PET computation methods are evaluated in this study. These methods are classified in temperature-based methods (HAM and MOH) and radiation-based methods (OUD, PT and PM). Table-A II-3 lists the conceptual differences between the five PET methods. In Sections II.3.2.1 and II.3.2.2, the PET algorithms are briefly described.

Table-A II-3 Conceptual differences between model components for the PET computation methods. The acronyms of models and methods used in this study are also specified; T = temperature, lat = latitude, R_e = extraterrestrial solar radiation, p = surface pressure, sH = specific humidity, w = wind speed, and R_n = net radiation

PET models	Method type	Input data	References
Hamon (HAM)	T-based	T, lat	Hamon (1961)
MOHYSE (MOH)	T-based	T, lat	Fortin et Turcotte (2007)
Oudin (OUD)	R-based	T, R_e	Oudin et al. (2005)
Penman-Monteith (PM)	R-based	T, p, sH, w, R_n	Monteith (1965)
Priestley-Taylor (PT)	R-based	T, p, sH, w, R_n	Priestley et Taylor (1972)

II.3.2.1 Temperature-based methods

II.3.2.1.1 Hamon

The Hamon PET method (Hamon, 1961) is computed as:

$$PET = \frac{29.8 \times DL \times ESAT}{\overline{T} + 273.3} \tag{II.14}$$

with

$$SVD = 0.6108 \exp\left(\frac{\overline{T} \times 17.27}{\overline{T} \times 237.3}\right) \tag{II.15}$$

where DL is the length of daytime corresponding to the sunrise to sunset in multiples of 12 hours for a given latitude; $\overline{T}$ is the mean temperature of air (°C); and SVD is the saturated vapor density (hPa) for a given mean air temperature.

II.3.2.1.2 MOHYSE

The PET computation method derived from MOHYSE (Fortin & Turcotte, 2007) is expressed by:

$$PET = \frac{ESAT}{0.6108 \cdot \pi} \cdot \cos^{-1}\left(-\tan\left(\frac{L \cdot \pi}{180}\right) \times \tan(S)\right) \quad \text{(II.16)}$$

with

$$S = 0.41 \times \sin\left(\frac{day - 80}{365} \times 2\pi\right) \quad \text{(II.17)}$$

where *S* is the daily solar declination.

II.3.2.2 Radiation-based methods

II.3.2.2.1 Oudin

The PET Oudin formula (Oudin et al., 2005) is based upon the daily extraterrestrial solar radiation (RE) as:

$$PET = \frac{R_e \times (T + 5)}{\lambda \times \rho_w \times 100} \quad \text{if } T + 5 > 0 \quad \text{(II.18)}$$

$$PET = 0 \quad \text{otherwise}$$

where T is the temperature of air (°C); λ is the vaporisation latent heat (MJ kg^{-1}); R_e is the extraterrestrial radiation (MJ m^{-2}); and ρ_w is the water density (kg m^{-3}).

II.3.2.2.2 Penman-Monteith

The Penman-Monteith formula (Monteith, 1965) is given by:

$$PET = \frac{sR_n + \rho_a \times c_p(e_d - e_a)/r_a}{\lambda \times \rho_w \times \left(s + \gamma\left(1 + \frac{r_s}{r_a}\right)\right)} \quad \text{(II.19)}$$

where R_n is the net solar radiation (MJ m^{-2} d^{-1}); s is the vapor pressure curve slope (kPa $°C^{-1}$); ρ_a is the density of air (kg m^{-3}); e_d is the air saturation vapor pressure (kPa); e_a is the water vapor pressure of air (kPa); c_p is the specific heat at constant pressure (MJ kg^{-1} $^0C^{-1}$); γ is the psychrometric constant (kPa $^0C^{-1}$); r_s is the plant canopy resistance (s m^{-1}); and r_a is the aerodynamic resistance (s m^{-1}).

II.3.2.2.3 Priestley-Taylor

The Priestley-Taylor's equation is expressed as:

$$PET = \frac{\alpha_a \times \Delta R_n}{\lambda \times \rho_w \times (s + \gamma)} \quad \text{with } \alpha_a = 1.26 \tag{II.20}$$

II.3.3 HM description

II.3.3.1 GR4J

The GR4J rainfall-runoff model is a conceptual model with four parameters, operating at the daily time step (Perrin et al., 2003). It has been used in many research works (Troin et al., 2015a; Velázquez et al., 2015b). GR4J is composed of a production and a routing store. Required inputs data are potential evapotranspiration and precipitation.

II.3.3.2 MOHYSE

The conceptual and lumped MOHYSE model (Fortin & Turcotte, 2007) was used in streamflow prediction in ungauged sites (Arsenault, Poissant, & Brissette, 2015) and in multi-model ensemble studies (Troin et al., 2015a). A production store and a routing store compose the model, with 10 parameters. Daily mean temperature and daily precipitation are necessary to run the model.

II.3.3.3 HMETS

The conceptual and lumped HMETS model uses two reservoirs for the saturated and vadose zones. The hydrological processes simulated by the model are infiltration, evapotranspiration, snow melting and accumulation, and flow routing towards the outlet of the catchment. HMETS with these 21 parameters requires daily maximum and minimum temperature and daily precipitation data. The model was used in multi-model projects (Arsenault, Gatien, Renaud, Brissette, & Martel, 2015) as well as in impact studies of climate change (Chen et al., 2011a). Readers can refer to Troin et al. (2015a) for additional information about the three HMs.

II.3.4 Models calibration, evaluation and uncertainty

The calibration of the 105 multi-model combinations is performed using all the CRCM ensemble simulations. For each SM-PET-HM combination, the parameters are calibrated based on the daily pseudo-observed discharge on even years, and the performance of the combination models is validated on odd years with the daily pseudo-observations for the 1961-1990 period. An evaluation of the simulations is then made over a period of 20 years for current (1971-1990) and future (2041-2060) climates, where daily simulated streamflows (Q) are compared to the pseudo-observations in the virtual world.

The uncertainties in the projections of future streamflow are evaluated by analysing the impacts of the model components (SM versus PET versus HM structure) and the uncertainty of the natural climate variability (CRCM ensemble simulations). To that end, three flow indicators are chosen: mean annual flows (MF), spring high flows which is the maximum flow observed in one day (March-June; HF) and summer low flows defined as the lowest 7-day average flow (July-October; LF). For each flow indicators, the pseudo-observed and simulated climate change signals (CCS) are calculated as:

$$CCS_{pseudo-obs(\%)} = \frac{(FI^{fut}_{pseudo-obs} - FI^{ref}_{pseudo-obs})}{FI^{ref}_{pseudo-obs}} \times 100 \qquad \text{(II.21)}$$

$$CCS_{sim(\%)} = \frac{(FI^{fut}_{sim} - FI^{ref}_{sim})}{FI^{ref}_{sim}} \times 100$$

where FI^{ref}_{sim} and $FI^{ref}_{pseudo-obs}$ are the simulated and pseudo-observed flow indicators (MF or HF or LF) over the 20-year reference period, respectively, and FI^{fut}_{sim} and $FI^{fut}_{pseudo-obs}$ are the simulated and pseudo-observed flow indicators (MF or HF or LF) over the 20-year future period, respectively.

The magnitudes of uncertainty sources in the climate change signals of the flow indicators in the virtual world are isolated and compared by using the variance decomposition method (Kendall, 1977; von Storch & Zwiers, 1999). The present approach follows Déqué et al. (2007); Ferro (2004); and Roberts et Snelgrove (2015).

Let I_{sphc} be the climate change signal on a given flow indicator for SM *s* varying from 1 to 7, PET method *e* varying from 1 to 5, HM *h* varying from 1 to 3, and natural climate variability *c* varying from 1 to 2 (2-member ensemble of the CRCM). The total variance of I_{sphc} is decomposed as:

$$V(I_{sphc}) = S + E + H + C + SE + SH + SC + EH + EC + HC + SEH + SEC + SHC + EHC + SEHC \qquad \text{(II.22)}$$

with, for example:

$$S = \frac{1}{7}\sum_{s=1}^{7}(IF_{s\bullet\bullet\bullet} - IF)^2 \quad \text{and} \qquad \text{(II.23)}$$

$$SE = \frac{1}{35}\sum_{s=1}^{7}\sum_{e=1}^{5}(IF_{se\bullet\bullet} - IF_{s\bullet\bullet\bullet} - IF_{\bullet e\bullet} + IF_{\bullet\bullet\bullet\bullet})^2$$

where • is the average of the index it has substituted. The terms of Equation (II.23) are expressed in percentages of variance. The total variance attributed to the SM can be expressed, for example, as:

$$V(S) = S + SE + SH + SC + SEH + SEC + SHC + SEHC \quad \text{(II.24)}$$

The magnitude of *V(S)*, *V(E)*, *V(H)* and *V(C)* indicates the effect of each uncertainty source on the value of the climate change signal for a given flow indicator.

II.4 Results

II.4.1 Evaluating the multi-model approach over current climate

Before comparing the reliability of the multi-model combinations at reproducing pseudo-observed streamflows under the historical period, the performance of the HMs during calibration and validation periods are analysed. For each HM, the 35-simulation NS values (7 SMs × 5 PET methods) are evaluated per catchment for the calibration and validation (Figure-A II-3). Models evaluations are performed on the daily hydrological data. Results show that the HMs differ in performance; the MOHYSE and HMETS ensembles of 35 simulations perform well, with NS scores from 0.61 to 0.91, for the calibration, and from 0.53 to 0.88, for the validation over the two basins. The GR4J ensemble performance is lower, with NS scores between 0.24 and 0.82 (median of 0.52) over the validation period. We then compare the pseudo-observed discharges in the virtual world with the observed discharges from the cQ2 database for the 1961-1990 period. The agreement between the pseudo-observed discharges and the observations is satisfactory, with NS values of 0.72 over

the Mistassini Basin and of 0.82 over the Arnaud Basin, giving credence to the virtual world approach for the present experiments.

Figures-A II-4 and II-5 present the 20-year pseudo-observed and simulated monthly mean discharges under current climate for the Mistassini Basin and the Arnaud Basin, respectively. The three HM ensemble simulations capture the timing of the spring peak flow over the two catchments. Overall, the MOHYSE and HMETS ensemble simulations provide a fairly good simulation of the magnitude of the spring peak flows over the basins; however an underestimation of the magnitude's spring peak flow is observed with the GR4J ensemble simulations over the two basins (mean relative bias of 16%). The performance is similar for the DD and DD/EB SMs in the SM-PET-HM combinations over the two basins, indicating a lack of sensitivity of the HMs to the SMs under the current climate. The PET methods have a strongest effect on the discharge simulations than the SMs for the catchments under investigation. Significant variations in the magnitude of the HM-simulated seasonal streamflows are obtained according to the PET methods, particularly when the PET methods are combined with HMETS and GR4J. The two approaches of the PET methods in the SM-PET-HM combinations result in contrasted seasonal streamflow simulations. For instance, the summer flows (JJA) are clearly overestimated in the GR4J ensemble simulations over the Mistassini Basin, when the two T-based PET methods (HAM and MOH) are combined with the SMs. This is confirmed with the analysis of percent deviation values in Supplementary materials Table-A II-S8. From Figures-A II-4 and II-5, we can see that the HM structure has an impact on the streamflow simulations, with noticeable variations throughout the year. However, all the HM structures lead to underestimated mean annual streamflows over the basins. We can note that distributions of the percent deviation values differ between the CRCM simulations under current climate for the study basins.

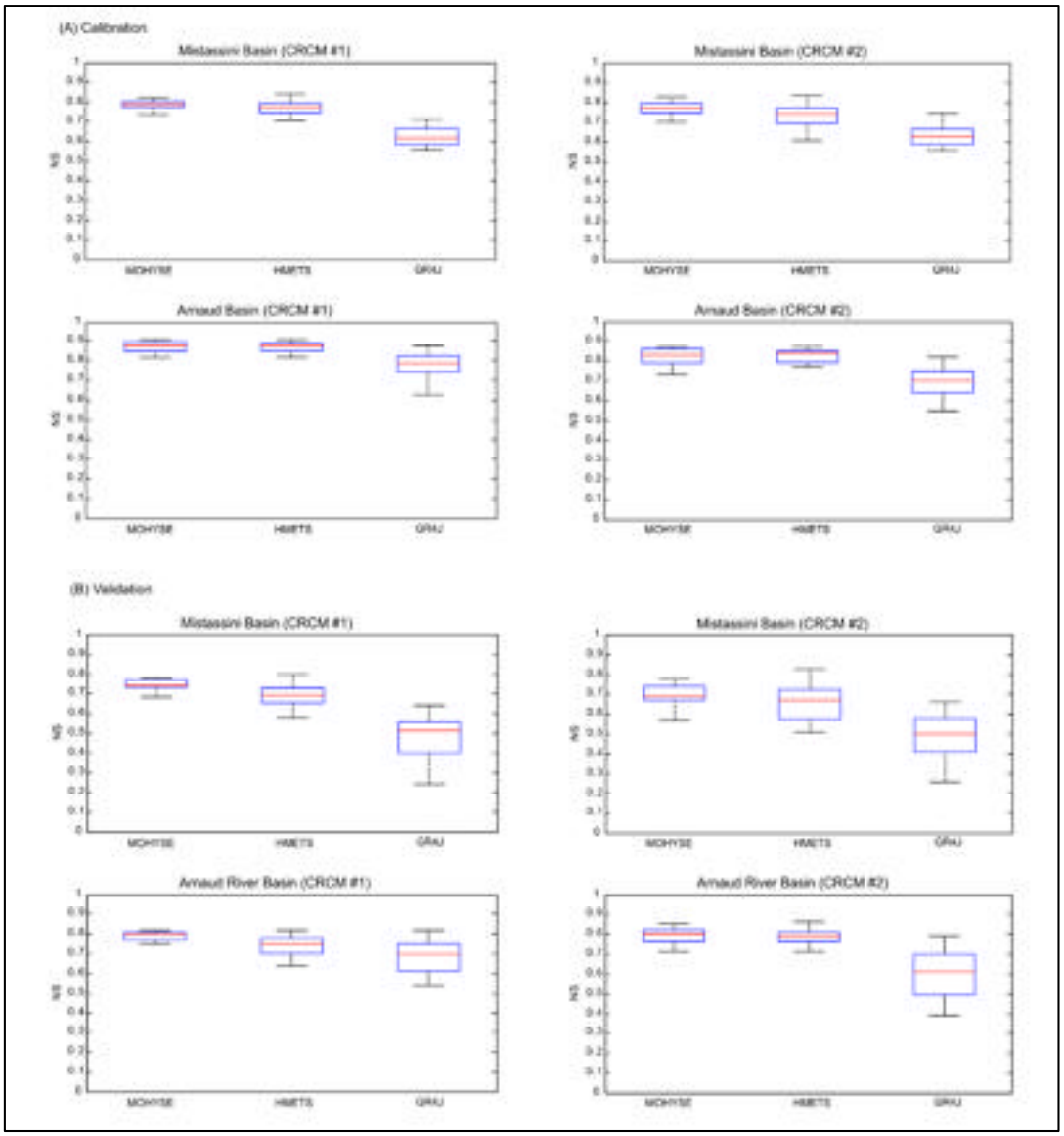

Figure-A II-3 Performance of the multi-models to simulate streamflow over the 15-year (A) calibration and (B) validation period over the study basins. The performance is expressed in terms of Nash Sutcliffe efficiency (NS) criterion based on the 35 SM-PET combinations per hydrological model (MOHYSE, HMETS and GR4J). For instance, on Figure 3A, for the Mistassini Basin (CRCM#1), the MOHYSE box represents the 35 NS values for all the SM-PET-MOHYSE combinations. The whiskers represent the higher and lower NS values among the 35 SM-PET-MOHYSE combinations, whereas the box presents the 25th, 50th, and 75th quantiles. The left column represent the results obtained with the CRCM#1 and the right column those obtained with the CRCM#2. Each line represents the results to a considered catchment

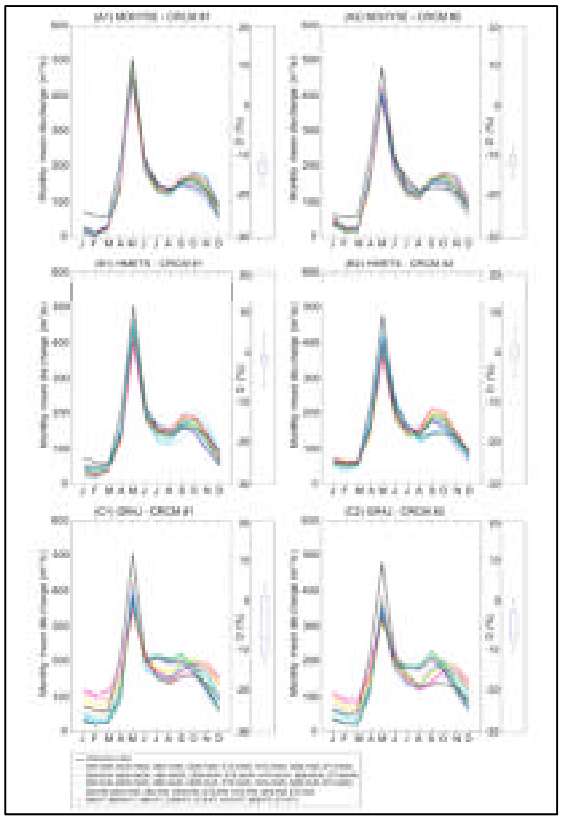

Figure-A II-4 Mean monthly discharges simulated using the CRCM ensemble simulations as input to the multi-models over the Mistassini Basin for the 1971-1990 period. The simulations are compared to the pseudo-observed discharges (PSEUDO-OBS) from the virtual world under the current climate. The box plots to the right present the percent deviation (D) of streamflow (%). For instance, on Figure 4A1, the box represents the 35 D values for all the SM-PET-MOHYSE combinations obtained with CRCM#1. The whiskers represent the higher and lower D values among the 35 SM-PET-MOHYSE combinations, whereas the box presents the 25th, 50th, and 75th quantiles. Same description for the others two models, (B) HMETS and (C) GR4J, with the two CRCM ensemble simulations (CRCM#1 and CRCM #2)

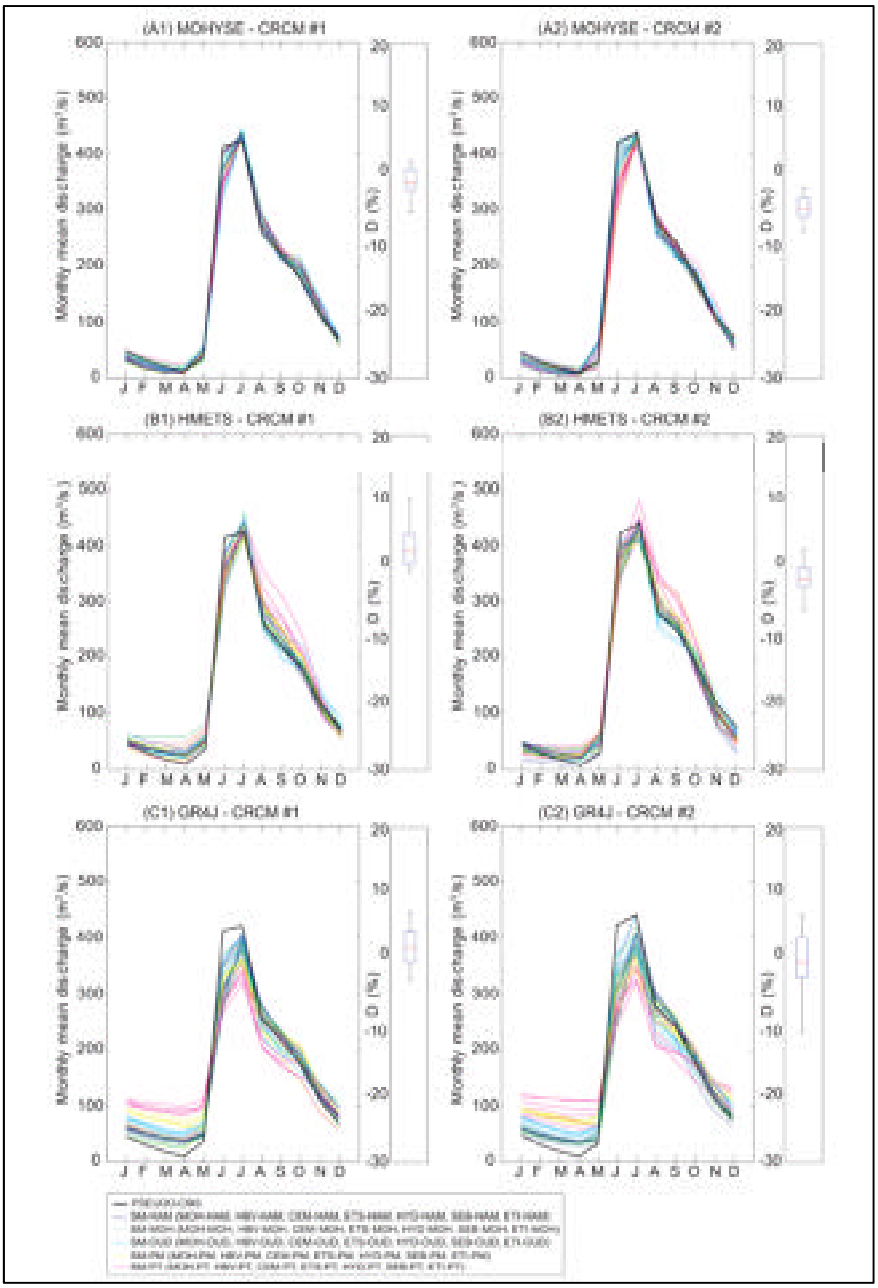

Figure-A II-5 As in Figure-A II-4, but for the Arnaud Basin

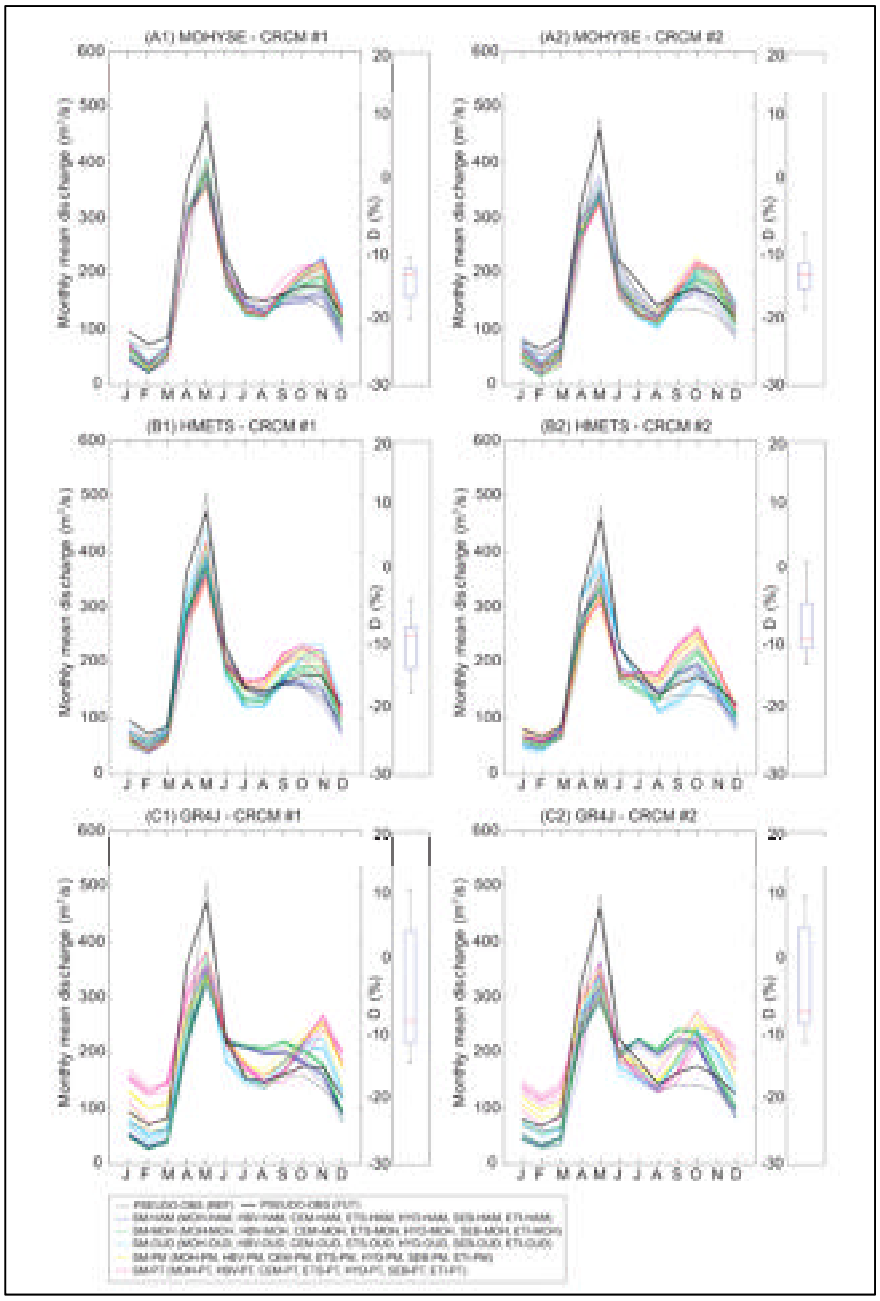

Figure-A II-6 As in Figure-A II-4, but for mean monthly discharges simulated using the CRCM ensemble simulations as input to the multimodels over the Mistassini basin for the 2041-60 period. The simulations are compared to the pseudo-observed simulations in the virtual world under the current climate are also plotted for comparison

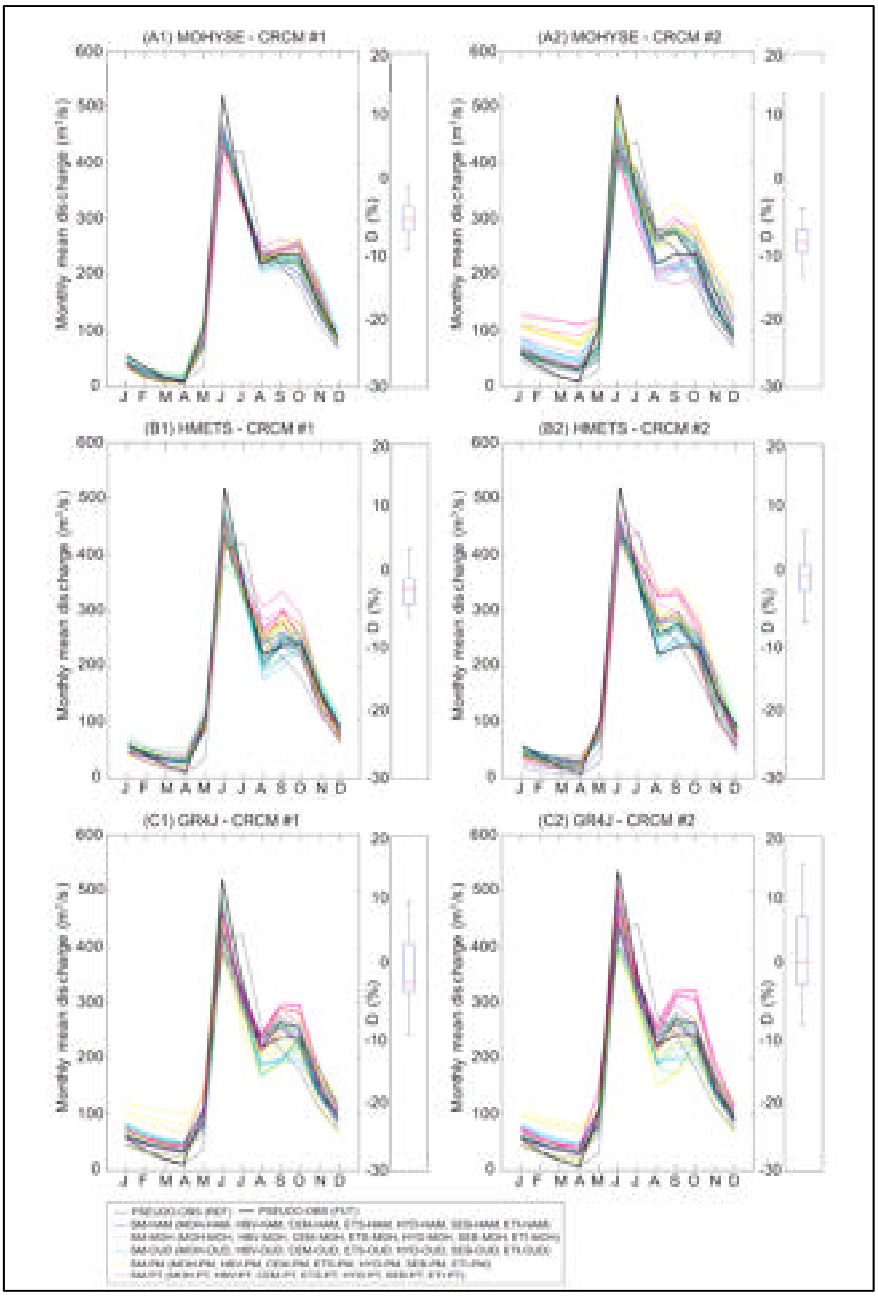

Figure-A II-7 As in Figure-A II-6, but for the Arnaud Basin

II.4.2 Evaluating the multi-model approach over future climate

The performance of the multi-models is then evaluated under the future climate by comparing the simulated streamflows from the SM-PET-HM combinations with the pseudo-observations.

Figures-A II-6 and II-7 show the 20-year pseudo-observed and simulated monthly mean discharges for the Mistassini Basin and the Arnaud Basin, respectively, over the future period. The timing of the spring peak flow is fairly well represented in the simulations over the two basins, but none of the HM ensemble simulations captures its magnitude. Again, this provides indications that the SMs cannot accurately represent the magnitude of the snowmelt peak under the future climate over the study basins. The GR4J ensemble simulations still give the worst representations of the magnitude of peak flows over both basins. As expected, the DD and DD/EB SM in the SM-PET-HM combinations exhibit similar performances at simulating streamflows under the future climate, with equal mean percent deviation values. The discreapancies between the simulated and pseudo-observed seasonal discharges also give a hint of the effect of the selected PET method in the HM; the seasonal cycle of streamflows under the future climate strongly differs according to the PET methods through the SM-PET-HM combinations (Figures-A II-6 and II-7). From an analysis of percent deviation values in Supplementary materials Table-A II-S9, we can see that the lowest performances of the ensemble simulations are based on the T-based PET methods. In contrast, when PET is computed with R-based equations, the performance of the SM-PET-HM combinations is more acceptable. The comparison of future pseudo-observed and simulated seasonal streamflows confirms the influence of the HM structure on the reliability of streamflow simulations, as previously detected under the current climate (Figures-A II-4 and II-5). Results are relatively homogenous between MOHYSE and HMETS, but GR4J systematically provides contrasted seasonal streamflow simulations, independently of the SM-PET combination. The spread of the ensemble simulations is larger for the GR4J ensemble than for the two other HMs' ensembles under the future climate (Figures-A II-6 and II-7) – and

also under the current climate (Figures-A II-4 and II-5). However, all the HM ensemble simulations lead to underestimated mean annual streamflows for the future period. Similarly to the historical period, the spread of the streamflow simulations is large according to the two CRCM simulations used to drive the SM-PET-HM combinations for the future period (Figures-A II-6 and II-7).

II.4.3 Impacts of climate change on virtual hydrological cycle

The climate change impacts on future hydroclimate conditions are analysed over the two study basins. From Table-A II-4, we can see that important changes are probably to occur for temperature and precipitation at the 2050 horizon, with a mean annual change of 3.2°C to 3.4°C for temperature and of 22% to 47% for precipitation over the basins. All the HM simulations show an increase in mean annual flows (MF) over the basins from 15% to 25%, although this trend seems to be more noticeable for the northernmost basin (Figure-A II-1).

For the Mistassini Basin, the projections of the timing of the spring peak flow at the 2050 horizon coincide with the present-day peak flow (Figures-A II-4 and II-6); its magnitude is likely to decrease by 11% to 19% over the Mistassini Basin according to the HM ensemble simulations. However, for the Arnaud Basin, an earlier spring peak flow is observed at the 2050 horizon, with an advance of one month and an increased magnitude of 9% to 58% (Figures-A II-5 and II-7). The summer (JJA) and autumn (SON) streamflows are projected to increase by 0% to 6% and by 15% to 17%, respectively, according to the HM ensemble simulations over the Mistassini Basin. A high increase in winter streamflows (DJF) of 37% to 60% is also expected over this basin. The Arnaud Basin exhibits more pronounced changes in seasonal streamflows, with increased winter and autumn flows by 19% to 69% and by 10% to 65%, respectively, according to the HM ensemble simulations. The summer streamflows are expected to vary between -12% and 59% according to the HM ensemble simulations over the basin. In addition, the figures show large differences in the seasonal streamflow projections between the two CRCM simulations within each HM ensemble

simulation over the two basins. This indicates a high sensitivity of the SM-PET-HM combinations to the climate model simulations. A detailed analysis of changes in mean annual flows as simulated by each SM-PET-HM combination using the CRCM ensemble simulations following the A2 scenario for the two basins can be found in Supplementary materials Table-A II-S10.

Table-A II-4. Changes in annual precipitation (%) and mean temperature (8C) as projected by the CRCM ensemble for the A2 scenario and in MF (%) as simulated by the multimodels over the study basins for the 2041–60 period relative to the 1971–90 period. The range of changes for the ensemble simulations is displayed in parentheses. The values in brackets correspond to the pseudo-observed changes of annual mean flows

	Mistassini basin	Arnaud basin
P	22 (20-24)	44 (39-48)
T	3.2 (3.0-3.3)	3.4 (3.2-3.6)
MF	15 (9-23) [15 (13-16)]	23 (8-44) [27 (25-29)]

II.5 Uncertainty analyses and discussion

II.5.1 Uncertainty in streamflow projections

This section evaluates the uncertainties in streamflow projections by analysing the impacts of model components (SM, PET and HM structure) and of natural climate variability (CRCM 2-member ensemble). The main interest is to estimate the role of the four sources of uncertainty on the projected streamflows. To that end, variance decomposition is performed on the signal of climate change for the three selected hydrologic indicators (MF, HF, LF; see Section II.3.4). The changes in the flow indicators at the 2050 horizon are given in Table-A II-5.

The 15 percentages of variance explained for the three flow indicators over the two basins are presented in Table-A II-6. Of the four primary terms, the HM structure (H) has the highest

values for the three indicators over the Arnaud Basin. The findings are somewhat contrasted for the Mistassini Basin, where the term of PET methods (E) dominates for MF variance and that of natural climate variability (C) dominates for HF and LF variances. The snow models (S) term do not contribute much to MF variance; however, the influence of snow model is slightly amplified for HF and LF variances over the two basins. The *EH* term, which is the interaction between the choice of the PET methods and the choice of the HM structure, is the dominant second interaction term for the three indicators across the basins. The third and fourth interaction terms (*SEH* to *SEHC* in Table-A II-6) are below 10%.

The 15 terms are then aggregated to give *V(S)*, *V(E)*, *V(H)*, *V(C)*. For example, the total variance attributed to the snow models is given by $V(S) = S + SE + SH + SC + SHE + SEC + SEHC$. In this case, the total variances are not equal to 100% because of the interaction terms. From the analysis of Table-A II-7, we can see that the results are not spatially homogenous between the study catchments. For the Mistassini Basin, the uncertainty due to the PET methods is the largest out of the four for the MF projections, whereas the main uncertainty for the HF and LF projections is the natural climate variability. The choice of the HM structure has an impact on the MF and HF projections over this basin, but with a slightly lower intensity. For the northernmost basin, the HM structure clearly dominates for the MF and HF projections. These findings can be partially extended to LF, where uncertainty due to the PET methods is also significant. Uncertainty associated with the natural climate variability is therefore marginal for these basins.

Table-A II-5. As in Table-A II-4, but for changes in MF, HF and LF (%). The values in brackets correspond to the pseudo-observed changes of the flow indicators

Flow indicators	Mistassini basin	Arnaud basin
MF	12 (7 to 22) [1 (-11 to 23)]	21 (7 to 38) [21 (-8 to 49)]
HF	2 (-13 to 34) [0 (-28 to 74)]	15 (-7 to 68) [24 (-25 to 117)]
LF	-5 (-38 to 16) [-17 (-58 to 32)]	27 (-11 to 60) [25 (-66 to 82)]

An explanation for these findings might be that the PET methods have not a strong effect on streamflow projections for the northern basin compared to the southern basin because of low precipitation and temperature (Table-A II-1). Given the low temperature and precipitation values for the northern basin, the variability of temperature and precipitation between the CRCM members can only be small and, by extension, the uncertainty related to the natural climate variability in streamflow projections insignificant. On the contrary, the natural climate variability should influence streamflow projections for the southern basin because of the higher climate variability. This should also imply that the signal of climate change on streamflows should be stronger over the northern basin in comparison to the southern basin (Tables-A II-4 and II-5). Note that the present experiments are based upon two members of the CRCM, limiting the evaluation of natural climate variability. As expected, the snow models do not contribute to uncertainty in the projections of the three hydrologic indicators over the study basins (Troin et al., 2016).

II.5.2 Implications for hydrological impact studies

Evaluating the effects of climate change on streamflows is challenging, as the various modeling stages are associated with uncertainties. Of the five uncertainties that need to be explored in impact studies, the uncertainty associated with hydrological modeling must be considered, since large differences in streamflow change projections are observed between models (Bae et al., 2011; Velázquez et al., 2015a). The uncertainty of model structure is not well understood to date, and additional works are necessary to evaluate the level of uncertainty attributed to each model component. Such an evaluation will contribute to reduce the hydrological modeling uncertainty in impact studies.

The design of the present experiment, conducted in the virtual world, allowed the testing of a larger ensemble of SM-PET-HM combinations. The results indicate that the dominant source of uncertainty for future streamflow simulations is not attributed to the selected SMs for the basins under investigation. Both the DD and DD/EB SMs through the SM-PET-HM

combinations lead to similar streamflow simulations at the 2050 horizon. Our findings are agreement with the previous works of Troin et al. (2015a); Troin et al. (2016). The influence of PET methods on streamflow projections is also explored. The various PET methods provide a large spread of streamflow changes that varies over the study basins. The differences in streamflow change are apparent between the T-based and R-based PET methods through the SM-PET-HM combinations. Higher streamflow changes are expected with the R-based PET methods. Note that the R-based PET methods are also affected by temperature (Table-A II-3). Consequently, the discrepancies on streamflow change following the PET methods are linked to the sensitivity of the PET methods to both solar radiation and temperature changes. This implies that a HM can give a different representation to climate simulations following the selected PET method (T-based versus R-based PET approach). The findings complement those of Bae et al. (2011), where the authors show that the changes on projected streamflows in the Chungju Dan Basin (Korea) is function of the choice of the PET methods. Three different HM structures are also used in this work. We show that the selection of the HM structure directly affects streamflow projections. However, only three lumped conceptual hydrological models are selected for the purpose of this study. If additional different models had been applied in terms of structural variations (e.g., semi-distributed and/or distributed models), then a better quantification of the role of the HM structure on the projected streamflows might have been achieved.

Table-A II-6. Percentage of variance explained by the 15 terms in the decomposition of the climate change signals on MF, HF, and LF over the Mistassini and Arnaud basins for the 2041-60 periods relative to the 1971-90 period

Flow indicators	*S*	*E*	*H*	*C*	*SE*	*SH*	*SC*	*EH*	*EC*	*HC*	*SEH*	*SEC*	*SHC*	*EHC*	*SEHC*
Mistassini basin															
MF	1	43	22	1	1	3	1	10	1	9	3	1	1	1	2
HF	4	7	18	38	1	1	1	14	0	1	4	3	2	1	5
LF	1	9	7	56	1	1	1	11	2	3	2	1	1	3	2
Arnaud basin															
MF	4	15	26	0	3	4	0	18	1	8	9	2	3	2	5
HF	5	3	50	0	4	3	5	7	3	1	6	2	2	5	4
LF	10	22	21	3	4	4	1	9	0	7	9	2	2	1	5

Table-A II-7. Percentage of the total variance explained by the snow models V(S), the PET methods V(E), the hydrological models V(H), and the natural climate variability V(C) in the decomposition of the climate change signals on MF, HF and LF over the Mistassini and Arnaud basins for the 2041–60 period relative to the 1971–90 period

Flow indicators	*V(S)*	*V(E)*	*V(H)*	*V(C)*
		Mistassini basin		
MF	13	61	51	19
HF	20	36	45	51
LF	9	31	30	68
		Arnaud basin		
MF	29	55	74	20
HF	30	34	78	22
LF	36	51	58	22

Considering the medium size of the simulations over two catchments, this work allows the role of the uncertainty sources on projected streamflows to be explored. The analysis of the variance decomposition demonstrated that the uncertainty associated with the choice of the HM structure dominates, followed by the uncertainty due to the choice of the PET methods or that of natural climate variability, depending on the hydrologic indicator and the study catchment. The choice of the HM structure causing more uncertainty than other model components shows the necessity of several numbers of model structures to include the uncertainty of the HM structure. In future studies, investigating the contribution of vertical and horizontal water flows in model structure would further enhance our understanding of structural uncertainty. The other HM component which is affected by a variation in climate conditions is permafrost dynamics. An additional work will be to explore the sensitivity of permafrost dynamics models to that of the SMs and PET methods on streamflow projections over Nordic catchments. Results also show that natural variability of climate has a non-negligible effect on the uncertainty of projected streamflows over the study basins, which is coherent with other multi-model impact works (Booij, 2005; Kay et al., 2009; Seiller & Anctil, 2014; Troin et al., 2015b). However, the results obtained from this work are limited

due to the implementation of the multi-models over just two catchments. More basins are required before the findings of the uncertainty analysis can be generalized to other regions.

II.6 Summary and conclusions

The present study compares multi-combinations of SMs-PET methods-HM structures at simulating pseudo-observed streamflows under current climate, and explores how these combinations of multi-models have an effect on the evaluation of the climate change impacts at the basin scale. Here, seven SMs are combined with five PET methods through three HM structures to simulate pseudo-observed streamflows over two Quebec basins. The investigation is conducted in the virtual world based upon meterological and hydrological data from the CRCM simulations. The relevant findings of this work are synthesized below.

1. Over the historical period (1961-1990), the 105 SM-PET-HM combinations per catchment provide satisfactory performances in simulating streamflows for the calibration and validation. The spring peak flow's timing is generally properly captured by all the ensemble simulations over the two catchments. However, the magnitude of the spring peak flows is more or less adequately simulated according to the HM ensemble simulation over the study basins.
2. Over the future period (2041-2060), the 105 SM-PET-HM combinations per catchment have acceptable performances at simulating streamflows, comparable to those obtained for the historical period. While all the ensemble simulations provide satisfactory simulations of the spring peak flow's timing, again, none however captures its magnitude.
3. The projections of future climate show that important changes on virtual streamflows are likely to happen in this century. For the Mistassini Basin, the timing of the future spring peak flow is supposed to coincide with that of the present day, but with a decreased magnitude. However, over the Arnaud Basin, it is expected an earlier spring peak flow than for the historical period, with an increased magnitude.

4. A ranking of the uncertainty sources related to model components and natural climate variability in streamflow projections is provided. Of the analysis of specific hydrologic indicators using the decomposition of variance, it is shown that the selection of the HM structure contributes the most to uncertainty on the projected streamflows, followed by the choice of the PET methods or the natural climate variability, depending on the hydrologic indicator and the catchment. The SMs lead to the lowest level of uncertainty to the total uncertainty linked to model components.

This work gives pathways to assess the uncertainty of HM structure in impact studies. Exploring the HM structure uncertainty in greater detail by analysing the uncertainty associated with both the horizontal and vertical water flow schemes will help to increase the understanding of model structure and, by extension, to reduce uncertainty on streamflow projections.

II.7 Acknowledgements

This study was funded by École de technologie supérieure (ETS). Ouranos supplied the CRCM simulations for this work. The authors also thank Hydro-Québec and IRSTEA for providing some snow models used in this study.

II.8 Supplemental material

The percent deviation (D) of daily streamflow (relative error) is computed in the reference period in the virtual world as:

$$D_{ref}(\%) = \frac{\bar{Q}_{sim}^{ref} - \bar{Q}_{pseudo-obs}^{ref}}{\bar{Q}_{pseudo-obs}^{ref}} \cdot 100 \qquad \text{(II.S1)}$$

where $\bar{Q}_{sim}^{ref}$ and $\bar{Q}_{pseudo-obs}^{ref}$ are the means of daily simulated and pseudo-observed streamflows over the 20-year reference period, respectively.

Table-A II-S8. Percent deviation (%) of streamflow for each SM-PET-HM combination using the CRCM ensemble simulations for the 1961-1990 period over the two basins

			The Mistassini Basin			The Arnaud Basin		
SMs	PET methods	SM-PET combination	HM structure			HM structure		
			MOHYSE	HMETS	GR4J	MOHYSE	HMETS	GR4J
DD	T-based	MOH-HAM	-16	-13	-11	-4	0	-2
	T-based	MOH-MOH	-13	-12	-10	-5	1	-2
	R-based	MOH-OUD	-10	-10	-8	-2	1	-1
	R-based	MOH-PM	-10	-8	2	-1	7	4
	R-based	MOH-PT	-12	-9	4	-2	6	2
DD	T-based	HBV-HAM	-11	-12	-11	-6	-1	-2
	T-based	HBV-MOH	-13	-11	-7	-6	5	-1
	R-based	HBV-OUD	-11	-10	-10	-2	-1	-2
	R-based	HBV-PM	-14	-9	2	-3	2	0
	R-based	HBV-PT	-13	-10	4	-3	2	1
DD	T-based	CEM-HAM	-16	-11	-12	-4	-2	-3
	T-based	CEM-MOH	-14	-10	-10	-5	-2	-2
	R-based	CEM-OUD	-12	-8	-7	-3	-2	-1
	R-based	CEM-PM	-13	-11	1	-3	-1	3
	R-based	CEM-PT	-13	-8	3	-3	3	3
DD	T-based	ETS-HAM	-15	-10	-12	-3	-1	-3
	T-based	ETS-MOH	-14	-10	-8	-3	-1	-2
	R-based	ETS-OUD	-11	-9	-9	-2	-1	-2
	R-based	ETS-PM	-13	-9	1	-3	-1	0
	R-based	ETS-PT	-13	-9	3	-3	4	0
DD/EB	T-based	HYD-HAM	-16	-8	-12	-4	-1	-2
	T-based	HYD-MOH	-16	-8	-8	-5	-1	-2
	R-based	HYD-OUD	-11	-9	-8	-2	4	-2
	R-based	HYD-PM	-13	-11	2	-3	-1	0
	R-based	HYD-PT	-11	-10	3	-2	5	0
DD/EB	T-based	SEB-HAM	-15	-15	-11	-6	-4	-2
	T-based	SEB-MOH	-12	-12	-10	-5	1	-2
	R-based	SEB-OUD	-10	-6	-9	-2	3	1
	R-based	SEB-PM	-12	-11	-1	-1	-2	2
	R-based	SEB-PT	-11	-9	2	-2	7	0
DD/EB	T-based	ETI-HAM	-15	-9	-13	-5	-3	-3
	T-based	ETI-MOH	-15	-7	-10	-5	-2	-2
	R-based	ETI-OUD	-11	-8	-8	-2	1	2
	R-based	ETI-PM	-14	-6	2	-1	3	1
	R-based	ETI-PT	-13	-4	2	-1	4	1
Summary	SMs	DD	-13^{-10}_{-16}	-10^{-8}_{-13}	-5^{4}_{-12}	-3^{-1}_{-6}	1^{7}_{-2}	-1^{4}_{-3}
		DD/EB	-13^{-10}_{-16}	-9^{-4}_{-15}	-5^{3}_{-13}	-3^{-1}_{-6}	1^{7}_{-4}	-1^{2}_{-3}
	PET methods	T-based	-14^{-11}_{-16}	-10^{-7}_{-15}	-10^{-7}_{-13}	-5^{-3}_{-6}	-1^{5}_{-4}	-2^{-1}_{-3}
		R-based	-12^{-10}_{-14}	-9^{-4}_{-11}	-1^{4}_{-10}	-2^{-1}_{-3}	2^{7}_{-2}	1^{4}_{-2}
	HM		-13^{-10}_{-16}	-9^{-4}_{-15}	-5^{4}_{-13}	-3^{-1}_{-6}	1^{7}_{-4}	-1^{4}_{-3}

Table-A II-S9. Percent deviation (%) of streamflow for each SM-PET-HM combination using the CRCM ensemble simulations for the 2041-2060 period over the two basins

			The Mistassini Basin			The Arnaud Basin		
SMs	PET methods	SM-PET combination	HM structure			HM structure		
			MOHYSE	HMETS	GR4J	MOHYSE	HMETS	GR4J
DD	T-based	MOH-HAM	-19	-15	-11	-7	-1	5
	T-based	MOH-MOH	-16	-13	-10	-4	-2	5
	R-based	MOH-OUD	-12	-11	-9	-5	-2	2
	R-based	MOH-PM	-12	-6	5	-4	9	18
	R-based	MOH-PT	-13	-5	9	-3	7	6
DD	T-based	HBV-HAM	-14	-15	-12	-12	-4	-2
	T-based	HBV-MOH	-17	-13	-7	-11	-3	-2
	R-based	HBV-OUD	-14	-11	-11	-10	-5	-1
	R-based	HBV-PM	-15	-7	5	-7	-3	15
	R-based	HBV-PT	-13	-7	11	-6	-3	-2
DD	T-based	CEM-HAM	-20	-15	-12	-20	-5	-2
	T-based	CEM-MOH	-17	-13	-10	-7	-4	-2
	R-based	CEM-OUD	-14	-10	-8	-7	-4	-4
	R-based	CEM-PM	-15	-9	3	-8	-3	7
	R-based	CEM-PT	-14	-5	11	-6	-1	3
DD	T-based	ETS-HAM	-19	-14	-13	-10	-3	4
	T-based	ETS-MOH	-17	-13	-8	-9	-2	3
	R-based	ETS-OUD	-13	-11	-10	-6	-4	5
	R-based	ETS-PM	-15	-8	2	-6	2	13
	R-based	ETS-PT	-13	-6	8	-5	2	5
DD/EB	T-based	HYD-HAM	-19	-12	-12	-7	-5	2
	T-based	HYD-MOH	-18	-12	-8	-9	-3	1
	R-based	HYD-OUD	-13	-11	-9	-7	-2	1
	R-based	HYD-PM	-14	-8	6	-6	-3	15
	R-based	HYD-PT	-12	-7	6	-5	9	10
DD/EB	T-based	SEB-HAM	-19	-16	-12	-9	-3	1
	T-based	SEB-MOH	-16	-14	-10	-10	-4	-1
	R-based	SEB-OUD	-12	-9	-10	-8	-4	4
	R-based	SEB-PM	-13	-9	7	-7	-9	7
	R-based	SEB-PT	-12	-5	7	-9	5	7
DD/EB	T-based	ETI-HAM	-18	-14	-13	-10	-8	0
	T-based	ETI-MOH	-17	-11	-10	-10	-5	1
	R-based	ETI-OUD	-13	-11	-9	-7	5	2
	R-based	ETI-PM	-15	-5	9	-6	-1	9
	R-based	ETI-PT	-12	-2	9	-5	5	9
Summary	SMs	DD	-15^{-12}_{-20}	-10^{-5}_{-15}	-3^{11}_{-13}	-8^{-3}_{-20}	-1^{9}_{-5}	4^{18}_{-4}
		DD/EB	-15^{-12}_{-19}	-10^{-2}_{-16}	-3^{9}_{-13}	-8^{-5}_{-10}	-1^{9}_{-9}	4^{15}_{-2}
	PET methods	T-based	-17^{-14}_{-20}	-13^{-11}_{-16}	-10^{-7}_{-13}	-10^{-4}_{-20}	-4^{-1}_{-8}	1^{5}_{-2}
		R-based	-13^{-12}_{-15}	-8^{-2}_{-11}	2^{11}_{-11}	-6^{-3}_{-10}	0^{9}_{-9}	6^{18}_{-4}
	HM		-15^{-12}_{-20}	-10^{-2}_{-16}	-3^{11}_{-13}	-8^{-3}_{-20}	-1^{9}_{-9}	4^{18}_{-4}

Table-A II-S10. Changes in annual mean flows (MF) (%) as simulated by each SM-PET-HM combination using the CRCM ensemble simulations for the A2 scenario over the two basins

			The Mistassini Basin			The Arnaud Basin		
SMs	PET methods	SM-PET combination	HM structure			HM structure		
			MOHYSE	HMETS	GR4J	MOHYSE	HMETS	GR4J
DD	T-based	MOH-HAM	11	12	14	18	24	25
	T-based	MOH-MOH	11	13	15	24	24	26
	R-based	MOH-OUD	13	13	14	19	22	22
	R-based	MOH-PM	14	18	19	22	26	37
	R-based	MOH-PT	14	20	21	22	25	30
DD	T-based	HBV-HAM	12	11	14	16	20	24
	T-based	HBV-MOH	12	12	15	16	20	25
	R-based	HBV-OUD	13	14	14	19	19	20
	R-based	HBV-PM	14	18	19	21	18	37
	R-based	HBV-PT	14	20	20	22	19	31
DD	T-based	CEM-HAM	10	10	14	11	23	25
	T-based	CEM-MOH	11	11	15	19	23	26
	R-based	CEM-OUD	13	13	14	22	22	21
	R-based	CEM-PM	14	18	21	20	22	37
	R-based	CEM-PT	14	19	20	21	20	33
DD	T-based	ETS-HAM	10	10	14	18	23	25
	T-based	ETS-MOH	12	11	13	18	23	26
	R-based	ETS-OUD	13	13	14	21	20	21
	R-based	ETS-PM	14	17	20	23	28	34
	R-based	ETS-PT	14	19	20	23	22	27
DD/EB	T-based	HYD-HAM	11	10	14	20	23	25
	T-based	HYD-MOH	12	11	15	20	22	26
	R-based	HYD-OUD	13	13	14	17	17	20
	R-based	HYD-PM	13	18	19	23	22	36
	R-based	HYD-PT	14	19	22	22	19	32
DD/EB	T-based	SEB-HAM	11	13	15	20	23	30
	T-based	SEB-MOH	11	13	15	18	24	26
	R-based	SEB-OUD	13	12	14	20	21	25
	R-based	SEB-PM	14	19	18	21	26	35
	R-based	SEB-PT	14	21	21	22	22	36
DD/EB	T-based	ETI-HAM	11	10	14	14	21	24
	T-based	ETI-MOH	12	11	15	18	22	25
	R-based	ETI-OUD	13	12	13	16	18	20
	R-based	ETI-PM	14	16	20	20	22	34
	R-based	ETI-PT	16	17	22	24	23	21
Summary	SMs	DD	17^{21}_{14}	15^{20}_{10}	13^{14}_{10}	20^{24}_{11}	22^{28}_{18}	28^{37}_{20}
		DD/EB	17^{22}_{13}	15^{21}_{11}	13^{16}_{11}	20^{24}_{14}	22^{26}_{17}	28^{36}_{20}
	PET methods	T-based	15^{15}_{14}	11^{13}_{10}	11^{12}_{10}	18^{24}_{11}	23^{24}_{20}	26^{30}_{24}
		R-based	18^{22}_{13}	17^{21}_{12}	14^{16}_{13}	21^{24}_{16}	22^{28}_{17}	29^{37}_{20}
	HM		17^{22}_{13}	15^{21}_{10}	13^{16}_{10}	20^{24}_{11}	22^{28}_{17}	28^{37}_{20}

ANNEXE III

THE HAZARDS OF SPLIT-SAMPLE VALIDATION IN HYDROLOGICAL MODEL CALIBRATION

Richard Arsenault[a], François Brissette[b], Jean-Luc Martel[c]

[a, b, c] Département de Génie de la construction, École de technologie supérieure

Article publié dans « Journal of Hydrology », octobre 2018.

Abstract:

This paper investigates the issues related to the use of validation in hydrological model calibration. Traditionally, models are calibrated and then assessed on an independent period (split-sample) to determine their adequacy in simulating streamflow as compared to observations. In this study, two hydrological models and three North American catchments are used to evaluate the effects of using validation to assess the model parameters' robustness on the model's actual simulation capabilities and accuracy in simulating streamflow. The length of the calibration period is increased from 1 to 16 years, and for each case a large number of randomly selected combinations of years are used for calibration and for validation using the Nash-Sutcliffe Efficiency metric. The calibrated model is then run on an independent 8-year test-period to assess the model's actual performance in simulation mode in unknown conditions. The process is bootstrapped 30 times to ensure the robustness of the results. The tests pit the calibration/validation methods on increasing calibration period lengths against a full calibration on the entire available dataset. Results show that the calibration on the full dataset is the optimal strategy as it generates the most robust parameter sets, provides the best model accuracy on an independent testing period and does not require assumption making on the modeler's part. The calibrated parameter sets for each test-case were evaluated using the relative bias and correlation metrics, which revealed that the method transfers well to these two other metrics. Results also demonstrate the pitfalls of the commonly used split-sampling strategy, where good parameter sets may be discarded due to model performance discrepancies between calibration and validation periods. The

conclusions point to the need to use as many years as possible in the calibration step and to entirely disregard the validation aspect under certain conditions.

Highlights

- Calibration on full time series is shown to be more robust than split-sample methods
- 30 Bootstrapping tests on 6 cases provide evidence towards this method being optimal
- Verification on 10 independent catchment-model pairs support the conclusions
- Caveats of split-sampling on model performance are demonstrated
- Length of the calibration period is proportional to the parameter set robustness

Keywords

Model validation, model calibration, model performance, hydrological modeling, split-sample testing

III.1 Introduction

Hydrological models are amongst the most commonly used environmental models. They are commonly used for streamflow forecasting (e.g. Arsenault, Latraverse, & Duchesne, 2016b; Day, 1985; Nash & Sutcliffe, 1970; Wood, Maurer, Kumar, & Lettenmaier, 2002), climate change impact studies (Bergström et al., 2001; Chen et al., 2011a; Chen et al., 2011b; Jiang et al., 2007; Middelkoop et al., 2001; Minville et al., 2008) and to better understand processes at various scales on catchments covering a wide array of sizes, from small urban parcels all the way up to the continental scale (Band, Patterson, Nemani, & Running, 1993; Fatichi et al., 2016; Gupta et al., 2014; Singh & Woolhiser, 2002). Many of the physical processes governing water and energy balance at the parcel or catchment scale are not well understood or in need of unavailable data and, as a result, all hydrological models require

some level of parameterization (Andreassian et al., 2006; Jakeman & Hornberger, 1993). Consequently, a calibration and validation process is generally needed to assess the hydrological model performance on any given catchment.

There is confusion in the literature about the precise meaning of model validation. Model calibration, validation, evaluation and even verification have been used interchangeably in some papers to describe if a model is adequate at representing observations. In particular, the term "validation" has been used to mean different things by different authors. In this paper the definitions of (Refsgaard, Henriksen, Harrar, Scholten, & Kassahun, 2005) are used since they are closest to the general agreement in the hydrological community.

III.1.1 Definitions

In the modeling community, model validation, model verification or model evaluation is the process by which a comparison is made between model outputs and observations to evaluate the adequacy of the model (Legates & McCabe, 1999). Some authors differentiate verification (computer code accuracy in solving equations) from validation (ability of the model to represent the underlying physical process). Model verification and validation should be seen as the process for which one verifies if a model is adequate and with the assumption of similar performance if the model is used in similar conditions (Klemeš, 1986). Furthermore, a distinction can be made between various types of validation (e.g. scientific vs performance validation as discussed in Biondi, Freni, Iacobellis, Mascaro, et Montanari (2012). Performance validation (ability to represent observations) is the one that is typically made for hydrological models. A review of environmental performance evaluation is presented in Bennett et al. (2013). In the hydrological modeling community, validation has a much stricter sense ascribed to the process of demonstrating the model ability to perform outside of its training period. This is the definition of Refsgaard et al. (2005) which is adopted in this paper for the sake of consistency with the work of many other authors in hydrology.

Model calibration consists in adjusting model parameters over a training period, either manually or automatically (Arsenault et al., 2014; Boyle, Gupta, & Sorroshian, 2000; Duan, Sorooshian, & Gupta, 1994; Gupta et al., 2014), so that model outputs match observations as closely as possible. The adequacy of parameters adjustment is typically based on a single objective function representing the similarity between model outputs and observations. An objective function is a mathematical equation to be minimized in order to insure maximal similarity between model outputs and observations. In hydrological modeling, the most commonly used objective function is the Nash-Sutcliffe Efficiency metric (NSE; Nash & Sutcliffe, 1970) which minimizes the root mean square error between modeled and observed streamflows. Many other objective functions have also been proposed (e.g. Garcia, Folton, & Oudin, 2017; Gupta et al., 2009; Moriasi et al., 2007). It should be noted that model calibration implies some level of model verification (or validation in its wide sense) since a bad performance would normally stop the modeling process and the need for validation (in its stricter sense as used in this paper).

Following calibration, a validation is typically performed over a different period to ensure parameter transferability and model robustness. This practice of using two different periods for calibration and validation is referred to as the split-sample approach.

III.1.2 Split-sample approach

The split-sample approach is a classic method which is central to Klemeš (1986) hierarchical scheme for validating hydrological models. It has been implemented in many ways over the years, with each variant aiming to target a specific calibration goal. The most common split-sample approach in the literature is the two-period method, by which the calibration and validation periods are split in two sections of approximately equal length. This method has the advantage of being easy to implement and minimizes the model runtime which is very helpful when the hydrological model is computationally intensive. For similar reasons, sometimes the calibration period is longer than the validation period or vice-versa. Tolson et

Shoemaker (2007) implemented a 6-year calibration and two independent validation series of 3 years and 1 year respectively on the Cannonsville Reservoir catchment using the SWAT2000 model. They used these two different validation periods to verify the model's adequacy on contrasting hydrological conditions and found that the validation *NSE* was higher than the calibration *NSE*, leading to the conclusion that the parameter optimization was robust. Arnold et al. (2012) recommends using a long-enough period to encompass varying conditions, e.g. dry/wet years. Larabi, St-Hilaire, et Chebana (2018) used eight calibration years followed by four validation years in an implementation of a Functional Data Analysis (FDA) based calibration scheme. Liu et al. (2018) specifically tackle the challenges related to sample-splitting and provide an in-depth analysis of the shortcomings of using this approach.

The second option is a mixed-bag approach by which calibration and validation years are sampled randomly throughout the dataset, sometimes using bootstrapping methods to resample and draw robust parameter sets, other times by identifying unusual events to use during calibration (Razavi & Tolson, 2013; Singh & Bárdossy, 2012). This method has the advantage of generating multiple parameter sets that can be analyzed against the calibration and validation performance, but the tradeoff is that the computing time for the calibration aspect grows linearly with the number of calibration years and the length of time over which they are spread out. In the case where some calibration years are discontinuous, the model must still be run on the entire time-series between the first and last calibration year to ensure that the model states are consistent in time. Wallner, Haberlandt, et Dietrich (2012) used this method to calibrate the HEC-HMS hydrological model on five catchments in Germany for evaluating regionalization approaches. They calibrated on the years 2004, 2007 and 2008, and the validation was performed on the 2005-2006 period. However, no reason was given as to why these years were selected instead of using a half-and-half method, for example. Singh et Bárdossy (2012) compare the calibration performance using the entire dataset to a subset of unusual events, and find that the results are only marginally worse than when using the full dataset. Razavi et Tolson (2013) implemented a method that selects surrogate years

representative of the entire time series and terminates the model simulation if the objective function does not seem promising after the first few evaluations, immediately switching to the new parameter set to be evaluated. Gharari, Hrachowitz, Fenicia, et Savenije (2013) divided the time series in multiple small blocks (e.g. by week, by month, by wetness, etc.) and performed calibration of the HyMod model on each block. The best parameter sets are discovered by analyzing the parameter distributions for each subset and are then validated on independent periods.

Finally, the last split-sample method presented here is the odd-even approach. This method was developed as a response to calibration under non-stationary conditions. In essence, the method is based on calibrating the hydrological model on the odd years of the dataset and the validation is performed on the even years (or vice-versa). In this setup, the hydrological model is exposed to the non-stationary trend in both the calibration and validation steps, which in theory should allow the model to integrate the trend information in the parameter set and then use this information on the validation period. Essou et al. (2016) and Arsenault et al. (2017) used the odd/even split-sample testing to take into consideration the trends arising in long climate time-series during calibration. Gowda et al. (2012) used the same technique when calibrating the ADAPT model because the first and last halves of the time series were respectively wetter and drier than the average. Using the odd-even method ensured sampling the entire spectrum of available values.

It is also important to note that in this paper, the split-sample approach is used solely for calibration and validation of a hydrological model. In some papers (i.e. Butts, Payne, Kristensen, & Madsen, 2004), split-sampling is used to compare different hydrological model structures. This application of split-sampling is out of the scope of this paper and is thus not investigated.

III.1.3 Length of the calibration period

The question of the length of the calibration period has been the subject of many studies (Razavi & Tolson, 2013). Vrugt et al. (2006) found that increasing the length of the calibration dataset improved the performance of the SAC-SMA hydrological model. Juston, Seibert, et Johansson (2009) and Perrin et al. (2007) investigated various sampling strategies to select a calibration dataset and found that sampling could be optimized so that calibration performance may reach that of calibrating over the full period. Razavi et Tolson (2013) came to similar conclusions using a surrogate shorter calibration period. In both papers, calibration over the entire dataset was considered the benchmark for comparison. van der Spek et Bakker (2017) looked at the length of the calibration period on the performance of a groundwater hydraulics model. They found the length of the calibration period to be more important than the frequency of observations. For an equal number of observations, a long record was clearly superior to a shorter one with a higher observation frequency.

Although a few authors imply that calibration over the full length of the available dataset may be preferable for parameter identifiability (e.g. Singh & Bárdossy, 2012) the overwhelming majority of published work still advocates the split sample strategy (e.g. Daggupati et al., 2015; Gaborit, Ricard, Lachance-Cloutier, Anctil, & Turcotte, 2015; Garavaglia et al., 2017; Gaur et al., 2017; Liu et al., 2018; Moriasi et al., 2015; Newman et al., 2015).

III.1.4 Research objectives

In this paper it is argued that in gray-box hydrological model (the type most commonly used in hydrology) where calibration is needed (whether manual or automatic) and which uses a standard “performance” model validation only, in those cases, validation is not needed and in fact detrimental to model performance. Refsgaard et al. (2005) called upon the community to investigate and develop better model validation methods, but in parallel, other authors hint to

the idea that model calibration on the entire period is the best option to obtain a robust, time-transferable parameter set (Singh & Bárdossy, 2012). In this paper, the issue is investigated in detail and the effect of using validation on model performance on an independent testing period is evaluated. More specifically, the question as to which option is the optimal choice between using a validation period and calibrating on the entire time-series (with no validation) is answered.

III.2 Study area and data

Three catchments' metadata (drainage area and outline) and streamflow time series from the MOPEX (Duan et al., 2006) and CANOPEX (Arsenault et al., 2016a) databases covering North America were considered in this study. The meteorological data needed as input for the hydrological models (i.e. precipitation and temperature) were obtained from different observed datasets. For the two catchments located in the United States, daily precipitation and temperature were extracted from the gridded dataset (interpolated on a 0.125° x 0.125° grid) of the University of Santa Clara (Maurer et al., 2002). Regarding the third catchment located within Canada, precipitation and temperature were obtained from Natural Resources Canada (NRCAN) gridded datasets (interpolated on a 0.083° x 0.083° grid; Hutchinson et al., 2009). Figure-A III-1 shows the location of the study sites as well as their elevation profiles. These catchments were selected based on four criteria:

1. At least 25 years of continuous data must be available with little to no missing values, i.e. less than 5%, in both the hydrometric and meteorological observation records;
2. There must be at least some snowfall on the catchment on a regular basis;
3. The hydrological models should perform to different levels based on the Nash-Sutcliffe Efficiency (NSE; Nash & Sutcliffe, 1970) to explore cases where the models would be more or less challenged to produce good hydrographs.
4. The observation period must not contain a trend in observed streamflow as detected by a Mann-Kendall test (Kendall, 1975) at a 95% confidence level.

The latter condition was imposed in order to allow random sampling of years whilst considering them as independent.

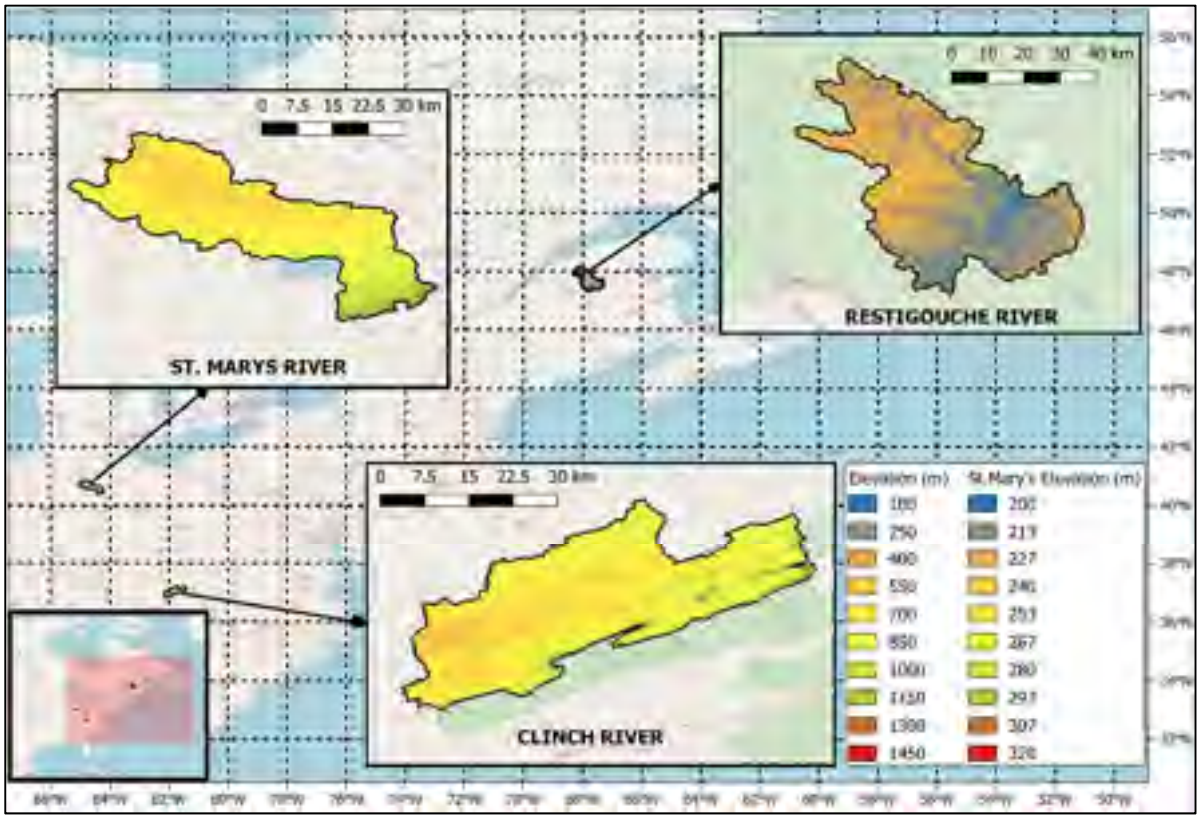

Figure-A III-1 The three selected catchments' locations and elevation profiles (m). The St. Mary's River catchment is much flatter than the other two catchments, which explains why it has its own elevation legend

Time and computing resources constraints limited the number of catchments that could be explored in this study; therefore, the catchments were selected in such a way as to maximize the variability in the test cases. Table-A III-1 shows the selected catchment physical and hydrometeorological characteristics.

Table-A III-1. The three selected catchments' characteristics and mean annual hydrometeorological variables

	Restigouche River	Clinch River	St. Mary's River
Catchment characteristics:			
Drainage area (km^2)	3160	1380	1608
Average altitude (m)	343	760	258
Latitude centroid (degrees)	47.7	37.0	30.6
Mean hydrometeorological variables:			
Annual total precipitation (mm)	1070	1146	975
Rainfall	693	1046	888
Snowfall	377	100	87
Annual daily temperature (°C)			
Minimum	-2.6	3.7	5.6
Maximum	8.2	17.5	16.1
Mean	2.8	11.1	10.8
Annual daily streamflow (m^3s^{-1})	70	19	17

For each catchment, a 25-year period (1986-2010) of data was selected to standardize the tests as detailed in the methodology. This was to allow a 1-year model warm-up period, an 8-year test period and 16 years for conducting the traditional model calibration and validation steps.

Finally, although the most intensive analysis in this paper was performed using these three catchments, the dataset was extended to eight catchments in the discussion section of this paper.

III.3 Methods

In this section, the hydrological models are described, as are the main methodological steps and underlying hypotheses.

III.3.1 Hydrological models

Two lumped hydrological models of varying complexity, namely the 6-parameter GR4J-CN model and the 21-parameter HMETS model were used in order to provide a stronger basis to generalize the results and evaluate the proposed methodology robustness. The GR4J-CN and HMETS models' structures are described in the following sections.

III.3.1.1 GR4J-CN

The GR4J (standing for modèle du Génie Rural à 4 paramètres Journaliers – daily rural engineering model with four parameters; Perrin et al., 2003) model is a simple parsimonious 4-parameter lumped model operating at a daily scale. This rainfall-runoff model structure is composed of a production and a routing store.

Since the simulation of the snow processes in a requirement in this work, the GR4J model has been coupled with the CemaNeige degree-day snow model (Valéry, 2010). CemaNeige is a 2-parameter module that simulates the evolution of the snow cover and the different processes leading to snowmelt.

This results in a 6-parameter hydrological model referred to as GR4J-CN in this work. The required inputs for the GR4J-CN model are continuous time series of daily precipitation, mean temperature and potential evapotranspiration. Potential evapotranspiration has been estimated based on the work of Oudin et al. (2005) and the resulting formula is based on extraterrestrial radiation.

This model structure (GR4J-CN with potential evapotranspiration estimated with the Oudin formula) has been used in numerous hydrological studies and has shown strong performance in the simulation of daily streamflow over North American catchments (Poissant et al., 2017; Troin et al., 2015a; Troin et al., 2018; Velázquez et al., 2015b).

III.3.1.2 HMETS

The HMETS (Hydrological Model – École de technologie supérieure; Martel et al., 2017) is a simple and efficient model that has been developed for educational applications. This lumped-conceptual model uses two connected reservoirs to simulate the vadose and phreatic zones. Streamflow is computed as the sum of surface, delayed, hypodermic and groundwater flows.

Here the snowmelt processes are simulated by the 10-parameter degree-day model developed by Vehviläinen (1992a). This model simulates the evolution of the snowpack notably through the melting and the refreezing processes. Similarly to the GR4J-CN model, the Oudin formula (Oudin et al., 2005) has been used to estimate the potential evapotranspiration required by HMETS.

HMETS has up to a total of 21 free parameters that can be optimized during the calibration. While several of these parameters can be initially fixed for a more parsimonious model, all 21 parameters were kept for the calibration in this work, to offer an opposite counterpart to the parsimonious GR4J-CN model. The needed inputs consist of daily precipitation, minimum and maximum temperature, as well as the potential evapotranspiration.

The HMETS model with the Oudin formula has also been applied over multiple hydrological studies using North American catchments and has shown to provide a good performance (Troin et al., 2015a; Troin et al., 2018).

III.3.2 Model calibration algorithm and objective function

In this project, the Covariance-Matrix Adaptation Evolution Strategy (CMA-ES; Hansen & Ostermeier, 1996, 2001) was implemented to automate the hydrological model parameter calibration step. CMA-ES was shown to be a robust algorithm that requires little to no

hyperparameter tuning, making it ideal for this project (Arsenault et al., 2014). CMA-ES's strength comes from the internal learning of a second-order model that represents the objective functions' response surface. It does not require any assumptions on response surface convexity or smoothness and performs well even on ill-conditioned functions. Its robustness was therefore a strong asset for this project due to the large impact that the optimization algorithm performance could have on the end-result. A budget of 10 000 model evaluations was selected based on the work of Arsenault et al. (2014) to ensure convergence towards an optimal parameter set.

In all cases, the *NSE* metric was used to calibrate, evaluate and test the hydrological model performance. While it does have some drawbacks, such as weighting the peak flows more heavily, it is still considered as a good overall objective function in hydrological model simulation (McCuen et al., 2006). The objective function used in CMA-ES, which attempts to minimize an objective function, was unity minus the *NSE* value (1-*NSE*). In practice, the optimization algorithm attempts to attain the minimum possible value of zero, returning the perfect *NSE* value of 1. The *NSE* itself is calculated as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{o,i} - Q_{s,i})^2}{\sum_{i=1}^{n}(Q_{o,i} - \overline{Q_{o,i}})^2} \quad \text{(III.1)}$$

where *NSE* is the Nash-Sutcliffe Efficiency metric, Q_o is the observed streamflow, Q_s is the simulated streamflow and the i index represents the simulation day. The optimization algorithm then used the [1-*NSE*] value to perform the optimization itself. The *NSE* was preferred over other metrics such as Kling-Gupta Efficiency (KGE; Gupta et al., 2009) simply due to the prevalence of *NSE* in the literature (Jain & Sudheer, 2008), which allows for a better visualization of the scores from the hydrological modeling community. The results are not expected to be influenced by this choice.

III.3.3 Strategy to evaluate the impact of calibration and validation

Typically, hydrologists will divide their observation database in two parts: a calibration period for parameter tuning, and a second period for the hydrological model validation. This separation follows many forms, including the split-sample or half-half method (Klemeš, 1986) and the odd/even year method (Arsenault et al., 2017; Gowda et al., 2012). Usually, the parameter set that is found during calibration is first evaluated on the calibration period, and on the validation period in a second step. The objective functions between both periods are then compared. An important drop in performance between the calibration and validation periods will raise flags and over-parameterization or data quality problems might be suspected. In any case, another parameter set might be evaluated using another sample until a robust parameter set is found for both the calibration and validation periods. It is also customary to switch the calibration and validation periods to verify if the underlying observed data are at cause. Once the hydrologist is satisfied with the calibrated parameter set, the model can then be used for simulation and forecasting applications.

The main drawback of using such sample-based calibration and validation methods is that the parameter set is conditioned on a subset of the available data, possibly depriving itself from important information that is lost in the validation phase. This is the research question this work aims to answer.

To see if that is truly the case, a strategy commonly used in neural-network calibration and testing was implemented for hydrological modeling (Guo, Pleiss, Sun, & Weinberger, 2017), in which the data are separated into three parts rather than two. The third part is reserved for testing the data on a period independent from the calibration and validation periods. In the case of this study, the 25 years of data were divided according to the method detailed in the pseudocode of Figure-A III-2.

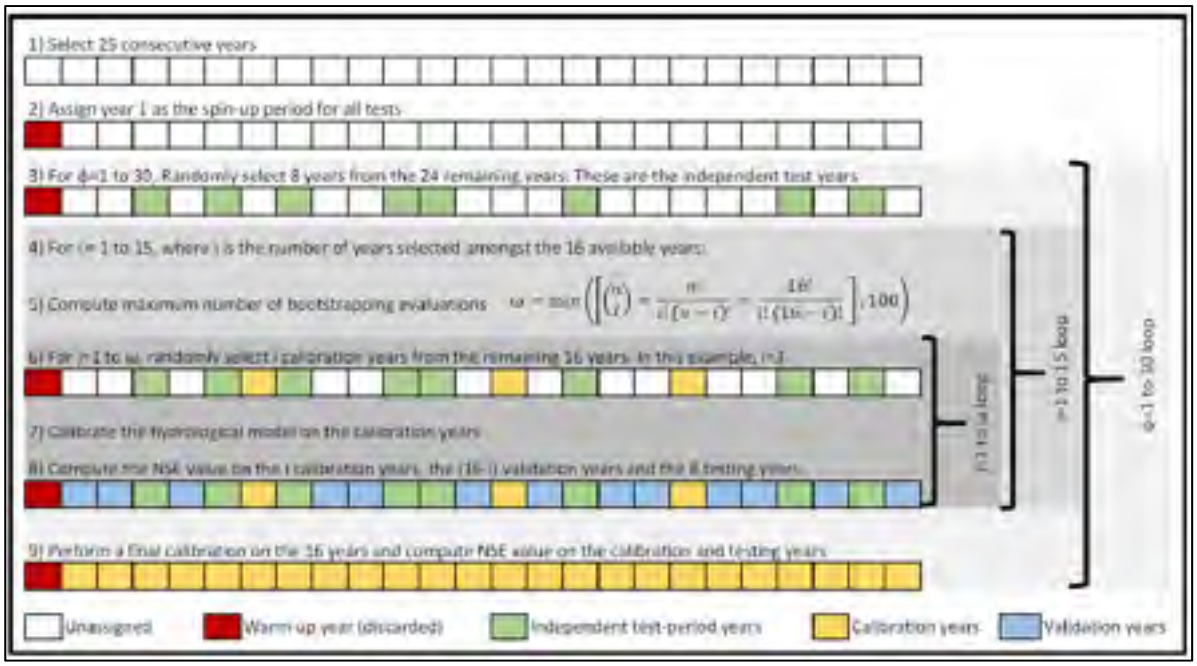

Figure-A III-2 Pseudocode of the methodology implemented in this study

The nine steps are carried out as follows:

STEP 1: For each catchment and hydrological model used in this study, 25 years of continuous streamflow and their corresponding input data are prepared.

STEP 2: The hydrological models are always run starting at the beginning of the first year and the *NSE* values never include the first year as it is considered a warm-up period to bring the model states to realistic internal variable states. Therefore, this year is always simulated by the hydrological model, but is never computed in the *NSE* scores. The warm up period of one year was considered sufficient so that the uncertainty of initial conditions becomes negligible compared to the uncertainty of parameters (Huard & Mailhot, 2008).

STEP 3: To ensure robustness of the results, the entire process (Steps 3-9) is repeated 30 times. For each $\varphi \in [1,30]$, a new random set of 8 independent test years is selected.

STEP 4: Given that there are 16 remaining years to perform calibration and validation, $i \in [1,15]$ years are selected from the 16 years to evaluate the impact of the number of calibration years on the performance over the independent test period. Here i is limited to 15 as the calibration on the entire 16-year set is performed later, at step 9.

STEP 5: To make sure that the process is not affected by random "luck of the draw", multiple trials are performed for each i. When $i = 1$, there are a total of 16 possible selections of 1 year from a set of 16, but when $i = 8$ years from 16 must be selected, then there are 12870 possible combinations. It would be extremely computationally intensive to perform the entire set of combinations in calibration, keeping in mind that the process must be repeated for each i and each φ for three catchments and two hydrological models. Therefore, to keep computing time reasonable and still explore the parameter space, the maximum number of bootstrap evaluations ω is fixed at 100.

STEP 6: At this step, i years are selected at random in the available 16 years. These will be used for calibrating the hydrological model. This step, along with steps 7-8, are repeated ω times.

STEP 7: With the i years selected, the model calibration is performed following the method described in Section III.3.2.

STEP 8: With the calibrated model parameters, compute the *NSE* metric values on the i calibration years, the 8 testing years and the (16-i) validation years.

STEP 9: Once all i and j are exhausted, the hydrological model is calibrated on the 16 years and compute the *NSE* scores on the calibration and testing periods. Obviously since all years are used in calibration, there are no remaining years to perform the validation step.

The results were then analyzed by comparing the *NSE* values on the independent testing periods as a function of the number of calibration years. In total, 239940 calibration experiments were performed in this work following the strategy described above.

III.4 Results

The first results aim to display the evolution of the *NSE* over the independent testing period as a function of the number of calibration years. Figure-A III-3 shows the distributions of *NSE* values obtained on the test period for one of the 30 independent testing periods φ. Each of the boxplots represents all of the *NSE* values ω for each combination of years *i* used during calibration. For example, the first boxplots hold 16 points (16 combinations of 1 year from a set of 16), the last boxplot also holds 16 points (16 combinations of 15 years from a set of 16) and all boxplots in between are limited to 100 combinations. The horizontal line that cuts through the boxplots presents the *NSE* value on the independent test period when the 16 years are used for calibration.

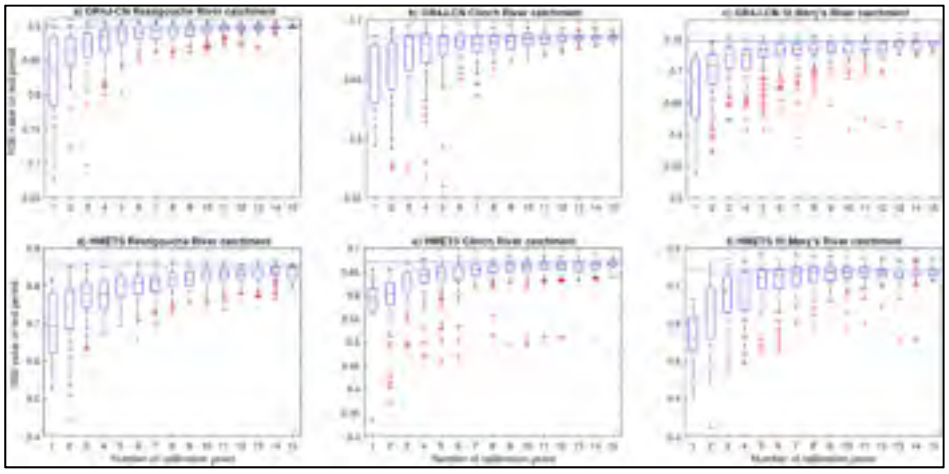

Figure-A III-3 Test-period *NSE* values with increasing number of contributing calibration years for the GR4J-CN model (top row) and HMETS model (bottom row) for the three catchments. The horizontal line represents the test-period *NSE* when all 16 remaining years are used during calibration. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

It can also be seen here that a small number of calibrations using fewer than the 16 years of available data (located in the upper whisker of most box plots) perform better on the test-period compared the calibration using all 16 years (the continuous line in Figure-A III-3). This is to be expected, as some of these calibrations end up using all the years that are more representative of the test-period, and leaving out the information that is not improving the simulation over this same test-period. Even though this increase in the *NSE* value seems like the optimal solution, it is impossible to know *a priori* which years will be more representative of the future conditions, therefore the optimal solution is the one that has the highest median value and the smallest variance.

Figure-A III-4 presents the calibration and validation *NSE* values for an increasing number of contributing calibration years. Once again, results for only one of the 30 testing periods φ are shown here but the same results are systematically found for all models and catchments.

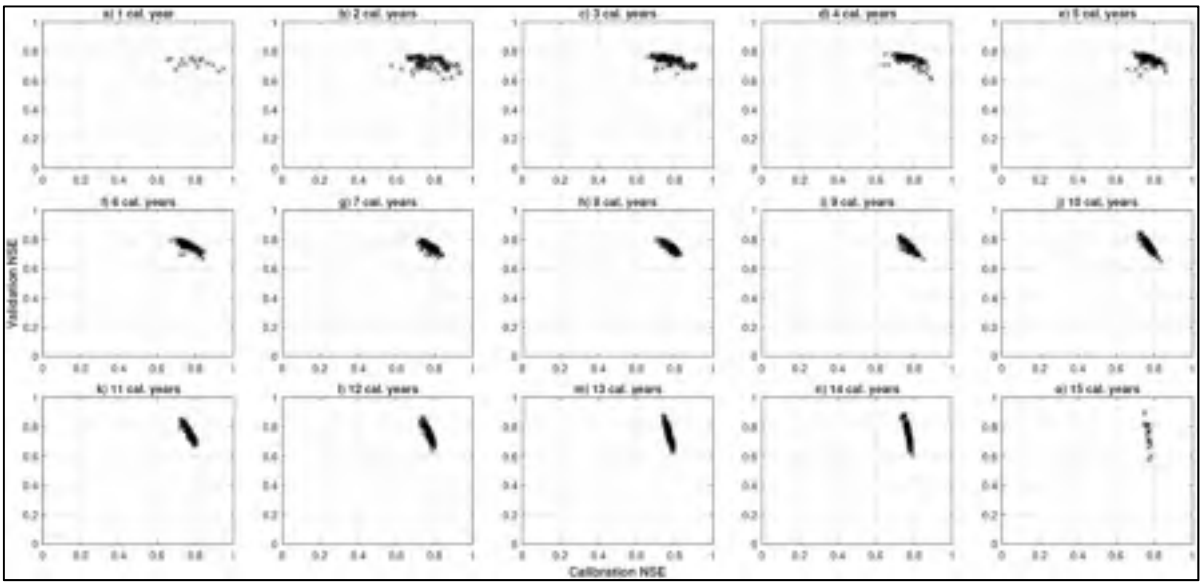

Figure-A III-4 Calibration (x-axis) and validation (y-axis) *NSE* values for independent testing period 1, on the Clinch River catchment and for the HMETS model. Each panel presents the results for a specific number of calibration years, increasing monotonically from 1 to 15. The case of 16 calibration years is not shown because there are no validation years left when calibrating on all 16 available years

It can be seen that there is a systematic evolution of the calibration and validation relationship, where at first the calibration skill is defined by the quality of the randomly selected year and the validation is relatively constant due to the much larger size, which gives it inertia and resistance to large changes in performance. As more years are used during the calibration, the trend reverses progressively to the point where all but one year is used in calibration and the validation skill becomes more volatile depending on that particular validation year's characteristics. This shows that selecting a good parameter set in calibration and discarding it if it does not perform as well in validation is a risky proposition that supposes that all years are hydrologically similar. Depending on the random draw of the calibration and validation periods, this could lead to false methodological hypotheses. For example, it could be decided to discard a parameter set that would have been "good" if the problematic year had been drawn in the calibration period rather than in the validation period.

The results were then compared on a more macro scale, looking at the 30 testing periods φ at the same time. Figure-A III-5 shows the results of the test-period *NSE* value when calibrating the model on 15 of the 16 available calibration years on the Clinch River catchment. Results are similar on the other two catchments and are not shown. Therefore, each boxplot has 16 points. The results can then be compared to the test-period *NSE* value obtained when all 16 years are used in calibration, identified by a cross. It can be seen that in most cases, calibrating on the full set of data is more robust than calibrating on 15 of the 16 years. Furthermore, there is no case in which calibrating on 16 years performs worse than the worst case amongst the 16 points in any boxplot.

It is important to note that the boxplots show large variability between each other. This is caused by the fact that the independent test-periods are composed of different years from the dataset and are thus considered as independent cases. Some of the selected years end up being more difficult to simulate, resulting more variability within the *NSE* values.

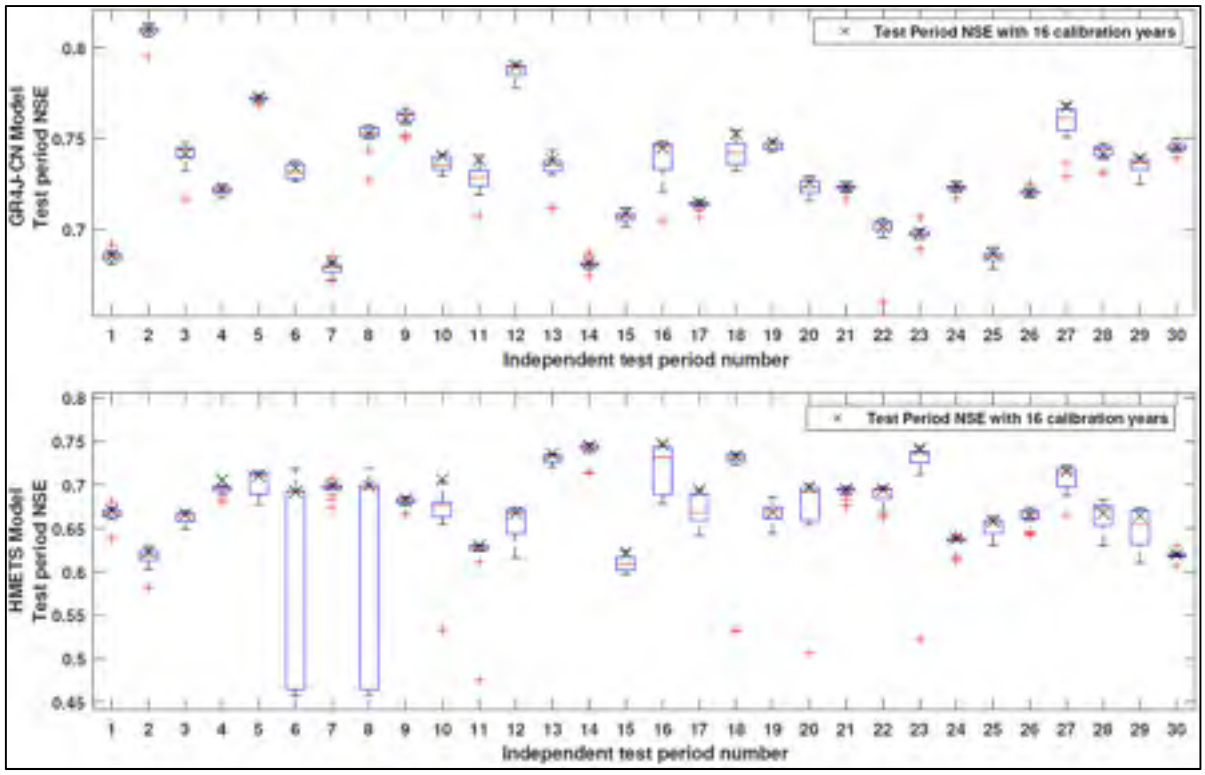

Figure-A III-5 Test-period *NSE* values for the 30 independent test runs on the Clinch River catchment. The boxplots show the test-period *NSE* when calibrating on 15 of the 16 available years (16 possible combinations in total, composing each box plot). The red line denotes the median value of the dataset and the cross represents the test-period *NSE* obtained when calibrating on all 16 available years. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

To further explore the impact of the number of calibration years on the test-period results, the difference in the *NSE* values between a calibration using all 16 available years and the median of all calibrations using a subsample of *n* years is investigated and shown in Figure-A III-6. In all cases, the *NSE* scores are the ones evaluated on the 30 independent testing periods.

A clear trend can be seen in Figure-A III-6, in which the shorter the calibration period, the lower the performance on the testing period as compared to when the 16 years are used for calibration. This is true for both models and all catchments, although only the Clinch River

catchment test-case is shown in Figure-A III-6. It is clear that the more years are used in calibration, the smaller the *NSE* differences are. Taken with the results from Figure-A III-3, it is clear that adding more calibration years allows the model to perform better on the independent testing period.

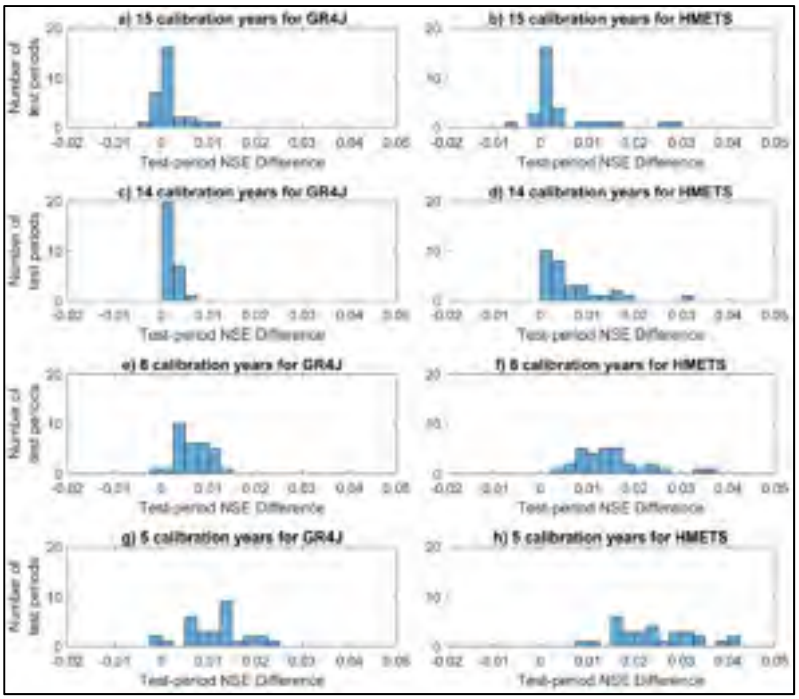

Figure-A III-6 Histogram of the Clinch River catchment difference in the *NSE* values between the 30 independent test-periods when using 1) all 16 years for the calibration and 2) the median of calibrations using 15, 14, 8 and 5 years (top to bottom). A positive *NSE* difference indicates that calibrating on 16 years is better than on a smaller number of years. The histograms display the number of test-cases that fall in each *NSE* difference bin

The results presented in Figures-A III-3-6 are not filtered based on performance, therefore one could argue that some of these results would not hold in an operational setting because if a calibration *NSE* is much higher than the validation period *NSE*, the parameter set could be rejected on the basis that the model was experiencing "overfitting" or that the parameter set was not robust. To put this hypothesis to the test, the same experiment as in Figure-A III-6

was performed, but this time only cases where the validation *NSE* is at least equal to the calibration *NSE*. Therefore, the validation *NSE* would theoretically be accepted as the skill is at least as good as on the calibration period. Figure-A III-7 presents the results.

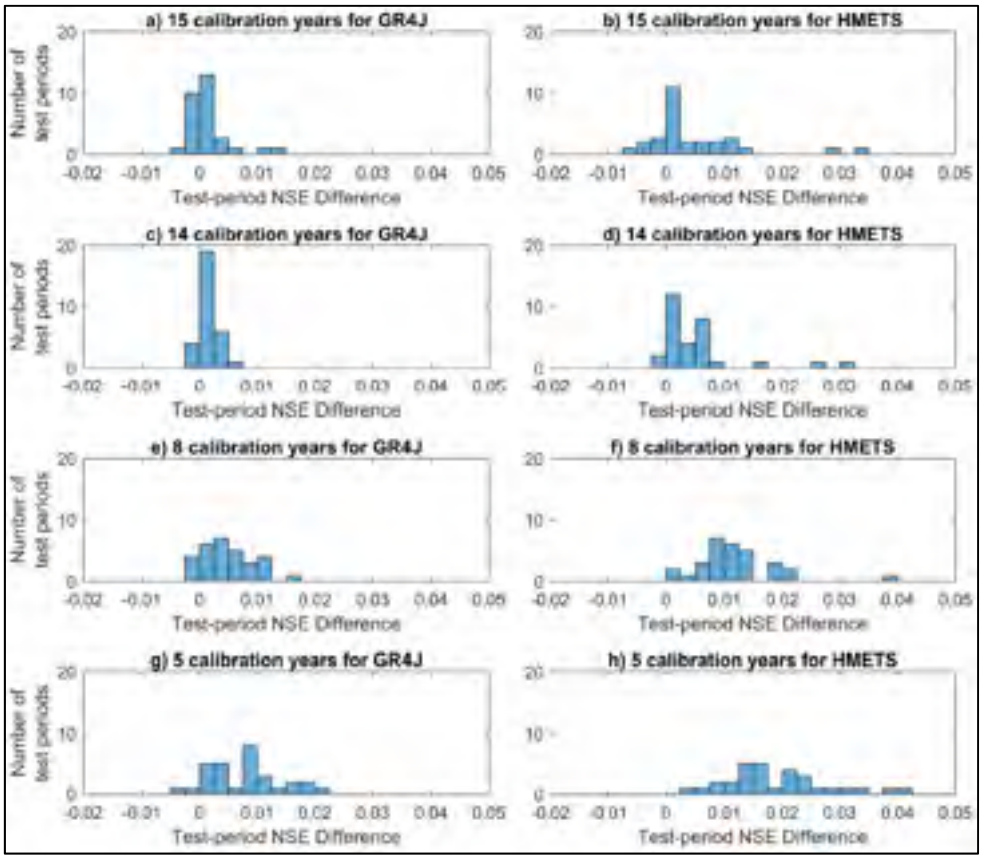

Figure-A III-7 Same as Figure-A III-6, but using only cases where the cross-validation *NSE* is at least equal to the calibration *NSE* are used to build the boxplots here

Figure-A III-7 clearly shows the same behavior as in Figure-A III-6, indicating that the skill in calibration and validation is not necessarily correlated, but can be linked to the characteristics of the calibration and validation period time series. A statistical test was performed to verify the significance of these results. The Wilcoxon test on ranks was used as a non-parametric replacement test for the 1-sided Student t-test. This was self-imposed

because the samples do not come from the same distribution given that each test has different test-period years. In all cases (filtered and unfiltered results from Figure-A III-6 and Figure-A III-7), the tests showed significant results. P-values ranged from 3x10-6 to 1x10-5 for the 15-out-of-16 years case, and all other results (14…1 out of 16 years) attained the lowest possible p-value limit for the Wilcoxon test with a sample size of 30, which is 1.73x10-6. It is therefore clear that the differences, although small, are significant in this context, although it is obvious that an NSE difference of 0.001 might not be particularly useful in an operational context. Nonetheless, the fact that some combinations of 15-in-16 years are better than the 16-in-16 year case is still theoretical, because in practice there is no way of knowing which years are to be used and which ones are to be discarded in the calibration process.

Calibration and validation were analyzed from another angle. This time, the calibrated parameter sets issued from the 16-year calibration for each of the 30 independent test-cases are used to simulate streamflows on all of the combinations of calibration and validation periods, as described in Figure-A III-2. In this graph, no use is made of the 8 independent years associated with each of the 30 independent test-case. This allowed investigating the *NSE* distributions on different periods and calibration time series lengths knowing that the parameter set is robust and performs well. Results are presented in Figure-A III-8.

In Figure-A III-8, it is possible to see the spread of *NSE* values that the hydrological model can return given randomly selected calibration years. The calibration-validation pairs, linked by dark lines in Figure-A III-8 show that high-performing *NSE* calibration is linked to low-performing *NSE* in validation, and vice-versa. This is an artifact of conditioning the results to a unique and robust parameter set. If a hydrologist were to calibrate on 15 out of 16 years and keep a single validation year, the odds are approximately 50% that the *NSE* value obtained on the validation period would be below the calibration skill. In some instances, such as for the 15-year calibration set, there exist a few years that will dramatically underperform in validation, even though the parameter set is perfectly acceptable and robust.

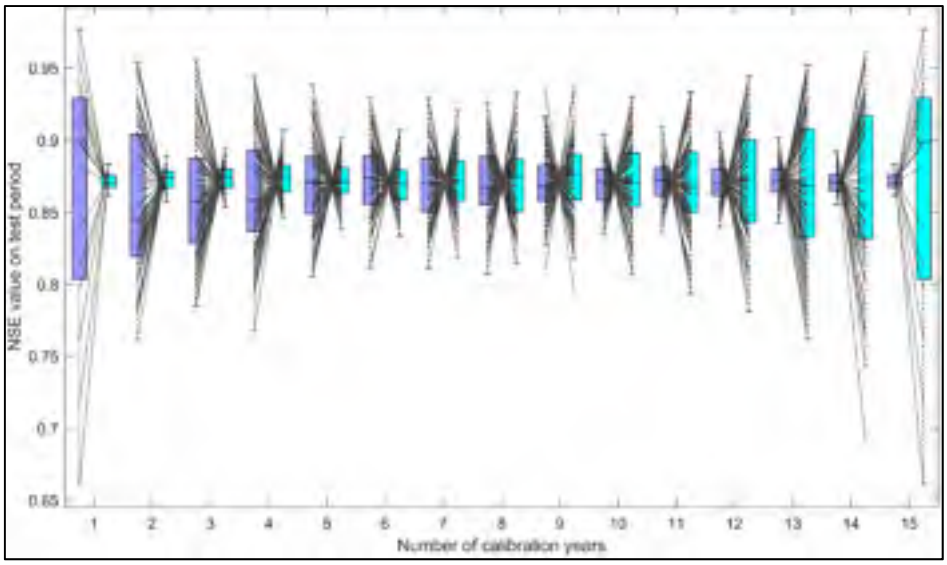

Figure-A III-8 Calibration and validation *NSE* values for independent testing period 1, using n = 1 to 15 calibration years. Dark and pale blue box-and-whisker plots represent the bootstrapped calibration and validation *NSE* values respectively. The x-axis values refer to the n calibration years used for the bootstrap. Grey lines link the calibration-validation pairs for the n calibration years. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

A final test was conducted to evaluate the robustness of calibrating over the 16 years. In this test, the parameter set derived from the calibration on the full 16-year time-series was used to run the hydrological models on the same calibration years of evaluation j using i calibration years from methodological steps 4-7 (see Figure-A III-2). The difference between the calibration and validation *NSE* values were computed for each of these cases. Furthermore, the same exercise was performed using the original calibration sets for each i and j. The idea was to measure the difference in calibration and validation *NSE* values using the original calibration sets and comparing those results to those obtained using a unique parameter set calibrated on the 16 years. Results are shown in Figure-A III-9.

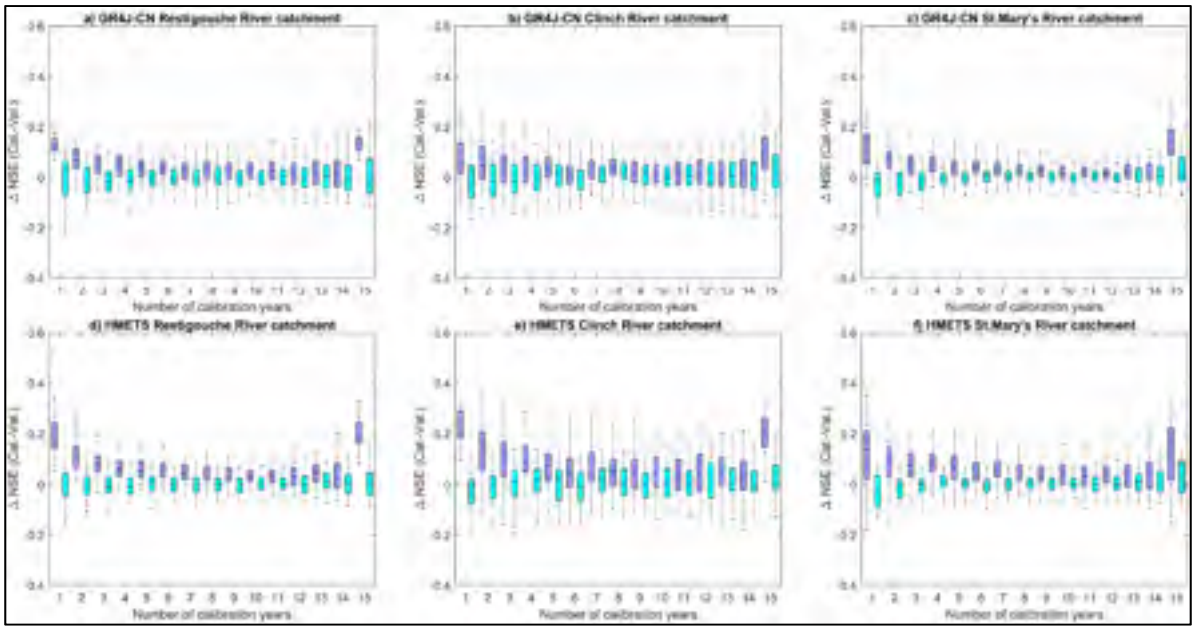

Figure-A III-9 Difference between calibration and validation *NSE* values for one independent test period using the original parameter sets (dark blue) and the full time-series calibrated parameter set (light blue) for the 15 increments of numbers of calibration years. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

It can be seen in Figure-A III-9 that there is a bias when using the original parameters calibrated on *i* years, and that this bias shrinks significantly (and oscillates around 0) for most cases with the parameter set calibrated on the entire period. This shows that the full calibration provides more robust results than calibrating on split-samples and that the validation *NSE* score is not a reliable predictor of calibration parameter performance.

III.5 Discussion

III.5.1 The hydrological model validation sacrifice

Ever since hydrological models have existed, parameter calibration and validation have been intertwined and considered an inseparable pair. Understandably, a hydrologist will want the hydrological model to be robust enough to warrant its use in forecasting or other

applications. A good validation skill comforts the user in the parameter set robustness for use in an operational setting. However, this comes at a cost on two levels.

First, the cross-validation is used purely as a parameter verification tool, and all the information contained within is withheld from the parameter set. Therefore, there might be some years in the cross-validation set that would be useful for future periods, but the parameter set would not be trained on those data and could thus not adequately react to the new inputs. This is a direct sacrifice of model performance for confidence in the model.

Second, it forces the hydrologist to use parameter sets that are perhaps not the best overall by sacrificing some skill in calibration to ensure calibration and validation are similar in terms of performance. In this case, a good parameter set could be rejected because of a larger-than-anticipated spread between the calibration and cross-validation *NSE*. As was shown in Figure-A III-8, this can happen quite easily and be costly from an operational point of view. Figures-A III-6 and III-7 also show the same reasoning.

This study shows that the cost can be high in terms of overall performance on an independent test period. The resampling bootstrapping method implemented in this work shows that even on randomly selected years, it is almost always a good idea to use all available information when calibrating the hydrological model. Of course, it is possible to tailor a set of calibration years that will outperform the entire series (for example, by removing data years that are too dissimilar to the testing period) but this cannot be performed *a priori*. Therefore, in absence of information on the future years, the most conservative approach is to calibrate on the entire available time-series in order to ensure that the parameter sets contain as much information as possible. Figures-A III-3, III-4 and III-5 demonstrate this point quite clearly.

III.5.2 The impact of the calibration objective function

In this study, only the *NSE* was used as the objective function during calibration. While it does have some drawbacks, as explained in section III-3.2, it is generally recognized as providing satisfactory results. Ideally, this study would have included more objective functions in order to validate the conclusions on other metrics, but the complexity of performing the sheer number of calibrations made it prohibitive. Instead, the parameter sets obtained through calibration on the *NSE* metric were evaluated using two other metrics, namely the relative bias (%) and the Pearson correlation coefficient. Figures-A III-10 and 11 show the results respectively for the relative bias and correlation for the three main catchments for a single test-case, although results are similar for all model-catchment pairs.

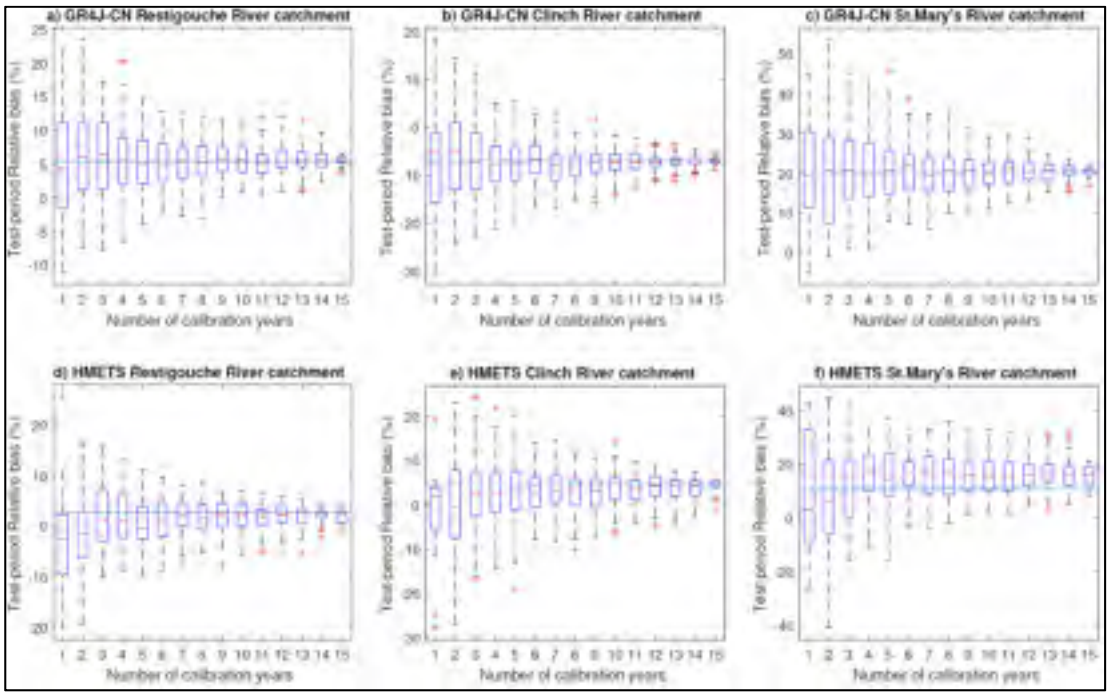

Figure-A III-10. Test-period Relative bias values with increasing number of contributing calibration years on the *NSE* objective function for the GR4J-CN model (top row) and HMETS model (bottom row) for the three catchments. The horizontal line represents the test-period Relative bias when all 16 remaining years are used during calibration. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

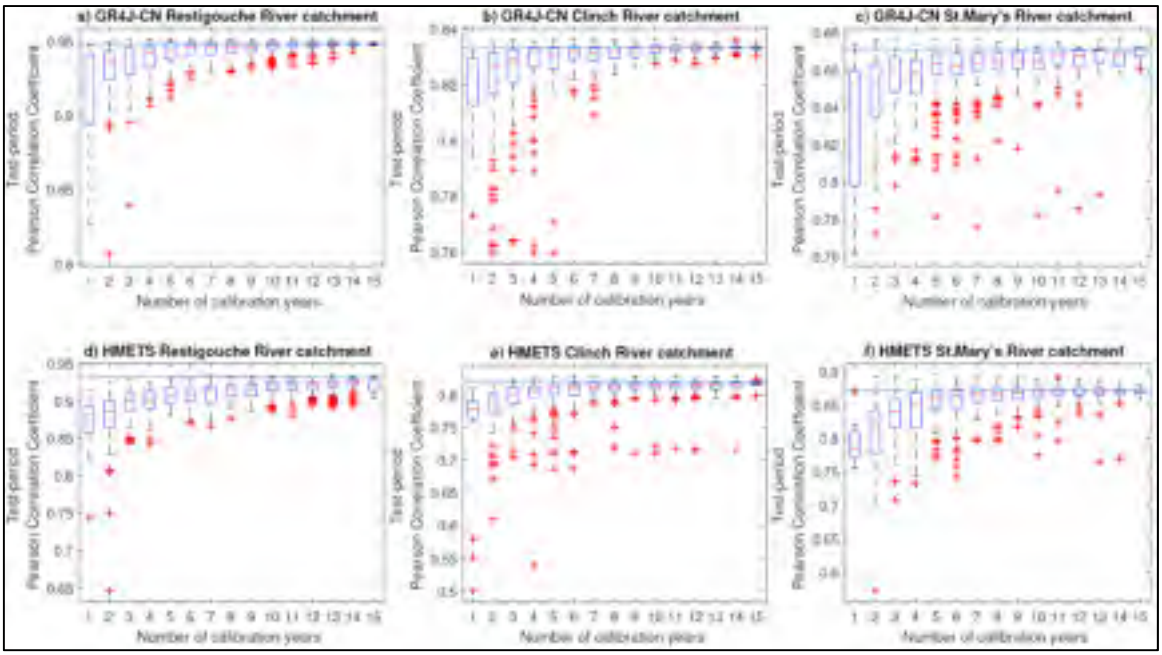

Figure-A III-11. Test-period *Pearson Correlation Coefficient* values with increasing number of contributing calibration years on the *NSE* objective function for the GR4J-CN model (top row) and HMETS model (bottom row) for the three catchments. The horizontal line represents the test-period Pearson Correlation Coefficient when all 16 remaining years are used during calibration. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

As can be seen in these figures, the bias and correlation values converge to a point as more and more calibration years are used. In Figure-A III-10, the bias levels converged when more calibration years were added, but they did not necessarily converge towards zero. This seems to be caused by the *NSE* preferring correlation to bias during calibration. Furthermore, it intuitively seems more likely that there could be a larger difference in bias when selecting years randomly between the calibration and validation period. However, the results seem to indicate that if the calibration were to be performed on the bias instead of the *NSE* as the objective function, the same converging results would be found, but would converge closer to zero.

The correlations (Figure-A III-11) follow the same patterns as *NSE* (Figure-A III-5), again probably due to the fact that *NSE* gives more importance to the correlation than the bias in this study, as explained in Gupta et al. (2009). A comparative study using the Kling-Gupta Efficiency (*KGE*) metric (Gupta et al., 2009), which allows distinguishing more easily between the correlation, bias and variance, could be performed to assess these points further.

III.5.3 The sense of security provided by parameter validation: False advertising?

The results in Figures-A III-3-5 and Figure-A III-8 indicate that the validation of a calibrated hydrological model parameters is very sensitive to the characteristics of the underlying data in both the calibration and validation periods. Both hydrological models used in this study displayed the same behavior, which is an increase in independent testing skill as more calibration years are included. However, validation is often used to ensure that the model performs adequately on an independent period and that the model is not overfitting the data instead of representing the hydrological processes. The results in Figures-A III-5 and III-8 show that the range of *NSE* values calculated on the independent testing period can vary wildly depending on the properties of the selected years that are selected for calibration and testing. In many studies and operational settings, a much lower validation *NSE* would be addressed by other means such as changing the calibration/validation periods to ensure similar performances, and until a so-called robust parameter set is obtained. It can therefore be difficult to distinguish between overparameterization, model inadequacy and simply the quality of a certain time period that renders its modeling difficult.

In this context, and in light of the results obtained in this study, it seems clear that the optimal solution is to include all available years in the dataset for the calibration process. Consequently, the responsibility of the hydrologist is to ensure that there is not a structural problem with the meteorological or hydrometric data that could lead to the model attempting to fit the model on unrealistic data. By selecting all years for calibration, the parameter sets will contain the maximum amount of information for the study site, therefore maximizing its

chances of performing adequately in future simulations and prediction. Of course, if calibrating the model on all years returns a low/poor *NSE* value, then there would be reason to suspect that the model is inadequate for the simulation conditions and would need to be investigated further, perhaps by changing the model altogether or inspecting the datasets for errors. The results in Figures-A III-3 and III-8 provide evidence that overparameterization cannot be distinguished from the seemingly poor results obtained from a randomly selected set of validation years. Ultimately, the *NSE* skill of a model calibrated on the entire available dataset should be the only criteria to determine if the model is adequate for simulating on the given catchment.

III.5.4 Is this concept generalizable?

This study used a contrasting set of three catchments with different hydrometeorological inputs and two hydrological models of different complexity. An analysis of calibrated parameter sets (results not shown) indicated that the GR4J-CN model displayed no equifinality when more than 12 years of data were included in the calibration set. Unlike GR4J-CN, HMETS' parameter sets spanned large sections of the parameter space (even when calibrating on 15 years), thus indicating presence of equifinality. In both cases, the optimal solution was to calibrate the hydrological model on all years, which resulted in superior performance on the independent testing periods. This evidence supports the idea that the parameter tuning, even when in presence of equifinality, integrates information from the learning set that can be applied afterwards. The fact that these conclusions hold for the six test-cases lends credibility to the idea that these findings should be applicable to a wide range of models and catchments.

To investigate the issue further, the experiments were performed on five independent catchments using a less restrictive methodology. Instead of setting a maximum of 100 combinations for each number of calibration years, only 64 were used in order to maximize available computing resources. Furthermore, only 10 independent test-cases were performed

instead of 30 as used in the full-scale experiment. Nonetheless, the results of this test show that the results also hold for these five catchments which gives more weight to the idea that the concept is generalizable. Table-A III-2 shows the characteristics of the five supplementary catchments, and Figure-A III-12 shows the same analysis as Figure-A III-5 applied to these verification catchments.

Table-A III-2. The five supplementary catchments' characteristics and mean annual hydrometeorological variables

	Indian Creek, CA	Yadkin River, NC	Trinity River, CA	Sinnemahoning Creek, PA	Neches River, TX
Catchment characteristics:					
Drainage area (km^2)	1913	5903	7386	1773	2964
Centroid Latitude (degrees)	40.1°N	35.9°N	41.1°N	41.3°N	31.9°N
Centroid Longitude (degrees)	120.9°W	80.4°W	123.7°W	78.1°W	95.4°W
Mean hydrometeorological variables:					
Annual total precipitation (mm)	806	1188	1355	999	1008
Rainfall	599	1135	1231	810	990
Snowfall	207	53	124	189	18
Annual daily temperature (°C)					
Minimum	0.3	7.1	3.7	1.8	12.2
Maximum	17.0	20.4	19.3	14.5	25.0
Mean	8.7	13.7	11.5	8.1	18.6
Annual daily streamflow (m^3s^{-1})	17	82	162	31	17

The Wilcoxon statistical test was once again applied on the new model-catchment pair results on the independent test period, and results are shown in Table-A III-3. In all but two cases, the differences between using 15 and 16 calibration years was significant. The difference was significant for all model-catchment pairs when using 14 calibration years. It is important to note that for these cases, only 10 samples (10 random independent test periods) were used; therefore, the power of the statistical test is reduced and the lowest possible p-value is 0.002.

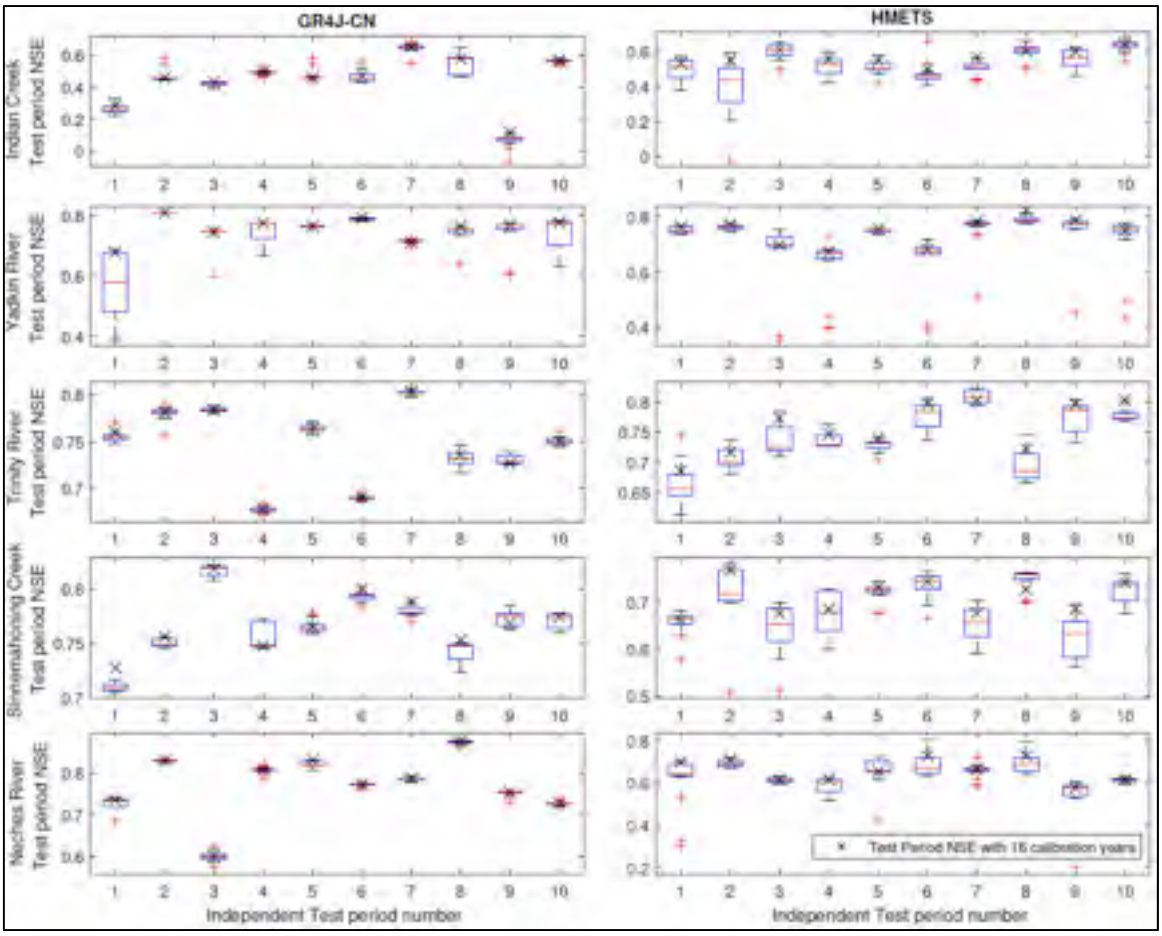

Figure-A III-12. Test-period *NSE* values for the 10 independent test runs on the five verification catchments. The boxplots show the test-period *NSE* when calibrating on 15 of the 16 available years (16 possible combinations in total, composing each box plot). The red line denotes the median value of the dataset and the cross represents the test-period *NSE* obtained when calibrating on all 16 available years. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

These results, along with the results in Figures-A III-10 and 11, seem to agree with the idea that the concept is generalizable at least to a certain extent. However, it is possible that the full-series calibration is not optimal for some other model-catchment setups, such as for catchments showing signs of non-stationarity as will be discussed in Section III.5.7. While we are confident that in most cases the optimal methodology is to calibrate on all available years, more work needs to be done to prove its generalizability on a wider scale.

Table-A III-3. P-values of *NSE* scores on the independent test period for the 5 verification catchments, for 14- and 15-in-16 calibration years as compared to the 16-in-16 year calibration case

Catchment	p-values			
	GR4J-CN		HMETS	
	15 years (of 16)	14 years (of 16)	15 years (of 16)	14 years (of 16)
(1) Indian Creek, CA	0.002	0.002	0.020	0.002
(2) Yadkin River, NC	0.232	0.002	0.027	0.002
(3) Trinity River, CA	0.027	0.002	0.004	0.002
(4) Sinnemahoning Creek, PA	0.242	0.002	0.105	0.002
(5) Neches River, TX	0.004	0.014	0.020	0.002

III.5.5 The compromise solution

It is understandable that modifying such an ingrained habit as using a split-sample calibration and validation is a tough sell, and that some, if not most, researchers and practitioners will prefer to keep the validation step as a robustness evaluation tool. However, it is also clear from the results analyzed in this work, that there is a cost to this solution. In an attempt to satisfy both sides of the equation, it is proposed to preserve the validation step (if so wanted by the operator), but that if the skill in validation is acceptable, then the calibration should be performed once again over the entire time series. In this way, the operator will know that the model is able to perform well on both independent periods, indicating that the processes are well simulated by the model. Then, by recalibrating over the entire dataset, the model will have been trained on more data, leading it to contain more information and making it more robust for future use.

This hybrid method can be seen as a compromise solution which would lead to the same conclusion as calibrating on all years and then judging the overall *NSE* score. In both cases, the entire time-series is used in calibration but the way that the model is deemed to be acceptable or not is performed in a separate step. The only drawback of the proposed solution is the required extra time to perform a split-sample calibration before the full calibration.

III.5.6 The issue of model complexity

There is a trend in hydrological science toward increasingly more complex process-based and/or distributed hydrological models with some models now implementing land surface schemes with complex formulations such as the Richards nonlinear differential equations of water movement in the non-saturated portion of the soil column (Paniconi & Putti, 2015). In such models, calibration over the full dataset length would not be possible without access to massive parallel computing facility. In these cases, the selection of a calibration subsample would still be necessary but should not follow typical splitting strategies but rather focus on strategic approaches such as proposed by Singh et Bárdossy (2012) or Razavi et Tolson (2013). However, if the eventuality that a full calibration would be possible to the user, the additional information included within the calibration would very likely increase the performance of future simulations. Furthermore, a final analysis was performed to evaluate the effects of the length of the available time series. In this test, 51 years of data were used from the Trinity River catchment in California, ranging from 1948-1998. A single test run was performed while keeping the first year as the warm-up period, 15 random years as the independent test period and performing the calibration on random subsets of 1-34 calibration years. A maximum of 500 combinations per case was used to explore the larger dimension of the problem. The following evaluation of the *NSE* on the independent test period for these trials was compared to that of calibrating on all 35 available years. Results are presented in Figure-A III-13.

It is clear that the length of the available dataset plays a crucial role in the effects of split-sample calibration and validation. Clearly, from Figure-A III-13, there is a limit to how much information a parameter set can contain. When there is sufficient information to represent the entirety of the time-series, then it seems that there is no gain in calibrating on all years, although there is a case to be made regarding the robustness of the parameter set. There are some combinations of 33 and 34 years that vastly underperform the 35-year case. Therefore, it is still recommended to use all available years during the calibration phase. Notice that the

last boxplot in panel a) (34 out of 35 years) seems different than the others. This is because it contains only 35 points, whereas the other boxplots contain 500 points each and thus the threshold for outlier detection is wider.

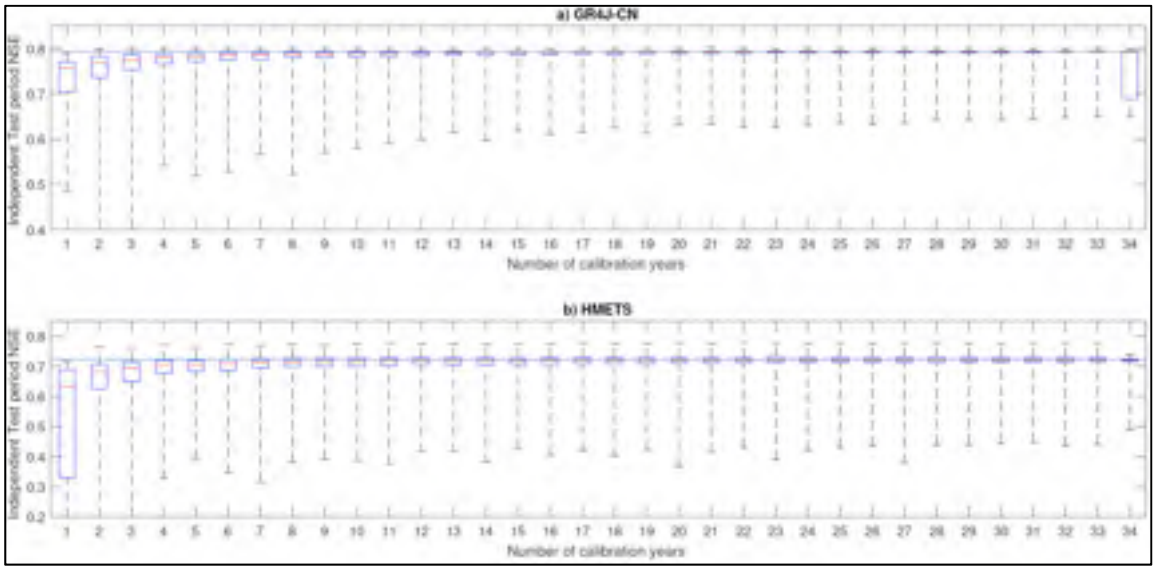

Figure-A III-13. Boxplots of independent test-period *NSE* scores when calibrating on random subsets of 1-34 calibration year combinations for the a) GR4J and b) HMETS models. The horizontal line represents the independent test-period NSE when calibrating on the 35 available calibration years. Boxplot box edges represent the 25th and 75th quantiles, and the whiskers represent +/-2.7 standard deviations from the mean

III.5.7 The question of stationarity

In this study, the effect of calibration and validation is investigated on three catchments which did not show signs of non-stationarity, i.e. the mean annual streamflow did not contain a trend over a 25-year period. This allowed randomly sampling from the database to generate calibration and validation sets. This raises the question as to how the method would fare on a catchment that is subject to non-stationarity. Obviously, in this scenario, the independent test period would need to be in the most recent years and those years could not be randomly selected from the entire time series. Furthermore, the calibration and validation would have

to make use of the remaining years, perhaps on a distinctly different climate. However, in an operational or research setting, there are two options to hedge one's bets.

First, the odd/even year split-sample method has been proposed to calibrate a hydrological model on non-stationary conditions. This method allows calibrating on the odd years and cross-validating on the even years (or vice-versa), thus including the same trend in the calibration and validation. While this gives confidence to the modeler (as explained in Section III.5.3), the reality is that the time series does not contain all of the information that is available and that the future simulations will be performed with a model that under-exploited. Furthermore, the drawback to this method is that it takes essentially as much time to perform this type of calibration, as the whole period still has to be simulated to implement such split-sample techniques.

Second, the logical conclusion would be to perform a calibration on the entire time-series to include the information contained in the even years. As stated in Section III.5.5, it would be possible to perform the validation and, given an acceptable *NSE*, recalibrate the parameters using the entire time series. In light of the results obtained in this study, it is posited that the method should perform as well as it does in stationary conditions, but this remains to be validated and should be explored in future works. Some of these issues are discussed in Thirel, Andréassian, et Perrin (2015).

III.6 Conclusion

In this study, an experimental approach is presented to investigate the added benefit of foregoing the traditional validation step in hydrological model calibration. The results presented above point towards validation being futile at best, detrimental at worst, and deceiving in all cases. The main hypothesis that calibrating a hydrological model on a certain period, and then validating the model's parameter set on another period, was shown to be flawed due to numerous possibilities of false-negatives. When a validation *NSE* is inferior to

a calibration *NSE*, typically the model is recalibrated or the calibration and validation years are inverted or shifted around. The end-result is a model whose parameters include only information from the calibration period, sacrificing the information contained in the rest of the time-series.

The method proposed in this study is to calibrate over the entire time-series and forego the validation step. If the calibration skill is acceptable, then the hypothesis is made that the model is able to simulate on that period. A hybrid method that conserves confidence in the model's robustness was proposed in which the model is recalibrated over the entire period if the typical calibration/validation step returns acceptable results.

The results are consistent over the three contrasting catchments and two models, which provides some evidence to the full time-series calibration to be the optimal strategy, although some calibration/validation combinations do outperform the full calibration skill in a small proportion. Statistical testing showed that the method was robust, and verification on 5 independent catchments showed similar results. Furthermore, the method translated well to other metrics such as the relative bias and the Pearson correlation coefficient.

Of course, the problem lies in the fact that the capacity of the model to perform better on an independent testing period cannot be evaluated a priori, therefore the optimal strategy is the one which is statistically more likely to occur given random inputs. In this study, in 30 independent runs and over six test-cases, the optimal strategy was convincingly to calibrate on the entire time-series and forgo validation. The same was found for the smaller verification run, with 10 independent runs over 10 model-catchment pairs.

Finally, it is important to note that this work has some limitations that should be addressed in future work. For example, non-stationary time-series should also benefit from this methodology but this remains to be validated. Furthermore, the generalizability of the method to other catchments and models should be investigated, although it is expected that

the same results will be found. For example, models operating on sub-daily time-steps could be tested using this methodology to see if the same behavior is observed. The fact that information is left out of the parameter set should manifest itself even further in more process-based hydrological models, but this also needs to be tested.

III.7 Acknowledgements

The authors would like to extend their gratitude to the anonymous reviewers and associate editor whose insightful comments were instrumental in shaping the paper in its current form. The authors would like to thank the École de technologie supérieure (ETS) and the Natural Sciences and Engineering Research Council of Canada (NSERC) for partial funding of this research project.

ANNEXE IV

RELATIVE IMPORTANCE OF INTERNAL CLIMATE VARIABILITY VERSUS ANTHROPOGENIC CLIMATE CHANGE IN GLOBAL CLIMATE CHANGE

Jie Chen[a], François Brissette[b], Jean-Luc Martel[c], Xunchang J. Zhang[d], Allan Frei[e]

[a] State Key Laboratory of Water Resources & Hydropower Engineering Science, Wuhan University.
[b, c] Département de Génie de la construction, École de technologie supérieure
[d] USDA-ARS Grazinglands Research Laboratory
[e] Institute of Sustainable Cities, and Department of Geography, Hunter College

Article accepté avec révisions majeures dans la revue Journal of Climate, décembre 2018.

Abstract:

The potential impacts of climate change have been extensively studied using the typical top-down approach. This approach estimates climate change impacts as the differences in environmental indicators (e.g. runoff or growth season) between the future and historical periods. In the majority of such studies, the role of internal climate variability (ICV) in a changing climate is not considered. However, recent studies indicate that the uncertainty related to ICV can be comparable to, or even larger than the anthropogenic climate change (ACC) at various time horizons. To better understand the role of ICV in climate change impact studies, this study separates these two components and quantifies the importance of ICV in relation to the ACC in global and regional climate change using a criterion of Time of Emergence (ToE). The ToE in this work is defined as the 30-year period when the ACC emerges from ICV. ToEs are investigated for both annual and seasonal mean precipitation and temperature using multiple climate models and multi-member ensemble of climate models. The ACC is defined as the multi-model ensemble mean and the ICV is defined as the inter-member variability of multi-member ensembles of climate models. The results show that the annual mean precipitation ACC are expected to emerge within this century over extra-tropical regions as well as along the equatorial band. However, ToEs are unlikely to occur, even by the end of this century, over intra-tropical regions outside of the equatorial band. In contrast, annual mean temperature ACC have already emerged from the temperature

ICV for most of the globe. Similar spatial patterns are observed at the annual and seasonal scales, while a weaker ACC for June-July-August precipitation and additional ICV for December-January-February temperature translate to later ToEs for some regions. In addition, similar ToE estimates are observed when using 5-, 10-, and 40-member ensembles, suggesting that a 5-member ensemble may be sufficient to estimate ICV for mean precipitation and temperature at the multi-decadal scale. Overall, this study implies that in many cases, medium term adaptation to changing mean precipitation should also focus on ICV. For temperature, ICV does not play a dominant role and mitigation strategies should therefore focus on ACC.

Keywords: anthropogenic climate change; internal climate variability; precipitation; temperature; time of emergence

IV.1. Introduction

Most of the world's regions have experienced a significant warming climate during the last half-century and global warming is very likely to continue during the 21st century (IPCC, 2013). The potential impacts of climate change have now been a concern for a few decades for many applications such as water resources and agriculture. Most climate change impact studies follow a typical top-down approach, which defines climate change impacts according to changes in the environmental indicators (e.g. runoff and growth season) of a future period relative to those over a historical reference period (Anandhi et al., 2013; Chen et al., 2011a; Matonse et al., 2013). Based on the above definition, climate change refers to a statistically pronounced change in the average state and/or variability of climate variables (e.g. precipitation and temperature) typically lasting for several decades or more (IPCC, 2013; Solomon, Qin, Manning, Averyt, & Marquis, 2007).

Climate change may be due to natural internal processes within the climate system (internal climate variability; ICV), or to variations in natural or anthropogenic external forcing

(external climate variability) (IPCC, 2013). ICV is the natural fluctuation intrinsic to a given climate state due to internal interactions within the complex nonlinear climate system. External forcing includes anthropogenic forcing such as greenhouse gas (GHG) emission and tropospheric aerosol loading acting on the atmosphere's composition, and natural external forcing ranging from radiation variations due to pulse-like events such as explosive volcanic eruptions and solar cycles to changes on geological time scales due to plate tectonics. Since natural external forcing affects climate on either short-term temporal scale (a few years for volcanic eruptions) or much longer temporal scale (tens of thousands of years for the Milankovitch cycles), they are not usually taken into account in decadal to multi-decadal climate change impact studies (Chen & Brissette, 2018). Consequently, the ICV and anthropogenic climate change (ACC) are the two main components that influence the overall climate change at the multi-decadal scale (Deser et al., 2012a; Hulme et al., 1999). In the case of climate projections from climate models, only these two components are typically considered. Top-down approaches, which define a difference between two simulations over future and historical periods do not typically separate the role of ICV from the overall climate change. Even though adaptation to climate changes includes changes caused by both ACC and ICV, the climate change mitigation may only focus on ACC. In other words, while ACC may be mitigated with global and regional emission reduction strategies, ICV, by definition, is irreducible. Thus, it is of great importance to investigate the role of ICV in climate change by separating both components of climate change (ACC and ICV).

The role of ICV in climate change has been given increasing attention during recent years (Deser et al., 2012a; Deser et al., 2012b; Hawkins & Sutton, 2009, 2011; Maraun, 2013b; Martel et al., 2018). In particular, climate model multi-member ensembles have been computed to investigate this issue. For example, the National Center for Atmospheric Research Community Climate System Model version 3 (CCSM3) was run 40 times to generate a 40-member ensemble to span a range of internal variability (Collins et al., 2006; Deser et al., 2012a). More recently, a 40-member large ensemble was designed by the Community Earth System Model (CESM) community (CESM1) with the explicit goal of

enabling assessment of climate change in the presence of ICV (Kay et al., 2015). In such ensembles, all members are simulated using the same model and external forcings, but with small differences in atmospheric initialization. Based on these multi-member ensembles, the role of ICV in climate change has been investigated in a few studies. Deser et al. (2012a) investigated the role of ICV in future North America climate change based on CCSM3 ensemble and found that ICV contributes substantial uncertainty to seasonal mean temperature and precipitation trends for the 2006-2060 period over North America on local, regional and continental scales. A follow-up study Deser et al. (2012b) further showed that the ICV of precipitation and temperature estimated using CCSM3 multi-member ensemble accounts for at least half of the inter-model spread in projected climate trends during 2006-2060 in the Coupled Model Intercomparison Project Phase 3 (CMIP3) multi-model ensemble.

Other studies (e.g. Giorgi & Bi, 2009; Hawkins & Sutton, 2011; King et al., 2015; Mahlstein, Knutti, Solomon, & Portmann, 2011; Maraun, 2013b; Martel et al., 2018) quantified the importance of ICV relative to ACC using various measures of Time of Emergence (ToE) as criteria. The ToE is defined as the time when ACC becomes greater than the range of ICV (Hawkins & Sutton, 2011). These studies estimated the range of ICV based on multi-model or multi-member ensembles or a combination of both. However, these two types of ensembles are independent from each other because of inter-model difference. The former is used to address a mix of ICV and inter-model uncertainties (typically associated with the understanding and parameterization of climate models), while the latter is used to specifically address the ICV in a given climate model (Fischer et al., 2013; Kay et al., 2015; Kim, Kwon, & Han, 2016). The use of multi-member ensembles seems to be more appropriate when investigating the ICV. Moreover, compared to most studies which estimated the ToE for a specific year or decade, it may be more reasonable and reliable to estimate the ToE for a multi-year period (e.g. a 20- or 30-year window), since climate change impact studies are commonly conducted at the multi-decadal time scale. For example, the World Meteorological Organization (WMO) suggested using 30 years as a classical period to

estimate climate change (IPCC, 2013), as done in most climate change impact studies (e.g. Chen & Brissette, 2018; Chen et al., 2013; Chen et al., 2011a; Mullan et al., 2017; Teutschbein & Seibert, 2012; Zhang, 2005).

Accordingly, this study quantifies the importance of ICV relative to ACC in global climate change in terms of the ToE using a combination of multi-model and multi-member ensembles. Specifically, three estimates of ICV for annual and seasonal mean precipitation and temperature are estimated using three different climate model multi-member ensembles, while the ACC is estimated using 20 Global Climate Models (GCMs) from the Coupled Model Intercomparison Project Phase 5 (CMIP5) ensemble (Taylor et al., 2011).

IV.2. Datasets

This study used both annual and seasonal total precipitation and mean temperature as simulated by 20 GCMs from the CMIP5 ensemble (Taylor et al., 2011). Basic information about these models is presented in Table-A IV-1. Three climate models (CanESM2, CSIRO-Mk3.6.0 and CESM1) consist of multi-member ensembles which are used to estimate the ICV. Only one member is used for all 20 GCMs (including CanESM2, CSIRO-Mk3.6.0 and CESM1) to assess the ACC. The Community Earth System Model version 1 (CESM1) taken from the Large Ensemble Community Project (Kay et al., 2015) includes a 40-member ensemble of fully-coupled simulations for the 1920-2100 period. Both CanESM2 and CSIRO-Mk3.6.0 taken from the CMIP5 database respectively include 5 and 10 members. All GCMs simulations cover the 1920-2100 period, with the 1920 to 2005 period being driven by historical climate forcing, and the Representative Concentration Pathway 8.5 (RCP8.5) forcing (Lamarque et al., 2011; Riahi et al., 2011) being used from 2006 on. The use of this high-emission scenario is expected to present a strong ACC, thus resulting in a relatively early ToE compared to other lower emission scenarios (e.g. RCP2.6 or RCP4.5).

Table-A IV-1. Basic information of the selected CMIP5 model

ID	Model name	Institution	No. of members	Horizontal resolution [lon. (º) x lat.(º)]
1	ACCESS1.0	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	1	1.875 x 1.25
2	ACCESS1.3		1	1.875 x 1.25
3	BCC-CSM1.1(m)	Beijing Climate Center, China Meteorological Administration	1	1.125 x 1.125
4	BCC-CSM1.1		1	2.8 x 2.8
5	CMCC-CM	Centro Euro-Mediterraneo per I Cambiamenti Climatici	1	0.75 x 0.75
6	CMCC-CMS		1	1.875 x 1.875
7	CNRM-CM5	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	1	1.4 x 1.4
8	GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory	1	2.5 x 2.0
9	GFDL-ESM2M		1	2.5 x 2.0
10	INMCM4	Institute for Numerical Mathematics	1	2.0 x 1.5
11	IPSL-CM5A-MR	Institut Pierre-Simon Laplace	1	2.5 x 1.25
12	IPSL-CM5B-LR			3.75 x 1.8
13	MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1	1.4 x 1.4
14	MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	1	2.8 x 2.8
15	MIROC-ESM		1	2.8 x 2.8
16	MRI-CGCM3	Meteorological Research Institute	1	1.1 x 1.1
17	IPSL-CM5A-LR	Institut Pierre-Simon Laplace	1	3.75 x 1.8
18	*CanESM2*	*Canadian Centre for Climate Modelling and Analysis*	*5*	*2.8 x 2.8*
19	*CSIRO-Mk3.6.0*	*Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence*	*10*	*1.8 x 1.8*
20	*CESM1*	*National Center for Atmospheric Research*	*40*	*1.0 x 1.0*

IV.3. Methodology

In order to quantify the importance of ICV relative to ACC for precipitation and temperature, both ICV and the ACC first have to be estimated. With respect to the ICV, the three multi-member ensembles (i.e. CanESM2, CSIRO-Mk3.6.0 and CESM1) are used, providing three different estimations. The ACC is estimated using the 20 GCMs ensemble, each providing a single simulation.

IV.3.1 Estimation of the internal climate variability

Three multi-member ensembles covering the 1951-2100 period were used to estimate ICV. Since climate model simulations may contain the signature of both ICV and ACC, the signature of ACC (represented by a trend) should first be removed in order to properly estimate ICV. Consistently with previous studies (e.g. Maraun, 2013b; Mehrotra, Sharma, Bari, Tuteja, & Amirthanathan, 2014; Zhuan et al., 2018), the hypothesis that the ACC follows a linear trend (for both mean precipitation and temperature) was made. Although a nonlinear regression may provide a better fit for the climate data over some regions, there is a danger of overfitting the time series and removing part of the ICV (Mehrotra et al., 2014; Zhuan et al., 2018). More details on the detrending method used in this work are presented below, followed by the definition of ICV.

IV.3.1.1 Detrending

A two-stage detrending method proposed by Zhuan et al. (2018) was used to remove the ACC from climate simulations for both historical (1920-2005) and future periods (2006-2100). This method has been detailed in Zhuan et al. (2018), and is only briefly described as follows. The first and second stages involve removing the ACC for the historical forcing (before 2005) and RCP8.5 forcing (after 2005) periods, respectively. The two sub-periods were treated separately because they are driven by different external climate forcing. The

trend was removed at the seasonal basis. A Mann-Kendall non-parametric test (Kendall, 1975) was used to check whether the ensemble mean of seasonal total precipitation or seasonal mean temperature has a significant trend over each grid point. The detrending was only performed if a significant trend (with a 95% confidence level) was detected, removing it using a linear method (Sen, 1968). Since all members of a multi-member ensemble were simulated with the same external forcing, they are expected to contain the same ACC signature. Therefore, the same trend estimated based on ensemble mean was removed from each individual member.

IV.3.1.2 Definition of the internal climate variability (ICV)

The ICV operates at various temporal (seasonal to multi-decadal) and spatial scales (local to global) for specific climate variables. However, it is defined in this study as the multi-decadal variability (i.e. variability over a 30-year) for both mean precipitation and temperature at the climate model local scale of the grid points. Even though the multi-decadal variability of mean precipitation and temperature is only one of many ICV components, it may be one of the most important ones in climate change impact studies. This is especially true when taking into account the fact that climate change trend is often defined over a 30-year period and climate change impact studies are usually conducted at the multi-decadal time scale. When using a multi-member ensemble, the multi-decadal variability was defined as the inter-member variability for a multi-year period. Previous studies (e.g. Chen & Brissette, 2018) have shown that multi-member ensembles perform reasonably well at representing multi-decadal variability. Specifically, for each multi-member ensemble of a climate model, the ICV was estimated using the following six steps.

1. The ACC was first removed using the above two-stage detrending method. The residual time series was used to estimate the ICV of annual and seasonal mean precipitation and temperature at the multi-decadal scale.
2. The whole period of 150-year (1951-2100) annual or seasonal mean precipitation and temperature was divided into 121 30-year periods (i.e. 1951-1980, 1952-1981, 1953-

1982, …, 2069-2099 and 2071-2100) with a one-year lag between each of two periods.

3. The 30-year running means of annual or seasonal mean precipitation and temperature were calculated for each member and each variable. For example, the annual mean precipitation of 1951-1980 was calculated for all of CESM1 40 members separately. Thus, a total of 121 values (one for each 30-year period) were obtained for each member of each climate model.
4. The standard deviation (σ) among all members of a given climate model was calculated for each period and each variable. For example, the σ of annual mean precipitation over all CESM1 40 members was calculated for each 121 30-year periods. A total of 121 σ were thus obtained for each multi-member ensemble and each climate variable.
5. The ICV was defined as ±2σ in this study. Since ICV of mean climate variables can be approximated as a normal distribution, the probability for climate change to exceed +2σ due to ICV is only 2.3%, with the same is for -2σ (Hansen, Sato, & Ruedy, 2012). A similar definition of ICV was also used in other studies (e.g. Hulme et al., 1999; IPCC, 2013; Mahlstein et al., 2011; Zhuan et al., 2018).
6. The same procedures were carried out for both annual and seasonal mean precipitation and temperature for all three multi-member ensembles separately and for all grid points.

IV.3.2 Estimation of the anthropogenic climate change

The ACC was defined as the difference (mm for precipitation and °C for temperature) of a specific climate variable between the reference period and a 30-year period of interests in terms of multi-model ensemble mean. The 1921-1950 period was used as a reference and the period of interests were calculated using a moving window method between the 1951-2100 periods. Specifically, the ACC was calculated based on the following five steps.

1. The multi-model ensemble means were calculated using the combination of 17 GCM's single member and the first member of the three multi-member ensembles, totaling 20 climate simulations (for annual or seasonal mean precipitation and temperature).
2. The 150-year (1951-2100) climate simulations were divided into 121 30-year periods (i.e. 1951-1980, 1952-1981, 1953-1982, …, 2069-2099 and 2071-2100) with a one-year lag for the time series of ensemble mean.
3. The 30-year running means of annual and seasonal mean precipitation and temperature were calculated for each variable based on the time series of the ensemble mean. A total of 121 mean values were obtained for each variable and one mean value for the reference period (1921-1950).
4. The ACC was defined as the difference between each of 121 mean values and the mean value of the reference period. A total of 121 ACC values were obtained for each variable.
5. The above procedures were carried out for annual and seasonal mean precipitation and temperature and all grid points.

This study defines the ACC as the change of the multi-model ensemble mean. However, a single climate model simulation may not be reliable in estimating the ACC due to the inter-model uncertainty and ICV. The ensemble mean was considered to be more reliable, since inter-model uncertainty and ICV are largely muted (Giorgi & Bi, 2009; Mahlstein et al., 2011; Maraun, 2013b). The reliability of using a multi-model ensemble mean to estimate ACC has been previously demonstrated in several studies (e.g. Hawkins & Sutton, 2009; Kay et al., 2015), and especially in IPCC assessment reports (IPCC, 2013).

IV.3.3 Estimation of the time of emergence

The importance of ICV relative to ACC was quantified using the ToE as a criterion. The ToE was defined in this study as the moment when ACC emerges from the ICV during a

particular 30-year period and all subsequent periods. In this study, 121 ICV values (±2σ of detrended multi-member ensemble) form one curve, and 121 ACC values (changes of multi-model ensemble mean) form another curve. The intersection of these two curves was identified as the ToE. Figure-A IV-1 shows examples of the estimation of ToE for four grid points with different ToE time evolutions. Each blue curve shows the temporal evolution of the mean annual precipitation change for a single climate simulation (20 in total – one for each GCM). Each curve was constructed using 121 ACC values. The red curve is the ensemble mean of ACC (mean of all blue curves). The green curve shows the ICV represented in terms of ±2σ with the solid curve for +2σ and the dash curve for -2σ. For the first example (Figure-A IV-1-A), an increasing ACC emerges from ICV (+2σ) during the twenty-first century and in the second example (Figure-A IV-1-B) a decreasing ACC emerges from ICV (-2σ) during the twenty-first century. Figure-A IV-1-C shows an example in which the ACC does not emerge from ICV during the study period. In the last example (Figure-A IV-1-D), the ACC emerges from ICV before the 1951-1980 period, because the ACC is strong at the beginning where the ICV is relatively small. Using this method, the ToEs were calculated for all grid points across the world. Since climate model simulations have different spatial resolutions, all simulations were linearly interpolated to a common resolution (1.0°×1.0° on latitude and longitude) before calculating ToEs.

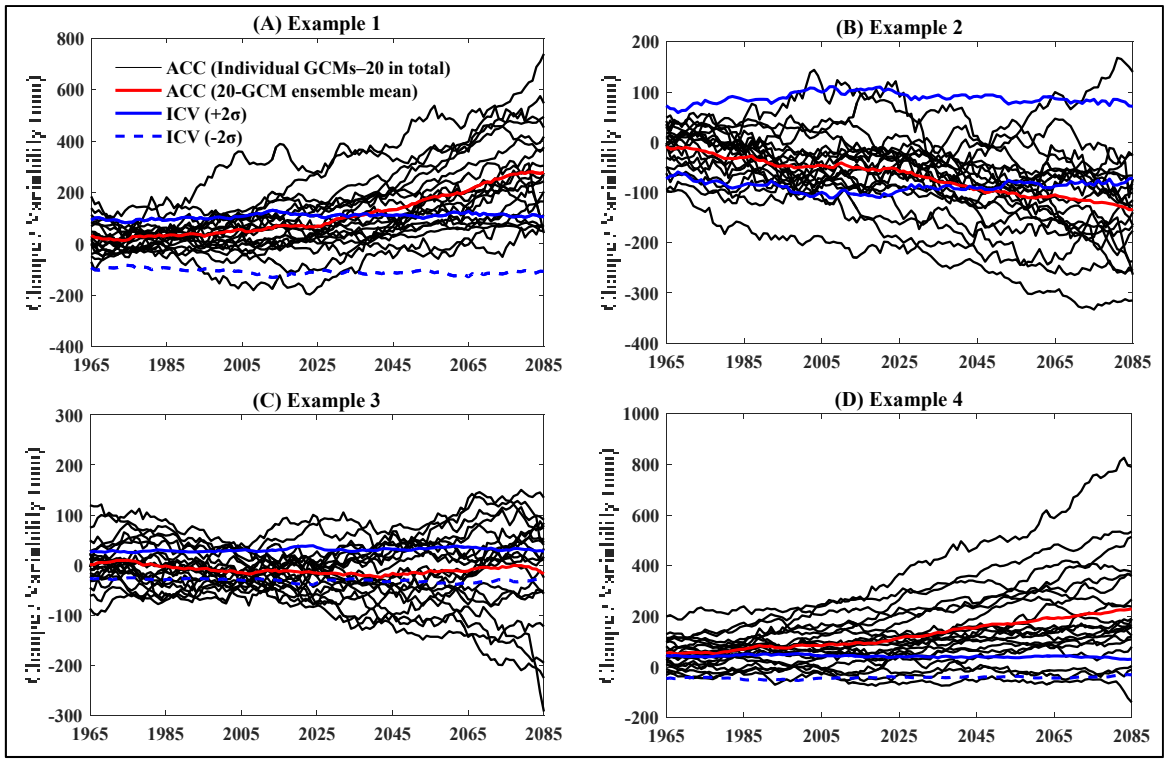

Figure-A IV-1. Determination of the Time of Emergence (middle of the 30-year period) for annual mean precipitation over four grid points. Each black curve shows the temporal evolution of the mean annual precipitation change for a single climate simulation (20 in total – one for each GCM). Each curve was constructed using 121 anthropogenic climate change (ACC) values. The red curve is the ensemble mean of ACC (mean of all black curves). The blue curve shows the internal climate variability (ICV) represented in terms of ±2 standard deviation (±2σ) with the solid curve for +2σ and the dash curve for -2σ. Examples 1 to 4 respectively show (1) an increasing ACC emerges from ICV during the twenty-first century, (2) a decreasing ACC emerges from ICV during the twenty-first century, (3) the ACC does not emerge from ICV during the study period, and (4) the ACC emerges from ICV before the 1951-1980 period

IV.4. Results

IV.4.1 Estimation of the internal climate variability and anthropogenic climate change

Figure-A IV-2 presents mean values of ICV for annual mean precipitation and annual mean temperature at the multi-decadal scale respectively estimated using the three climate model

multi-member ensembles. Following Chen et Brissette (2018), these mean values were obtained by averaging ICV (i.e. 2σ) of annual mean precipitation and temperature over 121 periods. Generally, the precipitation ICV at the multi-decadal scale is larger in the tropical than in temperate and arctic regions (see Figure-A IV-2). This is an expected pattern, as the precipitation ICV is directly related to its magnitude. In other words, larger precipitation amounts are associated to larger variability. The largest precipitation ICV occurs in Southeast Asia (SSA) region, which is one of the world's region with the most precipitation. In contrast to precipitation, temperature shows its highest ICV over arctic regions and lowest ICV for near-tropical regions and the Southern Hemisphere. At mid-latitudes in the northern hemisphere, the temperature ICV tends to be larger over land surfaces than over the oceans. This is partly because the land absorbs heat faster than the water in the summer, and releases heat faster in the winter. Consequently, land surfaces have higher temperature than ocean in the summer, with the opposite behavior observed in the winter. In addition, estimated ICV is very similar for the three ensembles, despite the wide range of members present (5, 10 and 40). This suggests that as few as 5 members may be sufficient to estimate ICV at the multi-decadal scale for annual mean precipitation and temperature.

Figure-A IV-3 presents the ACC for three 30-year future periods (2011-2040, 2041-2070 and 2071-2100), as represented by the multi-model ensemble mean of annual mean precipitation and temperature. Differently from Section IV.3.2, the ACC is calculated as the relative change (%) rather than the absolute change (mm) for precipitation in Figure-A IV-3. Generally, the ACC of precipitation and temperature consistently becomes stronger as the time period further progresses. The climate model ensemble predicts increases in annual mean precipitation for most regions to the south of the Tropics of Capricorn and to the north of the Tropics of Cancer, as well as along the equator. For other regions between Tropics of Capricorn and Cancer, the climate model ensemble predicts a possible decrease in annual mean precipitation. The Double Intertropical Convergence Zone (DITCZ) problem is clearly observed, which is characterized by the overproduction of DITCZ events (Wang, Lee, Chen, & Hsu, 2015; Zhang, Lin, & Zhang, 2007). Along the equator, the estimator of ACC may not

be reliable. Additionally, the climate model ensemble predicts an increasing trend in annual mean temperature for all grid points. The strongest trends are observed over Arctic Ocean areas with the exception of Greenland (GRL), followed by mid-latitude land areas. In addition, temperature trends tend to be larger over land areas than over adjacent oceans. Overall, not surprisingly, these results are very close to the changes presented in the fifth IPCC assessment report (IPCC, 2013).

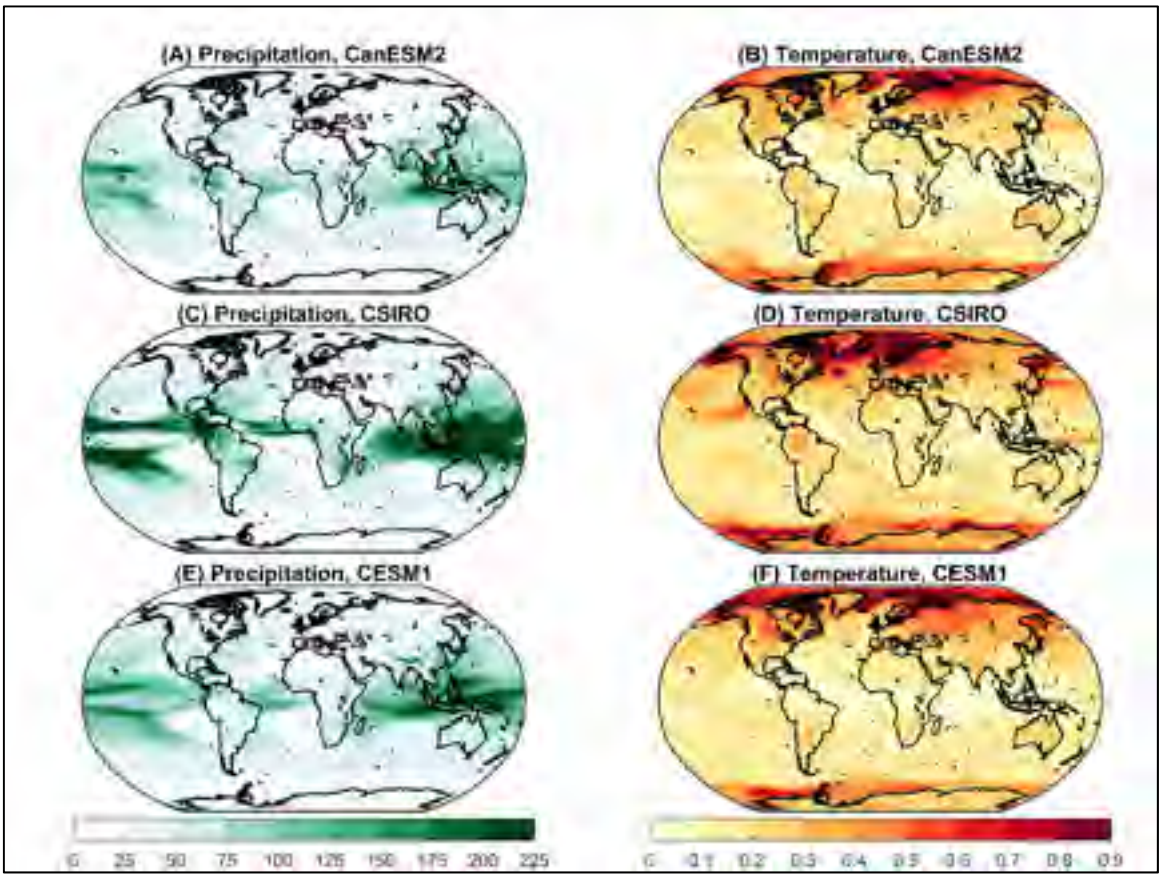

Figure-A IV-2. Averaged ICV (2σ) of annual mean precipitation (mm) and temperature (°C) estimated using three multi-member climate model ensembles (CanESM2, CSIRO and CESM1)

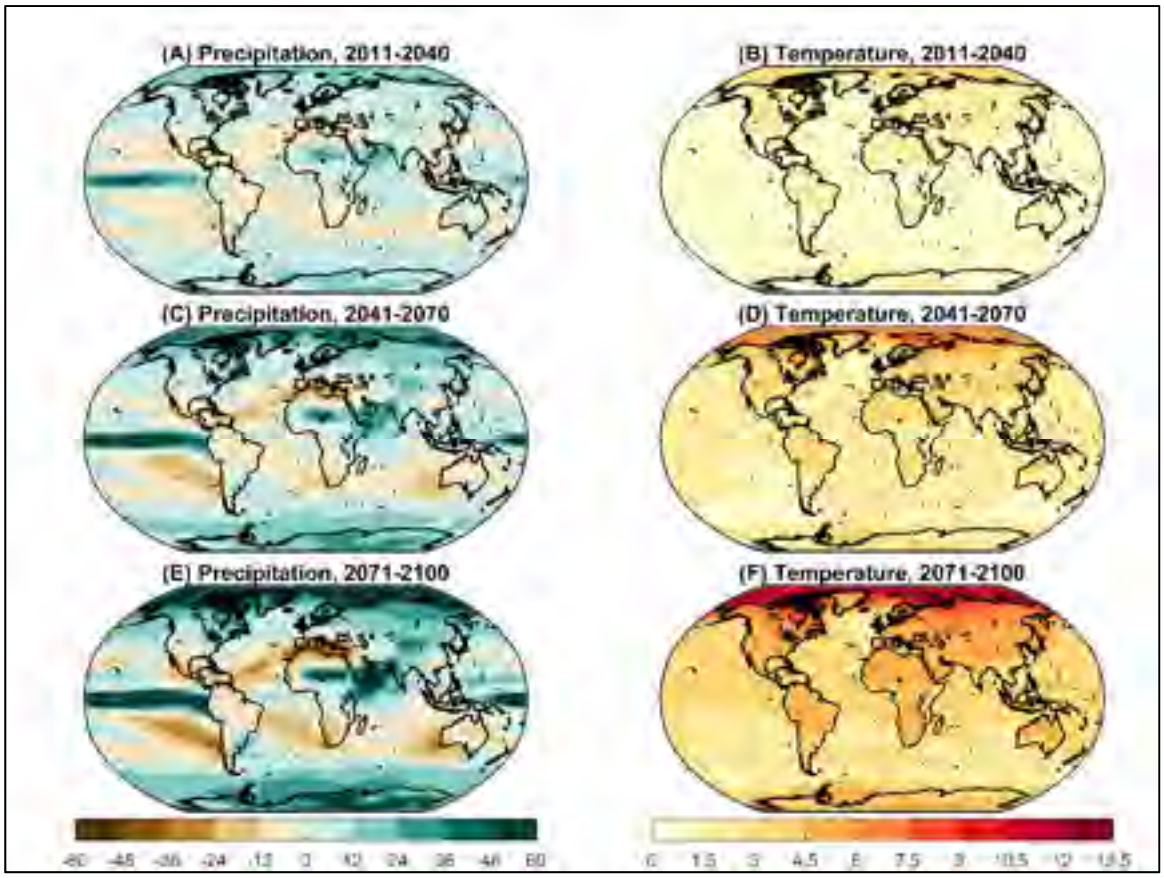

Figure-A IV-3. Anthropogenic climate change (ACC) in % for annual mean precipitation (left) and in °C for annual mean temperature (right) of three future periods (2011-2040, 2041-2070 and 2071-2100) estimated using multi-model ensembles

IV.4.2 Time of emergence of annual precipitation and temperature

Figure-A IV-4 presents the ToE for annual mean precipitation. The ICV is estimated using the three different multi-member ensembles (i.e. CanESM2, CSIRO and CESM1), while the ACC is identical for the three rows of Figure-A IV-4. Generally, all three multi-member ensembles show that the ToE of annual mean precipitation happens earlier for high-latitude and tropical regions than for mid-latitude regions. In other words, for extra-tropical regions, the precipitation ACC has or will emerge from precipitation ICV before the end this century. However, for most tropical regions (with the exception of the equatorial band), ICV will

dominate annual mean precipitation till the end the 21st century. This pattern is in line with the projected change of the precipitation magnitude as presented in Figure-A IV-3, as well as in the IPCC fifth Assessment Report (IPCC, 2013). This is expected, because a stronger climate change trend implies an earlier ToE for a similar ICV. In extra-tropical regions, the ICV of annual mean precipitation is relatively weak, while the ACC is relatively strong. When it comes to the equator, both the ICV and ACC are shown to be strong. However, ToEs along tropical lines may not be reliable, because of the DITCZ problem as mentioned earlier. CanESM2 predicts earlier ToEs for the middle of Atlantic Ocean, while CSIRO and CESM1 predict earlier ToEs for the Middle East and Western Africa (WAF). Overall, all three multi-member ensembles show a similar ToE pattern, implying that all three ensembles perform similarly in terms of simulating the ICV (Chen & Brissette, 2018).

At the regional scale, all three multi-member ensembles generally show similar patterns, with the exception of the Sahara (SAH) where CanESM2 shows larger ToE variability compared to the other two ensembles. All three ensembles predict that the ToE did not emerge by the 2071-2100 period for Australia (AUS). The ToE is also late in Southeast Asia (SEA). These two regions are both located in mid- and low-latitudes. For Europe (EU in Figure-A IV-4), the ToE mostly occurs before the middle of the twenty-first century with earlier emergence for Northern Europe (NEU) than the Mediterranean Basin (MED). Similar patterns are also observed for Eastern North America (ENA), Alaska (ALA), Greenland (GRL), Tibet (TIB) and North Asia (NAS). In particular, the ToE in TIB is earlier than other regions around it. This may be because the climate is more sensitive to GHG emission for the Qinghai-Tibet Plateau than for other areas (e.g. Liu & Chen, 2000; Luo, Jiang, & Tang, 2015). A lower climate model performance for high elevation area is also a possibility (Lee, Hong, Chang, Suh, & Kang, 2014). For all other regions, the ToE can be as earlier as the middle of this century or the later than the end of this century.

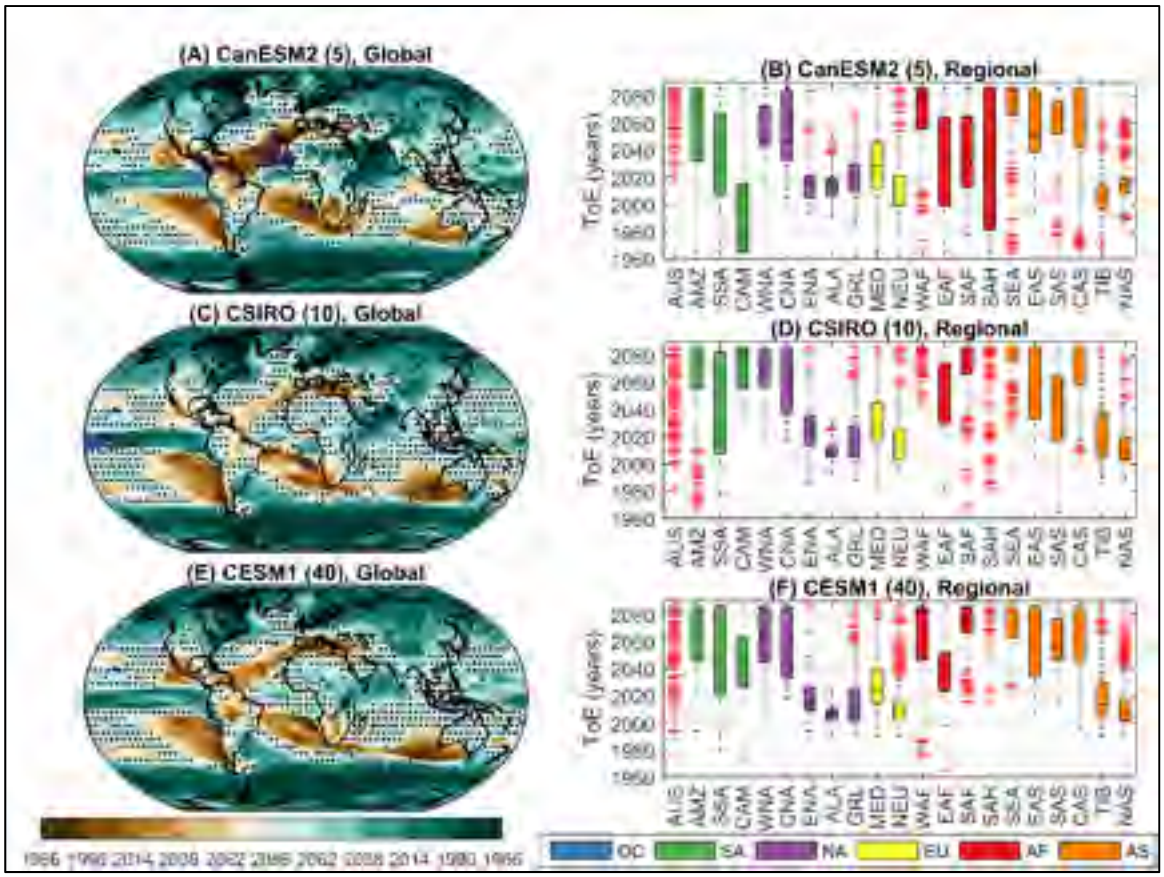

Figure-A IV-4. Time of emergence (middle of the 30-year period) of annual mean precipitation for three climate models (CanESM2, CSIRO and CESM1) (left) and that over 21 different regions (Giorgi and Francisco, 2000) in all land grid points of the world (right). In (A), (C) and (E), the brown colors show decreases in annual mean precipitation with darker color signifying earlier emergence, and blue-teal colors show increases in annual mean precipitation with darker color indicating earlier emergence. The black dots represent grid points where the ACC did not emerge from the ICV by the 2071-2100 period, while the blue crosses represent grid points where ACC emerged from the ICV before the 1951-1980 period. The full name and coverage of each region are presented in Table-A IV-2

Table-A IV-2. The list of 21 regions following the work of (Giorgi & Francisco, 2000)

Continent	Region	Acronym	Latitude (º)	Longitude (º)
Oceania (OC)	Australia	AUS	45S-11S	110E-155E
South America (SA)	Amazon Basin	AMZ	20S-12N	82W-34W
	Southern South America	SSA	56S-20S	76W-40W
	Central America	CAM	10N-30N	116W-83W
North America (NA)	Western North America	WNA	30N-60N	130W-103W
	Central North America	CAN	30N-50N	103W-85W
	Eastern North America	ENA	25N-50N	85W-60W
	Alaska	ALA	60N-72N	170W-103W
	Greenland	GRL	50N-85N	103W-10W
Europe (EU)	Mediterranean Basin	MED	30N-48N	10W-40E
	Northern Europe	NEU	48N-75N	10W-40E
Africa (AF)	Western Africa	WAF	12S-18N	20W-22E
	Eastern Africa	EAF	12S-18N	22E-52E
	Southern Africa	SAF	35S-12S	10W-52E
	Sahara	SAH	18N-30N	20W-65E
Asia (AS)	Southeast Asia	SEA	11S-20N	95E-155E
	East Asia	EAS	20N-50N	100E-145E
	South Asia	SAS	5N-30N	65E-100E
	Central Asia	CAS	30N-50N	40E-75E
	Tibet	TIB	30N-50N	75E-100E
	North Asia	NAS	50N-70N	40E-180E

Figure-A IV-5 presents the ToE of annual mean temperature for three climate model multi-member ensembles. For most regions in the world, the annual mean temperature ACC has emerged from the ICV during the last century. Regions with a later ToE are mostly located in the north Pacific and Atlantic Ocean, and especially south of GRL. For this region, the positive trend of annual mean temperature is not obvious even under the RCP8.5 emission scenario. Under lower emission scenarios (e.g. RCP2.6), this region may be cooler in the future as shown in IPCC Assessment Report (IPCC, 2013). In this case, the temperature change may emerge from ICV even later or not at all.

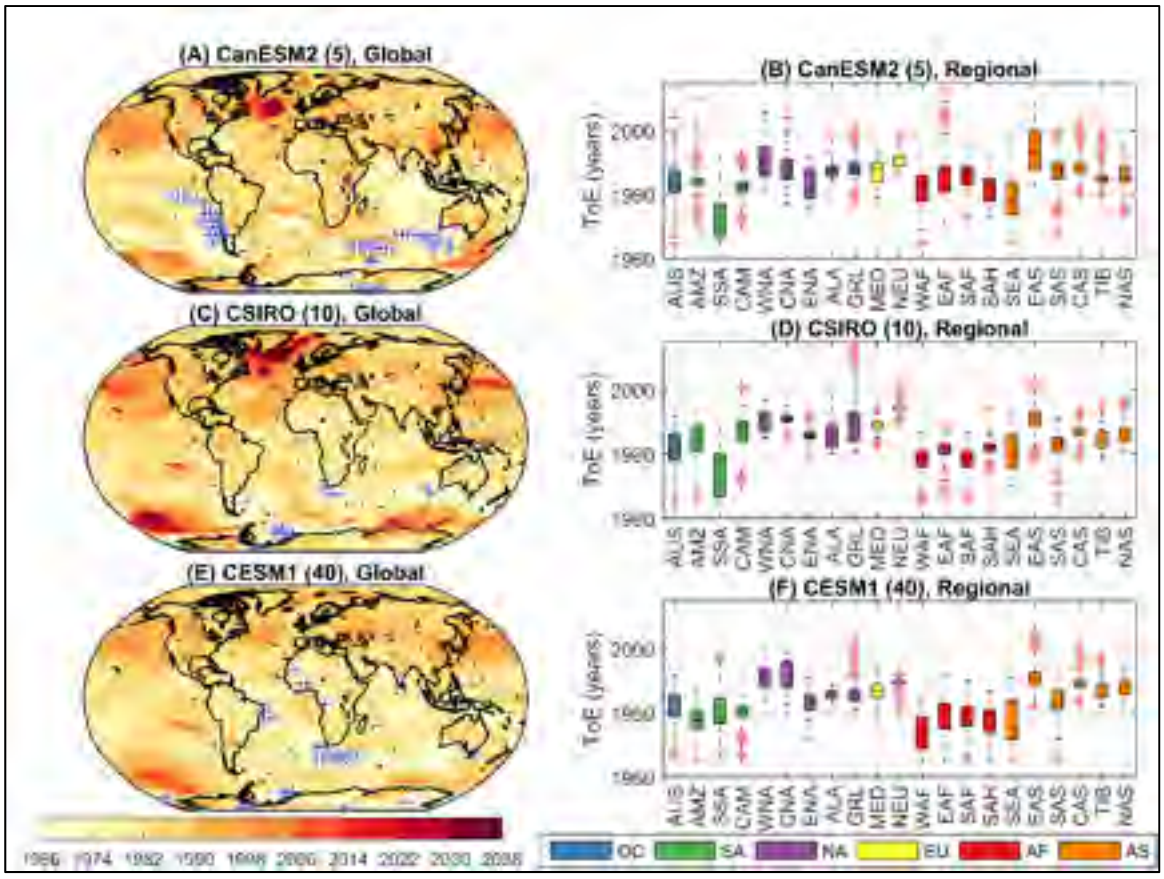

Figure-A IV-5. Time of emergence (middle of the 30-year period) of annual mean temperature for three climate models (CanESM2, CSIRO and CESM1) (left) and that over 21 different regions (Giorgi and Francisco, 2000) in all land grid points of the world (right). In (A), (C) and (E), the blue crosses represent grid points where ACC emerged from the ICV before the 1951-1980 period. The full name and coverage of each region are presented in Table-A IV-2

Overall, the annual mean temperature ACC projected by the RCP8.5 scenario will emerge from ICV before the 2023-2052 period for all grid points in all three multi-member ensembles. This implies that the annual mean temperature ICV is small compared to the increasing trend of temperature. As presented in Figure-A IV-2 and Figure-A IV-3, the annual mean temperature ICV (i.e. 2σ) is less than 1°C for most grid points. However, the temperature increase can be larger than 10°C under the RCP 8.5 scenario for some regions (e.g. arctic region). For example, some regions in the south Indian Ocean, Antarctic and to

the northwest of South America, have seen a ToE as early as the 1951-1980 period. The boxplots in Figure-A IV-5 show that the annual mean temperature ACC has emerged from the ICV for most land grid points. The ToE are mostly between the 1965-1994 and 1985-2014 periods. The ToE in Southern South America (SSA) and Southeast Asia (SEA) is generally earlier than for other regions. As it was with the case of annual mean precipitation, ICV estimates for annual mean temperature tend to be similar over all three ensembles (Figure-A IV-2), indicating that 5 members may be sufficient to reliably estimate multi-decadal temperature ICV.

IV.4.3 Time of emergence of seasonal precipitation and temperature

ToEs of precipitation and temperature were also calculated at the seasonal scale. Figure-A IV-6 presents the ToEs of June-July-August (JJA) and December-January-February (DJF) mean precipitation. A comparison of Figure-A IV-4 and Figure-A IV-6 shows that the JJA mean precipitation is predicted to decrease for more regions than the annual mean precipitation, while the opposite is observed for DJF mean precipitation. Again, all three multi-member ensembles display similar ICV of seasonal precipitation at the multi-decadal scale. Regarding to the ToE, the JJA and DJF precipitation changes will emerge from ICV later than the annual mean precipitation for some regions (e.g. mid-latitude region). This is especially true for JJA precipitation, due to a weaker ACC as presented Figure-A IV-7. Figure-A IV-7 shows that the ICV is similar between JJA and DJF, while the ACC in DJF is much stronger than that of JJA for most of regions. This explains the later JJA ToE compared to DJF. For most of the world's regions, the seasonal precipitation ACC will not emerge from ICV before the end of twenty-first century. Precipitation changes are expected to emerge during this century only for high-latitude regions in both hemispheres. Overall, this shows that ICV is larger at the seasonal scale than at the annual scale, which results in overall later ToE.

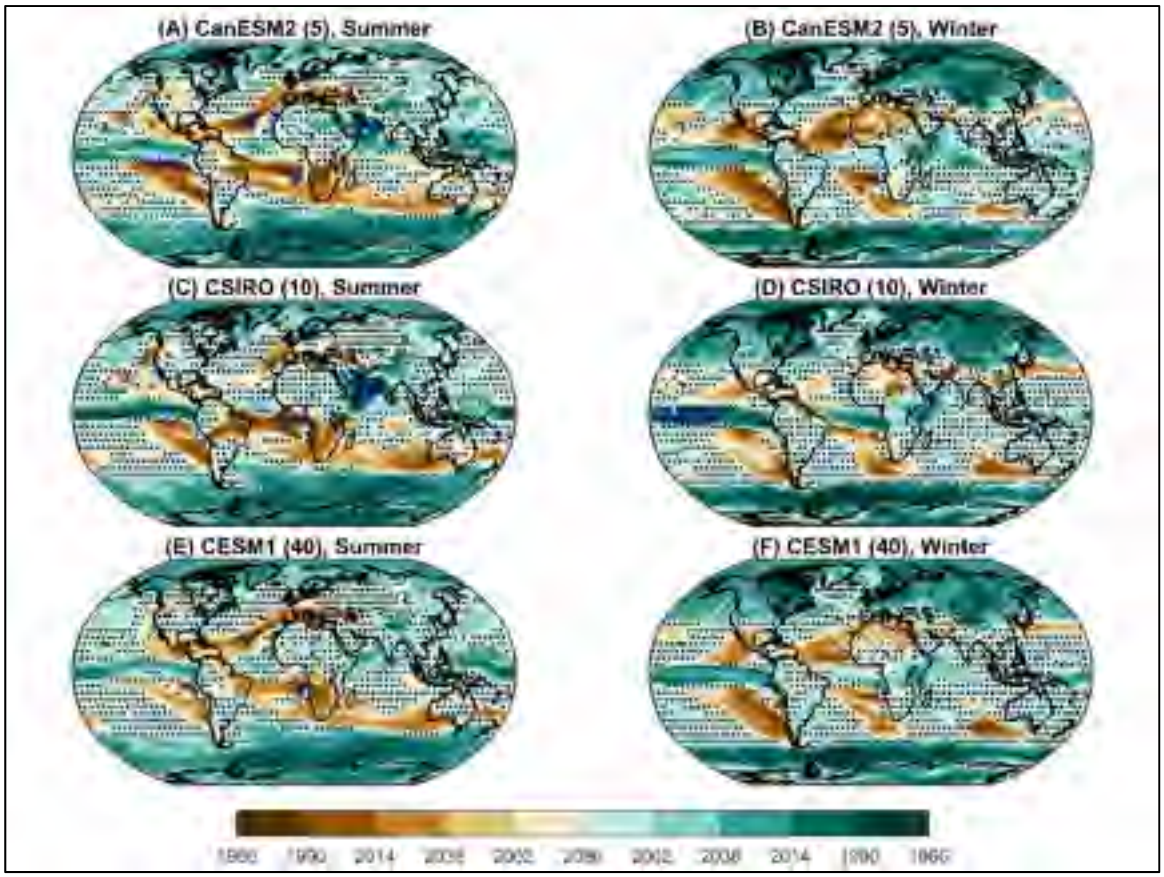

Figure-A IV-6. Time of emergence (middle of the 30-year period) of June-July-August (JJA, left) and December-January-February (DJF, right) mean precipitation for three climate models (CanESM2, CSIRO and CESM1). The brown colors show decreases in annual mean precipitation with darker colors indicating earlier ToE, and blue-teal colors show increases in annual mean precipitation with darker color showing earlier ToE, the black dots represent grid points where the ACC did not emerge from the ICV by the 2071-2100 period and the blue crosses represent regions where ACC emerges from the ICV during the 1951-1980 period

Figure-A IV-8 presents the ToEs of JJA and DJF mean temperature. For most regions of the world, the ACC has already emerged from the ICV for seasonal mean temperature. The ToE will not be later than the 2023-2052 period for most of global grid points. Generally, the ToE of seasonal temperature is later than that of annual temperature for most regions in the world. This is especially true for DJF temperature, which shows larger variability than for annual temperature (see Figure-A IV-3 for annual and Figure-A IV-9 for seasonal). Similar to

annual mean temperature, JJA and DJF mean temperatures are projected to increase for all grid points (Figure-A IV-9). Again, all three multi-member ensembles show similar behavior in terms of estimating the ICV of JJA and DJF temperatures at the multi-decadal scale. The more temperature increase in DJF than in JJA (and at the annual scale) cannot outweigh the larger temperature ICV in the DJF.

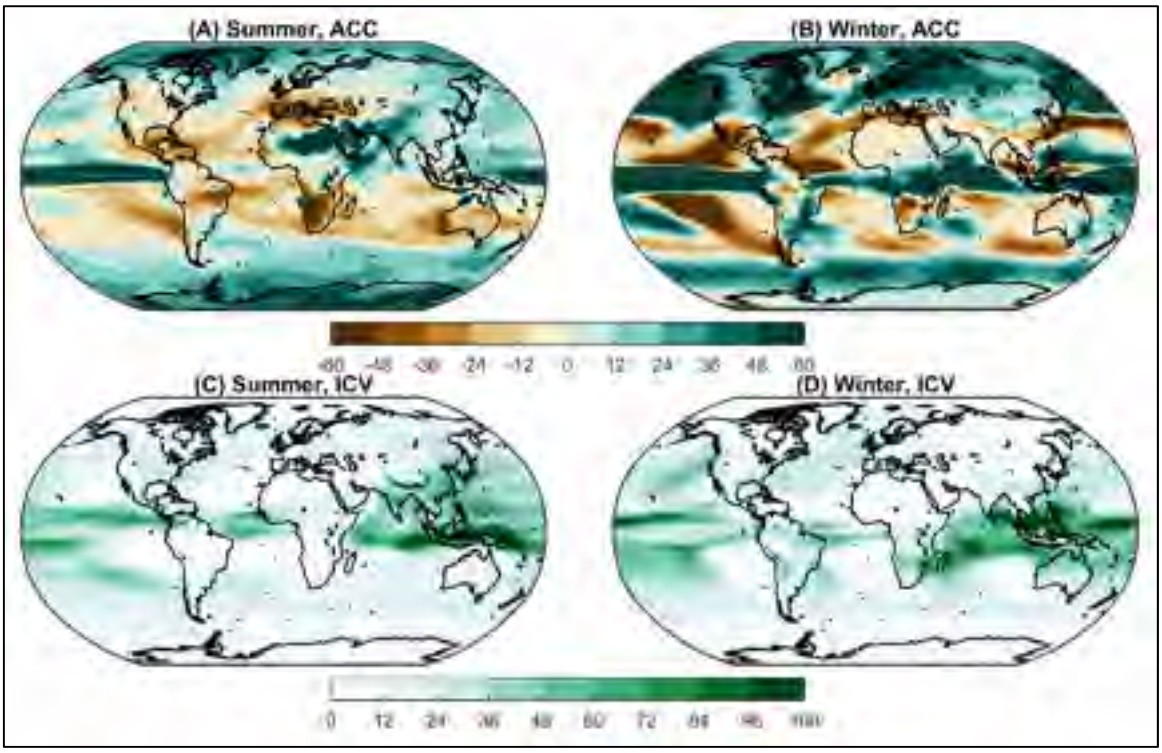

Figure-A IV-7. Anthropogenic climate change (ACC, %) and averaged internal climate variability (ICV, mm) June-July-August (JJA, left) and December-January-February (DJF, right) mean precipitation projected by CESM1. The ACC is only for the 2071-2100 period, while ICV is an average value for all periods

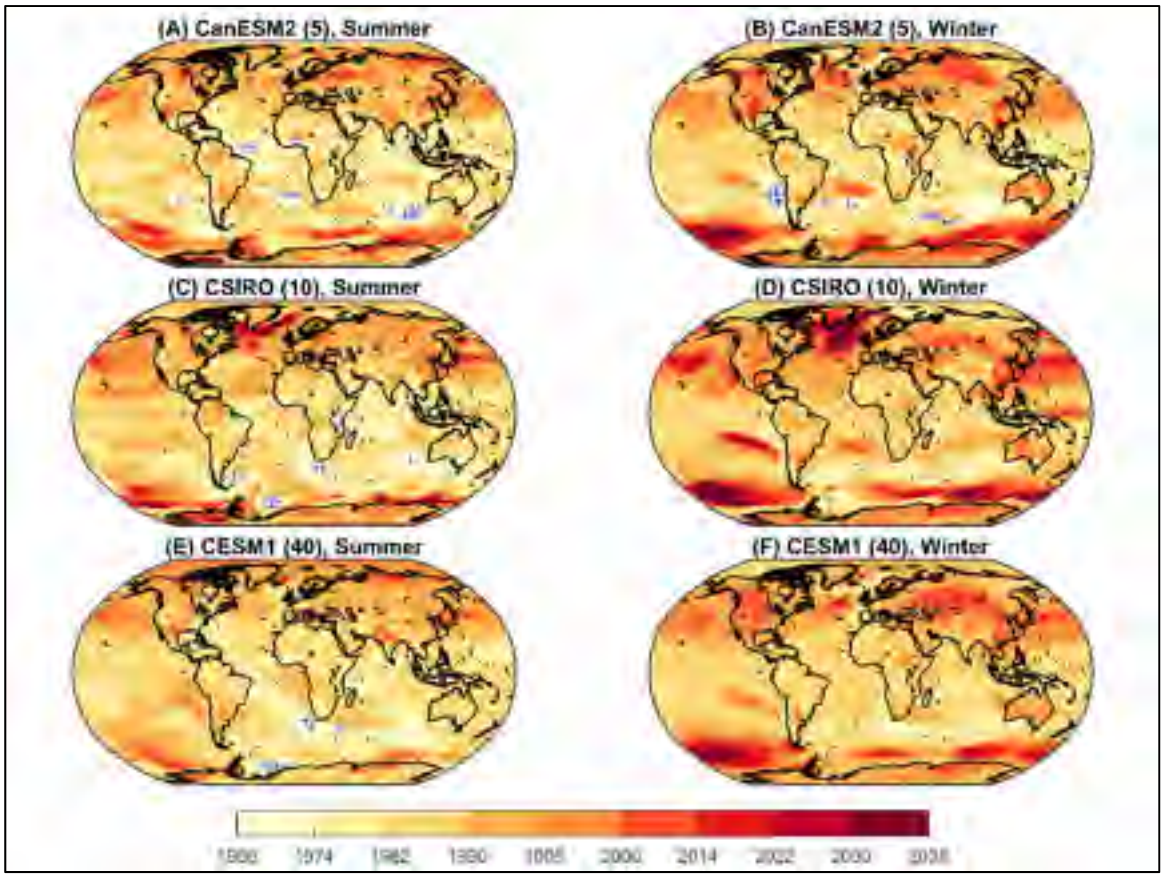

Figure-A IV-8. Time of emergence (middle of the 30-year period) of June-July-August (JJA, left) and December-January-February (DJF, right) mean temperature for three climate models (CanESM2, CSIRO and CESM1). The blue crosses represent regions where climate change signal emerges from the ICV during the 1951-1980 period

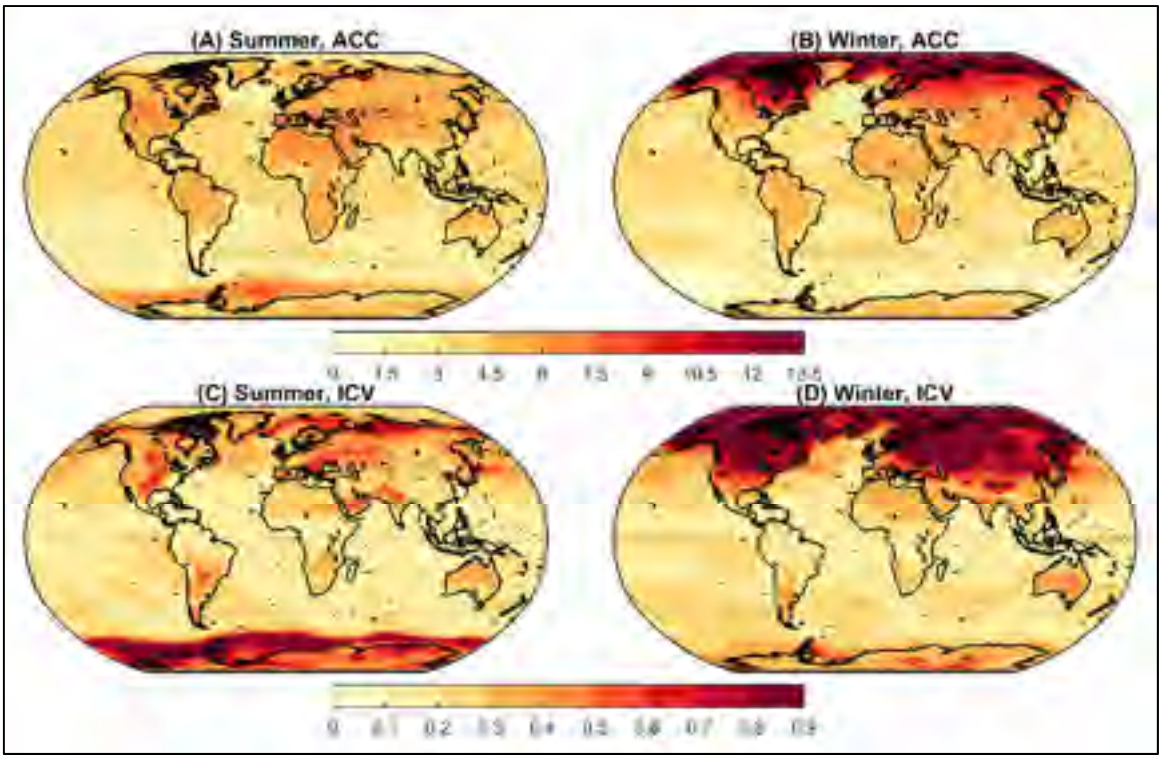

Figure-A IV-9. Anthropogenic climate change (ACC, °C) and averaged internal climate variability (ICV, °C) of June-July-August (JJA, left) and December-January-February (DJF, right) mean temperature projected by CESM1. The ACC is only for the 2071-2100 period, while ICV is an average value for all periods

IV.5. Discussion

The ICV and ACC are the two main components that drive the overall climate change at the multi-decadal scale. Quantifying the relative importance of both components allows for a better understanding of the effects of ICV and ACC on climate change, and their impacts on environmental vulnerability to climate. This information is critical to design more robust adaptation strategies and mitigation measures to lessen the impacts of climate change. A better knowledge of ICV and ensuing ToE provides key information to decision makers. For example, adaptation could be the key strategy for vulnerable regions with large ICV, whereas mitigation would be more useful to regions where ACC is shown to be dominant based on the ToE. Based on the combination of one member of multiple climate model and three

multi-member ensemble of a single climate model, this study contrasted the roles of ICV and ACC at the multi-decadal scale using the ToE as a criterion.

The ICV was defined as the inter-member variability of annual and seasonal precipitation and temperature at the multi-decadal scale using the detrended multi-member ensembles. In other words, the ICV was considered as the inter-member multi-decadal variability after the removal of ACC trends. Even though the multi-decadal variability of precipitation and temperature is only one manifestation of ICV for these two specific variables, it may be the most important one for climate change impact studies, as the climate change impact studies are usually conducted at the multi-decadal scale (e.g. 30 years). In order to estimate the ICV, the ACC was first removed using a two-stage linear regression approach. The reasonable performance of the two-stage detrending method has been verified in previous studies (e.g. Chen & Brissette, 2018; Zhuan et al., 2018). Even though the nonlinear approach may be better at fitting the climate time series (e.g. the use of fourth-order polynomial in Hawkins et Sutton (2009)), there is a danger of overfitting the data and removing the inherent ICV.

The use of multi-member ensembles to estimate the ICV is not new in climate change studies (Deser et al., 2012a; Deser et al., 2012b; Martel et al., 2018). Since all ensemble members are simulated by the same climate model under the same climate forcing, the only difference remain in the initial conditions. Thus, the variability or range of an ensemble provides insight into what could happen in a single realization that will occur in the real world (Deser et al., 2012a; Deser et al., 2012b). While ICV is purely evaluated in the world of climate models, Chen et Brissette (2018) showed that the multi-member ensemble performs reasonably well in terms of capturing the observed pattern of multi-decadal variability for annual and seasonal mean precipitation and temperature for all land grid points in the world. This study showed that the three estimates of ToEs are very similar, despite the wide range of members present (5, 10 and 40). This further indicates the reliability of the estimated ICV and suggests that as few as 5 members may be sufficient to estimate ICV at the multi-decadal scale for annual and seasonal mean precipitation and temperature. Martel et al. (2018) also showed

that the inter-annual variability exhibited by members of two large ensembles (CESM1 and CanESM2) was similar to that observed over the recent past for annual mean precipitation.

Other methods have also been used to estimate the ICV. For example, Hawkins et Sutton (2009, 2011) defined the ICV as the variances of precipitation and temperature time series which were calculated based on each detrended climate simulation at the decadal scale. In other words, the ICV was defined as the decadal variability of climate simulations. Previous studies (e.g. Chen & Brissette, 2018) have showed that the inter-member variability estimated using multi-member ensembles is similar to multi-decadal variability estimated using a single climate simulation for annual and seasonal mean precipitation and temperature. In addition, the estimation of ICV using a single climate simulation assumes that the ICV is constant over time, however, this assumption may not be hold due to the limited length of the climate simulations. The use of inter-member variability to represent the ICV releases the constant assumption. Other studies (e.g. Giorgi & Bi, 2009) also defined the ICV as the combination of inter-model and inter-member variability. However, the ICV and inter-model variability are different in origin, as the former reflects the chaotic nature of the climate system, while the latter reflects the model structural uncertainty associated with the understanding and parameterization of the climate models (Fischer et al., 2013; Hawkins & Sutton, 2011; Kim et al., 2016; Maraun, 2013b).

The ACC was defined as the multi-model ensemble mean, which is a common method used in IPCC fifth Assessment Report (IPCC, 2013) and many other studies (e.g. Hawkins & Sutton, 2009; Mahlstein et al., 2011; Maraun, 2013b). Although a single climate model simulation may be biased with respect to reproducing the real-world climate, the multi-model ensemble mean can be considered to be more reliable, since the inter-model uncertainty and ICV are largely averaged out (Mahlstein et al., 2011; Maraun, 2013b). Moreover, the calculation of 30-year running mean further filters out the short-term ICV, such as the inter-annual and decadal variability. The estimated ACC for the future period is dependent on the GHG emission scenario (i.e. the RCP). In other words, GHG emission scenarios may affect

the timing when ACC emerging from the ICV. Generally, higher emission scenarios are likely to result in larger increases in temperature and larger projected changes (increase or decrease) in precipitation than lower emission scenarios (Hawkins & Sutton, 2009, 2011). This implies that higher emission scenarios are expected to lead to earlier ToEs than lower emission scenarios. However, the use of different emission scenarios is not expected to change too much in the spatial pattern of ToEs, even though ToEs may appear later. In order to provide an earlier warning and make a conservative decision to adapt or mitigate climate change impacts, only one extreme GHG emission scenario is used to estimate ACC in this study.

With the estimated ICV and ACC, the ToE was defined as the intersection of these two components. Since both mean precipitation and temperature were calculated at the multi-decadal scale (30-year period) using a moving window approach in a one-year increment, the estimated ICV and ACC evolve smoothly overtime. Thus, the ToE can be easily identified. Moreover, the estimation of ToE at a multi-year period may be more reliable, as the 30-year moving window approach filters out the small fluctuations/changes and highlights the long-term climate change trend. Furthermore, the estimation of ToE at a 30-year period may be more useful for climate change impact studies, as the 30-year period is typically used for quantifying the climate change and its impacts.

Finally, results presented in this paper should not be extrapolated to other variables, and in particular for extremes. Several studies (e.g. Kharin et al., 2007; Wang et al., 2017) have shown that changes in precipitation and temperature extremes may be much larger than changes in mean values. As such, ToE may happen much earlier. As such, late ToE for mean precipitation should not be interpreted as climate change being inconsequential in the short-term, since trends in extreme precipitation, especially at the sub-daily time scale may be significantly more important.

IV.6. Conclusion

Based on multi-model and multi-member ensembles, this study quantified the importance of ICV relative to ACC at the multi-decadal scale. The ToE was used as a criterion to determine the 30-year period when the ACC emerges from ICV. The following conclusion can be drawn:

1. The multi-model climate ensemble predicts increases in annual and seasonal mean precipitation for most extra-tropical regions and a decrease for tropical regions, with the exception of the equatorial band. However, the annual and seasonal mean temperature is predicted to increase for all grid points. The ICV of mean precipitation is larger for tropical regions compared to extra-tropical regions. ICV of temperature is the largest for arctic regions. These results are fully consistent with the IPCC fifth assessment report (IPCC, 2013), which is not surprising since the ensemble used in this study comprises an important fraction of the CMIP5 ensemble.
2. The ACC of annual mean precipitation has already or will emerge this century for most extra-tropical regions. With the exception of the equatorial band, the ACC is unlikely to emerge from ICV during this century for most tropical regions. For annual mean temperature, the ACC has emerged from ICV during the last century or at the beginning of the current one for most regions in the world. The remaining regions (Northern Pacific and Atlantic Oceans), are likely emerge before the middle of this century.
3. Seasonal mean precipitation and temperature present similar ToE spatial pattern to those at the annual scale. However, the ToEs can be slightly later or earlier, due to different ICV and ACC. This is particularly obvious for JJA precipitation and DJF temperature. The weaker ACC of JJA precipitation and additional ICV of DJF temperature translate to later ToE compared to the annual scale.
4. This work showed that estimates of ICV from three multi-members ensemble were very similar, even though the number of members was widely different in the three

ensembles (5, 10 and 40). This appear to indicate that multi-decadal internal variability may be reliably estimated with as few as 5 members.

5. Overall, this study implies that adapting to ICV may be a good strategy to mean precipitation change in many regions of the world where ICV dominates the ACC, while mitigating ACC may be a good strategy to temperature change in many regions.

IV.7. Acknowledgement

This work was partially supported by the National Key Research and Development Program of China (Grant No. 2017YFA0603704), the National Natural Science Foundation of China (Grant No. 51779176, 51539009) and the Thousand Youth Talents Plan from the Organization Department of CCP Central Committee (Wuhan University, China). The authors would like to thank the contribution of the World Climate Research Program Working Group on Coupled Modelling for making available their climate model outputs as listed in Table-A IV-1, and the CEMS1 (CAM) large ensemble community project for providing the large ensemble of CESM1 data.

ANNEXE V

INDIVIDUAL AND COUPLED EFFECTS OF THE DOMINANT MODES OF NATURAL CLIMATE VARIABILITY ON SEASONAL PRECIPITATION AND TEMPERATURE OVER NORTH AMERICA

Magali Troin[a], Jean-Luc Martel[b], François Brissette[c]

[a, b, c] Département de Génie de la construction, École de technologie supérieure

Article en préparation pour soumission à la revue « Climate Dynamics ».

Abstract:

An evaluation of the relationships between large-scale oceanic and atmospheric oscillations on North America seasonal temperature and precipitation is performed to identify regions where observed climate variability is influenced by the natural climate variability. The natural variability of climate is explored on the basis of climatic indices on interannual and decadal time scales by considering individual and coupled effects of these indices on seasonal temperature and precipitation. The coupled effects of the climatic indices are evaluated for the two phases (positive versus negative) of the decadal (Arctic Oscillation – AO and North Atlantic Oscillation – NAO) and interannual (El Niño-Southern Oscillation – ENSO and Pacific North American pattern – PNA indexes within the two long-term phases of the interdecadal (Atlantic Multidecadal Oscillation – AMO and Pacific Decadal Oscillation – PDO) signals. The results show that the coupled impacts of AMO and PDO with NAO, AO, ENSO and PNA lead to stronger anomalies than individual impacts on seasonal temperature and precipitation. North America temperature variability is influenced by these climatic couplings, but the effects vary both in time following the phase of AMO and PDO (positive versus negative) and the season (winter versus summer), and in space (not the same regional patterns according to the AMO/PDO phase and the season). North America precipitation is identified to respond weaker than temperature to these climatic coupling, with lower anomaly values. The climatic indices modulate, to some extent, seasonal precipitation variability but their coupled effects interact temporally and spatially in more complex ways.

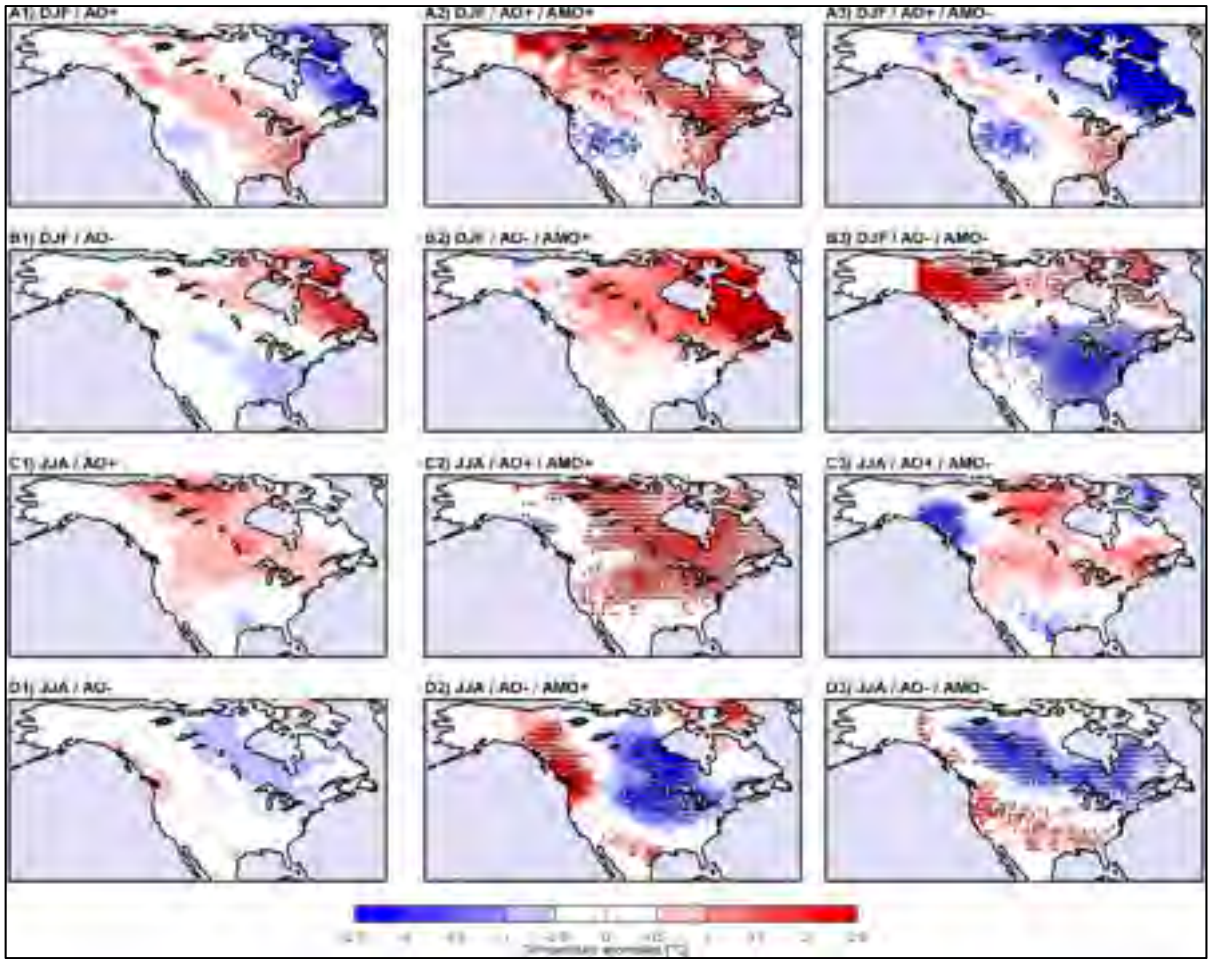

Figure-A V-1 Influence form AO and its coupled effect with AMO on North American mean temperature. Temperature anomalies (°C) are defined as the difference between the mean of the selected years based on the cold and/or warm phases of the climate indices and the one of all seasonal values. The black dots represent the statistical significance based on a bootstrap resampling approach with replacement using a total of 1 000 samples with a 95% confidence level

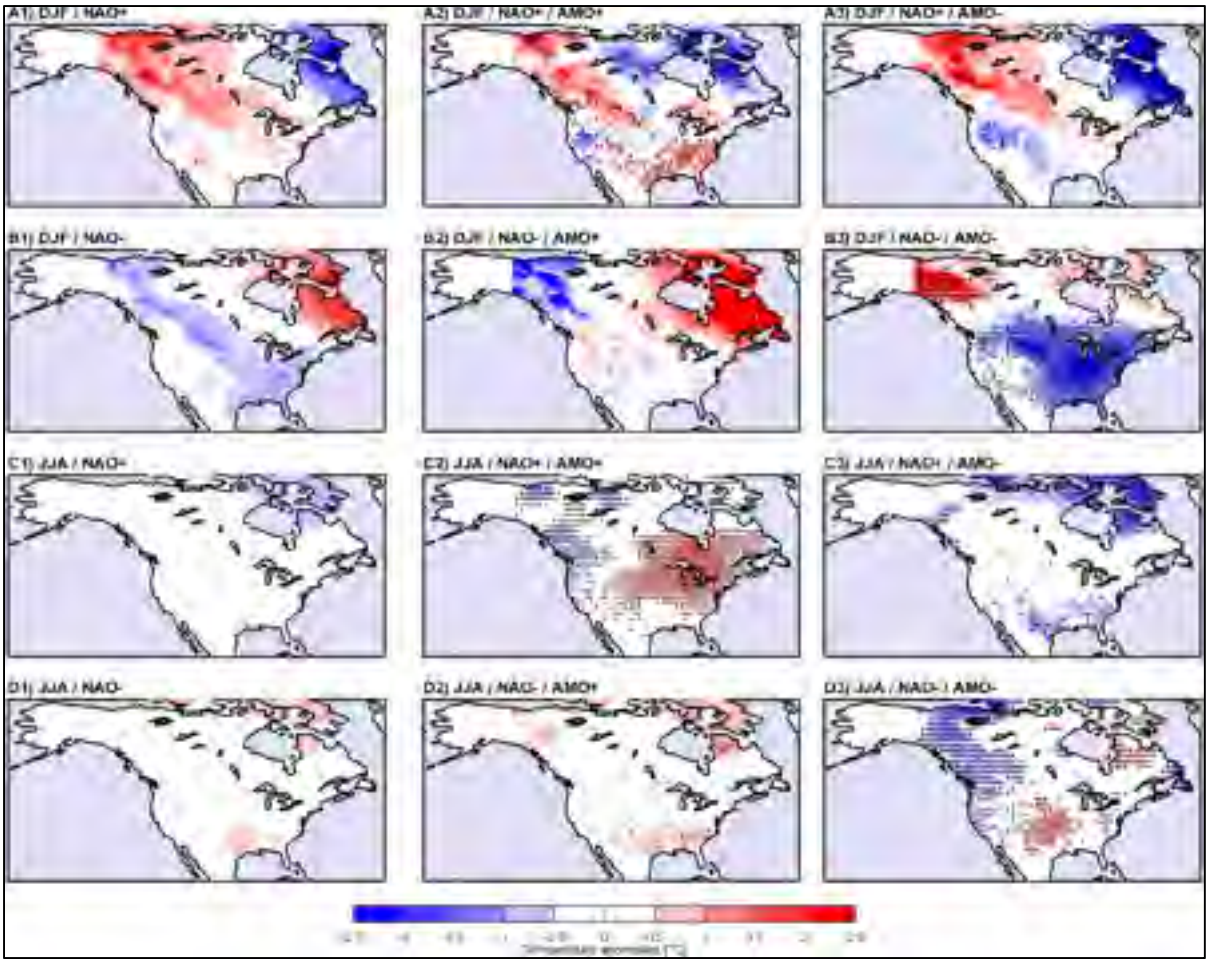

Figure-A V-2 Same as in Figure A V-1, but for AMO and NAO

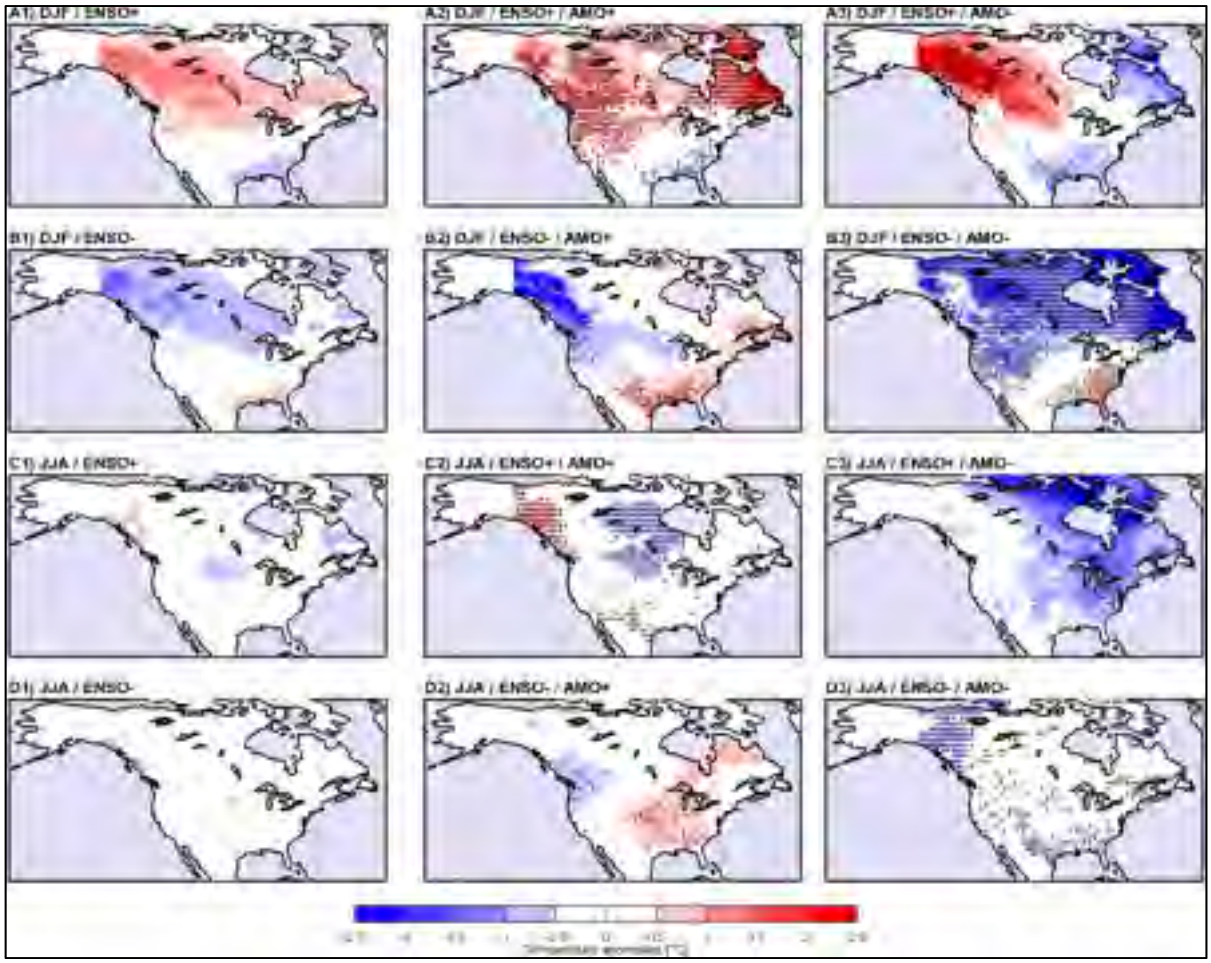

Figure-A V-3 Same as in Figure A-V-1, but for AMO and ENSO

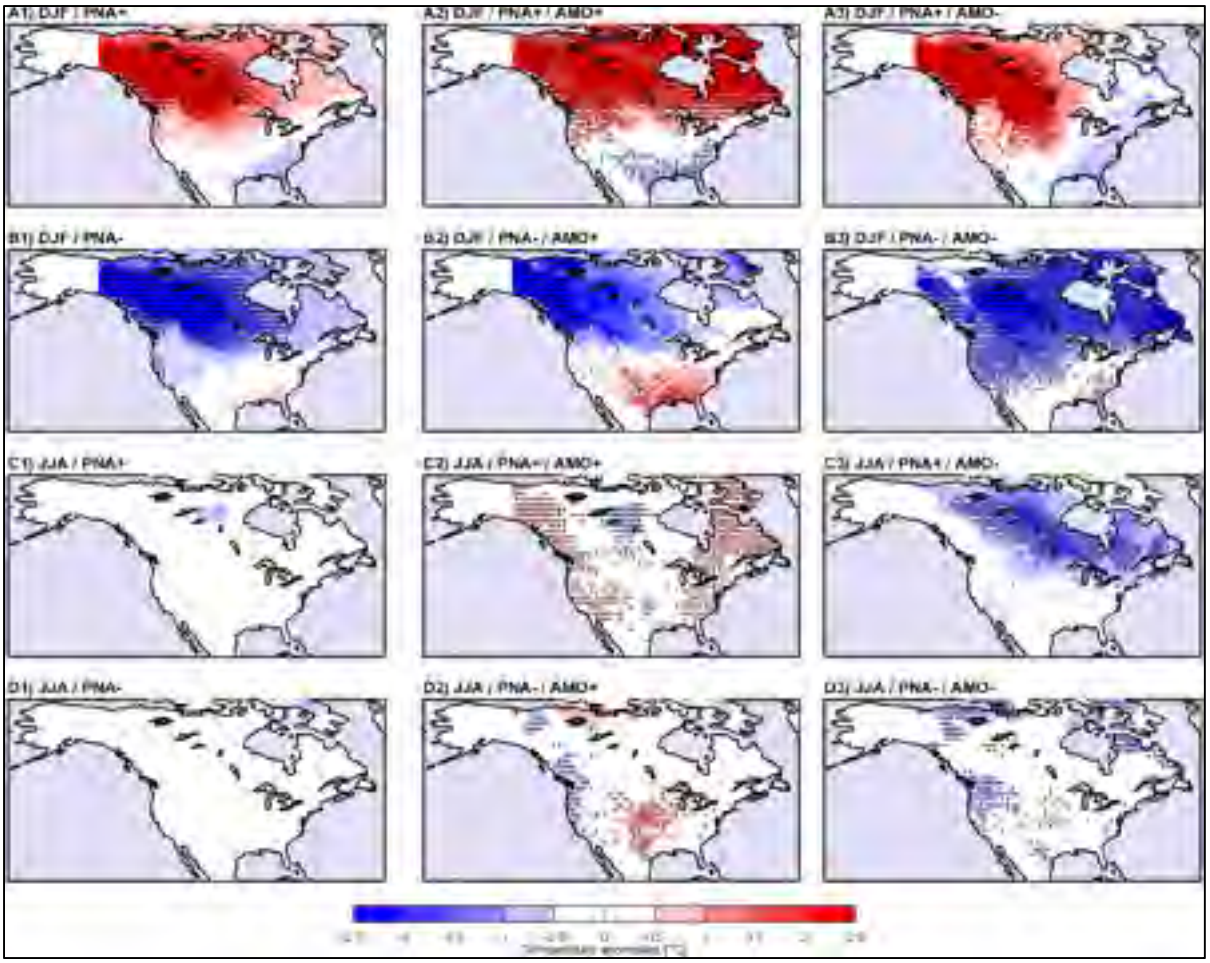

Figure-A V-4 Same as in Figure-A V-1, but for AMO and PNA

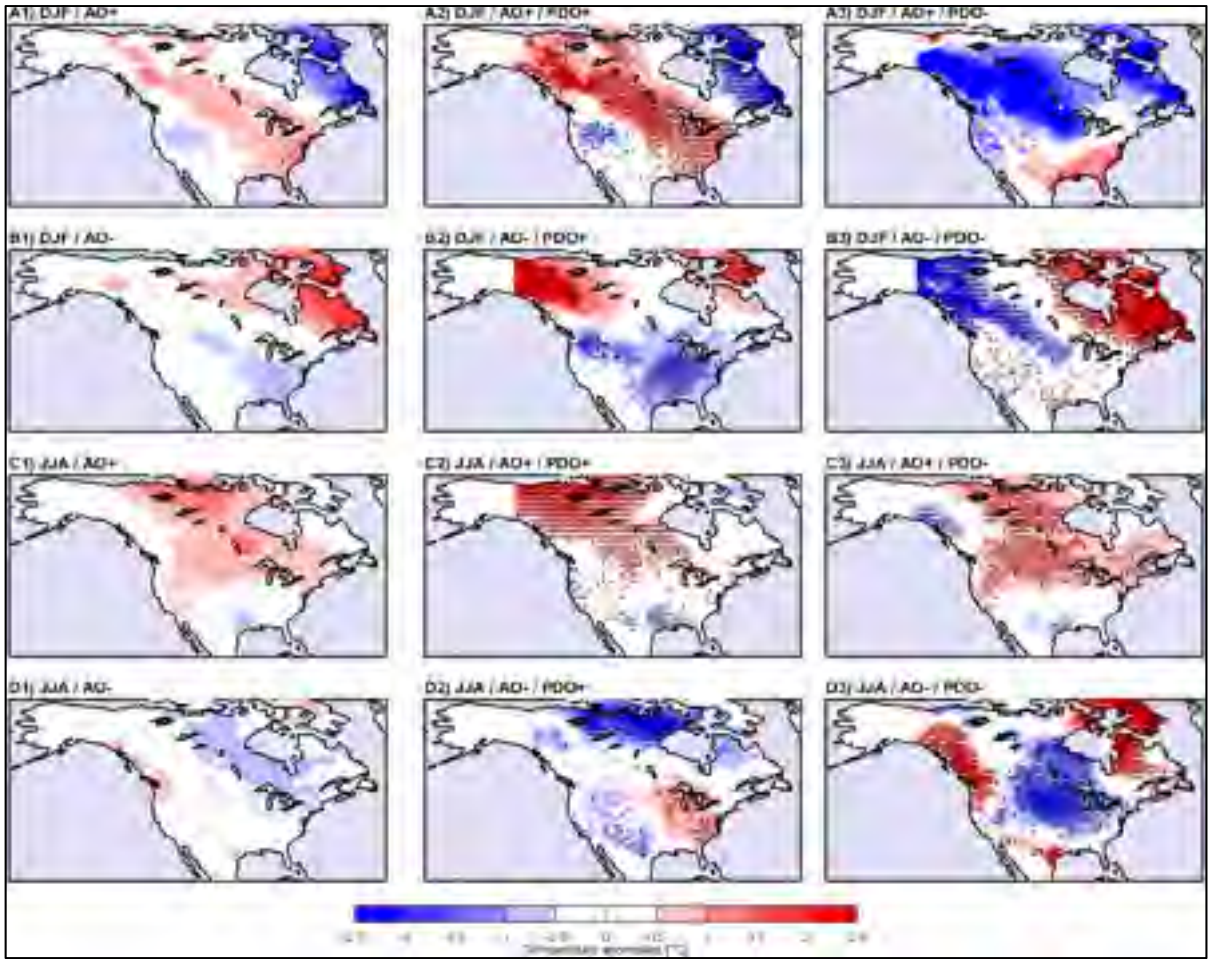

Figure-A V-5 Same as in Figure-A V-1, but for PDO and AO

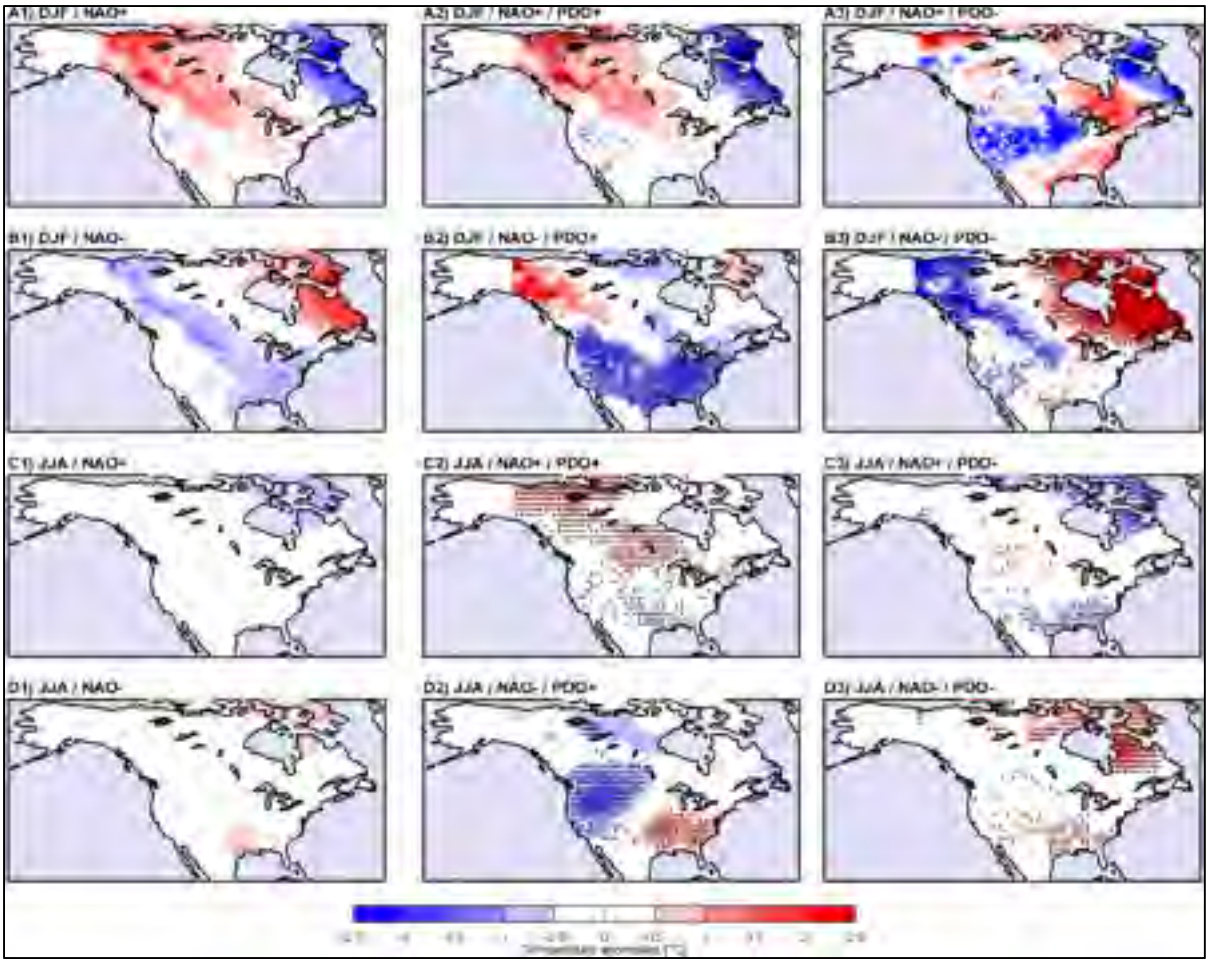

Figure-A V-6 Same as in Figure-A V-1, but for PDO and NAO

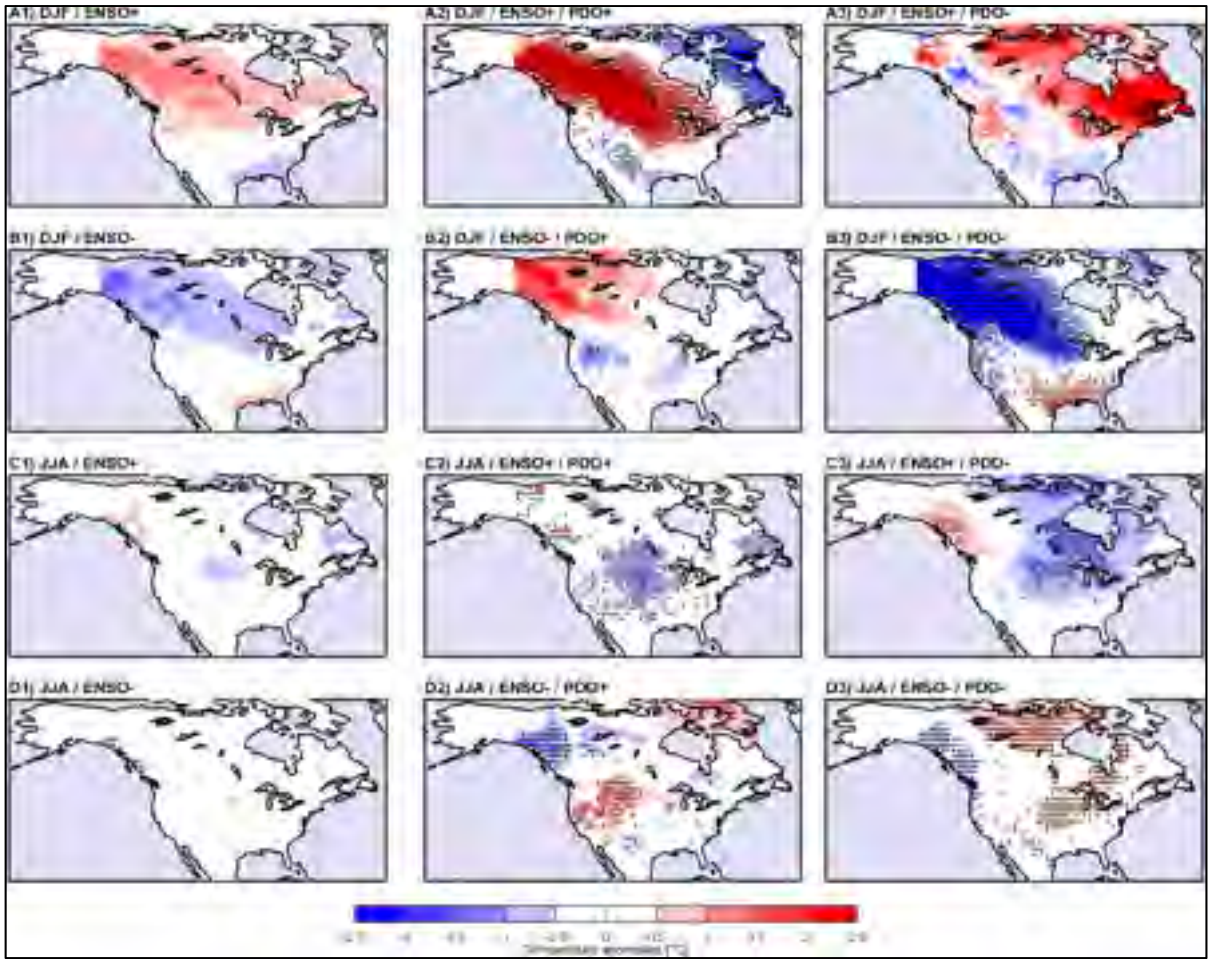

Figure-A V-7 Same as in Figure-A V-1, but for PDO and ENSO

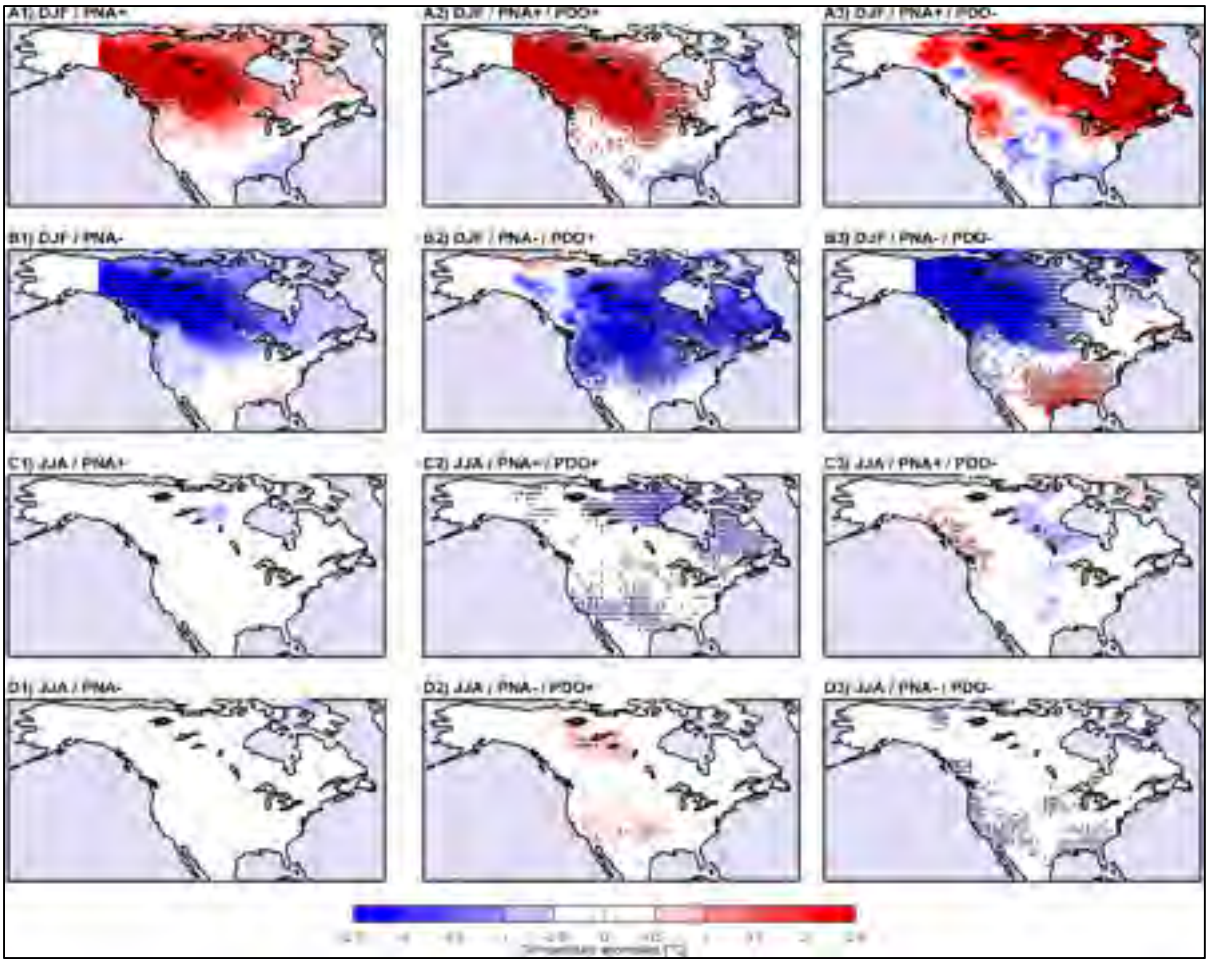

Figure-A V-8 Same as in Figure-A V-1, but for PDO and PNA

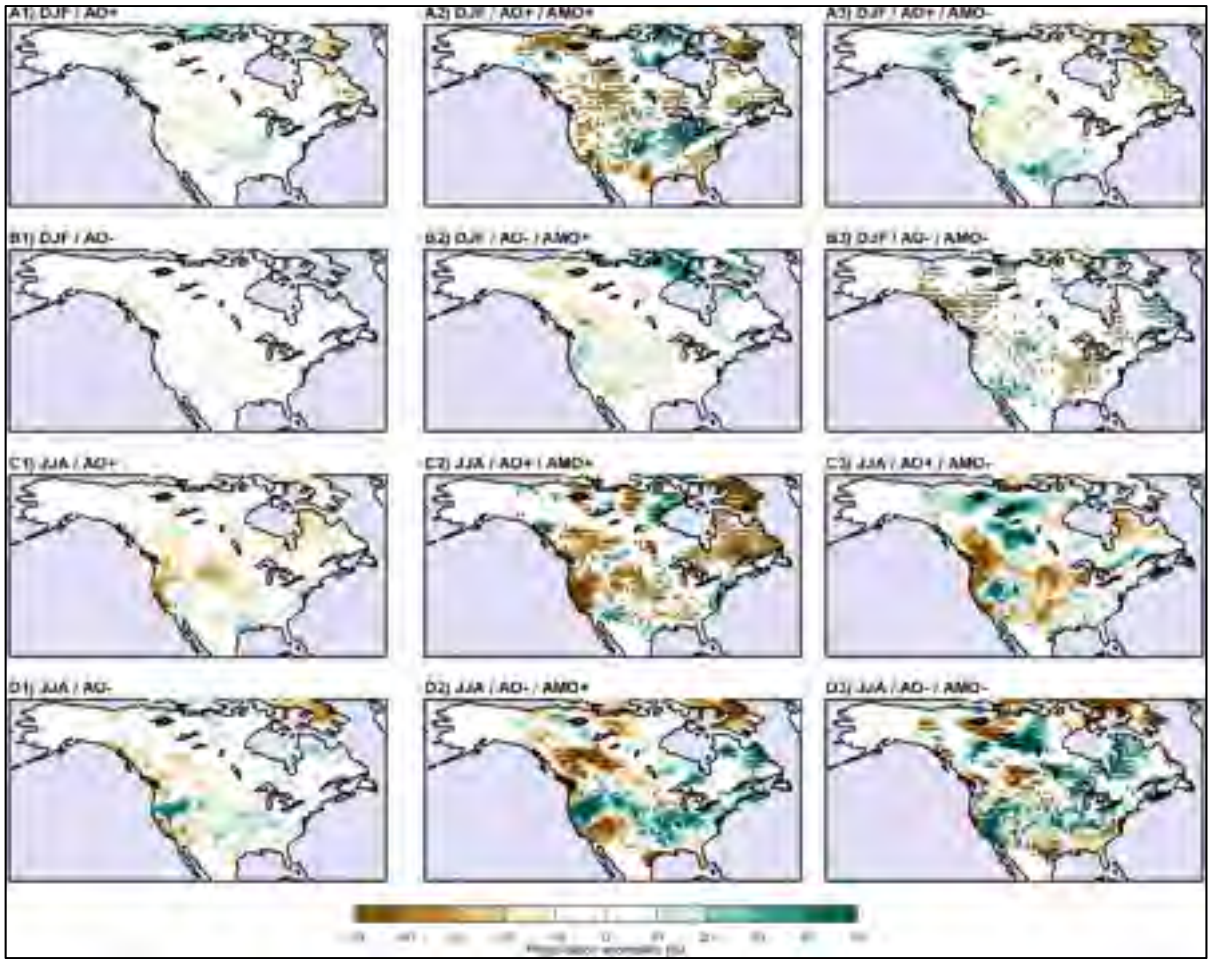

Figure-A V-9 Influence form AO and its coupled effect with AMO on North American total seasonal precipitation. Precipitation anomalies (%) are defined as the relative difference between the mean of the selected years based on the cold and/or warm phases of the climate indices and the one of all seasonal values. The black dots represent the statistical significance based on a bootstrap resampling approach with replacement using a total of 1 000 samples with a 95% confidence level

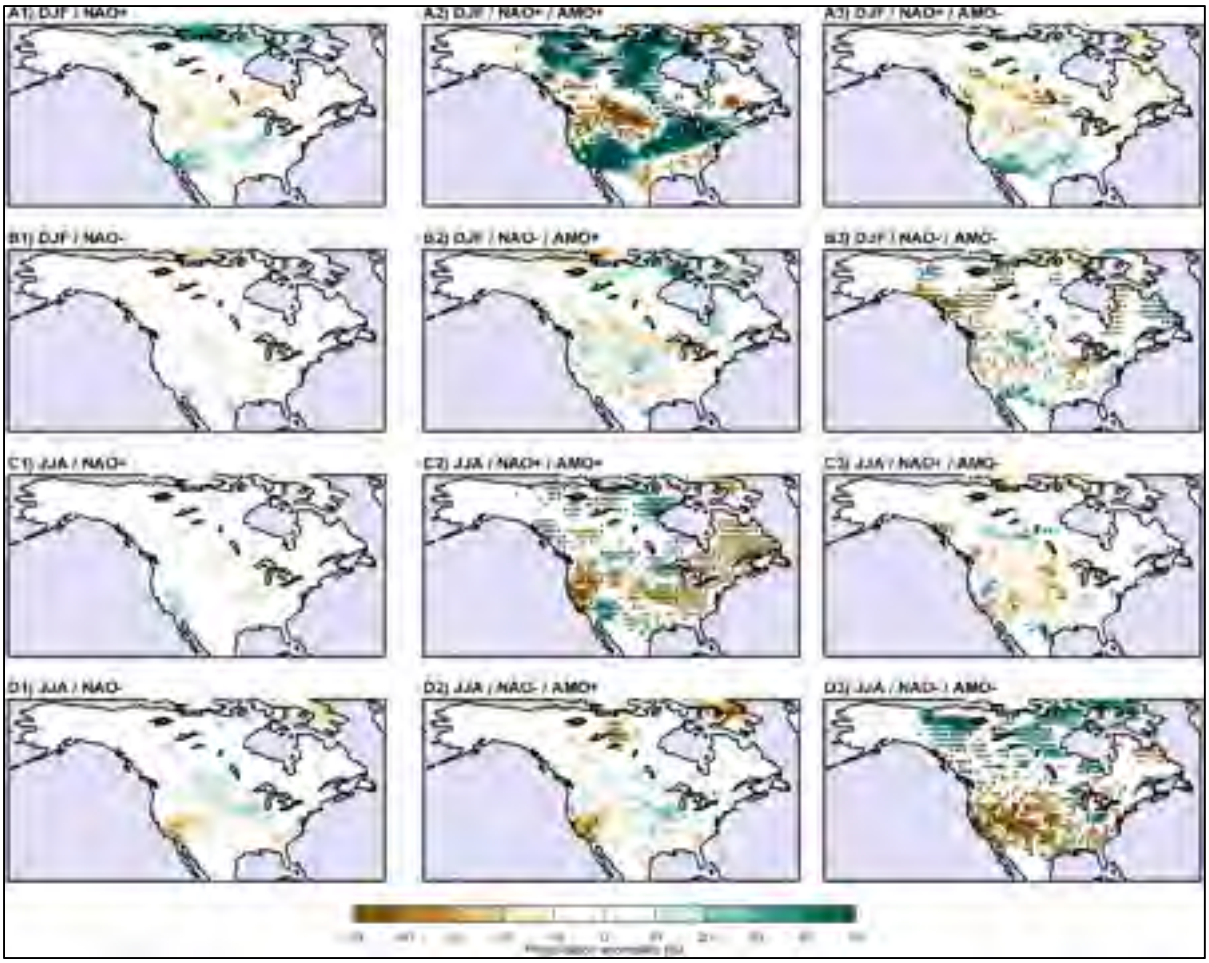

Figure-A V-10 Same as in Figure-A V-9, but for AMO and NAO

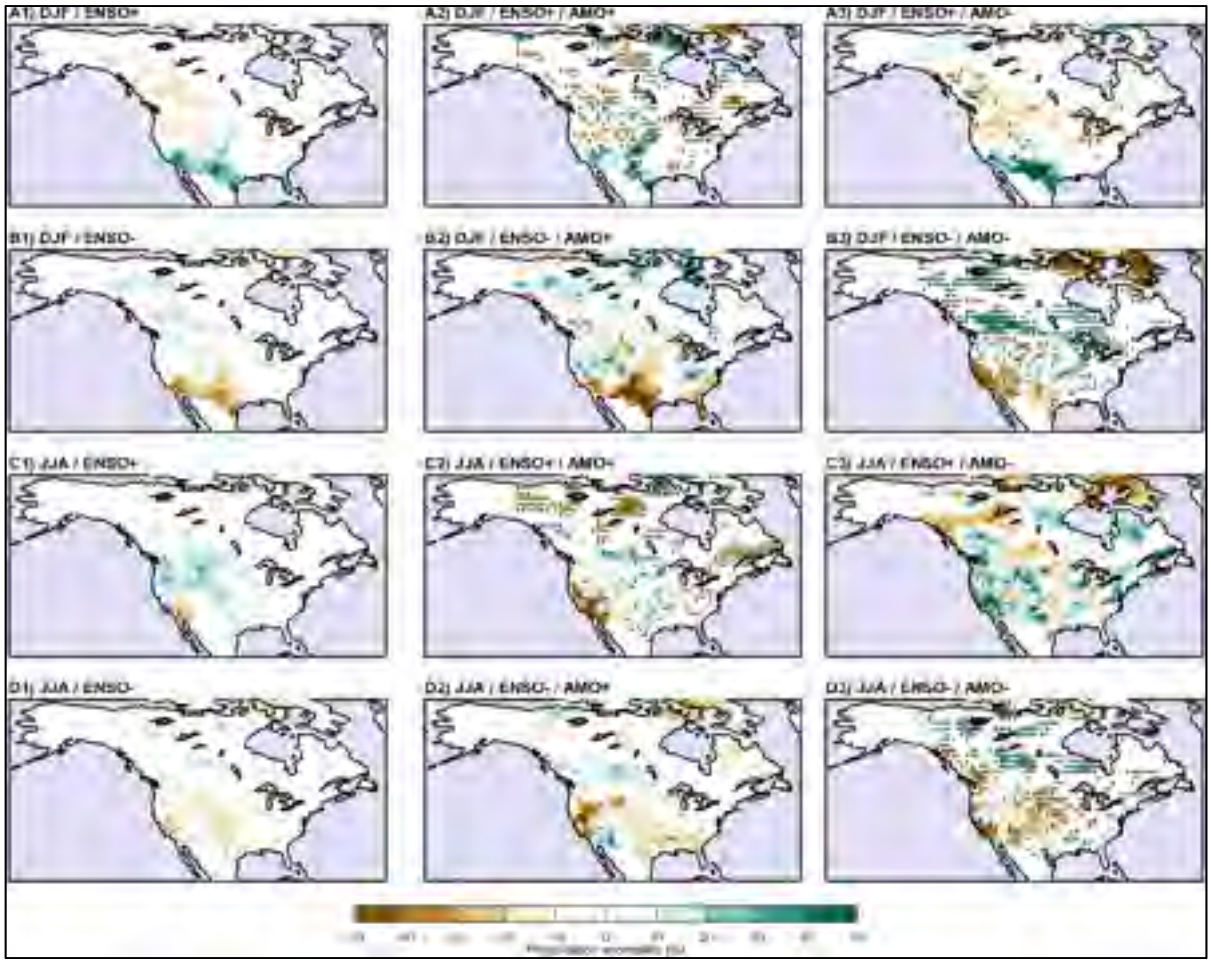

Figure-A V-11 Same as in Figure-A V-9, but for AMO and ENSO

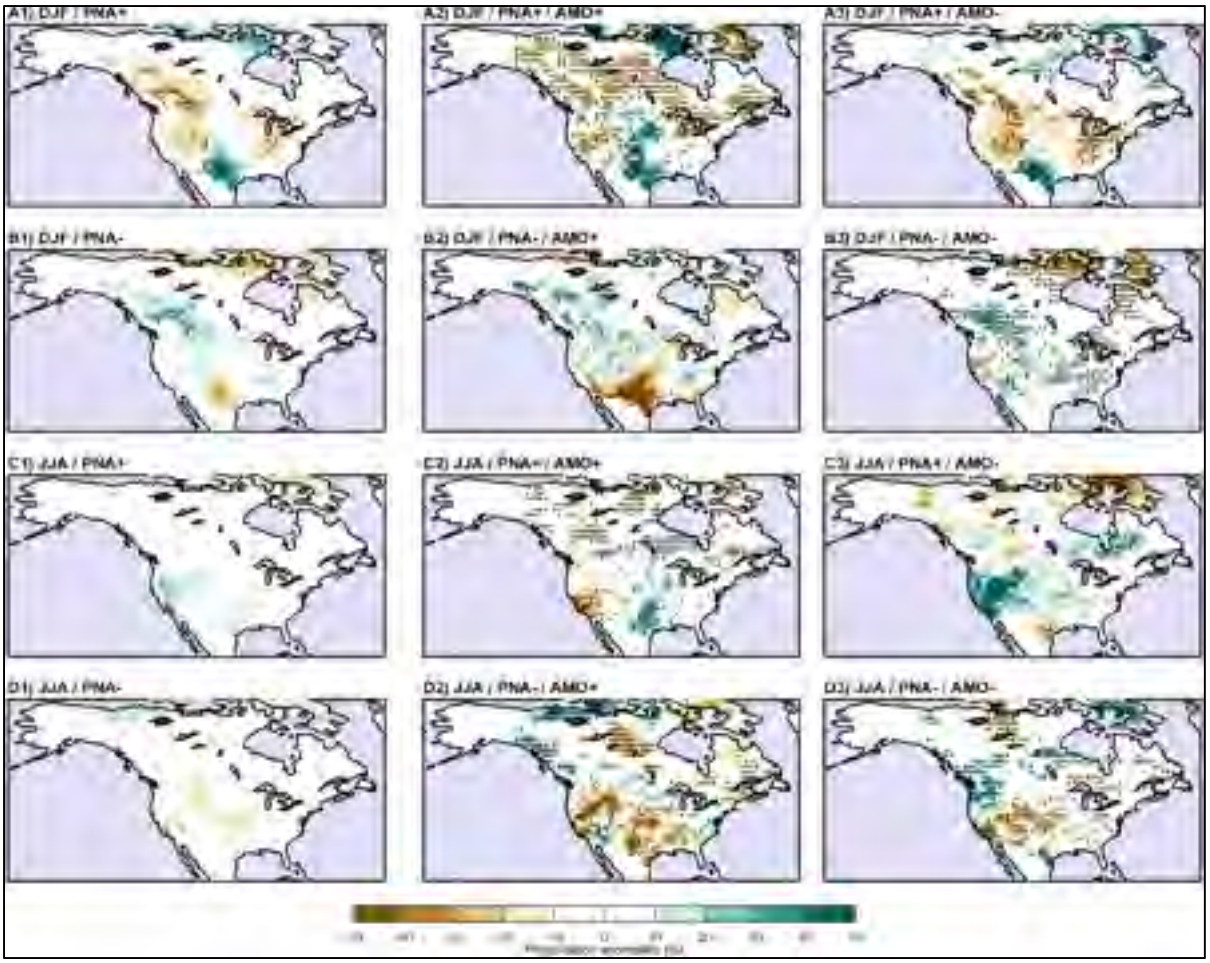

Figure-A V-12 Same as in Figure-A V-9, but for AMO and PNA

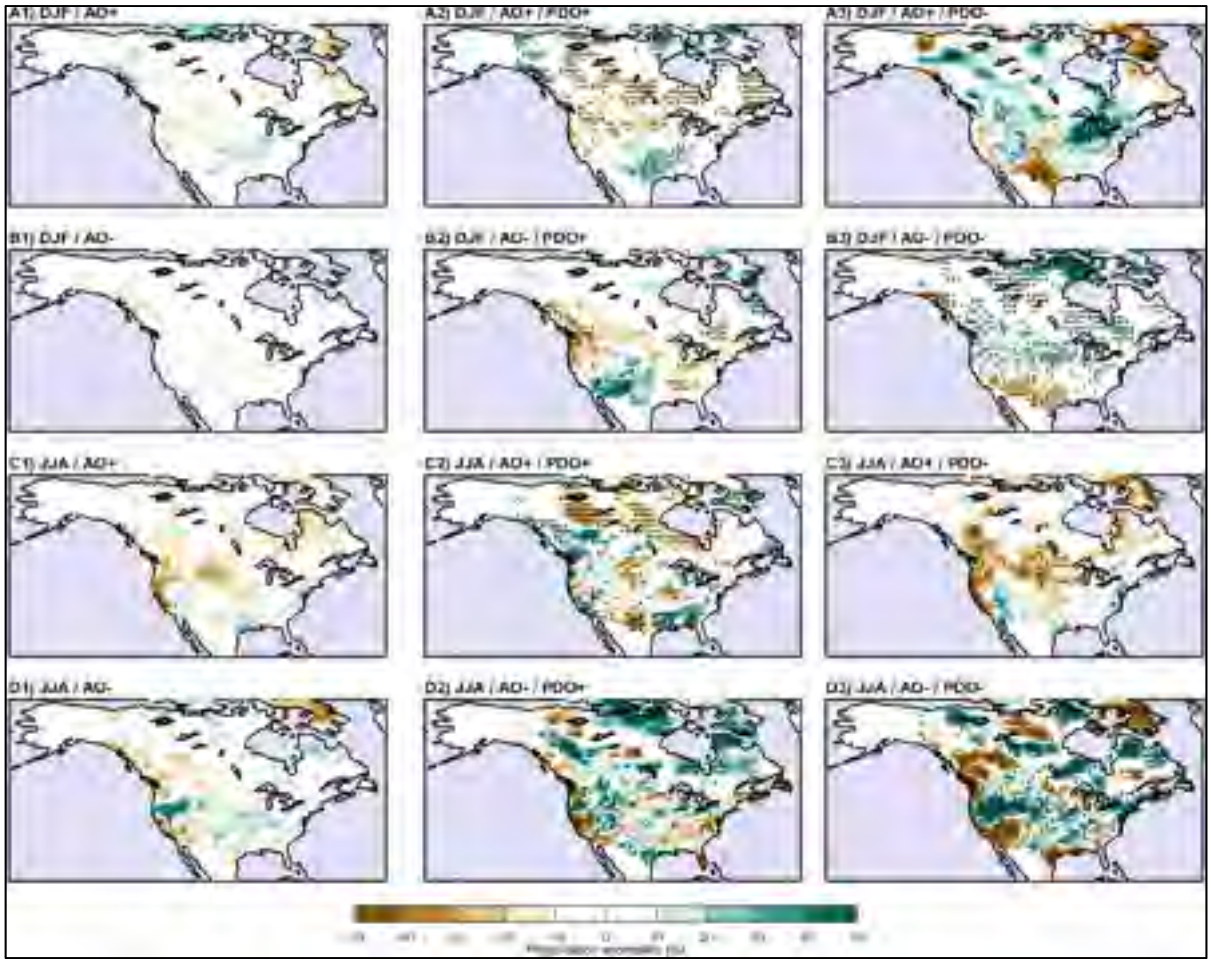

Figure-A V-13 Same as in Figure-A V-9, but for PDO and AO

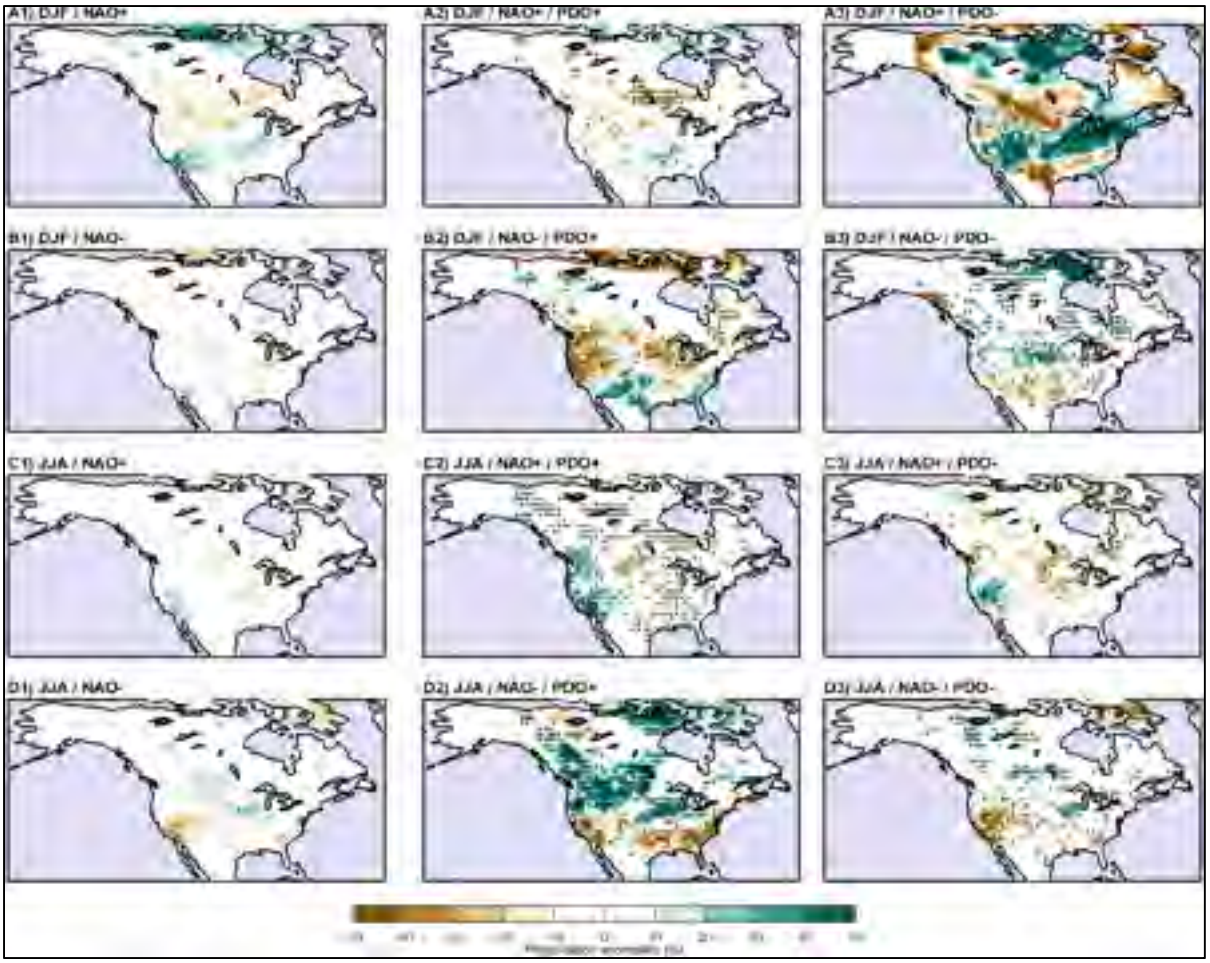

Figure-A V-14 Same as in Figure-A V-9, but for PDO and NAO

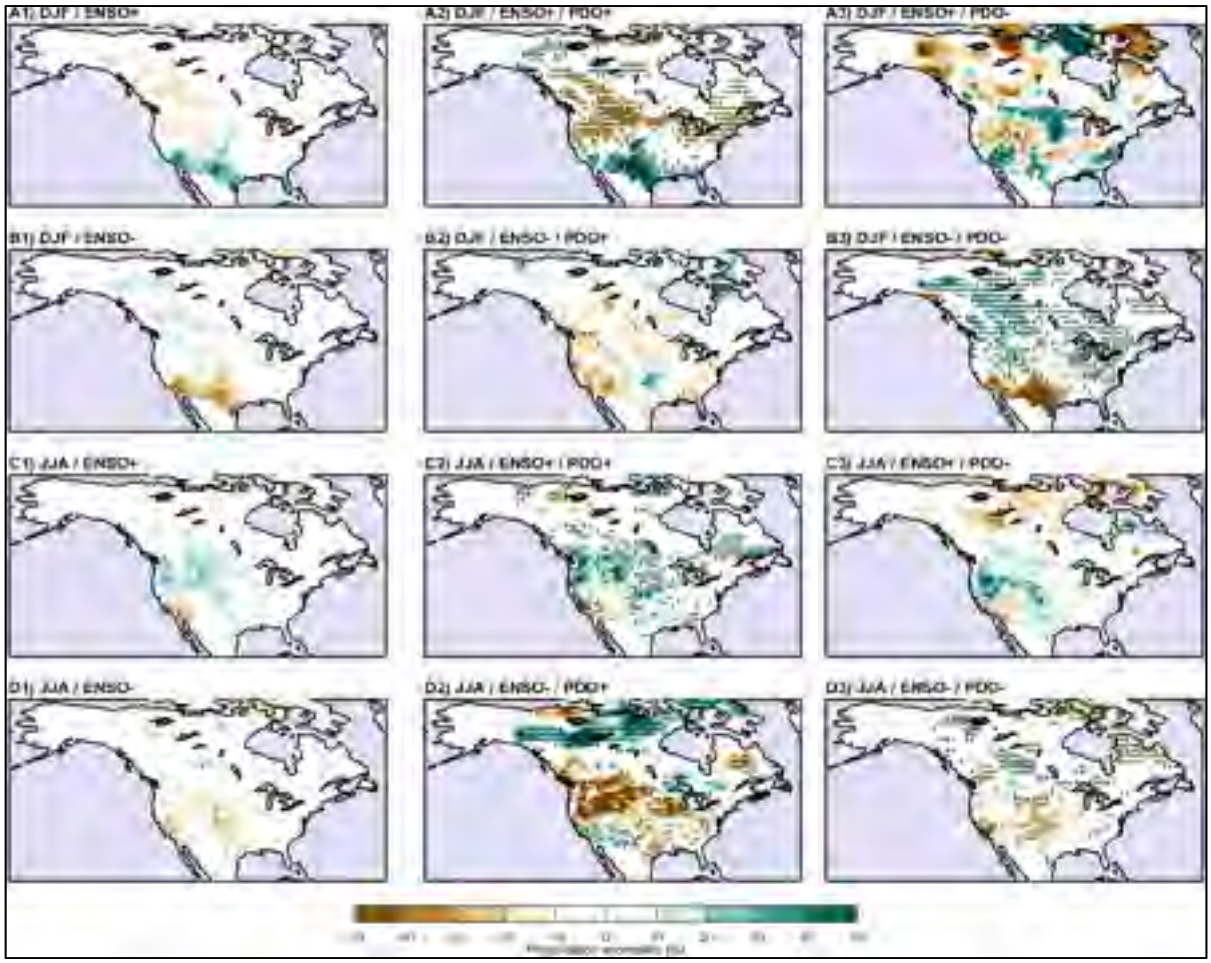

Figure-A V-15 Same as in Figure-A V-9, but for PDO and ENSO

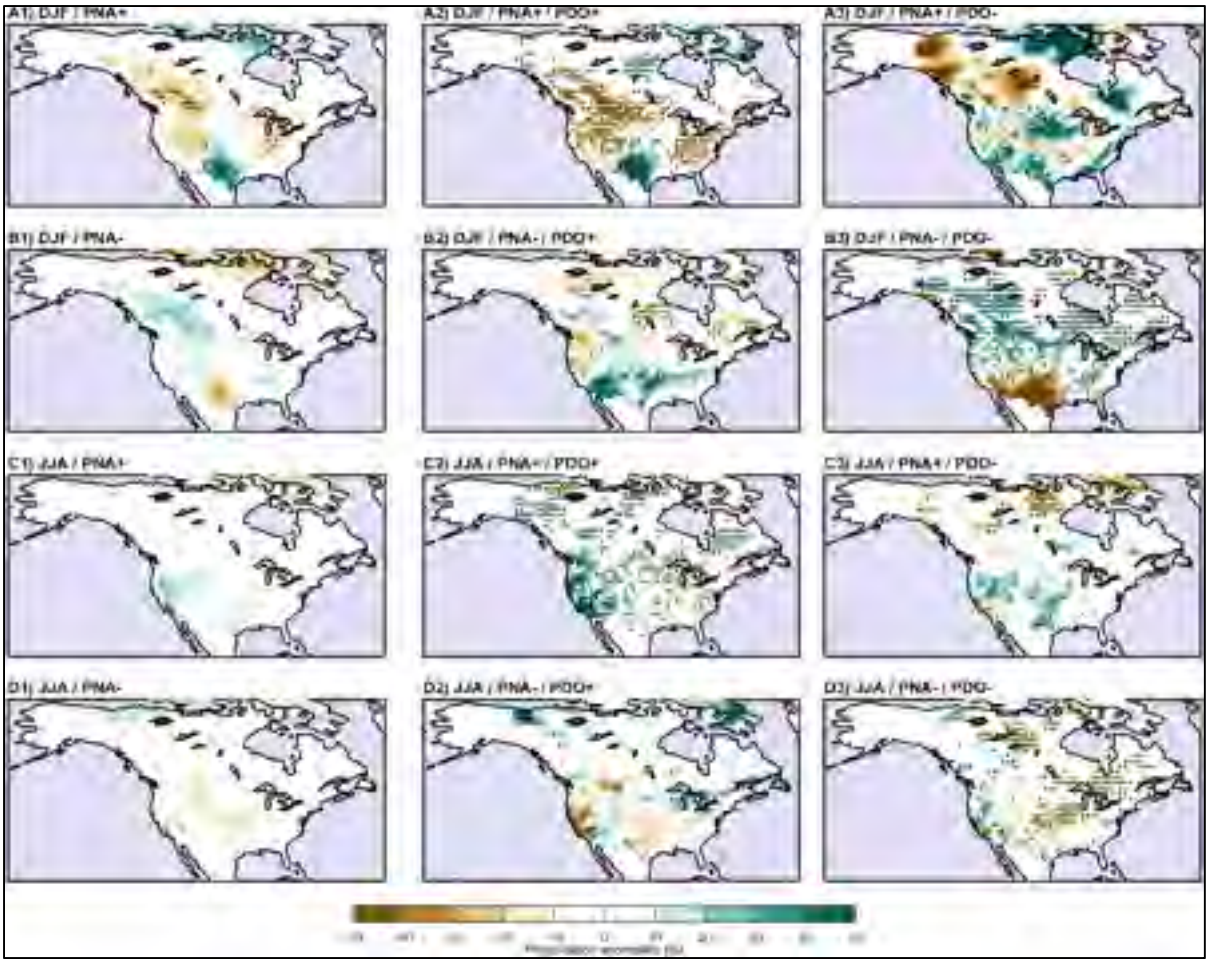

Figure-A V-16 Same as in Figure-A V-9, but for PDO and PNA

ANNEXE VI

UNCERTAINTY OF HYDROLOGICAL MODEL COMPONENTS OVER NORTH AMERICAN WATERSHEDS

Jean-Luc Martel[a], Magali Troin[b], François Brissette[c], Richard Arsenault[d]

[a, b, c, d] Département de Génie de la construction, École de technologie supérieure

Article en préparation pour soumission à la revue « Journal of Hydrometeorology ».

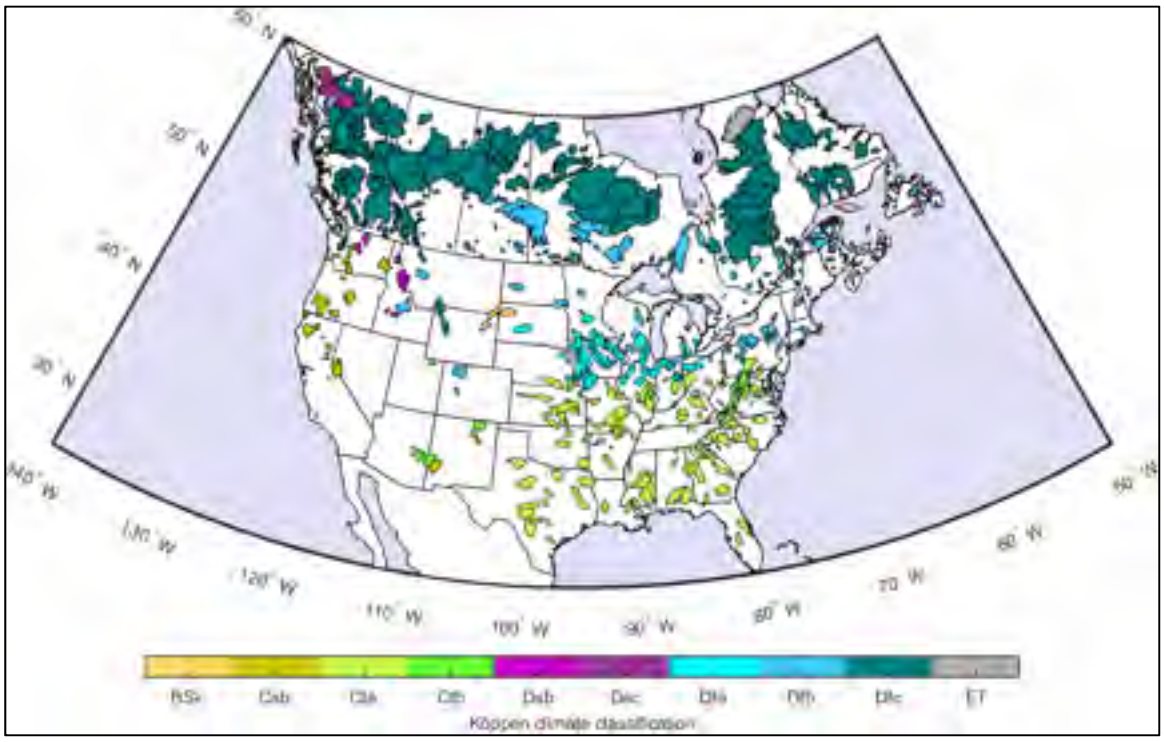

Figure-A VI-1 The 884 selected watersheds over North America used in this study based on the CANOPEX and MOPEX databases

Table-A VI-1 Köppen climate zones used in this study. The first column represent the main groups and the second and third columns represent the types and subtypes of climates

Accronym	1st letter	2nd letter	3rd letter	Number of watersheds
BSk	Arid	Steppe	Cold	7
Csb	Temperate	Dry Summer	Warm Summer	15
Cfa	Temperate	Without dry season	Hot Summer	101
Cfb	Temperate	Without dry season	Warm Summer	29
Dsb	Cold (Continental)	Dry Summer	Warm Summer	9
Dsc	Cold (Continental)	Dry Summer	Cold Summer	4
Dfa	Cold (Continental)	Without dry season	Hot Summer	49
Dfb	Cold (Continental)	Without dry season	Warm Summer	179
Dfc	Cold (Continental)	Without dry season	Cold Summer	293
ET	Polar	Tundra		12

Table-A VI-2 Summary of all snow, PET modules, vertical and horizontal flow modules used in his work

Snowmelt	Parameters	Model type	PET methods	Model type	Vertical flow	Parameters	Horizontal flow	Parameters
CEMANEIGE	2	Degree-day	Hamon	Temperature-based	GR4J	1	GR4J	3
HBV	1	Degree-day	Hydro-Québec	Temperature-based	HMETS	7	HMETS	4
HMETS	10	Degree-day	MOHYSE	Temperature-based	MOHYSE	6	MOHYSE	2
MOHYSE	2	Degree-day	Oudin	Radiation-based				
HYDROTEL	5	Degree-day/Energy balance						
Number of combinations		x5		x4		x3		x3 = 180

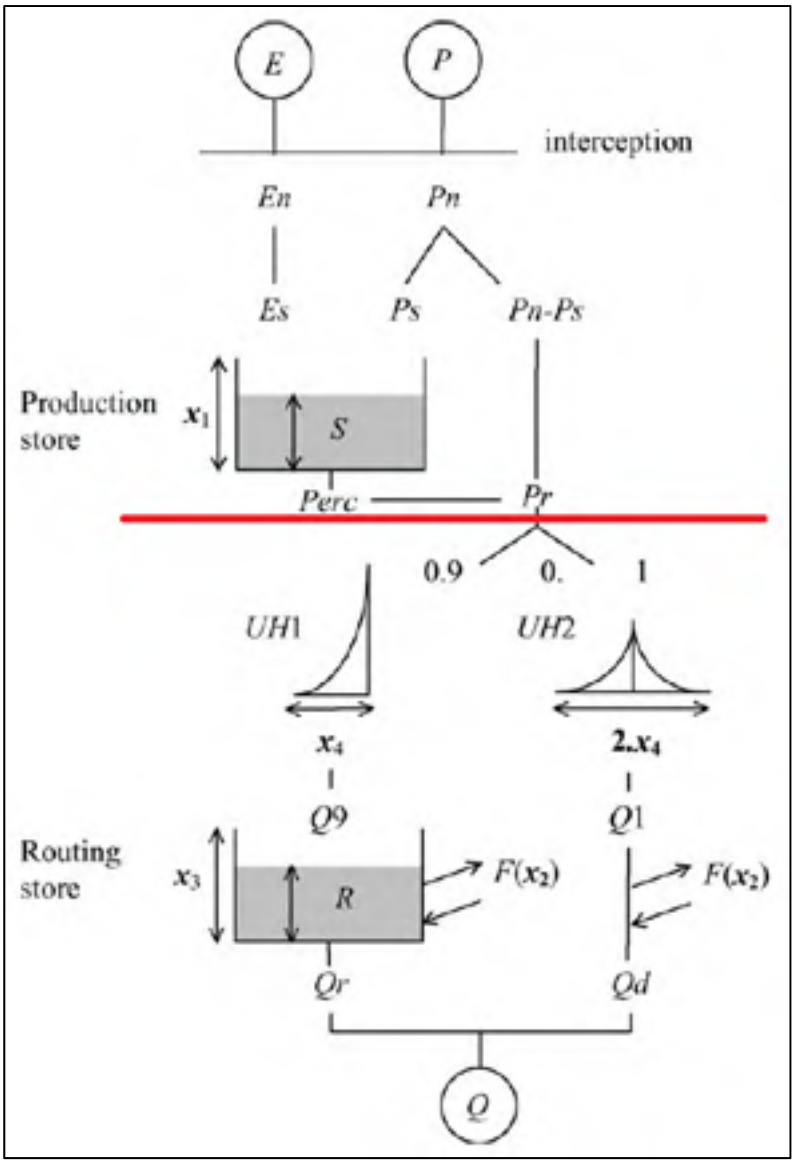

Figure-A VI-2 GR4J vertical and horizontal flows modules
(Figure scheme taken from Perrin et al. (2003))

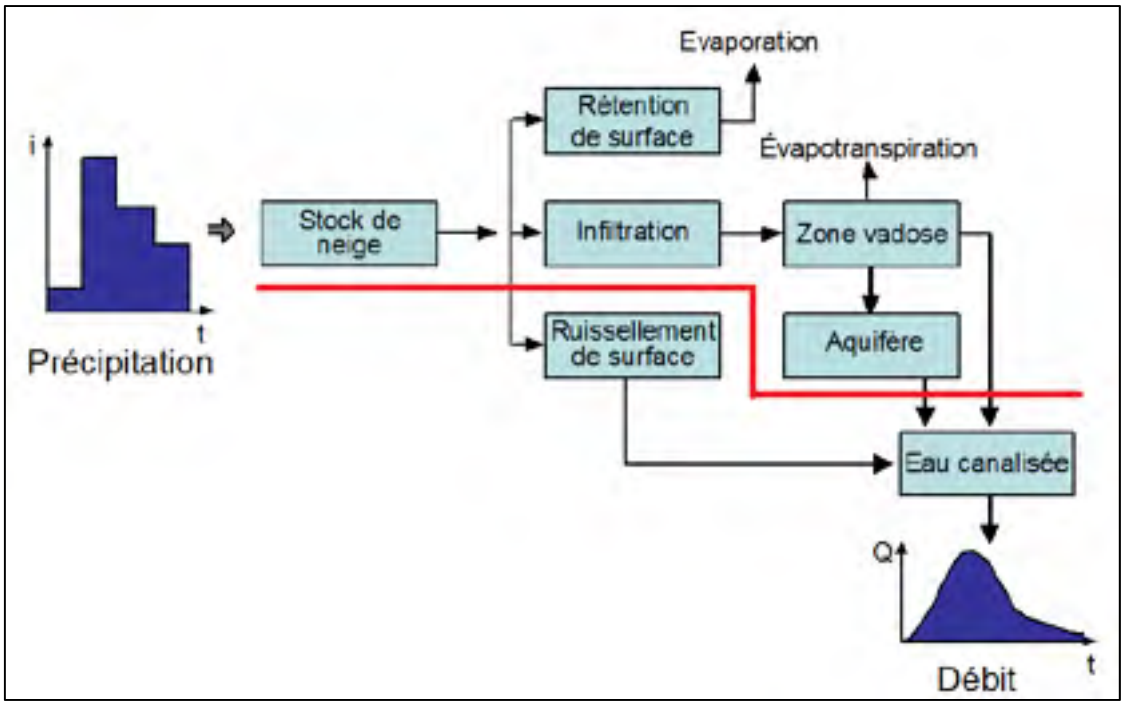

Figure-A VI-3 MOHYSE vertical and horizontal flows modules
(Figure scheme taken from Fortin et Turcotte (2007))

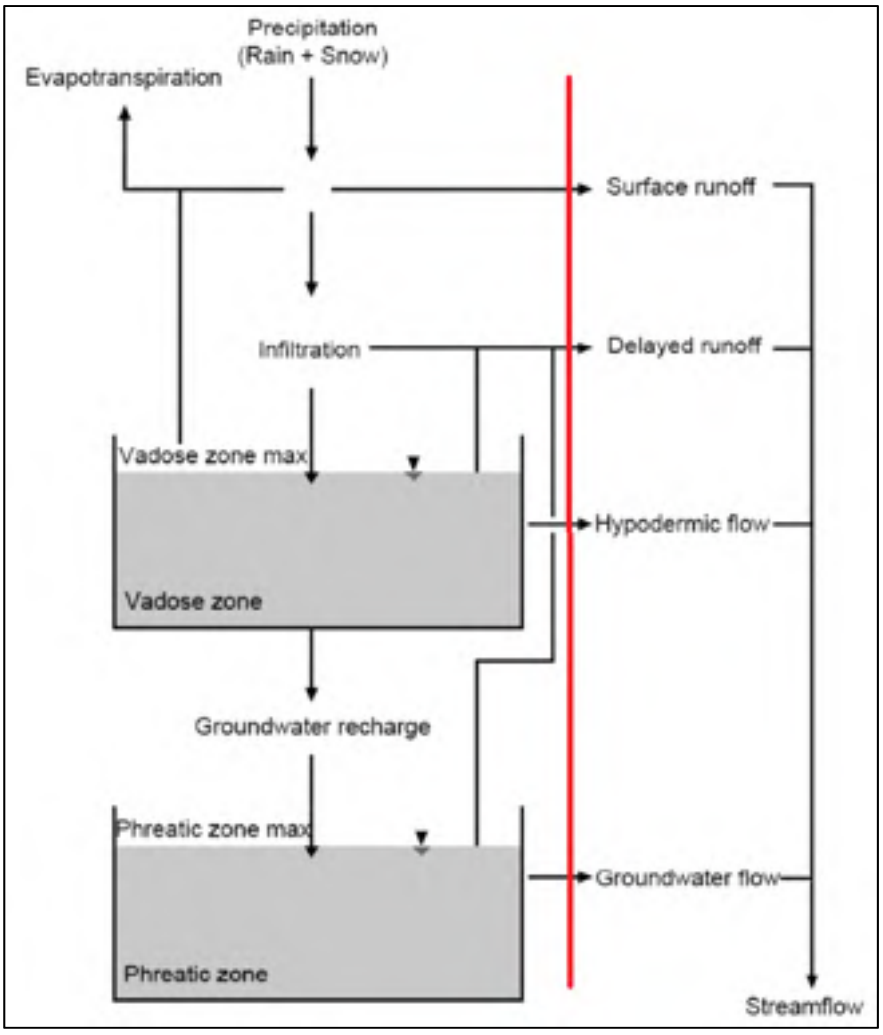

Figure-A VI-4 HMETS vertical and horizontal flows modules
(Figure scheme from Martel et al. (2017))

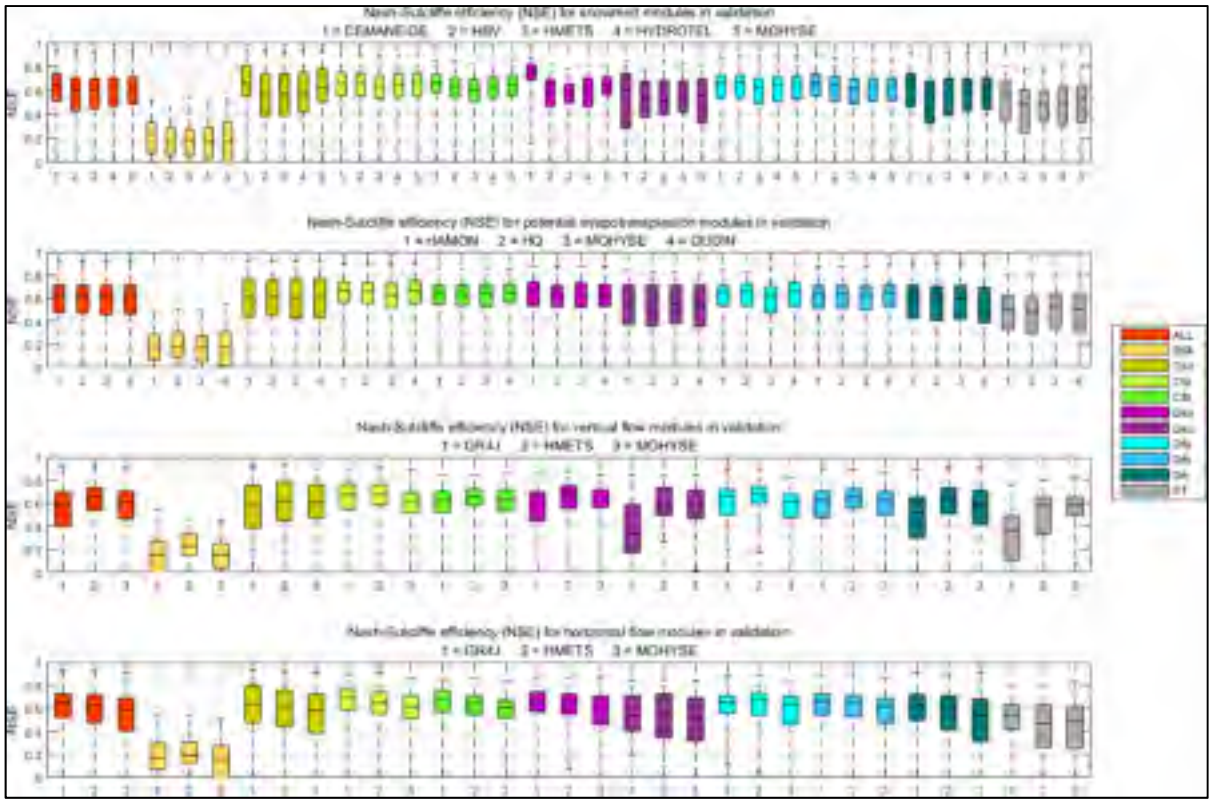

Figure-A VI-5 Boxplots of the NSE scores for the validation period on all watersheds. The four components (snowmelt module, PET module, vertical flow module and horizontal flow module) of the hydrological model are shown on each row respectively. Each boxplot represents all the watersheds of a given Köppen climate zone (defined by the color of the boxplot) for all model combination using a selected module. For instance, the snowmelt modules are shown on the first row and model number 1 corresponds to CEMANEIGE. In the first boxplot with a 1 on the abscissa show the NSE score for all watersheds in the BSk zone using all model combinations where CEMANEIGE is used

Table-A VI-3 Percentage of variance explained by the 15 terms in the decomposition over the climate zones for the Nash-Sutcliffe Efficiency (NSE)

	S	E	V	H	SE	SV	SH	EV	EH	VH	SEV	SEH	SVH	EVH	SEVH
ALL	10	1	23	25	0	8	3	3	1	20	1	0	4	1	0
BSk	0	5	8	7	2	1	2	12	15	12	3	3	2	22	6
Csb	19	1	8	14	2	10	4	4	8	9	1	2	4	10	4
Cfa	2	5	22	23	0	0	0	20	8	9	0	0	0	9	0
Cfb	15	2	9	33	0	2	2	7	3	21	1	1	0	4	1
Dsb	28	1	10	10	1	8	8	1	2	11	2	1	11	3	2
Dsc	3	1	66	3	1	6	4	3	0	7	1	0	2	0	1
Dfa	5	8	17	10	0	2	1	27	5	16	0	0	0	10	0
Dfb	14	3	11	28	0	9	2	7	1	22	0	0	2	2	0
Dfc	12	1	30	17	0	11	5	1	0	14	1	0	5	0	1
ET	7	1	55	4	1	9	8	3	0	7	1	0	4	0	0

Table-A VI-4 Percentage of the total variance explained by the snow models (V(S)), the PET modules (V(E)), the vertical flow (V(V)) and the horizontal flow (V(H)) in the decomposition over the climate zones for the Nash-Sutcliffe Efficiency (NSE)

	S	E	V	H
ALL	26	7	60	55
BSk	20	67	67	69
Csb	46	31	50	55
Cfa	3	43	61	50
Cfb	22	18	45	64
Dsb	62	12	49	49
Dsc	19	8	88	19
Dfa	8	50	71	42
Dfb	27	13	52	56
Dfc	36	5	64	43
ET	30	7	79	24

ANNEXE VII

IMPACTS OF FORCED AND UNFORCED CLIMATE VARIABILITY ON EXTREME FLOODS USING A LARGE CLIMATE ENSEMBLE

Jean-Luc Martel[a], François Brissette[b], Jie Chen[c], Alain Mailhot[d], Daniel Caya[e]

[a, b, c, e] Département de Génie de la construction, École de technologie supérieure
[d] Institut national de la recherche scientifique – Eau, Terre et Environnement

Présentation orale à l'European Geosciences Union General Assembly 2016, Vienne, Autriche, 17 – 22 avril 2016.

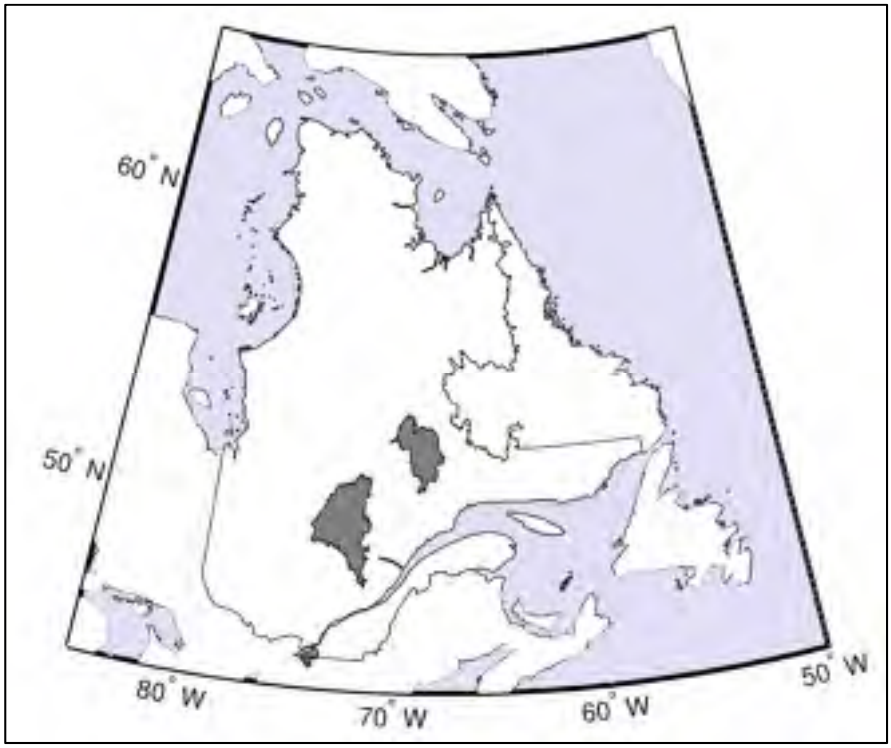

Figure-A VII-1 Outlines of Manic-5, Lac Saint-Jean and Châteauguay watersheds

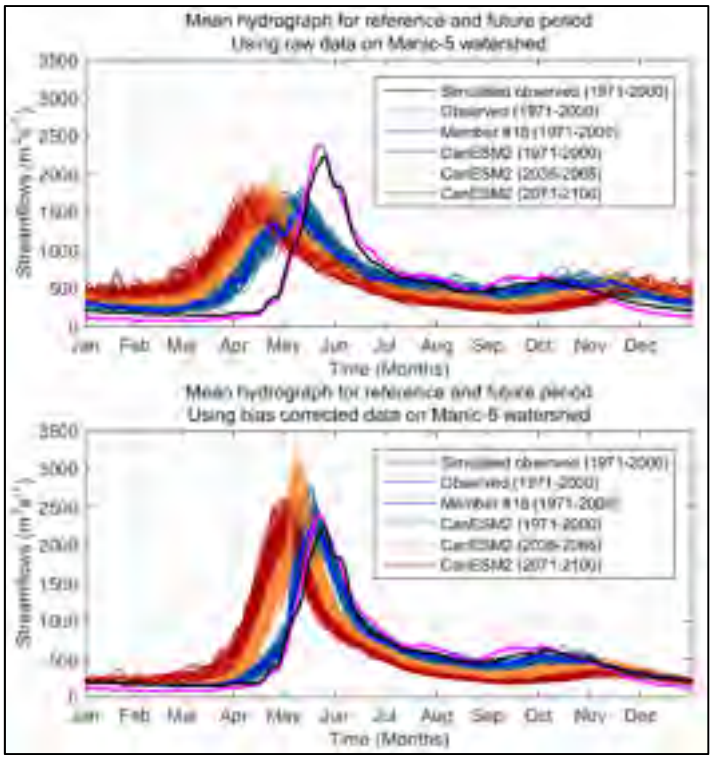

Figure-A VII-2 Mean annual hydrographs of the Manic-5 watershed for the simulated observed (black), observed (pink), median member used in the bias correction (blue bold line), and all 50 members for the 1971-2000 period (blue), 2036-2065 period (orange) and 2071-2100 period (red)

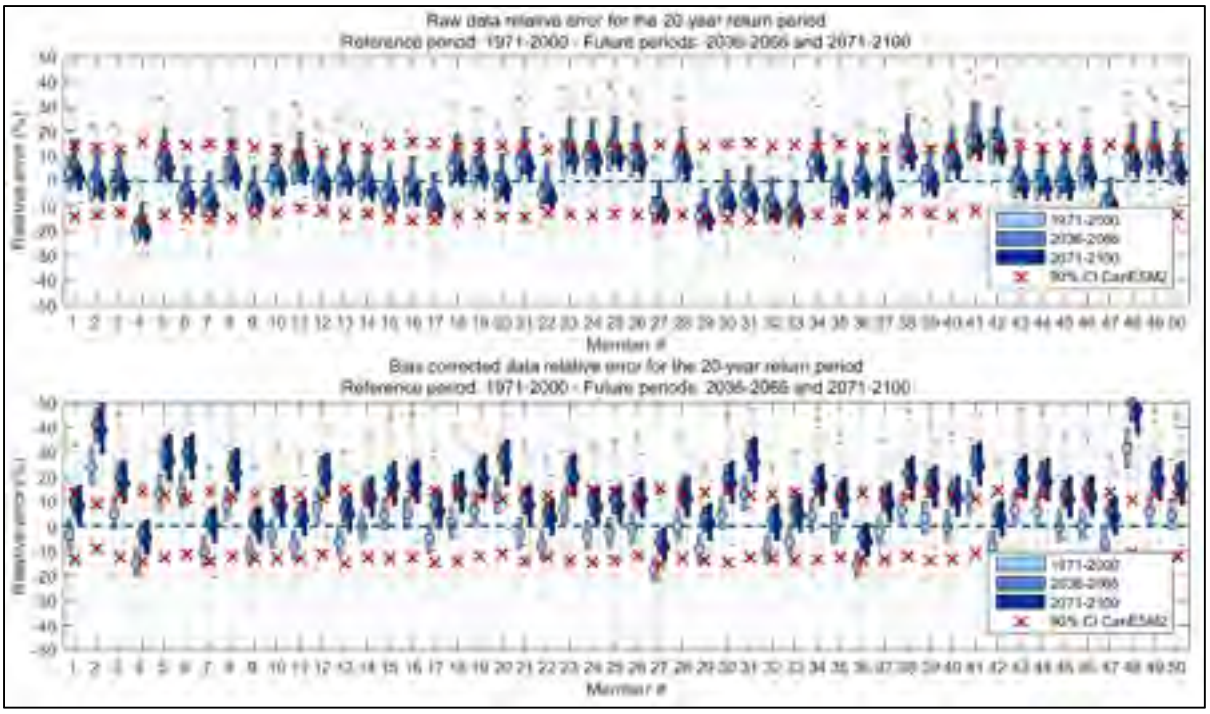

Figure-A VII-3 Box plots of the Manic-5 watershed relative error of the selected member's 20-year return period obtained with a Gumbel distribution against the other members for the three periods. The red "x" show the 90% confidence interval of the selected member. Results for the raw data are shown on the top panel and for the bias corrected data on the bottom panel. See Chapter 6 methodology for the bias correction methodology

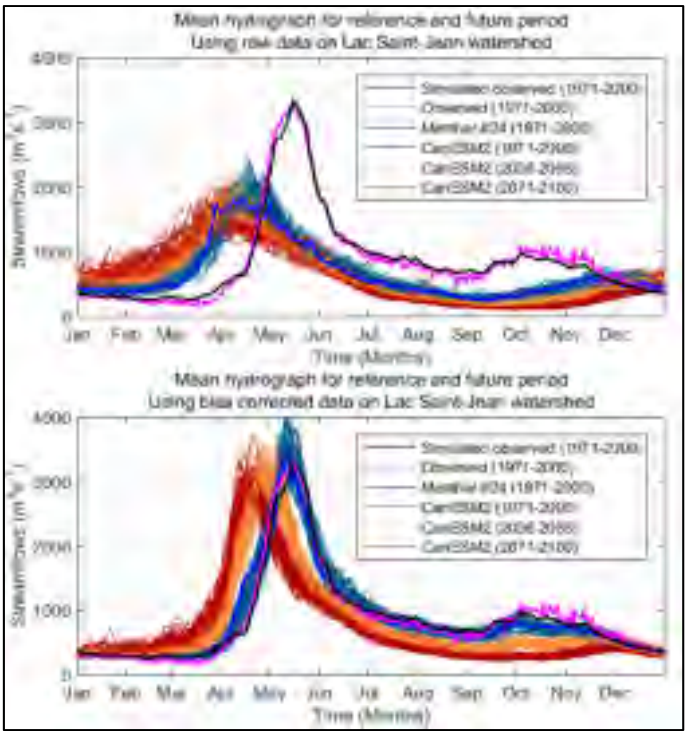

Figure-A VII-4 Same as Figure-A VII-3, but for the Lac Saint-Jean watershed

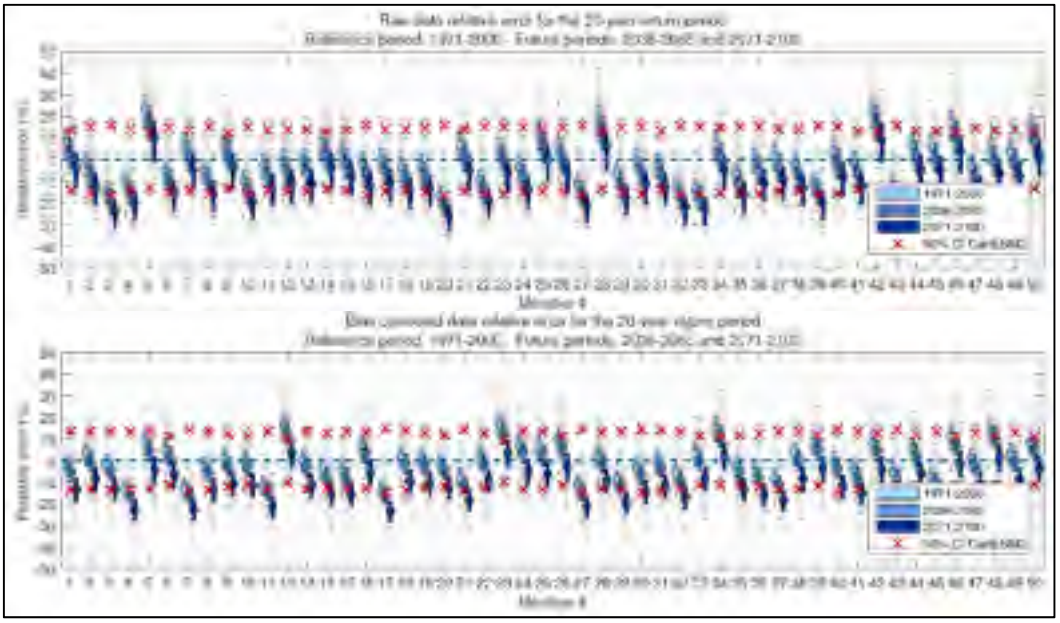

Figure-A VII-5 Same as Figure-A VII-4, but for the Lac Saint-Jean watershed

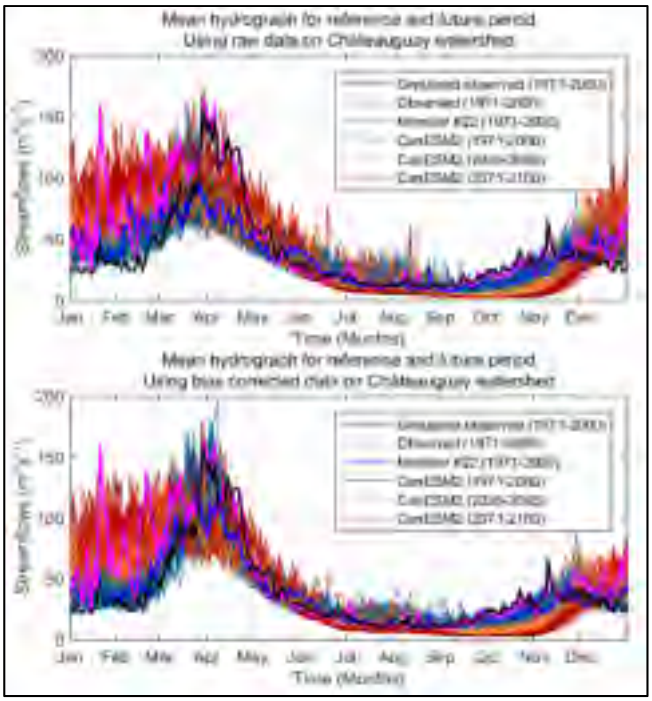

Figure-A VII-6 Same as Figure-A VII-3, but for the Châteauguay watershed

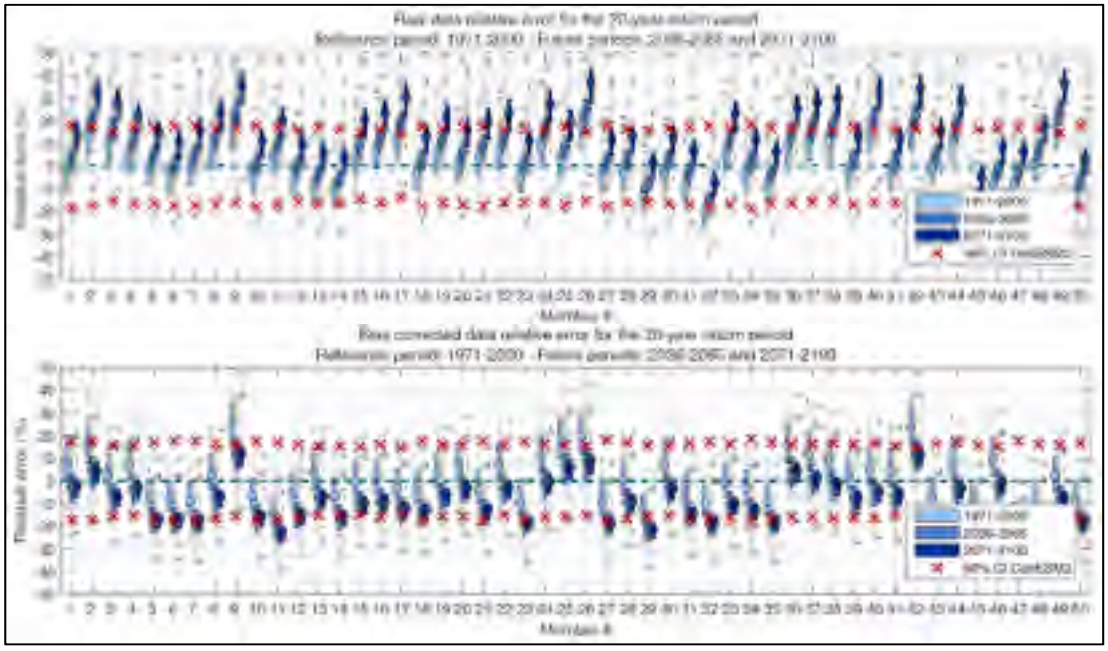

Figure-A VII-7 Same as Figure-A VII-4, but for the Châteauguay watershed

ANNEXE VIII

LISTE DES CONTRIBUTIONS SCIENTIFIQUES

Cette annexe fournie la liste complète des publications en lien avec les travaux de cette thèse:

VIII.1 Articles publiés dans des journaux avec comités de lecture :

- Leduc, M., Mailhot, A., Frigon, A., **Martel, J.-L.**, Ludwig, R., Brietzke, G. B., Giguère, M., Brissette, F., Turcotte, R., Braun, M. & Scinocca, J. (2019). 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*, 667-693. doi : 10.1175/JAMC-D-18-0021.1
- **Martel, J.-L.**, Mailhot, A., Brissette, F., & Caya, D. (2018). Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales. *Journal of Climate*, 31(11), 4241-4263. doi :10.1175/JCLI-D-17-0282.1.
- Arsenault, R., Brissette, F. & **Martel, J.-L.** (2018). The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology*, 566, 346-362. doi : 10.1016/j.jhydrol.2018.09.027
- Troin, M., Arsenault, R., **Martel, J.-L.** & Brissette, F. (2018). Uncertainty of hydrologic model components in climate change studies over nordic Quebec catchments. *Journal of Hydrometeorology*, 19(1), 27-46. doi : 10.1175/JHM-D-17-0002.1.

VIII.2 Articles en préparation ou soumis dans des journaux avec comité de lecture :

- **Martel, J.-L.**, Mailhot, A. & Brissette, F. (2018). Global and regional projected changes in 100-year sub-daily, daily and multi-day precipitation extremes estimated from three large ensembles of climate simulations. Accepté pour publication avec révisions majeures à *Journal of Climate*, 56p.

- **Martel, J.-L.**, Brissette, F. & Mailhot, A. (2018). Evolution of daily precipitation and streamflow extremes in two large ensembles of climate simulations over North American catchments. En préparation pour soumission à *Nature Climate Change*, 34p.
- **Martel, J.-L.**, Troin, M. & Brissette, F. (2018). Impacts of the dominant modes of natural climate variability on hydroclimatic variability over North American watersheds. Soumis à *Water Resources Research*, 58p.
- Chen, J., Brissette, F., **Martel, J.-L.**, Zhang, X. J. & Frei, A. (2018). Relative importance of internal climate variability versus anthropogenic climate change in global climate change. Accepté pour publication avec révisions majeures à *Journal of Climate*, 40p.

VIII.3 Présentations orales :

- Brissette, F.*, **Martel, J.-L.**, Mailhot, A., Ludwig, R., Turcotte, R., Frigon, A. & Leduc, M. (2018). ClimEx projection of extreme precipitation values. A unique high-resolution large ensemble workshop, Québec, Canada, 15 – 16 mai 2018.
- Leduc, M.*, Mailhot, A., Frigon, A., **Martel, J.-L.**, Bopp, M., Laprise, R. & Ludwig, R. (2018). Validation of ClimEx at the monthly scale and climate projections. A unique high-resolution large ensemble workshop, Québec, Canada, 15 – 16 mai 2018.
- Brissette, F., **Martel, J.-L.**, Bertrand, R., Vandal, M. & Arsenault, R. (2018). Recent trends in precipitation, temperature and streamflow extremes – from gauges to reanalysis to climate models – impact of spatial scale. 8e GEWEX (Global Energy and Water Exchanges project) Science Conference, Canmore, Canada, 6 – 11 mai 2018.
- Wood, R. R., **Martel, J.-L.**, von Trentini, F., Leduc, M., Frigon, A., Willkofer, F. & Ludwig, R. (2018). Analyzing natural variability and future changes of precipitation extremes in the new ClimEx single-model large-ensemble. European Geophysical Union (EGU) General Assembly 2018, Vienne, Autriche, 8 – 13 avril 2018.
- Ludwig, R., Braun, M., Brissette, F., Frigon, A., Komischke, H., Kranzlmueller, D., Leduc, M., Mailhot, A., **Martel, J.-L.**, Ricard, S., Schmid, F.-J., Turcotte, R., von Trentini, F., Weismueller, J., Willkofer, F. & Wood, R. (2018). Assessing the impacts of

climate change and natural variability on hydrological extreme events over Eastern North America and Europe. EGU General Assembly 2018, Vienne, Autriche, 8 – 13 avril 2018.

- **Martel, J.-L.**, Brissette, F., Mailhot, A., Wood, R. R., Ludwig, R., Frigon, A., Leduc, M. & Turcotte, R. (2017). Evolution of precipitation extremes in three large ensembles of climate simulations – Impact of spatial and temporal resolutions. American Geophysical Union (AGU) Fall Meeting 2017, Nouvelle-Orléans, États-Unis, 11 – 15 décembre 2017.
- Wood, R. R., **Martel, J.-L.**, Willkofer, F., von Trentini, F., Schmid, F.-J., Leduc, M., Frigon, A. & Ludwig, R. (2017). Extreme value analysis of hydro meteorological extremes in the ClimEx Large-Ensemble. AGU Fall Meeting 2017, Nouvelle-Orléans, États-Unis, 11 – 15 décembre 2017.
- **Martel, J.-L.***, Brissette, F., Mailhot, A., Wood, R. R., Ludwig, R., Frigon, A., Leduc, M. & Turcotte, R. (2017). Évolution des précipitations extrêmes dans trois grands ensembles de simulations climatiques – impact des résolutions spatiales et temporelles. 7e Symposium Ouranos, Montréal, Canada, 15 – 17 novembre 2017.
- **Martel, J.-L.**, Brissette, F., Mailhot, A., Wood, R. R., Ludwig, R., Frigon, A., Leduc, M. & Turcotte, R. (2017). Evolution of precipitation extremes in three large ensembles of climate simulations – Impact of spatial and temporal resolutions. Symposium ClimEx, Munich, Allemagne, 20 – 21 juin 2017.
- Leduc, M., Giguère, M., Ludwig, R., Frigon, A., Braun, M., Schmid, F.-J., Willkofer, F., Wood, R., von Trentini, F., Sushama, L., Scinocca, J., Turcotte, R., Brissette, F., Ricard, S., **Martel, J.-L.**, Brietzke, G. B., Valin, M. & Komischke, H. (2017). An overview of the CRCM5 large ensemble at 12-km. Symposium ClimEx, Munich, Allemagne, 20 – 21 juin 2017.
- Leduc, M., Giguère, M., Ludwig, R., Frigon, A., Braun, M., Schmid, F.-J., Willkofer, F., Wood, R. R., von Trentini, F., Sushama, L., Scinocca, J., Turcotte, R., Brissette, F., Ricard, S., **Martel, J.-L.**, Brietzke, G. B., Valin, M. & Komischke, H. (2017). A 50-member ensemble from a regional climate model at 12-km resolution to address local impacts from natural variability and extreme events under climate change. Canadian

Meteorological and Oceanographic Society's 51st Congress, Toronto, Canada, 4 – 8 juin 2017.

- **Martel, J.-L.**, Mailhot, A., Talbot, G., Brissette, F., Ludwig, R., Frigon, A., Leduc, M. & Turcotte, R. (2017). Evolution of precipitation extremes in two large ensembles of climate simulations. EGU General Assembly 2017, Vienne, Autriche, 23 – 28 avril 2017.
- Ludwig, R., Baese, F., Braun, M., Brietzke, G., Brissette, F., Frigon, A., Giguère, M., Komischke, H., Kranzlmueller, D., Leduc, M., **Martel, J.-L.**, Ricard, S., Schmid, F.-J., von Trentini, F., Turcotte, R., Weismueller, J., Willkofer, F. & Wood, R. R. (2017). ClimEx – Climate change and hydrological extreme events – risks and perspectives for water management in Bavaria and Québec. EGU General Assembly 2017, Vienne, Autriche, 23 – 28 avril 2017.
- Brissette, F., **Martel, J.-L.** & Zao, C. (2016). Changement climatique et variabilité naturelle climatique – les défis de l'adaptation dans la gestion des ressources en eau, XIi^{ème} journées internationales de technologie – JIT 2016 – Le changement climatique, l'environnement et la transition énergétique. Quelles implications sur l'enseignement et la recherche? Montréal, Canada, 24 – 26 août 2016.
- **Martel, J.-L.**, Brissette, F., Chen, J., Mailhot, A. & Caya, D. (2016). Impacts of forced and unforced climate variability on extreme floods using a large climate ensemble. EGU General Assembly 2016, Vienne, Autriche, 17 – 22 avril 2016.
- **Martel, J.-L.**, Brissette, F., Caya, D. & Mailhot, A. (2015). Évaluation de l'impact de la variabilité naturelle du climat sur l'estimation statistique des extrêmes hydrologiques. Colloque RHQ (La recherche hydrologique au Québec) 2015, Montréal, Canada, 9 – 10 juin 2015.
- **Martel, J.-L.**, Brissette, F., Caya, D. & Mailhot, A. (2015). Evaluation of the impact of natural climate variability on the estimation of hydrological extremes. 68e Conférence annuelle de l'Association canadienne des ressources hydriques, Winnipeg, Canada, 2 – 4 juin 2015.

* *Conférencier invité.*

VIII.4 Présentation par affiche :

- **Martel, J.-L.**, Brissette, F., Mailhot, A., Ludwig, R., Leduc, M. & Turcotte, R. (2018). Projected future changes in extreme flood events over North American catchments using two large ensembles of climate simulations. 8e GEWEX Science Conference, Canmore, Canada, 6 – 11 mai 2018.
- Willkofer, F., **Martel, J.-L.**, Leduc, M., Braun, M., Brissette, F., Ricard, S., Turcotte, R., Weismüller, J., Wood, R. R., von Trentini, F., Schmid, F.-J. & Ludwig, R. (2018). Assessing the impacts of climate change and natural climate variability on hydro-meteorological extreme events. 8th GEWEX Science Conference, Canmore, Canada, May 6 – 11, 2018.
- Wood, R. R., **Martel, J.-L.**, von Trentini, F., Leduc, M., Frigon, A. & Ludwig, R. (2018). Analyzing natural variability and future changes of precipitation extremes in the new ClimEx-Ensemble. 8th GEWEX Science Conference, Canmore, Canada, May 6 – 11, 2018.
- von Trentini, F., Schmid, F.-J., Braun, M., Brissette, F., Frigon, A., Leduc, M., **Martel, J.-L.**, Willkofer, F., Wood, R. R. & Ludwig, R. (2017). Evaluating the ClimEx single model large ensemble in comparison with EURO-CORDEX results of seasonal means and extreme precipitation indicators. AGU Fall Meeting 2017, Nouvelle-Orléans, États-Unis, 11 – 15 décembre 2017.
- Schmid, F.-J., von Trentini, F., Braun, M., Frigon, A., Leduc, M., **Martel, J.-L.**, Willkofer, F., Wood, R. R. & Ludwig, R. Evaluating the ClimEx single model large ensemble in comparison with EURO-CORDEX results of heatwave and drought indicators. AGU Fall Meeting 2017, Nouvelle-Orléans, États-Unis, 11 – 15 décembre 2017.
- **Martel, J.-L.**, Mailhot, A., Brissette, F. & Caya, D. (2017). Rôle de la variabilité naturelle du climat sur la détection du signal du changement climatique d'origine anthropique sur les précipitations extrêmes. Colloque RHQ 2017, Québec, Canada, 15 – 16 mai 2017.

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

Aalbers, E., Lenderink, G., van Meijgaard, E., & van den Hurk, B. J. J. M. (2017). Local-scale changes in mean and heavy precipitation in Western Europe, climate change or internal variability? *Climate Dynamics*. doi: 10.1007/s00382-017-3901-9

Addor, N., & Fischer, E. M. (2015). The influence of natural variability and interpolation errors on bias characterization in RCM simulations. *Journal of Geophysical Research: Atmospheres, 120*(19), 10,180-110,195. doi: 10.1002/2014JD022824

Addor, N., & Seibert, J. (2014). Bias correction for hydrological impact studies – beyond the daily perspective. *Hydrological Processes, 28*(17), 4823-4828. doi: 10.1002/hyp.10238

Ajami, N. K., Duan, Q., Gao, X., & Sorooshian, S. (2006). Multimodel combination techniques for analysis of hydrological simulations: Application to distributed model intercomparison project results. *Journal of Hydrometeorology, 7*(4), 755-768. doi: 10.1175/jhm519.1

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., . . . Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres, 111*(D5), D05109. doi: 10.1029/2005JD006290

Anandhi, A., Zion, M. S., Gowda, P. H., Pierson, D. C., Lounsbury, D., & Frei, A. (2013). Past and future changes in frost day indices in Catskill Mountain region of New York. *Hydrological Processes, 27*(21), 3094-3104. doi: 10.1002/hyp.9937

Anctil, F., & Coulibaly, P. (2004). Wavelet analysis of the interannual variability in southern Québec streamflow. *Journal of Climate, 17*(1), 163-173. doi: 10.1175/1520-0442(2004)017<0163:WAOTIV>2.0.CO;2

Andreassian, V., Hall, A., Chahinian, N., & Schaake, J. (2006). Large sample basin experiments for hydrological model parameterization: results of the Model Parameter Experiment (MOPEX). 346 p.

Andrews, E. D., Antweiler, R. C., Neiman, P. J., & Ralph, F. M. (2004). Influence of ENSO on flood frequency along the California coast. *Journal of Climate, 17*(2), 337-348. doi: 10.1175/1520-0442(2004)017<0337:IOEOFF>2.0.CO;2

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., . . . Jha, M. K. (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE, 55*(4), 1491. doi: 10.13031/2013.42256

Arnone, E., Pumo, D., Viola, F., Noto, L. V., & La Loggia, G. (2013). Rainfall statistics changes in Sicily. *Hydrology and Earth System Sciences, 17*(7), 2449-2458. doi: 10.5194/hess-17-2449-2013

Arora, V. K., Scinocca, J. F., Boer, G. J., Christian, J. R., Denman, K. L., Flato, G. M., . . . Merryfield, W. J. (2011). Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophysical Research Letters, 38*(5). doi: 10.1029/2010GL046270

Arsenault, R., Bazile, R., Ouellet Dallaire, C., & Brissette, F. (2016a). CANOPEX: A Canadian hydrometeorological watershed database. *Hydrological Processes, 30*(15), 2734-2736. doi: 10.1002/hyp.10880

Arsenault, R., & Brissette, F. (2014). Determining the optimal spatial distribution of weather station networks for hydrological modeling purposes using RCM datasets: An experimental approach. *Journal of Hydrometeorology, 15*(1), 517-526. doi: 10.1175/jhm-d-13-088.1

Arsenault, R., & Brissette, F. (2016). Analysis of continuous streamflow regionalization methods within a virtual setting. *Hydrological Sciences Journal, 61*(15), 2680-2693. doi: 10.1080/02626667.2016.1154557

Arsenault, R., Essou, G. R. C., & Brissette, F. P. (2017). Improving hydrological model simulations with combined multi-input and multimodel averaging frameworks. *Journal of Hydrologic Engineering, 22*(4), 04016066. doi: 10.1061/(ASCE)HE.1943-5584.0001489

Arsenault, R., Gatien, P., Renaud, B., Brissette, F., & Martel, J.-L. (2015). A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. *Journal of Hydrology, 529*, 754-767. doi: 10.1016/j.jhydrol.2015.09.001

Arsenault, R., Latraverse, M., & Duchesne, T. (2016b). An efficient method to correct under-dispersion in ensemble streamflow prediction of inflow volumes for reservoir optimization. *Water Resources Management, 30*(12), 4363-4380. doi: 10.1007/s11269-016-1425-4

Arsenault, R., Poissant, D., & Brissette, F. (2015). Parameter dimensionality reduction of a conceptual model for streamflow prediction in Canadian, snowmelt dominated ungauged basins. *Advances in Water Resources, 85*, 27-44. doi: 10.1016/j.advwatres.2015.08.014

Arsenault, R., Poulin, A., Côté, P., & Brissette, F. (2014). Comparison of stochastic optimization algorithms in hydrological model calibration. *Journal of Hydrologic Engineering, 19*(7), 1374-1384. doi: 10.1061/(ASCE)HE.1943-5584.0000938

Assani, A. A., Landais, D., Mesfioui, M., & Matteau, M. (2010). Relationship between the Atlantic Multidecadal Oscillation index and variability of mean annual flows for catchments in the St. Lawrence watershed (Quebec, Canada) during the past century. *Hydrology Research, 41*(2), 115-125. doi: 10.2166/nh.2010.055

Assani, A. A., Landry, R., & Laurencelle, M. (2012). Comparison of interannual variability modes and trends of seasonal precipitation and streamflow in southern Quebec (Canada). *River Research and Applications, 28*(10), 1740-1752. doi: 10.1002/rra.1544

Avila, F. B., Dong, S., Menang, K. P., Rajczak, J., Renom, M., Donat, M. G., & Alexander, L. V. (2015). Systematic investigation of gridding-related scaling effects on annual statistics of daily temperature and precipitation maxima: A case study for south-east Australia. *Weather and Climate Extremes, 9*(Supplement C), 6-16. doi: 10.1016/j.wace.2015.06.003

Bae, D.-H., Jung, I.-W., & Lettenmaier, D. P. (2011). Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea. *Journal of Hydrology, 401*(1), 90-105. doi: 10.1016/j.jhydrol.2011.02.012

Baede, A. P. M., Ahlonsou, E., Ding, Y., & Schimel, D. S. (2001). The climate system: an overview. Dans *Climate Change 2001: impacts, adaptation and vulnerability* (pp. 87-98). Cambridge University Press.

Ban, N., Schmidli, J., & Schär, C. (2014). Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research: Atmospheres, 119*(13), 7889-7907. doi: 10.1002/2014JD021478

Band, L. E., Patterson, P., Nemani, R., & Running, S. W. (1993). Forest ecosystem processes at the watershed scale: incorporating hillslope hydrology. *Agricultural and Forest Meteorology, 63*(1), 93-126. doi: 10.1016/0168-1923(93)90024-C

Barlow, M., Nigam, S., & Berbery, E. H. (2001). ENSO, pacific decadal variability, and U.S. summertime precipitation, drought, and stream flow. *Journal of Climate, 14*(9), 2105-2128. doi: 10.1175/1520-0442(2001)014<2105:Epdvau>2.0.Co;2

Barnett, T. P., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., . . . Tett, S. (1999). Detection and attribution of recent climate change: A status report. *Bulletin of*

the American Meteorological Society, 80(12), 2631-2659. doi: 10.1175/1520-0477(1999)080<2631:Daaorc>2.0.Co;2

Bates, B., Kundzewicz, Z., & Wu, S. (2008). *Climate change and water*. Intergovernmental Panel on Climate Change Secretariat.

Beauchamp, J., Leconte, R., Trudel, M., & Brissette, F. (2013). Estimation of the summer-fall PMP and PMF of a northern watershed under a changed climate. *Water Resources Research, 49*(6), 3852-3862. doi: 10.1002/wrcr.20336

Bell, G. D., & Chelliah, M. (2006). Leading tropical modes associated with interannual and multidecadal fluctuations in North Atlantic hurricane activity. *Journal of Climate, 19*(4), 590-612. doi: 10.1175/jcli3659.1

Benestad, R. E., Hanssen-Bauer, I., & Chen, D. (2008). *Empirical-statistical downscaling*. World Scientific Publishing Company.

Bennett, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., . . . Andreassian, V. (2013). Characterising performance of environmental models. *Environmental Modelling & Software, 40*, 1-20. doi: 10.1016/j.envsoft.2012.09.011

Bergstrom, S. (1976). Development and application of a conceptual runoff model for Scandinavian catchments. Swedish Meteorological and Hydrological Institute Rep. RH07, 134 pp.

Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Pettersson, A., & Rummukainen, M. (2001). Climate change impacts on runoff in Sweden - assessments by global climate models, dynamical downscaling and hydrological modelling. *Climate Research, 16*(2), 101-112.

Biondi, D., Freni, G., Iacobellis, V., Mascaro, G., & Montanari, A. (2012). Validation of hydrological models: Conceptual basis, methodological approaches and a proposal for a code of practice. *Physics and Chemistry of the Earth, Parts A/B/C, 42-44*, 70-76. doi: 10.1016/j.pce.2011.07.037

Biron, S., Assani, A. A., Frenette, J.-J., & Massicotte, P. (2014). Comparison of Lake Ontario and St. Lawrence River hydrologic droughts and their relationship to climate indices. *Water Resources Research, 50*(2), 1396-1409. doi: 10.1002/2012WR013441

Bohn, T. J., Sonessa, M. Y., & Lettenmaier, D. P. (2010). Seasonal hydrologic forecasting: Do multimodel ensemble averages always yield improvements in forecast skill? *Journal of Hydrometeorology, 11*(6), 1358-1372. doi: 10.1175/2010jhm1267.1

Bonsal, B., & Shabbar, A. (2008). Impacts of large-scale circulation variability on low streamflows over Canada: A review. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 33*(2), 137-154. doi: 10.4296/cwrj3302137

Bonsal, B., Shabbar, A., & Higuchi, K. (2001). Impacts of low frequency variability modes on Canadian winter temperature. *International Journal of Climatology, 21*(1), 95-108. doi: 10.1002/joc.590

Booij, M. J. (2005). Impact of climate change on river flooding assessed with different spatial model resolutions. *Journal of Hydrology, 303*(1), 176-198. doi: 10.1016/j.jhydrol.2004.07.013

Boyle, D. P., Gupta, H. V., & Sorroshian, S. (2000). Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research, 36*(12), 3663-3674. doi: 10.1029/2000WR900207

Brabets, T. P., & Walvoord, M. A. (2009). Trends in streamflow in the Yukon River basin from 1944 to 2005 and the influence of the Pacific Decadal Oscillation. *Journal of Hydrology, 371*(1), 108-119. doi: 10.1016/j.jhydrol.2009.03.018

Brayshaw, D. J., Woollings, T., & Vellinga, M. (2009). Tropical and extratropical responses of the North Atlantic atmospheric circulation to a sustained weakening of the MOC. *Journal of Climate, 22*(11), 3146-3155. doi: 10.1175/2008jcli2594.1

BRIC. (2016). Éclairer l'avenir : Bulletin de rendement des infrastructures canadiennes. 168.

Brigode, P., Oudin, L., & Perrin, C. (2013). Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology, 476*, 410-425. doi: 10.1016/j.jhydrol.2012.11.012

Brommer, D. M., Cerveny, R. S., & Balling, R. C. (2007). Characteristics of long-duration precipitation events across the United States. *Geophysical Research Letters, 34*(22), L22712. doi: 10.1029/2007GL031808

Brown, P. T., Ming, Y., Li, W., & Hill, S. A. (2017). Change in the magnitude and mechanisms of global temperature variability with warming. *Nature Climate Change, 7*, 743. doi: 10.1038/nclimate3381

Burn, D. H. (2008). Climatic influences on streamflow timing in the headwaters of the Mackenzie River Basin. *Journal of Hydrology, 352*(1), 225-238. doi: 10.1016/j.jhydrol.2008.01.019

Burn, D. H., & Hag Elnur, M. A. (2002). Detection of hydrologic trends and variability. *Journal of Hydrology, 255*(1), 107-122. doi: 10.1016/S0022-1694(01)00514-5

Burn, D. H., Mansour, R., Zhang, K., & Whitfield, P. H. (2011). Trends and variability in extreme rainfall events in British Columbia. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 36*(1), 67-82. doi: 10.4296/cwrj3601067

Burn, D. H., & Whitfield, P. H. (2016). Changes in floods and flood regimes in Canada. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41*(1-2), 139-150. doi: 10.1080/07011784.2015.1026844

Buser, C. M., Künsch, H. R., Lüthi, D., Wild, M., & Schär, C. (2009). Bayesian multi-model projection of climate: bias assumptions and interannual variability. *Climate Dynamics, 33*(6), 849-868. doi: 10.1007/s00382-009-0588-6

Butts, M. B., Payne, J. T., Kristensen, M., & Madsen, H. (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology, 298*(1), 242-266. doi: 10.1016/j.jhydrol.2004.03.042

Canada, E. (2014). National guideline on principles of climate change adaptation for professional engineers 18.

Caya, D., & Laprise, R. (1999). A semi-implicit semi-lagrangian regional climate model: The Canadian RCM. *Monthly Weather Review, 127*(3), 341-362. doi: 10.1175/1520-0493(1999)127<0341:Asislr>2.0.Co;2

Cayan, D. R., Redmond, K. T., & Riddle, L. G. (1999). ENSO and hydrologic extremes in the western United States. *Journal of Climate, 12*(9), 2881-2893. doi: 10.1175/1520-0442(1999)012<2881:Eaheit>2.0.Co;2

Chan, S. C., Kendon, E. J., Fowler, H. J., Blenkinsop, S., Roberts, N. M., & Ferro, C. A. T. (2014). The value of high-resolution met office regional climate models in the simulation of multihourly precipitation extremes. *Journal of Climate, 27*(16), 6155-6174. doi: 10.1175/JCLI-D-13-00723.1

Chandimala, J., & Zubair, L. (2007). Predictability of stream flow and rainfall based on ENSO for water resources management in Sri Lanka. *Journal of Hydrology, 335*(3), 303-312. doi: 10.1016/j.jhydrol.2006.11.024

Chen, C.-T., & Knutson, T. (2008). On the verification and comparison of extreme rainfall indices from climate models. *Journal of Climate, 21*(7), 1605-1621. doi: 10.1175/2007jcli1494.1

Chen, J., & Brissette, F. (2018). Reliability of climate model multi-member ensembles in estimating internal precipitation and temperature variability at the multi-decadal scale. *International Journal of Climatology, (Accepted)*.

Chen, J., Brissette, F. P., Chaumont, D., & Braun, M. (2013). Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins. *Journal of Hydrology, 479*, 200-214. doi: 10.1016/j.jhydrol.2012.11.062

Chen, J., Brissette, F. P., & Leconte, R. (2011a). Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *Journal of Hydrology, 401*(3), 190-202. doi: 10.1016/j.jhydrol.2011.02.020

Chen, J., Brissette, F. P., Liu, P., & Xia, J. (2017). Using raw regional climate model outputs for quantifying climate change impacts on hydrology. *Hydrological Processes, 31*(24), 4398-4413. doi: 10.1002/hyp.11368

Chen, J., Brissette, F. P., & Lucas-Picher, P. (2015). Assessing the limits of bias-correcting climate model outputs for climate change impact studies. *Journal of Geophysical Research: Atmospheres, 120*(3), 1123-1136. doi: 10.1002/2014JD022635

Chen, J., Brissette, F. P., Poulin, A., & Leconte, R. (2011b). Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research, 47*(12). doi: 10.1029/2011WR010602

Chen, J., Li, C., Brissette, F. P., Chen, H., Wang, M., & Essou, G. R. C. (2018). Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling. *Journal of Hydrology, 560*, 326-341. doi: 10.1016/j.jhydrol.2018.03.040

Chikamoto, Y., Timmermann, A., Stevenson, S., DiNezio, P., & Langford, S. (2015). Decadal predictability of soil water, vegetation, and wildfire frequency over North America. *Climate Dynamics, 45*(7), 2213-2235. doi: 10.1007/s00382-015-2469-5

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., . . . Brekke, L. D. (2016). Characterizing uncertainty of the hydrologic impacts of climate change. *Current Climate Change Reports, 2*(2), 55-64. doi: 10.1007/s40641-016-0034-x

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., . . . Krinner, G. (2013). Long-term climate change: projections, commitments and irreversibility. In: Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F. and Qin, D. and Plattner, G.-K. and Tignor, M. and

Allen, S.K. and Boschung, J. and Nauels, A. and Xia, Y. and Bex, V. and Midgley, P.M.]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., . . . Smith, R. D. (2006). The Community Climate System Model Version 3 (CCSM3). *Journal of Climate, 19*(11), 2122-2143. doi: 10.1175/jcli3761.1

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., . . . Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters, 8*(2), 024024. doi: 10.1088/1748-9326/8/2/024024

Cook, J., Oreskes, N., Doran, P. T., Anderegg, W. R. L., Verheggen, B., Maibach, E. W., . . . Rice, K. (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters, 11*(4), 048002. doi: 10.1088/1748-9326/11/4/048002

Coppola, E., Sobolowski, S., Pichelli, E., Raffaele, F., Ahrens, B., Anders, I., . . . Warrach-Sagi, K. (2018). A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. *Climate Dynamics*. doi: 10.1007/s00382-018-4521-8

Córdoba-Machado, S., Palomino-Lemus, R., Gámiz-Fortis, S. R., Castro-Díez, Y., & Esteban-Parra, M. J. (2016). Seasonal streamflow prediction in Colombia using atmospheric and oceanic patterns. *Journal of Hydrology, 538*, 1-12. doi: 10.1016/j.jhydrol.2016.04.003

Coulibaly, P., & Burn, D. H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resources Research, 40*(3), W03105. doi: 10.1029/2003WR002667

Coulibaly, P., & Burn, D. H. (2005). Spatial and temporal variability of Canadian seasonal streamflows. *Journal of Climate, 18*(1), 191-210. doi: 10.1175/jcli-3258.1

CSA. (2012). National survey of Canada's infrastructure engineers about climate change 39.

Cunnane, C. (1978). Unbiased plotting positions — A review. *Journal of Hydrology, 37*(3), 205-222. doi: 10.1016/0022-1694(78)90017-3

Daggupati, P., Pai, N., Ale, S., Douglas-Mankin, K. R., Zeckoski, R. W., Jeong, J., . . . Youssef, M. (2015). A recommended calibration and validation strategy for

hydrologic and water quality models. *Transactions of the ASABE, 58*(6), 1705. doi: 10.13031/trans.58.10712

Dankers, R., & Feyen, L. (2008). Climate change impact on flood hazard in Europe: An assessment based on high-resolution climate simulations. *Journal of Geophysical Research: Atmospheres, 113*(D19). doi: 10.1029/2007JD009719

Dankers, R., Feyen, L., & Christensen, O. B. (2009). On the benefit of high-resolution climate simulations in impact studies of hydrological extremes. *Hydrology and Earth System Sciences Discussions, 6*(2), 2573-2597. doi: 10.5194/hessd-6-2573-2009

Davies, H. C. (1976). A laterul boundary formulation for multi-level prediction models. *Quarterly Journal of the Royal Meteorological Society, 102*(432), 405-418. doi: 10.1002/qj.49710243210

Day, G. N. (1985). Extended Streamflow Forecasting Using NWSRFS. *Journal of Water Resources Planning and Management, 111*(2), 157-170. doi: 10.1061/(ASCE)0733-9496(1985)111:2(157)

de Elía, R., Biner, S., & Frigon, A. (2013). Interannual variability and expected regional climate change over North America. *Climate Dynamics, 41*(5), 1245-1267. doi: 10.1007/s00382-013-1717-9

de Elía, R., Laprise, R., & Denis, B. (2002). Forecasting skill limits of nested, limited-area models: A perfect-model approach. *Monthly Weather Review, 130*(8), 2006-2023. doi: 10.1175/1520-0493(2002)130<2006:Fslonl>2.0.Co;2

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., . . . Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society, 137*(656), 553-597. doi: 10.1002/qj.828

Déqué, M., Rowell, D. P., Lüthi, D., Giorgi, F., Christensen, J. H., Rockel, B., . . . van den Hurk, B. (2007). An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change, 81*(1), 53-70. doi: 10.1007/s10584-006-9228-x

Déry, S. J., Hernández-Henríquez, M. A., Owens, P. N., Parkes, M. W., & Petticrew, E. L. (2012). A century of hydrological variability and trends in the Fraser River basin. *Environmental Research Letters, 7*(2), 024019. doi: 10.1088/1748-9326/7/2/024019

Déry, S. J., & Wood, E. F. (2004). Teleconnection between the Arctic Oscillation and Hudson Bay River discharge. *Geophysical Research Letters, 31*(18), L18205. doi: 10.1029/2004GL020729

Déry, S. J., & Wood, E. F. (2005). Decreasing river discharge in northern Canada. *Geophysical Research Letters, 32*(10), L10401. doi: 10.1029/2005GL022845

Deser, C., Knutti, R., Solomon, S., & Phillips, A. S. (2012a). Communication of the role of natural variability in future North American climate. *Nature Clim. Change, 2*(11), 775-779. doi: 10.1038/nclimate1562

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012b). Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics, 38*(3), 527-546. doi: 10.1007/s00382-010-0977-x

Deser, C., Phillips, A. S., Alexander, M. A., & Smoliak, B. V. (2014). Projecting North American climate over the next 50 Years: Uncertainty due to internal variability. *Journal of Climate, 27*(6), 2271-2296. doi: 10.1175/JCLI-D-13-00451.1

Detzel, D. H. M., & Mine, M. R. M. (2014). Trends in hydrological series: methods and application. Dans *11th International Conference on Hydroscience & Engineering, Hamburg, Germany*.

Diaconescu, E. P., Gachon, P., Laprise, R., & Scinocca, J. F. (2016). Evaluation of precipitation indices over North America from various configurations of regional climate models. *Atmosphere-Ocean, 54*(4), 418-439. doi: 10.1080/07055900.2016.1185005

Dittus, A. J., Karoly, D. J., Lewis, S. C., & Alexander, L. V. (2015). A multiregion assessment of observed changes in the areal extent of temperature and precipitation extremes. *Journal of Climate, 28*(23), 9206-9220. doi: 10.1175/jcli-d-14-00753.1

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., & Caesar, J. (2013a). Global land-based datasets for monitoring climatic extremes. *Bulletin of the American Meteorological Society, 94*(7), 997-1006. doi: 10.1175/bams-d-12-00109.1

Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Dunn, R. J. H., . . . Kitching, S. (2013b). Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal of Geophysical Research: Atmospheres, 118*(5), 2098-2118. doi: 10.1002/jgrd.50150

Douglas, E. M., Vogel, R. M., & Kroll, C. N. (2000). Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology, 240*(1), 90-105. doi: 10.1016/S0022-1694(00)00336-X

Duan, Q., Ajami, N. K., Gao, X., & Sorooshian, S. (2007). Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources, 30*(5), 1371-1386. doi: 10.1016/j.advwatres.2006.11.014

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H. V., . . . Wood, E. F. (2006). Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology, 320*(1), 3-17. doi: 10.1016/j.jhydrol.2005.07.031

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research, 28*(4), 1015-1031. doi: 10.1029/91WR02985

Duan, Q., Sorooshian, S., & Gupta, V. K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology, 158*(3), 265-284. doi: 10.1016/0022-1694(94)90057-4

Easterling, D. R., Evans, J. L., Groisman, P. Y., Karl, T. R., Kunkel, K. E., & Ambenje, P. (2000). Observed variability and trends in extreme climate events: A brief review. *Bulletin of the American Meteorological Society, 81*(3), 417-425. doi: 10.1175/1520-0477(2000)081<0417:Ovatie>2.3.Co;2

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., & Liebert, J. (2012). HESS Opinions" Should we apply bias correction to global and regional climate model data?". *Hydrology and Earth System Sciences, 16*(9), 3391-3404. doi: 10.5194/hess-16-3391-2012

Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J., & Pappenberger, F. (2017). Complex picture for likelihood of ENSO-driven flood hazard. *Nature Communications, 8*, 14796. doi: 10.1038/ncomms14796

Enfield, D. B., Mestas-Nuñez, A. M., & Trimble, P. J. (2001). The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophysical Research Letters, 28*(10), 2077-2080. doi: 10.1029/2000GL012745

Essou, G. R. C., Arsenault, R., & Brissette, F. P. (2016). Comparison of climate datasets for lumped hydrological modeling over the continental United States. *Journal of Hydrology, 537*, 334-345. doi: 10.1016/j.jhydrol.2016.03.063

Fatichi, S., Rimkus, S., Burlando, P., & Bordoy, R. (2014). Does internal climate variability overwhelm climate change signals in streamflow? The upper Po and Rhone basin case studies. *Science of The Total Environment, 493*, 1171-1182. doi: 10.1016/j.scitotenv.2013.12.014

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., . . . Tarboton, D. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology, 537*, 45-60. doi: 10.1016/j.jhydrol.2016.03.026

Ferro, C. A. T. (2004). Attributing variation in a regional climate change modelling experiment. PRUDENCE Working Note, 21 pp. doi: http://prudence.dmi.dk/public/publications/analysis_of_variance.pdf

Fischer, E. M., Beyerle, U., & Knutti, R. (2013). Robust spatially aggregated projections of climate extremes. *Nature Clim. Change, 3*(12), 1033-1038. doi: 10.1038/nclimate2051

Fischer, E. M., & Knutti, R. (2014). Detection of spatially aggregated changes in temperature and precipitation extremes. *Geophysical Research Letters, 41*(2), 547-554. doi: 10.1002/2013GL058499

Fischer, E. M., Lawrence, D. M., & Sanderson, B. M. (2011). Quantifying uncertainties in projections of extremes—a perturbed land surface parameter experiment. *Climate Dynamics, 37*(7), 1381-1398. doi: 10.1007/s00382-010-0915-y

Fischer, E. M., Sedláček, J., Hawkins, E., & Knutti, R. (2014). Models agree on forced response pattern of precipitation and temperature extremes. *Geophysical Research Letters, 41*(23), 8554-8562. doi: 10.1002/2014GL062018

Fleming, S. W., Moore, R. D., & Clarke, G. K. C. (2006). Glacier-mediated streamflow teleconnections to the Arctic Oscillation. *International Journal of Climatology, 26*(5), 619-636. doi: 10.1002/joc.1273

Fleming, S. W., & Weber, F. A. (2012). Detection of long-term change in hydroelectric reservoir inflows: Bridging theory and practise. *Journal of Hydrology, 470-471*, 36-54. doi: 10.1016/j.jhydrol.2012.08.008

Fleming, S. W., & Whitfield, P. H. (2010). Spatiotemporal mapping of ENSO and PDO surface meteorological signals in British Columbia, Yukon, and southeast Alaska. *Atmosphere-Ocean, 48*(2), 122-131. doi: 10.3137/AO1107.2010

Fleming, S. W., Whitfield, P. H., Moore, R. D., & Quilty, E. J. (2007). Regime-dependent streamflow sensitivities to Pacific climate modes cross the Georgia–Puget transboundary ecoregion. *Hydrological Processes, 21*(24), 3264-3287. doi: 10.1002/hyp.6544

Fortin, V., & Turcotte, R. (2007). Le modèle hydrologique MOHYSE (in French). Université du Québec à Montréal Note SCA7420, 14 pp., .

Fowler, H. J., Blenkinsop, S., & 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. doi: 10.1002/joc.1556

Fu, C., James, A. L., & Wachowiak, M. P. (2012). Analyzing the combined influence of solar activity and El Niño on streamflow across southern Canada. *Water Resources Research, 48*(5), W05507. doi: 10.1029/2011WR011507

Fu, C., James, A. L., & Yao, H. (2015). Investigations of uncertainty in SWAT hydrologic simulations: a case study of a Canadian Shield catchment. *Hydrological Processes, 29*(18), 4000-4017. doi: 10.1002/hyp.10477

Fyfe, J. C., Derksen, C., Mudryk, L., Flato, G. M., Santer, B. D., Swart, N. C., . . . Jiao, Y. (2017). Large near-term projected snowpack loss over the western United States. *Nature Communications, 8*, 14996. doi: 10.1038/ncomms14996

Gaborit, É., Ricard, S., Lachance-Cloutier, S., Anctil, F., & Turcotte, R. (2015). Comparing global and local calibration schemes from a differential split-sample test perspective. *Canadian Journal of Earth Sciences, 52*(11), 990-999. doi: 10.1139/cjes-2015-0015

Garavaglia, F., Le Lay, M., Gottardi, F., Garçon, R., Gailhard, J., Paquet, E., & Mathevet, T. (2017). Impact of model structure on flow simulation and hydrological realism: from a lumped to a semi-distributed approach. *Hydrology and Earth System Sciences, 21*(8), 3937. doi: 10.5194/hess-21-3937-2017

Garcia, F., Folton, N., & Oudin, L. (2017). Which objective function to calibrate rainfall–runoff models for low-flow index simulations? *Hydrological Sciences Journal, 62*(7), 1149-1166. doi: 10.1080/02626667.2017.1308511

Gaur, S., Paul, P. K., Singh, R., Mishra, A., Gupta, P. K., & Singh, R. P. (2017). Operational testing of Satellite based Hydrological Model (SHM). Dans *EGU General Assembly Conference Abstracts* (Vol. 19, pp. 2543).

Gehne, M., Hamill, T. M., Kiladis, G. N., & Trenberth, K. E. (2016). Comparison of global precipitation estimates across a range of temporal and spatial scales. *Journal of Climate, 29*(21), 7773-7795. doi: 10.1175/jcli-d-15-0618.1

Georgakakos, K. P., Seo, D.-J., Gupta, H., Schaake, J., & Butts, M. B. (2004). Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. *Journal of Hydrology, 298*(1), 222-241. doi: 10.1016/j.jhydrol.2004.03.037

Gershunov, A., & Barnett, T. P. (1998). ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Observations and model results. *Journal of Climate, 11*(7), 1575-1586. doi: 10.1175/1520-0442(1998)011<1575:Eioier>2.0.Co;2

Gershunov, A., & Cayan, D. R. (2003). Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *Journal of Climate, 16*(16), 2752-2765. doi: 10.1175/1520-0442(2003)016<2752:Hdpfot>2.0.Co;2

Gharari, S., Hrachowitz, M., Fenicia, F., & Savenije, H. H. G. (2013). An approach to identify time consistent model parameters: sub-period calibration. *Hydrology and Earth System Sciences, 17*(1), 149-161. doi: 10.5194/hess-17-149-2013

Giorgi, F. (2002). Dependence of the surface climate interannual variability on spatial scale. *Geophysical Research Letters, 29*(23), 16-11-16-14. doi: 10.1029/2002GL016175

Giorgi, F., & Bi, X. (2009). Time of emergence (TOE) of GHG-forced precipitation change hot-spots. *Geophysical Research Letters, 36*(6). doi: 10.1029/2009GL037593

Giorgi, F., & Francisco, R. (2000). Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Climate Dynamics, 16*(2-3), 169-182.

Giorgi, F., & Gutowski, W. J. (2015). Regional dynamical downscaling and the CORDEX initiative. *Annual Review of Environment and Resources, 40*(1), 467-490. doi: 10.1146/annurev-environ-102014-021217

Giorgi, F., Jones, C., & Asrar, G. R. (2009). Addressing climate information needs at the regional level: the CORDEX framework. *World Meteorological Organization (WMO) Bulletin, 58*(3), 175.

Giorgi, F., & 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. doi: 10.1175/1520-0442(2002)015<1141:Coaura>2.0.Co;2

Giorgi, F., Torma, C., Coppola, E., Ban, N., Schär, C., & Somot, S. (2016). Enhanced summer convective rainfall at Alpine high elevations in response to climate warming. *Nature Geoscience, 9*, 584. doi: 10.1038/ngeo2761

Gobena, A. K., & Gan, T. Y. (2006). Low-frequency variability in Southwestern Canadian stream flow: links with large-scale climate anomalies. *International Journal of Climatology, 26*(13), 1843-1869. doi: 10.1002/joc.1336

Gobena, A. K., & Gan, T. Y. (2009). The role of Pacific climate on low-frequency hydroclimatic variability and predictability in southern Alberta, Canada. *Journal of Hydrometeorology, 10*(6), 1465-1478. doi: 10.1175/2009jhm1119.1

Gobena, A. K., Weber, F. A., & Fleming, S. W. (2013). The role of large-scale climate modes in regional streamflow variability and implications for water supply forecasting: A case study of the Canadian Columbia River basin. *Atmosphere-Ocean, 51*(4), 380-391. doi: 10.1080/07055900.2012.759899

Goldenberg, S. B., Landsea, C. W., Mestas-Nuñez, A. M., & Gray, W. M. (2001). The recent increase in Atlantic hurricane activity: Causes and implications. *Science, 293*(5529), 474-479. doi: 10.1126/science.1060040

Goly, A., & Teegavarapu, R. S. V. (2014). Individual and coupled influences of AMO and ENSO on regional precipitation characteristics and extremes. *Water Resources Research, 50*(6), 4686-4709. doi: 10.1002/2013WR014540

Gowda, P. H., Mulla, D. J., Desmond, E. D., Ward, A. D., & Moriasi, D. N. (2012). ADAPT: Model use, calibration, and validation. *Transactions of the ASABE, 55*(4), 1345. doi: 10.13031/2013.42246

Graham, L. P., Hagemann, S., Jaun, S., & Beniston, M. (2007). On interpreting hydrological change from regional climate models. *Climatic Change, 81*(1), 97-122. doi: 10.1007/s10584-006-9217-0

Groisman, P. Y., Knight, R. W., Easterling, D. R., Karl, T. R., Hegerl, G. C., & Razuvaev, V. N. (2005). Trends in intense precipitation in the climate record. *Journal of Climate, 18*(9), 1326-1350. doi: 10.1175/jcli3339.1

Groisman, P. Y., Knight, R. W., & Karl, T. R. (2001). Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century. *Bulletin*

of the American Meteorological Society, 82(2), 219-246. doi: 10.1175/1520-0477(2001)082<0219:Hpahsi>2.3.Co;2

Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. *arXiv preprint arXiv:1706.04599*.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology, 377*(1), 80-91. doi: 10.1016/j.jhydrol.2009.08.003

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: a need to balance depth with breadth. *Hydrology and Earth System Sciences, 18*(2), 463. doi: 10.5194/hess-18-463-2014

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., . . . von Storch, J. S. (2016). High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geoscientific Model Development, 9*(1), 4185-4208. doi: 10.5194/gmd-9-4185-2016

Halpert, M. S., & Ropelewski, C. F. (1992). Surface temperature patterns associated with the Southern Oscillation. *Journal of Climate, 5*(6), 577-593. doi: 10.1175/1520-0442(1992)005<0577:Stpawt>2.0.Co;2

Hamed, K. H. (2009). Exact distribution of the Mann–Kendall trend test statistic for persistent data. *Journal of Hydrology, 365*(1), 86-94. doi: 10.1016/j.jhydrol.2008.11.024

Hamed, K. H., & Rao, A. R. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology, 204*(1), 182-196. doi: 10.1016/S0022-1694(97)00125-X

Hamlet, A. F., & Lettenmaier, D. P. (1999). Columbia river streamflow forecasting based on ENSO and PDO climate signals. *Journal of Water Resources Planning and Management, 125*(6), 333-341. doi: 10.1061/(ASCE)0733-9496(1999)125:6(333)

Hamon, W. R. (1961). Estimating potential evaporation. *Journal of the Hydraulics Division, 87*, 107–120.

Hansen, J., Sato, M., & Ruedy, R. (2012). Perception of climate change. *PNAS, 109*(37), 2415-2423.

Hansen, N., & Ostermeier, A. (1996). Adapting arbitrary normal mutation distributions in evolution strategies: the covariance matrix adaptation. Dans *Proceedings of IEEE*

International Conference on Evolutionary Computation (pp. 312-317). doi: 10.1109/ICEC.1996.542381

Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation, 9*(2), 159-195. doi: 10.1162/106365601750190398

Haque, M. M., Rahman, A., Hagare, D., Kibria, G., & Karim, F. (2015). Estimation of catchment yield and associated uncertainties due to climate change in a mountainous catchment in Australia. *Hydrological Processes, 29*(19), 4339-4349. doi: 10.1002/hyp.10492

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology, 34*(3), 623-642. doi: 10.1002/joc.3711

Hawkins, E., Edwards, T., & McNeall, D. (2014). Pause for thought. *Nature Clim. Change, 4*(3), 154-156. doi: 10.1038/nclimate2150

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society, 90*(8), 1095-1108. doi: 10.1175/2009bams2607.1

Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics, 37*(1), 407-418. doi: 10.1007/s00382-010-0810-6

Hawkins, E., & Sutton, R. (2012). Time of emergence of climate signals. *Geophysical Research Letters, 39*(1). doi: 10.1029/2011GL050087

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres, 113*(D20). doi: 10.1029/2008JD010201

Hegerl, G. C., Black, E., Allan, R. P., Ingram, W. J., Polson, D., Trenberth, K. E., . . . Zhang, X. (2015). Challenges in quantifying changes in the global water cycle. *Bulletin of the American Meteorological Society, 96*(7), 1097-1115. doi: 10.1175/bams-d-13-00212.1

Hegerl, G. C., von Storch, H., Hasselmann, K., Santer, B. D., Cubasch, U., & Jones, P. D. (1996). Detecting greenhouse-gas-induced climate change with an optimal fingerprint

method. *Journal of Climate, 9*(10), 2281-2306. doi: 10.1175/1520-0442(1996)009<2281:Dggicc>2.0.Co;2

Herold, N., Behrangi, A., & Alexander, L. V. (2017). Large uncertainties in observed daily precipitation extremes over land. *Journal of Geophysical Research: Atmospheres, 122*(2), 668-681. doi: 10.1002/2016JD025842

Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., . . . Soares, P. M. M. (2019). Comparison of statistical downscaling methods with respect to extreme events over Europe: Validation results from the perfect predictor experiment of the COST Action VALUE. *International Journal of Climatology, 0*(0). doi: 10.1002/joc.5469

Hidalgo, H. G., & Dracup, J. A. (2003). ENSO and PDO effects on hydroclimatic variations of the upper Colorado River basin. *Journal of Hydrometeorology, 4*(1), 5-23. doi: 10.1175/1525-7541(2003)004<0005:Eapeoh>2.0.Co;2

Higgins, R. W., Leetmaa, A., & Kousky, V. E. (2002). Relationships between climate variability and winter temperature extremes in the United States. *Journal of Climate, 15*(13), 1555-1572. doi: 10.1175/1520-0442(2002)015<1555:Rbcvaw>2.0.Co;2

Holmes, C. R., Woollings, T., Hawkins, E., & de Vries, H. (2016). Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *Journal of Climate, 29*(6), 2221-2236. doi: 10.1175/jcli-d-14-00735.1

Hosseinzadehtalaei, P., Tabari, H., & Willems, P. (2018). Precipitation intensity–duration–frequency curves for central Belgium with an ensemble of EURO-CORDEX simulations, and associated uncertainties. *Atmospheric Research, 200*, 1-12. doi: 10.1016/j.atmosres.2017.09.015

Hu, Z.-Z., & Huang, B. (2009). Interferential impact of ENSO and PDO on dry and wet conditions in the U.S. Great Plains. *Journal of Climate, 22*(22), 6047-6065. doi: 10.1175/2009JCLI2798.1

Huard, D., & Mailhot, A. (2008). Calibration of hydrological model GR2M using Bayesian uncertainty analysis. *Water Resources Research, 44*(2). doi: 10.1029/2007WR005949

Hulme, M., Barrow, E. M., Arnell, N. W., Harrison, P. A., Johns, T. C., & Downing, T. E. (1999). Relative impacts of human-induced climate change and natural climate variability. *Nature, 397*, 688. doi: 10.1038/17789

Hurrell, J. W. (1996). Influence of variations in extratropical wintertime teleconnections on northern hemisphere temperature. *Geophysical Research Letters, 23*(6), 665-668. doi: 10.1029/96GL00459

Hurrell, J. W., Kushnir, Y., Ottersen, G., & Visbeck, M. (2003). An overview of the North Atlantic Oscillation. Dans *The North Atlantic Oscillation: Climatic Significance and Environmental Impact* (pp. 1-35). doi: 10.1029/134GM01. Repéré à https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/134GM01

Hutchinson, M. F. (1995). Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems, 9*(4), 385-403. doi: 10.1080/02693799508902045

Hutchinson, M. F., McKenney, D. W., Lawrence, K., Pedlar, J. H., Hopkinson, R. F., Milewska, E., & Papadopol, 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. doi: 10.1175/2008jamc1979.1

Hutchinson, M. F., & Xu, T. (2004). Anusplin version 4.2 user guide. *Centre for Resource and Environmental Studies, The Australian National University, Canberra*.

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, United Kingdom and New York, NY, USA: Cambridge University Press. doi: 10.1017/CBO9781107415324. Repéré à www.climatechange2013.org

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., . . . Yiou, P. (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. *Regional Environmental Change, 14*(2), 563-578. doi: 10.1007/s10113-013-0499-2

Jain, S. K., & Sudheer, K. P. (2008). Fitting of Hydrologic Models: A Close Look at the Nash-Sutcliffe Index. *Journal of Hydrologic Engineering, 13*(10), 981-986. doi: 10.1061/(ASCE)1084-0699(2008)13:10(981)

Jakeman, A. J., & Hornberger, G. M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research, 29*(8), 2637-2649. doi: 10.1029/93WR00877

Jiang, T., Chen, Y. D., Xu, C.-Y., Chen, X., Chen, X., & Singh, V. P. (2007). Comparison of hydrological impacts of climate change simulated by six hydrological models in the

Dongjiang Basin, South China. *Journal of Hydrology, 336*(3), 316-333. doi: 10.1016/j.jhydrol.2007.01.010

Jones, T. R., & Randall, D. A. (2011). Quantifying the limits of convective parameterizations. *Journal of Geophysical Research: Atmospheres, 116*(D8). doi: 10.1029/2010JD014913

Jung, I.-W., & Chang, H. (2011). Assessment of future runoff trends under multiple climate change scenarios in the Willamette River basin, Oregon, USA. *Hydrological Processes, 25*(2), 258-277. doi: 10.1002/hyp.7842

Juston, J., Seibert, J., & Johansson, P. O. (2009). Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment. *Hydrological Processes, 23*(21), 3093-3109. doi: 10.1002/hyp.7421

Kahya, E., & Dracup, J. A. (1992). The relationship between ENSO events and California streamflows. *AIP Conference Proceedings, 277*(1), 86-95. doi: 10.1063/1.43894

Kahya, E., & Dracup, J. A. (1993). U.S. streamflow patterns in relation to the El Niño/Southern Oscillation. *Water Resources Research, 29*(8), 2491-2503. doi: 10.1029/93WR00744

Kalra, A., & Ahmad, S. (2012). Estimating annual precipitation for the Colorado River basin using oceanic-atmospheric oscillations. *Water Resources Research, 48*(6), W06527. doi: 10.1029/2011WR010667

Katz, R. W. (2013). Statistical methods for nonstationary extremes. Dans *Extremes in a Changing Climate: Detection, Analysis and Uncertainty* (pp. 15-37). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-4479-0_2

Kay, A. L., Davies, H. N., Bell, V. A., & Jones, R. G. (2009). Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change, 92*(1), 41-63. doi: 10.1007/s10584-008-9471-4

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., . . . Vertenstein, M. (2015). The Community Earth System Model (CESM) Large Ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society, 96*(8), 1333-1349. doi: 10.1175/BAMS-D-13-00255.1

Kendall, M. G. (1975). *Rank correlation methods* (4th Edition éd.). London: Charles Griffin.

Kendall, S. (1977). The advanced theory of statistics 4th edn (UK: Griffin).

Kendon, E. J., Ban, N., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., . . . Wilkinson, J. M. (2017). Do convection-permitting regional climate models improve projections of future precipitation change? *Bulletin of the American Meteorological Society, 98*(1), 79-93. doi: 10.1175/bams-d-15-0004.1

Kendon, E. J., Roberts, N. M., Senior, C. A., & Roberts, M. J. (2012). Realism of rainfall in a very high-resolution regional climate model. *Journal of Climate, 25*(17), 5791-5806. doi: 10.1175/JCLI-D-11-00562.1

Khaliq, M. N., Ouarda, T. B. M. J., Gachon, P., & Sushama, L. (2008). Temporal evolution of low-flow regimes in Canadian rivers. *Water Resources Research, 44*(8), W08436. doi: 10.1029/2007WR006132

Kharin, V. V., Zwiers, F. W., Zhang, X., & Hegerl, G. C. (2007). Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate, 20*(8), 1419-1444. doi: 10.1175/jcli4066.1

Kharin, V. V., Zwiers, F. W., Zhang, X., & Wehner, M. (2013). Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change, 119*(2), 345-357. doi: 10.1007/s10584-013-0705-8

Kiffney, P. M., Bull, J. P., & Feller, M. C. (2002). Climatic and hydrologic variability in a coastal watershed of southwestern Bristish Columbia. *Journal of the American Water Resources Association, 38*(5), 1437-1451. doi: 10.1111/j.1752-1688.2002.tb04357.x

Kiktev, D., Sexton, D. M. H., Alexander, L., & Folland, C. K. (2003). Comparison of modeled and observed trends in indices of daily climate extremes. *Journal of Climate, 16*(22), 3560-3571. doi: 10.1175/1520-0442(2003)016<3560:comaot>2.0.co;2

Kim, K. B., Kwon, H.-H., & Han, D. (2016). Precipitation ensembles conforming to natural variations derived from a regional climate model using a new bias correction scheme. *Hydrology and Earth System Sciences, 20*(5), 2019-2034. doi: 10.5194/hess-20-2019-2016

Kim, T.-W., Valdés, J. B., Nijssen, B., & Roncayolo, D. (2006). Quantification of linkages between large-scale climatic patterns and precipitation in the Colorado River basin. *Journal of Hydrology, 321*(1), 173-186. doi: 10.1016/j.jhydrol.2005.07.043

King, A. D., Donat, M. G., Fischer, E. M., Hawkins, E., Alexander, L. V., Karoly, D. J., . . . Perkins, S. E. (2015). The timing of anthropogenic emergence in simulated climate extremes. *Environmental Research Letters, 10*(9). doi: 10.1088/1748-9326/10/9/094015

Klein Tank, A. M. G., Zwiers, F. W., & Zhang, X. (2009). *Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation*. World Meteorological Organization.

Klemeš, V. (1986). Operational testing of hydrological simulation models. *Hydrological Sciences Journal, 31*(1), 13-24. doi: 10.1080/02626668609491024

Knutti, R., Masson, D., & Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters, 40*(6), 1194-1199. doi: 10.1002/grl.50256

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., . . . Wulfmeyer, V. (2014). Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model Dev., 7*(4), 1297-1333. doi: 10.5194/gmd-7-1297-2014

Krasting, J. P., Broccoli, A. J., Dixon, K. W., & Lanzante, J. R. (2013). Future changes in northern hemisphere snowfall. *Journal of Climate, 26*(20), 7813-7828. doi: 10.1175/jcli-d-12-00832.1

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., . . . Sherstyukov, B. (2014). Flood risk and climate change: global and regional perspectives. *Hydrological Sciences Journal, 59*(1), 1-28. doi: 10.1080/02626667.2013.857411

Kundzewicz, Z. W., Mata, L. J., Arnell, N. W., Doll, P., Kabat, P., Jimenez, B., . . . Shiklomanov, I. (2007). Freshwater resources and their management. Climate change 2007: Impacts, adaptation and vulnerability. Contribution of Working Group II 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, 172-210.

Kunkel, K. E., Karl, T. R., Brooks, H., Kossin, J., Lawrimore, J. H., Arndt, D., . . . Wuebbles, D. (2013). Monitoring and understanding trends in extreme storms: State of knowledge. *Bulletin of the American Meteorological Society, 94*(4), 499-514. doi: 10.1175/bams-d-11-00262.1

Lamarque, J.-F., Kyle, G. P., Meinshausen, M., Riahi, K., Smith, S. J., van Vuuren, D. P., . . . Vitt, F. (2011). Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways. *Climatic Change, 109*(1), 191. doi: 10.1007/s10584-011-0155-0

Langlois, A., Bergeron, J., Brown, R., Royer, A., Harvey, R., Roy, A., . . . Thériault, N. (2014). Evaluation of CLASS 2.7 and 3.5 simulations of snow properties from the Canadian Regional Climate Model (CRCM4) over Québec, Canada. *Journal of Hydrometeorology, 15*(4), 1325-1343. doi: 10.1175/jhm-d-13-055.1

Langlois, A., Royer, A., Fillol, E., Frigon, A., & Laprise, R. (2004). Evaluation of the snow cover variation in the Canadian Regional Climate Model over eastern Canada using passive microwave satellite data. *Hydrological Processes, 18*(6), 1127-1138. doi: 10.1002/hyp.5514

Laprise, R., Caya, D., Frigon, A., & Paquin, D. (2003). Current and perturbed climate as simulated by the second-generation Canadian Regional Climate Model (CRCM-II) over northwestern North America. *Climate Dynamics, 21*(5), 405-421. doi: 10.1007/s00382-003-0342-4

Larabi, S., St-Hilaire, A., & Chebana, F. (2018). A new concept to calibrate and evaluate a hydrological model based on functional data analysis. *Journal of Water Management Modeling*. doi: 10.14796/JWMM.C442

Leahy, P. G., & Kiely, G. (2011). Short duration rainfall extremes in Ireland: Influence of climatic variability. *Water Resources Management, 25*(3), 987-1003. doi: 10.1007/s11269-010-9737-2

Leathers, D. J., Yarnal, B., & Palecki, M. A. (1991). The Pacific/North American teleconnection pattern and United States climate. Part I: Regional temperature and precipitation associations. *Journal of Climate, 4*(5), 517-528. doi: 10.1175/1520-0442(1991)004<0517:Tpatpa>2.0.Co;2

Leduc, M., Giguère, M., Ludwig, R., Frigon, A., Braun, M., Schmid, J., . . . Komischke, H. (2016a). The ClimEx Project: Dynamical downscaling of a GCM large ensemble at very high resolution for Bavaria and Quebec.

Leduc, M., & Laprise, R. (2009). Regional climate model sensitivity to domain size. *Climate Dynamics, 32*(6), 833-854. doi: 10.1007/s00382-008-0400-z

Leduc, M., Laprise, R., de Elía, R., & Šeparović, L. (2016b). Is institutional democracy a good proxy for model independence? *Journal of Climate, 29*(23), 8301-8316. doi: 10.1175/jcli-d-15-0761.1

Leduc, M., Laprise, R., Moretti-Poisson, M., & Morin, J.-P. (2011). Sensitivity to domain size of mid-latitude summer simulations with a regional climate model. *Climate Dynamics, 37*(1), 343-356. doi: 10.1007/s00382-011-1008-2

Leduc, M., Mailhot, A., Frigon, A., Martel, J.-L., Ludwig, R., Brissette, F., . . . Scinocca, J. (2019). 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*. doi: 10.1175/JAMC-D-18-0021.1

Leduc, M., Matthews, H. D., & de Elía, R. (2016c). Regional estimates of the transient climate response to cumulative CO2 emissions. *Nature Climate Change, 6*, 474. doi: 10.1038/nclimate2913

Lee, J.-W., Hong, S.-Y., Chang, E.-C., Suh, M.-S., & Kang, H.-S. (2014). Assessment of future climate change over East Asia due to the RCP scenarios downscaled by GRIMs-RMP. *Climate Dynamics, 42*(3), 733-747. doi: 10.1007/s00382-013-1841-6

Lee, T., & Singh, V. P. (2018). *Statistical downscaling for hydrological and environmental applications*. CRC Press.

Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation. *Water Resources Research, 35*(1), 233-241. doi: 10.1029/1998WR900018

Lenderink, G., & Fowler, H. J. (2017). Understanding rainfall extremes. *Nature Climate Change, 7*, 391. doi: 10.1038/nclimate3305

Lenderink, G., Mok, H. Y., Lee, T. C., & van Oldenborgh, G. J. (2011). Scaling and trends of hourly precipitation extremes in two different climate zones – Hong Kong and the Netherlands. *Hydrol. Earth Syst. Sci., 15*(9), 3033-3041. doi: 10.5194/hess-15-3033-2011

Levine, A. F. Z., McPhaden, M. J., & Frierson, D. M. W. (2017). The impact of the AMO on multidecadal ENSO variability. *Geophysical Research Letters, 44*(8), 3877-3886. doi: 10.1002/2017GL072524

Li, D., Wrzesien, M. L., Durand, M., Adam, J., & Lettenmaier, D. P. (2017). How much runoff originates as snow in the western United States, and how will that change in the future? *Geophysical Research Letters, 44*(12), 6163-6172. doi: 10.1002/2017GL073551

Lim, Y.-K., & Schubert, S. D. (2011). The impact of ENSO and the Arctic Oscillation on winter temperature extremes in the southeast United States. *Geophysical Research Letters, 38*(15), L15706. doi: 10.1029/2011GL048283

Lindenschmidt, K.-E., Fleischbein, K., & Baborowski, M. (2007). Structural uncertainty in a river water quality modelling system. *Ecological Modelling, 204*(3), 289-300. doi: 10.1016/j.ecolmodel.2007.01.004

Lins, H. F., & Slack, J. R. (1999). Streamflow trends in the United States. *Geophysical Research Letters, 26*(2), 227-230. doi: 10.1029/1998GL900291

Liu, D., Guo, S., Wang, Z., Liu, P., Yu, X., Zhao, Q., & Zou, H. (2018). Statistics for sample splitting for the calibration and validation of hydrological models. *Stochastic Environmental Research and Risk Assessment*. doi: 10.1007/s00477-018-1539-8

Liu, X., & Chen, B. (2000). Climatic warming in the Tibetan Plateau during recent decades. *International Journal of Climatology, 20*(14), 1729-1742. doi: 10.1002/1097-0088(20001130)20:14<1729::AID-JOC556>3.0.CO;2-Y

Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research, 43*(7). doi: 10.1029/2006WR005756

Liu, Z., Jian, Z., Yoshimura, K., Buenning, N. H., Poulsen, C. J., & Bowen, G. J. (2015). Recent contrasting winter temperature changes over North America linked to enhanced positive Pacific-North American pattern. *Geophysical Research Letters, 42*(18), 7750-7757. doi: 10.1002/2015GL065656

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences, 20*(2), 130-141. doi: 10.1175/1520-0469(1963)020<0130:Dnf>2.0.Co;2

Lucas-Picher, P., Laprise, R., & Winger, K. (2017). Evidence of added value in North American regional climate model hindcast simulations using ever-increasing horizontal resolutions. *Climate Dynamics, 48*(7), 2611-2633. doi: 10.1007/s00382-016-3227-z

Lucas-Picher, P., Riboust, P., Somot, S., & Laprise, R. (2015). Reconstruction of the spring 2011 Richelieu River flood by two regional climate models and a hydrological model. *Journal of Hydrometeorology, 16*(1), 36-54. doi: 10.1175/jhm-d-14-0116.1

Luo, Z., Jiang, Z., & Tang, S. (2015). Impacts of climate change on distributions and diversity of ungulates on the Tibetan Plateau. *Ecological Applications, 25*(1), 24-38. doi: 10.1890/13-1499.1

Lutz, A. F., Immerzeel, W. W., Gobiet, A., Pellicciotti, F., & Bierkens, M. F. P. (2013). Comparison of climate change signals in CMIP3 and CMIP5 multi-model ensembles

and implications for Central Asian glaciers. *Hydrol. Earth Syst. Sci., 17*(9), 3661-3677. doi: 10.5194/hess-17-3661-2013

Machguth, H., Paul, F., Hoelzle, M., & Haeberli, W. (2006). Distributed glacier mass-balance modelling as an important component of modern multi-level glacier monitoring. *Annals of Glaciology, 43*, 335-343. doi: 10.3189/172756406781812285

Madsen, H., Arnbjerg-Nielsen, K., & Mikkelsen, P. S. (2009). Update of regional intensity–duration–frequency curves in Denmark: Tendency towards increased storm intensities. *Atmospheric Research, 92*(3), 343-349. doi: 10.1016/j.atmosres.2009.01.013

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., & Kjeldsen, T. R. (2014). Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. *Journal of Hydrology, 519*, 3634-3650. doi: 10.1016/j.jhydrol.2014.11.003

Mahlstein, I., Knutti, R., Solomon, S., & Portmann, R. W. (2011). Early onset of significant local warming in low latitude countries. *Environmental Research Letters, 6*(3), 034009. doi: 10.1088/1748-9326/6/3/034009

Mailhot, A., Beauregard, I., Talbot, G., Caya, D., & Biner, S. (2012). Future changes in intense precipitation over Canada assessed from multi-model NARCCAP ensemble simulations. *International Journal of Climatology, 32*(8), 1151-1163. doi: 10.1002/joc.2343

Mailhot, A., & Duchesne, S. (2010). Design criteria of urban drainage infrastructures under climate change. *Journal of Water Resources Planning and Management, 136*(2), 201-208. doi: 10.1061/(ASCE)WR.1943-5452.0000023

Mailhot, A., Duchesne, S., Caya, D., & Talbot, G. (2007). Assessment of future change in intensity–duration–frequency (IDF) curves for Southern Quebec using the Canadian Regional Climate Model (CRCM). *Journal of Hydrology, 347*(1–2), 197-210. doi: 10.1016/j.jhydrol.2007.09.019

Mantua, N. J., & Hare, S. R. (2002). The Pacific Decadal Oscillation. *Journal of Oceanography, 58*(1), 35-44. doi: 10.1023/a:1015820616384

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., & Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society, 78*(6), 1069-1079. doi: 10.1175/1520-0477(1997)078<1069:Apicow>2.0.Co;2

Maraun, D. (2012). Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. *Geophysical Research Letters, 39*(6). doi: 10.1029/2012GL051210

Maraun, D. (2013a). Bias correction, quantile mapping, and downscaling: revisiting the inflation issue. *Journal of Climate, 26*(6), 2137-2143. doi: 10.1175/jcli-d-12-00821.1

Maraun, D. (2013b). When will trends in European mean and heavy daily precipitation emerge? *Environmental Research Letters, 8*(1). doi: 10.1088/1748-9326/8/1/014004

Maraun, D. (2016). Bias correcting climate change simulations - a critical review. *Current Climate Change Reports, 2*(4), 211-220. doi: 10.1007/s40641-016-0050-x

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., . . . Mearns, L. O. (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change, 7*, 764. doi: 10.1038/nclimate3418

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., . . . Thiele-Eich, I. (2010a). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of geophysics, 48*(3), RG3003. doi: 10.1029/2009RG000314

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., . . . Thiele-Eich, I. (2010b). Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of geophysics, 48*(3). doi: doi:10.1029/2009RG000314

Maraun, D., & Widmann, M. (2018). *Statistical downscaling and bias correction for climate research*. Cambridge University Press.

Martel, J.-L., Demeester, K., Brissette, F., Poulin, A., & Arsenault, R. (2017). HMETS—A simple and efficient hydrology model for teaching hydrological modelling, flow forecasting and climate change impacts. *International Journal of Engineering Education, 33*(4), 1307-1316.

Martel, J.-L., Mailhot, A., Brissette, F., & Caya, D. (2018). Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales. *Journal of Climate, 31*(11), 4241-4263. doi: 10.1175/jcli-d-17-0282.1

Martynov, A., Laprise, R., Sushama, L., Winger, K., Šeparović, L., & Dugas, B. (2013). Reanalysis-driven climate simulation over CORDEX North America domain using

the Canadian Regional Climate Model, version 5: model performance evaluation. *Climate Dynamics, 41*(11), 2973-3005. doi: 10.1007/s00382-013-1778-9

Massei, N., Laignel, B., Deloffre, J., Mesquita, J., Motelay, A., Lafite, R., & Durand, A. (2010). Long-term hydrological changes of the Seine River flow (France) and their relation to the North Atlantic Oscillation over the period 1950–2008. *International Journal of Climatology, 30*(14), 2146-2154. doi: 10.1002/joc.2022

Massei, N., Laignel, B., Rosero, E., Motelay-massei, A., Deloffre, J., Yang, Z. L., & Rossi, A. (2011). A wavelet approach to the short-term to pluri-decennal variability of streamflow in the Mississippi River basin from 1934 to 1998. *International Journal of Climatology, 31*(1), 31-43. doi: 10.1002/joc.1995

Matonse, A. H., Pierson, D. C., Frei, A., Zion, M. S., Anandhi, A., Schneiderman, E., & Wright, B. (2013). Investigating the impact of climate change on New York City's primary water supply. *Climatic Change, 116*(3), 437-456. doi: 10.1007/s10584-012-0515-4

Matte, D., Laprise, R., Thériault, J. M., & Lucas-Picher, P. (2017). Spatial spin-up of fine scales in a regional climate model simulation driven by low-resolution boundary conditions. *Climate Dynamics, 49*(1), 563-574. doi: 10.1007/s00382-016-3358-2

Maurer, E. P., Gibbard, S., & Duffy, P. B. (2006). Amplification of streamflow impacts of El Niño by increased atmospheric greenhouse gases. *Geophysical Research Letters, 33*(2), L02707. doi: 10.1029/2005GL025100

Maurer, E. P., Lettenmaier, D. P., & Mantua, N. J. (2004). Variability and potential sources of predictability of North American runoff. *Water Resources Research, 40*(9), W09306. doi: 10.1029/2003WR002789

Maurer, E. P., Wood, A. W., Adam, J. C., Lettenmaier, D. P., & Nijssen, B. (2002). A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *Journal of Climate, 15*(22), 3237-3251. doi: 10.1175/1520-0442(2002)015<3237:Althbd>2.0.Co;2

Mazouz, R., Assani, A. A., Quessy, J.-F., & Légaré, G. (2012). Comparison of the interannual variability of spring heavy floods characteristics of tributaries of the St. Lawrence River in Quebec (Canada). *Advances in Water Resources, 35*, 110-120. doi: 10.1016/j.advwatres.2011.10.006

McCabe, G. J. (1995). Relations between winter atmospheric circulation and annual streamflow in the western United States. *Climate Research, 5*(2), 139-148. doi: www.jstor.org/stable/24863412

McCabe, G. J., & Wolock, D. M. (2002). A step increase in streamflow in the conterminous United States. *Geophysical Research Letters, 29*(24), 38-31-38-34. doi: 10.1029/2002GL015999

McCuen, R. H., Knight, Z., & Cutter, A. G. (2006). Evaluation of the Nash-Sutcliffe efficiency index. *Journal of Hydrologic Engineering, 11*(6), 597-602. doi: 10.1061/(ASCE)1084-0699(2006)11:6(597)

McGuffie, K., & Henderson-Sellers, A. (2005). *A climate modelling primer*. John Wiley & Sons.

Mearns, L. O., Arritt, R., Biner, S., Bukovsky, M. S., McGinnis, S., Sain, S., . . . Snyder, M. (2012). The North American regional climate change assessment program: Overview of phase I results. *Bulletin of the American Meteorological Society, 93*(9), 1337-1362. doi: 10.1175/bams-d-11-00223.1

Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., . . . Mitchell, J. F. B. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society, 88*(9), 1383-1394. doi: 10.1175/bams-88-9-1383

Meehl, G. A., & Tebaldi, C. (2004). More intense, more frequent, and longer lasting heat waves in the 21st century. *Science, 305*(5686), 994-997. doi: 10.1126/science.1098704

Mehrotra, R., Sharma, A., Bari, M., Tuteja, N., & Amirthanathan, G. (2014). An assessment of CMIP5 multi-model decadal hindcasts over Australia from a hydrological viewpoint. *Journal of Hydrology, 519*, 2932-2951. doi: 10.1016/j.jhydrol.2014.07.053

Mehta, V. M., Wang, H., Mendoza, K., & Rosenberg, N. J. (2014). Predictability and prediction of decadal hydrologic cycles: A case study in Southern Africa. *Weather and Climate Extremes, 3*, 47-53. doi: 10.1016/j.wace.2014.04.002

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., . . . van Vuuren, D. P. P. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change, 109*(1), 213. doi: 10.1007/s10584-011-0156-z

Mendoza, P. A., Clark, M. P., Mizukami, N., Newman, A. J., Barlage, M., Gutmann, E. D., . . . Arnold, J. R. (2015). Effects of hydrologic model choice and calibration on the portrayal of climate change impacts. *Journal of Hydrometeorology, 16*(2), 762-780. doi: 10.1175/jhm-d-14-0104.1

Meylan, P., Favre, A.-C., & Musy, A. (2008). *Hydrologie fréquentielle: une science prédictive*. PPUR presses polytechniques.

Middelkoop, H., Daamen, K., Gellens, D., Grabs, W., Kwadijk, J. C. J., Lang, H., . . . Wilke, K. (2001). Impact of climate change on hydrological regimes and water resources management in the Rhine Basin. *Climatic Change, 49*(1), 105-128. doi: 10.1023/a:1010784727448

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008a). Stationarity Is Dead: Whither Water Management? *Science, 319*(5863), 573-574. doi: 10.1126/science.1151915

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008b). Stationarity is dead: Whither water management? *Science, 319*(5863), 573. doi: 10.1126/science.1151915

Min, S.-K., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. *Nature, 470*(7334), 378-381. doi: 10.1038/nature09763

Ministère du Développement durable, d. l. E. e. d. l. L. c. l. c. c. M. (2017). Manuel de calcul et de conception des ouvrages municipaux de gestion des eaux pluviales. 125 pp.

Minville, M., Brissette, F., & Leconte, R. (2008). Uncertainty of the impact of climate change on the hydrology of a nordic watershed. *Journal of Hydrology, 358*(1), 70-83. doi: 10.1016/j.jhydrol.2008.05.033

Minville, M., Cartier, D., Guay, C., Leclaire, L. A., Audet, C., Le Digabel, S., & Merleau, J. (2014). Improving process representation in conceptual hydrological model calibration using climate simulations. *Water Resources Research, 50*(6), 5044-5073. doi: 10.1002/2013WR013857

Mizuta, R., Murata, A., Ishii, M., Shiogama, H., Hibino, K., Mori, N., . . . Kimoto, M. (2017). Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bulletin of the American Meteorological Society, 98*(7), 1383-1398. doi: 10.1175/bams-d-16-0099.1

Mladjic, B., Sushama, L., Khaliq, M. N., Laprise, R., Caya, D., & Roy, R. (2011). Canadian RCM projected changes to extreme precipitation characteristics over Canada. *Journal of Climate, 24*(10), 2565-2584. doi: 10.1175/2010jcli3937.1

Mo, K. C., Schemm, J.-K. E., & Yoo, S.-H. (2009). Influence of ENSO and the Atlantic Multidecadal Oscillation on drought over the United States. *Journal of Climate, 22*(22), 5962-5982. doi: 10.1175/2009jcli2966.1

Monteith, J. L. (1965). Evaporation and the environment. The State and Movement of Water in Living Organisms, G. E. Fogg, Ed., Symposia of the Society for Experimental Biology, No. 19, Cambridge University Press, 205–234.

Mooney, P. A., Broderick, C., Bruyère, C. L., Mulligan, F. J., & Prein, A. F. (2017). Clustering of observed diurnal cycles of precipitation over the United States for evaluation of a WRF multiphysics regional climate ensemble. *Journal of Climate, 30*(22), 9267-9286. doi: 10.1175/jcli-d-16-0851.1

Mora, C., Frazier, A. G., Longman, R. J., Dacks, R. S., Walton, M. M., Tong, E. J., . . . Giambelluca, T. W. (2013). The projected timing of climate departure from recent variability. *Nature, 502*(7470), 183-187. doi: 10.1038/nature12540

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE, 50*(3), 885. doi: 10.13031/2013.23153

Moriasi, D. N., Zeckoski, R. W., Arnold, J. G., Baffaut, C., Malone, R. W., Daggupati, P., . . . Douglas-Mankin, K. R. (2015). Hydrologic and water quality models: key calibration and validation topics. *Transactions of the ASABE, 58*(6), 1609. doi: 10.13031/trans.58.11075

Mpelasoka, F. S., & Chiew, F. H. S. (2009). Influence of rainfall scenario construction methods on runoff projections. *Journal of Hydrometeorology, 10*(5), 1168-1183. doi: 10.1175/2009jhm1045.1

Mudryk, L. R., Kushner, P. J., & Derksen, C. (2014). Interpreting observed northern hemisphere snow trends with large ensembles of climate simulations. *Climate Dynamics, 43*(1), 345-359. doi: 10.1007/s00382-013-1954-y

Muerth, M., Gauvin St-Denis, B., Ricard, S., Velázquez, J., Schmid, J., Minville, M., . . . Turcotte, R. (2012a). On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff. *Hydrology and Earth System Sciences Discussions*, 10205-10243.

Muerth, M., Gauvin St-Denis, B., Ricard, S., Velázquez, J. A., Schmid, J., Minville, M., . . . Turcotte, R. (2012b). On the need for bias correction in regional climate scenarios to

assess climate change impacts on river runoff. *Hydrology and Earth System Sciences Discussions*, 10205-10243. doi: 10.5194/hessd-9-10205-2012

Mullan, D., Swindles, G., Patterson, T., Galloway, J., Macumber, A., Falck, H., . . . Pisaric, M. (2017). Climate change and the long-term viability of the World's busiest heavy haul ice road. *Theoretical and Applied Climatology, 129*(3), 1089-1108. doi: 10.1007/s00704-016-1830-x

Muschinski, T., & Katz, J. I. (2013). Trends in hourly rainfall statistics in the United States under a warming climate. *Nature Climate Change, 3*, 577. doi: 10.1038/nclimate1828

Music, B., & Caya, D. (2007). Evaluation of the hydrological cycle over the Mississippi River basin as simulated by the Canadian Regional Climate Model (CRCM). *Journal of Hydrometeorology, 8*(5), 969-988. doi: 10.1175/jhm627.1

Music, B., Caya, D., Frigon, A., Musy, A., Roy, R., & Rodenhuis, D. (2012). Canadian Regional Climate Model as a tool for assessing hydrological impacts of climate change at the watershed scale. Dans (pp. 157-165). Springer Vienna. doi: 10.1007/978-3-7091-0973-1_12

Musy, A., Hingray, B., & Picouet, C. (2014). *Hydrology: a science for engineers*. CRC Press.

Nahar, J., Johnson, F., & Sharma, A. (2017). Assessing the extent of non-stationary biases in GCMs. *Journal of Hydrology, 549*, 148-162. doi: 10.1016/j.jhydrol.2017.03.045

Najafi, M. R., Moradkhani, H., & Jung, I. W. (2011). Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrological Processes, 25*(18), 2814-2826. doi: 10.1002/hyp.8043

Nalley, D., Adamowski, J., Khalil, B., & Biswas, A. (2016). Inter-annual to inter-decadal streamflow variability in Quebec and Ontario in relation to dominant large-scale climate indices. *Journal of Hydrology, 536*, 426-446. doi: 10.1016/j.jhydrol.2016.02.049

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology, 10*(3), 282-290. doi: 10.1016/0022-1694(70)90255-6

National Research Council. (1998). *Decade-to-century-scale climate variability and change: a science strategy*. National Academies Press.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., . . . Arnold, J. R. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences, 19*(1), 209. doi: 10.5194/hess-19-209-2015

Nicholls, R. J., & Cazenave, A. (2010). Sea-level rise and its impact on coastal zones. *Science, 328*(5985), 1517-1520. doi: 10.1126/science.1185782

Nikiéma, O., Laprise, R., & Dugas, B. (2017). Energetics of transient-eddy and inter-member variabilities in global and regional climate model simulations. *Climate Dynamics*, 1-20. doi: 10.1007/s00382-017-3918-0

Notaro, M., Zarrin, A., Vavrus, S., & Bennington, V. (2013). Simulation of heavy lake-effect snowstorms across the Great Lakes basin by RegCM4: Synoptic climatology and variability. *Monthly Weather Review, 141*(6), 1990-2014. doi: 10.1175/mwr-d-11-00369.1

Nourani, V., Hosseini Baghanam, A., Adamowski, J., & Kisi, O. (2014). Applications of hybrid wavelet–Artificial Intelligence models in hydrology: A review. *Journal of Hydrology, 514*, 358-377. doi: 10.1016/j.jhydrol.2014.03.057

Ntegeka, V., & Willems, P. (2008). Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resources Research, 44*(7), W07402. doi: 10.1029/2007WR006471

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., . . . Lowe, J. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development, 9*(9), 3461. doi: 10.5194/gmd-2016-84

Ogden, F. L., & Julien, P. Y. (1993). Runoff sensitivity to temporal and spatial rainfall variability at runoff plane and small basin scales. *Water Resources Research, 29*(8), 2589-2597. doi: 10.1029/93WR00924

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology, 303*(1), 290-306. doi: 10.1016/j.jhydrol.2004.08.026

Paniconi, C., & Putti, M. (2015). Physically based modeling in catchment hydrology at 50: Survey and outlook. *Water Resources Research, 51*(9), 7090-7129. doi: 10.1002/2015WR017780

Papineau, J. M. (2001). Wintertime temperature anomalies in Alaska correlated with ENSO and PDO. *International Journal of Climatology, 21*(13), 1577-1592. doi: 10.1002/joc.686

Pechlivanidis, I. G., Jackson, B. M., McIntyre, N. R., & Wheater, H. S. (2011). Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST journal, 13*(3), 193-214.

Peixoto, J. P., & Oort, A. H. (1992). Physics of climate. 520 pp.

Pellicciotti, F., Brock, B., Strasser, U., Burlando, P., Funk, M., & Corripio, J. (2005). An enhanced temperature-index glacier melt model including the shortwave radiation balance: development and testing for Haut Glacier d'Arolla, Switzerland. *Journal of Glaciology, 51*(175), 573-587. doi: 10.3189/172756505781829124

Pendergrass, A. G. (2018). What precipitation is extreme? *Science, 360*(6393), 1072-1073. doi: 10.1126/science.aat1871

Pendergrass, A. G., Lehner, F., Sanderson, B. M., & Xu, Y. (2015). Does extreme precipitation intensity depend on the emissions scenario? *Geophysical Research Letters, 42*(20), 8767-8774. doi: 10.1002/2015GL065854

Pepin, N., Bradley, R. S., Diaz, H. F., Baraer, M., Caceres, E. B., Forsythe, N., . . . Yang, D. Q. (2015). Elevation-dependent warming in mountain regions of the world. *Nature Climate Change, 5*, 424. doi: 10.1038/nclimate2563

Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology, 279*(1), 275-289. doi: 10.1016/S0022-1694(03)00225-7

Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna, C., Michel, C., & Mathevet, T. (2007). Impact of limited streamflow data on the efficiency and the parameters of rainfall—runoff models. *Hydrological Sciences Journal, 52*(1), 131-151. doi: 10.1623/hysj.52.1.131

Peters, D. L., Atkinson, D., Monk, W. A., Tenenbaum, D. E., & Baird, D. J. (2013). A multi-scale hydroclimatic analysis of runoff generation in the Athabasca River, western Canada. *Hydrological Processes, 27*(13), 1915-1934. doi: 10.1002/hyp.9699

Piani, C., Haerter, J. O., & Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology, 99*(1), 187-192. doi: 10.1007/s00704-009-0134-9

Piechota, T. C., Dracup, J. A., & Fovell, R. G. (1997). Western US streamflow and atmospheric circulation patterns during El Niño-Southern Oscillation. *Journal of Hydrology, 201*(1), 249-271. doi: 10.1016/S0022-1694(97)00043-7

Plummer, D. A., Caya, D., Frigon, A., Côté, H., Giguère, M., Paquin, D., . . . de Elia, R. (2006). Climate and climate change over North America as simulated by the Canadian RCM. *Journal of Climate, 19*(13), 3112-3132. doi: 10.1175/jcli3769.1

Poissant, D., Arsenault, R., & Brissette, F. (2017). Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments. *Journal of Hydrology: Regional Studies, 12*, 220-237. doi: 10.1016/j.ejrh.2017.05.005

Pokhrel, P., & Gupta, H. V. (2011). On the ability to infer spatial catchment variability using streamflow hydrographs. *Water Resources Research, 47*(8), n/a-n/a. doi: 10.1029/2010WR009873

Poulin, A., Brissette, F., Leconte, R., Arsenault, R., & Malo, J.-S. (2011). Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. *Journal of Hydrology, 409*(3), 626-636. doi: 10.1016/j.jhydrol.2011.08.057

Prein, A. F., Holland, G. J., Rasmussen, R. M., Done, J., Ikeda, K., Clark, M. P., & Liu, C. H. (2013). Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado headwaters. *Journal of Climate, 26*(13), 4848-4857. doi: 10.1175/jcli-d-12-00727.1

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., . . . Feser, F. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of geophysics, 53*(2), 323-361. doi: 10.1002/2014RG000475

Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Clim. Change, 7*(1), 48-52. doi: 10.1038/nclimate3168

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review, 100*(2), 81-92. doi: 10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Prudhomme, C., & Davies, H. (2009a). Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate. *Climatic Change, 93*(1), 177-195. doi: 10.1007/s10584-008-9464-3

Prudhomme, C., & Davies, H. (2009b). Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate. *Climatic Change, 93*(1), 197-222. doi: 10.1007/s10584-008-9461-6

Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., . . . Midgley, P. M. (2014). Climate change 2013: The physical science basis. *Contribution of working group I to the Fifth Assessment Report of the Intergovernmental Panel on climate change*, 5-14.

Raftery, A. E., Zimmer, A., Frierson, D. M. W., Startz, R., & Liu, P. (2017). Less than 2 °C warming by 2100 unlikely. *Nature Climate Change, 7*, 637. doi: 10.1038/nclimate3352

Räisänen, J. (2002). CO2-induced changes in interannual temperature and precipitation variability in 19 CMIP2 experiments. *Journal of Climate, 15*(17), 2395-2411. doi: 10.1175/1520-0442(2002)015<2395:Ciciit>2.0.Co;2

Rajczak, J., & Schär, C. (2017). Projections of future precipitation extremes over Europe: A multimodel assessment of climate simulations. *Journal of Geophysical Research: Atmospheres, 122*(20), 10,773-710,800. doi: 10.1002/2017JD027176

Rasmussen, R., Liu, C., Ikeda, K., Gochis, D., Yates, D., Chen, F., . . . Gutmann, E. (2011). High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *Journal of Climate, 24*(12), 3015-3048. doi: 10.1175/2010jcli3985.1

Rathinasamy, M., Khosa, R., Adamowski, J., ch, S., Partheepan, G., Anand, J., & Narsimlu, B. (2014). Wavelet-based multiscale performance analysis: An approach to assess and improve hydrological models. *Water Resources Research, 50*(12), 9721-9737. doi: 10.1002/2013WR014650

Razavi, S., & Tolson, B. A. (2013). An efficient framework for hydrologic model calibration on long data periods. *Water Resources Research, 49*(12), 8418-8431. doi: 10.1002/2012WR013442

Redmond, K. T., & Cayan, D. R. (1994). El Niño/Southern Oscillation and western climate variability. Dans *Preprints, Sixth Conf. on Climate Variations, Nashville, TN, Amer. Meteor. Soc* (pp. 141-145).

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., & DMIP Participants, a. (2004). Overall distributed model intercomparison project results. *Journal of Hydrology, 298*(1), 27-60. doi: 10.1016/j.jhydrol.2004.03.031

Refsgaard, J. C., Henriksen, H. J., Harrar, W. G., Scholten, H., & Kassahun, A. (2005). Quality assurance in model based water management – review of existing practice and outline of new approaches. *Environmental Modelling & Software, 20*(10), 1201-1215. doi: 10.1016/j.envsoft.2004.07.006

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research, 46*(5). doi: 10.1029/2009WR008328

Renard, B., & Lall, U. (2014). Regional frequency analysis conditioned on large-scale atmospheric or oceanic fields. *Water Resources Research, 50*(12), 9536-9554. doi: 10.1002/2014wr016277

Renard, B., Lang, M., Bois, P., Dupeyrat, A., Mestre, O., Niel, H., . . . Gailhard, J. (2008). Regional methods for trend detection: Assessing field significance and regional consistency. *Water Resources Research, 44*(8). doi: 10.1029/2007wr006268

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., . . . Rafaj, P. (2011). RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Climatic Change, 109*(1), 33. doi: 10.1007/s10584-011-0149-y

Rice, J. S., Emanuel, R. E., Vose, J. M., & Nelson, S. A. C. (2015). Continental U.S. streamflow trends from 1940 to 2009 and their relationships with watershed spatial characteristics. *Water Resources Research, 51*(8), 6262-6275. doi: 10.1002/2014WR016367

Riette, S., & Caya, D. (2002). Sensitivity of short simulations to the various parameters in the new CRCM spectral nudging. - In: RITCHIE, H. (Ed.): Research activities in Atmospheric

and Oceanic Modeling, WMO/TD No. 1105, Report No. 32: 7.39-7.40. Tech. rep., (1105), 7.39-37.40.

Rizzi, J., Torresan, S., Zabeo, A., Critto, A., Tosoni, A., Tomasin, A., & Marcomini, A. (2017). Assessing storm surge risk under future sea-level rise scenarios: a case study in the North Adriatic coast. *Journal of Coastal Conservation, 21*(4), 453-471. doi: 10.1007/s11852-017-0517-5

Roberts, J., & Snelgrove, K. (2015). Uncertainty in regional climate model mean runoff projections under climate change: Case study of Labrador's Churchill River basin. *Atmosphere-Ocean, 53*(3), 319-331. doi: 10.1080/07055900.2015.1026872

Rogers, J. C., & Coleman, J. S. M. (2003). Interactions between the Atlantic Multidecadal Oscillation, El Niño/La Niña, and the PNA in winter Mississippi Valley stream flow. *Geophysical Research Letters, 30*(10), 1518. doi: 10.1029/2003GL017216

Rojas, R., Feyen, L., Bianchi, A., & Dosio, A. (2012). Assessment of future flood hazard in Europe using a large ensemble of bias-corrected regional climate simulations. *Journal of Geophysical Research: Atmospheres, 117*(D17). doi: 10.1029/2012JD017461

Rojas, R., Feyen, L., Dosio, A., & Bavera, D. (2011). Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations. *Hydrology and Earth System Sciences, 15*(8), 2599. doi: 10.5194/hess-15-2599-2011

Rood, S. B., Samuelson, G. M., Weber, J. K., & Wywrot, K. A. (2005). Twentieth-century decline in streamflows from the hydrographic apex of North America. *Journal of Hydrology, 306*(1), 215-233. doi: 10.1016/j.jhydrol.2004.09.010

Ropelewski, C. F., & Halpert, M. S. (1986). North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Monthly Weather Review, 114*(12), 2352-2362. doi: 10.1175/1520-0493(1986)114<2352:Napatp>2.0.Co;2

Ropelewski, C. F., & Halpert, M. S. (1996). Quantifying Southern Oscillation-precipitation relationships. *Journal of Climate, 9*(5), 1043-1059. doi: 10.1175/1520-0442(1996)009<1043:Qsopr>2.0.Co;2

Rossi, A., Massei, N., & Laignel, B. (2011). A synthesis of the time-scale variability of commonly used climate indices using continuous wavelet transform. *Global and Planetary Change, 78*(1), 1-13. doi: 10.1016/j.gloplacha.2011.04.008

Sanderson, B. M., Oleson, K. W., Strand, W. G., Lehner, F., & O'Neill, B. C. (2018). A new ensemble of GCM simulations to assess avoided impacts in a climate mitigation scenario. *Climatic Change, 146*(3), 303-318. doi: 10.1007/s10584-015-1567-z

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., & Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature, 427*, 332. doi: 10.1038/nature02300

Schmidli, J., Frei, C., & Vidale, P. L. (2006). Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology, 26*(5), 679-689. doi: 10.1002/joc.1287

Schmidt, N., Lipp, E. K., Rose, J. B., & Luther, M. E. (2001). ENSO influences on seasonal rainfall and river discharge in Florida. *Journal of Climate, 14*(4), 615-628. doi: 10.1175/1520-0442(2001)014<0615:Eiosra>2.0.Co;2

Schulte, J. A., Najjar, R. G., & Li, M. (2016). The influence of climate modes on streamflow in the Mid-Atlantic region of the United States. *Journal of Hydrology: Regional Studies, 5*, 80-99. doi: 10.1016/j.ejrh.2015.11.003

Schulz, K., & Bernhardt, M. (2016). The end of trend estimation for extreme floods under climate change? *Hydrological Processes, 30*(11), 1804-1808. doi: 10.1002/hyp.10816

Scoccimarro, E., Gualdi, S., Bellucci, A., Zampieri, M., & Navarra, A. (2013). Heavy precipitation events in a warmer climate: Results from CMIP5 models. *Journal of Climate, 26*(20), 7902-7911. doi: 10.1175/jcli-d-12-00850.1

Screen, J. A. (2014). Arctic amplification decreases temperature variance in northern mid- to high-latitudes. *Nature Climate Change, 4*, 577. doi: 10.1038/nclimate2268

Seaby, L. P., Refsgaard, J. C., Sonnenborg, T. O., & Højberg, A. L. (2015). Spatial uncertainty in bias corrected climate change projections and hydrogeological impacts. *Hydrological Processes, 29*(20), 4514-4532. doi: 10.1002/hyp.10501

Seiller, G., & Anctil, F. (2014). Climate change impacts on the hydrologic regime of a Canadian river: comparing uncertainties arising from climate natural variability and lumped hydrological model structures. *Hydrol. Earth Syst. Sci., 18*(6), 2033-2047. doi: 10.5194/hess-18-2033-2014

Selten, F. M., Branstator, G. W., Dijkstra, H. A., & Kliphuis, M. (2004). Tropical origins for recent and future Northern Hemisphere climate change. *Geophysical Research Letters, 31*(21). doi: 10.1029/2004GL020739

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association, 63*(324), 1379-1389.

Seneviratne, S. I., Lüthi, D., Litschi, M., & Schär, C. (2006). Land–atmosphere coupling and climate change in Europe. *Nature, 443*, 205. doi: 10.1038/nature05095

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., . . . Zhang, X. (2012). Changes in climate extremes and their impacts on the natural

physical environment. In: Managing the risks of extreme events and disasters to advance climate change adaptation, chap. 3, A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change., 109-230.

Šeparović, L., Alexandru, A., Laprise, R., Martynov, A., Sushama, L., Winger, K., . . . Valin, M. (2013). Present climate and climate change over North America as simulated by the fifth-generation Canadian regional climate model. *Climate Dynamics, 41*(11), 3167-3201. doi: 10.1007/s00382-013-1737-5

Separovic, L., de Elía, R., & Laprise, R. (2012). Impact of spectral nudging and domain size in studies of RCM response to parameter modification. *Climate Dynamics, 38*(7), 1325-1343. doi: 10.1007/s00382-011-1072-7

Shabbar, A., & Bonsal, B. (2004). Associations between low frequency variability modes and winter temperature extremes in Canada. *Atmosphere-Ocean, 42*(2), 127-140. doi: 10.3137/ao.420204

Shabbar, A., & Khandekar, M. (1996). The impact of el Nino-Southern oscillation on the temperature field over Canada: Research note. *Atmosphere-Ocean, 34*(2), 401-416. doi: 10.1080/07055900.1996.9649570

Shepard, D. S. (1984). Computer mapping: The SYMAP interpolation algorithm. Dans G. L. Gaile & C. J. Willmott (Éds.), *Spatial Statistics and Models* (pp. 133-145). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-017-3048-8_7

Shrestha, R. R., Cannon, A. J., Schnorbus, M. A., & Zwiers, F. W. (2017). Projecting future nonstationary extreme streamflow for the Fraser River, Canada. *Climatic Change, 145*(3), 289-303. doi: 10.1007/s10584-017-2098-6

Sigmond, M., & Fyfe, J. C. (2016). Tropical Pacific impacts on cooling North American winters. *Nature Clim. Change, 6*(10), 970-974. doi: 10.1038/nclimate3069

Sillmann, J., Kharin, V. V., Zhang, X., Zwiers, F. W., & Bronaugh, D. (2013a). Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres, 118*(4), 1716-1733. doi: 10.1002/jgrd.50203

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., & Bronaugh, D. (2013b). Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmospheres, 118*(6), 2473-2493. doi: 10.1002/jgrd.50188

Singh, S. L., & Bárdossy, A. (2012). Calibration of hydrological models on hydrologically unusual events. *Advances in Water Resources, 38*, 81-91. doi: 10.1016/j.advwatres.2011.12.006

Singh, V. P. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes, 11*(12), 1649-1669. doi: 10.1002/(SICI)1099-1085(19971015)11:12<1649::AID-HYP495>3.0.CO;2-1

Singh, V. P., & Woolhiser, D. A. (2002). Mathematical Modeling of Watershed Hydrology. *Journal of Hydrologic Engineering, 7*(4), 270-292. doi: 10.1061/(ASCE)1084-0699(2002)7:4(270)

Smith, M. B., Koren, V., Zhang, Z., Zhang, Y., Reed, S. M., Cui, Z., . . . Anderson, E. A. (2012). Results of the DMIP 2 Oklahoma experiments. *Journal of Hydrology, 418-419*(Supplement C), 17-48. doi: 10.1016/j.jhydrol.2011.08.056

Solomon, S., Qin, D., Manning, M., Averyt, K., & Marquis, M. (2007). *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC. Vol 4. Cambridge university press.* (Vol. 4). Cambridge university press.

Souza Filho, F. A., & Lall, U. (2003). Seasonal to interannual ensemble streamflow forecasts for Ceara, Brazil: Applications of a multivariate, semiparametric algorithm. *Water Resources Research, 39*(11), n/a-n/a. doi: 10.1029/2002WR001373

Spence, C. (2002). Streamflow variability (1965 to 1998) in five northwest territories and Nunavut rivers. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 27*(2), 135-154. doi: 10.4296/cwrj2702135

St. George, S. (2007). Streamflow in the Winnipeg River basin, Canada: Trends, extremes and climate linkages. *Journal of Hydrology, 332*(3), 396-411. doi: 10.1016/j.jhydrol.2006.07.014

Stahl, K., Moore, R. D., & McKendry, I. G. (2006). The role of synoptic-scale circulation in the linkage between large-scale ocean–atmosphere indices and winter surface climate in British Columbia, Canada. *International Journal of Climatology, 26*(4), 541-560. doi: 10.1002/joc.1268

Sterl, A., Severijns, C., Dijkstra, H., Hazeleger, W., van Oldenborgh, G. J., van den Broeke, M., . . . van Velthoven, P. (2008). When can we expect extremely high surface temperatures? *Geophysical Research Letters, 35*(14). doi: 10.1029/2008GL034071

Stewart, I. T., Cayan, D. R., & Dettinger, M. D. (2005). Changes toward earlier streamflow timing across western North America. *Journal of Climate, 18*(8), 1136-1155. doi: 10.1175/jcli3321.1

Stott, P. A., Jones, G. S., Lowe, J. A., Thorne, P., Durman, C., Johns, T. C., & Thelen, J.-C. (2006). Transient climate simulations with the HadGEM1 climate model: Causes of past warming and future climate change. *Journal of Climate, 19*(12), 2763-2782. doi: 10.1175/jcli3731.1

Stouffer, R. J., & Wetherald, R. T. (2007). Changes of variability in response to increasing greenhouse gases. Part I: Temperature. *Journal of Climate, 20*(21), 5455-5467. doi: 10.1175/2007jcli1384.1

Sudheer, K. P., Lakshmi, G., & Chaubey, I. (2011). Application of a pseudo simulator to evaluate the sensitivity of parameters in complex watershed models. *Environmental Modelling & Software, 26*(2), 135-143. doi: 10.1016/j.envsoft.2010.07.007

Sunyer, M. A., Sørup, H. J. D., Christensen, O. B., Madsen, H., Rosbjerg, D., Mikkelsen, P. S., & Arnbjerg-Nielsen, K. (2013). On the importance of observational data properties when assessing regional climate model performance of extreme precipitation. *Hydrology and Earth System Sciences, 17*(11), 4323. doi: 10.5194/hess-17-4323-2013.

Supharatid, S. (2015). Assessment of CMIP3-CMIP5 climate models precipitation projection and implication of flood vulnerability of Bangkok. *American Journal of Climate Change, 4*(1), 23. doi: 10.4236/ajcc.2015.41011

Swanson, K. L., Sugihara, G., & Tsonis, A. A. (2009). Long-term natural variability and 20th century climate change. *Proceedings of the National Academy of Sciences, 106*(38), 16120-16123. doi: 10.1073/pnas.0908699106

Tan, X., Gan, T. Y., & Shao, D. (2016). Wavelet analysis of precipitation extremes over Canadian ecoregions and teleconnections to large-scale climate anomalies. *Journal of Geophysical Research: Atmospheres, 121*(24), 14469-14486. doi: 10.1002/2016JD025533

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2011). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society, 93*(4), 485-498. doi: 10.1175/BAMS-D-11-00094.1

te Linde, A. H., Aerts, J. C. J. H., Bakker, A. M. R., & Kwadijk, J. C. J. (2010). Simulating low-probability peak discharges for the Rhine basin using resampled climate modeling data. *Water Resources Research, 46*(3). doi: 10.1029/2009WR007707

Tebaldi, C., Smith, R. L., Nychka, D., & Mearns, 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. doi: 10.1175/jcli3363.1

Teufel, B., Sushama, L., Huziy, O., Diro, G. T., Jeong, D. I., Winger, K., . . . Nguyen, V.-T.-V. (2018). Investigation of the mechanisms leading to the 2017 Montreal flood. *Climate Dynamics*. doi: 10.1007/s00382-018-4375-0

Teutschbein, C., & Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology, 456-457*, 12-29. doi: 10.1016/j.jhydrol.2012.05.052

Thirel, G., Andréassian, V., & Perrin, C. (2015). On the need to test hydrological models under changing conditions. *Hydrological Sciences Journal, 60*(7-8), 1165-1173. doi: 10.1080/02626667.2015.1050027

Thompson, D. W. J., Barnes, E. A., Deser, C., Foust, W. E., & Phillips, A. S. (2015). Quantifying the role of internal climate variability in future climate trends. *Journal of Climate, 28*(16), 6443-6456. doi: 10.1175/JCLI-D-14-00830.1

Thorne, R., & Woo, M.-K. (2011). Streamflow response to climatic variability in a complex mountainous environment: Fraser River basin, British Columbia, Canada. *Hydrological Processes, 25*(19), 3076-3085. doi: 10.1002/hyp.8225

Tolson, B. A., & Shoemaker, C. A. (2007). Cannonsville Reservoir watershed SWAT2000 model development, calibration and validation. *Journal of Hydrology, 337*(1), 68-86. doi: 10.1016/j.jhydrol.2007.01.017

Tootle, G. A., & Piechota, T. C. (2005). Interdecadal and interannual oceanic / atmospheric variability and United States seasonal streamflow. Dans *Impacts of Global Climate Change*. doi: 10.1061/40792(173)480

Tootle, G. A., Piechota, T. C., & Singh, A. (2005). Coupled oceanic-atmospheric variability and U.S. streamflow. *Water Resources Research, 41*(12), W12408. doi: 10.1029/2005WR004381

Trenberth, K. E. (1999). Conceptual framework for changes of extremes of the hydrological cycle with climate change. Dans T. R. Karl, N. Nicholls & A. Ghazi (Éds.), *Weather and Climate Extremes: Changes, Variations and a Perspective from the Insurance Industry* (pp. 327-339). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-015-9265-9_18. Repéré à https://doi.org/10.1007/978-94-015-9265-9_18

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society, 84*(9), 1205-1217. doi: 10.1175/bams-84-9-1205

Trenberth, K. E., Jones, P. D., Ambenje, P., Bojariu, R., Easterling, D., Tank, A. K., . . . Rusticucci, M. (2007). Observations: Surface and atmospheric climate change, chap. 3 of Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change: Cambridge University Press, Cambridge, UK and New York, NY, USA.

Tripathi, O. P., & Dominguez, F. (2013). Effects of spatial resolution in the simulation of daily and subdaily precipitation in the southwestern US. *Journal of Geophysical Research: Atmospheres, 118*(14), 7591-7605. doi: 10.1002/jgrd.50590

Troin, M., Arsenault, R., & Brissette, F. (2015a). Performance and uncertainty evaluation of snow models on snowmelt flow simulations over a nordic catchment (Mistassibi, Canada). *Hydrology, 2*(4), 289. doi: 10.3390/hydrology2040289

Troin, M., Arsenault, R., Martel, J.-L., & Brissette, F. (2018). Uncertainty of hydrological model components in climate change studies over two nordic Quebec catchments. *Journal of Hydrometeorology, 19*(1), 27-46. doi: 10.1175/jhm-d-17-0002.1

Troin, M., Poulin, A., Baraer, M., & Brissette, F. (2016). Comparing snow models under current and future climates: Uncertainties and implications for hydrological impact studies. *Journal of Hydrology, 540*, 588-602. doi: 10.1016/j.jhydrol.2016.06.055

Troin, M., Velázquez, J. A., Caya, D., & Brissette, F. (2015b). Comparing statistical post-processing of regional and global climate scenarios for hydrological impacts assessment: A case study of two Canadian catchments. *Journal of Hydrology, 520*, 268-288. doi: 10.1016/j.jhydrol.2014.11.047

Turcotte, R., Fortin, L.-G., Fortin, V., Fortin, J.-P., & Villeneuve, J.-P. (2007). Operational analysis of the spatial distribution and the temporal evolution of the snowpack water equivalent in southern Québec, Canada. *Hydrology Research, 38*(3), 211-234. doi: 10.2166/nh.2007.009

Twine, T. E., Kucharik, C. J., & Foley, J. A. (2005). Effects of El Niño–Southern Oscillation on the climate, water balance, and streamflow of the Mississippi River basin. *Journal of Climate, 18*(22), 4840-4861. doi: 10.1175/jcli3566.1

Valéry, A. (2010). Modélisation précipitations–débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants. Ph.D. thesis, Agro Paris Tech,

417 pp. https://webgr.irstea.fr/wp-content/uploads/2012/07/2010-VALERY-THESE.pdf.

van den Besselaar, E. J. M., Klein Tank, A. M. G., & Buishand, T. A. (2013). Trends in European precipitation extremes over 1951–2010. *International Journal of Climatology, 33*(12), 2682-2689. doi: 10.1002/joc.3619

van der Linden, P., & Mitchell, J. F. B. (2009). *ENSEMBLES: Climate Change and its Impacts - Summary of research and results from the ENSEMBLES project*. Repéré à http://mfkp.org/INRMM/article/14257308

van der Spek, J. E., & Bakker, M. (2017). The influence of the length of the calibration period and observation frequency on predictive uncertainty in time series modeling of groundwater dynamics. *Water Resources Research, 53*(3), 2294-2311. doi: 10.1002/2016WR019704

Van Griensven, A., Meixner, T., Srinivasan, R., & Grunwald, S. (2008). Fit-for-purpose analysis of uncertainty using split-sampling evaluations. *Hydrological Sciences Journal, 53*(5), 1090-1103. doi: 10.1623/hysj.53.5.1090

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., . . . Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change, 109*(1), 5. doi: 10.1007/s10584-011-0148-z

Vano, J. A., Kim, J. B., Rupp, D. E., & Mote, P. W. (2015). Selecting climate change scenarios using impact-relevant sensitivities. *Geophysical Research Letters, 42*(13), 5516-5525. doi: 10.1002/2015GL063208

Vehviläinen, B. (1992a). Snow cover models in operational watershed forecasting. *PhD Thesis, National Board of Waters and the Environment, Helsinki.*

Vehviläinen, B. (1992b). *Snow cover models in operational watershed forecasting. Ph.D. dissertation, Finnish Environment Institute, 112 pp.*

Velázquez, J. A., Anctil, F., Ramos, M. H., & Perrin, C. (2011). Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. *Advances in Geosciences, 29*, 33. doi: 10.5194/adgeo-29-33-2011

Velázquez, J. A., Troin, M., & Caya, D. (2015a). Modelación hidrológica del río Tampaón en el contexto del cambio climático. *Tecnología y ciencias del agua, 6*(5), 17-30.

Velázquez, J. A., Troin, M., Caya, D., & Brissette, F. (2015b). Evaluating the time-invariance hypothesis of climate model bias correction: Implications for hydrological impact studies. *Journal of Hydrometeorology, 16*(5), 2013-2026. doi: 10.1175/jhm-d-14-0159.1

Verseghy, D. L. (1991). Class—A Canadian land surface scheme for GCMS. I. soil model. *International Journal of Climatology, 11*(2), 111-133. doi: 10.1002/joc.3370110202

Verseghy, D. L., McFarlane, N. A., & Lazare, M. (1993). Class—A Canadian land surface scheme for GCMS, II. Vegetation model and coupled runs. *International Journal of Climatology, 13*(4), 347-370. doi: 10.1002/joc.3370130402

Vidal, J.-P., Hingray, B., Magand, C., Sauquet, E., & Ducharne, A. (2016). Hierarchy of climate and hydrological uncertainties in transient low-flow projections. *Hydrology and Earth System Sciences, 20*(9), 3651-3672.

Volosciuk, C., Maraun, D., Semenov, V. A., & Park, W. (2015). Extreme precipitation in an atmosphere general circulation model: impact of horizontal and vertical model resolutions. *Journal of Climate, 28*(3), 1184-1205. doi: 10.1175/jcli-d-14-00337.1

von Salzen, K., Scinocca, J. F., McFarlane, N. A., Li, J., Cole, J. N. S., Plummer, D., . . . Solheim, L. (2013). The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of physical processes. *Atmosphere-Ocean, 51*(1), 104-125. doi: 10.1080/07055900.2012.755610

von Storch, H., & Zwiers, F. W. (1999). Statistical analysis in climate research (pp. 484): Cambridge University Press.

Vrac, M., & Friederichs, P. (2015). Multivariate—intervariable, spatial, and temporal—bias correction. *Journal of Climate, 28*(1), 218-237. doi: 10.1175/jcli-d-14-00059.1

Vrugt, J. A., Gupta, H. V., Dekker, S. C., Sorooshian, S., Wagener, T., & Bouten, W. (2006). Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting model. *Journal of Hydrology, 325*(1), 288-307. doi: 10.1016/j.jhydrol.2005.10.041

Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment, 19*(6), 378-387. doi: 10.1007/s00477-005-0006-5

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V., . . . Wilson, J., S. (2010). The future of hydrology: An evolving science for a changing world. *Water Resources Research, 46*(5). doi: 10.1029/2009WR008906

Wallace, J. M., & Gutzler, D. S. (1981). Teleconnections in the geopotential height field during the northern hemisphere winter. *Monthly Weather Review, 109*(4), 784-812. doi: 10.1175/1520-0493(1981)109<0784:Titghf>2.0.Co;2

Wallner, M., Haberlandt, U., & Dietrich, J. (2012). Evaluation of different calibration strategies for large scale continuous hydrological modelling. *Advances in Geosciences 31 (2012), 31*, 67-74. doi: 10.5194/adgeo-31-67-2012

Wanders, N., & Wada, Y. (2015). Decadal predictability of river discharge with climate oscillations over the 20th and early 21st century. *Geophysical Research Letters, 42*(24), 10689-10695. doi: 10.1002/2015GL066929

Wang, C.-C., Lee, W.-L., Chen, Y.-L., & Hsu, H.-H. (2015). Processes leading to Double Intertropical Convergence Zone bias in CESM1/CAM5. *Journal of Climate, 28*(7), 2900-2915. doi: 10.1175/jcli-d-14-00622.1

Wang, J. Y., Whitfield, P. H., & Cannon, A. J. (2006). Influence of pacific climate patterns on low-flows in British Columbia and Yukon, Canada. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 31*(1), 25-40. doi: 10.4296/cwrj3101025

Wang, X. L., Wan, H., Zwiers, F. W., Swail, V. R., Compo, G. P., Allan, R. J., . . . Yin, X. (2011). Trends and low-frequency variability of storminess over western Europe, 1878–2007. *Climate Dynamics, 37*(11), 2355-2371. doi: 10.1007/s00382-011-1107-0

Wang, Y., Zhou, B., Qin, D., Wu, J., Gao, R., & Song, L. (2017). Changes in mean and extreme temperature and precipitation over the arid region of northwestern China: Observation and projection. *Advances in Atmospheric Sciences, 34*(3), 289-305. doi: 10.1007/s00376-016-6160-5

Ward, P. J., Beets, W., Bouwer, L. M., Aerts, J. C. J. H., & Renssen, H. (2010). Sensitivity of river discharge to ENSO. *Geophysical Research Letters, 37*(12), L12402. doi: 10.1029/2010GL043215

Ward, P. J., Eisner, S., Flörke, M., Dettinger, M. D., & Kummu, M. (2014a). Annual flood sensitivities to El Niño–Southern Oscillation at the global scale. *Hydrology and Earth System Sciences, 18*(1), 47-66. doi: 10.5194/hess-18-47-2014

Ward, P. J., Jongman, B., Kummu, M., Dettinger, M. D., Sperna Weiland, F. C., & Winsemius, H. C. (2014b). Strong influence of El Niño Southern Oscillation on flood risk around the world. *Proceedings of the National Academy of Sciences, 111*(44), 15659-15664. doi: 10.1073/pnas.1409822111

Ward, P. J., Kummu, M., & Lall, U. (2016). Flood frequencies and durations and their response to El Niño Southern Oscillation: Global analysis. *Journal of Hydrology, 539*, 358-378. doi: 10.1016/j.jhydrol.2016.05.045

Wehner, M. F. (2013). Very extreme seasonal precipitation in the NARCCAP ensemble: model performance and projections. *Climate Dynamics, 40*(1), 59-80. doi: 10.1007/s00382-012-1393-1

Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). Global increasing trends in annual maximum daily precipitation. *Journal of Climate, 26*(11), 3904-3918. doi: 10.1175/jcli-d-12-00502.1

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., . . . Roberts, N. M. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of geophysics, 52*(3), 522-555. doi: 10.1002/2014RG000464

Whan, K., & Zwiers, F. W. (2017). The impact of ENSO and the NAO on extreme winter precipitation in North America in observations and regional climate models. *Climate Dynamics, 48*(5), 1401-1411. doi: 10.1007/s00382-016-3148-x

Widmann, M., & Bretherton, C. S. (2000). Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. *Journal of Climate, 13*(11), 1936-1950. doi: 10.1175/1520-0442(2000)013<1936:Vompit>2.0.Co;2

Wigley, T. M. L., Jones, P. D., Briffa, K. R., & Smith, G. (1990). Obtaining sub-grid-scale information from coarse-resolution general circulation model output. *Journal of Geophysical Research: Atmospheres, 95*(D2), 1943-1953. doi: 10.1029/JD095iD02p01943

Wilby, R. L. (2005). Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrological Processes, 19*(16), 3201-3219. doi: 10.1002/hyp.5819

Wilby, R. L., & Harris, I. (2006). A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research, 42*(2). doi: 10.1029/2005WR004065

Wilks, D. S. (1997). Resampling hypothesis tests for autocorrelated fields. *Journal of Climate, 10*(1), 65-82. doi: 10.1175/1520-0442(1997)010<0065:rhtfaf>2.0.co;2

Wilks, D. S. (2006). On "field significance" and the false discovery rate. *Journal of Applied Meteorology and Climatology, 45*(9), 1181-1189. doi: 10.1175/jam2404.1

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.

Woldemeskel, F. M., Sharma, A., Sivakumar, B., & Mehrotra, R. (2016). Quantification of precipitation and temperature uncertainties simulated by CMIP3 and CMIP5 models. *Journal of Geophysical Research: Atmospheres, 121*(1), 3-17. doi: 10.1002/2015JD023719

Woo, M.-K., & Thorne, R. (2008). Analysis of cold season streamflow response to variability of climate in north-western North America. *Hydrology Research, 39*(4), 257-265. doi: 10.2166/nh.2008.102

Wood, A. W., Maurer, E. P., Kumar, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research: Atmospheres, 107*(D20), ACL 6-1-ACL 6-15. doi: 10.1029/2001JD000659

Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M., & Brayshaw, D. J. (2012). Response of the North Atlantic storm track to climate change shaped by ocean–atmosphere coupling. *Nature Geoscience, 5*, 313. doi: 10.1038/ngeo1438

Wuebbles, D., Meehl, G., Hayhoe, K., Karl, T. R., Kunkel, K., Santer, B., . . . Sun, L. (2013). CMIP5 climate model analyses: Climate extremes in the United States. *Bulletin of the American Meteorological Society, 95*(4), 571-583. doi: 10.1175/BAMS-D-12-00172.1

Xu, C.-Y. (1999). From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography: Earth and Environment, 23*(2), 229-249. doi: 10.1177/030913339902300204

Yu, B., Lin, H., Wu, Z. W., & Merryfield, W. J. (2016). Relationship between North American winter temperature and large-scale atmospheric circulation anomalies and its decadal variation. *Environmental Research Letters, 11*(7), 074001. doi: 10.1088/1748-9326/11/7/074001

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., . . . Zwiers, F. W. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change, 2*(6), 851-870. doi: 10.1002/wcc.147

Zhang, X., Harvey, K. D., Hogg, W. D., & Yuzyk, T. R. (2001). Trends in Canadian streamflow. *Water Resources Research, 37*(4), 987-998. doi: 10.1029/2000WR900357

Zhang, X., Lin, W., & Zhang, M. (2007). Toward understanding the double Intertropical Convergence Zone pathology in coupled ocean-atmosphere general circulation models. *Journal of Geophysical Research: Atmospheres, 112*(D12102). doi: 10.1029/2006JD007878

Zhang, X. C. (2005). Spatial downscaling of global climate model output for site-specific assessment of crop production and soil erosion. *Agricultural and Forest Meteorology, 135*(1), 215-229. doi: 10.1016/j.agrformet.2005.11.016

Zhu, J. (2013). Impact of climate change on extreme rainfall across the United States. *Journal of Hydrologic Engineering, 18*(10), 1301-1309. doi: 10.1061/(ASCE)HE.1943-5584.0000725

Zhuan, M.-J., Chen, J., Shen, M.-X., Xu, C.-Y., Chen, H., & Xiong, L.-H. (2018). Timing of human-induced climate change emergence from internal climate variability for hydrological impact studies. *Hydrology Research, 49*(2), 421-437. doi: 10.2166/nh.2018.059

Zolina, O. (2012). Changes in intense precipitation in Europe. In: Changes in flood risk in Europe. Special publication No.10. IAHS Press, Wallingford, Oxfordshire, UK.

Zolina, O., Simmer, C., Kapala, A., Shabanov, P., Becker, P., Mächel, H., . . . Groisman, P. (2014). Precipitation variability and extremes in central europe: new view from STAMMEX results. *Bulletin of the American Meteorological Society, 95*(7), 995-1002. doi: 10.1175/bams-d-12-00134.1