ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

THÈSE PAR ARTICLES PRÉSENTÉE À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DU
DOCTORAT EN GÉNIE
Ph. D.

PAR
Gilles René Comlan ESSOU

POTENTIEL DES DONNÉES DE PRÉCIPITATION ET TEMPÉRATURE DES RÉANALYSES ATMOSPHÉRIQUES EN MODÉLISATION HYDROLOGIQUE

MONTRÉAL, LE 1er AOÛT 2016

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. Christian Masson, président du jury
Département de génie mécanique à l'École de technologie supérieure

M. Saad Bennis, membre du jury
Département de génie de la construction à l'École de technologie supérieure

M. Alain N. Rousseau, examinateur externe indépendant
Centre Eau Terre Environnement à l'Institut National de la Recherche Scientifique

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

LE 13 JUIN 2016

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

AVANT-PROPOS

Cette thèse, au format d'une thèse par articles, a été supervisée par le Professeur François Brissette à l'École de technologie supérieure de l'Université du Québec. L'objectif principal était d'évaluer pour les études en modélisation hydrologique, le potentiel des données de précipitation et température provenant des réanalyses atmosphériques. Ainsi, cette thèse se place dans le contexte global de la recherche de sources de données météorologiques pouvant servir d'alternative aux traditionnelles stations météorologiques, afin de pallier le déficit d'information dans les régions où ces stations sont en nombre insuffisant ou inexistantes. Elle fait partie d'un projet de recherche plus étendu (dénommé «ALICE») qui vise à évaluer la «*Valeur scientifique et opérationnelle des données alternatives en science hydrologique*». Le projet ALICE est cofinancé par le Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG), Hydro-Québec, Rio-Tinto-Alcan, Ontario Power Generation et le Consortium Ouranos sur la Climatologie Régionale et l'Adaptation aux Changements Climatiques.

Cinq articles composent cette thèse. Quatre de ces articles présentés aux chapitres 3 à 6 constituent le corps de cette thèse, et le doctorant en est premier auteur. Le cinquième article présenté en annexe I est un article de collaboration pour lequel le doctorant est deuxième auteur. Le superviseur de thèse est l'un des co-auteurs de chacun de ces articles. Les cinq articles ont été publiés ou soumis dans des revues scientifiques avec comité de lecture. Bien que cette thèse soit en français, les articles y sont présentés en anglais. Leur structure est basée sur les règles imposées par les journaux où ils ont été soumis.

REMERCIEMENTS

Ce projet de recherche n'aurait pas abouti sans le soutien de plusieurs personnes qui y ont contribué de près ou de loin. Grand merci à chacun de vous !

J'aimerais plus particulièrement exprimer ma profonde gratitude à mon directeur de thèse, M. François Brissette, pour sa perspicacité, ses nombreuses idées scientifiques clés et conseils judicieux qui m'ont guidé à parachever cette thèse. François, je te remercie très sincèrement de m'avoir fait confiance tout au long de cette thèse, et d'avoir toujours été attentif à chacune de mes préoccupations. Malgré ton agenda bien rempli, tu as toujours su trouver une place pour lire mes travaux et émettre des commentaires très pertinents.

Je remercie profondément l'ensemble des membres du jury d'avoir accepté d'évaluer cette thèse, et pour toutes leurs remarques et suggestions constructives en vue de son amélioration. Je remercie M. Christian Masson d'avoir accepté de présider ce jury. Merci à M. Alain N. Rousseau du Centre Eau Terre Environnement de l'Institut National de la Recherche Scientifique (INRS), qui a accepté d'être l'examinateur externe indépendant de cette thèse. Merci à M. Saad Bennis pour sa lecture attentive et critique de la thèse. Merci également à M. Philippe Lucas-Picher pour son aide précieuse dans la compréhension des sciences du climat, et pour sa pertinente contribution scientifique à cette thèse. Merci aux professeurs du département de génie de la construction de l'ÉTS, en particulier Mme Annie Poulin.

Grand merci à la Banque Islamique de développement (BID) qui a soutenu financièrement cette recherche en m'octroyant une prestigieuse bourse de doctorat sans laquelle, il ne m'aurait pas été possible d'entreprendre ce projet. Je salue également tous les organismes subventionnaires de projet ALICE : le Conseil de Recherches en Sciences Naturelles et en Génie du Canada (CRSNG), Hydro-Québec, Rio-Tinto-Alcan, Ontario Power Generation et le Consortium Ouranos. Vos subventions ont grandement contribué à l'aboutissement de ce projet et je vous en remercie.

Merci à tous mes collègues du laboratoire DRAME, pour l'esprit d'équipe et la bonne ambiance et qui a toujours régné au sein du groupe. Il ne m'est pas possible d'énumérer tous les noms au risque d'en oublier involontairement quelques-uns. Toutefois, je voudrais saluer particulièrement Jie Chen, Richard Arsenault, Salam Bajamgnigni Gbambie, Florent Sabarly et Salomon Salumu. Votre collaboration a été pour moi une expérience très enrichissante, et je vous en remercie très sincèrement.

Je tiens aussi à remercier mes collègues de doctorat Eustache Gooré Bi et Landry Mballa pour leurs conseils et encouragements, et pour les échanges scientifiques que nous avons eus.

Je remercie grandement toute la communauté CMCI de Montréal, en particulier Calvin Wuntcha et Rodrigue Kéou pour leurs conseils et encouragements. Grand merci à Joël Wambé et son épouse Scholastique, pour leur soutien inconditionnel durant ce doctorat !

Je remercie profondément mes parents Guy et Jacqueline, pour leur soutien continu et leur affection. Grand merci à mes frères et sœurs, en particulier Roméo Essou, pour leurs conseils et soutien inconditionnel !

Enfin, mon plus grand merci va à l'endroit de mon épouse Kunteenne Jacqueline, qui n'a pas hésité à accepter de braver les hivers éprouvants de Montréal et à donner de son tout, afin de m'accompagner dans ce projet parsemé d'incertitudes et de questionnements. En plus d'être une merveilleuse épouse, tu es une excellente mère. Tu as toujours su entourer nos enfants de soins et d'affection. Je voudrais aussi te dire mille mercis d'avoir été patiente et compréhensive face à mes fréquentes veilles durant lesquelles je m'absorbais dans mes travaux de recherche. Tu as même très souvent, sacrifié volontiers ton sommeil pour me tenir compagnie à de telles occasions. Je t'en suis profondément reconnaissant. Je voudrais aussi dire à vous mes chers enfants (Johanna – Josué – Jéhoaddan), merci infiniment de votre patience durant toutes ces années de mon doctorat. Vous avez été une grande source de motivation pour moi. Je vous aime grandement !

POTENTIEL DES DONNÉES DE PRÉCIPITATION ET TEMPÉRATURE DES RÉANALYSES ATMOSPHÉRIQUES EN MODÉLISATION HYDROLOGIQUE

Gilles René Comlan ESSOU

RÉSUMÉ

La faible couverture spatiale des stations météorologiques dans plusieurs régions au monde limite la capacité des modèles hydrologiques à simuler les débits en rivière.

L'objectif de cette thèse est d'évaluer le potentiel des réanalyses atmosphériques comme alternative aux stations météorologiques afin de pallier le déficit d'information dans les régions où ces stations sont clairsemées ou inexistantes. Pour ce faire, des données de précipitation et température de trois réanalyses atmosphériques globales récentes (ERA-Intérim, CFSR et MERRA) ont été utilisées comme intrants météorologiques d'un modèle hydrologique pour simuler des débits moyens journaliers provenant de plus de 800 bassins versants situés dans différentes régions climatiques aux USA et au Canada.

Dans un premier temps, une pré-validation de jeux de données de précipitation et température des réanalyses a été faite en les comparant à des jeux de données d'observation à travers les USA, où la couverture spatiale des stations météorologiques est élevée. Chacun des jeux de données a ensuite été utilisé pour caler un modèle hydrologique et pour simuler des débits moyens journaliers provenant de 370 bassins versants aux USA. Les résultats ont montré que la température des réanalyses était similaire à celle des observations sur la majeure partie des USA. Par contre, la précipitation des trois réanalyses globales était biaisée, surtout durant l'été et l'hiver dans le Sud-Est des USA. En dépit de ces biais, les débits simulés contraints par les données des réanalyses étaient similaires à ceux forcés par les observations, sauf dans les régions climatiques Continentales et Subtropicales humides, où l'inadéquate saisonnalité de la précipitation des réanalyses a dégradé la qualité des débits simulés.

Ensuite, au Canada où la couverture spatiale des stations météorologiques est plus faible, la précision des débits simulés pour 316 bassins versants en utilisant des données de réanalyses a été comparée à celle des débits simulés en utilisant des données d'observation, en fonction de la densité de stations météorologiques des bassins versants. Les résultats ont montré que les débits simulés en utilisant les données de précipitation et température de CFSR étaient généralement similaires à ceux simulés en utilisant des données d'observation sur grille, quelle que soit la densité de stations météorologiques. Par contre, ERA-Interim et MERRA ont significativement mieux performé que les données d'observation dans la région de montagne, notamment lorsque la densité de stations météorologiques était inférieure à 1 station pour 1000km^2.

Enfin, les impacts de la combinaison des trois réanalyses atmosphériques globales et des données d'observation sur la précision des débits simulés ont été évalués. Deux approches de combinaison des bases de données ont été considérées. L'une consiste à utiliser une moyenne pondérée des intrants météorologiques (précipitation et température) de toutes les bases de données, pour caler le modèle hydrologique et pour simuler les débits. La seconde approche consiste à utiliser séparément les intrants météorologiques et à calculer une moyenne pondérée des différents hydrogrammes simulés. Les résultats ont montré des améliorations significatives sur la précision de débits simulés tant en combinant les intrants météorologiques qu'en combinant les hydrogrammes simulés, pour la plupart des bassins versants au Canada et aux USA. Par ailleurs, dans 100% des cas où la précision des débits simulés en utilisant uniquement des données d'observation est faible (correspondant à une valeurs du Nash-Sutcliffe < 0,5), la prise en compte des données des réanalyses a permis d'améliorer considérablement la précision des débits simulés (valeurs de Nash-Sutcliffe augmentées d'au moins 0,3).

Globalement, les résultats de cette thèse suggèrent que les données de précipitation et température provenant des réanalyses atmosphériques globales peuvent être utilisées pour les études de modélisation hydrologique dans les régions où il y a peu de stations météorologiques. Toutefois, puisque le potentiel des données de précipitation et température des réanalyses varie spatialement, elles doivent néanmoins être utilisées avec précaution dans les études hydrologiques.

Mots-clés : Réanalyses, données d'observation, modélisation hydrologique, calage, densité de stations météorologiques

POTENTIAL OF PRECIPITATION AND TEMPERATURE DATA FROM ATMOSPHERIC REANALYSES FOR HYDROLOGICAL MODELING

Gilles René Comlan ESSOU

ABSTRACT

The sparse coverage of weather stations over several regions of the world limits the ability of hydrological models to adequately simulate river flows.

The objective of this thesis is to evaluate the potential of atmospheric reanalyses as an alternative to weather stations to overcome the lack of information in areas where these stations are sparse or nonexistent. To do this, precipitation and temperature data from three recent global atmospheric reanalyses (ERA-Interim, CFSR and MERRA) were used as meteorological inputs to a hydrological model to simulate daily discharges on 800 watersheds located in different climatic regions of the USA and Canada.

First, a pre-validation of precipitation and temperature datasets from the global reanalyses was performed by comparing them to observational datasets over the USA where the spatial coverage of weather stations is high. Each dataset was then used to calibrate a hydrological model and to simulate daily river flows of 370 US watersheds. Results showed that temperatures from reanalyses were similar to that of observational data over most of the USA. On the other hand, precipitation from all three global reanalyses was biased, especially in summer and winter in south-eastern USA. Despite these biases, the simulated flows forced by the reanalysis datasets were similar to those forced by observations, except in the humid continental and subtropical climatic regions, where the poor precipitation seasonality of reanalyses degraded river flow simulations.

In Canada where the spatial coverage of weather stations is lower, the accuracy of the simulated streamflows of 316 watersheds using reanalysis data was compared to that of the flows simulated using observational data, according to the density of weather stations. Results showed that the simulated streamflows using precipitation and temperature data from CFSR were generally similar to those simulated using gridded observations, regardless of the weather station density. On the other hand, ERA-Interim and MERRA performed significantly better than the gridded observations in the Mountain region, especially when the density of weather stations is less than 1 station per 1000km^2.

Finally, the impacts of the combination of the three global atmospheric reanalyses and observational data on the accuracy of the simulated streamflows was evaluated. Two combination approaches were considered. The first consists of using a weighted average of meteorological inputs (precipitation and temperature) from all the databases, to calibrate the hydrological model and to simulate streamflow. The second approach consists of using all meteorological inputs separately to simulate hydrographs and to compute a weighted average of the simulated hydrographs. Results showed significant improvements of the accuracy of simulated streamflows in both combination cases over most watersheds. Moreover, in 100%

of the cases where the accuracy of the simulated streamflows using only observational data was low (corresponding to a Nash-Sutcliffe value < 0.5), taking into account reanalyses data greatly improved the accuracy of the simulated streamflows (Nash-Sutcliffe values increased by at least 0.3).

Overall, the results of this thesis suggest that precipitation and temperature from global atmospheric reanalyses can be used for hydrological modeling studies in regions where there are few weather stations. However, since the potential for precipitation and temperature data from reanalyses varies spatially, they should be used with caution in hydrological studies.

Keywords: Reanalyses, observational data, hydrological modeling, calibration, density of weather stations

TABLE DES MATIÈRES

Page

LISTE DES TABLEAUX

Page

LISTE DES FIGURES

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

CANOPEX	CANadian mOdel Parameter EXperiment
CFSR	Climate Forecast System Reanalysis
CMAES	Covariance Matrix Adaptation Evolution Strategy
CPC	Climate Prediction Center
CRU	Climatic Research Unit
ECMWF	European Center for Medium-Range Weather Forecasts
ERA-40	ECMWF Re-Analysis 40 years
ERA-Interim	ECMWF Re-Analysis Interim
GPCC	Global Precipitation Climatology Centre
GRC	Granger – Ramanathan Averaging C
HYDAT	HYdrometric DATabase
MERRA	Modern Era Reanalysis for Research and Applications
MOPEX	MOdel Parameter EXperiment
NARR	North American Regional Reanalysis
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NRCan	Natural Resource Canada
NSE	Nash-Sutcliffe Efficiency
RMSE	Root Mean Squared Error
USA	United States of America
WATCH	WATer and global CHange project

WFDEI	WATCH Forcing Data methodology applied to ERA-Interim reanalysis data
3D-VAR	Three Dimensional Variational scheme
4D-VAR	Four Dimensional Variational scheme

INTRODUCTION

0.1 Contexte

L'eau est essentielle à la vie et est indispensable à d'innombrables activités socio-économiques et récréatives. Son cycle et ses interactions avec les milieux terrestre et atmosphérique sont étudiés par l'hydrologie. L'hydrologie intervient dans plusieurs applications dont la planification et la gestion des ressources en eau (irrigation, production hydroélectrique, etc.) et la prévision des événements extrêmes (tels que les inondations et sécheresses) (Hingray et al. 2009; Singh and Woolhiser 2002). Très souvent, les hydrologues utilisent des modèles hydrologiques pour simuler la réponse hydrologique d'un hydrosystème aux sollicitations météorologiques telles que la précipitation et la température (Payraudeau et al. 2002; Pechlivanidis et al. 2011). L'hydrosystème le plus couramment considéré en modélisation hydrologique est le *bassin versant*. Sa réponse hydrologique est souvent exprimée sous forme de débit d'écoulement qui peut servir entre autres à: reconstituer des débits historiques non mesurés, générer des scénarios hydrologiques en vue de la gestion et l'aménagement du territoire (e.g. fournir l'information hydrologique nécessaire au dimensionnement d'ouvrages hydrauliques et de protection contre les crues), évaluer l'impact des changements futurs sur les ressources en eau, notamment en ce qui a trait au climat et à l'occupation du sol.

En modélisation hydrologique, certaines données sont généralement requises. Les principales peuvent être regroupées en trois catégories. La première catégorie est constituée de données sur les caractéristiques physiques du bassin versant telles que son contour, sa superficie, sa topographie et l'occupation du sol, et de données sur son réseau hydrographique telles que les profils en long et en travers des biefs principaux. La deuxième catégorie se compose de données sur les conditions initiales telles que l'humidité du sol, l'accumulation initiale de neige et le débit initial. La troisième catégorie se compose de données hydrométriques telles que les débits et les niveaux d'eau dans la rivière, et de données météorologiques telles que la précipitation et la température.

En général, les données météorologiques proviennent des stations météorologiques qui ne fournissent que des mesures ponctuelles. Cependant, dans plusieurs régions à travers le monde, les réseaux de stations météorologiques sont clairsemés, ce qui entraîne un déficit de l'information météorologique requise par les modèles hydrologiques. D'où l'importance de trouver d'autres sources de données météorologiques pouvant contribuer à pallier ce déficit.

Les publications antérieures sur le sujet exposent différentes approches d'estimation de données météorologiques à des sites où les stations météorologiques sont clairsemées ou inexistantes, pour des applications hydrologiques. Les principales approches rapportées dans la littérature sont l'estimation de données soit par interpolation spatiale de stations météorologiques (exemples, Herrera et al. (2012); Wibig et al. (2014); Sun et al. (2014)), par télédétection au moyen de radar (exemples, Yilmaz et al. (2005); Rango (1994)) ou de satellites (exemples, Schmugge et al. (2002); Tang et al. (2009); Gleason and Smith (2014)), ou encore, par combinaison du peu de données de stations disponibles et de données de télédétection (exemples, Xiaoyang et al. (2003); Yu et al. (2011)). Bien que ces différentes approches aient connu un certain succès, quelques défis majeurs demeurent. D'un autre côté, grâce aux progrès technologiques des dernières décennies, de nouvelles sources de données météorologiques appelées *réanalyses atmosphériques* ont été développées. Elles ont l'avantage d'offrir une couverture globale pour plusieurs centaines de variables climatiques, et leurs données sont spatialement et temporellement cohérentes (Dee et al. 2011; Mesinger et al. 2006; Saha et al. 2010). Ainsi donc, les données provenant des réanalyses atmosphériques pourraient être une alternative intéressante aux traditionnelles données de stations météorologiques, pour des études hydrologiques dans des régions où les données de stations météorologiques sont insuffisantes ou inexistantes. Pourtant, à ce jour, le potentiel des réanalyses pour la modélisation hydrologique a été peu évalué dans la littérature, comparativement à celui des données estimées par les approches énumérées précédemment.

Dans ce contexte, la présente recherche vise à évaluer le potentiel des données de précipitation et température des réanalyses en modélisation hydrologique. Afin de prendre en considération différents régimes hydroclimatiques, l'étude se réalise sur plusieurs centaines de bassins versants répartis à travers les États-Unis et le Canada.

0.2 Structure de la thèse

La présente thèse comporte sept chapitres (figure 0.1). Ces chapitres sont précédés d'une introduction générale, et sont suivis d'une conclusion et de quelques recommandations.

Le chapitre 1 expose la problématique et les objectifs de la recherche.

Le chapitre 2 est une revue bibliographique des principales approches suggérées dans la littérature en lien avec la problématique de recherche. Il permet aussi de positionner les travaux de la présente thèse par rapport à l'ensemble de ces approches.

Le chapitre 3 présente l'article intitulé : *Comparison of climate datasets for lumped hydrological modeling over the continental United States*. Cet article soumis pour publication dans le *Journal of Hydrology*, compare quatre bases de données interpolées des États-Unis en modélisation hydrologique globale. À partir des résultats de cet article, l'une de ces bases de données (celle de Santa Clara) a été retenue pour être comparée aux réanalyses dans la suite des travaux, sur les bassins versants des États-Unis.

Le chapitre 4 est constitué de l'article intitulé : *Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling?* Cet article soumis pour publication dans le *Journal of Hydrometeorology*, compare d'une part, les données de température et précipitation des réanalyses à celles de la base de données interpolées de Santa Clara aux États-Unis. D'autre part, il compare leurs performances comme forçages météorologiques d'un modèle hydrologique global, pour la simulation des débits moyens journaliers de 370 bassins versants des États-Unis.

Le chapitre 5 présente l'article intitulé : *The use of reanalyses and gridded observations as weather input data for a hydrological model: comparison of performances of simulated river flows according to the density of weather stations*. Cet article soumis pour publication dans le Journal of Hydrometeorology, compare pour 316 bassins versants canadiens, la précision de débits simulés en utilisant des données provenant de trois réanalyses atmosphériques globales (ERA-Interim, CFSR et MERRA) à celle de débits simulés en utilisant des données interpolées sur grille à partir des données de stations météorologiques (NRCan). Cette comparaison a été faite suivant la densité de stations météorologiques.

Le chapitre 6 est constitué de l'article intitulé : *Impacts of combining reanalyses and weather station-based data on the accuracy of discharge modeling*. Cet article soumis pour publication dans le Journal of Hydrology, évalue les impacts de la combinaison réanalyses-observations sur la précision des débits simulés, pour 830 bassins versants situés au Canada et aux États-Unis, suivant deux approches de combinaison.

Le chapitre 7 présente une discussion générale de la thèse.

Un article connexe à cette thèse est présenté en annexe I. Il s'agit d'un travail duquel je suis co-auteur et qui est directement lié à la présente thèse. En effet, cet article intitulé «*Use of four reanalysis datasets to assess the terrestrial branch of the water cycle over Quebec, Canada*», compare les composantes du cycle hydrologique des réanalyses atmosphériques évaluées dans cette thèse, et étudie la fermeture de leur bilan hydrologique.

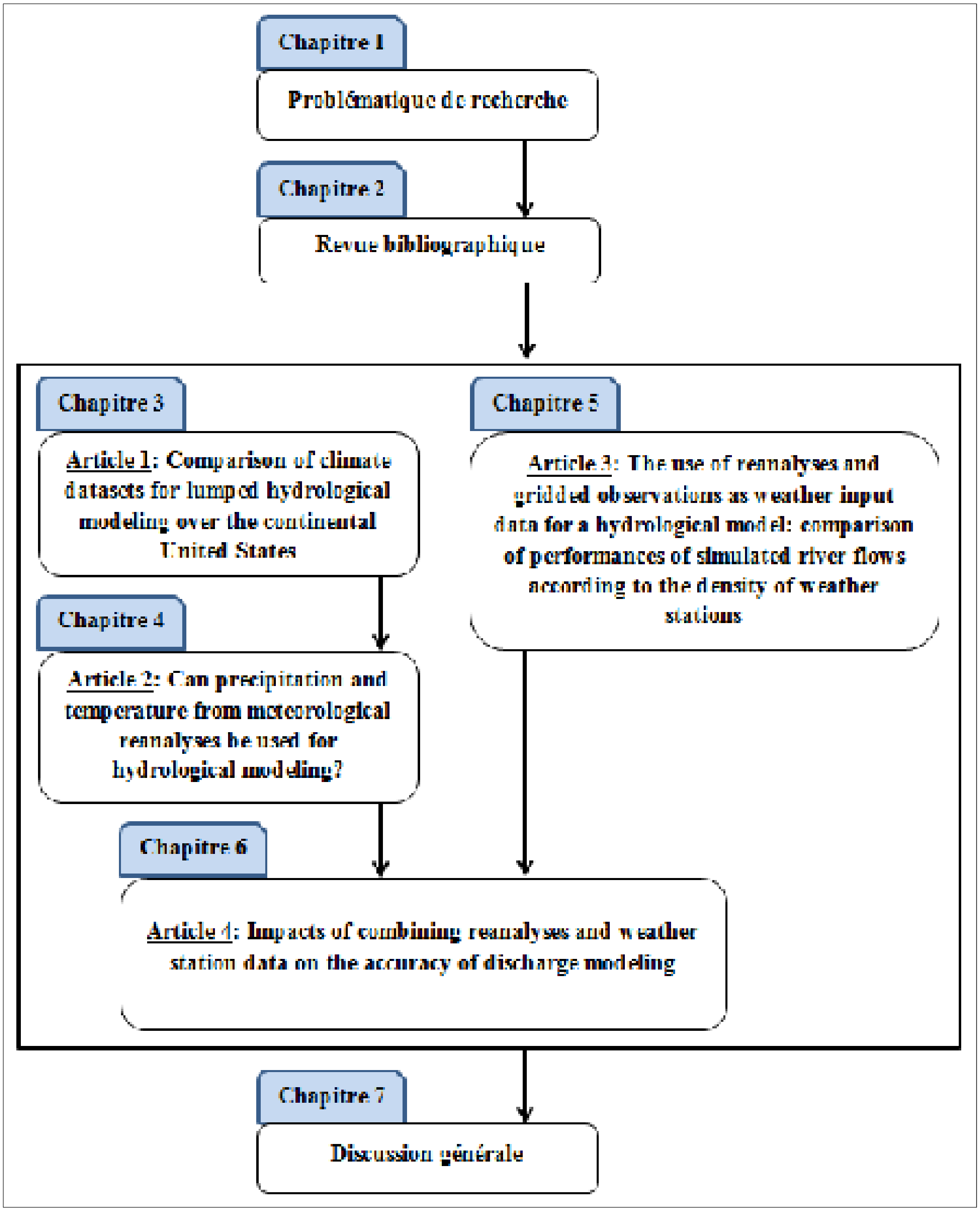

Figure 0.1 Structure de la thèse

CHAPITRE 1

PROBLÉMATIQUE DE RECHERCHE

Les études hydrologiques reposent fortement sur l'information météorologique. Parmi les variables météorologiques, l'une des plus cruciales en modélisation hydrologique est la précipitation (Krajewski et al. 1991; Lopes 1996; Obled et al. 1994). Cependant, obtenir des données de précipitation de qualité adéquate et en quantité suffisante est très souvent un défi aux ingénieurs hydrologues. En général, les données de précipitation utilisées dans les études hydrologiques proviennent des stations météorologiques. Mais, l'un des points faibles des enregistrements des stations météorologiques est qu'elles sont des mesures ponctuelles. Ainsi, même lorsqu'un réseau de stations est de forte densité, il ne peut fournir qu'une vue partielle de la variabilité spatiale de la précipitation.

Par ailleurs, les coûts d'acquisition, d'installation et d'entretien des stations météorologiques sont souvent très élevés. Pour cela, les réseaux de stations météorologiques sont de faible densité spatiale dans plusieurs régions (exemple, voir figure 1.1). Cela est particulièrement vrai pour les régions faiblement peuplées et celles difficiles d'accès comme les régions montagneuses. Dans de telles régions, les quelques stations existantes ne sont généralement pas suffisantes pour fournir l'information nécessaire aux études hydrologiques. De plus, les données ne sont souvent pas disponibles aux sites où elles sont requises, et même lorsqu'elles y sont disponibles, elles sont généralement peu représentatives des conditions météorologiques réelles.

En plus d'être des mesures ponctuelles, les données des stations météorologiques sont souvent lacunaires, de courte longueur, et ne couvrent pas toujours la période d'intérêt visée par un projet donné (Chvíla et al. 2005; Goodison et al. 1981; Nystuen 1999; Sevruk 1996; Upton and Rahimi 2003). Par exemple au Canada, sur 7968 stations météorologiques d'Environnement Canada (figure 1.2a), 82% ont plus de 50% de données de précipitation manquantes entre 1979 et 2014 (figure 1.2b), et seulement 0,73% des stations ont des séries de données de précipitation longues de 10 ans ou plus entre 1979 et 2014, avec moins de 10% de lacunes (figure 1.2c).

Le déficit d'information météorologique lié aux données des stations météorologiques entrave souvent sérieusement la faisabilité des analyses hydrologiques, ou affecte grandement leur fiabilité.

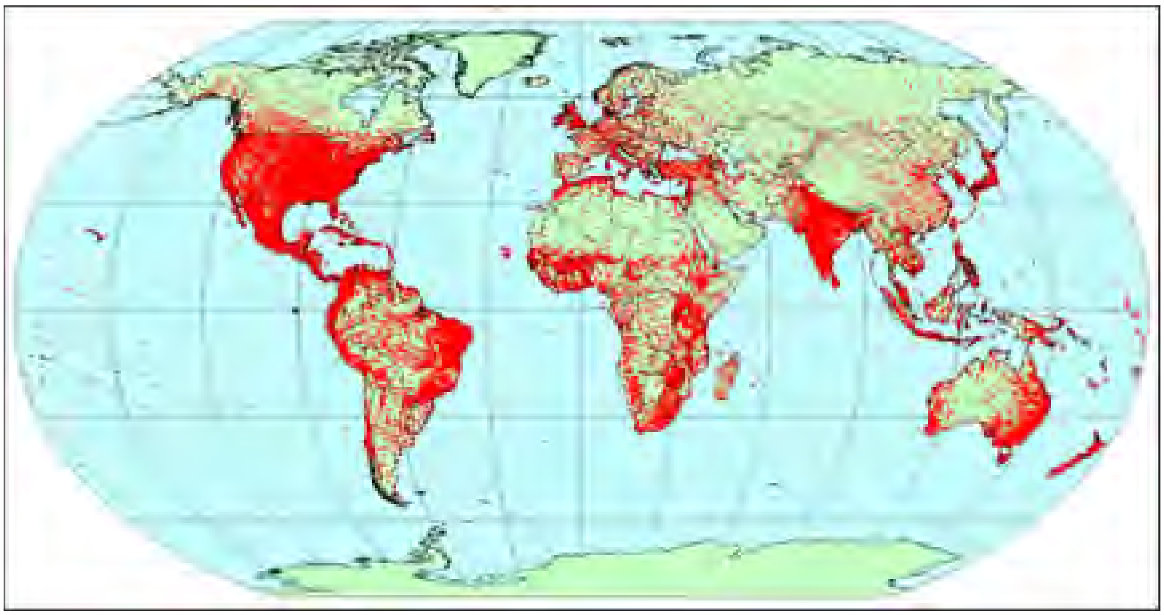

Figure 1.1 Localisation des stations météorologiques (points rouges) provenant d'une variété de bases de données à travers le monde, et ayant des mesures de précipitation (avec ou sans lacunes) sur la période 1950-2000
Adaptée de Hijmans et al. (2005)

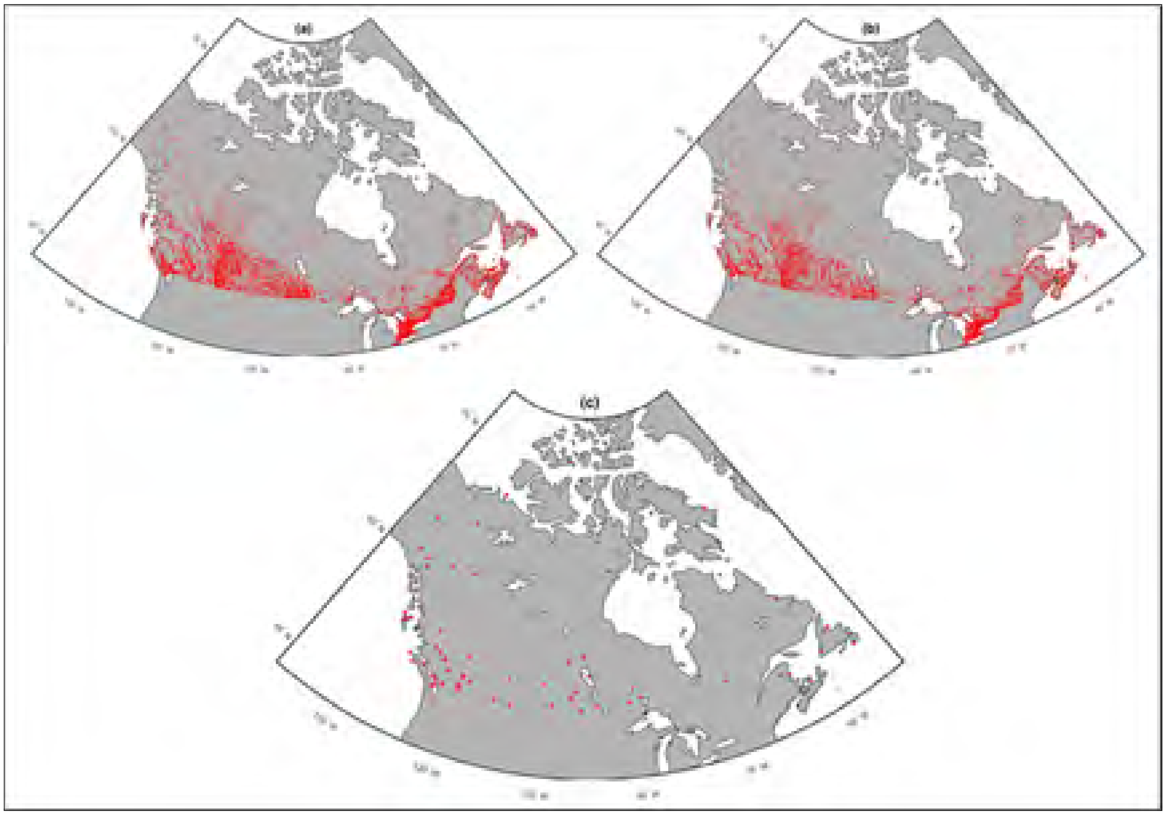

Figure 1.2 Localisations de 7968 stations météorologiques d'Environnement Canada (a); stations ayant plus de 50% de données de précipitation manquantes entre 1979 et 2014 (b); stations ayant au moins 10 ans de données de précipitation entre 1979 et 2014, avec moins de 10% de lacunes (c)

CHAPITRE 2

REVUE BIBLIOGRAPHIQUE, OBJECTIFS DE L'ÉTUDE ET HYPOTHÈSES

La présente revue bibliographique présente l'état des connaissances sur les différentes approches de solutions proposées dans la littérature pour pallier le déficit d'information découlant de données des stations météorologiques requises, notamment dans les régions où le réseau de stations météorologiques est de faible densité. Les forces et faiblesses de chacune de ces approches sont mises en évidence. Ce qui, d'une part, permet de situer l'approche envisagée dans la présente thèse par rapport à l'ensemble de ces autres approches, et d'autre part, d'évaluer son originalité.

2.1 Utilisation des données tirées de sources documentaires historiques

La connaissance des débits de pointe des crues historiques est d'une grande importance pour la planification et la prévention des risques pouvant être associés aux événements hydrologiques extrêmes. Mais, puisque les séries chronologiques de données de stations météorologiques sont souvent de courte longueur (et lacunaires), il n'est souvent pas possible d'estimer les débits de pointe des crues historiques en utilisant un modèle hydrologique. Alors, pour estimer de tels débits, plusieurs travaux effectués au cours des dernières décennies ont envisagé l'utilisation de diverses sources documentaires historiques (Glade and Albini 2001; Thorndycraft and Benito 2003). Ces sources documentaires se composent entre autres de journaux publiés, de documents picturaux, de correspondances personnelles, de sources narratives écrites, sources épigraphiques (exemple, figure 2.1), etc. (Brázdil et al. 2006; Payrastre et al. 2006). Il existe différentes méthodes pour la collecte et le traitement des informations contenues dans de tels documents (Cœur et al. 2002; Coeur and Lang 2000; Naulet et al. 2001). Les débits historiques estimés à partir des informations extraites sont utilisés comme compléments aux données hydrométéorologiques systématiques (ou mesurées).

Figure 2.1 Marque d'inondation du 9 Juillet 1736 sur un bâtiment historique
du marché Old Town à PoznaĚ (en Pologne)
Adaptée de Brázdil et al. (2006)

Le principal avantage de cette approche est qu’elle permet de reconstituer dans le temps et l’espace, les informations historiques concernant les conditions d’écoulement des cours d’eau, ainsi que les événements hydrologiques extrêmes observés (inondations et étiages sévères), en vue de leur analyse fréquentielle (Berger et al. 2010; Lang et al. 1998; Naulet 2002; Ouarda et al. 1998). Cette approche a été largement utilisée à travers le monde pour reconstituer la fluctuation, la fréquence et la saisonnalité des inondations. C’est le cas par exemples, au Canada (Shrubsole et al. 1993), en République Tchèque (Brázdil and Bukáček 2000; Brázdil et al. 2006) , en France (Cœur 2003; Naulet et al. 2005), en Allemagne (Glaser and Stangl 2004; Jacobeit et al. 2003; Mudelsee et al. 2004), en Italie (Alessandroni and

Remedia 2002; Guidoboni and Guidoboni 1998), aux Pays-Bas (Tol and Langen 2000), en Slovaquie (Svoboda et al. 2000), en Espagne (Benito et al. 2003; Llasat et al. 2005), en Suisse (Gees 1996; Pfisterer 1998), au Royaume-Uni (Archer 1999; Williams and Archer 2002), pour ne citer que ceux-là.

Le principal défi avec cette approche est que les sources documentaires historiques ne sont pas toujours disponibles. De plus, lorsqu'elles sont disponibles, il y a toujours le risque qu'elles contiennent des informations incomplètes, incohérentes et incertaines (Payrastre et al. 2006; Thorndycraft and Benito 2003).

En dépit de ces défis, cette approche semble être la seule alternative possible pour reconstituer les débits extrêmes du passé lointain, en vue de l'estimation de l'intensité ou de la fréquence de crues de projet pour la planification et la conception d'ouvrages hydrauliques. Elle demeure toutefois inappropriée pour plusieurs autres applications hydrologiques telle que la gestion des réservoirs hydriques qui nécessite l'estimation en temps réel ou quasi-réel des débits, et donc des volumes en stock. Pour de telles applications, l'utilisation des données systématiques est incontournable.

2.2 Utilisation des données obtenues par interpolation spatiale

Pour surmonter le problème de la faible couverture spatiale des réseaux de stations météorologiques, la spatialisation des données mesurées a été envisagée. Des techniques d'interpolation ont été développées à cet effet, en vue d'estimer les champs réels de variables météorologiques (généralement la précipitation et la température) à partir des mesures ponctuelles des stations météorologiques, pour obtenir une couverture spatiale plus complète de ces variables (figure 2.2).

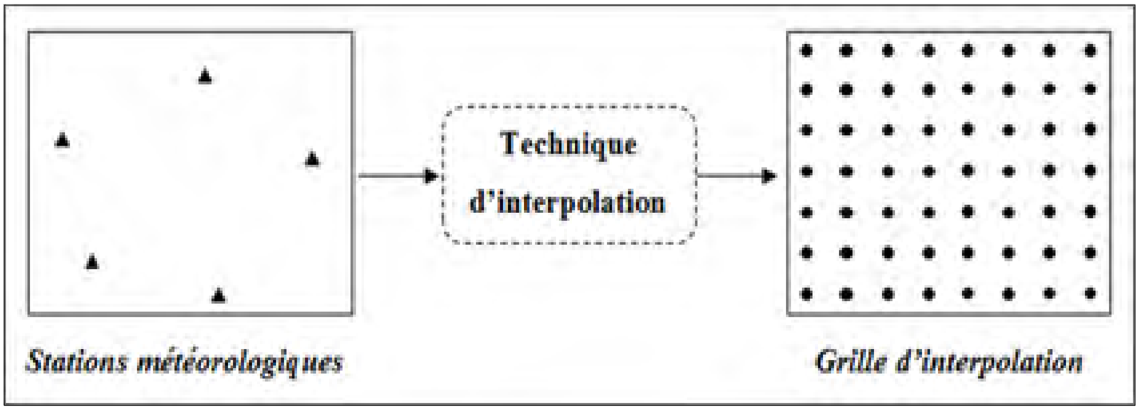

Figure 2.2 Illustration du principe de l'interpolation spatiale des stations météorologiques

Il existe une grande variété de techniques d'interpolation spatiale (Creutin and Obled 1982). Selon Degré et al. (2013), ces techniques peuvent être classées en deux groupes : les *techniques déterministes* et les *techniques géostatistiques*. Les techniques déterministes les plus fréquemment utilisées sont la méthode des polygones de Thiessen et la méthode de l'inverse de la distance (Konan et al. 2010; Skaugen and Andersen 2010; Vente 1964). Ces deux méthodes sont fréquemment utilisées en hydrologie opérationnelle car les poids affectés aux différentes stations restent invariants dans le temps, tant que le réseau de mesure n'est pas modifié. Toutefois, ces méthodes ne permettent pas une interpolation optimale du champ de précipitation (Hingray et al. 2009). Les méthodes géostatistiques (exemple krigeage) quant à elles permettent une interpolation optimale, mais sont moins aisées à appliquer et par conséquent, sont peu utilisées en mode opérationnel (Baillargeon et al. 2004; Mahdian et al. 2009; Perry and Hollis 2005).

Quelle que soit la méthode d'interpolation considérée, la précision des estimations dépend fortement du réseau de stations utilisé (Taupin 2003; Taupin et al. 1998), et moins le réseau de stations est dense, plus le choix de la méthode d'interpolation est déterminant (Christensen et al. 1998; Renard and Comby 2006). Toutes les techniques d'interpolation spatiale induisent des incertitudes dans les données interpolées, et l'incertitude est d'autant plus grande que le réseau est peu dense, et que le pas de temps considéré est petit (Avila et al. 2015; Daly 2006; Huiyi and Shaofeng 2010; Tozer et al. 2012). De plus, l'incertitude sur les estimés de précipitation varie en général, d’un événement pluvieux à un autre, suivant leur variabilité et extension spatiale.

Les données interpolées sont souvent produites sur grilles régulières et sont aussi appelées «données maillées» (Haylock et al. 2008; Rudolf and Schneider 2005). Il existe plusieurs bases de données interpolées à travers le monde. Certaines bases de données interpolées telles que NRCan (Hutchinson et al. 2009), Daymet (Thornton et al. 2012), CPC (Higgins et al. 2000), Santa Clara (Maurer et al. 2002), APHRODITE (Yatagai et al. 2012), etc., sont locales ou régionales. D’autres telles que CRU (Harris et al. 2014; New et al. 1999, 2000), GPCC (Rudolf and Schneider 2005; Schneider et al. 2014), etc., sont mondiales. Les différences entre les bases de données interpolées se situent non seulement dans leurs résolutions spatiale et temporelle, mais aussi dans la technique d’interpolation utilisée et dans la densité du réseau de stations météorologiques exploité.

Au cours des dernières décennies, plusieurs travaux ont comparé les données interpolées à celles observées, afin d’évaluer leur utilité pour les études hydrologiques (Gallo and Xian 2014a; Herrera et al. 2012; Pai et al. 2014; Sun et al. 2014; Turco et al. 2013; Wibig et al. 2014; Yin et al. 2014). Il ressort globalement de ces travaux que lorsque le réseau de stations météorologiques utilisé pour produire les données interpolées est dense et opérationnel, ces données interpolées se comparent bien aux observations et leur précision est généralement acceptable pour les études de modélisation hydrologique. Par exemple, Vu et al. (2012) et Lauri et al. (2014) ont observé qu’en utilisant les données de précipitation interpolées de la base de données APHRODITE, développée à partir d’un réseau dense de stations météorologiques, les débits moyens journaliers des fleuves Dak Bla et Mekong au Vietnam pouvaient être simulés d’une façon adéquate. En fait, la densité du réseau de stations météorologiques utilisé lors de l’interpolation a une influence sur la précision des données interpolées, et par conséquent sur la précision des simulations hydrologiques réalisées à partir de ces données, même lorsque le modèle hydrologique utilisé est global (Duncan et al. 1993; Ruelland et al. 2008; St-Hilaire et al. 2003).

Le principal défi de l'interpolation spatiale des données de stations concerne les terrains à topographie complexe, tels que les régions montagneuses, où la variabilité spatiale de la précipitation est forte. En effet, dans de telles régions, à cause de la forte variabilité spatiale de la précipitation, les champs de précipitation sont difficiles à estimer avec précision, même par l'intermédiaire d'un réseau assez dense de stations météorologiques (Buytaert et al. 2006; Johnson and Hanson 1995). Cette forte variabilité spatiale de la précipitation est même susceptible d'influencer grandement la réponse hydrologique de petits bassins versants (Andréassian et al. 2001; Merz and Bárdossy 1998). La forte variabilité spatiale de la précipitation en terrains à topographie complexe est engendrée par l'effet orographique (Chow et al. 1988). Pour plus d'efficacité de l'interpolation en terrain à topographie complexe, des techniques d'interpolation spatiale ont été conçues en tenant spécifiquement compte de l'effet orographique. Certaines d'entre-elles se basent simplement sur la relation précipitation-altitude alors que d'autres sont plus complexes et tiennent compte de certaines caractéristiques du sol (pente, aspect, etc.) (Daly et al. 1994; Garen et al. 1994; Goovaerts 2000; Hay et al. 1998; Nakama and Risley 1993; Risley 1994). Mais, quelle que soit la technique utilisée, les quantités de précipitation interpolées en région montagneuse à partir d'un réseau de stations épars, sont souvent très incertaines (Mizukami and Smith 2012). Ces incertitudes ont un impact considérable sur la précision des simulations hydrologiques (Biemans et al. 2009; Fekete et al. 2004; Moulin et al. 2009). Par exemples, Muñoz et al. (2011) ont constaté que des biais secs dans des quantités de précipitation interpolées pour la région montagneuse des Andes au Sud-central du Chili s'étaient traduits, en simulation hydrologique, en sous-estimations significatives des débits de pointes observés en saison pluvieuse. Tozer et al. (2012) ont comparé trois bases de données interpolées de précipitation mensuelle en Australie, et ont constaté que non seulement ces bases de données étaient très différentes l'une de l'autre, mais aussi elles étaient très différentes des données de stations. En modélisation hydrologique, ces différences s'étaient traduites par d'importantes incertitudes sur les débits simulés.

En plus des incertitudes, les champs réels de variables météorologiques sont souvent lissés dans les données interpolées, et même sur-lissés dans les régions à forte variabilité spatiale de précipitation (Hofstra et al. 2010). Cela les rend inappropriées pour l'étude des événements extrêmes. Le lissage des données peut engendrer des erreurs significatives dans les analyses hydrologiques car les processus hydrologiques sont fortement non-linéaires (Vischel 2006).

Au total, bien que l'interpolation des données ponctuelles des stations météorologiques soit une approche intéressante pour améliorer la couverture spatiale des observations dans les régions où le réseau de stations est épars, la fiabilité des données interpolées dépend fortement de la densité du réseau de stations utilisé. Pour cette raison, dans les régions ayant une faible densité de stations et où la variabilité spatiale de la précipitation est élevée, l'utilisation des données interpolées en modélisation hydrologique peut s'avérer inadéquate. De plus, puisque l'interpolation spatiale n'est logiquement réalisable que lorsque des données de stations existent, alors dans des régions entièrement dépourvues de stations météorologiques, il serait impossible ou irréaliste d'envisager une quelconque interpolation spatiale des données. D'où l'importance d'autres approches qui préconisent l'utilisation de données indirectement dépendantes des données des stations météorologiques.

2.3 Utilisation des données obtenues par télédétection

À l'instar des données spatialisées par interpolation, les données provenant de la télédétection offrent une couverture spatiale complète d'un site ou d'une région donnée. Cependant, contrairement aux données interpolées, les données obtenues par télédétection ne découlent pas des données de stations, mais sont des mesures indirectes des variables météorologiques. La télédétection offre un aperçu général de la distribution spatiale et de la dynamique des processus hydrologiques, qui n'est généralement pas disponible avec les relevés terrestres.

Depuis plusieurs décennies, la télédétection a été utilisée dans de nombreuses études hydrologiques (Bastola and François 2012; Engman 1996; Engman and Gurney 1991; Fortin et al. 2001; Gleason and Smith 2014; Houser et al. 1998; Kite and Pietroniro 1996; Murray et al. 2013; Pandey 2013; Pietroniro and Leconte 2000; Schmugge et al. 2002; Schultz 1996; Tang et al. 2009; Xu et al. 2014). Selon plusieurs auteurs dont Engman (1999) et Schultz and Engman (2000), l'attrait de la télédétection en hydrologie est principalement lié au fait que : (1) les techniques de télédétection ont la capacité de mesurer l'information spatiale à l'opposé des données ponctuelles des stations météorologiques, et (2) ces techniques peuvent évaluer l'état hydrique de la surface du sol sur de vastes domaines. Généralement, les données de télédétection utilisées dans des applications hydrologiques proviennent soit des radars météorologiques, ou des satellites (Alsdorf and Lettenmaier 2003; D'souza et al. 1990; Dribault 2012; Puech 2000; Rango 1994; Rango and Shalaby 1998; Yilmaz et al. 2005).

Les radars météorologiques émettent des ondes radios ou impulsions électromagnétiques dans le spectre des fréquences micro-ondes, et ce dans toutes les directions, puis captent en retour le rayonnement réfléchi par les hydrométéores (gouttes d'eau, flocons, de neige, grêlons) (Lee et al. 2007), (figure 2.3). Certains types de radars permettent d'accéder à la forme des hydrométéores détectés, et donc de déterminer leur type (gouttes, flocons, grêlons) (Sauvageot 2000). Comparativement à un réseau classique de pluviographes, le radar présente un intérêt certain en hydrologie, car il permet d'estimer, avec une haute résolution, la variabilité spatio-temporelle de la précipitation dans un rayon de quelques dizaines à centaines de kilomètres (Hingray et al. 2009). Ceci offre des potentialités opérationnelles pour la prévision hydrologique, ainsi que pour la gestion de certains ouvrages hydrauliques (exemple, réseaux d'assainissement). Les données radar sont généralement utilisées pour soutenir les efforts de modélisation des bassins versants ayant une faible couverture de stations météorologiques, en particulier lorsque des pas de temps courts sont considérés (Price et al. 2014). En hydrologie, l'un des points critiques de la mesure radar de précipitation est de pouvoir transformer la réflectivité (Z) des cibles obtenue par le radar en intensité (R) de précipitation au niveau du sol (Fulton et al. 1998; Saltikoff et al. 2000).

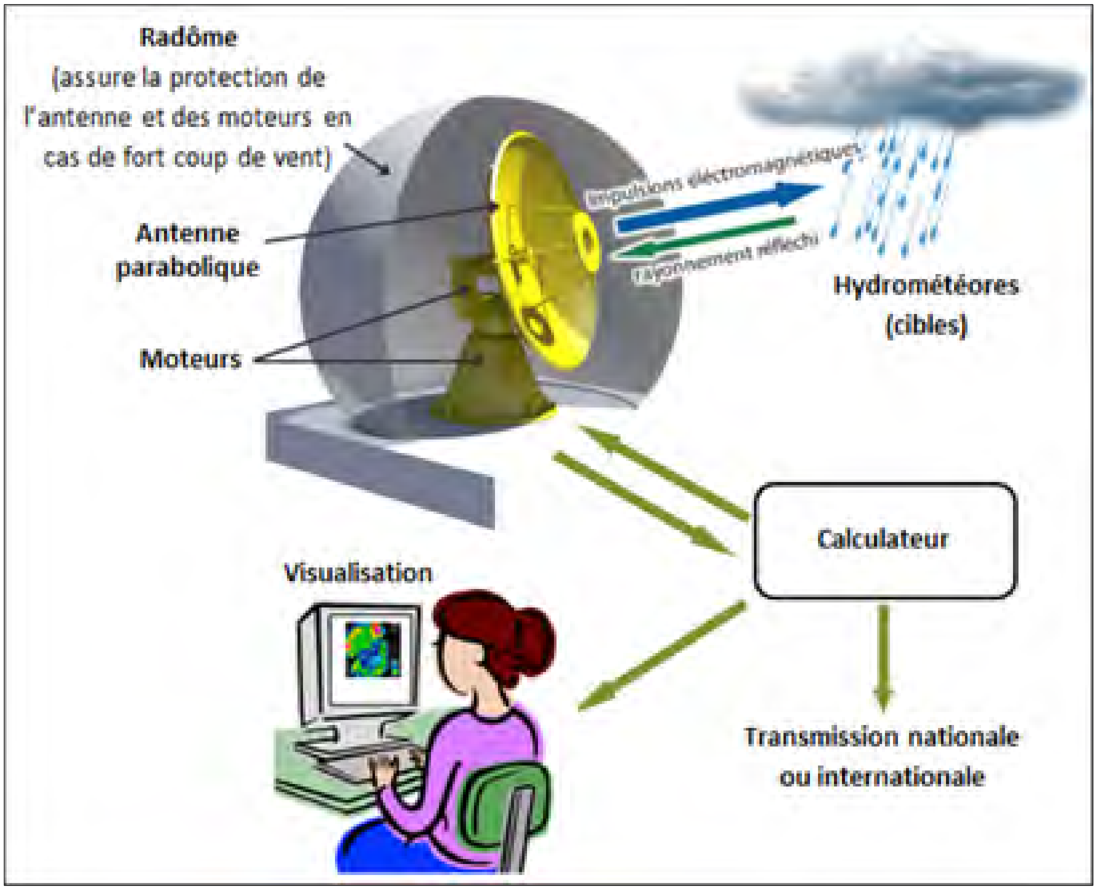

Figure 2.3 Fonctionnement du radar météorologique

Les satellites météorologiques effectuent des mesures sur un grand nombre de canaux, ce qui permet de les utiliser pour différencier divers phénomènes météorologiques (nuages, précipitation, vents, brouillard, etc.) (Rango and Shalaby 1998). La résolution spatiale des données satellite varie en général, entre 100km et moins de 5km (Forman et al. 2008). Les satellites offrent une alternative intéressante pour l'estimation de la précipitation surfacique dans les régions où le réseau de stations météorologiques est peu dense. De plus, les satellites météorologiques peuvent estimer la précipitation à des endroits inaccessibles pour les radars météorologiques, comme les régions à topographie complexe. De récentes études ont suggéré que les données estimées par satellite ont du potentiel en modélisation hydrologique. Par exemple, Artan et al. (2007) ont utilisé des estimations satellite de précipitation pour modéliser les débits d'écoulement issus de sous-bassins des fleuves Nil et Mékong et ont conclu que la précision des débits simulés était raisonnable. De même, les travaux de

Shrestha et al. (2008) ont montré que les débits du fleuve Bagmati au Népal pouvaient être simulés de façon raisonnable à l'aide d'un modèle hydrologique forcé par des données satellite de précipitation.

Malgré un certain succès, les données estimées par télédétection (radar et satellite) posent toujours des défis aux hydrologues et sont en général considérées comme peu fiables pour les études hydrologiques. Les observations satellite et radar ne sont pas utilisables directement; elles nécessitent toujours un traitement adéquat pour l'extraction de variables physiques convenables aux modèles hydrologiques (Hingray et al. 2009). De plus, les paramètres extraits des images satellites ou radar et ceux requis par les modèles hydrologiques sont rarement en accord, en précision et en échelles spatiale et temporelle (Puech 2000). Les données satellites et radar sont souvent entachées d'erreurs qui peuvent provenir de plusieurs sources (Joss and Germann 2000; Ramli and Tahir 2012). Par exemple, la rugosité et la végétation affectent l'émissivité du sol et peuvent ainsi, induire des erreurs dans les estimations de l'humidité du sol par satellite. En effet, la rugosité de la surface du sol augmente l'émissivité du sol et diminue la sensibilité à l'humidité du sol, réduisant ainsi l'amplitude de température de brillance entre des sols humides et des sols secs (Tran 2010; Van de Griend and Engman 1985). La végétation quant à elle, peut absorber une partie du rayonnement qui provient du sol et émettre des rayonnements elle-même (Walker 1999). Bien que la rugosité et la végétation affectent l'émissivité du sol, l'effet de la végétation est plus important car cette dernière peut totalement obscurcir la surface du sol si elle est présente en assez grande quantité. C'est la raison pour laquelle selon Rango (1994), le succès de la télédétection par satellite en ce qui concerne l'estimation de l'état hydrique des sols, est plus grand dans les régions arides et semi-arides où le couvert végétal clairsemé n'obscurcit que faiblement la surface du sol. Les données radar sont particulièrement sensibles aux phénomènes d'atténuation des ondes électromagnétiques (par les hydrométéores, les nuages, les gaz, ou la précipitation) (Delrieu et al. 2000), aux forts gradients de réflexivité avec l'altitude (Andrieu and Creutin 1995; Parent du Châtelet et al. 2005), et aux échos de terrain du fait d'obstacles tels que les bâtiments (surtout en ville), les arbres et les topographies accidentées (surtout en régions montagneuses) (Pellarin et al. 2002). L'obstruction des ondes électromagnétiques par de tels obstacles peut parfois dégrader de façon significative la portée et la précision des mesures radar, les rendant ainsi davantage incertaines (Warner et al. 2000; Westrick et al. 1999). Selon Hunter (1996), les volumes réels de précipitation sont souvent sous-estimés par les estimations radar, notamment en saison froide. Rasmussen et al. (2003) ont par ailleurs constaté que l'estimation des taux de chute de neige à partir des signaux radar engendre souvent des incertitudes supplémentaires à cause de la large gamme des formes de flocons de neige et de vitesses de chute. Il est par conséquent, toujours nécessaire d'ajuster les données radar avant leur utilisation pour les études hydrologiques car, la correction de biais améliore en général leur précision et leur performance en modélisation hydrologique (Krajewski and Smith 2002; Xiaoyang et al. 2003). Toutefois, cette correction repose fortement sur l'utilisation des mesures des stations météorologiques (Bastola and François 2012; Eleuch et al. 2010; Fulton et al. 1998; McKee and Binns 2015; Seo 1998; Steiner et al. 1999; Tridon 2011). Ainsi, la validation des données radar n'est possible que dans des régions où le réseau de stations météorologiques est de forte densité, surtout en présence de terrains à topographie complexe (Diederich et al. 2015; Eleuch et al. 2010). Bitew et al. (2012) soutiennent que la correction de biais est aussi nécessaire pour les données satellitaires de précipitation, avant leur utilisation dans des études de modélisation hydrologique. Pour Teo and Grimes (2007), les incertitudes associées aux hauteurs de précipitation estimées à partir de satellites sont toujours difficiles à évaluer correctement. C'est possiblement pour cette raison que Gagnon (2012) affirme qu'il n'existe pas de jeux de données de précipitation fiables estimées par satellite qui couvrent de vastes étendues sur une longue période de temps, à une résolution spatio-temporelle adaptée à l'échelle du bassin versant. Selon Tang et al. (2009), bien que la plupart des variables météorologiques soient observables par télédétection, les capteurs et les plates-formes actuels ne sont pas encore en mesure de fournir des observations hydrologiques cohérentes des composantes du bilan d'eau de surface, peu importe l'échelle spatiale considérée.

2.4 Utilisation des données estimées par combinaisons des observations de surface et des observations télédétectées

De telles combinaisons visent à exploiter à la fois la fiabilité des observations de surface et la couverture spatiale des données télédétectées. Cette approche permet de fournir aux modèles hydrologiques des données à une précision et une résolution spatio-temporelle adéquates (Krajewski and Smith 2002; Turk et al. 2008). Certains travaux récents ont montré que de telles données ont un fort potentiel en modélisation hydrologique. Par exemples, Xiaoyang et al. (2003) et plus récemment Yu et al. (2011), ont constaté que les hauteurs de précipitation issues de la combinaison entre données de télédétection et données de stations météorologiques permettaient d'obtenir une meilleure simulation des débits pour des bassins versants en Chine, comparativement à chacun des deux jeux de données considérés séparément. La principale faiblesse de telles données combinées est que souvent, elles reproduisent mal les extrêmes (Curtis et al. 2007).

2.5 Utilisation des données provenant des réanalyses atmosphériques

Une réanalyse atmosphérique (ou réanalyse météorologique) est une méthode utilisée pour obtenir un portrait exhaustif de l'état du système terrestre en utilisant le plus d'observations possible. Cette méthode consiste à combiner un schéma d'assimilation de données immuable et un modèle de prévision météorologique. Pour produire des prévisions météorologiques de qualité, le modèle de prévision de la réanalyse a besoin de connaître l'état de l'atmosphère et de la surface terrestre à l'instant initial de la prévision à produire. Pour cela, des observations de diverses sources et régions du globe sont intégrées dans le modèle de prévision grâce au cycle d'assimilation des données de la réanalyse, tous les 6 ou 12 heures. À la fin de ce cycle, le modèle de prévision établit le portrait le plus fidèle possible de l'atmosphère à l'instant initial, et un tel portrait est appelé « analyse ». Cette analyse est constituée de plusieurs dizaines de variables cohérentes sur la grille de calcul du modèle de prévision et regroupe deux catégories de variables : 1) les variables pour lesquelles le modèle a pu tenir compte des observations et 2) celles qui sont de purs produits du modèle. Le modèle de prévision utilise ensuite l'analyse établie pour initialiser la simulation qui permettra de produire les prévisions météorologiques de l'instant suivant (figure 2.4). Actuellement, environ 7 à 9 millions d'observations sont ingérés à chaque pas de temps dans les modèles de prévision des réanalyses (Bosilovich 2008; Poli et al. 2010a; Whitaker et al. 2009). Dans le système d'une réanalyse, les seuls éléments variables dans le temps sont les sources de données d'entrée brutes. Cela est inévitable dans la mesure où les réseaux d'observation sont en constante évolution (figure 2.5). Les principales sources d'observation sont les réseaux de mesure de surface, les radiosondes, les aéronefs et les satellites (figure 2.6) (Mesinger et al. 2006; Suarez et al. 2008; Wang et al. 2011). Les réseaux de mesure de surface comprennent les stations météorologiques, les bouées et les rapports de navires qui fournissent des données de surface pour les variables telles que la température, l'humidité, la pression, la direction et la vitesse du vent. Les radiosondes, les aéronefs et les satellites fournissent diverses données atmosphériques, tels que le rayonnement, le vent, l'humidité et la pression atmosphérique à différentes hauteurs.

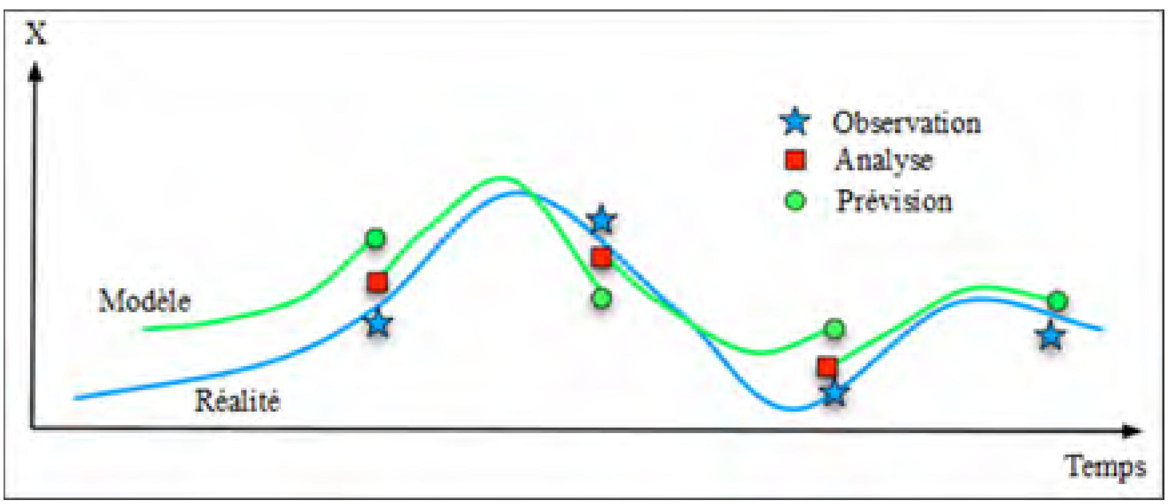

Figure 2.4 Illustration du processus de prévision du système d'une réanalyse

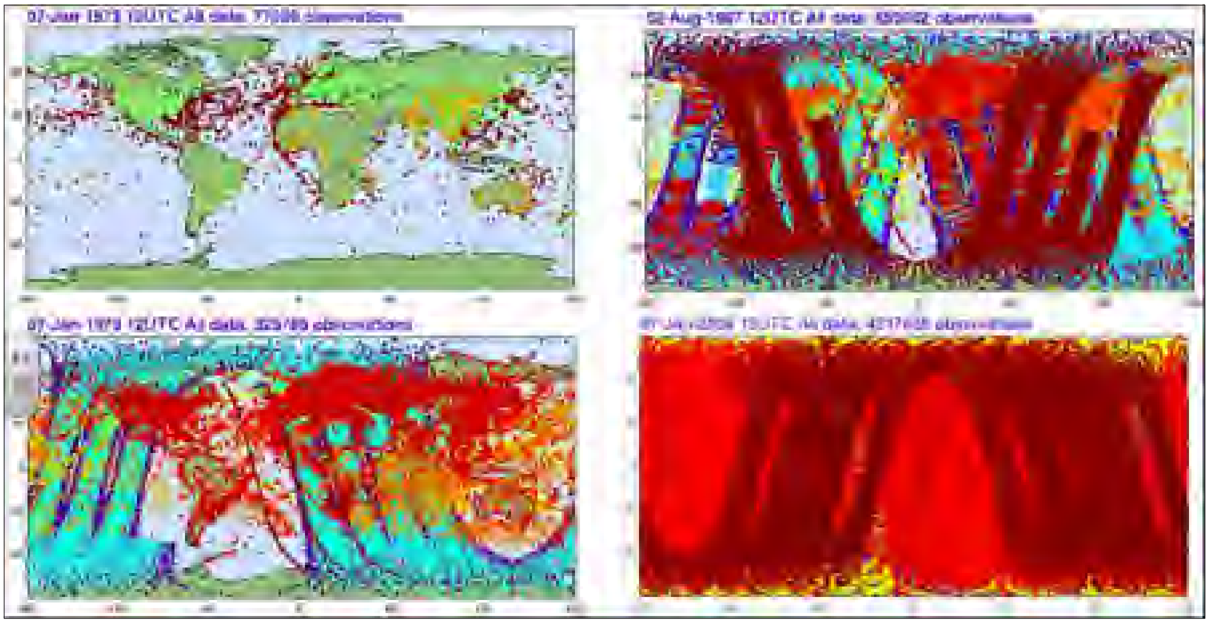

Figure 2.5 Évolution des systèmes d'observation de 1973 (pré-satellite) à 2006. Chaque couleur représente un système d'observation différent et les titres indiquent le nombre de points d'observation pour une période de 6 heures
au cours de chacune des années considérées
Adaptée de Bosilovich (2008)

Figure 2.6 Illustration des principales sources de données d’observation assimilées dans les réanalyses

Les principales forces des réanalyses atmosphériques sont : couverture spatiale généralement mondiale; résolution temporelle de quelques heures sur plusieurs décennies ; cohérence spatiale et temporelle des données ; résolutions de plus en plus fines et biais de plus en plus faibles ; estimations des données pour des centaines de variables météorologiques ; données de plusieurs réanalyses mises à jour en temps réel ou quasi-réel à l'usage du public ; utilisation de millions d'observations dans un système d'assimilation stable et utilisation relativement simple des données de réanalyses (du point de vue de leur traitement) (Kalnay et al. 1996; Rienecker et al. 2011; Uppala et al. 2005).

Les réanalyses présentent toutefois quelques points faibles dont les principaux sont : variation du nombre et de la qualité des observations assimilées pouvant introduire de fausses variabilités et tendances, de sorte que la fiabilité de certaines variables peut varier dans le temps et dans l'espace (Bosilovich et al. 2008; Lorenz and Kunstmann 2012), nouvelle analyse à chaque 6 ou 12h, et mélange observation/modèle pouvant entraîner la non fermeture du bilan d'eau (voir article en annexe I). La précipitation de surface de certaines réanalyses provient de leur modèle de prévision de sorte que les erreurs du modèle dans la représentation de la précipitation (Janowiak et al. 1998; Serreze and Hurst 2000; Trenberth and Guillemot 1998) peuvent se traduire en erreurs dans d'autres variables telles que l'évapotranspiration, le ruissellement et l'humidité du sol (Lenters et al. 2000; Maurer et al. 2001).

En dépit de ces faiblesses, les données provenant des réanalyses ont été largement utilisées dans les études climatiques incluant entre autres, la validation des modèles climatiques dans les régions où les données d'observation n'existent pas (Shiu et al. 2012; Sillmann et al. 2013a; Sillmann et al. 2013b), et la surveillance du climat (Bosilovich 2013; Lorenz and Kunstmann 2012; Manzanas et al. 2014; Rusticucci et al. 2014; Vose et al. 2012; Zhang et al. 2013). La conclusion générale de ces études est que les réanalyses sont bien utiles aux études climatiques à condition d'être utilisées avec les soins appropriés. Par exemple, les travaux de Nigam and Ruiz-Barradas (2006) ont montré que la variance spatiale de la précipitation de NARR (Mesinger et al. 2006) en été et en hiver était similaire à celle des observations aux

États-Unis, en raison d'une assimilation indirecte de la précipitation de surface dans la réanalyse NARR. Ces auteurs ont également constaté qu'en été, dans la région des Grandes Plaines, la précipitation dans NARR et celle observée ont des variabilités interannuelles similaires (Ruiz-Barradas and Nigam 2006). Rusticucci et al. (2014) ont montré que la variabilité interannuelle de la précipitation observée dans le sud des Andes centrales en Amérique du Sud est bien représentée dans la réanalyse ERA-Interim (Dee et al. 2011). Bien que généralement, la température et la précipitation des réanalyses se comparent bien aux observations, elles sont parfois significativement biaisées, notamment dans le cas de la précipitation.

Contrairement aux données estimées par interpolation spatiale ou par télédétection, le potentiel des données provenant des réanalyses pour les études de modélisation hydrologique a été peu exploré et peu de travaux sur le sujet ont été rapportés dans la littérature. Par exemples, Woo and Thorne (2006) ont utilisé des données de température et de précipitation des réanalyses ERA-40 (Uppala et al. 2005), NCEP-NCAR (Kalnay et al. 1996) et NARR (Mesinger et al. 2006) pour simuler l'apport de la fonte des neiges aux débits d'écoulement provenant d'un grand bassin versant de la région subarctique du Canada. Ils ont constaté qu'en raison d'un biais froid dans ces réanalyses, les pics de fonte des neiges étaient simulés avec retard. Choi et al. (2009) ont évalué l'utilité des données de précipitation et température de la réanalyse NARR pour la modélisation hydrologique de trois bassins versants dans le Nord du Manitoba au Canada. Ils ont constaté que les débits observés avaient été raisonnablement simulé en utilisant ces données. Vu et al. (2012) ont essayé de simuler les débits d'écoulement provenant d'un bassin versant du fleuve Dak Bla au Vietnam en utilisant des données de la réanalyse NCEP-NCAR, mais ont constaté que les débits simulés étaient significativement différents de ceux mesurés.

Il convient de noter que la plupart des travaux portant sur l'évaluation du potentiel des réanalyses en modélisation hydrologique n'ont examiné que des variables provenant d'«anciennes» réanalyses. La plupart de ces réanalyses ont évolué avec le temps, et des versions plus élaborées sont actuellement disponibles. Par exemple, la réanalyse ERA-

Interim (Dee et al. 2011) est une version améliorée de ERA-40 citée ci-haut. Dans le même sens, la réanalyse CFSR (Saha et al. 2010) est un produit plus récent du NCEP comparativement à la réanalyse NCEP-NCAR.

Au total, le potentiel des récentes réanalyses pour les études de modélisation hydrologique a été peu examiné dans la littérature. Un autre constat est que les études publiées sur l'utilisation des données des réanalyses pour la modélisation hydrologique sont toutes basées sur un nombre réduit de bassins versants (en général 1 à 5 bassins versants), et les conclusions qui en découlent sont difficiles à généraliser. Il serait plus pertinent de considérer un plus grand nombre de bassins versants dans les études, pour ainsi tenir compte de divers régimes hydroclimatiques. Il importe aussi d'évaluer l'impact de la combinaison des données des réanalyses avec celles des observations (données des stations météorologiques) sur la précision des simulations hydrologiques. Ce point n'a pas encore été abordé dans la littérature. Ainsi, bien que les réanalyses atmosphériques soient d'importantes sources des données météorologiques (en particulier pour les régions ayant peu ou pas de stations météorologiques), leur potentiel pour les études de modélisation hydrologique est encore très peu connu, et constitue de ce fait, une intéressante piste de recherche.

2.6 Objectifs de l'étude

2.6.1 Objectif général

L'objectif général de cette thèse est d'évaluer le potentiel des données de précipitation et température des réanalyses comme alternative aux données de stations météorologiques, pour les études de modélisation hydrologique. À travers cet objectif général, trois objectifs spécifiques sont visés.

2.6.2 Objectifs spécifiques

Objectif spécifique 1

Évaluer la précision des débits moyens journaliers simulés en utilisant des données de précipitation et température des réanalyses comme forçages météorologiques d'un modèle hydrologique global. Cet objectif a été atteint à travers l'article 2 (chapitre 4).

Objectif spécifique 2

Comparer pour diverses densités de stations météorologiques, la précision des hydrogrammes simulés à partir de données de précipitation et température des réanalyses à celle des hydrogrammes simulés à partir de données d'observation. L'article 3 (chapitre 5) répond à cet objectif.

Objectif spécifique 3

Évaluer l'impact de la combinaison de données de précipitation et température de réanalyses et celles des observations, sur la précision des débits moyens journaliers simulés. Cet objectif a été atteint à travers l'article 4 (chapitre 6).

2.7 Hypothèses de recherche

Pour explorer le potentiel des données de précipitation et température des réanalyses en modélisation hydrologique, les hypothèses suivantes ont été énoncées :

Hypothèse 1

Les biais des données de précipitation et température des réanalyses seraient suffisamment faibles pour permettre leur utilisation directe comme intrants aux modèles hydrologiques, sans un traitement au préalable.

Hypothèse 2

L'utilisation des données de précipitation et température des réanalyses en modélisation hydrologique, seraient plus pertinente que celle des données d'observation, dans les régions où la densité du réseau de stations météorologiques est faible.

Hypothèse 3

L'utilisation conjointe des données de réanalyses et des données d'observation contribuerait à améliorer la précision des simulations hydrologiques.

Pour atteindre les objectifs de cette thèse, un ensemble de plusieurs centaines de bassins versants localisés aux USA et au Canada a été considéré. En prélude au choix de la base de données d'observation à utiliser comme référence dans les analyses hydrologiques en vue (pour les bassins versants aux USA), les travaux présentés dans l'Article 1 (Chapitre 3) de cette thèse ont été réalisés. Les résultats ont conduit au choix de la base de données d'observation interpolée de Santa Clara.

De plus, quatre réanalyses atmosphériques parmi les plus récentes et dont les résolutions spatiales sont les plus fines, ont été considérées dans cette thèse. Ces réanalyses atmosphériques sont : European Re-Analysis Interim (ERA-Interim) (Dee et al. 2011), Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010), Modern Era Reanalysis for Research and Applications (MERRA) (Rienecker et al. 2011) et North American Regional Reanalysis (NARR) (Mesinger et al. 2006). Un résumé de leurs principales caractéristiques se présente comme suit :

European Re-Analysis Interim (ERA-Interim) (Dee et al. 2011)

- Réanalyse de 3^{ième} génération : utilise un long historique d'améliorations
- Réanalyse globale
- Schéma d'assimilation de qualité supérieure (4D-VAR)
- Assimilation de la température de l'air (à 2m du sol)
- Cycle hydrologique trop intense surtout au-dessus des océans

- En Arctique, biais chaud et trop humide dans la couche limite
- Ne capture pas les inversions dans les bas niveaux
- Sur la période 2003-2010, l'analyse de neige est affectée par une erreur de géo-location de l'information satellite du couvert de neige
- L'utilisation des observations de surface est limitée par la mauvaise représentativité près de la surface et les défauts de la méthode d'analyse
- L'analyse 4DVAR de l'atmosphère est faite séparément de l'analyse de surface qui est plus simple et qui utilise les observations près de la surface
- Dans le 4DVAR, les conditions de la surface sont fixées (alors, lorsqu'il pleut, la surface et l'humidité du sol ne changent pas)

Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010)

- Réanalyse de 3^{ième} génération : utilise un long historique d'améliorations
- Réanalyse globale
- Schéma d'assimilation de moins bonne qualité (3D-VAR)
- Assimilation des radiances des satellites
- Fine résolution spatiale
- Utilisation d'un modèle couplé atmosphère-océan
- Utilisation d'un modèle interactif glace de mer
- Analyse de surface réalisée à partir d'un modèle hydrologique utilisant la précipitation observée
- Prise en compte de la variation du CO_2, des aérosols et des variations solaires
- Analyses de la surface, de l'océan et de l'atmosphère effectuées séparément (analyse non couplée).
- Discontinuités sol profond et océan profond (Streams)

    Stream 1: 1 December 1978 to 31 December 1986
    Stream 2: 1 November 1985 to 31 December 1989
    Stream 5: 1 January 1989 to 31 December 1994
    Stream 6: 1 January 1994 to 31 March 1999

Stream 3: 1 April 1998 to 31 March 2005

Stream 4: 1 April 2004 to 31 December 2009

Modern Era Reanalysis for Research and Applications (MERRA) (Rienecker et al. 2011)

- Réanalyse de 3^{ième} génération : utilise un long historique d'améliorations
- Réanalyse globale
- Schéma d'assimilation de moins bonne qualité (3D-VAR)
- Assimilation des radiances des satellites
- Utilisation des observations de la NASA's Earth Observing System satellites et amélioration de la branche atmosphérique du cycle hydrologique.
- Intégrale verticale et incrément d'analyse fournis pour fermer le budget atmosphérique
- Changement du système d'observations au fil du temps, ce qui affecte les tendances (par exemple, la différence Précipitation – Évaporation augmente due à l'assimilation des radiances de satellites dès 1998)
- Précipitation trop intense en après-midi puis ré-évaporation immédiate entraînant faible infiltration

North American Regional Reanalysis (NARR) (Mesinger et al. 2006)

- Réanalyse de 2^{ième} génération
- Réanalyse régionale (couvre seulement l'Amérique du Nord)
- Schéma d'assimilation de moins bonne qualité (3D-VAR)
- Assimilation des observations de précipitation converties en chaleur latente (ce rend la précipitation modélisée durant l'assimilation similaire à celle observée)
- Inhomogénéités (temps et espace) de la précipitation assimilée
- L'analyse de la précipitation est désagrégée en une analyse à chaque heure
- Assimilation des vents à 10m
- Très fine résolution spatiale
- N'assimile pas la température de surface (température l'air à 2m du sol)

Une description complémentaire de chacune de ces réanalyses est présentée dans les articles 2-4 et dans l’annexe I.

CHAPITRE 3

ARTICLE 1. COMPARISON OF CLIMATE DATASETS FOR LUMPED HYDROLOGICAL MODELING OVER THE CONTINENTAL UNITED STATES

Gilles R. C. Essou[1], Richard Arsenault[1] and François P. Brissette[1]

[1] Département de Génie de la Construction, École de technologie supérieure, 1100 rue Notre-Dame Ouest, Montréal, Québec, Canada, H3C 1K3.

Article publié dans la revue « Journal of Hydrology » en 2016

3.1 Abstract

Climate data measured by weather stations are crucially important and regularly used in hydrologic modeling. However, they are not always available due to the low spatial density and short record history of many station networks. To overcome these limitations, gridded datasets have become increasingly available. They have excellent continuous spatial coverage and no missing data. However, these datasets are usually interpolated using station data, with little new information besides elevation. Furthermore, minimal validation has been done on most of these datasets. This study compares three such datasets covering the continental United States to evaluate their differences and their impact on lumped hydrological modeling. Three daily time step gridded datasets with resolutions varying between 0.25° and 1km were used in this study - Santa-Clara, Daymet and CPC. The hydrological modeling evaluation of these datasets was performed over 424 basins from the MOPEX database. Results show that there are significant differences between the datasets, even though they were essentially all interpolated from almost the same climate databases. Despite those differences, the hydrological model used in this study was able to perform equally well after a specific calibration to each dataset. While there were a few exceptions, by and large, Nash-Sutcliffe efficiency metrics obtained in validation were not statistically different from one database to the other for most basins. It appears that there are no reasons

to favour one dataset versus another for lumped hydrological modelling, and that these datasets perform just as well as using the original station data.

Keywords: Interpolated data, bias, hydrological modeling, calibration, performance comparison

3.2 Introduction

Climate data obtained from ground weather stations are the main inputs to hydrological models. However, spatial coverage of weather stations is often limited in mountain areas and low-population areas. In addition, short temporal coverage and missing data are typical of many station records.

To overcome these problems, many water management agencies have been using gridded datasets obtained by interpolating station data onto a regular grid. Such datasets have continuous spatial and temporal coverage and are much simpler to use than their station dataset counterparts. Several competing interpolation methods have been proposed. The simplest interpolate between stations (i.e. Thiessen polygons, simple kriging) (Hartkamp et al. 1999; Skaugen and Andersen 2010), whereas the more complex use additional information from other sources or integrate physical properties such as the atmospheric lapse rate. This is the case for local (Daly et al. 1994; Hasenauer et al. 2003; Taylor et al. 1997; Thornton et al. 1997) and regional regression methods (Chen et al. 2007; Mahdian et al. 2009; Perry and Hollis 2005). On the other hand, even though gridded datasets offer good spatial coverage, their reliability may be questionable in areas with a sparse weather station network (Mizukami and Smith 2012). Gridded datasets also contain uncertainties linked to each specific interpolation scheme (Tozer et al. 2012).

In the United-States (as in many other countries throughout the world) there has been a widespread effort to produce robust interpolated datasets. Several such datasets have been made freely available to the scientific community by several groups such as the University of

Santa-Clara (Maurer et al. 2002), the Climate Prediction Center (CPC) (Higgins et al. 2000), and Daymet (Thornton et al. 2012; Thornton et al. 1997). These databases have been used in various recent hydrological studies (Ali et al. 2014; Elsner et al. 2014; Gallo and Xian 2014b; McEvoy et al. 2014; Neiman et al. 2014; Singh et al. 2014; Ye et al. 2014; Zurita-Milla et al. 2014). However, to date, there exists little validation work as to the ability of these datasets for hydrological studies. Accordingly, the present study aims at comparing various precipitation and temperature gridded datasets at the watershed scale, and to evaluate the differences for lumped hydrological modeling.

3.3 Study area and datasets

3.3.1 Study area

The study area is a group of 424 catchments in the continental United-States, located from 67°W to 125°W longitude and 25°N to 50°N, as shown in figure 3.1. The catchments are dispersed in 5 climatic zones according to the Köppen-Geiger classification system (Kottek et al. 2006). There are 236 basins classified as Humid continental, 107 as Humid subtropical, 13 in the Marine west-coast region, 24 as Mediterranean and 44 as Semi-arid. The watersheds range between 66 km^2 and 10,325 km^2 in size.

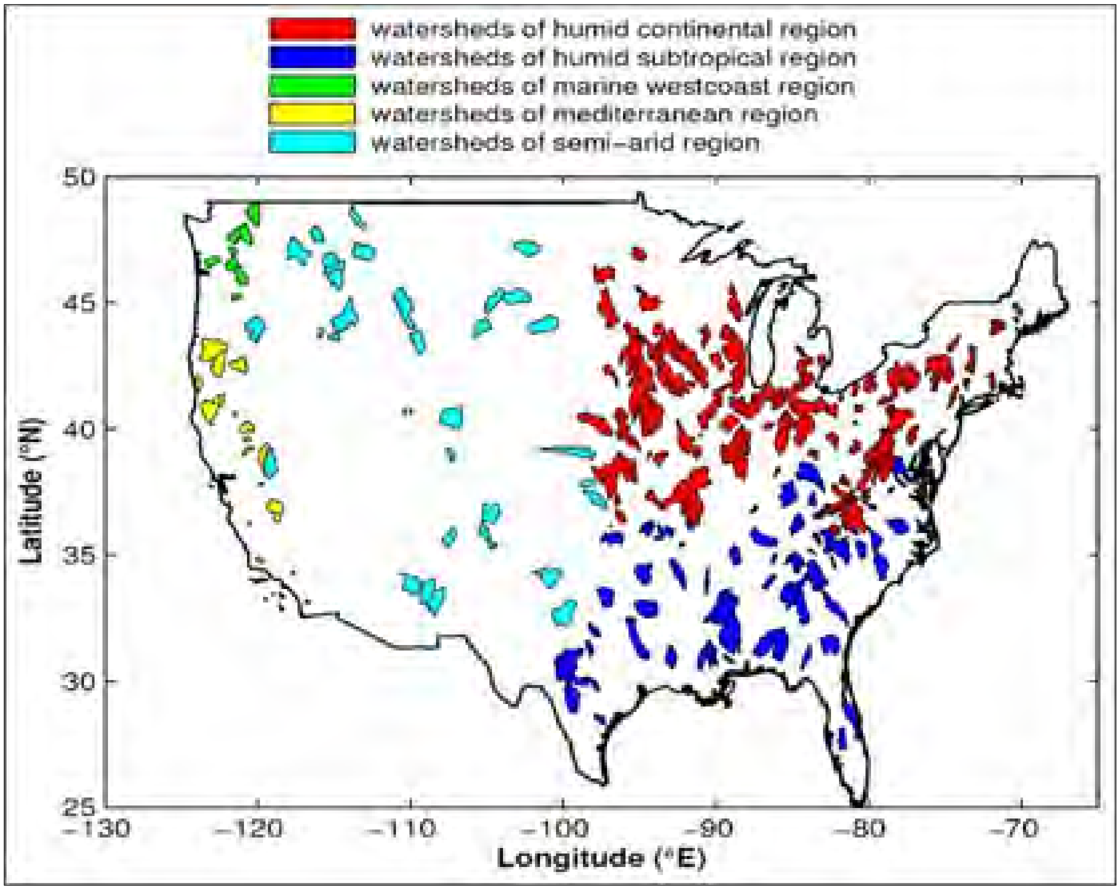

Figure 3.1 Location and climate classification of the 424 watersheds used in this study

3.3.2 Datasets

All the comparisons and simulations were performed with daily climate data as well as daily discharge time series. The four databases used in this study are as described below:

3.3.2.1 MOPEX area averaged data

The ***MO****del* ***P****arameter* ***E****stimation e****X****periment* (MOPEX) database contains precipitation, temperature (minimum and maximum), potential evaporation and streamflows on a daily time step. The database covers years 1949-2003. Its conception stems from the National Climatic Data Center (NCDC) weather station observations (about 16,139 stations) (Duan et al. 2006). The MOPEX climate data are averaged observation values on the different catchments. An inverse distance weighting method was implemented to estimate the final MOPEX climate data. A detailed description of this data source is available in Schaake et al. (2006). It is important to note that each catchment in the database requires a minimal density of weather stations, which is determined by the size of the catchment as explained in Schaake et al. (2000). Furthermore, only time series of length greater than 10 years were admitted in the database. The reference discharge data is also taken from this database. The MOPEX dataset is available online: ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data.

3.3.2.2 Santa-Clara gridded data

The University of Santa-Clara gridded datasets were initially developed in Washington, but they were formatted into their current form at the University of Santa-Clara. The daily precipitation and temperatures (minimum and maximum) are available for years 1949-2010. They were interpolated on a 0.125° x 0.125° grid using the weather measurement data provided by the *National Oceanic and Atmospheric Administration* (NOAA) cooperative network, averaging 1 station per 700 km^2 (Maurer et al. 2002). The interpolation algorithm is based on the *Synergraphic Mapping System* (SYMAP) by Shepard (1984) and implemented as proposed by Widmann and Bretherton (2000). Particularly, the precipitations were downscaled to correspond to the long-term means of the precipitations from the *Parameter-elevation Regressions on Independent Slopes Model* (PRISM) (Daly et al. 1994; Daly et al. 1997). More precisely, it relies on 12 monthly means for the 1961-1990 period, which are statistically adjusted to capture the local variations on complex terrain. The Santa-Clara dataset is available online: http://hydro.engr.scu.edu/ files/gridded_obs/ daily/ncfiles_2010.

3.3.2.3 Climate Prediction Center gridded data

The *Climate Prediction Center (CPC)* data contains precipitation data for years 1949-2013 with a spatial resolution of 0.25° x 0.25°. The interpolation uses three main sources of observation data (Higgins et al. 2000). The first is the CPC cooperative network stations for the 1996-1999 period (15622 stations). The second is daily observations from the NCDC for years 1948-1998 (approximately 16139 stations). The third is from the Hourly Precipitation

Dataset (HPD) (approximately 5933 stations) (Higgins et al. 1996). The interpolation uses the Cressman method (Cressman 1959). Information on the location of weather stations used to build the CPC data can be found at: http://www.cpc.ncep.noaa.gov /products/Precip_Monitoring/Figures/NAMS/NAMS_curr.p.gnum.gif. The CPC dataset is available online: http://www.esrl.noaa.gov/psd/ data/gridded/data.unified.daily.conus.html.

3.3.2.4 Daymet gridded data

The Daymet dataset includes maximum and minimum temperatures and precipitation on a daily scale for the period 1980-2013. They were produced using the Daymet suite, an ensemble of algorithms and software designed to interpolate and extrapolate values at grid points with a 1km x 1km resolution (Thornton et al. 2012). Daymet uses observation network data to perform the interpolation with a Gaussian weighting scheme. A detailed description of Daymet is available in Thornton et al. (1997). Information on the location of weather stations used to build the Daymet data can be found at: https://daymet.ornl.gov/ overview.html. The Daymet dataset is available online: http://daymet.ornl.gov/.

A summary of the dataset characteristics is presented in table 3.1.

Table 3.1 Characteristics of datasets used in this study

Dataset	Spatial resolution	Temporal domain	Variables	Interpolation algorithm/ technique	Source	Reference
MOPEX	---	1949-2003	P, Tmin, Tmax, PE, Q	Inverse distance	ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data	(Duan et al. 2006)
Santa-Clara	0.125° x 0.125°	1949-2010	P, Tmin, Tmax, W	synergraphic mapping system (SYMAP)	http://hydro.engr.scu.edu/files/gridded_obs/daily/ncfiles_2010	(Maurer et al. 2002)
CPC	0.25° x 0.25°	1949-2013	P	Optimal interpolation (OI)	http://www.esrl.noaa.gov/psd/data/gridded/data.unified.daily.conus.html	(Higgins et al. 2000)
Daymet	1km x 1km	1980-2013	P, Tmin, Tmax, Srad, Vp, Swe	---	http://daymet.ornl.gov/	(Thornton et al. 2012; Thornton et al. 1997)

P=Precipitation, Tmin=minimum temperature; Tmax=maximum temperature; PE=potential evaporation, Q=streamflow; W=wind speed; Srad=shortwave radiation; Vp=vapor pressure; Swe=snow-water equivalent.

3.4 Methodology

3.4.1 Dataset comparison

The interpolated data grid points located inside each of the catchments were averaged using the inverse distance weighting method calculated with respect to the catchment centroid (Dirks et al. 1998). This method was shown to be amongst the best interpolation methods for such uses (Baillargeon et al. 2004; Ruelland et al. 2008). The comparison was performed on daily, seasonal and extreme data. Moreover, the daily data was compared by climatic zone.

The first comparison criterion used in this study is the well-known Root Mean Squared Error (*RMSE),* which is defined as:

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}(\boldsymbol{X_i} - \boldsymbol{Y_i})^{\mathbf{2}}} \tag{3.1}$$

where $\boldsymbol{X_i}$ and $\boldsymbol{Y_i}$ represent data values for day i, from X and Y datasets, and N is the length of the time series.

The RMSE gives an indication on the average difference amplitude between two series. An RMSE value of 0 is a perfect fit, and larger values indicate larger differences.
The second comparison criterion is the bias ($\boldsymbol{B}$), defined as:

$$\boldsymbol{B} = \frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}(\boldsymbol{X_i} - \boldsymbol{Y_i}) \tag{3.2}$$

The bias allows estimating how much one series underestimates or overestimates a second series. A positive bias indicates an overestimation of the observations, while the opposite is true for negative biases.

For temperature, biases were directly computed between each of the datasets. For precipitation, since the number of datasets is higher, biases were computed against a reference value equal to the average value of the 4 datasets in the studies. This was done to simplify the interpretation of results.

The third criterion is the correlation between the daily time series. It is simply defined by the linear correlation coefficient.

The fourth criteria for the comparative analyses are intended to gain insight in comparing extreme values. They are the 99th percentile of daily precipitation (mm/day), the 99th percentile of daily maximum temperature (°C) and 5th percentile of daily minimum temperature (°C).

3.4.2 Hydrological model

The hydrological model used in this study is the HSAMI model (Fortin 2000; Minville et al. 2008). It is a lumped conceptual rainfall-runoff model developed and used operationally by Hydro-Québec for over 30 years. It is used to predict streamflow values on over 100 catchments in the province of Québec on hourly and daily time scales. The HSAMI model has also been used extensively in streamflow prediction applications, climate change impact studies and rainfall-runoff modeling research projects (Arsenault and Brissette 2014b; Arsenault et al. 2013; Chen et al. 2011a, 2012; Chen et al. 2011b; Minville et al. 2008, 2009; Poulin et al. 2011). It simulates the main hydrological cycle processes such as vertical and horizontal water transfer, evapotranspiration, snowmelt and soil freezing. It has up to 23 parameters that must be calibrated: 10 for the various production function processes, 5 for the horizontal transfer through reservoir-type soil layers, 2 for evapotranspiration and 6 for snow-related processes. There are four interconnected reservoirs that contribute to the vertical water transfer balance: Snow on ground, surface runoff, saturated soil layer and unsaturated soil layer. The horizontal water transfer is based on two unit-hydrographs (one for surface runoff and one for underground runoff) and a linear reservoir. HSAMI requires spatially averaged minimum and maximum temperatures as well as rainfall and snowfall depths. The cloud cover fraction and snow on ground may also be used if they are available.

Because of the large number of catchments, an automatic optimization algorithm was chosen to perform the model calibrations. Arsenault et al. (2014) showed that the CMAES

(Covariance Matrix Adaptation Evolution Strategy) (Hansen and Ostermeier 1996, 2001) algorithm was the optimal choice for calibrating the HSAMI model on 10 catchments, 8 of which were from the MOPEX database. Thus the CMAES optimization algorithm was used to perform the many calibrations in this project.

The calibration metric was computed on the odd years and cross-validated on the even years, and vice-versa. This allowed taking into account any climatic trends (such as decadal or multi-decadal natural variability) or modifications in underlying data from the addition or removal of weather stations. However there is a drawback to this method: the model must be run for the entire period in order to select the odd years for calibration, thus doubling the computational requirements compared to traditional block-type calibration. Also, 10 calibrations were performed in the odd/even approach, as well as 10 other calibrations in the even/odd approach, for a total of 20 calibrations. Only the best parameter set was taken for each case. This reduces the likelihood of having the calibration algorithm not converge during the optimization process.

The Nash-Sutcliffe Efficiency (NSE) metric (Nash and Sutcliffe 1970) was used to compare hydrologic simulation performance levels. Other metrics could have been used, but the NSE is the most widely used metric and was the obvious choice for this study.

The NSE values were compared and the non-parametric Wilcoxon test was used to identify statistically significant differences in between results (Rakotomalala 2008).

Furthermore, the precipitation and temperature datasets were then mixed and recombined to produce a total of 12 distinct datasets, and the calibration, validation and comparison aspects were also performed on the newly created datasets. Table 3.2 shows all of the resulting datasets used in this study.

From table 3.2, it is clear that the common period to all groups is 1980-2003. For this reason the entire study will be performed with these years to avoid any biases that could be caused by using different periods between the datasets.

Table 3.2 List of datasets used in this study and coverage periods

Components		**Period**
Temperatures	**Precipitation**	
MOPEX	MOPEX	1949 – 2003
Santa-Clara	Santa-Clara	1949 – 2003
MOPEX	Santa-Clara	1949 – 2003
Santa-Clara	MOPEX	1949 – 2003
MOPEX	CPC	1949 – 2003
Santa-Clara	CPC	1949 – 2003
Daymet	Daymet	1980 - 2003
Daymet	MOPEX	1980 - 2003
Daymet	Santa-Clara	1980 - 2003
Daymet	CPC	1980 - 2003
MOPEX	Daymet	1980 - 2003
Santa-Clara	Daymet	1980 - 2003

3.5 Results

3.5.1 Temperature comparison

3.5.1.1 Mean daily temperature

The results of the RMSE, bias and correlation coefficients between the mean daily temperature values of the MOPEX, Daymet and Santa Clara datasets are presented in figure 3.2. From figure 3.2A, it can be seen that comparatively to Daymet, the Santa-Clara mean daily temperatures deviate more from the MOPEX daily temperatures as approximately 71% of the catchments reflect a higher RMSE for the Santa-Clara dataset. However, the Santa-Clara and Daymet values are closer to one another than with MOPEX in the sense that RMSE values are smaller on 83% of the watersheds.

With respect to mean daily temperature, when compared to MOPEX, the Santa-Clara dataset has a median bias of -0.2°C, whereas Daymet shows a median bias equal to -0.1°C (figure 3.2B). However, both datasets have a cold bias relatively to the MOPEX dataset on the majority of catchments (75% and 65% of watersheds respectively). As a general rule, Santa-Clara and Daymet temperatures are colder than MOPEX, but Daymet is globally warmer than Santa-Clara.

The three datasets are strongly correlated to one another with correlation coefficients between 0.93 and 1 for all basins (figure 3.2C).

There are, however, some statistically significant differences between the temperature datasets. The Wilcoxon test (95% confidence interval) showed that the MOPEX dataset is different from its Santa-Clara and Daymet counterparts on 38% of the watersheds. Daymet and Santa-Clara are statistically different from one another in 36% of watersheds.

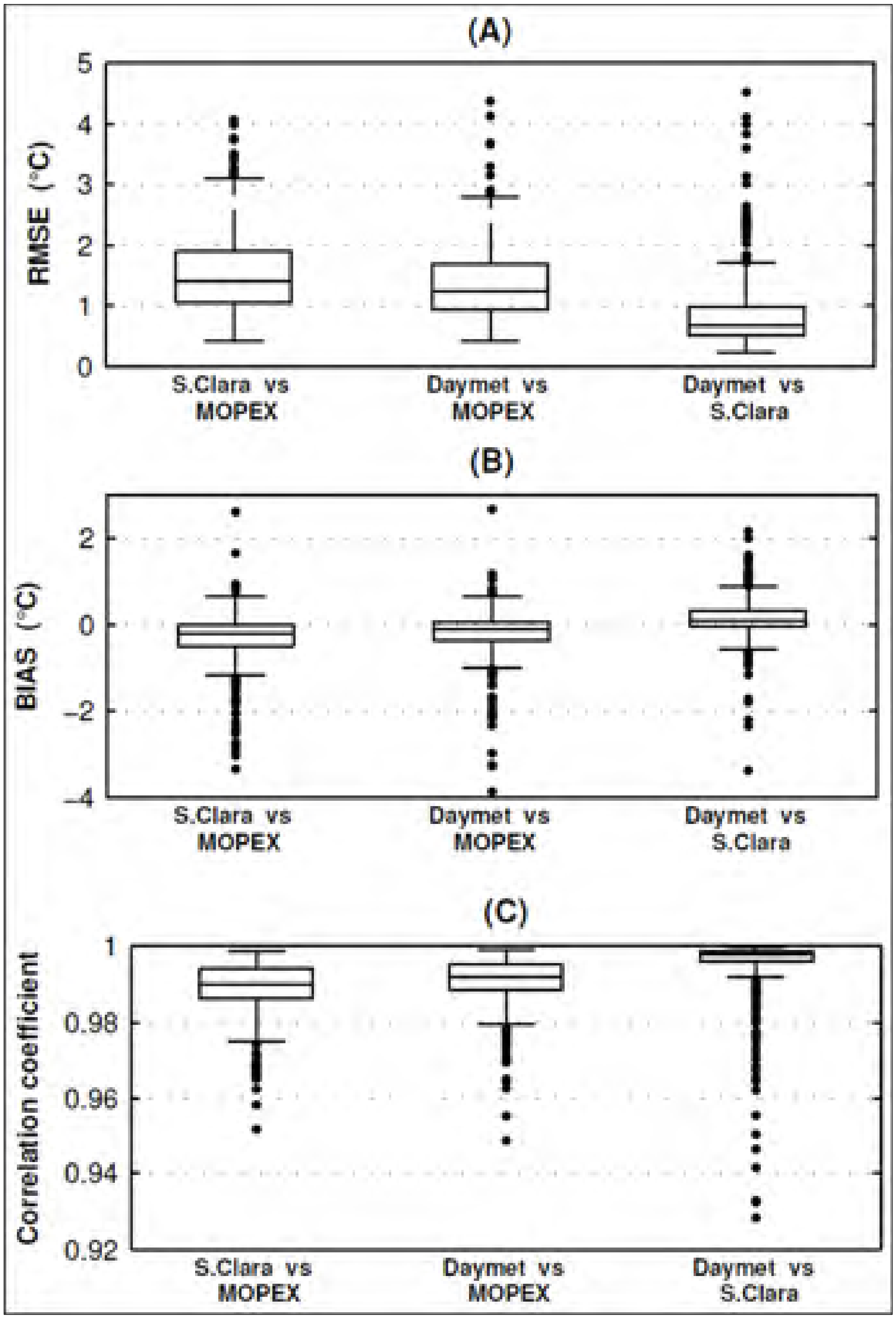

Figure 3.2 RMSE (A), bias (B) and correlation coefficients (C) between the mean daily temperatures of the Santa-Clara, Daymet and MOPEX datasets. The lower and upper limits of each boxplot represent the 25th and 75th percentiles, respectively. The middle line represents the median (50th percentile). The limit values of the whiskers correspond to (u+2.7σ) and (u-2.7σ) where u=average of the plotted points and σ=standard deviation. The outliers are points higher or smaller than the whiskers limits

3.5.1.2 Mean daily temperature by climatic zone

The results of the RMSE between the mean daily temperature values of the MOPEX, Daymet and Santa-Clara datasets for each of the climatic zones are presented in figures 3.3A – 3.3E. Results clearly show that RMSE values between Santa-Clara and MOPEX are larger than the ones between Daymet and MOPEX for all climatic zones with the exception of the Mediterranean catchments. Moreover, for the Santa-Clara dataset, the RMSE in Semi-arid climate are relatively larger (median = 2.2°C) but the RMSE in Humid subtropical climate are lower (median = 1.1°C). As for Daymet, the largest RMSE values were found in the Mediterranean region (median = 1.7°C), and the lowest, in the Humid subtropical climate (median = 1.0°C). In all climatic zones, with the exception of the oceanic and Semi-arid zones, Daymet and Santa-Clara temperatures are more similar to one another than to the MOPEX dataset. In the oceanic and Semi-arid zones, Daymet values are closer to MOPEX than Santa-Clara.

The results for the bias are presented in figures 3.3F–3.3J. The results show that in all the climatic zones, the Santa-Clara temperature biases relative to the MOPEX temperature are mainly cold (median bias <0°C). However, these biases are colder in the Marine/west-coast climate region (median = -1.4°C) and approximately nil in the Humid subtropical climate (median = -0.01°C). The Daymet biases (when compared to MOPEX) are mainly cold as well in all climate zones except for the Humid subtropical climate where the median bias is also approximately nil (median = 0.02°C). The Mediterranean climate is relatively colder with a median bias of -0.6°C. In all climate zones, the Daymet biases are mostly warmer than Santa-Clara, and particularly so in the oceanic climatic zone (mean bias = 1.03°C).

Generally, the differences between the three datasets are smaller in the Humid continental and subtropical climatic zones. The Wilcoxon test indicates that in those two climatic zones, the MOPEX dataset differs from the other two datasets in only 30% of the basins, compared to 66-100% in the other climatic zones.

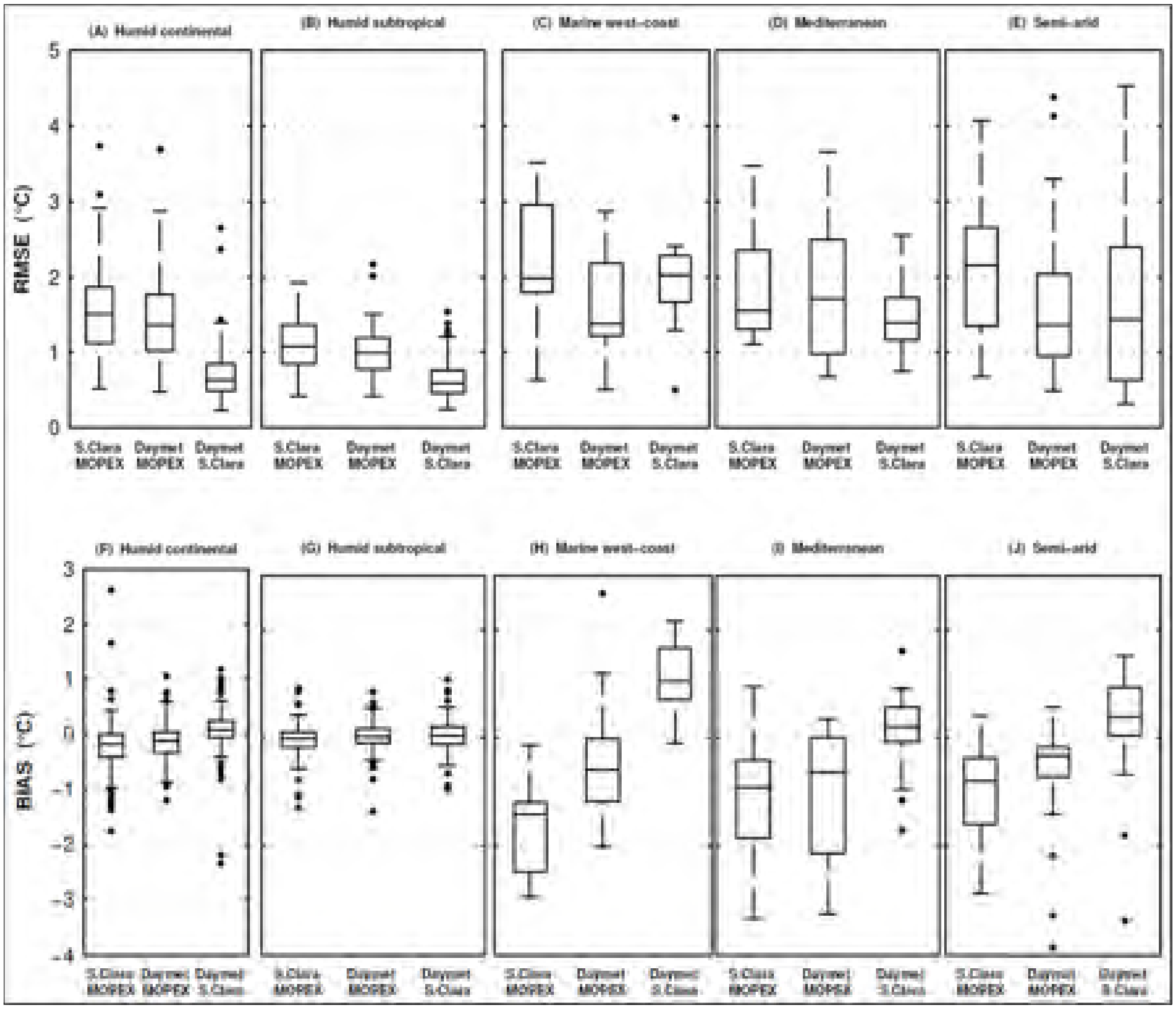

Figure 3.3 Mean daily temperature RMSE and bias for the Santa-Clara, Daymet and MOPEX datasets for the 5 climate zones

3.5.1.3 Mean seasonal temperatures

Results are similar for seasonal temperatures and are not shown. RMSE values between mean seasonal temperatures are relatively small for both Santa-Clara and Daymet when compared to MOPEX. However, for all seasons, Santa-Clara displays larger RMSE values than Daymet when compared to MOPEX. Also, in both cases, the temperature RMSE values are generally larger in winter (median RMSE = 0.3°C) and lower in summer (median RMSE = 0.1°C).

As expected, mean seasonal biases follow the cold biases of the daily mean temperature for both Daymet and Santa-Clara, and for all seasons. The coldest biases are experienced in winter (median = -0.4°C for Santa-Clara and -0.2°C for Daymet) and the least cold biases in the summer (median = -0.06°C for Santa-Clara and -0.01°C for Daymet). In all cases, biases related to the MOPEX seasonal temperatures are colder for the Santa-Clara dataset than for Daymet.

3.5.1.4 Extreme temperatures: 99th percentile of daily maximum temperatures and 1st percentile of daily minimum temperatures

The relative differences of temperature extremes for all three datasets are presented in figure 3.4. For the daily maximum temperature (figure 3.4A) the median biases are relatively small in all cases. When compared to MOPEX, the 99th percentiles of daily maximum temperatures of Santa-Clara have a warm bias on 57% of basins (median bias of 0.1°C) whereas those of Daymet have a cold bias on 65% of basins (median of -0.1°C). This implies a cold bias for the 99th percentiles of daily maximum temperatures of Daymet compared to their Santa-Clara counterparts on 74% of the basins (median of -0.2°C). Globally, the Santa-Clara dataset has the highest maximum temperatures, followed by MOPEX, with Daymet having the lowest maximum temperatures.

For the daily minimum temperatures (represented with the 1st quantile), when compared to MOPEX, Santa-Clara and Daymet show differences between -13.0°C and 4.0°C. These biases are cold on respectively 91% and 88% of basins, with median biases of -3.1°C and -2.7°C. In other words, MOPEX minimum extreme temperatures are much warmer whereas the Santa-Clara and Daymet datasets tend to produce lower minimum temperatures. Compared to Santa-Clara, Daymet minimum temperatures show biases between -1°C and 1°C on 80% of the basins.

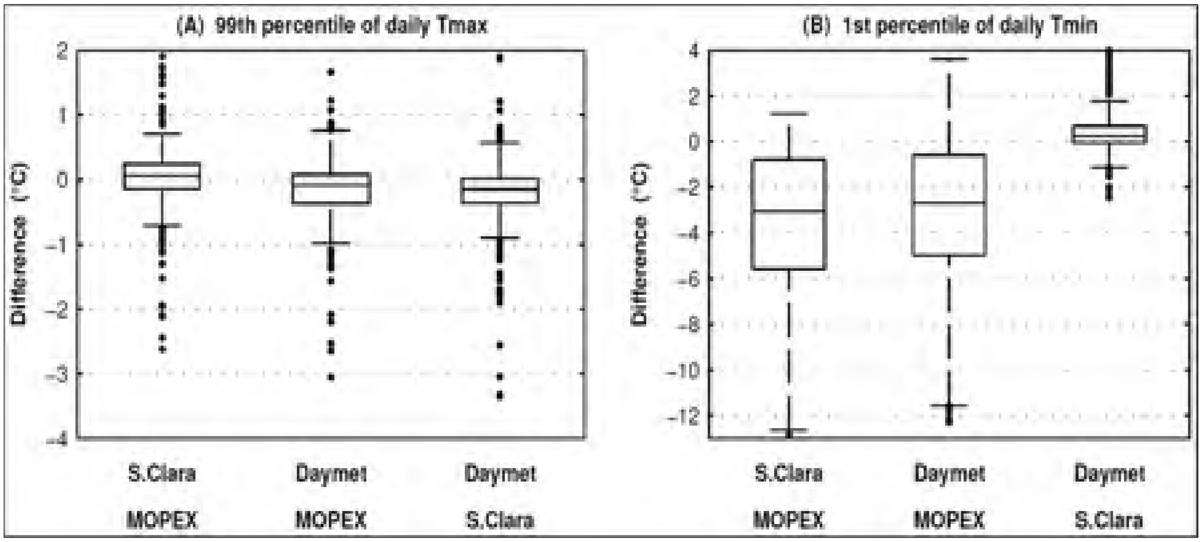

Figure 3.4 Differences between extremes temperatures of the Santa-Clara, Daymet and MOPEX datasets

3.6 Precipitation comparison

3.6.1 Daily precipitation

The results of the daily precipitation RMSE, bias and correlation coefficients between the MOPEX, Santa-Clara, Daymet and CPC datasets are presented in figure 3.5. Results in figure 3.5A show that the, Santa Clara precipitation is most similar to the MOPEX reference. On 97% of the catchments, the daily precipitation RMSE of the Santa-Clara dataset is lower than that of the CPC dataset. In turn, the CPC daily precipitation RMSE is lower than for Daymet in 75% of the catchments.

With respect to the mean of all four datasets (used as a reference), mean daily precipitation of the MOPEX, Santa-Clara and CPC show dry biases on respectively 75%, 66% and 70% of basins with median values of -1.3%, -0.7% and -1.2% (figure 3.5B). At the other end of the spectrum, Daymet values have a wet bias on 86% of catchments with a median bias of 3.1%.
Overall, The CPC, Santa-Clara and Daymet datasets are wetter than the MOPEX dataset.
Correlation coefficients of daily precipitation between MOPEX and the other datasets display varying levels of correlation. Figure 3.5C shows that the daily precipitation values are the

most correlated to the MOPEX reference dataset. Furthermore, Daymet and CPC show very strong correlation, indicating that the precipitation is treated similarly in both interpolation methods.

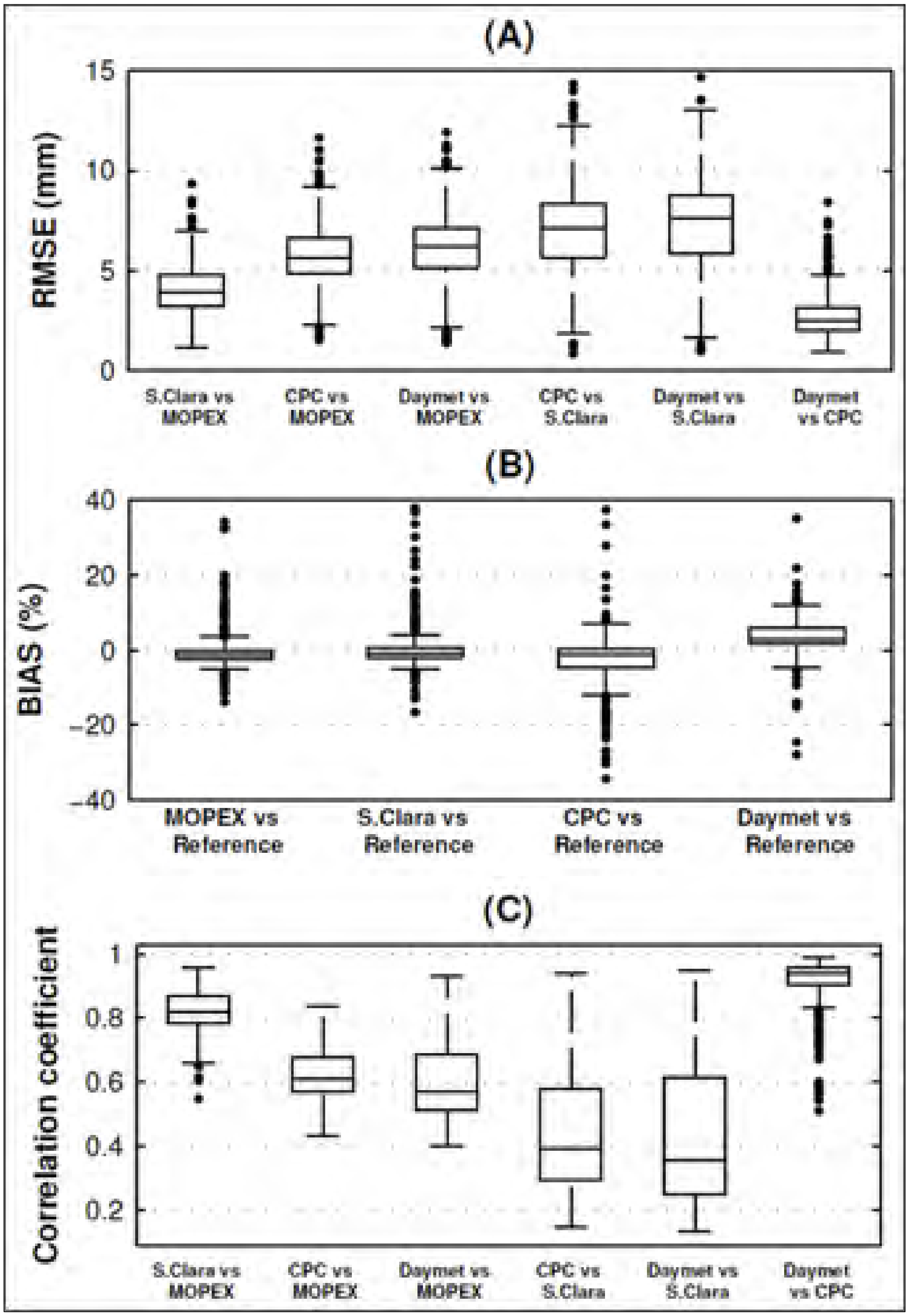

Figure 3.5 RMSE (A), bias (B) and correlation coefficients (C) of the daily precipitation of the MOPEX, Santa-Clara, Daymet and CPC datasets

3.6.2 Daily precipitation by climatic zone

The RMSE between daily precipitation datasets are presented by climatic zones in figures 3.6A–3.6E. In general the tendencies discussed in the preceding section mostly apply to all climate zones. In particular, in the Humid continental and subtropical zones, Santa-Clara and MOPEX are closer to one another. All four datasets are in best agreement in the Semi-arid climatic zone (median RMSE of 2.5mm).

The biases of daily precipitation are presented in figures 3.6F–3.6J. As mentioned earlier, biases are compared against reference values computed as the mean of all four datasets. Biases are smaller in the Humid continental and subtropical zones. In both of those zones, biases are mostly dry for MOPEX (median bias = -1.8% and -1.1%), Santa-Clara (-0.9% and -0.5%) and CPC (-1.1% and -0.8%). Conversely, Daymet has mostly humid biases for the same two climatic zones (3.8% and 2.5%). For all other climatic zones, MOPEX is the most humid dataset and CPC is the driest.

Correlation coefficients of daily precipitation between all four datasets are presented in figures 3.6K–3.6O. An analysis of the results reveals very similar conclusions to those of RMSE. This can easily be seen as figures 3.6K–3.6O are practically mirror images of figures 3.6A–3.6E. All four datasets are in best agreement in the Oceanic and Mediterranean climatic zones (median correlation coefficient of 0.67). The correlation coefficients are lower than 0.5, especially between CPC and Santa-Clara, and between Daymet and Santa-Clara, in the Humid continental and subtropical zones. In these two regions, Daymet and CPC correlate well, possibly because both use the NCDC weather stations network which is quite dense in the eastern United States. The Santa Clara database is built from a different weather stations network, and that may explain why it weakly correlates with Daymet and CPC. In the Humid continental and subtropical regions, precipitation is unevenly distributed over the year and most rainfall occurs as convective storms in the summer because of the tropical atmospheric flow from the Gulf of Mexico. These local events may be differently represented in Santa Clara and Daymet (or CPC) because of the use of different weather stations networks, and lead to low correlation. Although in general Santa Clara is not well correlated with Daymet and CPC, the correlation coefficients in CPC and Santa-Clara, and in Daymet and Santa-Clara, are overall higher in the Oceanic and Mediterranean regions, where precipitation has a low spatial variability because it is influenced by the proximity to the Pacific Ocean.

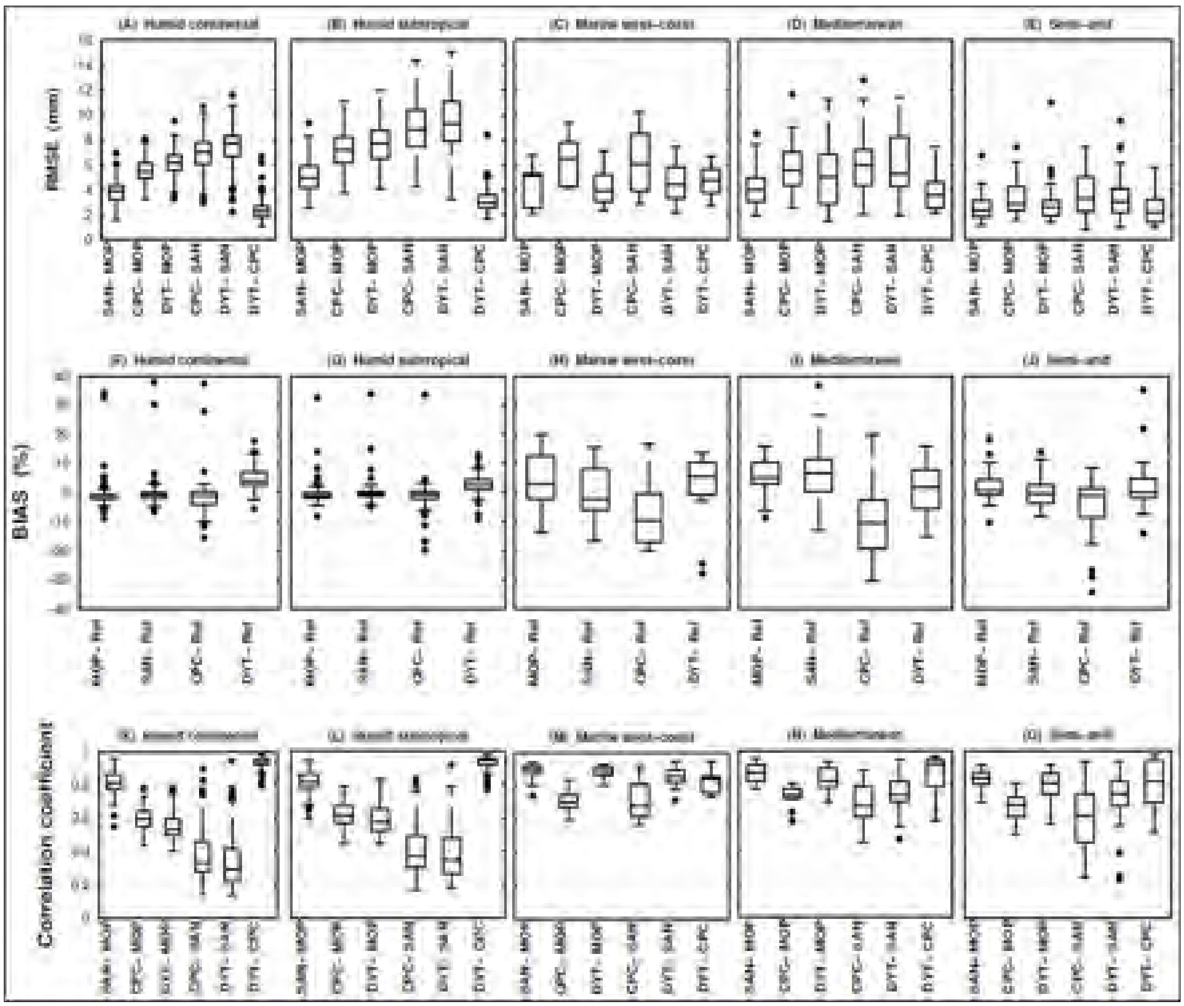

Figure 3.6 Daily precipitation correlation coefficient (A-E), RMSE (F-J) and bias (K-O) for all the datasets for the 5 climate zones. (**MOP = MOPEX; SAN = Santa-Clara; DYT = Daymet; Ref=Reference** (mean of all four datasets))

3.6.3 Total seasonal precipitation

Trends for seasonal precipitation are similar to annual ones (results not shown). When compared to MOPEX, for all seasons, the smallest RMSE belongs to the Santa-Clara dataset followed by CPC and Daymet. CPC and Daymet are closest to one another. RMSE values for all databases are larger in the summer and lower in the winter. The median biases of total seasonal precipitation compared against reference values (the mean of all four datasets) are shown in table 3.3. Results indicate that biases are mostly dry in winter, spring and fall for MOPEX, Santa-Clara and CPC. In the summer, CPC biases are small and mostly humid whereas MOPEX and Santa-Clara datasets are dry. Daymet is the wettest dataset with wet biases for all seasons.

Table 3.3 Median biases of total seasonal precipitation compared against the mean of all four datasets (%)

	MOPEX	Santa Clara	CPC	Daymet
Winter	-0.7%,	-1.1%	-2.1%	4.1%
Spring	-1.6%	-0.7%	-2.6%	5.1%
Summer	-1.9%	-0.6%	0.1%	2.3%
Fall	-1.6%,	-1.0%	-2.4%	4.6%

3.6.4 Extreme precipitations: 99^{th} percentile of daily precipitation distribution

The distributions of extreme precipitation biases between the four datasets are presented in figure 3.7. Results show that when compared to MOPEX, the biases of extreme precipitation for the Santa-Clara, CPC and Daymet datasets have median values respectively equal to 0.7%, -3.4% and 2.6%. These biases are humid on respectively 55%, 28% and 64% of basins. This implies that extreme precipitations from the Santa-Clara and Daymet datasets are larger overall than those of the MOPEX dataset.

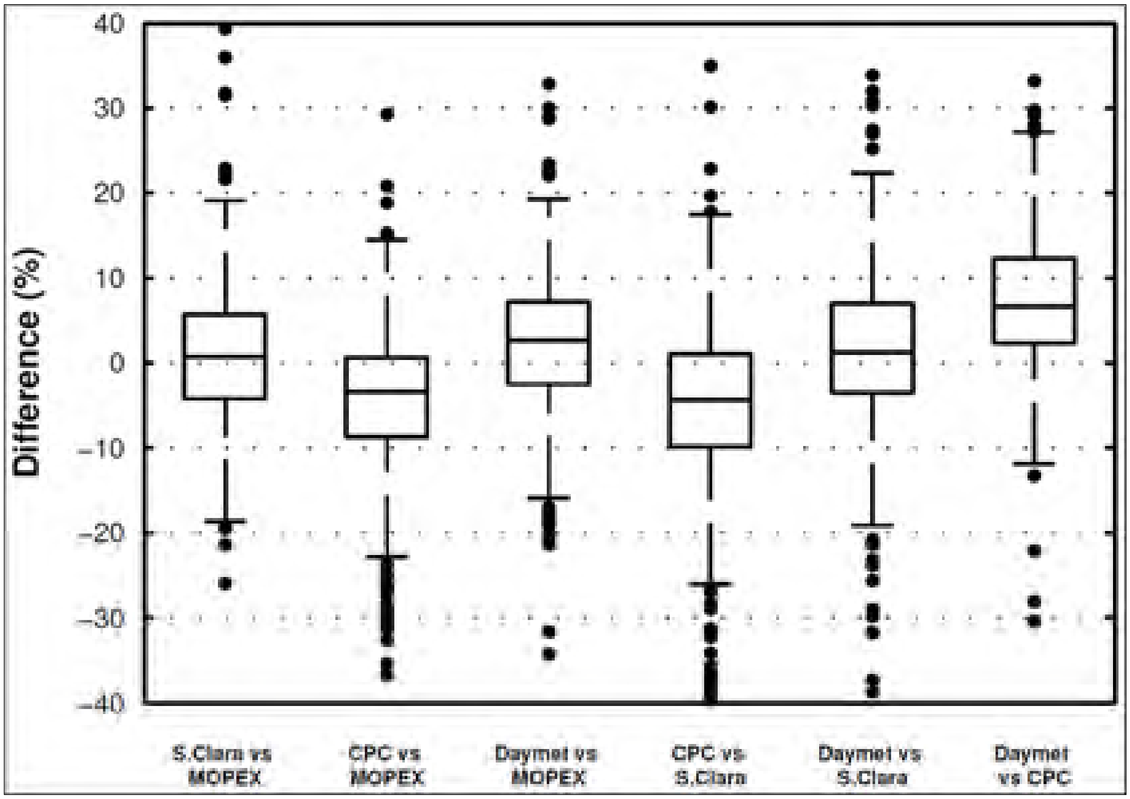

Figure 3.7 Differences between the 99th percentile of daily precipitation distribution of the MOPEX, Santa-Clara, CPC and Daymet datasets

3.7 Hydrological performance

The performance of the HSAMI hydrological model is first assessed using the MOPEX database. Results are shown in figure 3.8 and indicate that the hydrological model performs reasonably well, with a NSE median value of 0.783. The model performs well over most of the United States with the exception of the semi-arid climate (see figure 3.1) where several catchments have a NSE value inferior to 0.6. This is not surprising considering that the hydrological model used in this study was developed for temperate climates and is not well adapted to the specific conditions of more arid landscapes. However, since the goal of this study is an inter-comparison of datasets, this relative lack of performance in semi-arid regions is of minimal concern.

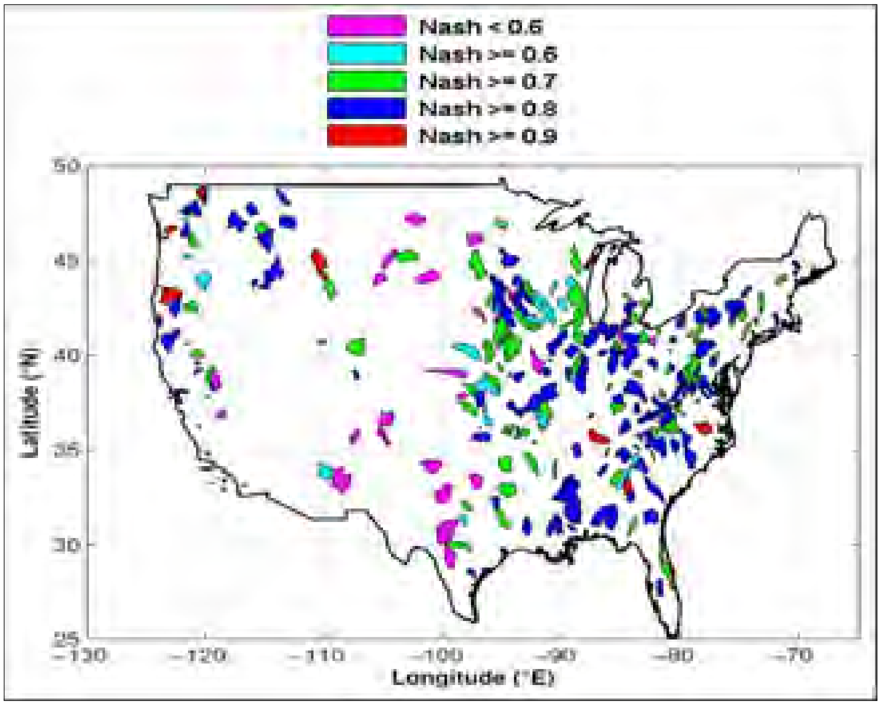

Figure 3.8 Validation results (NSE) of the HSAMI hydrological model using the MOPEX database (Flow discharge, precipitation and temperature)

The distribution of hydrological model performances using the various datasets is presented in figure 3.9.

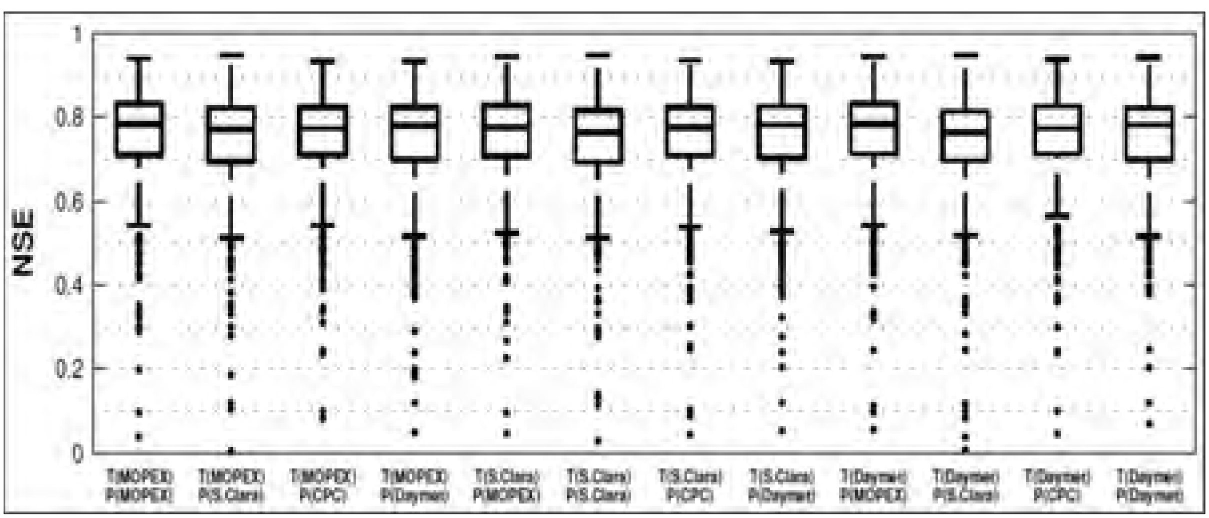

Figure 3.9 Validation NSE distributions for the 12 climate datasets

It is clear from that figure that the performance level is similar overall. Median validation NSE values from the MOPEX, Santa-Clara and Daymet datasets were found to be respectively equal to 0.783, 0.762 and 0.780. For all hybrid combinations of precipitation and temperature data, the median NSE values were between 0.763 and 0.783. A comparison was made catchment-by-catchment to determine the frequency with which each climate combination shows superior performance. The results are shown in table 3.4.

Table 3.4 Frequency with which each climate combination shows superior performance

	PRECIPITATION (P)				
TEMPERATURE (T)	MOPEX (%)	Santa-Clara (%)	CPC (%)	Daymet (%)	**Total (%)**
MOPEX	14.07	4.77	6.03	10.30	**35.17**
Santa-Clara	7.79	7.79	8.04	11.81	**35.43**
Daymet	5.28	8.04	6.53	9.55	**29.40**
Total (%)	**27.14**	**20.60**	**20.60**	**31.66**	**100**

Table 3.4 indicates that all datasets perform at a very similar level. Still, it indicates that the $T_{(MOPEX)}$-$P_{(MOPEX)}$ dataset performs better on average, followed by $T_{(S.Clara)}$-$P_{(Daymet)}$.

A Wilcoxon test was performed between each of the groups in figure 3.9 to determine which ones were statistically different. Results reveal statistically significant differences between some combinations of temperature and precipitation datasets:

1- The combinations $T_{(MOPEX)}$-$P_{(MOPEX)}$ and $T_{(Daymet)}$-$P_{(MOPEX)}$ are statistically different from all combinations using Santa-Clara's precipitation ($T_{(MOPEX)}$-$P_{(S.Clara)}$, $T_{(S.Clara)}$-$P_{(S.Clara)}$ and $T_{(Daymet)}$-$P_{(S.Clara)}$);

2- The Santa-Clara dataset ($T_{(S.Clara)}$-$P_{(S.Clara)}$) differs from any combination dataset obtained by substituting $P_{(S.Clara)}$ with either $P_{(MOPEX)}$, $P_{(CPC)}$ or $P_{(Daymet)}$;

3- The combination $T_{(S.Clara)}$-$P_{(S.Clara)}$ is different from any dataset containing Daymet precipitation.

These results unsurprisingly indicate that precipitation datasets are more critical than temperature datasets for hydrological modeling.

The performance of all 12 combination datasets was also analyzed with respect to seasonal discharge and annual maximum discharge. The results for seasonal values (not shown) are similar to those presented in figure 3.9. All datasets perform better in winter (median NSE values between 0.626 and 0.737) and spring (median NSE values between 0.716 and 0.759), but not as well in summer (median NSE values between 0.631 and 0.705) and fall (median NSE values between 0.546 and 0.694). The different seasonal performances may partly be due to the different seasonal biases of precipitation (as shown in table 3.3). Furthermore, they may be caused by the different hydrological regimes that prevail and that vary from a season to another. For example, in spring, the streamflows are mainly influenced by snowmelt and are more easily simulated because they vary gradually. In summer, the streamflows are mainly influenced by rainfall which has high spatial and temporal variability, thus making modeling more challenging.

Further analyses based on catchment size and climate zone classifications were also performed. Following these tests, it was shown that basin size had no impact on the relative performances of the groups, while the climate dataset only played a role only on the Mediterranean climate basins. The NSE distributions for the 12 climate datasets on the Mediterranean climate catchments are presented in figure 3.10. It can be seen in figure 3.10A that for the 24 Mediterranean climate catchments, using Daymet precipitation results in much better simulations, independently of the temperature datasets used. The spread is also smaller. The MOPEX precipitation is the least adequate for this climate zone resulting in a lower median performance value and a larger spread. It is not clear as to why this is the case. These catchments are located in mountainous regions, but so are the catchments from the west coast climate zone which do not exhibit a similar pattern.

At the seasonal scale, the only noticeable differences between the 12 datasets were all observed in the Mediterranean climatic zone, and only in winter and spring. In this climatic zone, the best spring and winter modeling results always used the CPC precipitation, independently of which temperature dataset was used (figures 3.10B and 3.10C).

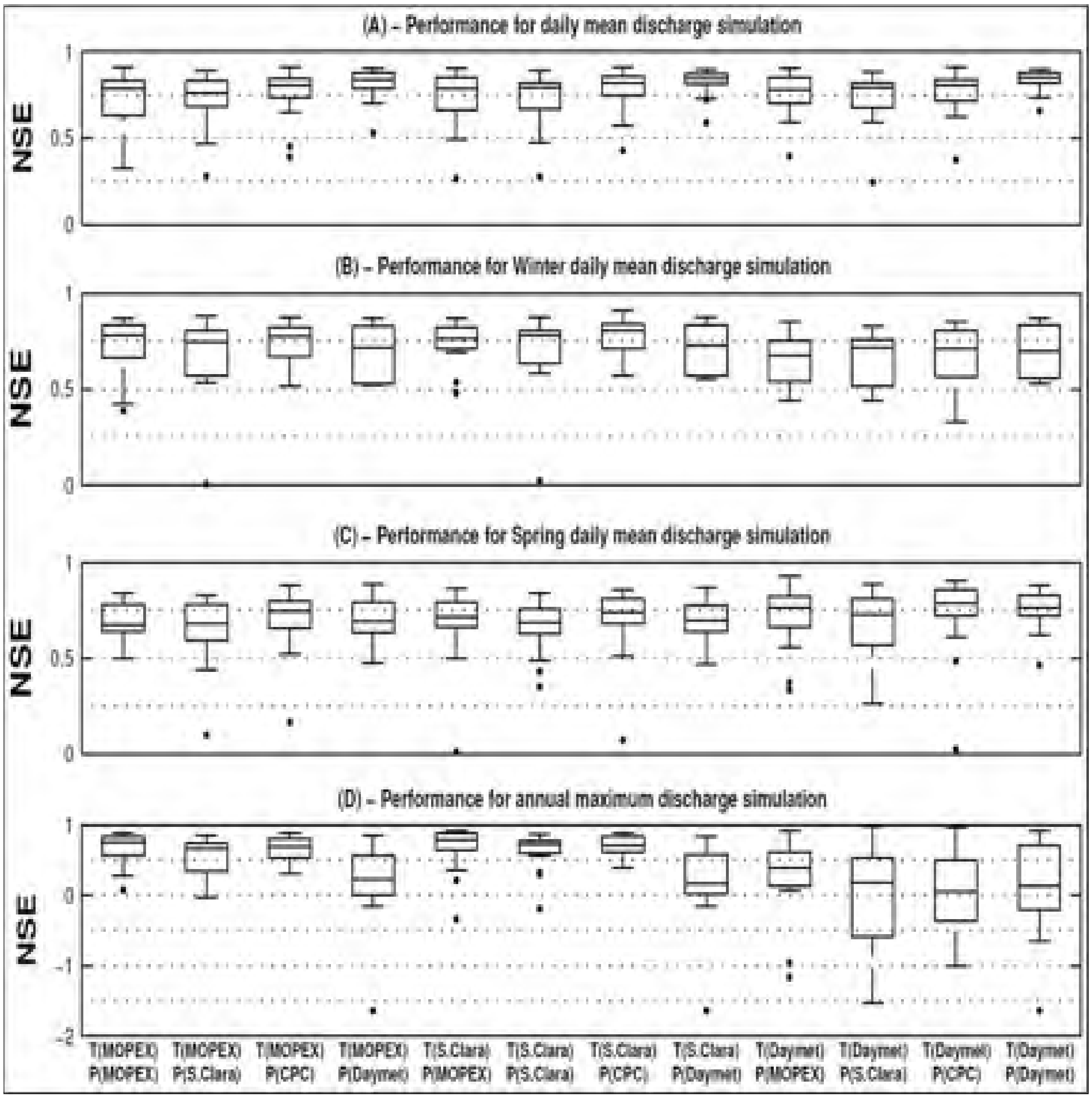

Figure 3.10 Validation NSE distributions on the Mediterranean catchments for the 12 climate datasets. There are 24 catchments under a Mediterranean climate

3.8 Discussion

While weather station networks remain the most important source of information for hydrological modeling, their often low spatial resolution can sometimes lead to unrepresentative and poor model performance (Arsenault and Brissette 2014a). The need to improve this resolution has been the driving force behind gridded and interpolated climate datasets. However, such datasets have limitations with respect to hydrological modeling (Mizukami and Smith 2012; Muñoz et al. 2011).

Gridded datasets have the important advantage of having no missing data and the potential ability to generate valuable information in areas not densely covered by weather stations, especially when taking into account external variables such as elevation (Tapsoba et al. 2005). On the other hand, interpolating algorithms are limited in this potential ability, and «spreading» very sparse station data onto a fine grid may results in artifacts not anchored in any real physics. The uncertainties resulting from the interpolation algorithm manifest themselves in the sometimes large differences between the datasets. The variability in the contributing observational networks also plays a role in generating variability in the gridded datasets. These differences sometimes remain large even at the basin scale as seen in this study.

To shed light on these issues, fours interpolated datasets were compared in this study with an emphasis on hydrological modeling. By mixing the 4 precipitation and 3 temperature datasets, flow discharge was simulated on the 424 catchments of the MOPEX database using the HSAMI hydrological model, resulting in 12 flow discharge time series for each catchment. A common 24-year period (1980-2003) was used for all datasets.

Results indicate that there are differences, sometimes significant, between all four datasets. They all display biases when compared amongst themselves. There is a good agreement between datasets for mean daily temperatures, especially in the Humid continental and subtropical climatic zones. For mean daily temperatures, there are two distinct grouping with

MOPEX-Santa-Clara and CPC-Daymet being close to one another. The sheer number of climate stations and their attributes makes it all but impossible to find correlations between underlying observation climate data and the final gridded product. Moreover, most publicly available weather stations are used by all datasets (some use a subset of the entire set, while others used all available data). It is beyond the scope of this paper to explain the differences between the gridded products; readers are encouraged to read (Duan et al. 2006; Higgins et al. 2000; Maurer et al. 2002; Thornton et al. 2012) for more details on each dataset. However, the differences between the gridded products may largely be attributable to the interpolation schemes which differ substantially from one dataset to another.

Despite the observed differences, the use of each dataset as the driving meteorological input to a lumped hydrological model led to equally good modeling results. Consequently, within the limits of this study, all datasets appear to be similar and equally good for hydrological studies.

The resolution of the gridded dataset and the complexity of the interpolation scheme do not appear to have any effect in the results. This is likely partly due to the fact that a lumped model was used in the assessment and that all grid points were averaged at the catchment scale, perhaps hiding some potential advantages of the higher-resolution dataset. It is possible that advantages of higher resolution grids could be uncovered using distributed models on the larger catchments. But this would be a time-consuming and computationally-intensive task, especially to set-up and calibrate distributed hydrological models on a large number of catchments. In this study 101760 (424 catchments x 12 hybrid datasets x 20 odd/even year calibrations) individual model calibrations were performed. This would be a daunting task for a complex distributed hydrological model, even on a subset of the catchments used in this study.

In this work, precipitation and temperature datasets were mixed and matched to form 12 different combinations. No ill-effects were observed in doing so, presumably because precipitation and temperature datasets are usually interpolated independently. As such, there is likely little physical coherence between values of precipitation and temperature in interpolated datasets. This is an aspect that could be better investigated through a comparison against high-resolution climate model or reanalysis of data, where physical consistency between datasets should arguably be much better preserved.

Using statistics averaged over the 424 catchments, this study showed that all gridded datasets behaved similarly for hydrological modeling. However, this study could not evaluate the impact of network density even though it is one of the most interesting scientific problems. The MOPEX database contains catchment-averaged temperature and precipitation data. Information about the number of stations used to generate the catchment-averaged data (which would be needed to estimate network density for each catchment) is not present in the database. Only a rough density estimate can be calculated from the MOPEX dataset requirements, which state “desired” minimums of 1, 2, 3, 6 and 12 stations for basins less than 1, 10 , 100, 1000 and 10000 square miles respectively. Network density could also be estimated using the existing NCDC stations. However, since watersheds in the MOPEX database were contributed by many different parties, such an estimation would be error-prone since stations from the CPC cooperative network could also have been used in some catchment and not in some others. Questions related to network density, such as whether or not gridded datasets offer benefits in areas with poor station coverage (as opposed to densely-covered regions where all datasets are expected to converge) would be better tackled using a small subset of carefully chosen watersheds for which precipitation and temperature data would be recalculated using NCDC stations for example.

Also worth noting is that the results are mostly similar from one climate region to the next, except in the Mediterranean climate zone where some differences are visible. However we must take into account the number of catchments in each zone. There are 24 Mediterranean and 13 Marine/West-coast catchments, whereas there are 343 catchments in the humid regions. The comparison between these groups is illustrative at best since there are an insufficient number of catchments for proper statistical significance testing in the small groups.

An advantage of using gridded datasets is that they are much easier to use than station data. They have uniform coverage and no missing data. Catchment-averaging can be done using a simple arithmetic mean instead of using weight-based averaging as is commonly done, with weights constantly changing depending on which stations are reporting data on any given day. However, gridded datasets are not available in real-time, or near real-time like station data. As such they cannot be used in forecasting mode unless the interpolation is also done in near real-time. This is a process that is now done in-house by many water resources managers, but not yet available to the general public. It is however foreseeable that such data will be available in the near future. For example, such a product is currently in development by Environment Canada (Choi et al. 2013).

Finally this study opens the door to a more in-depth investigation of other gridded datasets. For example, more complex datasets such as PRISM (Daly et al. 1994; Daly et al. 1997) and even reanalysis datasets could be included in such a study. Reanalysis datasets offer the advantage of a much larger set of variables that could be useful for hydrological modeling.

3.9 Conclusion

This study compared four different interpolated precipitation and temperature datasets (MOPEX, Santa-Clara, Daymet and CPC), the last three being interpolated on a regular grid. The comparison was based on basin-averaged data. Their performance in hydrological modeling over 424 catchments in the continental US was analyzed. The spatial heterogeneity of the catchments allowed comparing the HSAMI model performance relative to catchment size and climate attributes.

The comparison was two-fold. First, the climate characteristics were compared to one another with various metrics, and the correlation coefficients, RMSE and bias were compared between the groups. It was shown that there are non-negligible biases between the interpolated datasets for many catchments. Second, each interpolated dataset was used as direct input to a specifically calibrated hydrological model. Although there are important differences between the various precipitation and temperature datasets, their hydrological performances in validation was not statistically different for most of the watersheds. It appears that there is no reason to favour one dataset versus another for lumped hydrological modeling, and that these datasets perform just as well as using the basin-averaged original station data.

3.10 Acknowledgements

This work was funded through a Natural Science and Engineering Research Council collaborative research grant (NSERC-CRD) with Hydro-Québec, Rio-Tinto-Alcan, Ontario Power Generation and the Ouranos consortium on regional climatology and adaptation to climate change as industrial partners. These industrial research partners are greatly acknowledged for their direct and indirect contributions to this work. We also sincerely thank all the individuals and institutions that developed the datasets used in this work, and made them available to the scientific community. We hope that this work is a small contribution to their important effort.

CHAPITRE 4

ARTICLE 2. CAN PRECIPITATION AND TEMPERATURE FROM METEOROLOGICAL REANALYSES BE USED FOR HYDROLOGICAL MODELING?

Gilles R.C. Essou[1], Florent Sabarly[1], Philippe Lucas-Picher[2], François Brissette[1] and Annie Poulin[1]

[1] Département de Génie de la Construction, École de technologie supérieure, 1100 rue Notre-Dame Ouest, Montréal, Québec, Canada, H3C 1K3.
[2] Département des Sciences de la Terre et de l'Atmosphère, Université du Québec à Montréal, 405 Rue Sainte-Catherine Est, Montréal, Québec, Canada, H2L 2C4.

Article publié dans la revue « Journal of Hydrometeorology » en 2016

4.1 Abstract

The sparse coverage of weather stations over several regions of the world limits the ability of hydrological models to simulate river flows. The lack of stations could be filled by global meteorological reanalyses, but their use for hydrological modeling has barely been investigated. This paper investigates the potential of reanalyses for use as proxies of observed surface precipitation and temperature to force hydrological models. To that end, three global atmospheric reanalyses (ERA-Interim, CFSR, MERRA), one regional reanalysis (NARR) and one global meteorological forcing dataset obtained by bias-correcting ERA-Interim (WFDEI) were compared to one gridded observation database over the contiguous US. Results showed that all temperature datasets were similar to the gridded observation over most of the USA. On the other hand, precipitation from all three global reanalyses was biased, especially in summer and winter in south-eastern USA. The regional reanalysis precipitation was closer to observations since it indirectly assimilates surface precipitation. The WFDEI dataset was generally less biased than the reanalysis datasets. All datasets were then used as inputs to specifically calibrate a global conceptual hydrological model on 370 watersheds of the MOPEX database. River flows were computed for each watershed, and results showed that the flows simulated using NARR and gridded observations forcings were

very similar to the observed flows. Despite their biases, the simulated flows forced by the global reanalysis datasets were also similar to observations, except in the humid continental and subtropical climatic regions, where their poor precipitation seasonality degraded river flow simulations. The WFDEI dataset performed better than reanalysis in the humid continental and subtropical climatic regions, but was no better than reanalysis - and sometimes worse - in other climatic zones. Overall, the results indicate that global reanalyses could be used as proxies of surface temperature and precipitation to force hydrological models in regions with few weather stations.

Keywords: Reanalyses, river flow simulation, gridded observations, hydrological modeling, calibration

4.2 Introduction

Precipitation and temperature are important for hydrological studies. They are commonly used as meteorological forcings for hydrological models, and often come from weather stations. However, weather station coverage is sparse over several regions in the world, and recent years have seen a decline in surface observational networks in most countries.

This results in a shortage of data which sometimes severely limits our ability to conduct hydrological studies. It is therefore important to investigate the capacity to use meteorological data from others sources as proxies of station data, to overcome this deficit of observations.

Several studies have examined the contribution of remote sensing data for hydrological modeling (Bastola and François 2012; Cole and Moore 2009; Sagintayev et al. 2012). Overall, results of these works showed that remote sensing data have potential, but are not precise enough to allow hydrological models to adequately simulate river flows. Indeed, the precipitation rates estimated by radar often contain errors due to the difficulty faced by radars in distinguishing types and diameters of precipitation particles (Hunter 1996). Moreover, the scope and accuracy of radar measurements are often significantly degraded by the obstruction of electromagnetic waves caused, for example, by the rugged topography and trees (Warner et al. 2000; Westrick et al. 1999). Similar problems also plague satellite data since instruments cannot directly measure rainfall, and this necessitates the use of rainfall estimation techniques which have physical limitations (Barrett 1970; Grimes et al. 1999; Kidd et al. 2003; Vicente et al. 2002). Error correction in the remote sensing data still relies mostly on the use of data observed from ground weather stations. Therefore, remote sensing data can only be validated in regions with a dense stations network (Seo 1998; Steiner et al. 1999; Turk et al. 2008).

Meteorological reanalyses constitute another source of meteorological data. They make use of a wide variety of observation databases assimilated in a complex fashion into a numerical weather prediction model to produce a spatially and temporally coherent synthesis of various meteorological variables over the recent historical period. The reanalysis forecast model remains unchanged for consistency of simulated weather data. Data assimilated by reanalysis come from measurements recorded for decades throughout the world; these measurements themselves are derived from different sources. The main sources are terrestrial measurement networks, radiosondes, aircrafts, satellites and floats (Mesinger et al. 2006; Suarez et al. 2008; Wang et al. 2011). Terrestrial measurement networks are composed of weather stations, buoys and ships, and provide surface data for variables such as temperature, humidity, pressure, wind direction and speed. Radiosondes, aircrafts and satellites provide various atmospheric data, such as radiance, wind, humidity and pressure at different atmospheric heights. Reanalyses also assimilate data from several autonomous profiling floats (argo floats) which measure real-time temperature and the salinity of the first 2000 meters of ocean water. Although reanalyses are not direct observations, they provide variables throughout the world, including in areas where weather stations are non-existent or scattered (Bosilovich 2013).

Many studies have compared data from reanalyses to weather station data in several regions of the world (Bosilovich 2013; Lorenz and Kunstmann 2012; Manzanas et al. 2014;

Rusticucci et al. 2014; Vose et al. 2012; Zhang et al. 2013). These studies generally conclude that in many cases, reanalyses are comparable to observations. For instance, Nigam and Ruiz-Barradas (2006) showed that the spatial variance of summer and winter precipitation of NARR reanalysis (Mesinger et al. 2006) was similar to observations over the United States due to NARR's assimilation of surface precipitation. They also found that in the Great Plains region, summer precipitation in NARR and observations showed similar interannual variability (Ruiz-Barradas and Nigam 2006). Rusticucci et al. (2014) found that the interannual variability of observed precipitation in the southern Central Andes in South America was well represented in the ERA-Interim reanalysis (Dee et al. 2011).

Comparatively to remote sensing data, the potential of reanalysis data for hydrological modeling studies has been less explored. In general, studies on this topic have been based on a reduced number of watersheds, and their conclusions are therefore difficult to generalize. For instance, Woo and Thorne (2006) used temperature and precipitation data from ERA-40 (Uppala et al. 2005), NCEP-NCAR (Kalnay et al. 1996) and NARR reanalyses to simulate flows on a large subarctic mountain watershed in Canada. They found a cold bias in these reanalyses that produced a late snowmelt. Furthermore, Choi et al. (2009) evaluated the applicability of NARR data for hydrological modeling on three watersheds in northern Manitoba in Canada. They found that river flows simulated from NARR data adequately represented observed hydrographs. Vu et al. (2012) tried to simulate the Dak Bla River discharges in Vietnam using data from NCEP-NCAR reanalysis and found that simulated discharges significantly differed from those observed. It should be noted that most of the reanalyses examined in these studies have been improved, and their ability to be used as observation proxies for hydrological modeling studies is yet to be investigated.

To address biases present in reanalysis, global forcing datasets have been constructed using post-processing techniques (e.g. bias correction) based on observations (Sheffield et al. 2006; Weedon et al. 2011; Weedon et al. 2014). These global forcing datasets offer long-term consistent time series of near-surface meteorological variables that can be used for the study of seasonal and interannual variability. Some of these datasets are based on older reanalyses

– such as NCEP–NCAR and ERA-40 – (Ngo-Duc et al. 2005; Sheffield et al. 2006; Weedon et al. 2011) and could, therefore, be less accurate than the ones based on more recent reanalyses – such as ERA-Interim – (Weedon et al. 2014). Although a bias-corrected dataset is intended to be more accurate than the reanalysis on which it is based, in regions where weather stations are sparse or non-existent, bias correction may not bring improvement, and could even introduce additional errors in the corrected data. In addition, since precipitation and temperature are usually post-processed separately, some coherency between both variables could be lost in the process, with potential adverse effects on impact models.

This study aims to evaluate the use of three global atmospheric reanalyses – ERA-Interim, CFSR and MERRA – and a regional reanalysis – NARR – for hydrological modeling. The importance of biases and bias correction is further investigated by including the WFDEI dataset.

Specifically this study has two objectives: (1) compare temperature and precipitation datasets from these datasets to an observationally-based gridded dataset, and, (2) test their ability to serve as inputs for the hydrological modeling of 370 watersheds of the MOPEX database located in different climatic regions of the contiguous US. These watersheds were selected because of their relatively high density of weather stations. This study will be useful in validating the use of reanalyses for hydrological modeling in regions with relatively abundant surface weather stations. If it is successful, the next step will be to evaluate the potential use of reanalysis data in regions with sparse or low density of stations. Ultimately, the main interest of reanalyses for hydrological studies is twofold: to provide proxy data in regions not well covered with surface weather stations (e.g. Northern Canada, Africa) and to provide additional variables less commonly measured (e.g. wind, humidity, real evapotranspiration). The inclusion of the WFDEI dataset will allow the impacts of bias correction on the performance of ERA-Interim to be assessed for hydrological modeling. Contrasting the performance of bias-correction in regions with dense weather networks (Eastern US) versus that of regions less well covered (mid-West) should yield important information as to the applicability of reanalysis in remote regions.

4.3 Region of interest and datasets

4.3.1 Region of interest

This study was conducted over the contiguous United States and the selected watersheds for hydrological simulations were derived from the MOdel Parameter Estimation eXperiment database (MOPEX) (Duan et al. 2006). A total of 370 watersheds in 5 climatic regions (figure 4.1) according to the Köppen-Geiger climate classification (Kottek et al. 2006) were used. The watershed areas range between 104 and 10,325 km^2. The daily mean precipitation, temperature and discharge of the watersheds from each climatic region are presented in table 4.1.

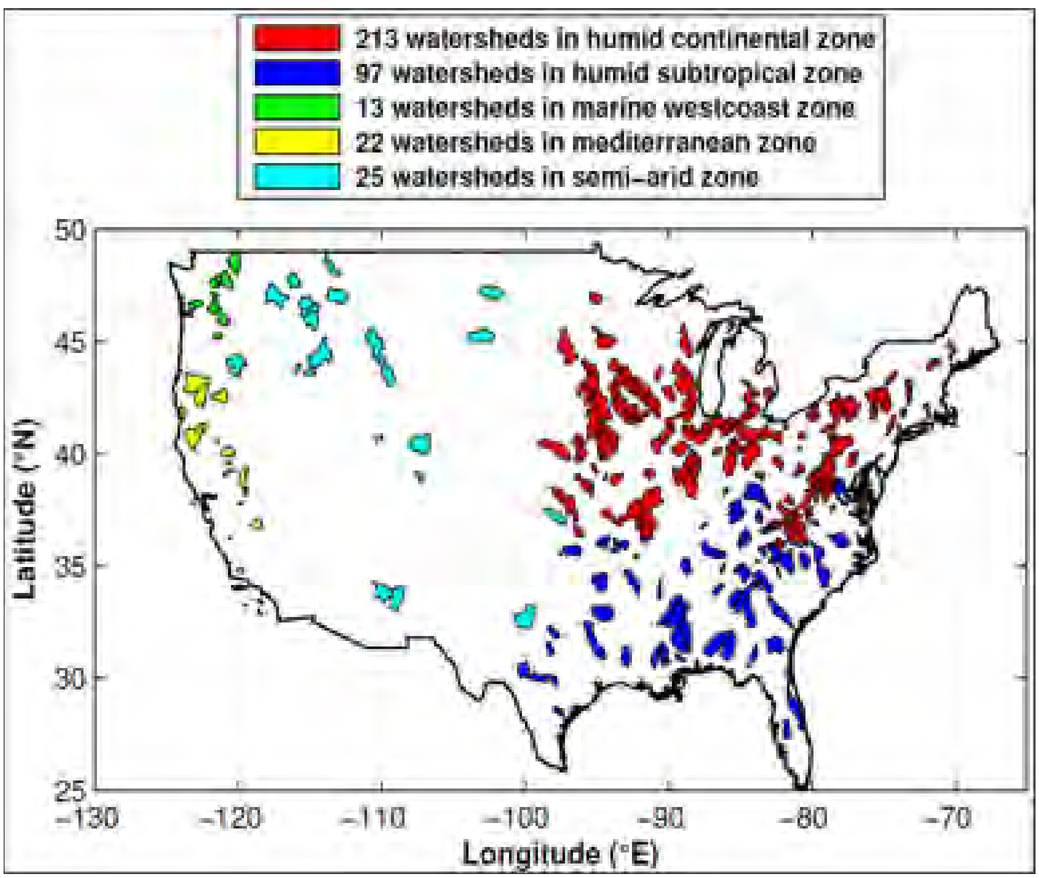

Figure 4.1 The 370 watersheds in the region of interest. The different colors show all watersheds from the same climatic zones

Table 4.1 Range of watershed-averaged daily mean precipitation, temperature and discharge for each climate zone

Climate zone	Number of watersheds	Area (km^2)	Daily mean precipitation (mm/d)	Daily mean temperature (°C)	Daily mean discharge ($mm/d/km^2$)
Humid continental	213	172 – 10091	1.7 – 4.1	5.2 – 15.2	0.1 – 2.7
Humid subtropical	97	179 – 9882	1.7 – 5.5	10.4 – 22.7	0.1 – 3.1
West coast	13	831 – 4588	2.3 – 7.5	4.9 – 10.5	0.8 – 6.9
Mediterranean	22	104 – 9535	1.3 – 7.4	4.5 – 15.4	0.1 – 5.7
Semi-arid	25	293 – 10325	1.0 – 3.8	1.0 – 17.9	0.1 – 2.3

4.3.2 Datasets

This study covers the 1979-2003 period, which is the longest common period of all the databases. All datasets are based on a daily sample.

4.3.2.1 Reanalysis and WFDEI datasets

ERA-Interim

ERA-Interim is the latest global reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al. 2011). It covers the period from 1979-present, and is produced by the December 2006 integrated forecast model of ECMWF (IFS Cy31r2). ERA-Interim uses a 4D variational data assimilation approach. The observations assimilated before 2002 come mainly from the data used for ERA-40 (Uppala et al. 2005). ERA-Interim is updated in near real time, using data from the operational ECMWF forecast system (Dee et al. 2011). ERA-Interim temperature results from the assimilated surface temperature, while precipitation is produced by the weather forecast model. The horizontal resolution of ERA-Interim is 0.75° x 0.75°. The ERA-Interim dataset is available for free online at: http://apps.ecmwf.int/datasets/.

Climate Forecast System Reanalysis (CFSR)

The CFSR global reanalysis is produced by the National Centers for Environmental Prediction (NCEP) from a coupled climate system atmosphere-ocean-land surface with an interactive sea-ice component. It covers the period from 1979-present, and uses a 3D variational data assimilation approach (Saha et al. 2010). CFSR assimilates satellite radiance rather than estimated temperature and humidity values (Wang et al. 2011). Estimates of greenhouse gas concentration changes, aerosols and solar variations are used as forcings in CFSR; CFSR also assimilates hydrological quantities of a land surface parallel model forced by the NOAA Climate Prediction Center (CPC) pentad merged analysis of precipitation (Xie and Arkin 1997) and the CPC unified daily gauge analysis (Wang et al. 2011). The horizontal resolution of CFSR is 0.313° (lon) x 0.312° (lat), and the CFSR dataset is available for free online at: http://cfs.ncep.noaa.gov/cfsr/.

Modern Era Reanalysis for Research and Applications (MERRA)

The MERRA global reanalysis is developed by the Global Modeling and Assimilation Office (GMAO) of the National Aeronautics and Space Administration (NASA) in order to allow the use of the GMAO satellites observations in a climate context, and to improve the hydrological cycle represented in the first generation of reanalyses (Rienecker et al. 2011). MERRA covers the satellites era (1979-present), and is generated from the 5.2.0 version of the Goddard Earth Observing System (GEOS) atmospheric model and a data assimilation system based on a 3D variational approach. The data assimilation system (DAS), the input data flux and their sources, observations, and error statistics are well documented in Suarez et al. (2008). The primary performance drivers for the GEOS DAS are temperature, humidity and wind fields (Schubert et al. 1993). The horizontal resolution of MERRA is 2/3° (lon) x 1/2° (lat). The datasets are available for free online at: http://disc.sci.gsfc.nasa. gov/mdisc/ overview/index.shtml.

North American Regional Reanalysis (NARR)

The NARR regional reanalysis is a product of the National Centers for Environmental Prediction (NCEP), developed to produce high-resolution data for North America. NARR was developed from major improvements of the NCEP-NCAR global reanalyses (Kalnay et al. 1996; Kistler et al. 2001), both in terms of resolution and precision. In the light of these improvements, NARR adequately represents extreme events, such as droughts and floods. For more details about these improvements, see Mesinger et al. (2006). NARR covers the period from 1979-present. NARR initially covered the 1979-2003 period. A real-time extension of the NARR called the Regional Climate Data Assimilation System or R-CDAS covers the more recent period. The NARR system uses the Eta 32-km atmospheric model with 45 vertical layers and a 3D variational data assimilation approach (Mesinger et al. 2006). That model uses the convection scheme of Betts–Miller–Janjic (BMJ) (Betts and Miller 1986; Janjic 1994). Surface precipitation is assimilated in NARR as latent heat profiles (Mesinger et al. 2006). Precipitation data used for assimilation come from different sources. A 1° rain gauge analysis is used for Mexico and Canada. A 1/8° daily rain gauge data analysis from the U.S. Climate Prediction Center (CPC) is used for the contiguous United States (CONUS) (Shafran et al. 2004); CONUS is orographically adjusted using the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) approach (Daly et al. 1994). Over oceans south of 27.5°N and land south of Mexico, the CPC Merged Analysis of Precipitation (CMAP) global 2.5° analysis is used, and no data are assimilated for oceans north of 43.5°N (Mesinger et al. 2006). NARR also updates the simulated snowpack using the daily global snow analysis (SNODEP) of the U.S. Air Force Weather Agency (Kopp and Kiess 1996). The horizontal resolution of NARR is 32km x 32km. The datasets are available for free online at: ftp.cdc.noaa.gov/NARR.

WATCH-Forcing-Data-ERA-Interim (WFDEI)

The WFDEI is a global meteorological forcing dataset produced using the WATCH Forcing Data (WFD) methodology (Weedon et al. 2011) applied to ERA-Interim data. It covers the period 1979-2012 and contains eight meteorological variables at a 3-hourly time step. Bias correction was performed on a monthly basis. For two of the variables – rainfall rate and snowfall rate – biases were corrected using the CRU TS3.101 data (TS3.21 for 2010-2012) (Harris et al. 2014; New et al. 1999, 2000) and the GPCCv5 data (v6 for 2010) (Rudolf and Schneider 2005; Schneider et al. 2014). The horizontal resolution of the WFDEI datasets is 0.5° x 0.5°. The WFDEI dataset is available online at: ftp.iiasa.ac.at.

4.3.2.2 Observations datasets

Gridded datasets of Santa Clara

On the basis of the results presented in Article 1 (Chapter 3), the gridded datasets from Santa Clara database are used in this work. The Santa Clara observed gridded dataset was produced at the University of Washington by interpolating observed data from weather cooperative stations of the National Oceanic and Atmospheric Administration (NOAA) (average of 1 station per 700km^2) (Maurer et al. 2002). The Synergraphic Mapping System algorithm (SYMAP) of Shepard (1984) implemented by Widmann and Bretherton (2000) was used for the interpolation. The Santa Clara gridded precipitation was scaled to match the long-term average precipitation of the Parameter-elevation Regressions on Independent Slopes Model (PRISM) (Daly et al. 1994; Daly et al. 1997). The Santa Clara gridded dataset covers the 1949-2010 period, and its horizontal resolution is 0.125° x 0.125°. They are available at: http://hydro.engr.scu.edu/files/gridded_ obs/daily/ncfiles_2010.

Discharge datasets of the MOPEX database

The MOPEX database contains daily mean precipitation and temperature (minimum and maximum) for 400 watersheds. Streamflows at each watershed outlet are also provided. The watershed-averaged precipitation and temperature data is derived from the National Climatic Data Center (NCDC) weather stations (Duan et al. 2006). An inverse distance weighting method was implemented to estimate the final MOPEX climate data. A detailed description of this data source is available in (Schaake et al. 2006). Only time series of length greater than 10 years were admitted in the database. The MOPEX database covers the 1949-2003 period and is available at: ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data.

Table 4.2 summarizes general information on all the databases described above.

Table 4.2 Description of all databases used in this study

Dataset	Acronym	Horizontal resolution	Assimilation scheme	Source	Reference
Model Parameter Estimation eXperiment	MOPEX	---	---	ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data	(Duan et al. 2006)
Santa Clara gridded datasets	SClara	0.125° x 0.125°		http://hydro.engr.scu.edu/files/gridded_obs/daily/ncfiles_2010	(Maurer et al. 2002)
North American Regional Reanalysis	NARR	32km × 32km	3DVAR	http://www.esrl.noaa.gov/psd/data/gridded/data.narr.html	(Mesinger et al. 2006)
ERA-Interim reanalysis	ERAI	0.75° x 0.75°	4DVAR	http://data-portal.ecmwf.int/	(Dee et al. 2011)
Climate Forecast System Reanalysis	CFSR	0.313° (lon) x 0.312° (lat)	3DVAR	http://cfs.ncep.noaa.gov/cfsr/	(Saha et al. 2010)
Modern Era Reanalysis for Research and Applications	MERRA	2/3° (lon) x 1/2° (lat)	3DVAR	http://disc.sci.gsfc.nasa.gov/mdisc/overview/index.shtml.	(Rienecker et al. 2011)

4.4 Methodology

Initially, the quality of precipitation and temperature from all datasets was assessed through a comparison against observations, as represented by the Santa-Clara gridded dataset. In a second step, hydrological simulations were performed with all datasets to further assess the ability of the different reanalyses to capture the complex precipitation-temperatures interactions needed to adequately simulate watershed hydrology.

4.4.1 Data comparison: Temperature and Precipitation

Prior to the comparison, the Santa-Clara gridded dataset was aggregated to the resolution of each reanalysis and WFDEI datasets. This aggregation was achieved by averaging the data from the Santa Clara grid toward each target grids (spatial average).

The statistics used for comparison include the bias, root-mean-square error (RMSE), variances ratio and correlation using daily time series. The mean annual cycles were also calculated and compared for each climate region. The bias is the difference between a dataset mean precipitation (or temperature) over a given period and that of the corresponding observations. It indicates how much a given dataset overestimates or underestimates the observed data. Thus, a positive (or negative) bias corresponds to an overestimation (or an underestimation). The RMSE is a measure of the absolute fit between each dataset and observations. Low RMSE values indicate a better fit. The variance ratio compares the dataset variability to that of observations, and thus, a ratio of 1 indicates equal variability. The temporal correlation coefficient shows the intensity of the link between daily time series from each dataset and observed data. A zero correlation coefficient corresponds to an absence of correlation, while a correlation coefficient of 1 (or -1) indicates a perfect positive (or negative) dependence between the time series.

Results are presented for each season: winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

Extreme values are not analyzed because they are beyond the scope of this study.

4.4.2 Hydrological modelling: Input data and Model calibration

The sizes of the watersheds considered in this study are relatively small. Therefore, the lumped conceptual hydrological model, HSAMI (Fortin 2000; Minville et al. 2008), is used to simulate discharges. HSAMI has been used to predict the hourly and daily flows of more than a hundred watersheds in Quebec. It has also been used operationally by Hydro-Québec over 100 watersheds for more than 30 years, as well as in climate change impact projects (Chen et al. 2012; Poulin et al. 2011). HSAMI simulates the main hydrological cycle processes, such as vertical and horizontal water transfers, evapotranspiration, snowmelt and soil freezing. HSAMI has 23 calibration parameters: 10 for the different production function processes, 5 for the horizontal transfer through reservoir-type soil layers, 2 for evapotranspiration, and 6 for snow-related processes. There are four interconnected reservoirs that contribute to the vertical water transfer balance: Snow on ground, surface runoff, saturated soil layer and unsaturated soil layer. The horizontal water transfer is based on two unit-hydrographs (one for surface runoff and one for delayed runoff) and one linear reservoir for groundwater flows. HSAMI requires spatially averaged daily minimum and maximum temperatures as well as daily rainfall and snowfall depths. Precipitation and temperature data from all databases were averaged over each watershed using the Thiessen polygon method (Thiessen 1911). Other methods were tested (e.g., weighting by the inverse of the distance), but had no impact on the conclusions of this study.

Because of the large number of watersheds used in this study, an automatic optimization algorithm was used to calibrate the hydrological model. Arsenault and Brissette (2014a) showed that the Covariance Matrix Adaptation Evolution Strategy (CMAES) algorithm (Hansen and Ostermeier 1996, 2001) was the optimal choice for calibrating the HSAMI model. Thus, the CMAES optimization algorithm was used to perform all calibrations in this study.

The Nash-Sutcliffe Efficiency (NSE) metric (Nash and Sutcliffe 1970) was used to evaluate the performance of the different databases. Other performance metrics could have been used, but the NSE is by far the most widely used in hydrology, and it was deemed adequate for the needs of this study.

In calibration, the NSE was calculated based on the even years, with cross-validation on odd years, and vice versa. This allows different climatic trends to be taken into account (e.g., natural decadal or multi-decadal variability). However, this method has a disadvantage because the hydrological model has to be executed over the entire study period to select the odd years or pairs to calculate the NSE. This therefore doubles the computational cost. For each watershed, 10 calibrations in the even/odd approach and 10 calibrations in the odd/even approach were achieved for a total of 20 calibrations. This approach reduces the likelihood of the calibration algorithm not converging during a single optimization process. For each watershed, only the best parameter set was selected.

The HSAMI hydrological model was calibrated particularly to each specific dataset. The non-parametric Wilcoxon test was performed to test the null hypothesis of equal NSE median values of simulated discharges, between the Santa Clara gridded database and each reanalysis at the 95% confidence level (Rakotomalala 2008). To avoid any issue due to equifinality and overfitting, all results presented in the next section cover only the validation period.

4.5 Results

4.5.1 Data comparison: Temperature and Precipitation

4.5.1.1 Temperature

The spatial distributions of the mean temperature biases are similar from one reanalysis to another, especially in spring and autumn (figures 4.2a–4.2d). In general, all reanalyses tend to overestimate the observed temperatures, which results in a warm bias over most of the USA. NARR biases are relatively low in the Eastern USA. Over the Western USA, NARR warm biases are more important. In winter, NARR displays a cold bias in the Midwest. During winter and autumn, ERA-Interim overestimates the temperature over most of the USA, except in Florida. In summer, ERA-Interim is cooler than observations in the Southern USA.

In spring, in the Northeastern USA, ERA-Interim agrees well with observations (bias between -0.5°C and 0.5°C). CFSR is warmer than observations in the South and Midwest, regardless of the season. MERRA is warmer than observations in the Western USA, but cooler in the Midwest and New England during winter and in the South during summer.

In the case of WFDEI, the spatial distribution of the mean temperature biases is similar for all seasons. In general, these biases are between -0.5°C and 1°C in the Eastern USA (figure 4.2e). They are warmer in the Northwest USA (particularly during winter, bias > 2°C) and cooler in the Southwest USA (particularly during spring and summer). Overall, results show that bias correction reduced the biases of the mean temperature compared to ERA-Interim. This is particularly true in Eastern USA, likely due to a denser network of weather stations.

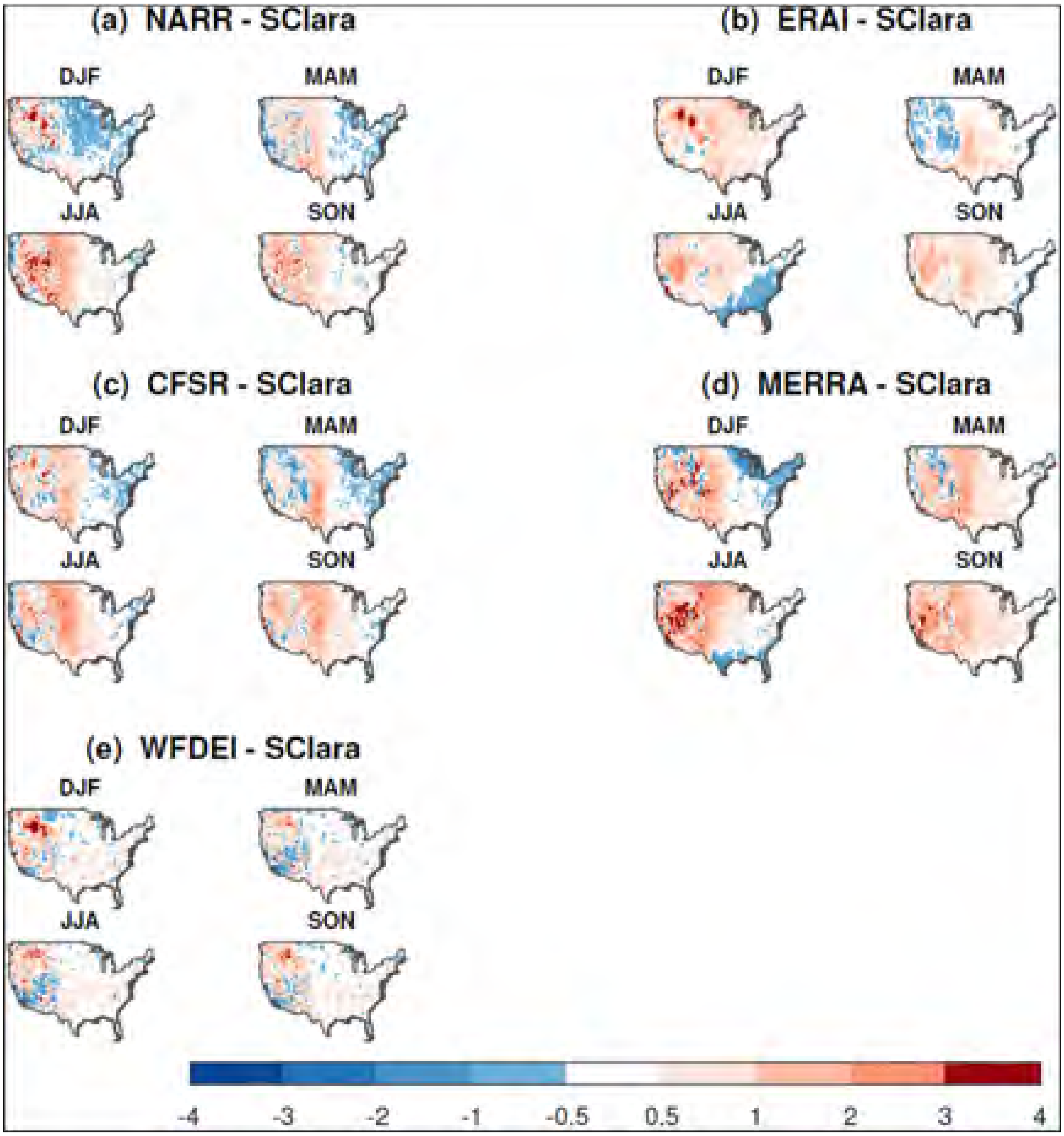

Figure 4.2 1979-2003 mean seasonal temperature difference (°C) between reanalyses, WFDEI and the observed gridded dataset from Santa Clara

In general, RMSE values indicate that the gaps between reanalysis mean temperatures and observations are higher in winter (figures 4.3a–4.3d). In the Eastern USA, the RMSE values of NARR are less than 2°C, and are lower than in the Western part. Furthermore, in the Midwest and the Rockies, the RMSE values of ERA-Interim, CFSR and MERRA are higher than in the South and the East.

Most of the RMSE values of WFDEI are between 1°C and 1.5°C in the Eastern USA, except in winter (figure 4.3e). In the Western USA, the RMSE values are higher (RMSE > 1.5°C) especially during winter in the Midwest where RMSE > 2°C. Overall, the RMSE values of the mean temperature from WFDEI are lower than their ERA-Interim counterpart.

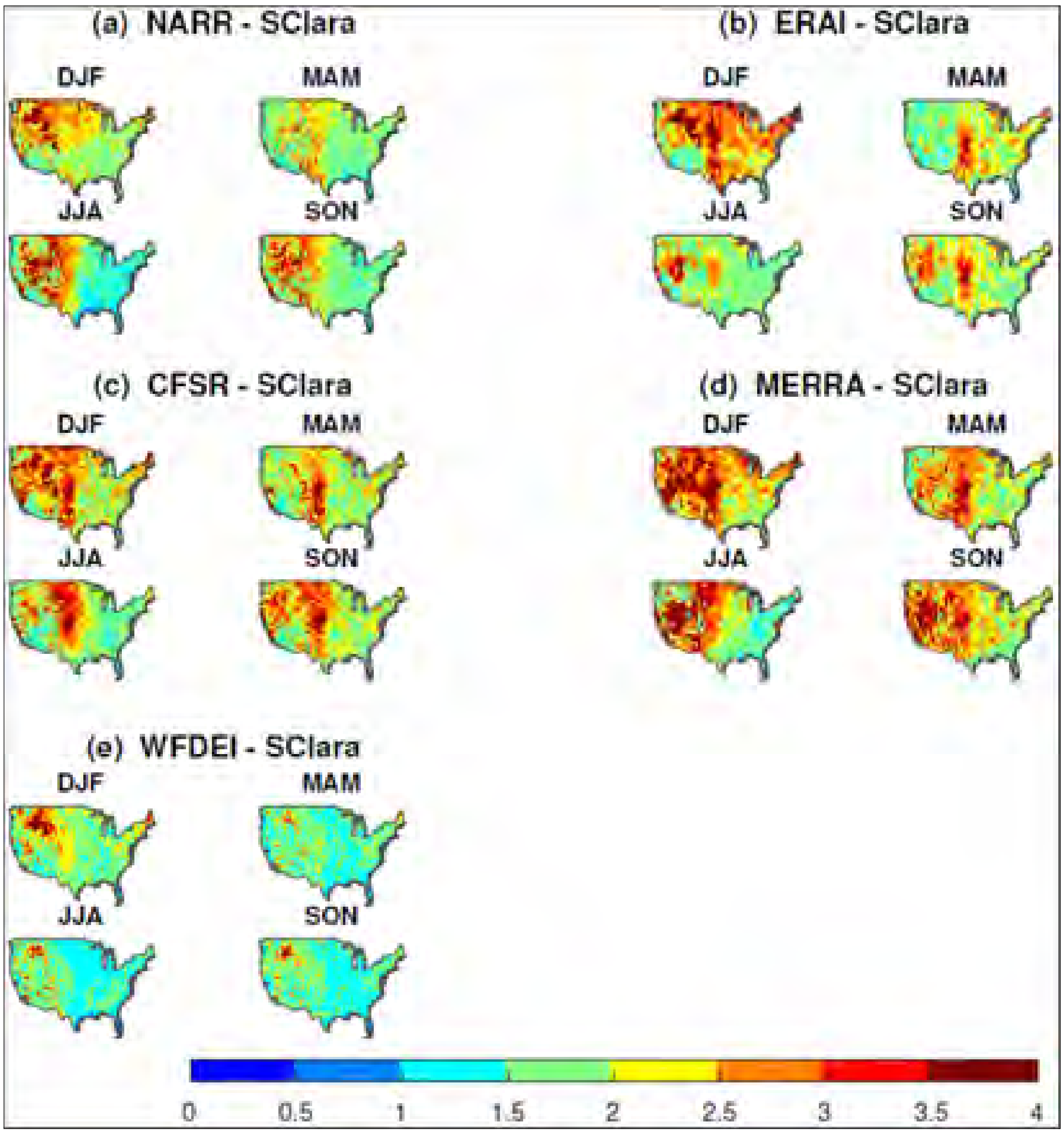

Figure 4.3 1979-2003 mean seasonal temperature RMSE (°C) between the reanalyses, WFDEI and the observed gridded dataset of Santa Clara

The variances of daily temperature from the reanalyses are generally similar to the observations, particularly in the humid Continental and Subtropical regions (figures 4.4a–4.4d). Compared to precipitation (shown below), temperature variance is generally lower, and is more easily simulated by the reanalyses. The four reanalyses tend to show less variance than observed during winter in the Western USA. The highest variance ratios (ratio > 1.5) are obtained during the summer, in the Northwestern USA for NARR and Western USA for ERA-Interim and MERRA. CFSR temperature variance is greater than observations over most of the USA during summer. Reanalysis model uncertainty plays an important role in the representation of the temperature variance (Willett et al. 2012). That uncertainty is higher in summer, and leads to a larger difference between reanalyses and observed temperature variance.

The variances of WFDEI daily temperature are generally very similar to the observations in the Eastern USA during spring and autumn with variance ratios varying between 0.75 and 1.25 (figure 4.4e). The patterns of the variance ratios of WFDEI and ERA-Interim are similar. However, ratios are lower with WFDEI than with ERA-Interim.

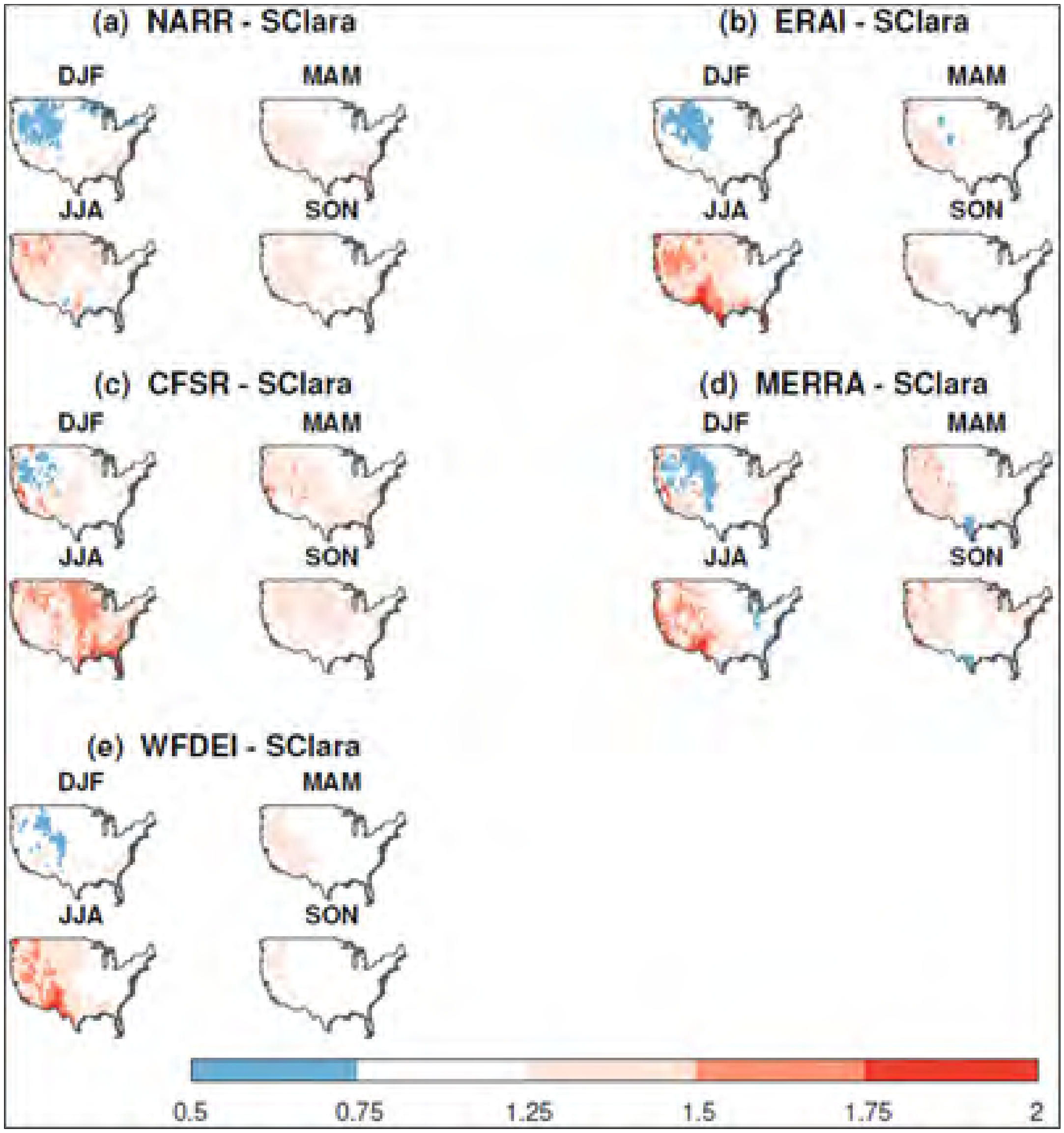

Figure 4.4 1979-2003 ratio of variance of mean daily temperature between reanalyses, WFDEI and the observed gridded dataset from Santa Clara for each season

Correlations between reanalysis daily temperature time series and observations are much larger than for precipitation. During spring and fall, reanalysis correlation coefficients are higher than 0.9 throughout the USA (figures 4.5a–4.5d). In winter, the correlations are lower in the Rockies. Lower correlation coefficients are observed in the Southern USA during summer. In general, correlation is higher between observations and NARR. Of the three global reanalyses, ERA-Interim temperature has the best overall correlation with observations. Indeed, the mean area averages correlation coefficients in winter, spring, summer and autumn, are respectively 0.92, 0.97, 0.87 and 0.98 for ERA-Interim, 0.91, 0.94, 0.89 and 0.97 for CFSR, and 0.89, 0.96, 0.85 and 0.97 for MERRA. This is likely due to the assimilation of land surface temperatures in ERA-Interim (Dee et al. 2011; Simmons et al. 2010). The correlation spatial patterns of WFDEI and ERA-Interim daily temperature time series are similar (figure 4.5e). However, correlation coefficients are higher for WFDEI.

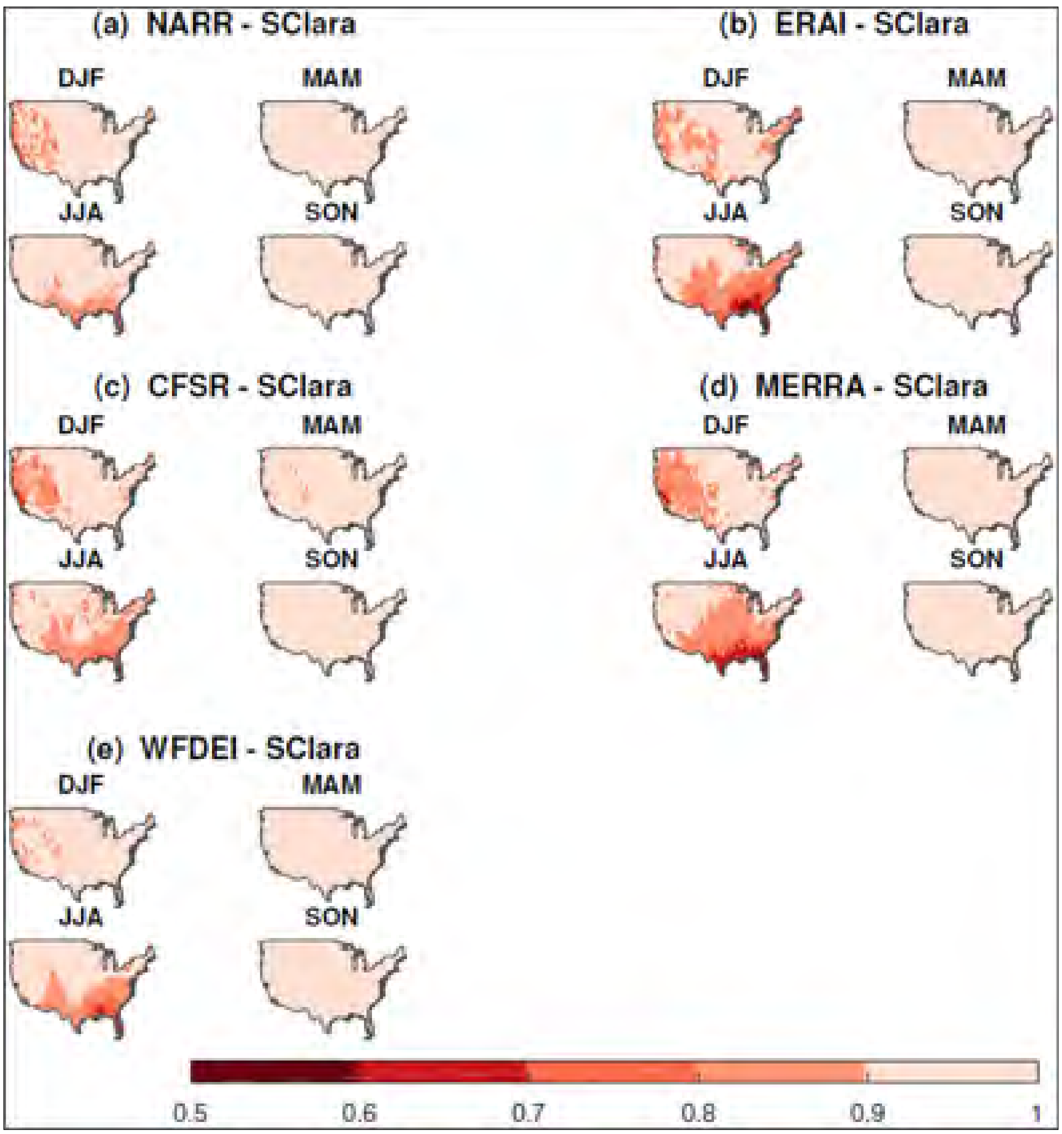

Figure 4.5 Daily temporal correlation between reanalyses, WFDEI and Santa Clara temperature. Results are shown by season and are based on daily data for the 1979–2003 period

Overall, reanalyses and WFDEI adequately reproduce the mean annual cycle of observed temperatures in the five climate regions of this study (Figure not shown). This result is not surprising because reanalysis temperature is well correlated with observed temperature, and their biases are low.

In summary, the representation of temperature in reanalyses is robust, probably because the atmospheric temperature from radiosondes and satellites are regularly assimilated in the reanalysis systems. Bias correcting reanalysis (as represented by the WFDEI dataset) results in improved values of all considered statistical criteria. The absolute improvement is however, relatively small, since reanalysis perform quite well with respect to temperature.

4.5.1.2 Precipitation

Differences between the mean seasonal precipitation of reanalyses and observations (figures 4.6a – 4.6d) show that NARR data is much closer to the observations compared to the other datasets, including WFDEI. This is due to the fact that unlike other reanalyses, the NARR atmospheric model precipitation is forced by observed precipitation through latent heat profiles. In general, NARR is slightly drier than observations for each season over the USA, except in the Midwest region, where NARR precipitation biases are relatively low (±10%). ERA-Interim and MERRA tend to be drier than observations in the Southeastern and the West coast in the winter, spring and fall, but they are wetter in the Northern High Plains, especially in winter. Overall, ERA-Interim wet biases are higher than those of MERRA, while MERRA dry biases are higher than those of ERA-Interim. In summer, these two reanalyses are wetter than observations in the South of the USA, but are drier in the Midwest. Bosilovich (2013) obtained similar results while studying the ability of reanalyses to reproduce changes in summer precipitation and temperature in the United States. CFSR shows similar biases as ERA-Interim and MERRA. However, in the Midwest and Western US, CFSR wet biases are significantly higher in winter (bias > 130%), and its dry biases are considerably higher in the summer (bias < -40 %).

Differences between the mean seasonal precipitation of WFDEI-GPCC/WFDEI-CRU and observations are shown in figures 4.6e – 4.6f. Results show that the precipitation of WFDEI is similar to that of observations in the Eastern USA during spring, summer and autumn and in the Southeastern USA during winter (bias ±10%). In the Western USA, WFDEI-GPCC and WFDEI-CRU are dryer than observations. Overall, bias correction seems to have improved the precipitation from ERA-Interim only in the Eastern USA, possibly because of the higher density of weather stations in this region.

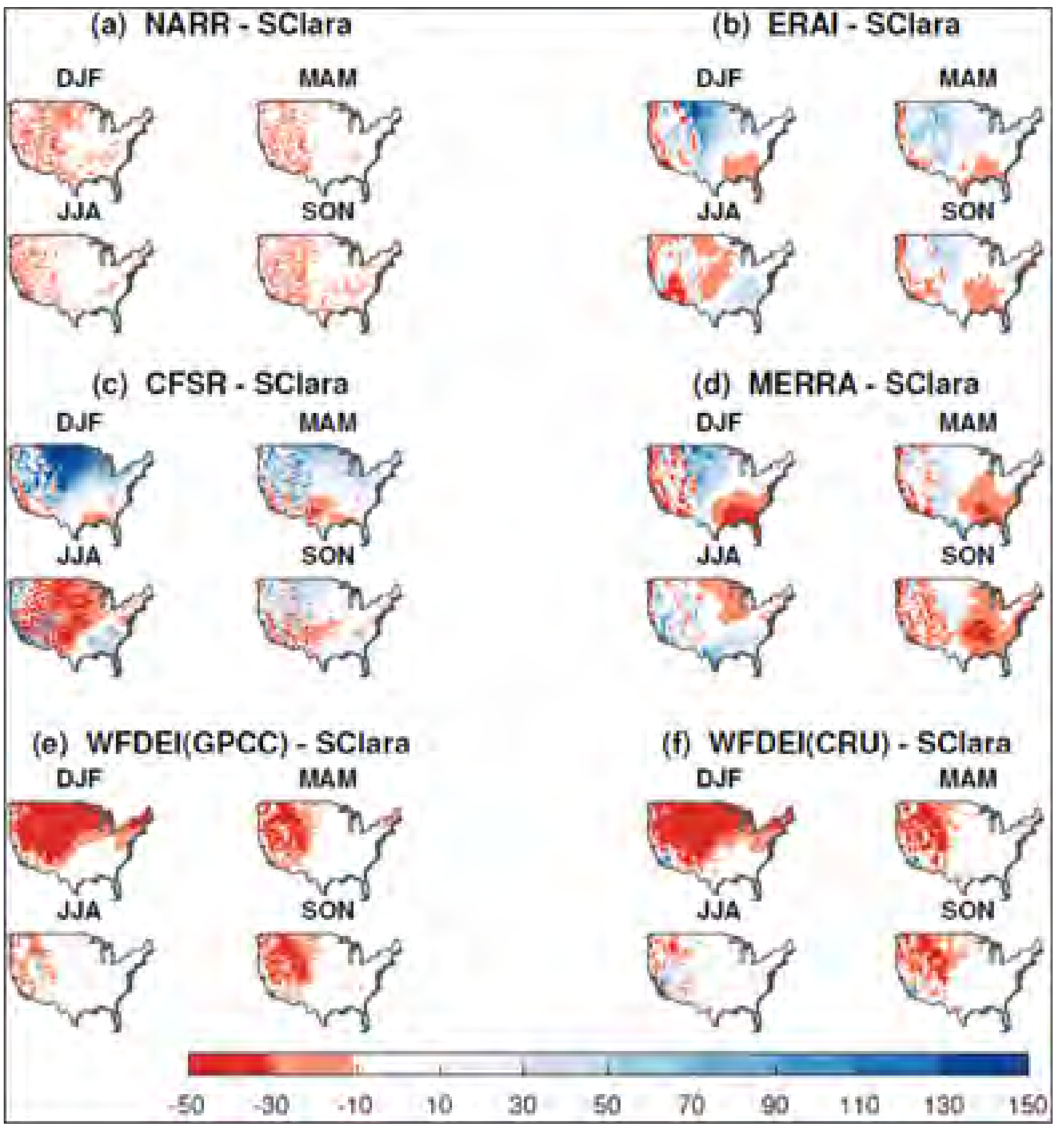

Figure 4.6 1979-2003 mean seasonal precipitation relative difference between reanalyses, WFDEI and observed gridded dataset from Santa Clara

Overall, the RMSE spatial distributions of the reanalyses are similar to one another (figures 4.7a – 4.7d). In general, RMSE values are high (RMSE > 6mm/day) in the Southeastern USA, where land-atmosphere interactions strongly affect the reanalysis forecast model. In fact, the land–atmosphere interactions influence the physical parameterizations in the forecast model (Bosilovich 2013). Moreover, the atmospheric moisture fluxes and the land surface soil moisture affect local precipitation (Wei et al. 2015). In the Southeastern USA, there is a strong humidity gradient and an intense moisture flux, which increases the influence of land-atmosphere interactions on the forecast model of the reanalysis. Conversely, in the Western half of the USA, RMSE values are lower (RMSE < 2mm/day). However, in the ocean and Mediterranean regions, where precipitation is abundant, regular, and largely influenced by the Pacific Ocean, reanalyses face difficulties in adequately estimating quantities of precipitation in the winter, spring and autumn (RMSE > 5mm/day). In general, NARR and MERRA RMSE values are lower than those of the other two reanalysis. The highest RMSE values are obtained with CFSR (RMSE > 9mm/day in the Southern USA).

The pattern of the RMSE values for WFDEI-GPCC and WFDEI-CRU is similar to that of ERA-Interim. However, in the Southeastern USA – and in the Midwest during summer – the RMSE values of precipitation from WFDEI are slightly higher than those of ERA-Interim (figures 4.7e – 4.7f). In this region, bias correction of ERA-Interim has not resulted into any improvement.

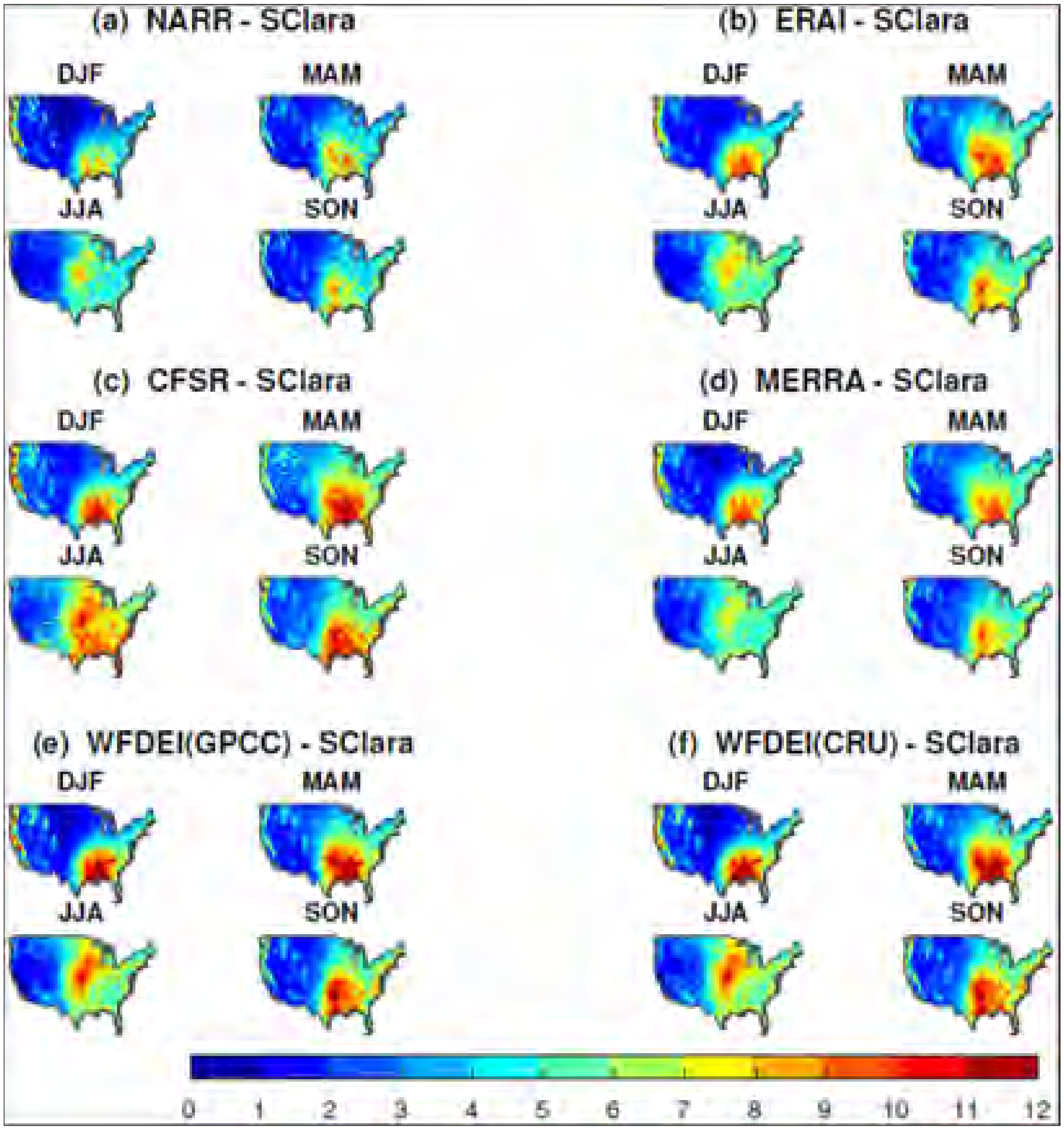

Figure 4.7 1979-2003 seasonal precipitation RMSE (mm/day) between reanalyses, WFDEI and observed gridded dataset from Santa Clara

NARR precipitation variance is globally similar to the observations (variances ratios between 0.8 and 1.2) (figure 4.8a). In general, ERA-Interim and MERRA precipitation variances are higher than observed values, except in the Southern USA (figures 4.8b – 4.8d). Over most of the USA, CFSR precipitation variance is higher than that of the observations.

The variance of WFDEI precipitation is higher than observed values in the Eastern USA during winter, spring and autumn and lower in the Western USA, particularly in the Midwest during winter (figures 4.8e – 4.8f). In summer, WFDEI and ERA-Interim perform similarly.

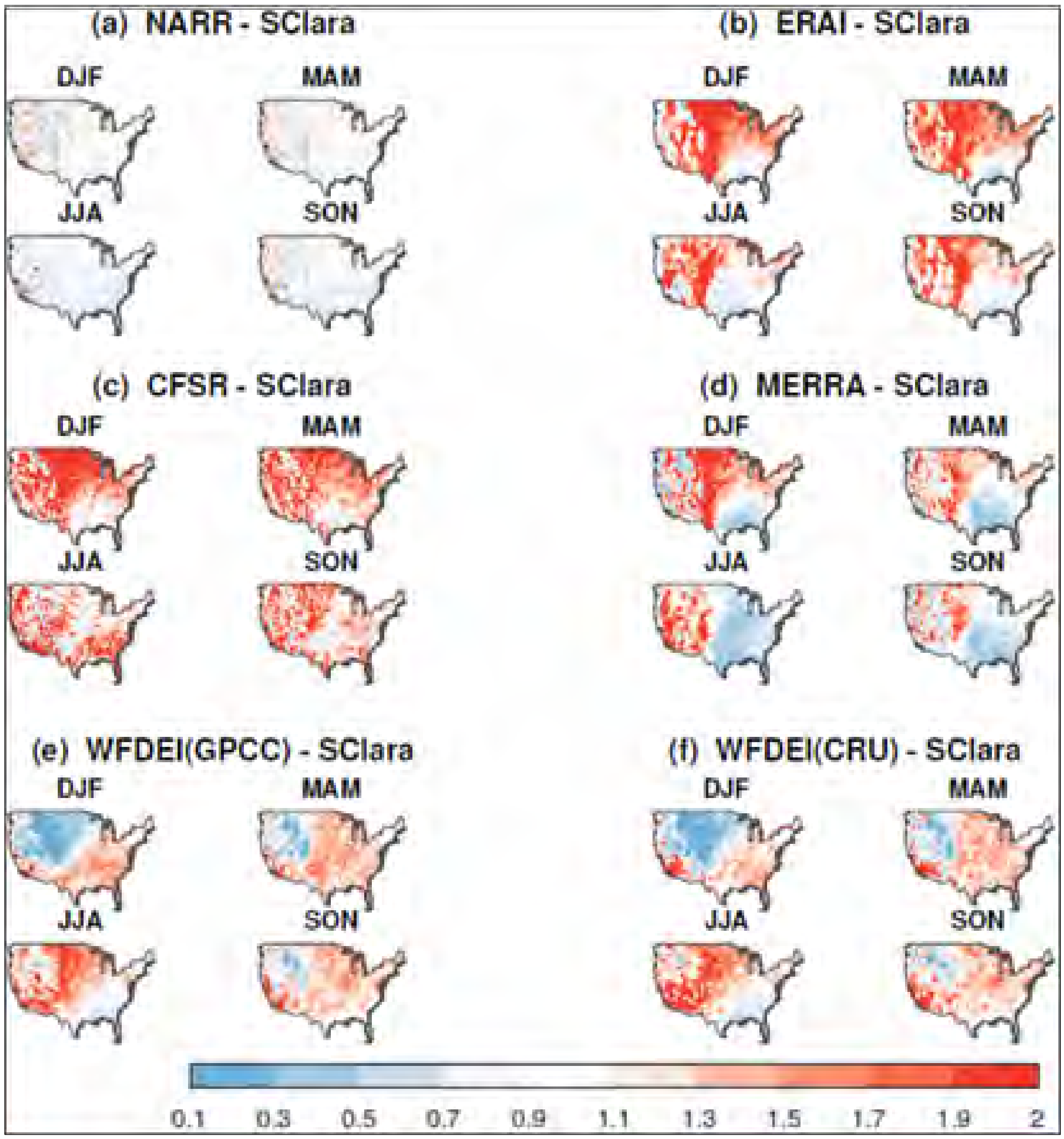

Figure 4.8 Variance ratios of daily precipitation between reanalyses, WFDEI and observed gridded dataset from Santa Clara for each season

Daily precipitation series from reanalysis correlate generally well with the observations over the USA (figures 4.9a – 4.9d). The lowest correlations are obtained in the summer because during that season, reanalysis models have a lower predictive ability, which is explained by local and stochastic weather conditions (Bosilovich 2013; Bosilovich et al. 2009). Compared to the other reanalyses, NARR displays the best correlations with coefficients above 0.8 over the entire USA, with the exception of the Central part, where the correspondence with observations is lower (corr. coef. < 0.6). Fuller (2012) also observed a low correlation between NARR and observed precipitation from two weather stations in the Central USA. During winter, spring and fall, ERA-Interim, MERRA and CFSR correlate well with observations in the Eastern USA and in the mountainous regions of the west USA (corr. coef. > 0.7). In Southern and Central USA, these three reanalyses are less well correlated with observations. During winter, the atmospheric dynamics over the United States generally plays an important role for precipitation, and there is a high correlation between precipitation from reanalyses and observations, except in the Lower Mississippi River Valley region in southern United States where precipitation is mainly convective in winter. However, in summer, the correlation between the precipitation from reanalyses and observations are lower than in winter because convection processes are more important due to the strong land-atmosphere interactions (Higgins et al. 2010). Despite its coarser resolution, ERA-Interim correlates better with the observations as compared to the other two global reanalyses. Furthermore, despite its higher resolution, CFSR displays the lowest correlations.

The correlation between WFDEI precipitation and observations is similar to that between ERA-Interim and observations (figures 4.9e – 4.9f), except in Western USA, during winter, spring and summer, where WFDEI is slightly worse. Overall, bias correcting ERA-Interim precipitation has not improved its correlation against observations.

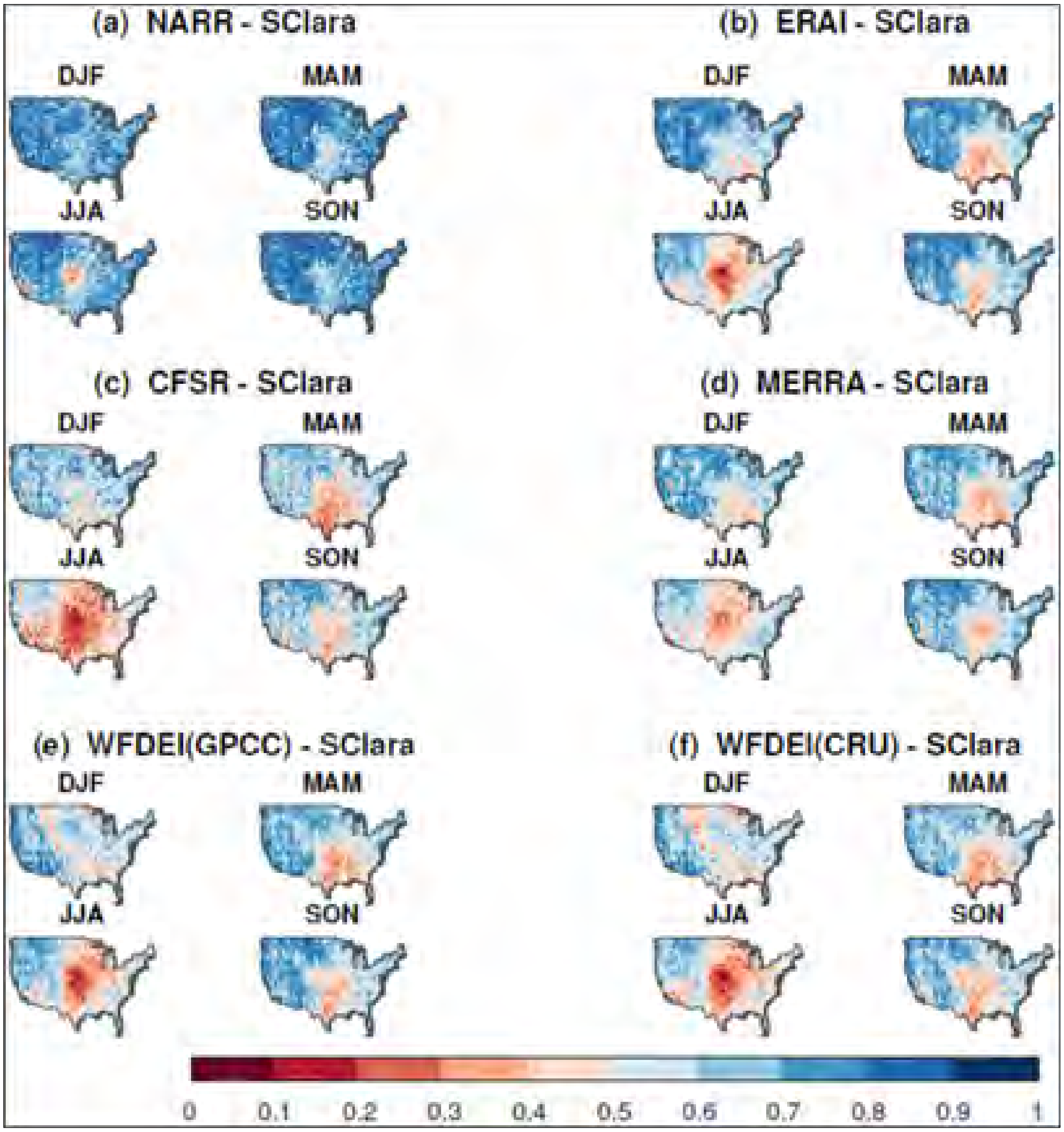

Figure 4.9 Daily correlation between reanalyses, WFDEI and Santa Clara precipitation. Results are shown by season and are based on daily data for the 1979–2003 period

The mean annual cycle of precipitation from all datasets is similar to the one observed in the West coast and Mediterranean regions, especially during the summer and autumn (figures 4.10a – 4.10b). This is consistent with previous results. Indeed, for these regions in the summer and autumn, reanalysis precipitations correlate well with observations, and their biases are relatively low. Results also show that in these climate regions, CFSR precipitation amounts are greater than observations during the winter and spring. ERA-Interim and

MERRA precipitation are lower than observations during the winter and autumn. In the Semi-arid region (figure 4.10c), the reanalysis annual cycles are similar to the observed one. In the humid Continental region, CFSR weakly reproduces the annual observed cycle (figure 4.10d); the same is seen in the humid Subtropical region for all three global reanalyses (figure 4.10e). Unlike these global reanalyses, NARR adequately simulates the annual precipitation cycle in this region. This result is not surprising since NARR assimilates surface precipitation. Bukovsky and Karoly (2007) also obtained similar results for NARR. Comparatively to ERA-Interim, WFDEI adequately simulates the annual precipitation cycle in the humid Subtropical region. This implies that precipitation biases present in the ERA-Interim dataset have been adequately corrected in that climate region, as discussed earlier.

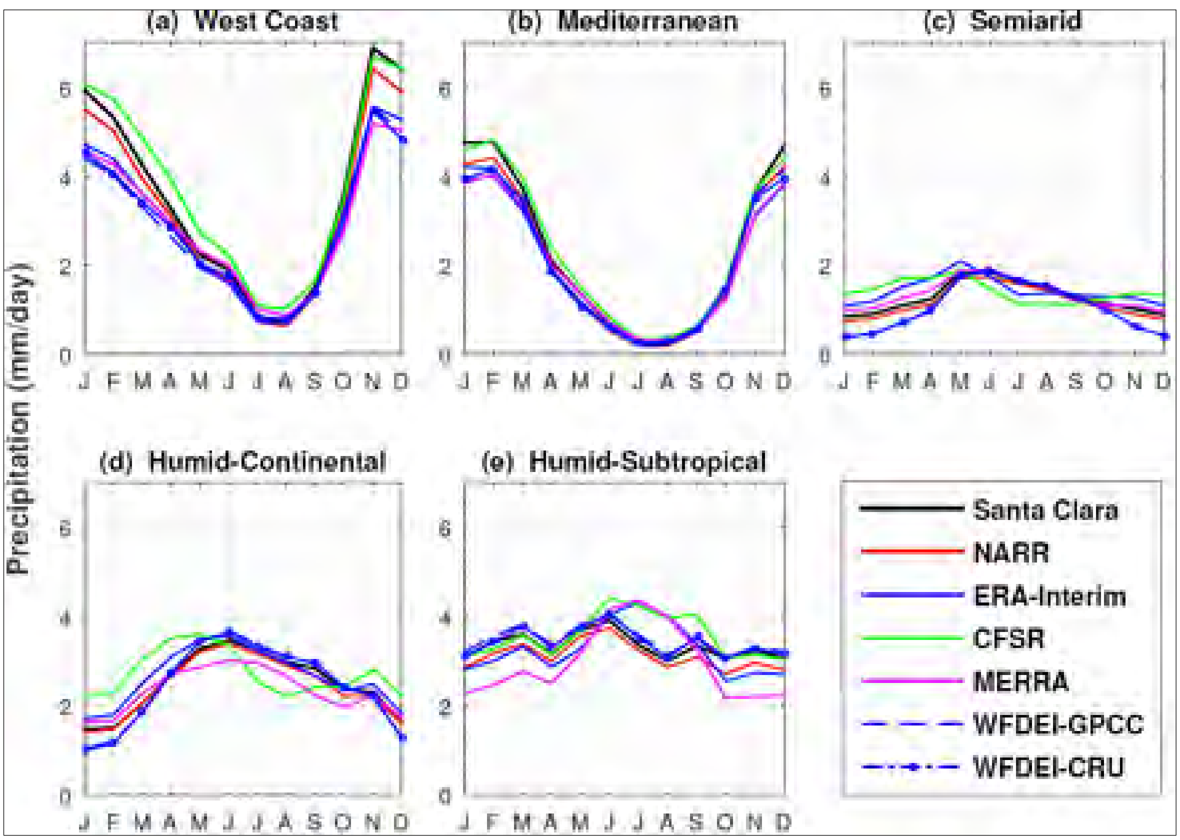

Figure 4.10 1979-2003 mean annual cycle of precipitation for the 5 climatic zones in which the 370 watersheds are distributed

4.5.2 Hydrological simulations: Performance statistics

The validation performance of different precipitation and temperature combinations are shown in figure 4.11. Results show that precipitation from NARR, combined with the temperature of any of the reanalyses – except MERRA – leads to a similar performance. The same is true for the precipitation of other reanalyses. This indicates that NARR, ERA-Interim and CFSR temperatures are equally good with respect to hydrological modeling. In other words, the differences existing between these reanalysis temperature datasets are not significant. Results also show that the use of the MERRA temperature leads to a slight statistically significant drop in performance. In almost all the cases, the drop in performance appears to be significant according to the Wilcoxon statistical test performed, and stands at a 95% level of significance. This result is consistent with the temperature comparison results, which showed that the MERRA temperature deviates the most from observations. It also appears from the results presented in figure 4.11 that reanalysis performance in hydrological simulation is mainly determined by the quality of its precipitation field.

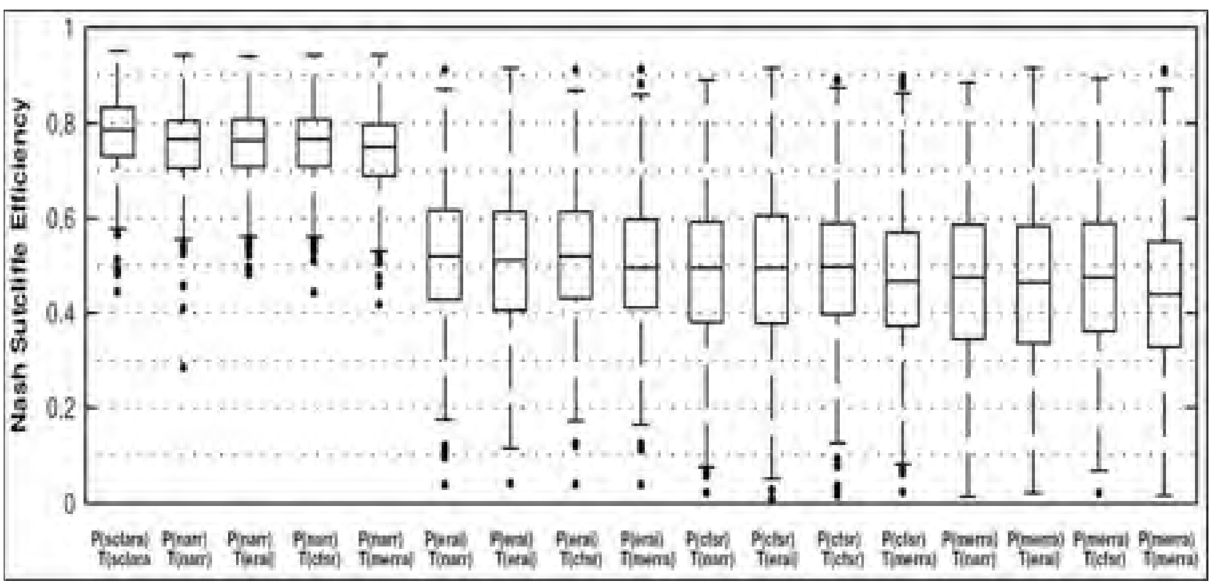

Figure 4.11 Distribution of the NSE values of the different datasets for 370 watersheds. Results are based on daily discharges simulated in the validation period. The lower and upper limits of each boxplot represent the 25th and 75th percentiles, respectively. The middle line represents the median (50th percentile). The ends of the whiskers represent extreme values

Generally, the hydrological performance of observations (Santa Clara) is slightly superior to that of NARR (figure 4.12a). The NSE median values are respectively 0.784 for observations and 0.764 for NARR when considering all of the 370 watersheds. Performances obtained from the global reanalyses and WFDEI are significantly lower than those of NARR. Their NSE median values are equal to 0.512 for ERA-Interim, 0.496 for CFSR 0.441 for MERRA, 0.590 for WFDEI-GPCC and 0.519 for WFDEI-CRU respectively. Hydrologists generally consider 0.6 as an acceptable NSE value (Chiew et al. 2009; Kouame et al. 2013; Kralisch et al. 2007; Pappenberger and Buizza 2009). The performances of the global reanalyses and those of the global WFDEI data are below that threshold for at least 73% of the watersheds with the exception of WFDEI-GPCC (53% of the watersheds). Individual comparison show that the NSE values of WFDEI-GPCC and WFDEI-CRU are respectively superior to those of ERA-Interim for 74% and 52% of the 370 watersheds. A more complex portrait emerges when each of the five climate regions is considered (figures 4.12b – 4.12f).

Hydrological performances are also compared by climate regions. Results show that in both the Humid Continental and Humid Subtropical regions (figures 4.12b – 4.12c), the global reanalysis performance is particularly low. For example, from the 213 catchments considered in the Humid Continental region, 78% display a NSE value lower than 0.6 for ERA-Interim, 86% for CFSR and 90% for MERRA. In that climatic zone, the NSE values of WFDEI-GPCC (median NSE = 0.554) and WFDEI-CRU (median NSE = 0.490) are superior to those of ERA-Interim (median NSE = 0.512) for 72% and 44% of the watersheds respectively. Thus, WFDEI-GPCC is significantly better than ERA-Interim whereas the latter is slightly better than WFDEI-CRU for most of the watersheds in the Humid Continental region. In the humid Subtropical region, both WFDEI-GPCC (median NSE = 0.601) and WFDEI-CRU (median NSE =0.530) are also significantly better than ERA-Interim (median NSE = 0.445).

On the other hand, all the reanalyses performed very well over the West coast, Mediterranean and Semi-arid regions. The NSE median value of each reanalysis is greater than 0.7 in these climate regions, and their performances are similar to (and sometimes better than) those obtained from observations. WFDEI-GPCC and WFDEI-CRU perform similarly to ERA-

Interim, except in the Semi-arid region where ERA-Interim (median NSE = 0.768) perform significantly better than those of WFDEI-GPCC (median NSE = 0.630) and WFDEI-CRU (median NSE = 0.580). Thus in the Semi-arid region, bias correction made ERA-Interim worse, possibly because of the relatively sparse network of weather stations in that region.

As mentioned previously in the data comparison section, the three global reanalyses do not adequately reproduce the observed annual cycle of precipitation in the humid Continental and Subtropical regions of the USA (see figures 4.10d – 4.10e). These same behaviours have been observed at the watershed scale. The inadequate representation of the seasonal precipitation cycle is the main cause leading to the poor performance of these reanalyses for hydrological modeling over both the Continental and humid Subtropical regions. Indeed, the hydrological model is not able to adequately simulate flows observed when precipitation seasonality is not well represented in input datasets.

The performance of the reanalyses was also assessed on the basis of simulated mean monthly flows over the calibration-validation period. The distribution of performances is similar to those shown in figure 4.12, but with slightly improved NSE values. This implies that the monthly bias structure of the reanalyses, not deficiencies at the daily scale, is the main obstacle to hydrological modelling.

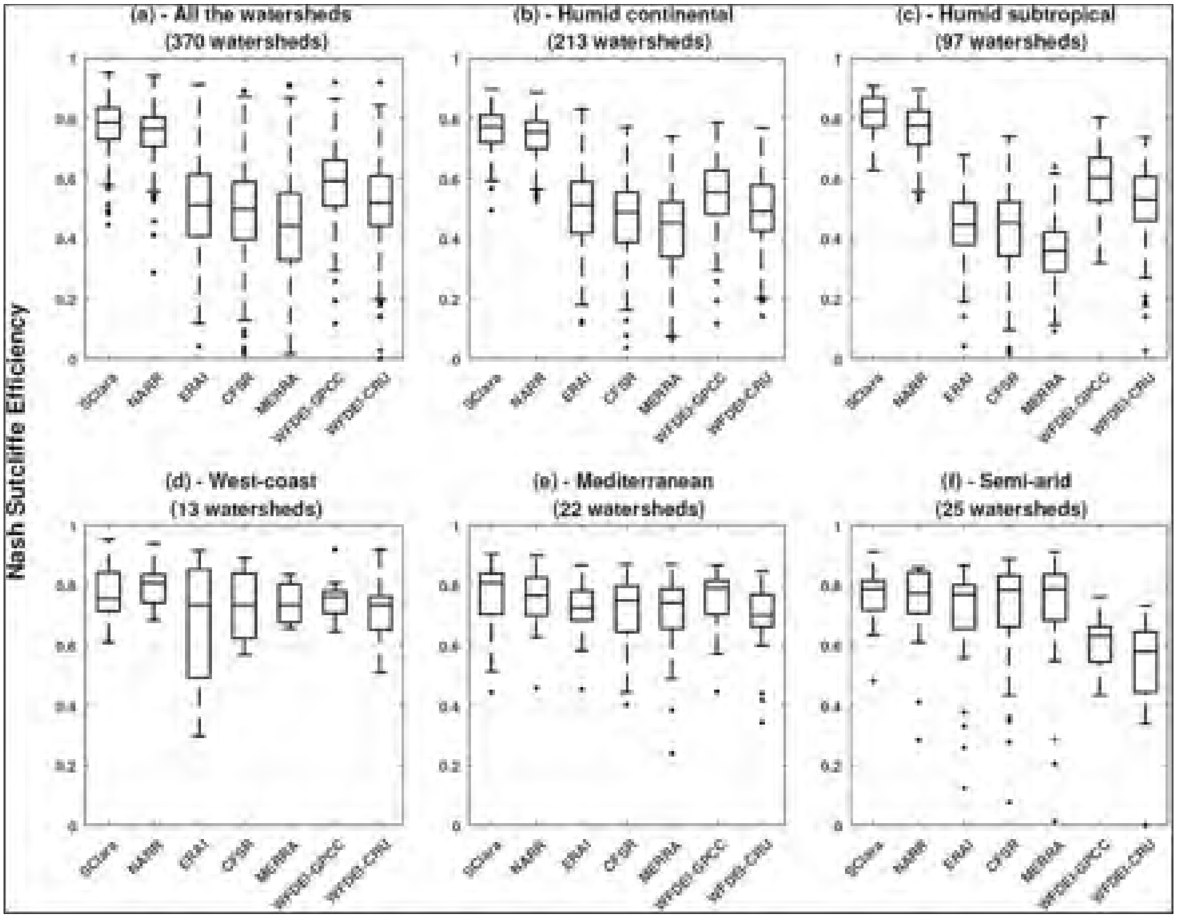

Figure 4.12 Distribution of performances based on daily discharges simulated in the validation period

The measured and simulated average annual hydrographs computed on the basis of mean daily flows over the calibration-validation period were compared in the Humid Continental and Humid Subtropical regions. Results are shown for 2 rivers: the East Fork White River at Columbus located in Indiana (in the Humid Continental region) (figure 4.13a), and the Flint River at Montezuma located in Georgia (in the Humid Subtropical region) (figure 4.13b). The simulated discharge NSE values (over the validation period) for the East Fork White River and the Flint River are respectively 0.791 and 0.828 for Santa Clara, 0.776 and 0.813 for NARR, 0.583 and 0.455 for ERA-Interim, 0.464 and 0.472 for CFSR, 0.547 and 0.376 for MERRA, 0.613 and 0.634 for WFDEI-PGCC, 0.592 and for 0.628 for WFDEI-CRU. These results are typical of most watersheds in each of the two humid regions. For both rivers, the dry biases of global reanalyses in winter led to a significant underestimation of discharges whereas the wet biases in summer led to a considerable overestimation of discharge. In general, the simulated discharges using the Santa Clara dataset also led to an underestimation of winter-spring discharges and an overestimation of summer-autumn discharges, but to a much lesser extent than those due to global reanalyses. This explains the low performances of global reanalyses for both rivers, and more generally, the low performances of global reanalyses in both Humid Continental and Subtropical regions.

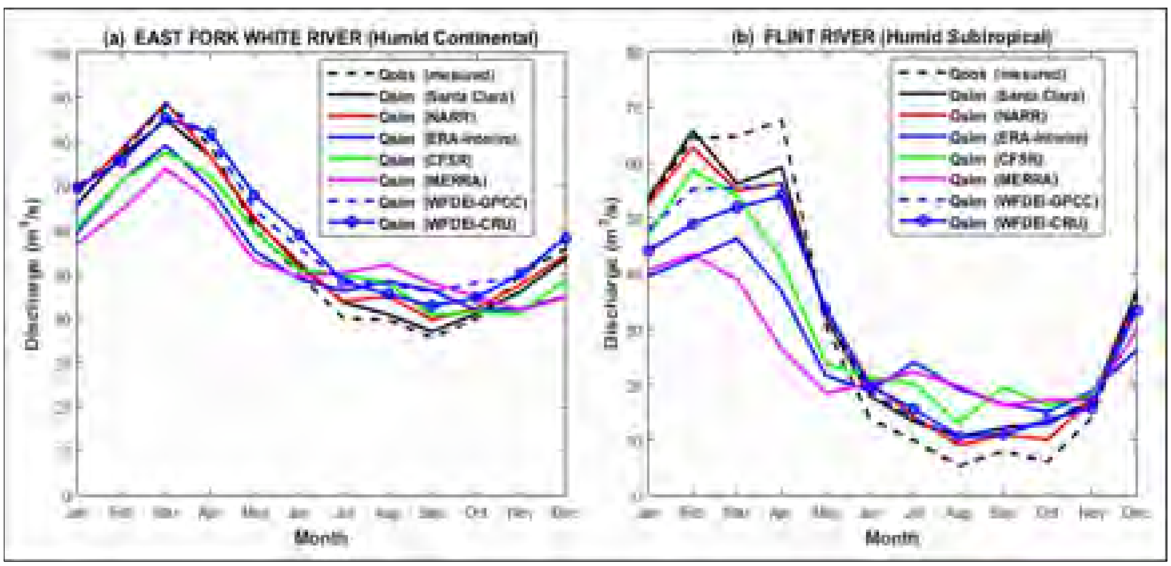

Figure 4.13 Comparison of the measured and simulated mean annual cycle hydrographs of the East Fork White River (a) and of the Flint River (b). Hydrographs are based on daily mean discharges during 1979–2003

4.6 Discussion

Precipitation and temperature are the two principal meteorological inputs for hydrological modeling. These data are sourced mainly from weather stations, although many regions in the world have a sparse network of monitoring stations, resulting in a severe limitation of hydrological studies.

In regions with sparse weather station coverage, reanalyses may offer a good alternative to station data; reanalyses in fact offer global coverage, and may be good proxies in the absence of surface observations, since they rely on global observations from multiple sources that are assimilated in a weather forecast model. However, the spatial resolution of reanalyses is relatively coarse, and the quality of their precipitation and temperature has to be validated in detail before being used for hydrological modeling.

To investigate the potential of reanalysis for use as proxies of surface observations of precipitation and temperature, four different atmospheric reanalyses were evaluated and compared to observations. In this work, observations are represented by the Santa-Clara gridded dataset. A comparison of precipitation from reanalyses to the gridded dataset showed that reanalyses are generally biased, especially in the Midwest and Humid subtropical regions. Temperature biases from reanalyses vary from season to season and from a reanalysis to another. Generally, temperatures exhibits smaller biases than precipitation, but MERRA temperature biases are consistently high in the Western USA during summer.

Overall, differences between reanalyses and observed gridded precipitation and temperature were judged to be sufficiently small to allow reanalysis outputs to be used directly for hydrological modeling, without any sort of bias correction needed. To that end, temperature and precipitation from reanalyses were averaged at the watershed scale on 370 watersheds in the USA. The HSAMI hydrological model was then calibrated to each dataset (reanalyses and gridded observations), and the river flow simulated by the hydrological model was evaluated against observed flows over a validation period. This approach provides an indirect validation of reanalyses that are used to force the hydrological model. It measures the differences between reanalyses and observations, by taking into account the consistency between precipitation and temperature, which is key for hydrological modeling. Proceeding without bias correction implies that all the differences between the observed and reanalysis fields are small enough to be taken into account through the adjustment of the hydrological model parameters. It also recognizes the fact that differences between gridded observations and reanalyses may be the result of biases in the reanalyses, gridded observations, or a combination of both. While weather surface observations are commonly recognized as constituting the most accurate representation of reality, they do suffer from biases, especially at the watershed scale, due to observational errors and, more importantly, to inhomogeneous coverage of weather stations, especially in the case of mountainous watersheds.

The results showed that adequately representing precipitation seasonality is critical, and that simulated river flows using NARR forcing are similar to the simulated streamflows using the gridded observations. This is linked to the NARR surface precipitation assimilation in its atmospheric model (Mesinger et al. 2006; Sheffield et al. 2012). Although this assimilation is done indirectly through latent heat profiles, it seems to be effective. It should be mentioned that the good capacity of NARR over the continental USA does not extend to Canada, where weather station coverage is much lower, especially in Northern Canada (Bukovsky and Karoly 2007; Langlois et al. 2009).

In the Humid Continental and Subtropical regions, the precipitation from ERA-Interim, CFSR and MERRA are significantly different from the gridded observation and NARR. These reanalyses do not assimilate surface precipitation data and rely on the physics of their weather forecast models to simulate precipitation, which they often do rather poorly, especially in the summer (Bosilovich 2013; Higgins et al. 2010). Indeed, in these climatic regions, summers are hot and humid because of the tropical atmospheric flow from the Gulf of Mexico. Most rainfall occurs as convective storms in the summer. These local events are not well simulated in global reanalyses, mainly because of their coarse resolutions. Moreover, precipitation is unevenly distributed over the year in both climatic regions, and their seasonality is highly sensitive to daily precipitation because of a weak mean annual cycle. For these reasons, global reanalyses are unable to adequately reproduce the seasonality of precipitation in Humid Continental and Subtropical regions. This explains the relatively poor ability of the three global reanalyses to produce an adequate simulation of the river flow by the hydrological model, in both the Humid Continental and Subtropical regions. Despite a specific calibration to each dataset, the hydrological model’s parameters were not able to compensate for the inadequate representation of the seasonality of precipitation by the global reanalyses.

For the other three climatic regions, the streamflows simulated using the global reanalyses were similar to those obtained from the gridded observation. This suggests that surface precipitation assimilation is not always essential for a good river flow simulation by hydrological modeling forced by global reanalyses. In the three western US climatic regions, the frequency and intensity of precipitation are both lower than in the Eastern climatic regions. In particular, for the West-coast (or Oceanic) and Mediterranean regions, precipitation is influenced by the proximity to the Pacific Ocean. Moreover, precipitation generating weather systems in the western US during the cold/wet season are much more dynamic; that is, precipitation is more strongly forced and likely more predictable. For these reasons, despite their coarse resolutions, global reanalyses manage to adequately represent precipitation seasonality, and therefore, lead to river flow simulations that are comparable to when gridded observations are used to force hydrological models.

Overall, the results confirm the potential of reanalyses as adequate forcings to hydrological models, despite some known weaknesses, such as their coarse resolutions and non-closure of the water budget due to the mixture of model data and observed data during each analysis process performed about every 12h (Lorenz and Kunstmann 2012; Trenberth et al. 2011). In addition, the separate analysis of the surface, the atmosphere and the ocean, as well as the change in time and space of the quantity and quality of assimilated data (Poli et al. 2010b; Wang et al. 2011), may introduce false variabilities, sudden changes, and trends in the reanalysis datasets. However, reanalyses will continue to improve in the future, and should result in even better accuracy for river flow simulations from hydrological models.

The goal of this paper was to present a comprehensive evaluation of reanalysis precipitation and temperature for streamflow simulations from hydrological models. The continental US was chosen due to its overall relatively good station coverage, thus allowing a robust validation benchmark for reanalysis temperature and precipitation. The real interest of reanalyses for hydrological studies lies in regions not well covered with surface weather stations, such as Northern Canada and the Arctic (Lindsay et al. 2014). Thus, further work should compare the accuracy of simulated streamflow using reanalyses to those using gridded observations, as a function of the density of surface weather stations. It is expected that the accuracy of simulated streamflows using gridded observations will decrease with the reduction of the density of the weather stations. On the other hand, global reanalyses should be less affected by the lack of weather stations since they simulate their own precipitation and rely on global data from different sources in their assimilation process.

For regions where precipitation in the global reanalyses are known to be not very good, a combination of reanalyses with the few available observations may be an interesting approach to develop better datasets. Such global datasets have been developed by post-processing (bias correction) global reanalysis with global observation-derived gridded datasets (Sheffield et al. 2006; Weedon et al. 2011; Weedon et al. 2014). However, global observation datasets also contain spatially dependent biases (Adam and Lettenmaier 2003; Cherry et al. 2007; Goodison et al. 1998). In addition, bias correcting precipitation and temperature independently can impact the spatial and temporal correlation between those variables (Li et al. 2014). Therefore, uncertainties also exist in these global forcing data.

Indeed, the comparison of ERA-Interim to its bias-corrected counterpart (WFDEI) shows that in the Western USA, ERA-Interim was just as good as or better than WFDEI. That is likely due to the fact that in that region, precipitation is more dynamic and thus well reproduced by reanalyses. Moreover, the relatively low density of weather stations in the semi-arid region might have reduced the efficiency of bias-correction and even possibly introduced errors leading to a degraded performance compared to the original ERA-Interim dataset.

Overall, results suggest that post-processing (bias correction) global reanalyses with global observation-derived gridded datasets will not automatically result into improved river flow simulation. This suggests that the quality of the underlying observational dataset is critical. This has important implications for the use of such datasets in remote regions.

Future work should compare the accuracy of simulated streamflow using reanalyses to those using global forcing data associated (e.g. compare ERA-Interim to WFDEI) over regions with sparse weather stations such as the Northern Canada.

A key advantage of reanalyses (ERA-Interim, CFSR and MERRA) is that they are updated on a regular basis (in near real time in some cases) which is important for many water resources management applications, which is not the case for global forcing databases (e.g. WFDEI does not extend beyond 2012).

4.7 Conclusion

In this study, precipitation and temperature data from the NARR regional reanalysis and from the ERA-Interim, CFSR and MERRA global reanalyses were compared to gridded Santa Clara observations over the contiguous USA. The potential use of precipitation and temperature data from reanalyses as direct inputs for hydrological modeling was investigated. Precipitation and temperature series were used to calibrate a lumped hydrological model and to simulate river flows over 370 watersheds located in five climatic regions over the continental USA. The Nash-Sutcliffe values of simulated river flows using reanalysis forcings were compared against simulated streamflows using gridded observations.

Results showed that the temperatures from reanalyses are generally comparable to those observed over the continental USA, except in the Western USA during summer for MERRA. Furthermore, there were some notable differences between reanalyses precipitation and observations, especially in the summer and winter.

Hydrological simulation was then used to indirectly validate the reanalyses. Over the five chosen climatic regions, the simulated river flows using the NARR forcing were as good as when the gridded observations were used. Overall, the Nash-Sutcliffe values of the simulated river flows using the global reanalyses were equal to those of the simulated river flows using the gridded observations, with the exception of the humid continental and subtropical regions, where precipitation seasonality is not well reproduced.

This study shows that reanalyses have a strong potential for use as proxies to weather station data, despite various differences between reanalyses and gridded observations. This potential is particularly promising in regions where weather station coverage is limited.

4.8 Acknowledgments

This work was funded through a Natural Science and Engineering Research Council collaborative research grant (NSERC-CRD) with Hydro-Québec, Rio-Tinto-Alcan, Ontario Power Generation and the Ouranos consortium on regional climatology and adaptation to climate change as industrial partners. These industrial research partners are greatly acknowledged for their direct and indirect contributions to this work. We also sincerely thank all the individuals and institutions that developed the datasets used in this work, and made them available to the scientific community. We hope that this work is a small contribution to their important effort.

CHAPITRE 5

ARTICLE 3. THE USE OF REANALYSES AND GRIDDED OBSERVATIONS AS WEATHER INPUT DATA FOR A HYDROLOGICAL MODEL: COMPARISON OF PERFORMANCES OF SIMULATED RIVER FLOWS ACCORDING TO THE WEATHER STATIONS

Gilles R.C. Essou[1], François Brissette[1] and Philippe Lucas-Picher[2]

[1] Département de Génie de la Construction, École de technologie supérieure, 1100 rue Notre-Dame Ouest, Montréal, Québec, Canada, H3C 1K3.
[2] Département des Sciences de la Terre et de l'Atmosphère, Université du Québec à Montréal, 405 Rue Sainte-Catherine Est, Montréal, Québec, Canada, H2L 2C4.

Article soumis à la revue Journal of Hydrometeorology en mars 2016

5.1 Abstract

Precipitation forcing is critical for hydrological modelling as it has a strong impact on the accuracy of simulated river flows. In general, precipitation data used in hydrological modelling are provided by weather stations. However, in regions with sparse weather station coverage, the spatial interpolation of the individual weather stations provides a rough approximation of the real precipitation fields. In such regions, precipitation from interpolated weather stations is generally considered unreliable for hydrological modelling. Precipitation estimates from reanalyses could represent an interesting alternative in regions where the weather station density is low because the accuracy of reanalyses does not depend directly on weather station data. This article compares the performances of river flows simulated by a watershed model using estimates of precipitation and temperature from reanalyses and gridded observations. The comparison was carried out according to the density of surface weather stations for 316 Canadian watersheds located in three climatic regions. Three state-of-the-art atmospheric reanalyses – ERA-Interim, CFSR and MERRA – and one gridded observations database over Canada – NRCan – were used. Results showed that the Nash-Sutcliffe values of simulated river flows using precipitation and temperature data from CFSR and NRCan were generally equivalent regardless of the weather station density. ERA-Interim

and MERRA performed significantly better than NRCan for watersheds with weather station densities of less than 1 station per 1000km^2 in the Mountain region. Overall, these results indicate that for hydrological modelling in regions with high spatial variability of precipitation such as Mountain regions, reanalyses perform better than gridded observations when the weather station density is low.

Keywords: Weather station density, reanalyses, gridded observations, hydrological modelling, climatic regions

5.2 Introduction

Precipitation forcing is critical for hydrological modelling, and has a strong impact on the accuracy of simulated river flows (Fekete et al. 2004; Lopes 1996). Precipitation measured by traditional weather stations often provides relatively accurate estimates at a few locations of a given region. However, measured precipitation is sometimes biased because of measurement errors, which have been shown to be as high as 10% for liquid precipitation (Adam and Lettenmaier 2003) and 100% for solid precipitation (Cherry et al. 2007; Goodison et al. 1998). In many parts of the world, weather station density is low. Recent studies have assessed the impact of precipitation station density on the accuracy of predictions by watershed models (Andréassian et al. 2001; Faurès et al. 1995). Duncan et al. (1993) studied the impact of weather station density on the accuracy of the flow predictions of a rural watershed in southern Quebec, and found this influence to be very strong, not only in terms of the accuracy of total runoff estimates, but also on the accuracy of peak flow and simulated peak time. Chaplot et al. (2005) studied the effect of the accuracy of spatial rainfall information on river flow modelling for two small watersheds located in the US and found that a small station density often led to inaccurate discharge estimates. In the light of these studies, it is clear that the density of weather stations has a strong impact on hydrological model predictions.

The spatial interpolation of individual weather stations provides a good approximation of precipitation in an area of interest. However, in regions with sparse weather station coverage, the interpolated precipitation fields will provide a rough estimate of precipitation levels. In fact, spatial interpolation methods always introduce some artifacts in interpolated datasets, and it is difficult to verify their realism in regions with low weather station densities (Daly 2006; Tozer et al. 2012). Therefore, in such regions, interpolated precipitation – and temperature – are generally considered unreliable for hydrological modelling (Mizukami and Smith 2012).

Reanalyses may represent a good alternative dataset of precipitation and temperature data for regions where weather stations are sparsely distributed or nonexistent. Reanalyses use a constant data assimilation scheme and numerical forecasting model, which ingest millions of available observations at a given time step over a given period (Dee et al. 2011; Saha et al. 2010). The observation sources include, but are not limited to, radiosonde, satellite, buoy, aircraft and ship reports. A constant reanalysis framework provides a dynamically consistent estimate of the climate state at each time step (Mesinger et al. 2006; Rienecker et al. 2011). In addition, reanalyses offer a global coverage and span three or more decades. They also provide hundreds of climate variables (Kalnay et al. 1996; Suarez et al. 2008; Uppala et al. 2005; Wang et al. 2011). Observations used in reanalyses vary because the observational network is constantly changing. Changes in the mix of observations over the duration of each reanalysis can produce spurious trends and artificial variability. However, spatial reanalysis resolutions and biases improve continuously with time. Commonly, reanalysis products are used to validate climate models in regions where weather stations are not available (Shiu et al. 2012; Sillmann et al. 2013a; Sillmann et al. 2013b).

Recent studies have also examined the potential of precipitation and temperature data from reanalyses for hydrological studies (Choi et al. 2009; Vu et al. 2012; Woo and Thorne 2006). Fuka et al. (2014) tested the value of CFSR precipitation and temperature as weather inputs for hydrological modelling of five watersheds representing different hydro-climatic regimes in the United States. They found that using CFSR precipitation and temperature data to force a watershed model provides river discharge simulations that are as good as or better than river discharges computed from models that are forced using traditional weather gauging stations, especially when these stations are more than 10 km away from the watershed outlet. More recently, Essou et al. (2016b) used precipitation and temperature series from the NARR, ERA-Interim, CFSR and MERRA reanalyses and from the gridded Santa Clara observations to calibrate a lumped hydrological model and to simulate river flows over 370 watersheds in the continental USA. They found that the simulated river flows using NARR forcing were as good as when gridded observations were used. Moreover, the Nash-Sutcliffe values of the river flows simulated using the other three reanalyses were equal to those from the gridded observations, with the exception of the Humid Continental and Subtropical regions, where precipitation seasonality is not well reproduced by the three reanalyses.

The objective of this study is to compare the accuracy of river flows simulated by a watershed model using precipitation and temperature estimates from reanalyses and gridded observations, as a function of the density of surface weather stations. To achieve this goal, 316 Canadian watersheds with different weather station densities and located in three climatic regions are considered. Three state-of-the-art atmospheric reanalyses – ERA-Interim, CFSR and MERRA – and one gridded observations database over Canada - NRCan - are used. For watersheds with low weather station densities, reanalysis-based estimates are expected to be more reliable, and therefore more efficient, for hydrological modelling than the gridded observations.

5.3 Watersheds of interest and data

5.3.1 Watersheds of interest

In this study, we consider 316 watersheds derived from the CANadian mOdel Parameter EXperiment (CANOPEX) database (Arsenault et al. 2015b). The size of the watersheds varies between 460 and 127,635 km^2. For the analysis, the watersheds are distributed into three climatic regions according to the Köppen-Geiger climate classification (Kottek et al. 2006): Mountain (144 watersheds), Boreal (149 watersheds) and Humid Continental or Atlantic Canada (23 watersheds) (figure 5.1a). The mean annual precipitation of the watersheds varies between 1 and 6.5 mm/day, and is more abundant over the coastal regions (figure 5.1b). The mean annual temperature varies between -6 °C and 7 °C, and decreases as we go from south to north (figure 5.1c). The daily mean precipitation and temperature (based on NRCan gridded estimates), and the daily mean discharge of the watersheds from each climatic region, are presented in table 5.1.

Table 5.1 Range of the watershed-average daily mean precipitation, temperature and discharge for each climate region

Climate region	Number of watersheds	Area (km^2)	Daily mean precipitation (mm/d)	Daily mean temperature (°C)	Daily mean discharge (mm/d/km^2)
Mountain	144	510 – 111,305	1.1 – 6.5	-4.5 – 6.6	0.1 – 6.4
Boreal	149	460 – 127,635	1.0 – 3.6	-6.0 – 5.0	0.1 – 2.6
Humid Continental	23	565 – 5,585	2.8 – 3.9	2.0 – 7.0	1.7 – 3.8

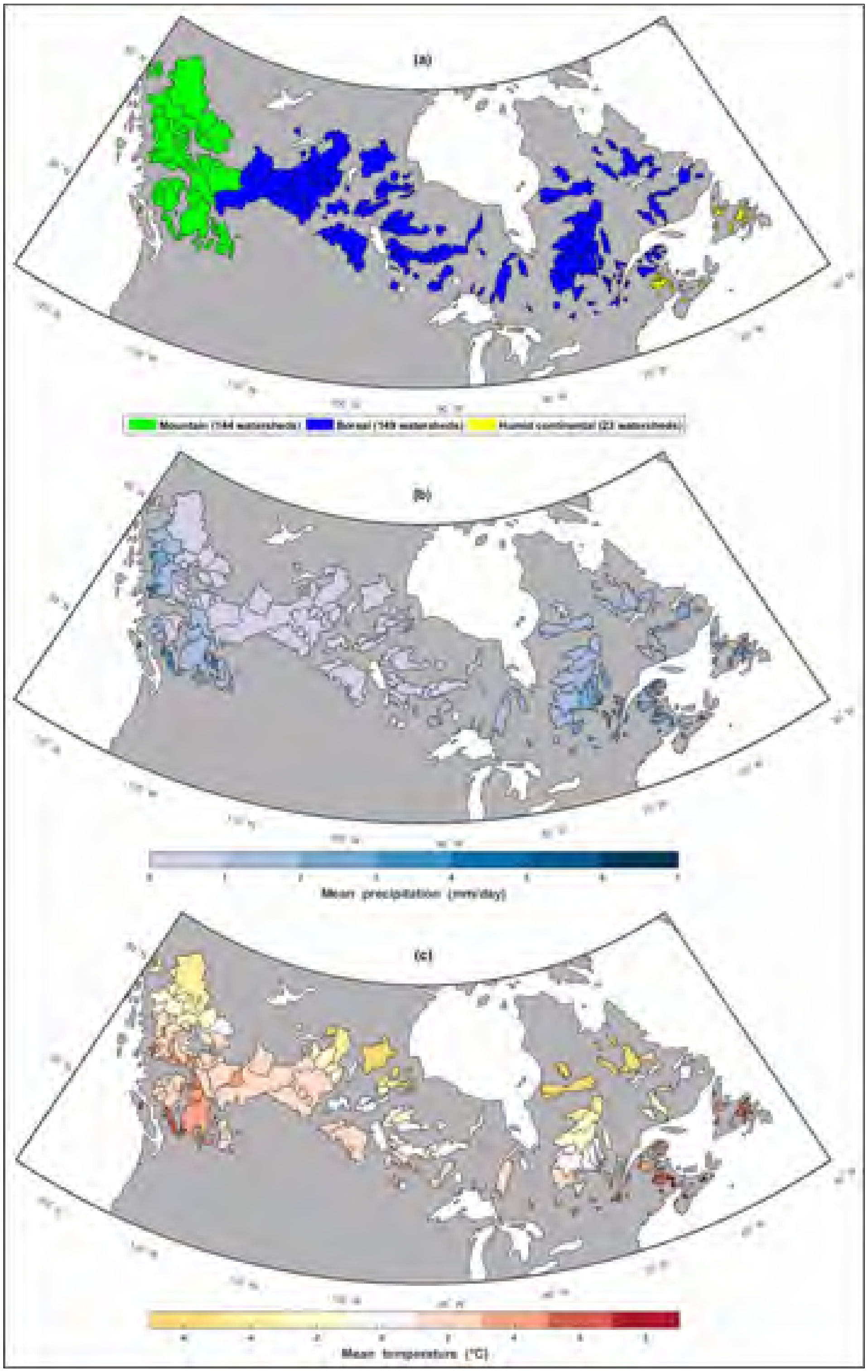

Figure 5.1 (a) The 316 watersheds of interest according to three climatic regions, (b) 1979-2010 mean annual precipitation (mm/day) and (c) 1979-2010 mean annual temperature (°C) based on NRCan gridded estimates

5.3.2 Data

Meteorological data: Daily precipitation and temperature data from NRCan and reanalyses are considered. NRCan is a 10-km gridded database developed by Natural Resource Canada (NRCan) from the interpolation of daily precipitation, and maximum and minimum air temperature data over the period of 1950-2010, using thin plate-smoothing splines (ANUSPLIN) (Hopkinson et al. 2011; Hutchinson 1995, 2004; Hutchinson et al. 2009). The number of active Environment Canada weather stations used varies from 2,000 to 3,000 for precipitation, and 1,500 to 3,000 for air temperature.

ERA-Interim is the latest global reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al. 2011). It covers the period of 1979-present, and uses a 4D-Variational (4D-VAR) data assimilation approach. The observations assimilated before 2002 are derived mainly from the data used for ERA-40 (Uppala et al. 2005). ERA-Interim is updated in near real time using data from the operational ECMWF forecast system (Dee et al. 2011). The temperature from ERA-Interim results from the assimilated surface temperature, while precipitation is produced by the weather forecast model. The horizontal resolution of ERA-Interim is 0.75° x 0.75° (about 80km x 80km).

The Climate Forecast System Reanalysis (CFSR) is produced by the National Centers for Environmental Prediction (NCEP). It is the first reanalysis produced from a coupled climate atmosphere-ocean-land surface system with an interactive sea ice component. It covers the period of 1979-present, and uses a 3D-VAR data assimilation approach (Saha et al. 2010). CFSR assimilates satellite radiance data rather than estimated temperature and humidity values (Wang et al. 2011). CFSR uses the Noah land surface model, which is forced with the NOAA Climate Prediction Center (CPC) pentad merged analysis of precipitation (Xie and Arkin 1997) and the CPC unified daily gauge analysis (Wang et al. 2011) instead of using the precipitation generated by the atmospheric model, which is considered too biased (Saha et al.

2010). The horizontal resolution of CFSR is 0.31° (longitude) x 0.31° (latitude) (about 35km x 35km).

The Modern Era Reanalysis for Research and Applications (MERRA) is developed by the Global Modelling and Assimilation Office (GMAO) of the National Aeronautics and Space Administration (NASA) in order to maximise the use of GMAO satellite observations in a climate context and to improve the closure of the hydrological cycle (Rienecker et al. 2011). MERRA covers the satellites era (1979-present), and is generated from version 5.2.0 of the Goddard Earth Observing System (GEOS) atmospheric model and a data assimilation system based on a 3D-VAR approach. The data assimilation system (DAS), the input data flux and their sources, observations, and error statistics are well documented in Suarez et al. (2008). The main specificity of MERRA consists in the use of an incremental analysis update (IAU) procedure to improve the closure of the water budget. The horizontal resolution of MERRA is 2/3° (longitude) x 1/2° (latitude) (about 75km x 55km).

Hydrometric data: Daily mean river flow data for each of the 316 watersheds from the HYDAT database (Coulibaly et al. 2013; Winkler 1993) are used. Data from HYDAT come from about 7,000 hydrometric stations across Canada.

A description of the databases used is presented in table 5.2.

Table 5.2 Description of the databases used in this study

Database	**Variable used**	**Acronym**	**Horizontal resolution**	**Assimilation scheme**	**Source**	**Reference**
NRCan	P, Tmin, Tmax	NRCan	10km x 10km	---	---	(Hutchinson et al. 2009)
ERA-Interim reanalysis	P, Tmin, Tmax	ERAI	80km x 80km	4DVAR	http://data-portal.ecmwf.int/	(Dee et al. 2011)
Climate Forecast System Reanalysis	P, Tmin, Tmax	CFSR	35km x 35km	3DVAR	http://cfs.ncep.noaa.gov/cfsr/	(Saha et al. 2010)
Modern Era Reanalysis for Research and Applications	P, Tmin, Tmax	MERRA	75km x 55km	3DVAR	http://disc.sci.gsfc.nasa.gov/mdisc/overview/index.shtml.	(Rienecker et al. 2011)
HYdrometric DATabase	Q	HYDAT	---	---	ftp://arccf10.tor.ec.gc.ca/wsc/software/HYDAT/	(Coulibaly et al. 2013; Winkler 1993)

P = Precipitation; Tmin = minimum temperature; Tmax = maximum temperature; Q = streamflow

5.4 Methodology

The mean annual precipitation and temperature from reanalyses are compared to those from NRCan to determine the systematic biases present. Each database is then used as input data for the calibration of a lumped hydrological model that simulates river flows. The accuracy of the simulated flows is then assessed with observations over a validation period. Simulated river flows are sorted according to the density of weather stations for each watershed. The period analyzed varies from one watershed to another according to the common period between hydrometric and meteorological data within the period of 1979-2010. The shortest period is 1979-1989 and the longest one is 1979-2010.

5.4.1 Data comparison

The precipitation and temperature data over each watershed are computed using the Thiessen polygon method (Thiessen 1911). Differences between reanalyses and NRCan mean seasonal temperature are assessed using bias statistics. Relative values – RBIAS– are computed for precipitation. NRCan is considered as the reference data in bias calculations. The bias is the difference between a reanalysis and NRCan for a given period. It indicates how much a reanalysis overestimates or underestimates relative to NRCan.

5.4.2 Weather stations densities

The weather stations considered in calculating station densities are those used by Environment Canada to develop the NRCan database. The station densities of each watershed are calculated in 3 steps. First, all the stations within each watershed or its surroundings (within a buffer of 20km) are preselected. The Thiessen polygons method was used to determine the weight of each weather station. Secondly, for each day of the period of interest, the number of operational weather stations (that provide data for the considered day) is determined. Third, the density of stations is computed as the ratio between the number of operational weather stations and the area of the watershed over the period analyzed.

5.4.3 Hydrological model and calibration strategy

The lumped conceptual hydrological model HSAMI (Fortin 2000) is used to simulate river discharges. HSAMI has been used to predict the hourly and daily flows of more than one hundred watersheds in Quebec. It has also been used operationally by Hydro-Québec over 100 watersheds for more than 30 years, as well as in climate change impact projects (Chen et al. 2012; Poulin et al. 2011). The HSAMI model has 23 calibration parameters (table 5.3): two for evapotranspiration, six for snowmelt, ten for infiltration and percolation, and five for the routing of surface runoff and interflow (figure 5.2). In HSAMI, four interconnected reservoirs contribute to the vertical water transfer balance: snow on the ground, surface water, unsaturated zones, and saturated zones. The horizontal water transfer is based on two unit hydrographs (one for surface runoff and one for delayed runoff) and one linear reservoir for groundwater flows.

The calibration was performed on the even years of each watershed study period, while the validation was based solely on the odd years. All the calibrations were performed using the Covariance Matrix Adaptation Evolution Strategy (CMAES) algorithm (Hansen and Ostermeier 1996, 2001). Arsenault et al. (2014) showed that CMAES is able to find optimal parameter sets for the HSAMI model.

The Nash-Sutcliffe Efficiency (NSE) metric (Nash and Sutcliffe 1970) was computed as an objective function based on the even years, with cross-validation on the odd years to take into account the different climatic trends (e.g., natural decadal or multi-decadal variability). The NSE is commonly used, and despite drawbacks such as heavily weighting peak flows, it is found to be the best objective function for reflecting the overall fit of a hydrograph (Servat and Dezetter 1991). The HSAMI model is calibrated for each dataset. For each watershed, 20 calibrations were performed and the best parameter set selected.

The non-parametric Wilcoxon test (Rakotomalala 2008) was performed to statistically evaluate simulated discharge at the 95% confidence level.

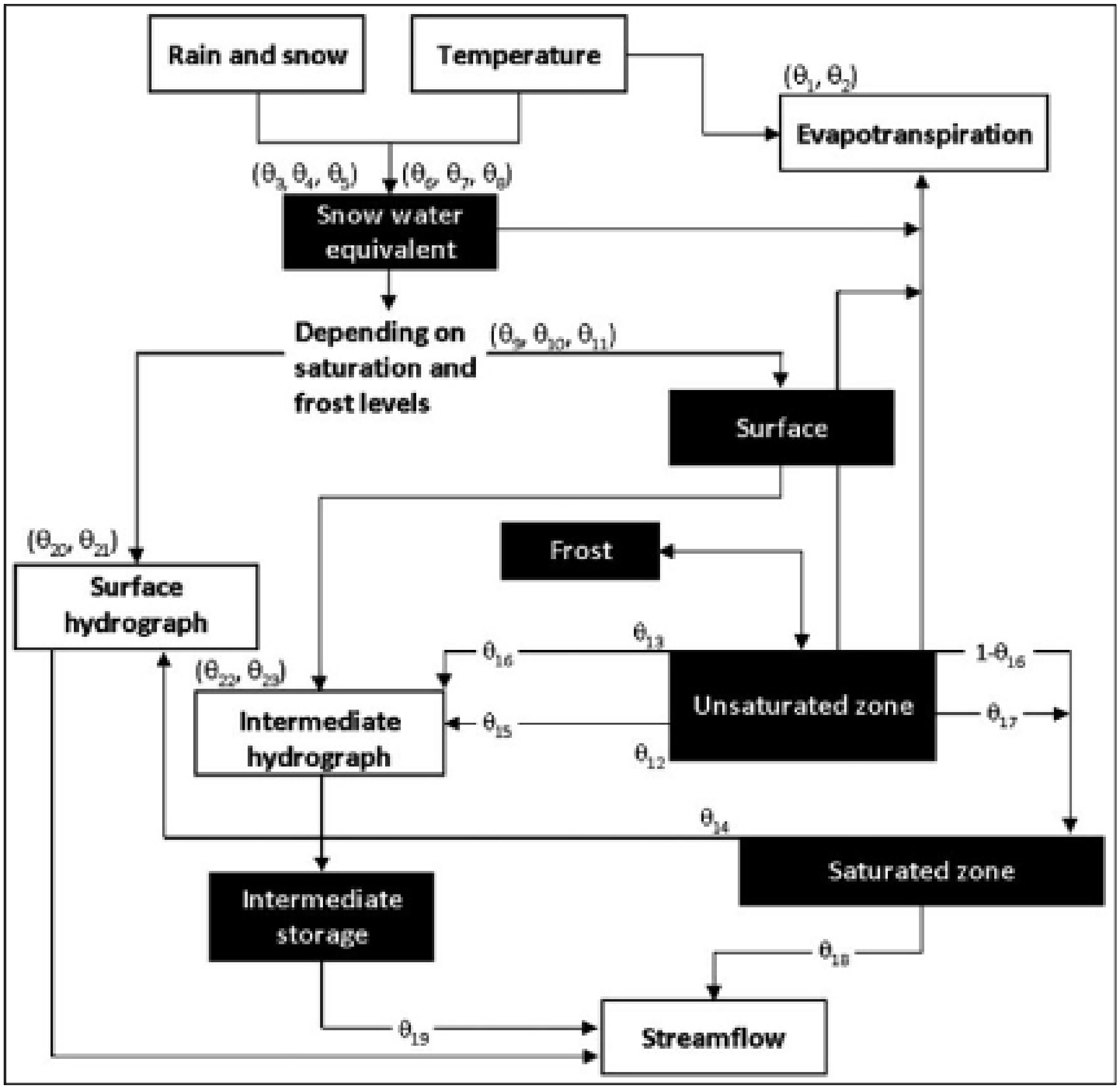

Figure 5.2 Flow chart of the HSAMI model. Black boxes represent conceptual reservoirs
Tirée de Minville et al. (2014)

Table 5.3 HSAMI model parameters

Process	Parameter	Unit	Description
Evapotranspiration	θ_1		Summer potential evapotranspiration (PET) multiplying factor
	θ_2		Winter proportion of summer PET
Snowmelt	θ_3	(cm/°C d)	Degree-day factor for daily snowmelt (based on maximum temperature)
	θ_4	(cm/°C d)	Degree-day factor for nighttime snowmelt (based on minimum temperature)
	θ_5	(°C)	Threshold temperature for daily snowmelt initiation
	θ_6	(°C)	Threshold temperature for nighttime snowmelt initiation
	θ_7	(°C)	Minimum temperature threshold to accelerate rain induced snowmelt
	θ_8		Parameter that relates snowmelt conditions to snow covered proportion of watershed
Infiltration and Percolation	θ_9		Parameter that relates freezing conditions to the proportion of surface runoff
	θ_{10}	(cm/d)	Daily rain rate necessary for 50% of surface runoff when the soil is completely dry
	θ_{11}	(cm/d)	Daily rain rate necessary for 50% of surface runoff when the soil is saturated
	θ_{12}	(cm)	Water in unsaturated zone which cannot flow by gravity

Process	Parameter	Unit	Description
Infiltration and Percolation	θ_{13}	(cm)	Maximum water quantity which can be contained in the unsaturated zone
	θ_{14}	(cm)	Maximum water quantity which can be contained in the aquifer before turning into surface runoff
	θ_{15}		Surface water proportion which flows through the intermediary hydrograph in normal conditions
	θ_{16}		Surface water proportion which flows through the intermediary hydrograph when unsaturated zone is full
	θ_{17}	(cm/d)	Outflow rate from the unsaturated zone to the saturated zone
	θ_{18}	(cm/d)	Outflow rate from the saturated zone which constitutes the base flow rate
Routing of Surface and Subsurface Runoff	θ_{19}	(cm/d)	Outflow rate from the interflow storage which constitutes the interflow rate
	θ_{20}	(day)	Time to peak of surface unit hydrograph
	θ_{21}		Shape parameter of surface unit hydrograph
	θ_{22}	(day)	Time to peak of interflow unit hydrograph
	θ_{23}		Shape parameter of interflow unit hydrograph

5.4.4 Evaluation of simulated flows accuracy

The simulated hydrographs from each dataset were compared to the observed hydrographs over the validation period. The performances of the simulations were computed using the Nash-Sutcliffe Efficiency metric values. First, the validation NSE values were compared considering all the 316 watersheds. Second, an evaluation as a function of the density of weather stations was carried out. Third, in each of the three climatic regions, performances obtained using NRCan were compared to those using reanalyses for watersheds where the density of weather stations was considered low. In the Mountain region, the spatial variability of precipitation is known to be great (Bailey et al. 1997). It is expected that in such a region, the considerable spatial variability of precipitation will greatly affect the accuracy of the interpolated datasets when the density of the available weather stations is low. Therefore, in such a region, NRCan might not perform as well as the reanalyses.

5.5 Results

5.5.1 Data comparison: Temperature and Precipitation

The difference between mean seasonal temperatures from reanalyses and NRCan varies and lies around ± 2.5°C (figures 5.3a-5.3f). Generally, reanalyses are warmer than NRCan in winter. Still in winter, ERA-Interim is mainly colder than NRCan in the Mountain region (biases of around -1.5°C) and warmer in the northern part of the Mountain region and in the Boreal and Humid Continental regions. The warmest biases are obtained in the Eastern Boreal region where warm biases reach 2.5°C (figure 5.3a). CFSR is warmer than NRCan in the Western Boreal where biases reach +2.5°C. Moreover, CFSR is colder than NRCan in the northern part of the Mountain region, but is generally similar to NRCan in the Eastern Boreal region (biases between ± 0.5°C) (figure 5.3b). MERRA is warmer than NRCan throughout Canada, and biases are mainly between +2°C and +2.5°C (figure 5.3c).

In summer, ERA-Interim is colder than NRCan in the northern part of the Mountain region (biases around -1°C) and slightly warmer in the southern part of the Mountain region, and in the Boreal and Humid Continental regions (biases between 0 and +1°C) (figure 5.3d). CFSR is warmer than NRCan in the Western Boreal region, where biases are between +0.5°C and +1°C. In the western part of the Mountain region, CFSR is generally colder than NRCan. However, in the northern part of the Mountain region and in the Eastern Boreal, CFSR is similar to NRCan (biases between ± 0.5°C) (figure 5.3e). MERRA is generally colder than NRCan in the Mountain region. However, in the Western and North Eastern Boreal regions, MERRA is warmer than NRCan (biases between +1.0°C and +2.5°C). In the South Eastern Boreal region, the temperature of MERRA is similar to that of NRCan in summer (biases between ± 0.5°C) (figure 5.3f).

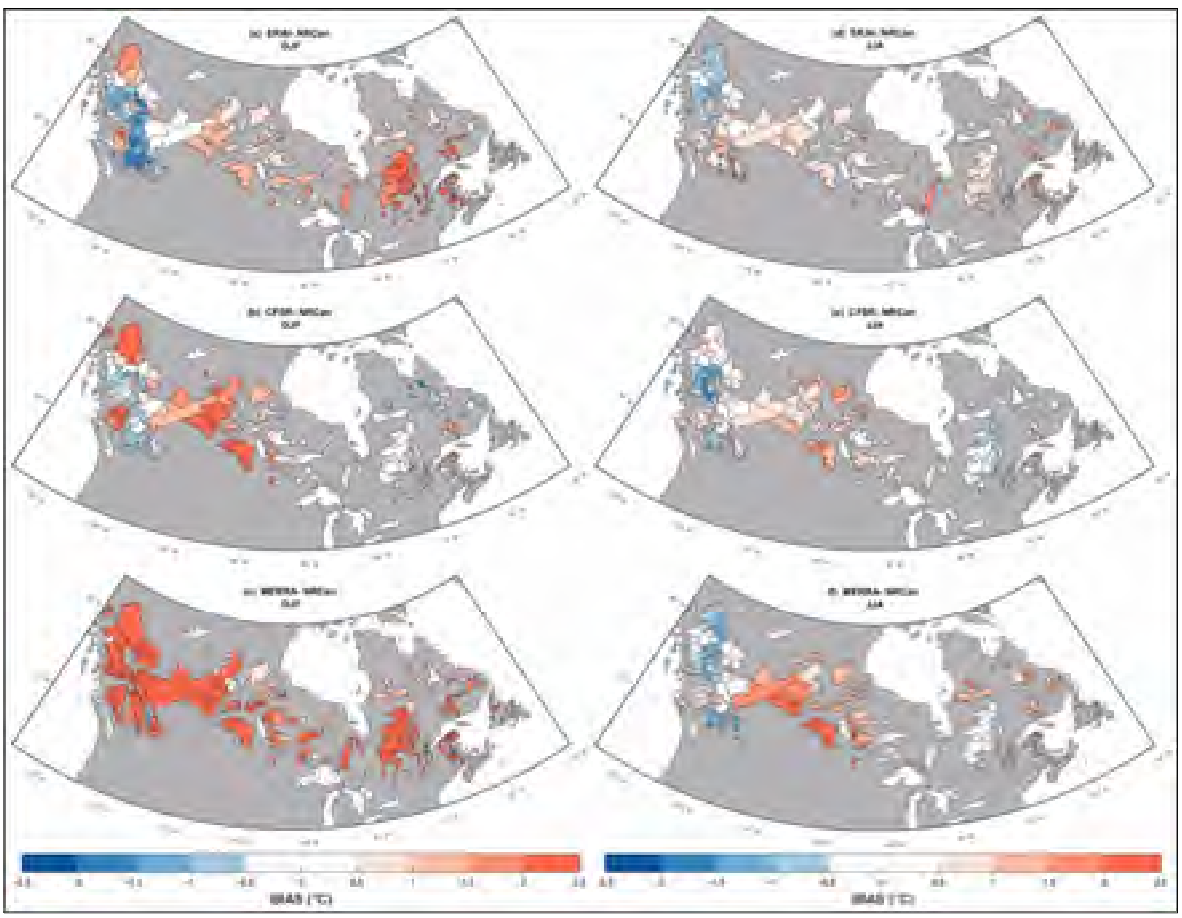

Figure 5.3 1979-2010 winter (DJF) and summer (JJA) temperature biases (°C) between reanalyses and observed gridded NRCan. (DJF = December-January-February; JJA = June-July-August)

For precipitation, the mean winter biases between the reanalyses and NRCan vary between -20% and +120% (figures 5.4a-5.4c). ERA-Interim is wetter than NRCan in the Mountain region (biases mainly between +30% and +60%) and dryer in the Eastern Boreal region (biases between -20% and -10%). In the Western Boreal and Humid continental regions, ERA-Interim is similar to NRCan, with biases between ± 10% (figure 5.4a). CFSR and MERRA are wetter than NRCan for almost all the watersheds. The wettest CFSR biases (bias > +90%) are obtained over 21% of the watersheds located in the Mountain and Western Boreal regions. For about 32% of the watersheds located in the Mountain and Boreal regions, the MERRA biases exceed +100%.

In the summer, all three reanalyses are wetter than NRCan over 91% of the watersheds for ERA-Interim and CFSR, and over 62% of the watersheds for MERRA. The wettest reanalysis biases are obtained in the Mountain region.

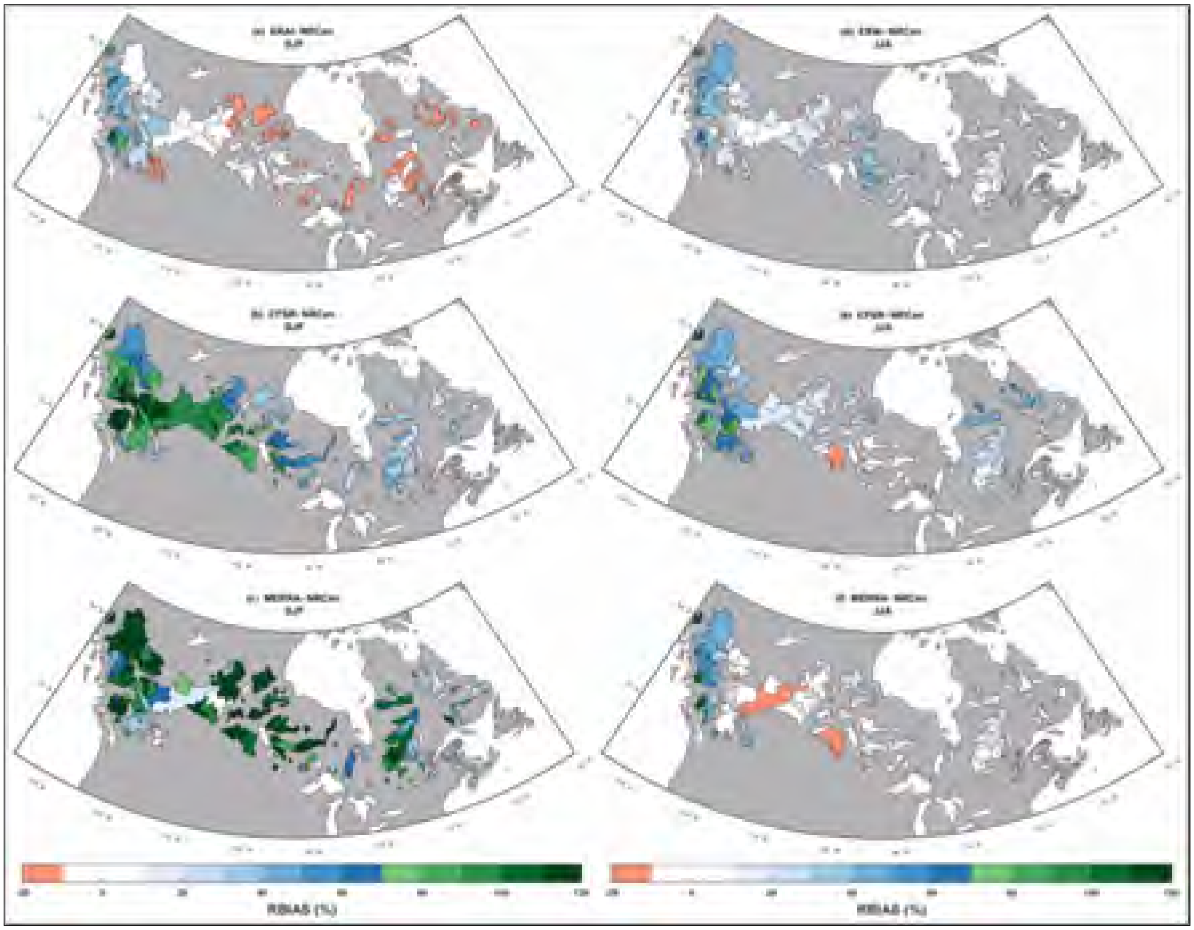

Figure 5.4 1979-2010 DJF and JJA precipitation relative bias (%) between reanalyses and NRCan gridded observations

5.5.2 Weather station density

The spatial distributions of the weather stations and their densities are shown in figures 5.5.a and 5.5.b (and table 5.4) respectively. Results show that densities are less than 10 stations per 1000km^2. Watersheds with a weather station density greater than 1 station per 1000km^2 are located mainly in Southwestern and Southeastern Canada. Results in figure 5.5.c show that for 67% of the watersheds, the stations density is less than 1 station per 1000km^2 and only 7% of them have 3 stations or more per 1000km^2. Overall, we consider that most of the 316 watersheds have a low weather station density.

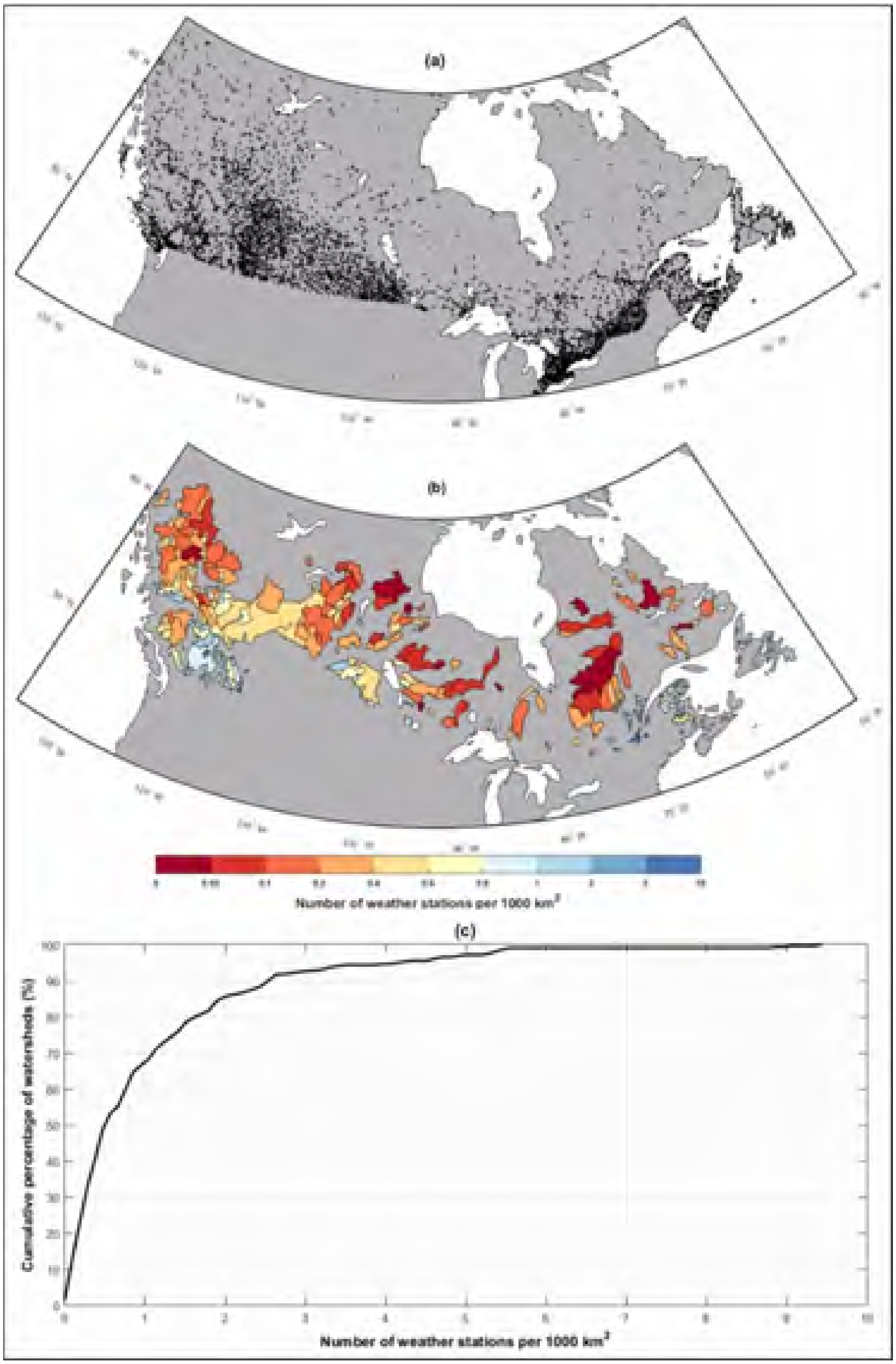

Figure 5.5 (a) Location of the weather stations, (b) spatial distribution of density of weather station for each watershed and (c) cumulative percentage of the number of watersheds according to the density of weather stations

Range of the densities of weather stations by climatic regions is presented in table 5.4.

Table 5.4 Range of the densities of weather stations

Climate region	**Number of watersheds**	**Density of weather stations (stations / 1000km^2)**	**Mean density of weather stations (stations / 1000km^2)**
Mountain	144	0 - 5	1
Boreal	149	0 – 10	1
Humid Continental	23	1 – 6	2

5.5.3 Comparison of the simulated river discharge

The NSE values obtained over the validation period using different inputs are shown in figure 5.6. The results from figure 5.6.a show that when all the 316 watersheds are considered, the median NSE values are 0.8 for NRCan, 0.81 for ERA-Interim, 0.8 for CFSR and 0.77 for MERRA. Moreover, the NSE values are greater than 0.6 for 90% of the watersheds for NRCan and ERA-Interim, 85% for CFSR and 81% for MERRA. Globally, all the datasets perform satisfactorily. The performance of NRCan seems to be better than that of MERRA. A comparison of each watershed shows that the NSE values for ERA-Interim are greater than those for NRCan for only 52% of the watersheds. The CFSR and MERRA NSE values are greater than those for NRCan for 39% of the watersheds. Results from the Wilcoxon statistical test showed that the performance of NRCan is generally equivalent to those for the reanalyses.

However, some significant differences appear when the results are sorted according to climatic regions. In the Mountain region, the median NSE values for the 144 watersheds are 0.83 for NRCan; 0.86 for ERA-Interim; 0.83 for CFSR, and 0.85 for MERRA (figure 5.6.b). In addition, the NRCan NSE values are lower than those for ERA-Interim, CFSR and

MERRA, respectively for 72%, 47% and 59% of watersheds. Thus, except for CFSR, reanalysis NSE values are greater than those for NRCan on most of the watersheds in the Mountain region. However, the results of the statistical test indicate that only the ERA-Interim NSE values are significantly greater than those for NRCan.

In the Boreal region, the median NSE values for the 149 watersheds are 0.76 for NRCan and ERA-Interim, 0.72 for CFSR, and 0.70 for MERRA (figure 5.6.c). From the NSE values for each watershed, it can be seen that the values for NRCan are greater than those for ERA-Interim, CFSR and MERRA, respectively for 59%, 64% and 72% of the watersheds. Although the NRCan NSE values are greater than those for the reanalyses for most of the watersheds, the statistical test shows that differences between NRCan and reanalyses NSE values are not statistically significant, except for MERRA.

The results in figure 6.d show that for the 23 watersheds in the Humid Continental region, the NSE median values for NRCan, ERA-Interim, CFSR and MERRA are 0.79, 0.76, 0.74 and 0.71, respectively. In addition, the NSE values for NRCan are greater than those for the reanalyses for at least 87% of the watersheds. The results of the statistical test show that the NSE values for NRCan are significantly greater than those for the reanalyses, except for ERA-Interim.

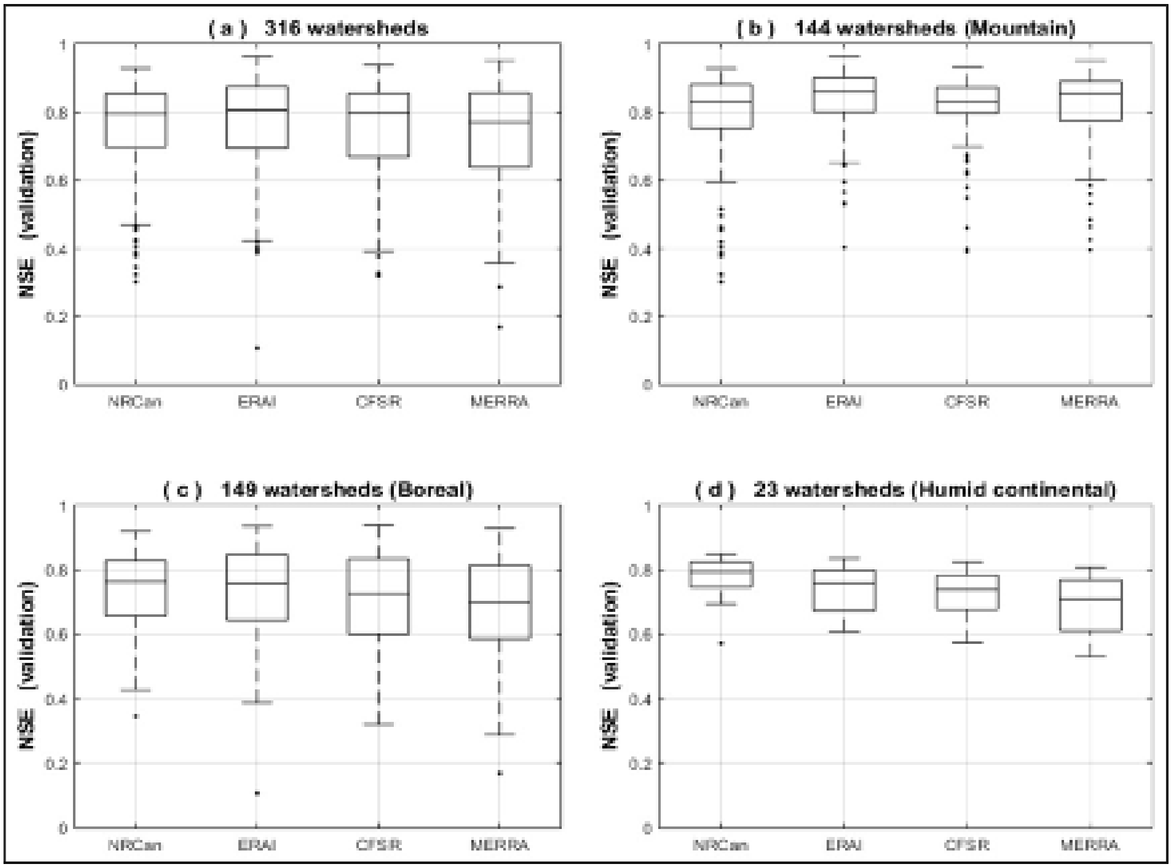

Figure 5.6 NSE of the simulated river flow with HSAMI over the validation period using NRCan, ERA-Interim, CFSR and MERRA

Results in figure 5.7 show the influence of the watershed size and the density of weather stations on the NSE, computed using the reanalyses and NRCan. Overall, the performance of the reanalyses increases with the watershed size, while the performance of NRCan is not dependent on the density (figure 5.7.a). Conversely, the performance of NRCan increases (but weakly) with the density of weather stations, while the performance of reanalyses is not influenced by the density, although a downward trend is seen starting from a density of 1.56 stations per 1000 km^2 and above (figure 5.7.b).

These results are not surprising since it is known that the accuracy of precipitation from gridded observations depends on the density of weather stations, which has an impact on the hydrological simulation performance (Chaplot et al. 2005; Tozer et al. 2012; Vischel 2006).

However, it is important to note that the performance of the hydrological simulations will not increase perpetually with the density of weather stations. Indeed, recent studies have shown that beyond an optimal threshold of the density of weather stations, the addition of more stations has no impact on the performance of watershed river flows simulated by a hydrological model (Arsenault and Brissette 2014a). On the other hand, precipitation from the reanalyses is produced by weather forecast models, and is therefore not directly dependent on precipitation measured by surface weather stations. For this reason, the accuracy of precipitation from these reanalyses, and hence their hydrological modelling performance, is not directly influenced by the density of weather stations, but among other things, is more dependent on the spatial resolution of the weather forecast models. This explains the influence of the size of the watersheds on the hydrological modelling performance of the three reanalyses.

Consequently, the drop in the performance of the reanalyses observed beyond 1.56 stations per $1000km^2$ is not actually related to the density of the weather stations, but rather, on the size of the watersheds. Indeed, all the watersheds with at least 1.56 weather stations per $1000km^2$ are small in size (with sizes below $32000km^2$, and more than 64% of them are less than $1000km^2$). According to the results of figure 7.a, the weakest reanalyses performance is generally obtained for watersheds with such areas.

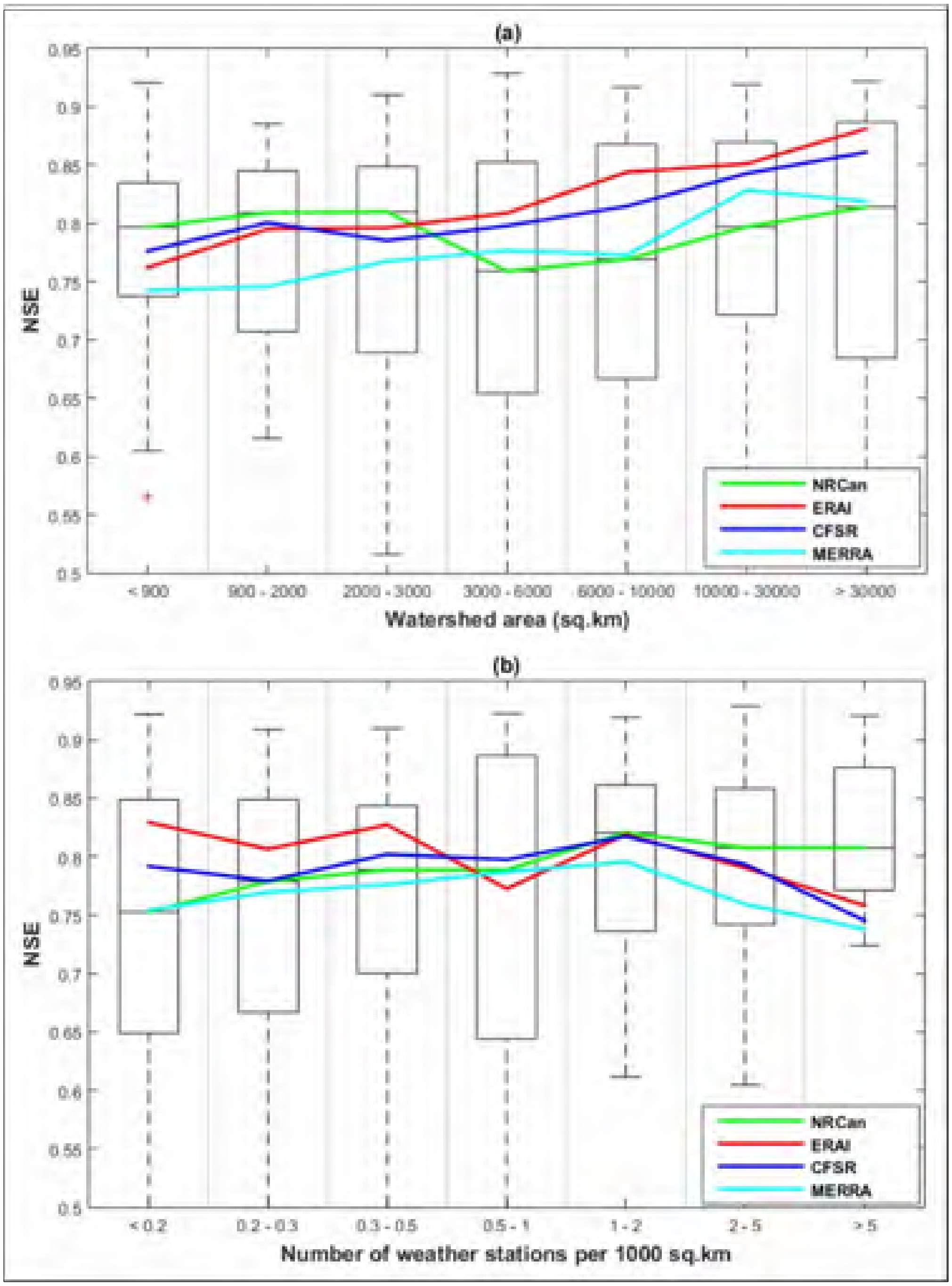

Figure 5.7 Distribution of reanalyses and NRCan NSE values (a) according to the size of the watershed and (b) the density of weather stations. The boxplots show the distribution of the NRCan NSE values. The bins were selected such that each boxplot would include 50 watersheds, except for the one on the extreme right, which includes only 16 watersheds. The median of the NRCan NSE values are connected by the green line. The other lines connect the median of the NSE values for the reanalyses, but their corresponding boxplots are not shown in order to avoid overloading the figure

The performance of the reanalyses and NRCan are compared according to the density of weather stations in the different climatic regions. The results are shown in figure 5.8.

Mountain region (144 watersheds)

The weather station density is greater than 3 stations per 1000km^2 for 7 watersheds in the Mountain region (figure 5.8.a). For these watersheds, the median NSE values are 0.87, 0.83, 0.8 and 0.85 for NRCan, ERA-Interim, CFSR and MERRA, respectively. The NRCan NSE values are greater than those for the reanalyses for 71% of these watersheds. However, there is statistically no significant difference between the performances of NRCan and the reanalyses.

Similarly, there is no significant difference between the NRCan NSE values and those for the reanalyses for the watersheds where the density of weather stations is between 2 and 3 stations per 1000km^2. The median NSE values are 0.76 for NRCan and CFSR, 0.78 for ERA-Interim, and 0.7 for MERRA (figure 5.8.b).

Where the station density is between 1 and 2 stations per 1000km^2, the median NSE values are 0.84 for NRCan and MERRA, 0.85 for ERA-Interim, and 0.83 for CFSR, and there is no significant difference between the performance of the reanalyses and that of NRCan (figure 5.8.c).

About 67% of the watersheds in the Mountain region have a density of weather stations lower than 1 station per 1000km^2. For these watersheds, the median NSE values are 0.84 for NRCan, 0.88 for ERA-Interim, 0.85 for CFSR, and 0.86 for MERRA (figure 5.8.d). The NSE values for ERA-Interim and MERRA are greater than those for NRCan on most of the watersheds (73% and 60%, respectively). However, the NSE values for CFSR are greater than those for NRCan for only 44% of the watersheds. The statistical test shows that the performance of ERA-Interim and MERRA is significantly greater than that for NRCan, whereas CFSR and NRCan perform equally.

Overall, in the Mountain region, all the reanalyses, except for CFSR, significantly outperform NRCan when the station density is lower than 1 station per 1000km^2. For higher weather station densities, the performances of the reanalyses and NRCan are similar.

Boreal region (149 watersheds)

In the Boreal region, the median NSE values are 0.8, 0.7, 0.74 and 0.7 for NRCan, ERA-Interim, CFSR and MERRA, respectively (figure 5.8.e). Moreover, the NSE values for NRCan are greater than those for the reanalyses for at least 67% of the watersheds, and the statistical test reveals that the performance of NRCan is significantly better than that of the reanalyses, except for CFSR.

Where the weather station density is between 2 and 3 stations per 1000km^2, the NRCan NSE values (median = 0.82) are greater than these of the reanalyses (median = 0.80) for at least 71% of the watersheds (figure 5.8.f). However, the NRCan NSE values are statistically similar to those for reanalyses.

Considering the watersheds where the weather station density is between 1 and 2 stations per 1000km^2, the median NSE values are 0.81 for NRCan, 0.77 ERA-Interim, 0.8 for CFSR, and 0.74 for MERRA (figure 5.8.g). Although the NSE values for NRCan are greater than those for the reanalyses for at least 71% of the watersheds, there is no significant difference between the performance of NRCan and those for reanalyses, except for MERRA.

Similarly, the NSE values for NRCan (median = 0.74) are significantly greater than those for MERRA (median = 0.67), but are statistically equivalent to the NSE values for ERA-Interim (median = 0.75) and CFSR (median = 0.71) when the weather station density is lower than 1 station per 1000km^2 (figure 5.8.h).

It is clear from this analysis that for watersheds considered in the Boreal region, NRCan and CFSR have statistically equivalent performances, regardless of the station density. Moreover, NRCan is significantly better than ERA-Interim when the weather station density is high (more than 3 stations per 1000 km^2).

Humid continental region (23 watersheds)

In the Humid continental region, the NRCan NSE values (median = 0.78) are statistically equivalent to those for ERA-Interim (median = 0.76) and CFSR (median = 0.74), but are significantly greater than those for MERRA (median = 0.74), when the weather station density is greater than 3 stations per 1000km^2 (figure 5.8.i).

Similar results were obtained for watersheds with weather station densities between 2 and 3 stations per 1000km^2. For these watersheds, the median NSE values are 0.81 for NRCan, 0.76 for ERA-Interim, 0.77 for CFSR and 0.73 for MERRA (figure 5.8.j).

When the weather station density is between 1 and 2 stations per 1000km^2, NRCan and the reanalyses perform similarly. The NSE values for NRCan, ERA-Interim, CFSR and MERRA are respectively 0.7, 0.66, 0.67 and 0.61 (figure 5.8.k).

Similarly, for lower weather station densities (less than 1 station per 1000km^2), the NSE values for NRCan (median = 0.82) are statistically equivalent to those for ERA-Interim (median = 0.82), CFSR (median = 0.76) and MERRA (median = 0.77) (figure 5.8.l).

Thus, in the Humid continental region, the performance of NRCan is statistically equivalent to those for ERA-Interim and CFSR, regardless of the weather station densities. However, NRCan performs significantly better than MERRA when the weather station density is greater than 2 stations per 1000km^2.

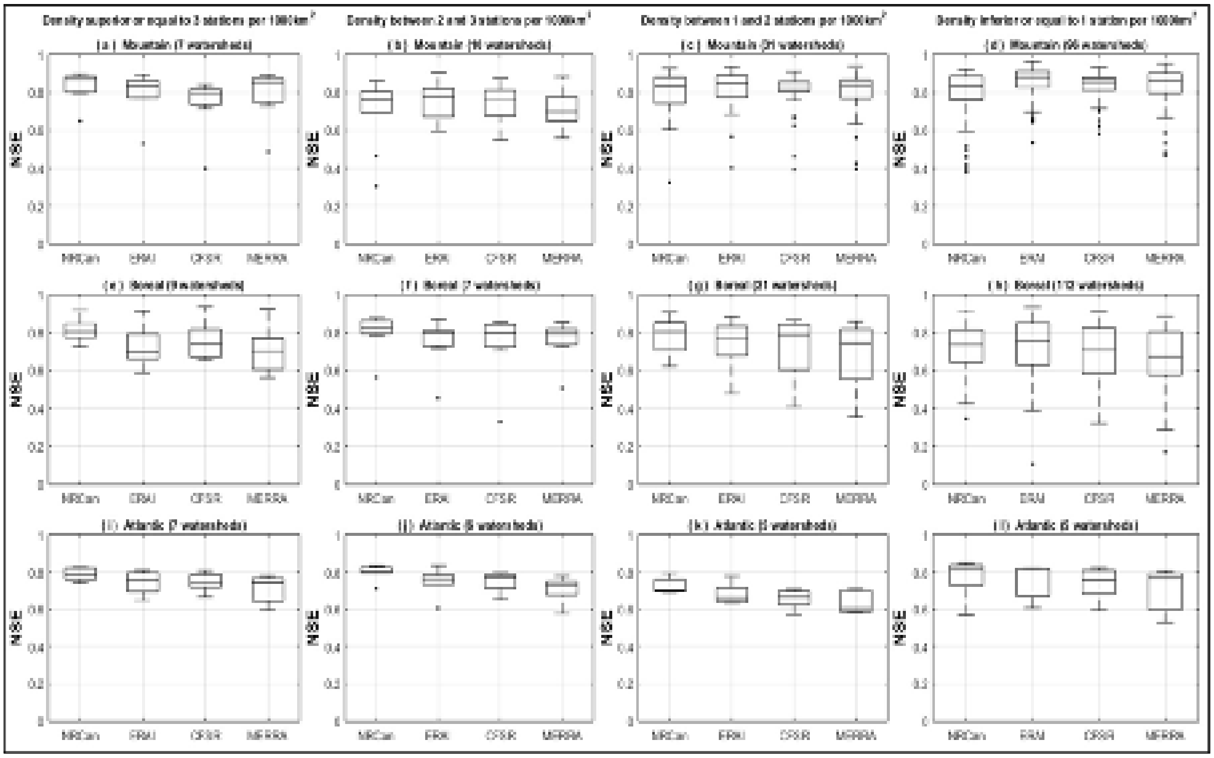

Figure 5.8 Distribution of the NSE values for reanalyses and NRCan according to the weather stations density and climatic regions

5.6 Discussion

Gridded observation datasets developed from the spatial interpolation of weather stations are usually useful for hydrological modelling. However, their credibility is questionable in regions where weather stations are sparsely distributed. Meteorological data from global reanalyses can represent a good alternative to gridded observations in forcing hydrological models in regions where weather stations are sparsely distributed. Global reanalyses provide a physically consistent estimate of weather events and rely on global observations from multiple sources that are assimilated in a weather forecast model. However, spatial resolutions of reanalyses are relatively coarse, and their data are generally biased. Nevertheless, biases and model resolutions are steadily improving.

This work compares the use of reanalyses instead of gridded observations to force hydrological models in regions with few conventional weather stations. To investigate the potential of the reanalyses as proxies of temperature and precipitation from weather stations, three atmospheric reanalyses – ERA-Interim, CFSR and MERRA – were evaluated and compared to gridded observations, NRCan. First, temperature and precipitation from reanalyses were spatially averaged over 316 Canadian watersheds and compared to NRCan. Second, for each of the watersheds, the HSAMI lumped hydrological model was calibrated to each reanalysis and NRCan dataset. The river discharges simulated by the hydrological model was evaluated against observed discharges over a validation period. The performances of the discharges simulated using precipitation and temperature from reanalyses and NRCan were compared according to the density of weather stations. About 67% of the watersheds have less than 1 station per 1000km^2, and therefore provide a good representation of regions with a sparse distribution of weather stations.

Results from the temperature and precipitation comparison showed some difference between the reanalyses and NRCan. Overall, mean seasonal temperature differences between the reanalyses and NRCan are relatively low, especially in the summer. Generally, the differences are lower between NRCan and both CFSR and ERA-Interim. This is possibly linked to satellite radiance assimilated by CFSR (Wang et al. 2011) and the land surface temperature assimilated by ERA-Interim (Dee et al. 2011; Simmons et al. 2010). In general, precipitation from ERA-Interim is closer to that of NRCan as compared to the other reanalyses. However, the three reanalyses tended to be wetter than NRCan. Differences between precipitation from NRCan and that from the reanalyses are particularly great in the Mountain region, where orographic precipitation is predominant (Bailey et al. 1997; Gervais et al. 2014). These differences are likely explained by biases in the reanalyses, or in NRCan, or in both. In fact, the reanalyses are possibly unable to adequately represent orographic precipitation because of their coarse resolutions, which smooth topography. On the other hand, the orographic precipitation is possibly smoothed in NRCan by the spatial interpolation of the few available weather stations. Overall, with such uncertainty in the reanalyses and the gridded observations, it is difficult to determine which one is the most accurate. However, hydrological modelling results indirectly highlight the quality of the reanalyses and NRCan datasets.

The calibration of the hydrological model for each dataset by finding the optimum fit of the 23 parameters of the model filters – to some extent – the errors of the datasets. However, the calibration has its own limits and the simulated discharge will always be dependent on the quality of the forcing data.

Globally, for the 316 watersheds, similar performances by HSAMI were obtained when forced by NRCan or by the reanalyses. Results showed that the density of weather stations has an impact on the performance of NRCan, but not on the performance of the reanalyses. This is in line with expectations. However, the performance of NRCan is not directly proportional to the weather station density. This is explained by the fact that for a given watershed, there is a threshold beyond which an increase in the density of weather stations will stop improving the performance of hydrological simulations using a lumped hydrological model (Arsenault and Brissette 2014a). Results also showed that when the density of weather stations is greater than 3 stations per 1000km^2, the performance of NRCan tends to be statistically similar or better than those for reanalyses. Conversely, when the weather station density is low (less than 1 station per 1000km^2), the performances of the reanalyses tends to be statistically equal to or greater than those for NRCan; this is particularly the case in the Mountain region, where the differences in precipitation between the reanalyses and NRCan are the largest, and ERA-Interim and MERRA perform significantly better than NRCan.

These results validate the fact that precipitation and temperature from reanalyses are globally more accurate than those from gridded observations, especially in the Canadian Mountain region, when few surface weather stations are available. This means that in such regions, reanalyses should be used instead of gridded observations, to force hydrological models. Moreover, these results also suggest that reanalyses should be of a great interests for hydrological modelling in regions such as Northern Canada, which are not well covered with surface weather stations (Lindsay et al. 2014).

Although this study was performed on Canadian watersheds, it can be repeated for other regions of the world, and similar results should be obtained.

One of the limitations of this study is the use of a lumped hydrological model for river flow simulations. This is due to the large number of watersheds considered and to the high computational costs that would result from the use of a distributed hydrological model. However, if a distributed hydrological model was used, the individual performance of each database may be different for large watersheds, but the general trend of performances and the main findings of this study may not change.

5.7 Conclusion

This study compared precipitation and temperature data from the ERA-Interim, CFSR and MERRA global reanalyses to gridded NRCan observations over 316 watersheds located in three climatic regions in Canada. Moreover, these precipitation and temperature data were used to force a lumped hydrological model, and the Nash-Sutcliffe values of the simulated river flows were compared as a function of the density of surface weather stations.

Results showed that temperature data from reanalyses are similar to that of the gridded observations of NRCan in the summer. Nevertheless, significant temperature differences were found between reanalyses and NRCan in the winter. Reanalyses tend to be considerably wetter than NRCan during winter and summer in Western Canada, mainly in the Mountain region.

Over the 316 watersheds, the Nash-Sutcliffe values of the reanalyses were statistically equivalent to those for NRCan. However, as expected, the analysis according to the weather station density showed that in the Mountain region, the performances of the reanalyses, especially ERA-Interim and MERRA, were significantly better than those for NRCan when the surface weather station density is less than 1 station per 1000km^2.

Overall, this study showed that compared to the gridded observations, reanalyses represent a reliable proxy to data from weather stations in complex terrain regions and where surface weather stations are sparsely distributed.

5.8 Acknowledgments

This work was funded through a Natural Science and Engineering Research Council collaborative research grant (NSERC-CRD) with Hydro-Québec, Rio-Tinto-Alcan, Ontario Power Generation and the Ouranos consortium on regional climatology and adaptation to climate change as industrial partners. These industrial research partners are greatly acknowledged for their direct and indirect contributions to this work. We also sincerely thank all the individuals and institutions that developed the datasets used in this work, and made them available to the scientific community.

CHAPITRE 6

ARTICLE 4. IMPACTS OF COMBINING REANALYSES AND WEATHER STATION DATA ON THE ACCURACY OF DISCHARGE MODELING

Gilles R.C. Essou[1], François Brissette[1] and Philippe Lucas-Picher[2]

[1] Département de Génie de la Construction, École de technologie supérieure, 1100 rue Notre-Dame Ouest, Montréal, Québec, Canada, H3C 1K3.
[2] Département des Sciences de la Terre et de l'Atmosphère, Université du Québec à Montréal, 405 Rue Sainte-Catherine Est, Montréal, Québec, Canada, H2L 2C4.

Article soumis à la revue Journal of Hydrology en mars 2016

6.1 Abstract

Reanalyses are important sources of meteorological data. Recent studies have shown that precipitation and temperature data from reanalysis present a strong potential for hydrological modelling, especially in regions with a sparse observational network. The objective of this study is to evaluate the impacts of the combination of three global atmospheric reanalyses – ERA-Interim, CFSR and MERRA – and one gridded observation dataset on the accuracy of hydrological model discharge simulations. Two combination approaches were used. The first one combined reanalyses and the observational database using a weighted average of the precipitation and temperature inputs. The second one consisted in using all meteorological inputs separately and combining the simulated hydrographs. The combinations were performed over 460 Canadian watersheds (representing regions with a low density of weather stations) and 370 US watersheds (representing regions with a higher density of weather stations). Results showed significant improvements in the simulated discharges for 68% and 92% of the Canadian watersheds for the input combinations and output combinations, respectively. Moreover, both approaches led to significant improvements in the simulated discharges for 72% of the US watersheds studied. For all watersheds where simulated discharges using observational data had a Nash Sutcliffe efficiency (NSE) lower than 0.5, the combination with reanalyses resulted in a median NSE increase of 0.3. This indicates that

reanalysis can successfully compensate for deficiencies in the surface observation record and provide significantly better hydrological modelling performance.

Keywords: Reanalyses, observations, hydrological modelling, calibration, data combination

6.2 Introduction

It is well known that the quality of weather data used as input for hydrological models has a strong influence on the accuracy of river flow predictions (Duncan et al. 1993; Fekete et al. 2004). However, for many regions such as Northern Canada, available surface weather stations are sparsely distributed, and the quality of historical measurements is often questionable. Therefore, finding adequate data for hydrological modelling is a real challenge in such areas.

In recent decades, significant effort has been dedicated to producing global datasets for climate monitoring and research using weather forecasting models and a complex assimilation of observations called "reanalyses". Reanalyses use a constant data assimilation scheme and a numerical forecasting model, which for their part use several observations every 6-12 hours, over a given period (Dee et al. 2011; Saha et al. 2010). Available observations include radiosondes, satellites, buoys, aircraft and ship reports. While the assimilation scheme is constant, the observational network changes constantly. Nonetheless, reanalyses provide a physically consistent estimate of the climate state at each time step. In addition to global coverage, reanalyses typically cover several decades, and provide a large array of climate variables (Mesinger et al. 2006; Rienecker et al. 2011). Despite the spatial and temporal consistency of reanalyses, the observational database, which changes constantly over the duration of each reanalysis can produce spurious trends and artificial variability. Reanalyses outputs are often biased, and especially so for surface fields, but have steadily improved in time. Part of the biases involved is due to the relatively coarse grid resolution, as well as to the parameterization of many physical processes such as convective storms.

Recent reanalysis outputs have been gradually made available at higher spatial and temporal scales, thus potentially reducing biases.

Recent studies have assessed the usefulness of reanalysis data for climate monitoring, and have found them to be extremely useful if used with care (Bosilovich 2013; Lorenz and Kunstmann 2012; Manzanas et al. 2014; Nigam and Ruiz-Barradas 2006; Rusticucci et al. 2014; Zhang et al. 2013). Moreover, reanalyses have demonstrated good potential to drive hydrological models (Choi et al. 2009; Essou et al. 2016a; Fuka et al. 2014; Vu et al. 2012). Recently, Essou et al. (2016b) used precipitation and temperature series from the NARR, ERA-Interim, CFSR and MERRA reanalyses and from one gridded observation database (Santa Clara dataset) to calibrate a lumped hydrological model and to simulate river flows over 370 watersheds in the continental USA. They found that the river flows obtained using NARR forcing were as good as when gridded observations were used. Moreover, the Nash-Sutcliffe values of the river flows simulated using the other three reanalyses were equal to or better than those from the gridded observations, with the exception of the Humid continental and subtropical regions, where the precipitation seasonality was not well reproduced. Reanalyses may thus be useful for hydrological modelling, especially in areas with a sparse weather station density. However, instead of using either reanalyses or surface weather stations alone, a more optimal scenario may consist in combining both data sources. Such a multi-model approach involves the combination of several hydrological models to simulate river flows more accurately than the models taken individually (Ajami et al. 2006; Arsenault et al. 2015c; Cavadias and Morin 1986; Diks and Vrugt 2010; Shamseldin et al. 1997). Recently, Arsenault et al. (2015a) combined three hydrological models and four climate datasets to produce multi-input averaged flows and found that this approach provides better results than the classical multi-model averaging. In their work, all the datasets used came from the same data source type (gridded databases). The combination of different data sources has the potential to improve the accuracy of simulated river flows. For instance, Sun et al. (2000) evaluated flood estimation combining radar and raingauge data for the Finniss River catchment in Darwin, Australia, and found that rainfall estimated by coKriging both data sources considerably improved flood estimates. They showed that an optimal combination of both databases improves the estimation of subcatchment rainfall. To our knowledge, the potential presented by combining reanalyses and weather stations for hydrological modelling has never been investigated.

This study will focus on the impacts of combining three global atmospheric reanalyses – ERA-Interim, CFSR and MERRA – and one gridded observation dataset on hydrological model simulations. More specifically, it aims to assess the impacts of such a combination on the ability to simulate river discharges: (1) in the presence of a high density of surface weather stations (US watersheds); (2) in the presence of a low density of surface weather stations (Canadian watersheds); and (3) over watersheds where hydrological models perform poorly, thus calling into question the quality of surface observations. The results of this study will determine whether the combination of reanalyses and observations in regions with a sparse density of weather stations, such as Northern Canada, is impactful for hydrological modelling.

6.3 Study area

The study area consists of 830 watersheds in North America, 370 of which are located in the United States, and 460 in Canada. The watersheds are located in various hydro-climatic regimes. The US watersheds were selected because of their relatively high density of weather stations compared to their Canadian counterparts. They were derived from the MOdel Parameter Estimation eXperiment database (MOPEX) (Duan et al. 2006), and their total areas ranged between 104 and 10,325 km^2. The Canadian watersheds were selected because of their relatively low density of weather stations. They were derived from the CANadian mOdel Parameter EXperiment (CANOPEX) database (Arsenault et al. 2015b). Their total areas ranged between 450 and 127,635 km^2. The Canadian watersheds tend to be larger since major rivers are the only ones typically gauged in remote areas.

Over the study area, the mean annual precipitation is between 0 and 5 mm/day (figure 6.1a). The highest precipitation (> 4mm/day) area is located in the Southeastern US, and the lowest precipitation (< 1mm/day) area is located in Northern Canada. The mean annual temperature varies between -5°C and 20°C (figure 6.1b). The temperature decreases from South to North in the study area. Consequently, watersheds in the Southeastern US are the warmest, while the colder ones are located in Northern Canada.

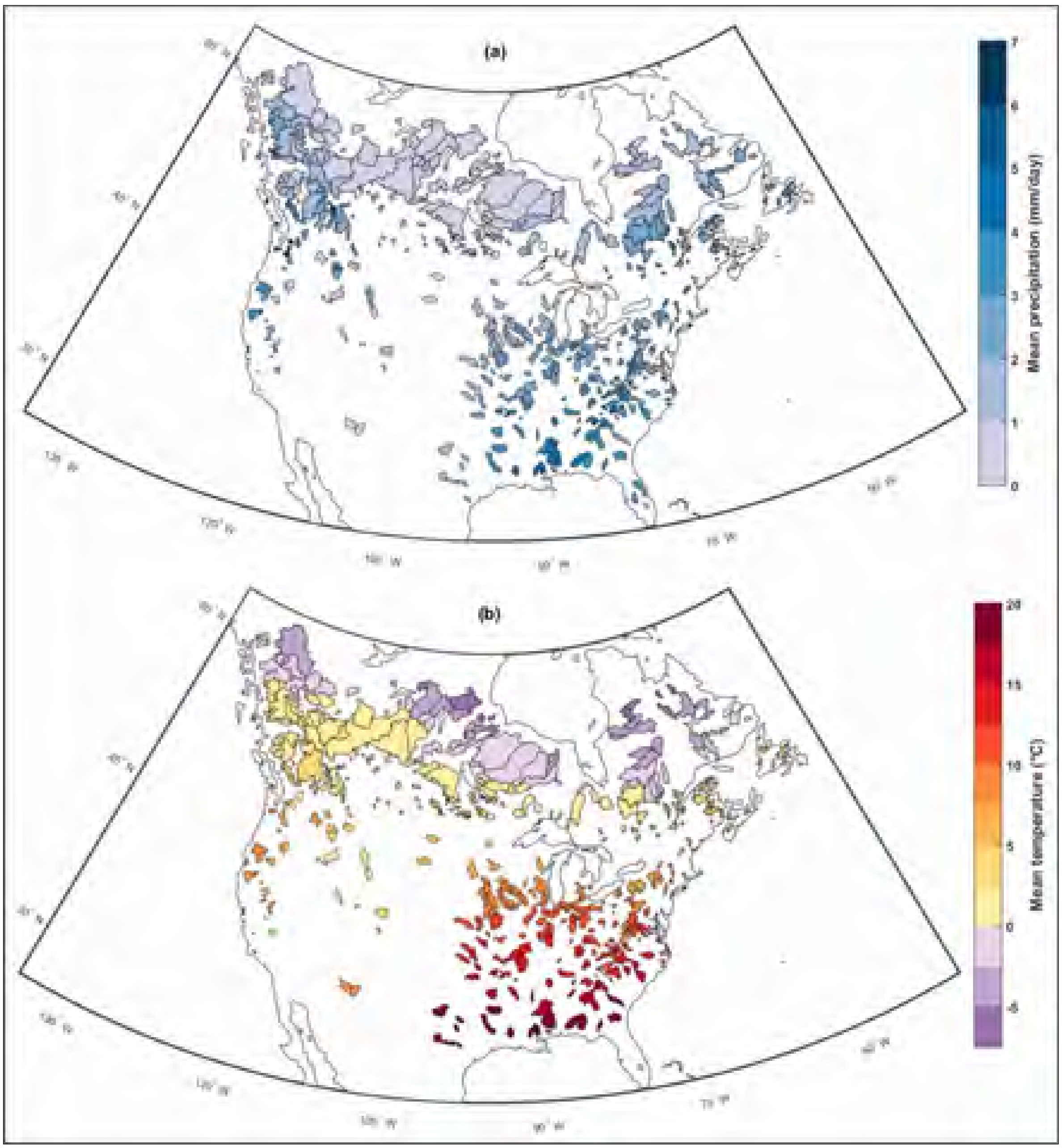

Figure 6.1 Mean annual (a) precipitation (mm/day) and (b) temperature (°C) of the 830 watersheds analyzed in this study

6.4 Datasets

6.4.1 Observational datasets

The observational datasets consisted of daily meteorological (minimum and maximum temperature and precipitation) and hydrometric datasets derived from the Santa Clara and MOPEX databases, for the US watersheds, and from the NRCan and HYDAT databases, for the Canadian watersheds.

The Santa Clara dataset (SClara) is a gridded database based on the high-density network of weather cooperative stations of the National Oceanic and Atmospheric Administration (NOAA) (average of 1 station per 700km^2) (Maurer et al. 2002). The SClara gridded database consists of daily precipitation and maximum and minimum air temperature at a 0.125° x 0.125° spatial resolution (about 12km x 12km) for the period of 1949-2010. The Synergraphic Mapping System (SYMAP) algorithm was used for the data interpolation (Shepard 1984; Widmann and Bretherton 2000). The SClara gridded precipitation was scaled to match the long-term average precipitation of the Parameter-elevation Regressions on Independent Slopes Model (PRISM) (Daly et al. 1994; Daly et al. 1997).

The MOPEX database contains daily mean hydrometric and meteorological data covering the period of 1949-2003 (Duan et al. 2006). This study only used the MOPEX hydrometric data.

NRCan is a gridded database based on the low-density network of Environment Canada weather stations. The NRCan dataset consists of daily precipitation and 2-m temperature at a 10-km spatial resolution over the period of 1950-2010. The Interpolation was performed using the thin plate-smoothing splines (ANUSPLIN) method (Hopkinson et al. 2011; Hutchinson 1995, 2004; Hutchinson et al. 2009).

The HYDAT database contains daily mean discharge data from about 7000 hydrometric stations across Canada (Coulibaly et al. 2013; Winkler 1993). Both of the previous datasets were brought together within the watershed-based coherent CANOPEX database (Arsenault et al. 2015b).

6.4.2 Reanalysis datasets

The daily mean precipitation and 2-m temperature from the ERA-Interim, CFSR and MERRA reanalyses were used to force the hydrological model described later in this article.

ERA-Interim is a 0.75° x 0.75° (about 80km x 80km) global reanalysis of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al. 2011). It covers the period of 1979-present, and uses a 4D-VAR data assimilation scheme. ERA-Interim runs in near real time, using data from the operational ECMWF forecast system. The observations assimilated before 2002 are derived mainly from the data used for the ERA-40 dataset (Uppala et al. 2005). The 2-m temperature from ERA-Interim results from assimilated observations, while precipitation is produced by the weather forecast model.

The Climate Forecast System Reanalysis (CFSR) is a 0.3° x 0.3° (about 35km x 35km) global reanalysis from the National Centers for Environmental Prediction (NCEP). It covers the period of 1979-present and uses a 3D-VAR data assimilation scheme (Saha et al. 2010). CFSR is the first reanalysis produced with a coupled atmosphere-ocean-land surface climate system with an interactive sea ice component. The land surface component does not use the precipitation generated by the atmospheric model, which is considered too biased (Saha et al. 2010). Instead, CFSR uses precipitation from the NOAA Climate Prediction Center (CPC) pentad merged analysis of precipitation (Xie and Arkin 1997) and the CPC unified daily gauge analysis (Wang et al. 2011). CFSR assimilates satellite radiance data rather than measured temperature and humidity values (Wang et al. 2011).

The Modern Era Reanalysis for Research and Applications (MERRA) is a 2/3° (lon) x 1/2° (lat) (about 75km x 55km) global reanalysis. It is developed by the Global Modeling and Assimilation Office (GMAO) of the National Aeronautics and Space Administration (NASA) to maximize the use of GMAO satellite observations in a climate context and to improve the hydrological cycle represented in the first generation of reanalyses (Rienecker et al. 2011). MERRA covers the satellites era (1979-present), and is generated from version 5.2.0 of the Goddard Earth Observing System (GEOS) atmospheric model and a data assimilation system based on a 3D-VAR approach (Suarez et al. 2008). The main specificity of MERRA consists in its use of an incremental analysis update (IAU) procedure that improves the closure of the water cycle.

A description of the databases is presented in table 6.1.

Table 6.1 Description of the databases used in this study

Database	**Acronym**	**Horizontal resolution**	**Assimilation scheme**	**Variables**	**Source**	**Reference**
Santa Clara	SClara	0.125° x 0.125°		P, Tmin, Tmax	http://hydro.engr.scu.edu/files/gridded_obs/daily/ncfiles_2010	(Maurer et al. 2002)
Model Parameter Estimation eXperiment	MOPEX	---	---	Q	ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/US_Data	(Duan et al. 2006)

Database	Acronym	Horizontal resolution	Assimilation scheme	Variables	Source	Reference
NRCan	NRCan	10km x 10km		P, Tmin, Tmax	---	Hutchinson et al, 2009
HYdrometric DATabase	HYDAT	---	---	Q	ftp://arccf10.tor.ec.gc.ca/wsc/software/HYDAT/	(Coulibaly et al. 2013; Winkler 1993)
ERA-Interim reanalysis	ERAI	0.75° x 0.75°	4DVAR	P, Tmin, Tmax	http://data-portal.ecmwf.int/	(Dee et al. 2011)
Climate Forecast System Reanalysis	CFSR	0.3° (lon) x 0.3° (lat)	3DVAR	P, Tmin, Tmax	http://cfs.ncep.noaa.gov/cfsr/	(Saha et al. 2010)
Modern Era Reanalysis for Research and Applications	MERRA	2/3° (lon) x 1/2° (lat)	3DVAR	P, Tmin, Tmax	http://disc.sci.gsfc.nasa.gov/mdisc/overview/index.shtml.	(Rienecker et al. 2011)

P = Precipitation; Tmin = minimum temperature; Tmax = maximum temperature; Q = streamflow.

This study covers the period of 1979-2003, for the US watersheds, and various periods between 1979 and 2010, for the Canadian watersheds in order to account for the availability of observed discharge datasets – the shortest period is 11 years long (1979-1989), but for most of the watersheds, streamflows cover the entire period (1979-2010).

6.5 Methods

6.5.1 Hydrological model and calibration strategy

A lumped conceptual hydrological model, HSAMI (Fortin 2000), was used because of the large number of watersheds included in this study. The setup and calibration of a distributed model over such a large database would be a daunting task. The HSAMI model has been used by Hydro-Québec, Quebec's power utility company, for over two decades to forecast daily flows at about 100 watersheds over the province of Quebec. The model has been used in various flow simulations and climate change impact studies (Chen et al. 2011b; Mareuil et al. 2007; Minville et al. 2008; Riboust and Brissette 2015). It simulates the main hydrological cycle processes, such as vertical and horizontal water transfers, evapotranspiration, snowmelt and soil freezing. The model has 23 calibration parameters: 2 for evapotranspiration, 6 for snowmelt, 10 for infiltration and percolation, and 5 for the routing of surface runoff and interflow. There are four interconnected reservoirs contributing to the vertical water transfer balance: snow on ground, surface water, unsaturated zones, and saturated zones. The horizontal water transfer is based on two unit hydrographs (one for surface runoff and one for delayed runoff) and one linear reservoir for groundwater flows.

Calibrations were performed on the even years for each watershed, and the odd years were retained for validation. This approach typically results in similar values of the objective function over both calibration and validation periods, thus avoiding the trap of selecting a non-representative calibration period. All the calibrations were performed using the Covariance Matrix Adaptation Evolution Strategy (CMAES) algorithm (Hansen and Ostermeier 1996, 2001). Arsenault et al. (2014) showed that CMAES was the most appropriate optimisation algorithm for the HSAMI model.

The Nash-Sutcliffe Efficiency (NSE) metric (Nash and Sutcliffe 1970) was computed as an objective function based on the even years, with cross-validation on the odd years. The NSE is recognized to be one of the best objective functions for reflecting the overall fit of a hydrograph (Servat and Dezetter 1991) and the most commonly used. For each watershed, 20 calibrations were performed (to study equifinality within another project). For this work, the best parameter set was selected, although all parameter sets displayed a similar performance, a known consequence of equifinality.

The non-parametric Wilcoxon test (Rakotomalala 2008) was performed to statistically evaluate simulated discharge in validation at the 95% confidence level.

6.5.2 Combination strategy

Two combination approaches were used. The first one combines reanalyses and the observational database – i.e., it uses a weighted average of precipitation and temperature inputs. The second approach consists in using all meteorological inputs separately and combining the simulated hydrographs. Each approach is described below.

6.5.2.1 Combination of precipitation and temperature

In this approach, a weighted average using precipitation and temperature data from each reanalysis and the observational database is computed as follows:

$$X_r = \sum_{i=1}^{n} \beta_i * X_i \tag{6.1}$$

where n is the number of databases to combine (for each watershed, n = 4 to account for 3 reanalyses and 1 observed database), the weight of the database *i* is β_i and X_i is the variable whose average, X_r, is needed (either temperature or precipitation).

The strategy used to determine the weights β_i is illustrated in figure 6.2. The determination is done during the calibration process, with the CMAES optimization algorithm, which is used to identify the parameters of the hydrological model. Thus, the weights are consistent with the values of the hydrological model parameters. Consequently, the optimal weights for each meteorological input vary from one watershed to another. As for the hydrological model parameters, the weights determined over the calibration period are the same ones used over the validation period.

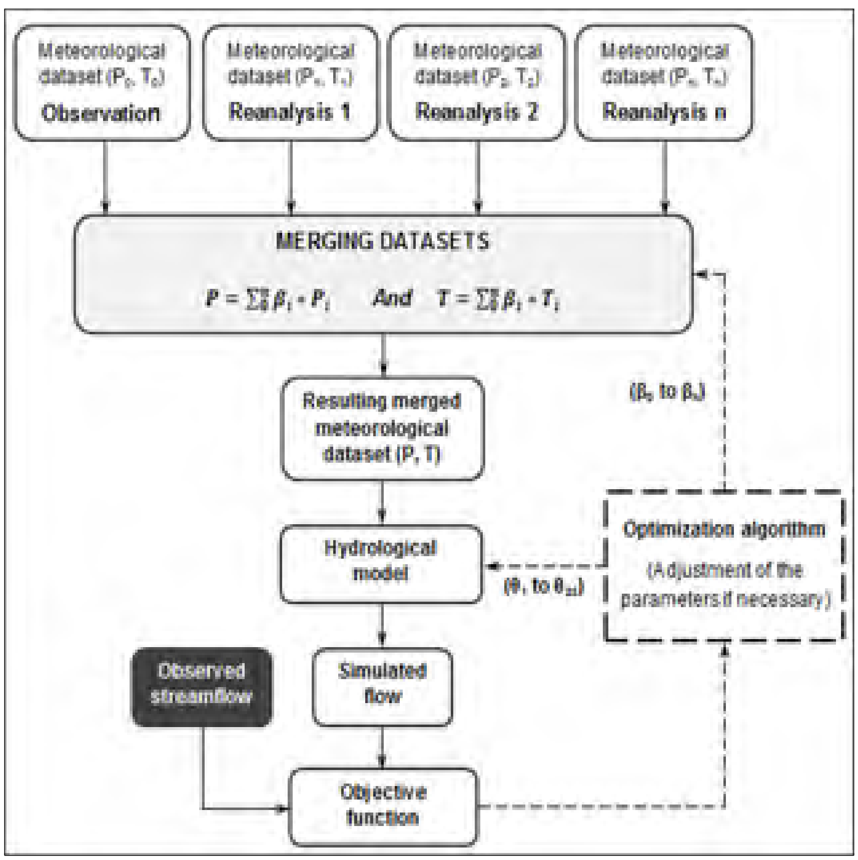

Figure 6.2 Strategy to determine the weights and to combine precipitation and temperature values during the calibration of the hydrological model

6.5.2.2 Combination of simulated hydrographs

The hydrological model is first calibrated using each of the four input datasets. Four separate hydrographs are produced. Subsequently, the weight of each simulated hydrograph is determined over the calibration period to minimize the objective function – the Nash-Sutcliffe Efficiency – computed using the weighted average of all four hydrographs.

The simulated hydrograph weights are determined using the Granger Ramanathan (GR) algorithm (Granger and Ramanathan 1984). Recent studies have shown that this method is robust for hydrological modelling applications (Diks and Vrugt 2010). There are three algorithms in the GR method – GRA, GRB and GRC – but the latter is the most robust, and was therefore used in this study. The approach minimizes the RMSE and corrects the bias of the weighted averaged discharge values with a constant term. It determines unconstrained weights based on the ordinary least squares (OLS) algorithm. The resulting weighted average (Q_{GRC}) flow is calculated as follows:

$$W_{GRC} = (Q_{sim}{}^{T} Q_{sim})^{-1} Q_{sim}{}^{T} Q_{obs} - W_0 (Q_{sim}{}^{T} Q_{sim})^{-1} Q_{sim}{}^{T} l$$

$$W_0 = \frac{l^T (Q_{obs} - Q_{sim} \times ((Q_{sim}{}^{T} Q_{sim})^{-1} Q_{sim}{}^{T} Q_{obs}))}{n - l^T Q_{sim} (Q_{sim}{}^{T} Q_{sim})^{-1} Q_{sim}{}^{T} l} \tag{6.2}$$

$$Q_{GRC} = Q_{sim} W_{GRC} + W_0$$

where W_{GRC} are the GRC weights, W_0 is the constant term used to correct bias, Q_{sim} is an n*m matrix of n daily simulated streamflow values from m databases, Q_{obs} is an n*1 vector of the observed streamflow, *l* is a unit vector of length n, the superscript T indicates that the matrix is transposed, and Q_{GRC} is the combined streamflow. In this approach, weights are not constrained to be either positive or to sum to unity (see Arsenault et al. (2015a) for details).

6.6 Results

6.6.1 Comparison of the discharge simulation from the two combination approaches

The validation performance of the hydrological model is shown in figure 6.3 for each individual dataset, as well as for the two combination methods. Results for the 460 Canadian watersheds show that the NSE median values are 0.78 for NRCan, 0.78 for ERA-Interim, 0.77 for CFSR and 0.74 for MERRA (figure 6.3a). In addition, the NRCan NSE values are higher than those for ERA-Interim, CFSR and MERRA for 52%, 65% and 66% of watersheds, respectively. Thus, for most of the watersheds, the observation NRCan database is essentially equivalent to ERA-Interim, but performs better than CFSR and MERRA. According to the Wilcoxon statistical test performed at the 95% level of significance, the differences between the individual datasets are not significant, with the exception of MERRA, which consistently performs slightly worse. Results in figure 6.3a also show that the median NSE values are 0.82, when combined inputs are used, and 0.85, when combined hydrographs (or combined outputs) are used. Moreover, the NSE values obtained from using combined inputs and combined outputs are higher than NRCan-only NSE values for 68% and 92% of the watersheds, respectively (figures 6.4a and 6.4b). The statistical test shows that the performances obtained from both combination approaches are significantly different from what obtains with each individual dataset. Thus, the use of combined precipitation and temperature data from the reanalyses and NRCan or combined hydrographs significantly improves discharge estimates for the Canadian watersheds.

The NSE performances obtained from both combination approaches are significantly different from one another. Furthermore, the NSE values obtained from using combined hydrographs are higher than those obtained using combined inputs for 73% of the watersheds. Thus, the hydrograph combination approach is significantly better than the input combination.

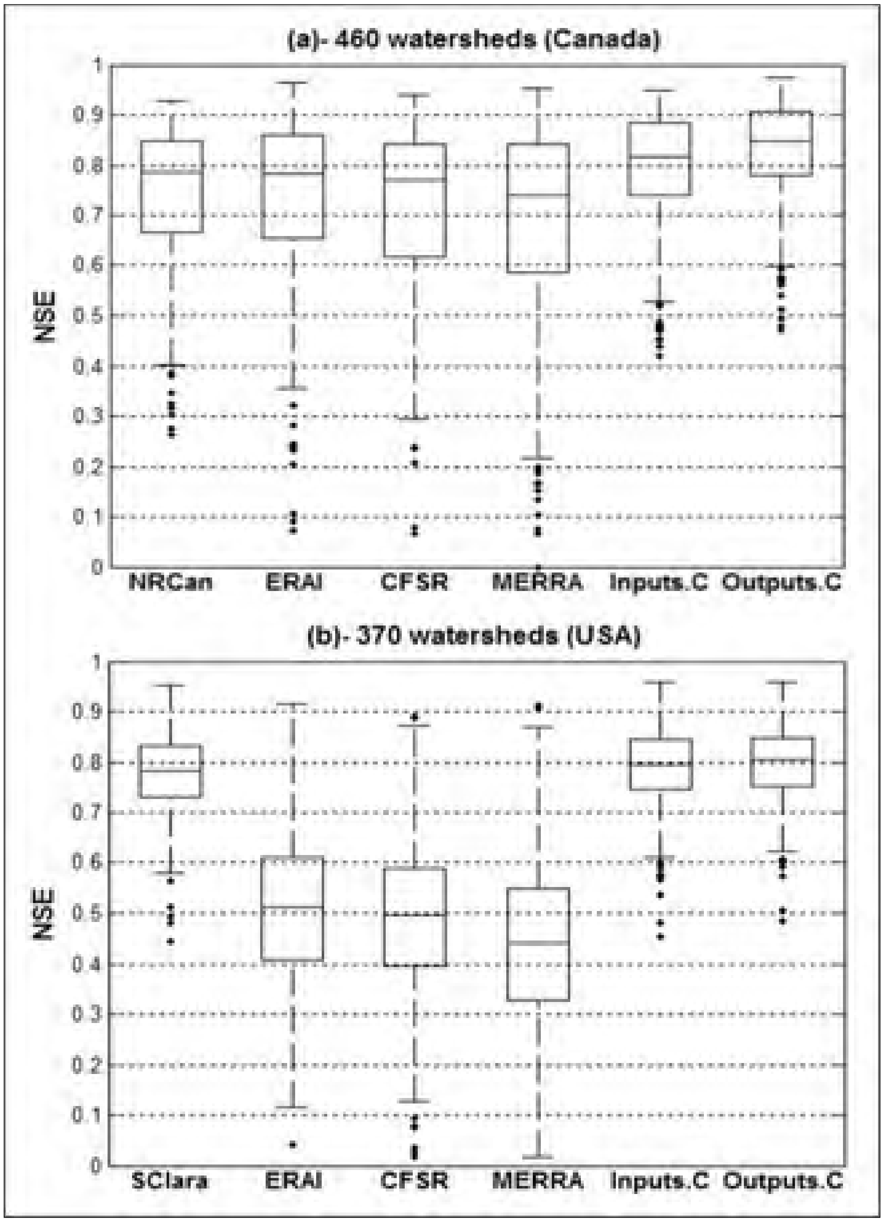

Figure 6.3 Validation NSE performances of the discharges simulated using Santa Clara, ERA-Interim, CFSR, MERRA, combined inputs (Inputs.C) and combined outputs (Outputs.C) for (a) 460 Canadian watersheds and (b) 370 USA watersheds

The results are different over the US watersheds (figure 6.3b). The NSE performances for the 370 US watersheds show median NSE values of 0.78 for the observation-based Santa Clara dataset, 0.51 for ERA-Interim, 0.5 for CFSR, and 0.44 for MERRA. The Santa Clara dataset performs significantly better than each of the individual reanalysis datasets. Individually, the

Santa-Clara dataset performed better than the ERA-Interim, CFSR and MERRA datasets for 94%, 95% and 92% of the watersheds, respectively.

The results in figure 6.3b also show that both combination approaches result in a modest improvement in the NSE median value (0.8 in both cases). In addition, in both cases, an improvement is observed for 72% of the watersheds (figures 6.4c and 6.4d). The small increase in performance obtained from both combination methods was nonetheless found to be statistically significant. Thus, despite the relatively poor performance of the three global reanalysis datasets over the US, they contribute to improve streamflow simulations for most of the US watersheds.

The NSE values obtained using combined hydrographs were compared to those obtained from combined inputs, and were found to be higher for 54% of the US watersheds. This difference was not statistically significant.

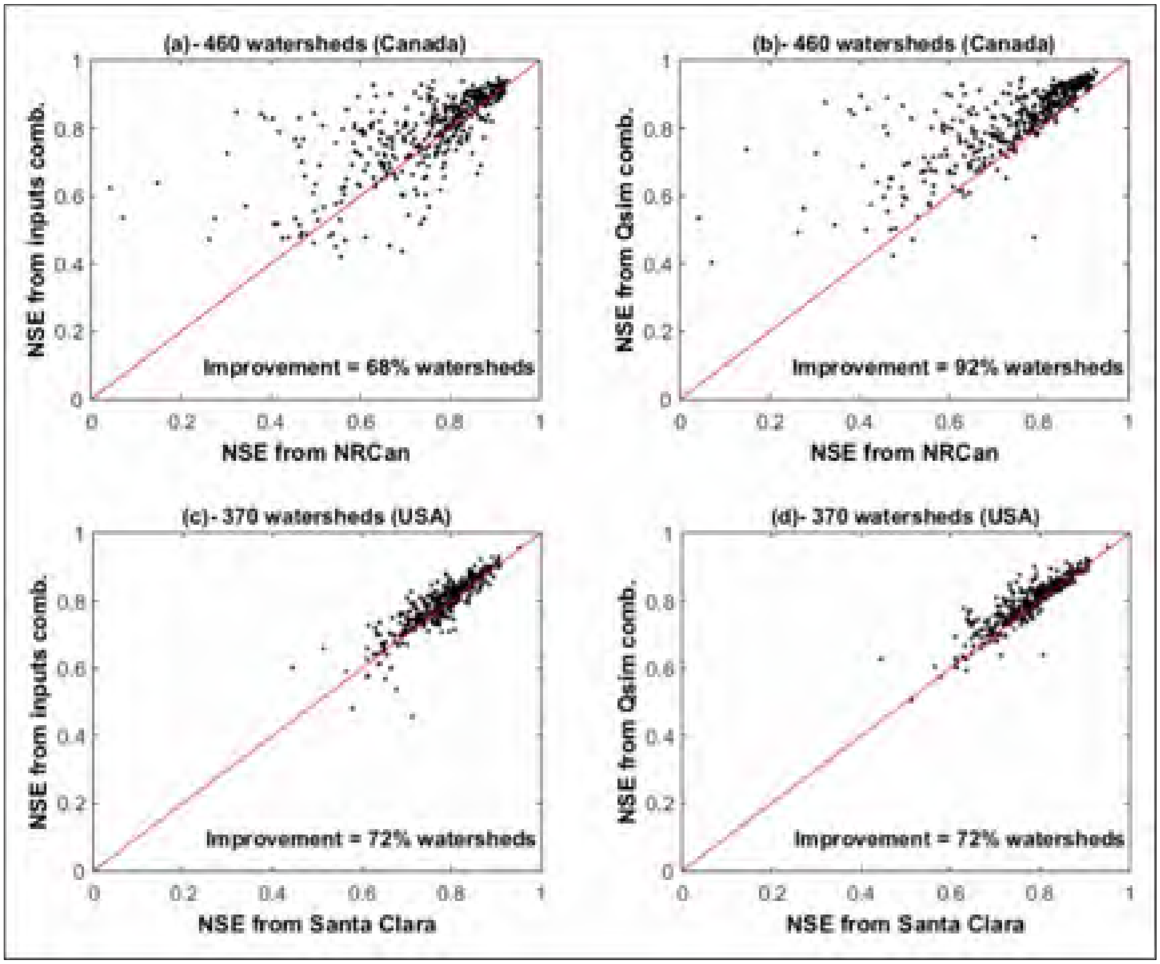

Figure 6.4 Improvement in the NSE values using combined inputs (a and c) and combined outputs (b and d). Results are based on validation performances

6.6.2 Analysis of performances that are not improved

It can be inferred from the results presented in figures 6.4a and 6.4c that for 32% of the Canadian watersheds and 28% of the US watersheds, the use of combined inputs from reanalyses and NRCan led to a drop in the NSE values, as compared to those obtained from NRCan alone. When combining hydrographs, a drop in performance over 8% of Canadian watersheds and 28% of the US watersheds was observed when compared to the best individual dataset.

To determine whether the drops in performance were significant, a comparison was made between the NSE values obtained using observations (NRCan or Santa Clara) and those obtained using combinations.

Results show that for the Canadian watersheds, when combined inputs were used, a 0.04 drop versus the NRCan median NSE value (0.81 to 0.77) occurred (for 32% for the watersheds) (figure 6.5a). On the other hand, the use of combined outputs led to a 0.01 drop versus the median NRCan NSE value (0.75 to 0.74) (for only 8% for the watersheds) (figure 6.5b). In the case of combined inputs, the drops were significant, but not so for the combined outputs.

For the US watersheds, the results indicate that when inputs are combined, a 0.02 drop occurs, as compared to the NSE median value for Santa Clara (0.79 to 0.77) (Fig. 6.5c). Moreover, the results in Fig 6.5d show that when outputs are combined, a 0.01 drop occurs versus the median NSE value for Santa Clara (0.80 to 0.79). In the case of combined inputs, the drops are statistically significant, but not so for the combined outputs.

Thus, although drops in performance may be caused by the use of a direct combination (inputs) or an indirect combination (outputs) of reanalyses and observations, the drops are generally small when outputs are combined, and larger when inputs are combined.

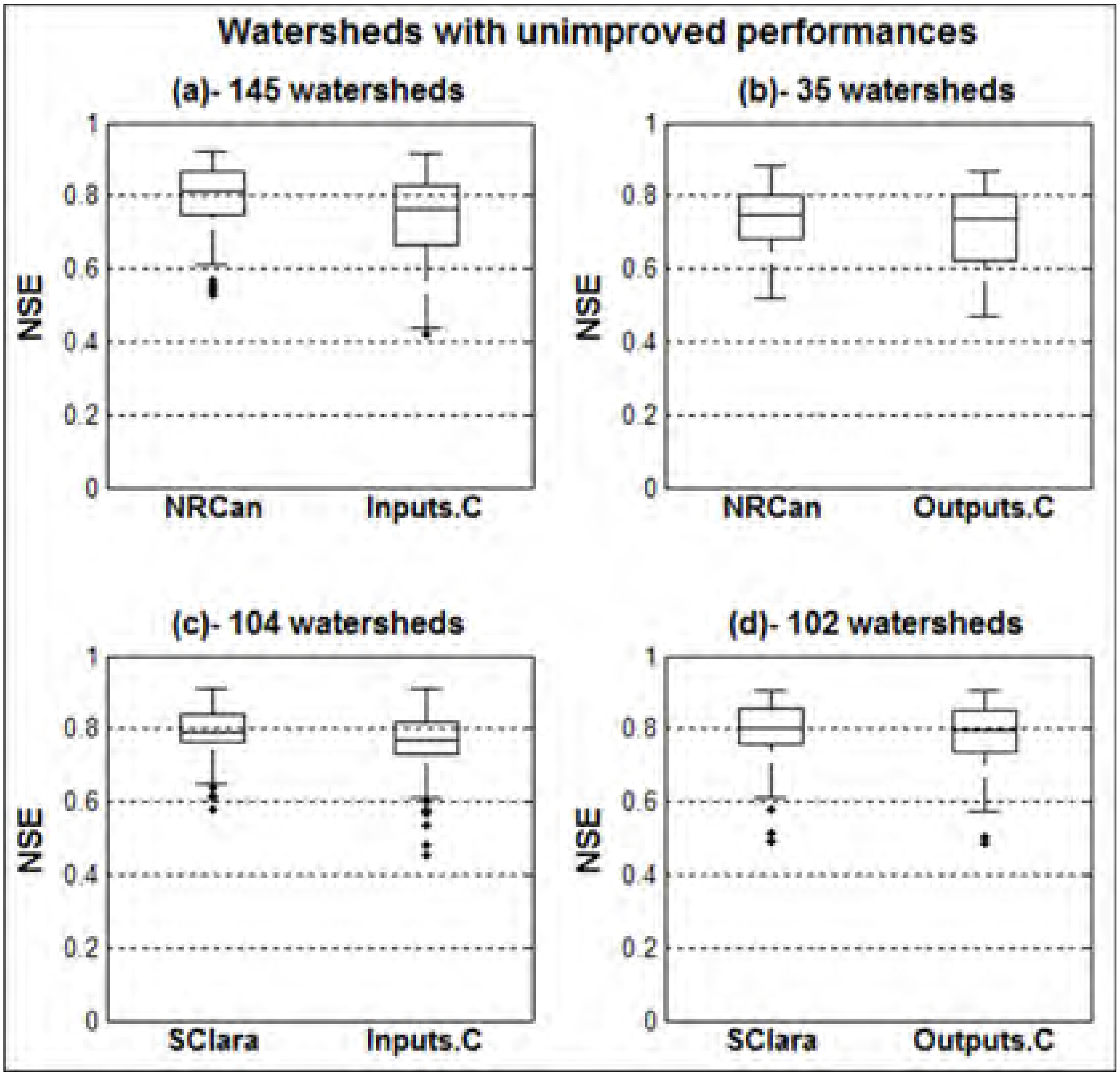

Figure 6.5 Distribution of the NSE values when there are no improvements due to the input combinations (a, c) and to the output combinations (b, d). Results are based on validation performances

6.6.3 Comparison of the performances for watersheds with a low performance using the observational databases

Typically, a NSE value below 0.5 is considered low or unsatisfactory (Moriasi et al. 2007; STREAMFLOW 2009). Such NSE values were obtained for 26 Canadian watersheds when NRCan alone was used. Such results were also found for 3 US watersheds when the Santa Clara database was used. For those watersheds, a comparison of the NSE values was performed to determine the impact of taking reanalyses into account in the simulation performance of discharges.

The results obtained for the 26 Canadian watersheds are similar to those obtained for the 3 previously mentioned US watersheds. For that reason, the results are presented for the 29 Canadian/US watersheds together (figure 6.6). The results indicate that the NSE median values are 0.42 for NRCan/SClara, 0.52 for ERA-Interim, 0.53 for CFSR, and 0.56 for MERRA. It can be inferred from these median NSE values that globally, the NSE using reanalyses are barely satisfactory for the 29 watersheds. However, a NSE comparison shows that NSE values using ERA-Interim, CFSR and MERRA are higher than the NRCan/SClara values, for 81% of the 29 watersheds. The statistical test indicates that for those watersheds, reanalyses provide significantly better NSE values than when NRCan/SClara is used.

The results in figure 6.6 also show that the median NSE values are 0.74 when inputs are combined, and 0.71 when outputs are combined. Thus, the NSE values obtained using those combinations are much higher than those obtained by using each database specifically (NRCan or reanalysis alone). The results also show that there is an improvement in the NSE values for all 29 watersheds. This illustrates the potential of using reanalyses to improve the simulated discharge for watersheds where observational databases result in low performances.

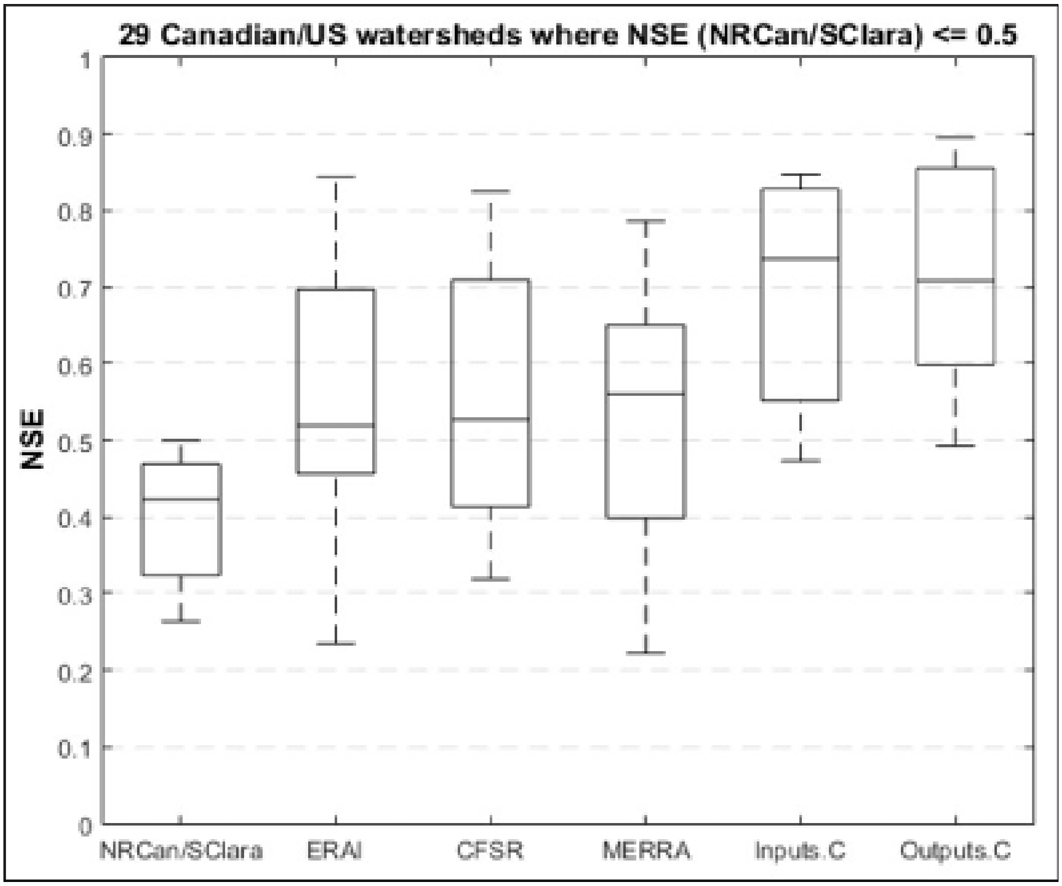

Figure 6.6 Impact of the combination of reanalyses and NRCan on the discharge simulation performance for the 26 Canadian watersheds with an NSE below 0.5 when driven by NRCan. Results are based on the validation NSE

6.6.4 Analysis of the weight distribution

This analysis aims to determine whether weights assigned to each of the four databases are significantly different. The distribution of weights (figures 6.7a-6.7b) indicates that both observational databases are more heavily weighted than reanalysis, particularly in the US. Among the three reanalyses, ERA-Interim tends to have the highest weights, while MERRA has the lowest.

For the 460 Canadian watersheds, the median weight values for input combinations are 0.37 for NRCan, 0.21 for ERA-Interim, 0.18 for CFSR, and 0.08 for MERRA (figure 6.7a). In the case of output combinations, the median weight values are 0.42 for NRCan, 0.26 for ERA-Interim, 0.17 for CFSR, and 0.14 for MERRA (figure 6.7b). Statistically, the NRCan weights are significantly higher than those obtained with the reanalyses, and this is true for both the input and output combinations.

The results from the 370 US watersheds give, in the case of input combinations, a median weight of 0.84 for Santa Clara, against 0.06 for ERA-Interim and CFSR, and 0 for MERRA (figure 6.7c). When outputs are combined, the median weights are 0.8 for Santa Clara, 0.11 for ERA-Interim, 0.1 for CFSR and 0.02 for MERRA (figure 6.7d). In both cases, the weights for Santa Clara are significantly higher than those obtained with the reanalyses. This means that for hydrological modelling, the data of the three reanalyses are less realistic, as compared to those for Santa Clara. Nevertheless, the meteorological data from reanalyses lead to a statistically significant improvement in the simulated discharge for most of the 370 watersheds.

The spatial distributions of the weights of the different databases are presented in Figure 6.8 and Figure 6.9.

For the Canadian watersheds, in the case of input combinations, no trend is observable for the NRCan, CFSR and MERRA weights. The ERA-Interim weights tend to be considerably higher than those for NRCan in Northern Canada. Since precipitation is the most critical meteorological variable in hydrological modelling, these results therefore mean that in Northern Canada, precipitation from ERA-Interim is of higher quality than those from the observation-based NRCan database.

In the case of output combinations, no trend is observable as well for the weights of CFSR and MERRA. The ERA-Interim weights tend to be dominant in Eastern Canada, while those for NRCan are mostly dominant in Western Canada (except in the Mountain region). This trend with the NRCan weights probably results from the proper filtering of biases in precipitation from NRCan due to the specific calibration of the hydrological model.

In both the input and output combination cases, the Santa Clara weights are significantly higher than those from reanalyses in the Eastern US, while in the Western US, reanalyses and Santa Clara have similar weights. These results mean that precipitation from ERA-Interim, CFSR and MERRA are of similar quality to those from observations in the Western US, and of significantly lower quality in the Eastern US. These results are consistent with those of Essou et al. (2016b) (Chapter 4). Indeed, these authors showed that unlike in the Western US, precipitation from the three reanalyses do not adequately reproduce the seasonality of precipitation observations in the Eastern US.

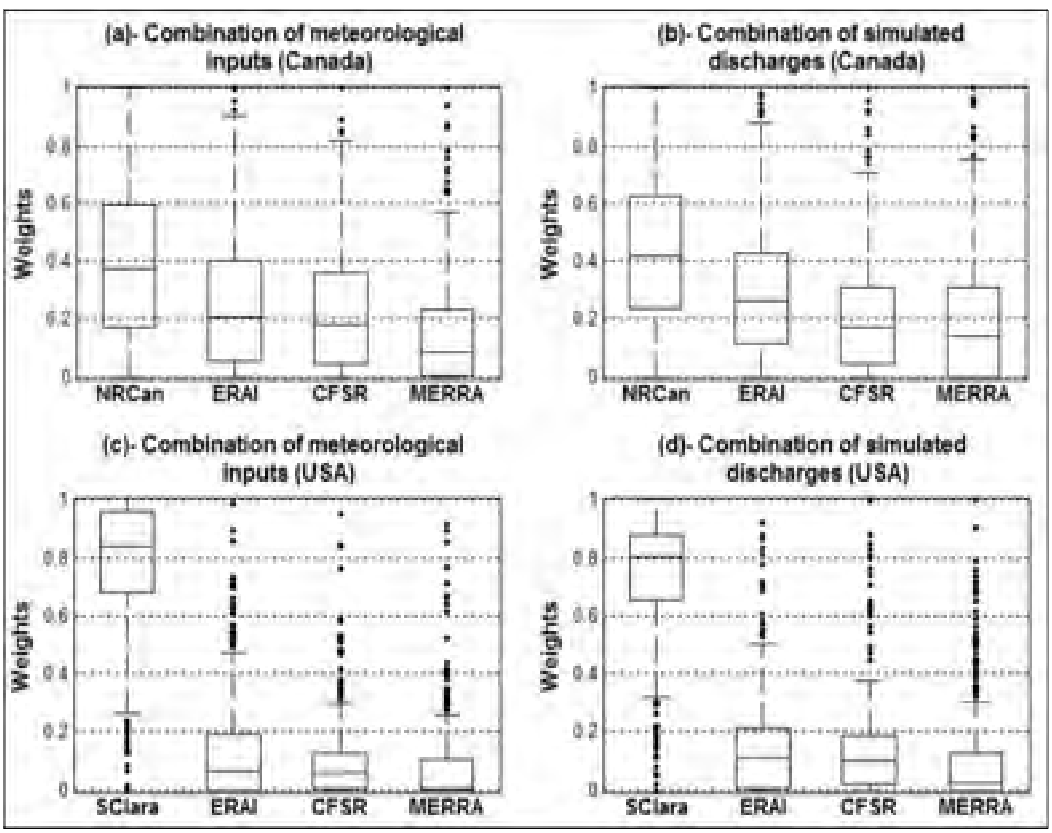

Figure 6.7 Distribution of the weights of the observational databases and of reanalyses for input combinations (a, c) and for output combinations (b, d) over Canada and the USA

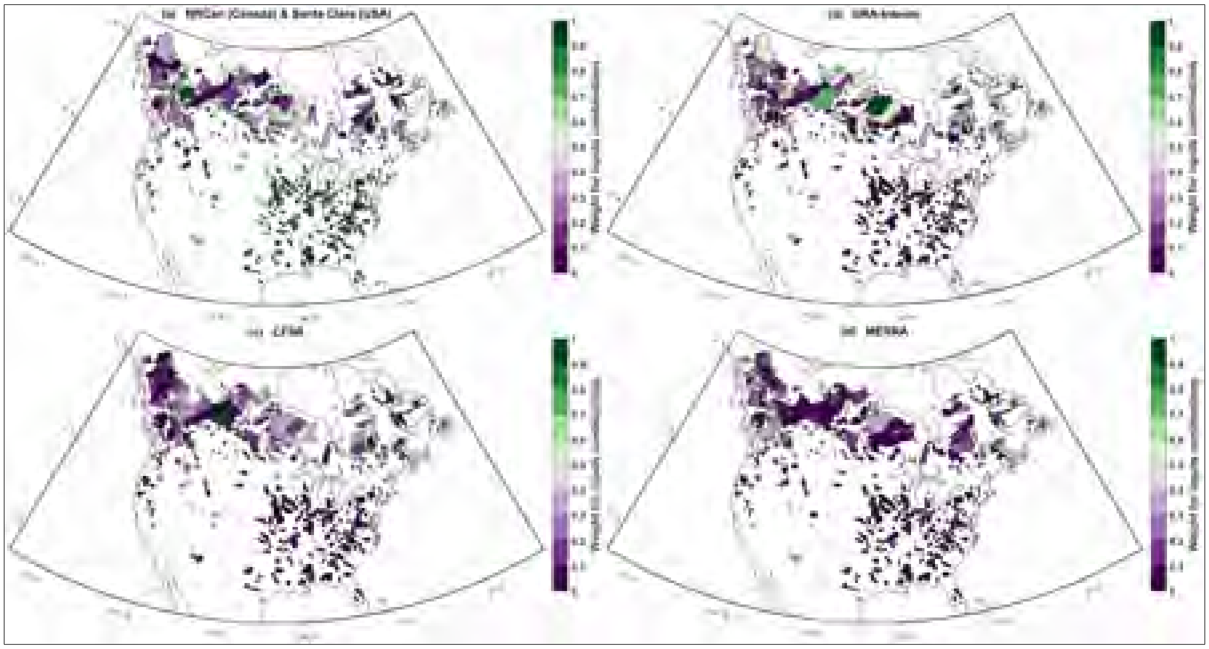

Figure 6.8 Spatial distribution of the weights of the observational databases and of reanalyses for input combinations over Canada and the USA

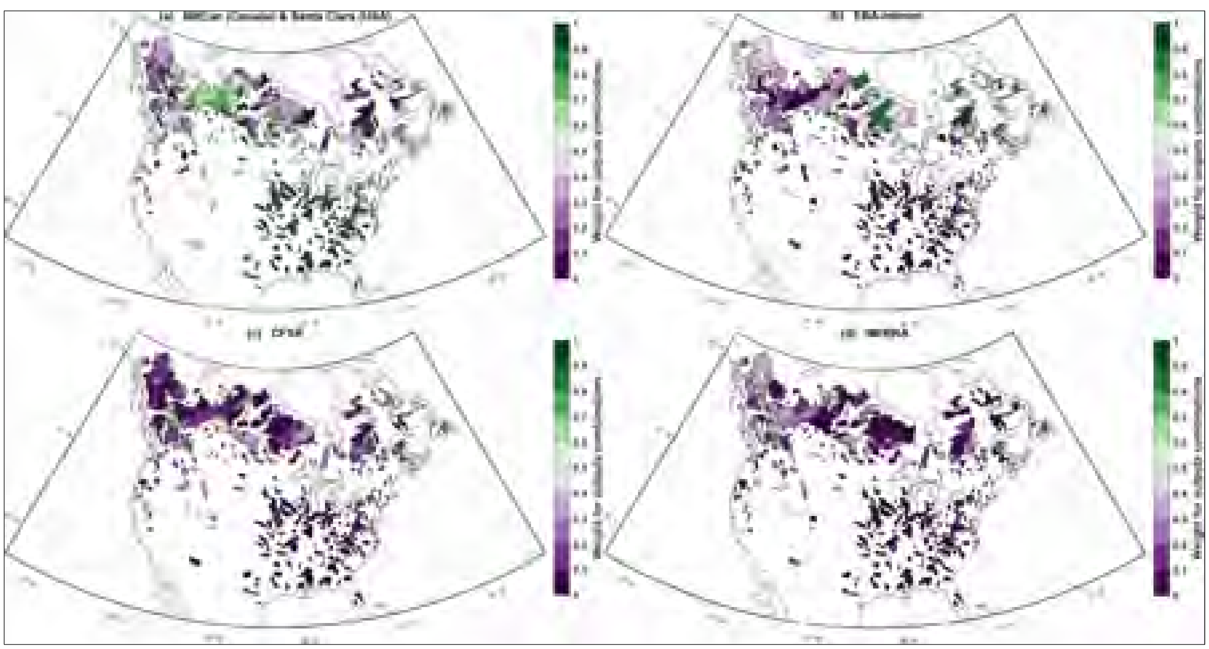

Figure 6.9 Spatial distribution of the weights of the observational databases and of reanalyses for output combinations over Canada and the USA

6.7 Discussion

Global atmospheric reanalyses are important sources of climate data that have the advantage of providing a global coverage and spatially and temporally consistent data in near real time. They have proven to be quite useful for climate monitoring when used with appropriate care. Moreover, although precipitation and temperature data from reanalyses are often biased, they have shown high potential for hydrological modelling under various hydro-climatic regimes (Choi et al. 2009; Essou et al. 2016b; Fuka et al. 2014; Sabarly et al. 2016; Vu et al. 2012). This means that reanalyses can be useful inputs for hydrological modelling. However, it seems obvious that in most cases, reanalyses cannot replace observational data; these data are commonly used in hydrological modelling, but can contain significant uncertainties, especially in regions with a sparse weather station network (Chvíla et al. 2005; Mizukami and Smith 2012; Tozer et al. 2012). As a result, it could therefore be quite useful to combine reanalyses and observational data in regions with a sparse distribution of weather stations. This fact needs first to be validated in regions with a high density of weather stations.

Three state-of-the-art atmospheric reanalyses – ERA-Interim, CFSR and MERRA – as well as a gridded observational database were combined as inputs for hydrological modelling of 460 Canadian and 370 US watersheds. An analysis of the impact of combining reanalyses and observational databases was also carried out, specifically considering the watersheds for which observational databases led to low performances.

Two combination approaches were used. One consisted in computing a weighted average of the meteorological inputs (precipitation and temperature) from four databases. The second consisted in a weighted averaging of four simulated hydrographs resulting from each separate meteorological database. The CMAES algorithm (Hansen and Ostermeier 1996, 2001) was used to determine the weights in the inputs combination approach, whereas the C-variant of the Granger and Ramanathan (1984) method was used to calculate the weights for the output combination approach.

The main advantage of the input combination approach lies in its computational ease since the hydrological model is calibrated once, and only needs to be run once following the calculation of the weighted inputs. The output combination approach requires more work as the hydrological model must be calibrated and subsequently run for each of the datasets prior to the outputs being averaged. However, this approach has the advantage of being more robust, as shown in the results, possibly because the hydrological model better exploits the unique coherence between precipitation and temperature in each of the datasets. When averaging precipitation and temperature separately, this coherency may be lost, leading to a loss of performance for the input combination. Another reason could be the fact that the input combination approach is based on the assumption that weights determined during the calibration period will be valid for the validation period. This may not be true in some cases because biases in the different databases can vary from one period to another. This could be the case for reanalyses since their observational constraints, and therefore their reliability, can vary depending on the period and location considered. Furthermore, the changing mix of observations and biases in the assimilated observations could introduce spurious trends into reanalysis outputs, which could also vary from one period to another (Bosilovich 2013; Marshall 2003).

Generally, the results showed that a combination of reanalyses and observed data led to significantly improved performances for most of the Canadian and US watersheds. Overall, the performance of reanalysis for the simulation of flow over Canadian watersheds was superior to that for their US counterparts. This is likely due to the fact that Canadian watersheds are snow-dominated, and hydrological models can simulate river discharges more easily than when snow is not a dominant factor in the river discharge, as in the US. Most of the US watersheds considered in this work are dominated by rainfall – 84% of the US watersheds are located in Humid continental and Humid subtropical regions, in the Eastern US, where the performance of reanalyses is lower, as shown in Essou et al. (2016b). The higher performance of reanalysis for Canadian watersheds could also be due to the watershed sizes, which are generally larger than over the US.

Despite the relatively low performance of reanalysis over US watersheds, the combination of reanalyses and the Santa Clara database (input and output combinations) proved to be quite useful in significantly improving the simulation performance over the validation period for 72% of the watersheds. The improvements in simulation performance were also significant for the Canadian watersheds. These results show that despite biases, precipitation and temperature data from reanalyses are complementary to surface observations. It is also interesting to note that for all the watersheds where observational data presented a low performance (NSE <= 0.5), the combination with reanalyses led to significantly improved performances – an increase of about 0.3 versus the median NSE value of the observational databases alone.

In this work, three reanalyses were systematically used for the combinations. However, as shown in figure 6.7, they do not all necessarily improve the simulation performances – weights are sometimes equal to zero. Further investigation is therefore needed in order to determine the conditions under which the contribution of each reanalysis is critical. However, from the analysis of the weight distributions, it appears that contributions by ERA-Interim and CFSR are usually more useful, as compared to those by MERRA. Thus, despite its coarser resolution, ERA-Interim would appear to be globally more accurate than the two other reanalyses for river flow modelling. That is likely due to the fact that ERA-Interim has the most sophisticated assimilation scheme.

Although significant improvements in performance were globally observed from both combination approaches, it is important to note that some loss of performance was also observed in some cases, for both combination approaches. This performance loss was typically small, and especially present in the output combination approach.

6.8 Conclusion

This study assessed the impacts of the combination of three global atmospheric reanalyses and gridded observations on the accuracy of hydrological model discharge simulations. These impacts were assessed for three cases: (1) in the presence of a high density of surface weather stations (US watersheds); (2) in the presence of a low density of surface weather stations (Canadian watersheds); and (3) over watersheds where hydrological models perform poorly, thus calling into question the quality of surface observations. Two combination approaches were used: a combination of reanalyses and the observational database using a weighted average of precipitation and temperature inputs, and another consisting in using all meteorological inputs separately and combining the simulated hydrographs.

For both approaches, the results showed significant improvements in the discharge simulation performance, for most of the Canadian and US watersheds. However, overall, the improvements in performance were greater for the Canadian watersheds than for their US counterparts. Moreover, for all the watersheds where observational data presented a low performance (Nash Sutcliffe Efficiency of simulated discharge <= 0.5), the combination with reanalyses led to significantly improved performances. This indicates that reanalysis provides an added value for hydrological modelling, and can be useful in overcoming shortcomings in the surface observation record.

6.9 Acknowledgments

This work was funded through a Natural Science and Engineering Research Council collaborative research grant (NSERC-CRD) with Hydro-Québec, Rio-Tinto-Alcan, Ontario Power Generation and the Ouranos consortium on regional climatology and adaptation to climate change as industrial partners. These industrial research partners are greatly acknowledged for their direct and indirect contributions to this work. We also sincerely thank all the individuals and institutions that developed the datasets used in this work, and

made them available to the scientific community. We hope that this work represent a small contribution to their important effort.

CHAPITRE 7

DISCUSSION GÉNÉRALE

7.1 Synthèse des principaux résultats

Cette thèse se plaçait dans le contexte global de la recherche de sources de données météorologiques pouvant servir d'alternative aux traditionnelles stations météorologiques, afin de pallier le déficit d'information dans les régions où les stations sont en nombre insuffisant ou inexistantes.

Dans ce cadre, l'objectif de la thèse était d'évaluer le potentiel des données de précipitation et de température provenant des réanalyses atmosphériques pour les études de modélisation hydrologique. Plus précisément, cette thèse visait d'une part à comparer des données de précipitation et température provenant des réanalyses à celles d'observation, et d'autre part à évaluer leur utilité comme intrants météorologiques de modèles hydrologiques pour simuler des débits en rivière.

Pour atteindre cet objectif, trois réanalyses atmosphériques globales parmi les plus récentes et dont les résolutions spatiales sont parmi les moins grossières (ERA-Interim, CFSR et MERRA) (Dee et al. 2011; Rienecker et al. 2011; Saha et al. 2010) ont été considérées. Une réanalyse atmosphérique régionale (NARR) (Mesinger et al. 2006) a aussi été considérée dans la partie initiale de cette recherche. Les simulations hydrologiques ont été réalisées pour plus de 800 bassins versants situés aux USA et au Canada. Ces bassins versants représentent différents régimes hydroclimatiques. Le modèle hydrologique global HSAMI a été utilisé pour les simulations hydrologiques (Fortin 2000). Des données d'observation (données de stations météorologiques interpolées sur grilles) ont également été considérées dans cette recherche.

Pour les analyses hydrologiques effectuées aux USA, les données d'observations utilisées provenaient de la base de données interpolées de Santa Clara. Ces données ont préalablement été comparées à d'autres jeux de données d'observation aux USA (Article 1, Chapitre 3). Les résultats de cette comparaison ont mis en lumière l'existence d'écarts significatifs entre les différents jeux de données bien que ces derniers aient été interpolées quasiment à partir des mêmes bases de données climatologiques. Ce qui signifie que les méthodes d'interpolation utilisées ont introduit des incertitudes dans ces jeux de données malgré la forte densité du réseau de stations météorologiques aux USA. Toutefois, en dépit des incertitudes, les simulations hydrologiques réalisées à partir d'un modèle hydrologique global calé spécifiquement sur chacun de ces jeux de données étaient statistiquement équivalentes. Ainsi, du point de vue de la modélisation hydrologique globale, il n'y avait pas de raison de préférer un jeu de données d'observations à un autre. Toutefois, parmi les bases de données d'observation comparées, seule celle de Santa Clara fournissait à la fois des données de précipitation et température (minimale et maximale) sur la plus longue période d'intérêt considérée dans cette thèse (1979 – 2010). C'est alors ce qui a motivé le choix d'utiliser la base de données de Santa Clara.

Au Canada, la base de données de NRCan a été utilisée. D'autres bases de données d'observations auraient pu être utilisées, mais de récents travaux (exemple Gbambie et al. (2016)) ont montré que les performances des simulations hydrologiques basées sur NRCan étaient globalement similaires à celles d'autres bases de donnes interpolées au Canada, malgré d'importantes différences les bases de données. Ainsi, l'utilisation d'une base de données interpolée autre que NRCan n'aurait pas nécessairement changé les principales conclusions de cette thèse.

La comparaison des données présentée dans l'Articles 2 (Chapitres 4) a montré qu'en général, les données de température et de précipitation provenant des réanalyses globales étaient différentes de celles des observations. Les différences variaient d'une saison à une autre, et d'une région climatique à une autre. Toutefois, en général, les différences étaient plus élevées entre les données de précipitation comparativement aux données de température. Ceci signifie qu'il y a plus d'incertitude dans les données de précipitation, principaux intrants météorologiques en modélisation hydrologique (Duncan et al. 1993; Fekete et al. 2004). Tel que discuté dans la thèse, ces incertitudes proviennent à la fois des réanalyses et des données d'observation.

Aux USA (Article 2, Chapitre 4), les plus larges écarts entre les réanalyses globales et les données d'observation ont été observés dans les régions Continentale et Subtropicale humides (dans l'Est des USA). Cela est possiblement dû au fait que dans cette partie des USA, la saison estivale très humide est fortement dominée par des systèmes convectifs qui génèrent des orages que les réanalyses globales ne reproduisent pas adéquatement, à cause de leur résolution spatiale grossière. Par contre, dans l'Ouest des USA, la précipitation des réanalyses était globalement similaire à celle des observations. Cette similarité pourrait s'expliquer par la prédominance de systèmes météorologiques dynamiques (dans l'Ouest des USA) pendant la saison froide / humide, ce qui rend la précipitation générée plus prévisible par les réanalyses en dépit de leur résolution spatiale grossière.

Bien que le modèle hydrologique ait été préalablement calé de façon spécifique en utilisant les données de chaque réanalyse, les débits mesurés n'avaient pas été adéquatement simulés dans l'Est des USA. Par contre, dans l'Ouest des USA, les débits mesurés avaient été simulés avec une précision satisfaisante (Nash-Sutcliffe médian > 0,7).

La correction de biais pourrait être envisagée en vue d'améliorer la performance des réanalyses globales dans les régions (comme l'Est des USA) où les biais des réanalyses sont élevés. Ce point a été examiné en considérant une base de données globale WFDEI (Weedon et al. 2014), basée sur la réanalyse ERA-Interim, et qui utilise des données d'observation à l'échelle mensuelle de CRU (Harris et al. 2014; New et al. 1999, 2000) et de GPCC (Rudolf and Schneider 2005; Schneider et al. 2014) pour corriger les biais des précipitation et température d'ERA-Intérim. Les résultats de modélisation hydrologique ont effectivement montré que dans l'Est des USA (où le réseau de stations météorologiques utilisées pour la correction de biais est très dense), les débits simulés en utilisant les données de WFDEI avaient globalement une précision significativement supérieure à celle des débits simulés à partir des données d'ERA-Interim. Par contre, dans l'Ouest des USA où la précipitation est plus facilement prédictible, les performances des simulations hydrologiques réalisées à partir des données d'ERA-Interim étaient statistiquement équivalentes à celles basées sur l'utilisation des données de WFDEI. En outre, dans la région Semi-aride (où le réseau de stations météorologiques est moins dense que dans le reste des USA), les débits forcés par les données d'ERA-Intérim étaient significativement plus précis que ceux forcés par les données de WFDEI. Ces résultats suggèrent que la correction de biais des réanalyses n'est pas toujours nécessaire. Ils suggèrent aussi qu'en présence d'un réseau clairsemé de stations météorologiques, la correction de biais introduit des erreurs supplémentaires dans les données.

À l'instar de la correction de biais, l'assimilation de la précipitation de surface est souvent vue comme une approche pour réduire les incertitudes de la précipitation des réanalyses. Ce point a aussi été examiné en considérant la réanalyse régionale NARR. Contrairement à la précipitation des réanalyses globales, celle de NARR basée sur l'assimilation de la précipitation de surface, était plus comparable à la précipitation des observations, sur l'ensemble des USA. Par ailleurs, l'utilisation de la précipitation de NARR a permis de simuler adéquatement les débits mesurés tant dans l'Ouest des USA que dans l'Est des USA. Ces résultats montrent l'importance de l'assimilation de la précipitation de surface. Toutefois, de récents travaux ont montré que la bonne capacité de NARR aux USA où le réseau de stations météorologiques est très dense, ne s'étend pas au Canada (en particulier au Nord du Canada) où le réseau de stations météorologiques est beaucoup moins dense (Bukovsky and Karoly 2007; Langlois et al. 2009).

Au Canada (Article 3, Chapitre 5), l'utilisation des données provenant des réanalyses globales a globalement permis de simuler adéquatement les débits moyens journaliers provenant de 316 bassins versants. Pour l'ensemble de ces bassins versants, les valeurs médianes du critère d'efficience de Nash-Sutcliffe étaient 0.81 pour ERA-Interim, 0.80 pour CFSR et 0.77 pour MERRA (figure 5.6a, Chapitre 5). De plus, la précision des débits simulés en utilisant les données des réanalyses globales étaient dans l'ensemble, similaire à celle des débits simulés à partir des données d'observation (Nash-Sutcliffe médian de 0.80). Les résultats ont par ailleurs montré que, lorsque la densité de stations météorologiques décroît, la performance des simulations hydrologiques réalisées à partir des données d'observation tend aussi à décroître, ce qui n'est pas le cas avec les données des réanalyses. Cela peut s'expliquer par le fait que les réanalyses utilisent une variété de sources de données météorologiques, et ne sont donc pas directement dépendantes des stations météorologiques. C'est d'ailleurs l'une des raisons pour lesquelles l'utilisation des réanalyses semble intéressante pour les régions où le réseau de stations météorologiques est de faible densité. De façon générale, dans la région montagneuse de l'Ouest du Canada où la précipitation a une forte variabilité spatiale (Bailey et al. 1997), les débits simulés à partir des données des réanalyses (notamment ERA-Interim) étaient significativement plus précis que ceux simulés à partir des observations, lorsque la densité de stations météorologiques est inférieure à 1 station pour 1000km^2.

Parmi les trois réanalyses atmosphériques globales considérées dans la thèse, ERA-Interim s'est globalement montré la plus performante, tant aux USA qu'au Canada. Pourtant, sa résolution spatiale est plus grossière que celle des deux autres réanalyses globales. Ce qui indique que la résolution spatiale d'une réanalyse ne détermine pas à elle seule, le potentiel de cette réanalyse en modélisation hydrologique. La précision du modèle numérique de prévision météorologique, la qualité des données assimilées et le schéma d'assimilation des données sont tous des facteurs pouvant influencer le potentiel d'une réanalyse en modélisation hydrologique. En effet, ERA-Interim a un schéma d'assimilation de données plus sophistiqué que celui des deux autres réanalyses globales (Dee et al. 2011; Saha et al. 2010; Suarez et al. 2008).

En général, au Canada, les débits simulés à partir des données des réanalyses globales sont plus précis que ceux simulés aux USA à partir des données des mêmes réanalyses. Cela est possiblement dû à la taille plus élevée des bassins versants Canadiens comparativement à celle des bassins versants des USA. En effet, il est ressorti des résultats de l'article 3 (figure 5.7, Chapitre 5) que la précision des débits simulés à partir des données provenant des réanalyses tend à croître avec la taille du bassin versant (possiblement à cause de leur résolution spatiale grossière). Une autre raison pouvant aussi expliquer la bonne performance des réanalyses globales au Canada est la prédominance de la neige et à la forte contribution de l'eau de fonte au ruissellement total, ce qui rend les débits en rivière relativement plus facile à simuler par un modèle hydrologique.

Tant aux USA qu'au Canada, la combinaison des trois réanalyses globales avec les données d'observation a permis d'améliorer significativement la précision des débits simulés pour la plupart des bassins versants (Article 4, Chapitre 6). Les améliorations étaient plus remarquables au Canada où la couverture spatiale de stations météorologiques est plus faible. De plus, pour tous les bassins versants où la précision des débits simulés en utilisant uniquement des données d'observation est faible (correspondant à une valeurs du Nash-Sutcliffe < 0,5), la prise en compte des données des réanalyses a permis d'améliorer considérablement la précision des débits simulés (valeurs de Nash-Sutcliffe augmentées d'au moins 0,3) (figure 6.6, Chapitre 6).

7.2 Contribution à l'avancement des connaissances et originalité de la recherche

Les travaux antérieurs sur l'étude du potentiel des réanalyses en modélisation hydrologique se sont tous focalisés sur un nombre très réduit de bassins versants (en général 1 à 5 bassins versants), ce qui ne permettait pas de généraliser leurs conclusions (Choi et al. 2009; Fuka et al. 2014; Vu et al. 2012; Woo and Thorne 2006). Contrairement à ces études, la présente recherche a exploré plus large, en analysant le potentiel des réanalyses pour plus de 800 bassins versants de tailles différentes, répartis dans plusieurs régions climatiques aux USA et au Canada. Ainsi, le potentiel de chaque réanalyse a été analysé séparément sur la base des résultats obtenus pour plusieurs dizaines de bassins versants par régions climatiques. Ce qui rend plus généralisables les résultats obtenus, et fait de cette recherche, une contribution sur laquelle les travaux scientifiques futurs pourront s'appuyer. Par ailleurs, les résultats présentés dans cette thèse pourront aider les scientifiques et gestionnaires à prendre des décisions plus éclairées, notamment en ce qui concerne le choix et l'utilisation des réanalyses en modélisation hydrologique, dans des régions telles que le Nord du Canada.

De plus, comparativement aux études antérieures, les réanalyses globales évaluées dans la présente recherche figurent parmi les plus récentes et les moins grossières. Les résultats qui en découlent contribueront à identifier les améliorations futures à prévoir, en vue de rendre les réanalyses plus performantes en modélisation hydrologique.

La comparaison de la performance des réanalyses à celle des observations en fonction de la densité des stations météorologiques, a aussi été un point important de cette recherche. Elle a permis de vérifier l'hypothèse selon laquelle les réanalyses seraient plus performantes que les données d'observation lorsque le réseau de stations météorologiques en place est de faible densité. Cette hypothèse a été validée dans l'Ouest canadien, notamment pour les réanalyses MERRA et ERA-Interim.

Cette recherche a aussi montré que la combinaison des réanalyses avec les données d'observation permet d'améliorer significativement la précision des débits simulés. À notre connaissance, une telle approche (impliquant des réanalyses) et son impact en modélisation hydrologique n'a pas encore été rapportée dans la littérature. C'est donc une approche novatrice qui contribuera à améliorer la précision des simulations hydrologiques dans les régions où il y a peu de stations météorologiques.

Les réanalyses utilisées dans cette thèse sont mises à jour à l'usage du public en temps quasi-réel ou réelle (environ 8 semaines de retard pour ERA-Intérim, 2-4 semaines pour NARR, 2-3 semaines pour MERRA, quelques heures pour CFSR). Cette disponibilité des données sur une période de 30 ans et plus (1979 – présent) est intéressante pour plusieurs études en lien avec la gestion des ressources hydriques, notamment pour la planification hydrologique. Dans le cas particulier de la réanalyse CFSR où les données du Climate Forecast System Version 2 (CFSv2) (Saha et al. 2014) sont disponibles en temps réel (moins de 24 heures de retard), il est possible d'envisager une utilisation de ces données en mode opérationnel, pour les prévisions hydrologiques à court terme par exemple.

L'originalité de cette thèse réside principalement dans :

- l'approche méthodologique pour évaluer le potentiel des réanalyses en modélisation hydrologique, incluant la prise en considération de plusieurs centaines de rivières représentant différents régimes hydroclimatiques.
- la combinaison des réanalyses avec les observations pour améliorer la précision des simulations hydrologiques dans les régions où il y a peu de stations météorologiques.

7.3 Limites de la recherche et travaux futurs

L'une des limites de cette recherche est l'utilisation d'un modèle hydrologique global bien que les données météorologiques utilisées comme intrants (réanalyses et observations) soient sous forme de grilles. L'utilisation d'un tel modèle a nécessité le calcul des moyennes spatiales des grilles, préalablement à leur utilisation dans le modèle hydrologique. Ainsi, certains avantages potentiels liés à la résolution spatiale des grilles plus fines, en particulier dans les régions où la variabilité spatiale de la précipitation est élevée, ont possiblement été cachés. Il aurait peut-être fallu utiliser un modèle hydrologique distribué ou semi-distribué afin de faire ressortir les éventuels avantages des grilles plus fines. Toutefois, la calibration d'un modèle distribué ou semi-distribué pour un si grand nombre de bassins versants aurait été un travail ardu qui n'aurait pas nécessairement conduit à des conclusions différentes de celles de cette recherche.

Le potentiel des réanalyses en rapport avec la simulation des débits extrêmes n'a pas été examiné dans cette recherche, car cela est au-delà des objectifs visés. Quelques travaux antérieurs ont évalué la représentation de précipitation et température extrêmes dans les réanalyses (Donat et al. 2014; Pitman and Perkins 2009; Sun and Barros 2010). Mais, la question de savoir si les données des réanalyses peuvent être utilisées pour simuler ou prévoir les débits extrêmes, notamment dans les régions où les mesures de terrains sont déficitaires, demeure un point intéressant à examiner.

Seules deux variables des réanalyses (précipitation et température) ont été évaluées. Toutefois, hormis ces deux variables, les réanalyses fournissent plusieurs dizaines d'autres variables climatiques. Considérant les résultats présentés en annexe I, certaines de ces variables telles que l'évapotranspiration et le ruissellement pourraient être intéressantes pour les études en modélisation hydrologique.

Le critère de Nash-Sutcliffe a été le seul critère utilisé dans cette recherche. Son choix a été déterminé par le fait qu'il est le plus largement utilisé dans la littérature, ce qui facilite la comparaison des résultats obtenus dans cette recherche à ceux d'autres travaux publiés. Toutefois, il pondère plus fortement les débits de crues. L'utilisation d'autres critères tels que le biais relatif, l'erreur RMSE normalisée, etc. (Gupta et al. 1998; Moriasi et al. 2007; STREAMFLOW 2009) pourrait éventuellement conduire à des résultats légèrement différents.

Sur la base de ce qui précède, les avenues de recherche suivantes sont proposées pour de futurs travaux :

- reprendre les simulations hydrologiques pour des bassins versants de tailles moyennes à grandes (plus de 1000 km^2) en utilisant un modèle hydrologique distribué ou semi-distribué;
- évaluer l'utilité des données des réanalyses en rapport avec la simulation des débits extrêmes dans les régions où les stations météorologiques sont clairsemées;
- évaluer l'utilité d'autres variables climatiques des réanalyses (exemples, évapotranspiration, ruissellement, humidité du sol, etc.) en modélisation hydrologique;
- reproduire les travaux présentés dans cette recherche dans d'autres régions hydroclimatiques au monde, où il y a peu de stations météorologiques, en utilisant en plus du critère de Nash-Sutcliffe, d'autres critères de performance (exemple, biais relatif, l'erreur RMSE normalisé, etc.).

Par ailleurs, dans cette thèse, la combinaison réanalyses-observations a été utilisée en mode statique (c'est-à-dire les poids ne variaient pas dans le temps). Toutefois, compte tenu de la mise à jour en temps réelle des données de certaines réanalyses (exemple CFSR), il serait intéressant pour les travaux futurs, d'envisager la combinaison réanalyses-observations en mode dynamique (c'est-à-dire faire varier dans le temps, les poids utilisés pour les pondérations). Des algorithmes tels que le filtre de Kalman pourraient être utilisés à cette fin.

CONCLUSION

Les stations météorologiques sont les principales sources terrestres des données météorologiques utilisées en modélisation hydrologique. Cependant, ces stations ne fournissent que des mesures ponctuelles de sorte que dans les régions où leur couverture spatiale est faible, les données météorologiques requises pour les études hydrologiques (notamment les données de précipitation et de température) sont souvent déficitaires.

L'objectif global de cette thèse était d'évaluer le potentiel des réanalyses comme alternative aux stations météorologiques dans les régions où les stations sont clairsemées ou inexistantes. Plus précisément, il s'agissait d'évaluer le potentiel des données de précipitation et de température provenant des réanalyses atmosphériques pour les études de modélisation hydrologique. Dans ce cadre, le projet de recherche a essayé de répondre aux trois objectifs spécifiques ci-après, tout en considérant plus de 800 bassins versants répartis aux USA et au Canada, dans différentes régions climatiques :

- évaluer les performances des simulations hydrologiques réalisées en utilisant des données de précipitation et température de réanalyses globales (ERA-Interim, CFSR et MERRA), comme forçages météorologiques d'un modèle hydrologique;
- comparer pour diverses densités de stations météorologiques, les performances de simulations hydrologiques basées sur l'utilisation de données de réanalyses, à celles basées sur l'utilisation de données d'observation;
- évaluer l'impact de la combinaison de données de précipitation et température de réanalyses et celles des observations, sur la précision des débits moyens journaliers simulés.

Les résultats ont montré pour la plupart des bassins versants considérés, que les débits mesurés étaient adéquatement simulés par le modèle hydrologique en utilisant des données de réanalyses. Par ailleurs, la précision des débits simulés à partir des données de réanalyses était équivalente (et parfois supérieure) à celle des débits simulés en utilisant des données d'observation, sauf dans l'Est des USA.

Les réanalyses ont tendance à être plus performantes que les observations interpolées lorsque la densité de stations météorologiques diminue. En particulier, dans les montagnes de l'Ouest canadien, pour des densités de stations météorologiques inférieures à 1 station pour 1000km^2, la précision des débits simulés à partir des données de réanalyses (notamment ERA-Interim et MERRA) était considérablement supérieure à celle des débits simulés en utilisant des données d'observation.

La combinaison des données provenant des réanalyses avec celles provenant des observations a permis d'améliorer significativement la précision des débits simulés. Cette amélioration était particulièrement remarquable au Canada où les stations météorologiques sont clairsemées.

En définitive, les résultats de cette recherche ont montré que les données de précipitation et température provenant des réanalyses atmosphériques ont effectivement un fort potentiel pour les études en modélisation hydrologique. Par ailleurs, il a été montré que ce potentiel est plus élevé que celui des données d'observation dans des régions canadiennes à faible densité de stations météorologiques. Sur la base de ces résultats, les réanalyses peuvent effectivement servir comme alternative aux stations météorologiques dans les régions où les stations sont clairsemées. Toutefois l'utilisation des données de réanalyses en modélisation hydrologique devrait se faire avec précaution compte tenu des biais spatialement variables et de la non fermeture du bilan d'eau ce certaines réanalyses (annexe I). Enfin, il ressort de cette recherche que l'utilisation combinée des données provenant des réanalyses et de celles provenant des quelques stations météorologiques disponibles est une bonne approche pour améliorer la précision des débits simulés dans les régions où les stations météorologiques sont éparses.

ANNEXE I

ARTICLE DE COLLABORATION. USE OF FOUR REANALYSES DATASETS TO ASSESS THE TERRESTRIAL BRANCH OF THE WATER CYCLE OVER QUEBEC, CANADA

Florent Sabarly[1], Gilles Essou[1], Philippe Lucas-Picher[1,2], Annie Poulin[1] and François Brissette[1]

[1] Département de Génie de la Construction, École de technologie supérieure, 1100 rue Notre-Dame Ouest, Montréal, Québec, Canada, H3C 1K3.
[2] Département des Sciences de la Terre et de l'Atmosphère, Université du Québec à Montréal, 405 Rue Sainte-Catherine Est, Montréal, Québec, Canada, H2L 2C4.

Article publié dans la revue « Journal of Hydrometeorology » en 2016

Abstract

Reanalyses have the potential to provide meteorological information in areas where few or no traditional observation records are available. The terrestrial branch of the water cycle of the reanalyses CFSR, MERRA, ERA-Interim and NARR, is examined over Quebec, Canada, for the 1979-2008 time period. Precipitation, evaporation, runoff and water balance are studied using observed precipitation and streamflows, according to three spatial scales: (1) the entire province of Quebec, (2) five regions derived from a climate classification, and (3) eleven river basins. The results reveal that MERRA provides a relatively closed water balance, while a significant residual was found for the other three reanalyses. MERRA and ERA-Interim seem to provide the most reliable precipitation over the province. On the other hand, precipitation from CFSR and NARR do not appear to be particularly reliable, especially over southern Quebec, as they almost systematically showed the highest and the lowest values, respectively. Moreover, the partitioning of precipitation into evaporation and runoff from MERRA and NARR does not agree with what was expected, particularly over southern, central and eastern Quebec. Despite the weaknesses identified, the ability of reanalyses to reproduce the terrestrial water cycle of the recent past (i.e. 1979 - 2008) remains globally satisfactory. Nonetheless, their potential to provide reliable information

must be validated by comparing reanalyses directly with weather stations, especially in remote areas.

I.1 Introduction

Since the end of the last century, interest in the study of climate change has grown considerably. Looking at observations or data produced by global climate models representing the recent past, and analyzing them using statistics could be relevant to detecting possible trends regarding climate change. In the '80s and '90s, considerable improvements in weather forecasting and numerous upgrades of model and data assimilation methods contributed to the notion of reanalysing the recent past (Bengtsson and Shukla 1988). Briefly, a reanalysis aims to provide the best estimation of the state of the atmosphere, ocean and land surface from the recent past. The term reanalysis stands for Retroactive Analysis, since the analysis is done on a past period extending up to the near present. Reanalyses are three-dimensional gridded datasets produced by a weather forecasting model. Two main characteristics can define these datasets. Firstly, a data assimilation scheme is used to integrate observations from different sources in order to provide the most coherent state of the atmosphere. Secondly, the assimilation scheme and the forecasting model of a reanalysis remain unchanged during the entire simulation period. As such, inconsistencies that might be induced by continuous updates of the data production system are avoided.

Many observations measured from different sources, such as radio sounding, aircrafts, boats, satellites, surface sensors, buoys, etc., are assimilated in the production of a reanalysis. Moreover, the large range of available reanalysis data products provides information such as radiative fluxes, wind, temperature, humidity, precipitation, albedo, snow, vegetation and land cover, to name just a few. The global coverage and the huge range of available variables with a consistent time and space resolution during the simulated period represent some of the main benefits that reanalyses provide to climate studies, including a smaller bias and a finer spatial resolution from one generation of reanalyses to the next. Nevertheless, reanalyses also show some limitations. Indeed, the reliability of some variables may significantly vary in time and space. Moreover, the evolution of the number and quality of assimilated observations may introduce some artificial variability and trends. As well, the water balance is rarely conserved, and reanalyses sometimes show substantial biases between variables, such as precipitation, that are not directly constrained by assimilation.

In the '90s, the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP) began producing the first global reanalyses, the ECMWF ERA-15 reanalysis (Gibson et al. 1997) and the NCEP National Center for Atmospheric Research Reanalysis (NCEP-NCAR), also known as R1 (Kalnay et al. 1996). Armed with awareness about the limitations of these first two datasets, a second generation of reanalyses was produced, namely, the Department Of Energy reanalysis (NCEP-DOE), also known as R2 (Kanamitsu et al. 2002), the North American Regional Reanalysis (NARR) (Mesinger 2004), the ECMWF ERA-40 reanalysis (Uppala et al. 2005) and the Japanese 25-yr. reanalysis JRA-25 (Onogi et al. 2007) developed by the Japan Meteorological Agency (JMA). Recently, a third generation of reanalysis was developed, including the Climate Forecast System Reanalysis (CFSR) (Saha et al. 2010) from NCEP, the ECMWF ERA-Interim reanalysis (Dee et al. 2011) and the Modern-Era Retrospective analysis for Research and Applications (MERRA) produced by NASA (Rienecker et al. 2011).

In addition to their application to climate studies, reanalyses have the potential to provide climate information in areas that are sparsely inhabited or with limited surface observations (in space or time), for applications such as water resource management and hydrological modelling. For the latter, examining precipitation, evaporation, runoff and the water balance for the terrestrial branch of the water cycle (Peixoto and Oort 1992) obtained through reanalyses should provide useful information.

Bukovsky and Karoly (2007) compared the precipitation of NARR, R2 and ERA-40 using a set of gridded observations. Exploring the spatial distribution and the diurnal and annual cycles, they revealed that NARR showed better results than the other two reanalyses over the continental United States. Nevertheless, these authors recommended that users should proceed cautiously when looking at the rest of the North American domain, especially along the US borders and in South-eastern Canada, where the overall NARR precipitation is strongly underestimated. Among all existing reanalyses, NARR is the only one that assimilates precipitation, and in which the quality of the simulated precipitation strongly depends on the quality of observed data and on the assimilation process (Mesinger 2004). Zhang et al. (2012) focused on the change in the global average of precipitation from CFSR during the period 1998-2001. They demonstrated that an interaction between the bias of the data assimilation model and the nonstationarity in the ingestion of some observed data is the source of the global average increase in the CFSR precipitation after 1998. Bosilovich et al. (2011) evaluated the water balance of MERRA over the entire globe. Since this reanalysis was configured to include in the water balance the residual generated by the data assimilation process, the water balance over the land and oceans is closed. However, a notable shift in annual water balance was identified, starting in 1999, which coincides with the beginning of Advanced Microwave Sounding Unit (AMSU) radiance assimilation. Sheffield et al. (2012) identified noteworthy differences in NARR evaporation and runoff, compared to two offline land surface model simulations (Noah v2.7.1 and Variable Infiltration Capacity - VIC), using observational runoff estimates over the continental United States. NARR (which uses a previous version of Noah) and Noah simulations present an overestimation of annual evaporation, and runoff ratios (simulated runoff divided by observed runoff) that are 50% lower than the VIC simulation results. Regarding NARR, the authors identified these differences as being mainly related to the evaporation component of the Noah model, versus other factors such as atmospheric forcings or biases induced by precipitation assimilation into NARR. Lorenz and Kunstmann (2012) investigated the closure of the water balance (i.e. balance between precipitation, evaporation, surface runoff and moisture flux) of ERA-Interim, MERRA and CFSR over the entire globe, and found that ERA-Interim is the reanalysis that likely provides the most reliable rainfall estimates globally, especially over regions with a dense network of observations. Moreover, they showed that, in the long-term mean, ERA-Interim and MERRA show a reasonable closure of the global surface water balance, as P - E (precipitation minus evaporation) over land equals the divergence of moisture *E* - *P* over the oceans. Furthermore, the change in the number of assimilated data in CFSR and MERRA around 1998, revealed by Zhang et al. (2012) for the case of CFSR, leads to a substantial imbalance between *P* - *E* over the land and oceans.

The province of Quebec in Canada has abundant freshwater resources. Its numerous lakes and rivers play a fundamental role in local wildlife and flora sustainability. On the other hand, there is great interest in managing water resources, particularly in terms of hydroelectricity production. Quebec has a surface area greater than 1.5 million km^2, and is characterized by five different climate regimes (Bukovsky 2011). However, the spatial distribution of meteorological stations varies across the province, dense in the south, decreasing to the north, to almost non-existent in the far north. Therefore, reanalyses should be useful in providing substantial information, particularly in these remote regions with little observational data. This study focuses on the assessment of the components of the terrestrial branch of the water cycle (Peixoto and Oort 1992) of four recent reanalyses, CFSR, ERA-Interim, MERRA and NARR, over the province of Quebec, and especially on the assessment of the reliability of the four reanalyses in representing the terrestrial water cycle components in the northern regions of the province. The analysis is divided into three parts. In the first part, the long-term mean water balance is examined over the entire territory of Quebec (the water balance will be further defined in the Data and Methods section using equation A I-1). Secondly, the mean annual cycle (expressed through long term mean monthly values) of the terrestrial water cycle components and annual water balance are analysed according to the climate classification of Bukovsky. Thirdly, the precipitation, runoff and *E/P* ratio (evaporation divided by precipitation) are examined over river basins representing the different hydrological regimes of the province. Section I.2 details the methodology and the main characteristics of the datasets used in this study. Section I.3 presents the results and discussion and section 4 provides the concluding remarks.

I.2 Data & methods

The reanalysis and observational datasets that were used in this study are first described, and then the methodology follows.

I.2.1 Reanalysis datasets

The widely used NARR, CFSR, MERRA and ERA-Interim reanalyses were chosen to evaluate the terrestrial branch of the water cycle over the province of Quebec. These reanalyses benefited from improvements obtained from the preceding generation of reanalyses, and thus represent the most advanced and suitable products available for carrying out this study. This section describes the relevant properties of each reanalysis as used in this study, including their similarities and differences. For the analysis, we selected the 1979-2008 period since the 30-year period is widely used in climate studies. Moreover, the selected period starts in 1979, as it is the first year covered by the four selected reanalyses. All the references and main properties of the datasets used in this study are summarized in table-A I-1.

Table-A I-1 Main properties summary of reanalysis datasets

	CFSR	**ERA-Interim**	**MERRA**	**NARR**
Organization	NCEP	ECMWF	NASA	NCEP
Dates	1979-present	1979-present	1979-present	1979-present
Surface resolution	0.3° (~33 km)	0.75° (~83 km)	2/3° x 1/2° (~74 x 56 km^2)	~32 km
Assimilation approach	3D-VAR	4D-VAR	3D-VAR	3D-VAR

	CFSR	ERA-Interim	MERRA	NARR
Particularity	Surface model forced by observed P	Uses 4D-VAR scheme	Designed to balance the water cycle	Assimilates precipitation
Surface output times	6h	3h	1h	3h
References	Saha et al. (2010)	Dee et al. (2011)	Rienecker et al. (2011)	Mesinger et al. (2004)

The recent global ECMWF ERA-Interim reanalysis includes an assimilation system based on a 4D-VAR approach, which is more complex and computationally intensive than the 3D-VAR used by CFSR, NARR and MERRA, but uses more efficiently available observations (Dee et al. 2011). Most of the surface fields are available every 3 hours and every 6 hours for the atmospheric fields, on a 0.75-degree regular grid (about 83-km horizontal resolution).

NARR, which covers the North American continent and parts of the North Atlantic and Pacific Oceans, has the particularity of assimilating observed precipitation, contrary to the other three reanalyses. In fact, before being assimilated, the observed precipitation in NARR is converted into latent heat. Additional details about how NARR is generated and about precipitation assimilation can be found in (Mesinger 2004). NARR outputs are available every 3 hours, with a 32-km horizontal resolution (about 0.3-degree).

CFSR is the first coupled global atmosphere, ocean, land surface and sea ice system reanalysis (Saha et al. 2010). As for ERA-Interim and MERRA, the atmosphere, ocean, land surface and sea ice models share boundary data during the forecasting process. However, the analysis of each model component in CFSR, ERA-Interim and MERRA is processed separately. Both CFSR and NARR use the Noah land surface model, but in CFSR, the Noah model has the particularity to be forced with observed precipitation, instead of using the precipitation generated by the atmospheric model, which is considered too biased (Saha et al.

2010). Furthermore, Meng et al. (2012) explained that "previous studies have shown nontrivial biases in the Global Data assimilation System precipitation (Gottschalck et al. 2005). Such a bias over land often leads to biases in many simulated land surface variables". Surface fields of CFSR are available at a resolution of 0.3-degree (about 32-km horizontal resolution), and every 6 hours.

MERRA is a global reanalysis simulated on a 2/3-degree longitude and 1/2-degree latitude regular grid, with hourly outputs. The assimilation process of MERRA is similar to that of CFSR, which is the Grid point Statistical Interpolation scheme, developed at NCEP. In addition, the main specificity of MERRA consists in the use of an incremental analysis update (IAU) procedure to improve water balance conservation (for further explanations, see Rienecker et al. (2011)).

I.2.2 Observational datasets

This section describes the NRCAN observationally-based gridded precipitation, as well as the $(cQ)^2$ observed streamflow time series used as reference in this study. The NRCAN dataset provides daily Canada-wide precipitation, minimum and maximum temperatures on a 10-km resolution regular grid (Hutchinson et al. 2009). It was originally produced to support studies requiring daily data, for instance, hydrological modelling, agricultural and forestry applications, extreme event analysis, etc. The gridded precipitation was obtained from the interpolation of measurements at individual stations, whose number varies in time from about 2000 to 3000 for the 1961-2010 period. The ANUSPLIN model was applied for the interpolation. It uses a trivariate thin plate smoothing splines method to model the spatial distribution of precipitation as functions of latitude, longitude and elevation, across Canada. Figure-A I-1 shows the locations of the 638 weather stations available to generate the daily precipitation for the period 1979-2008 over the province of Quebec, provided by Environment Canada.

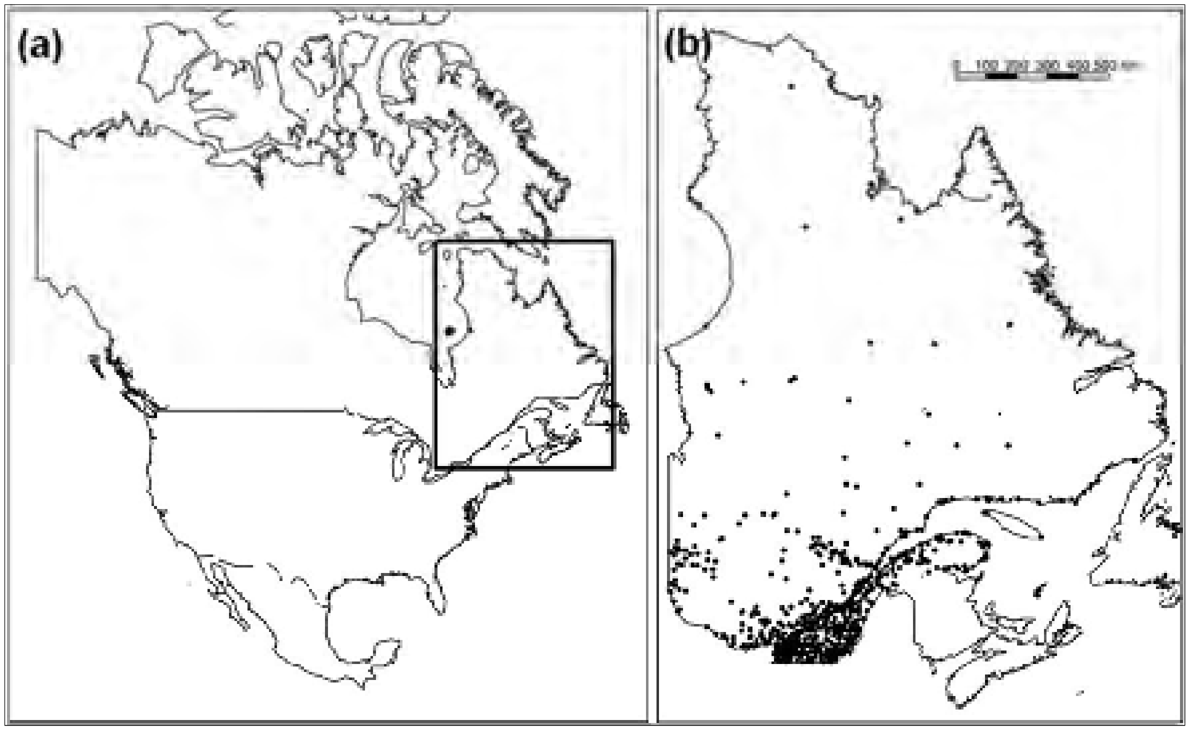

Figure-A I-1 (a) Location of the Quebec province in North America. (b) Locations of the 638 Environment Canada weather stations used to generate the daily NRCAN dataset for the period 1979-2008 (Hutchinson et al. 2009) over the eastern provinces

The $(cQ)^2$ database (Impact des Changements Climatiques sur l'hydrologie(Q) au Québec) contains daily outlet streamflows, names and surface areas, as well as center and contour coordinates of 306 river basins over the province of Quebec (Arsenault and Brissette 2014b). This database was jointly produced by Hydro-Quebec, Rio Tinto Alcan and the Centre d'Expertise Hydrique du Québec, in order to unify water resources information and simplify access to hydrometric data.

Both observational datasets have some known limitations mainly related to streamflow measurements in $(cQ)^2$ and to decreasing spatial density of weather stations used to produce the NRCAN dataset from southern to northern Quebec. For instance, over the Saguenay-Lac Saint Jean and west of the Côte Nord regions, NRCAN shows inconsistent precipitation between 1995 and 1996, and significant underestimation from 2004 to 2008, due to the malfunction of a weather station. Moreover, measuring consistent streamflow time series in northern remote areas is challenging and may introduce some errors. Therefore, some of the $(cQ)^2$ time series have been post-processed to correct some inconsistencies; for instance, natural streamflow reconstruction over regulated river basins or measurement correction due to inaccurate measurements of river flows under ice cover. Despite the application of post-processing techniques, some biases remain, and streamflow time series in the southern part of the province may be more reliable than those in the northern regions. Moreover, as the density of observed data integrated into the NRCAN decreases from South to North, the reliability of this dataset may be more questionable over northern regions.

I.2.3 Methodology

The terrestrial water balance of the reanalyses can be written following equation A I-1:

$$\frac{\partial W}{\partial t} = P - E - R - RES \qquad \text{(A I-1)}$$

where W is the surface water storage (mm), P is the precipitation (mm day^{-1}), E is the total surface evaporation (mm day^{-1}) and R is the total runoff (mm day^{-1}), including the subsurface runoff. The term RES (mm day^{-1}) stands as a residual that comes from the assimilation process. During this process, a new state of the atmosphere is generated, which is different from the one simulated by the model, but closer to observations. Atmospheric states are then discontinuous, which unbalances the water balance. In other words, the RES term may also be viewed as an estimate of the overall error in the water balance (Roads et al. 2003). Roads et al. (2003) computed annual mean (1996-1999) surface variables of NCEP-NCAR and NCEP-DOE reanalyses over the Mississippi river basin, and presented the RES values of these two reanalyses which were respectively equal to 0.592 and 0.255 mm day^{-1}. In this study, equation A I-1 is rearranged into equation A I-2:

$$B = \frac{P - E - R}{P} \times 100 \qquad \text{(A I-2)}$$

where B is the relative water balance (% of P), which includes both RES and the change of surface water storage. Considering climatic time scales, the temporal change in surface water storage (soil water and snow water equivalent) is assumed to be negligible (Kleidon and Schymanski 2008). In this case, B equals RES. This assumption is not completely verified on an annual time scale. Nevertheless, Roads et al. (1998) concluded that the RES term should be the most important contribution in comparison with the change in water storage.

Precipitation, evaporation and runoff from CFSR, ERA-Interim, MERRA and NARR were first downloaded from their respective websites for the period going from 1979 to 2008 (30 years). Each variable is aggregated on a daily basis, and computed in millimeters per day. The evaporation includes evaporation from the bare soil, transpiration from vegetation, interception loss and snow sublimation. For CFSR, the evaporation component is derived from the latent heat flux (in W m^{-2}), as the evaporation is not directly available. Equation A I-3 was used to compute the evaporation:

$$E = \frac{H_f}{L_e(T)} \times C \tag{A I-3}$$

where E is the evaporation (mm day^{-1}), H_f the latent heat flux (W m^{-2}), $L_e(T)$ is the latent heat of vaporization (J kg^{-1}), and C is a constant to convert the evaporation to a daily time step (C=86400 s day^{-1}). Although the latent heat of vaporization depends on the surface temperature T, its influence on the evaporation can be neglected (Lorenz and Kunstmann 2012). Therefore, the latent heat of vaporization is approximated to 2.5 MJ kg^{-1} and equation A I-3 can be simplified as follows (equation A I-4):

$$E = H_f \times 0.03456 \tag{A I-4}$$

In the case of the NARR and MERRA datasets, baseflow and surface runoffs, which are provided separately, are summed to compute the total runoff. The analysis that was carried out in this paper is divided into three parts, according to the three different spatial scales that are considered. Long-term and annual time scales have also been computed, as introduced hereafter.

Long-term mean of the water cycle components over the province of Quebec

In the first part of this study, precipitation *P*, evaporation *E*, runoff *R*, and the relative water balance *B* from the four reanalyses are averaged at a daily time scale over the period 1979 - 2008 for each reanalysis tile. Firstly, *P*, *E* and *R* are averaged annually (mm day^{-1}) from October 1st to September 30th of the following year. As such, the snowpack is supposed to completely accumulate and melt, and then to be transferred to runoff during the year. Therefore, there are 29 annual values, since the periods of January 1st to September 30th of 1979, and of October 1st to December 31th of 2008 are not considered. Values of *P*, *E* and *R* are then averaged during the entire period on each tile following equation A I-5:

$$X = \sum_{i=1}^{n} \frac{x_i}{n} \tag{A I-5}$$

where X (mm day^{-1}) is the averaged variable *P*, *E* or *R*, x_i the value of *P*, *E* or *R* at year *i* and *n* the total number of years in the period 1979 - 2008 (29 values, as explained above). The relative water balance *B* is computed from the averaged values of *P*, *E* and *R* using equation A I-2. To consider only data over the land, land-sea masks of the four reanalyses are applied. As MERRA natively provides land cover fractions, a threshold of 0.5 was used to distinguish the land from the water surfaces.

Water cycle components over the climatic regions

In the second part of this study, spatial averages of *P*, *E*, *R* and *B* are computed over subregions of Quebec, according to the climate classification of Bukovsky (2011). This classification was created to provide a consistent climate division, as part of the North American Regional Climate Change Assessment Program (NARCCAP). The climate classification is based on a simplification of ecoregions of Ricketts (1999). According to this classification, Quebec is divided into five climatic regions: Great Lakes (GL), North Atlantic

(NA), East Boreal (EB), East Taïga (ETA) and East Tundra (ETU). Figure-A I-2a shows the five climatic regions over the province of Quebec, at the resolution of the CFSR dataset. The annual cycles of P, E and R are produced (one value per month per variable) following equation A I-6 :

$$Y = \frac{1}{p \times k} \sum_{j=1}^{k} \sum_{i=1}^{p} x_{ij} \tag{A I-6}$$

where x_{ij} is the value of P, E or R (mm day^{-1}) at day i and on tile j, p is the number of reanalysis tiles within a climatic region, k is the total number of days in a particular month throughout the entire 1979 - 2008 period (for example, k=930 for January). Finally, one value (Y, in mm day^{-1}) is computed for each month. The NRCAN precipitation is used as the reference for precipitation, and its mean annual cycle computed in the same way as for the reanalyses.

The relative water balance B of reanalyses is averaged annually (one value per year) for each climatic region, following the same methodology as for the preceeding calculation over the entire Quebec: annual average values of P, E and R are first computed using equation A I-6, except that k is now the number of days within the period from October 1st to September 30th of the following year. The relative water balance is then calculated using equation A I-2 for each year.

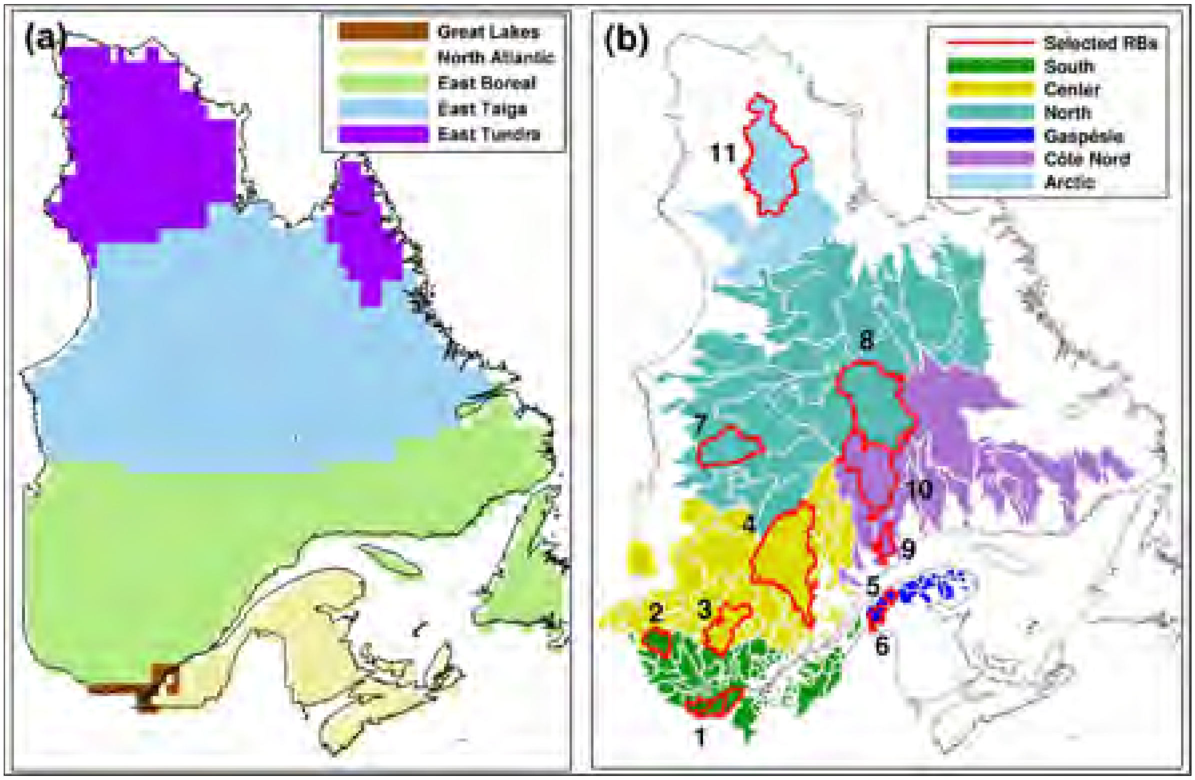

Figure-A I-2 (a) Climatic regions of Bukovsky for CFSR, and (b) Hydrologic regions and selected river basins (with red contours) from $(cQ)^2$

Water cycle components within river basins

In the third part of this study, *P, R* and the *E/P* ratio (evaporation divided by precipitation) are computed annually over eleven selected river basins from the $(cQ)^2$ database using equation A I-6 in which *p* is the number of dataset gridpoints within a river basin and *k* is the number of days within the period from October 1st to September 30th of the following year. Spatial averages are computed as follows: (1) when at least four dataset gridpoints are located inside the contour of a river basin (or aggregated river basin), a simple arithmetic mean is used; (2) otherwise, the Thiessen's polygons method is applied (Rhynsburger 1973), attributing specific weights to the four closest gridpoints inside or outside the river basin(s) contour. Table-A I-2 presents the number of gridpoints of the five datasets located inside the contour of the selected river basins. The river basins were selected according to three main criteria: (1) their representativeness of the six major hydrologic regimes of Quebec, namely South, Center, North, Gaspésie, Côte Nord and Arctic; since the $(cQ)^2$ streamflow time series contain missing data, especially during the winter low-flow period, (2) the basins with the longest period without any missing data, and (3) the longest temporal coverage were favoured. Rivers in Arctic and North regions flow to the west and north, whereas those in Côte Nord, Gaspésie, South and Center flow into the St. Lawrence River. These regimes correspond fairly well with the Bukovsky climatic regions, with a distinction being made between the Center and Côte Nord, as compared to the unique East Boreal climate region. Each hydrological regime is represented by two river basins, except for the Arctic, where only one river basin showed consistent observed streamflows. Furthermore, one of the two river basins in the South is formed by three small aggregated river basins; their streamflow time series were spatially aggregated, while their surface areas were summed up, which increases the number of reanalysis gridpoints within the river basin, providing more relevant information.

Table-A I-2 Number of grid points within the 11 river basins for CFSR, ERA-Interim, MERRA, NARR, and NRCAN

	Hydrologic region	**CFSR**	**ERA-Interim**	**MERRA**	**NARR**	**NRCAN**
(1) Aggregated RBs	South	12	3*	3*	5	169
(2) Kipawa	South	7	1*	2*	5	107
(3) Baskatong	Center	15	1*	6	13	221
(4) Lac St-Jean	Center	58	12	18	44	812
(5) Mitis-1	Gaspésie	2*	0*	1*	1*	22
(6) Temiscouata	Gaspésie	4	0*	1*	2*	47
(7) Opinaca	North	19	4	5	13	271
(8) Caniapiscau	North	55	8	13	33	737
(9) Manic-2	Côte Nord	6	1*	0*	4	75

	Hydrologic region	CFSR	ERA-Interim	MERRA	NARR	NRCAN
(10) Manic-5	Côte Nord	34	5	10	24	465
(11) Arnaud	Arctic	45	8	16	24	614
** Thiessen's polygons method has been applied on these river basins, using the four closest grid points.*						

Figure-A I-2b shows the hydrologic regions and the selected river basins of $(cQ)^2$ over Quebec. Among the 306 river basins, some are subbasins, and are not represented in figure-A I-2b. To further investigate precipitation from the four reanalyses, the temporal correlation coefficient between the daily time series of precipitation from each reanalysis and from NRCAN is computed over the river basins of interest. Moreover, distribution of the daily precipitation intensities from CFSR, ERA-Interim, MERRA, NARR and NRCAN over the eleven river basins for the period 1979-2008 for winter (DJF), and summer (JJA) is also computed. Each daily precipitation is categorised into 8 bins of precipitation from 0.25-1 mm day^{-1} to 64-128 mm day^{-1} and the bins are presented in percentage of their contribution to the total amount of seasonal precipitation. A threshold of 0.25 mm day^{-1} is applied to dissociate dry days from wet days. For the five dataset, no aggregation is applied, each gridpoint is considered as a daily precipitation within a day. Regarding the runoff analysis, runoff estimates are calculated using the $(cQ)^2$ streamflows relatively to the surface area of each of the 11 river basins. Moreover, NRCAN precipitation is used as reference data, as well as an estimated E/P ratio derived from NRCAN precipitation and $(cQ)^2$ streamflows. The reference ratio is calculated by total annual streamflow volumes (m^3 s^{-1}) being converted into runoff (mm day^{-1}) according to the surface area of each river basin. Assuming that the water balance of observed data is closed, an estimation of evaporation is computed using the equation A I-7:

$$E = P - R \tag{A I-7}$$

where E is the evaporation estimate derived from observations (mm day^{-1}), P is the NRCAN precipitation (mm day^{-1}) and R is the runoff (mm day^{-1}) derived from observed streamflows. As for the preceding calculations, equation A I-7 is used on an annual basis from October 1st to September 30th of the following year. The E/P ratio of estimated observations is then computed, dividing E by the NRCAN precipitation. Since streamflow time series differ from one river basin to another, E/P is calculated only during the available streamflow periods. Finally, those computations are supplemented with the analysis of the distribution of the yearly mean annual precipitation, runoff and ratio (mm day^{-1}) previously calculated from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the eleven selected river basins.

I.3 Results

I.3.1 Long-term mean of the water cycle components over the province of Quebec

Figure-A I-3 shows the long-term mean of precipitation, evaporation and runoff for the four reanalyses over the province of Quebec, as well as the contour of the river basin Manic-5, in which the Manicouagan Lake is located. Globally, the spatial distribution of precipitation from ERA-Interim and MERRA are quite similar, whereas precipitation from CFSR generally reaches the highest values among the four datasets (up to 5 mm day^{-1} over the south of the province). Moreover, the underestimation of precipitation from NARR highlighted by Bukovsky and Karoly (2007) at the US-Canadian border is clearly obvious in figure-A I-3. Regarding evaporation, MERRA shows higher values in the center and in the south relatively to the other three reanalyses (2.5 to 3 mm day^{-1}), while NARR and ERA-Interim show the lowest values in the north of the province (0 to 1 mm day^{-1}). CFSR and NARR show high runoff values close to the Manicouagan Lake, especially NARR, which reveals values above 3 mm day^{-1} (blue spot on the NARR runoff map in figure-A I-3), while the three other reanalysis runoffs range from 0.5 to 2.5 mm day^{-1}. Furthermore, ERA-Interim shows the highest runoff values in the center and the south, while runoff values from MERRA and NARR drop to almost zero in the west and south, respectively.

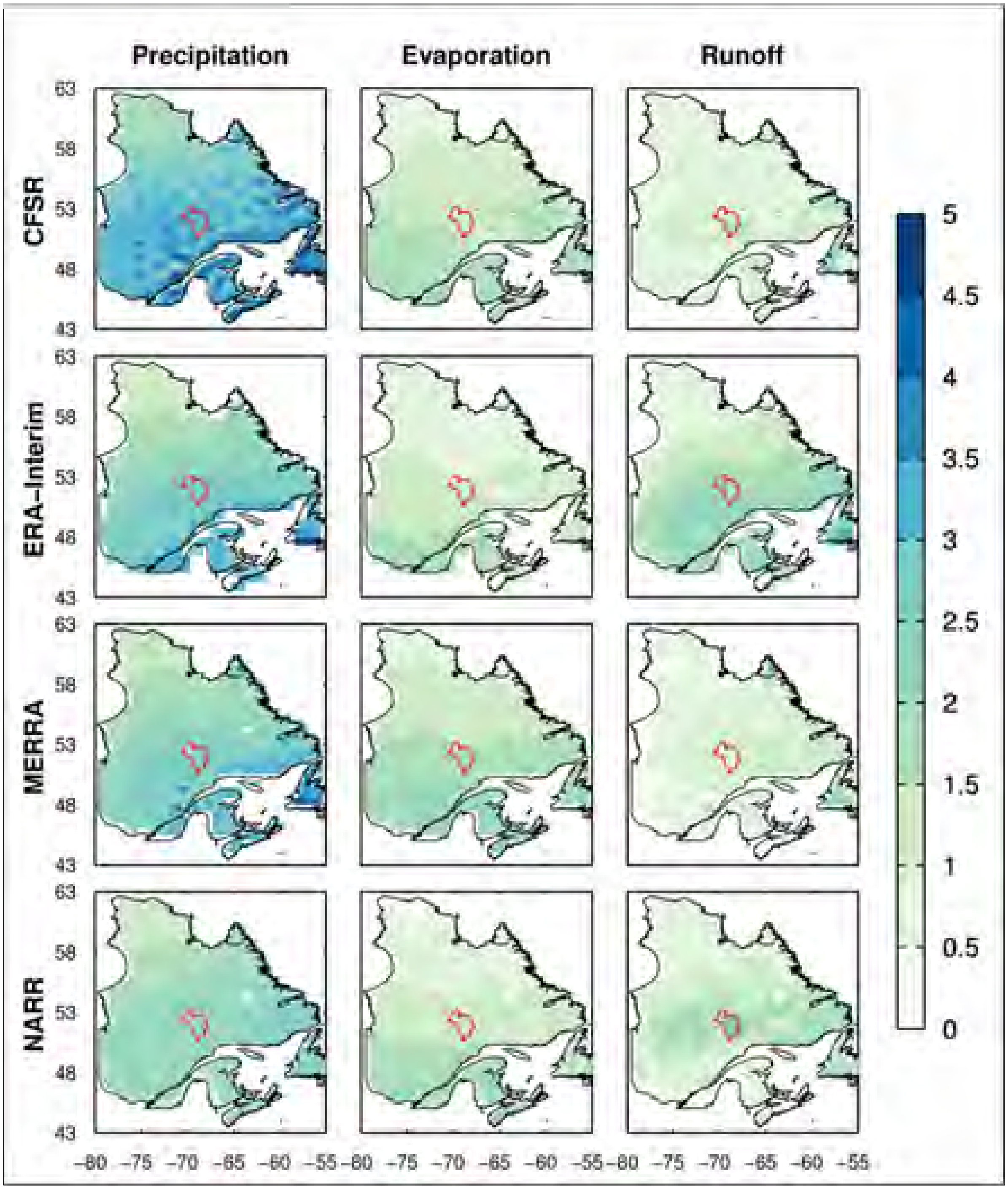

Figure-A I-3 Long-term mean precipitation, evaporation and runoff (mm day^{-1}) of CFSR, ERA-Interim, MERRA and NARR over the province of Quebec. The red line represents the contour of the river basin Manic-5 in which the Manicouagan Lake is located

Figure-A I-4 shows the closure of the water balance using the long-term mean *B* values computed for CFSR, ERA-Interim, MERRA and NARR over Quebec. As expected, the water balance of MERRA is closed, with values of *B* almost equal to 0. Conversely, the water balances of CFSR, NARR and ERA-Interim are not closed, with values of *B* different from 0. Over the entire province, CFSR presents positive values of *B* from 30 to 50%. These noticeable results are mainly related to the high values of CFSR precipitation illustrated in figure-A I-3. It is worth recalling that CFSR uses observed precipitation instead of that simulated to force its surface scheme, which introduces some imbalance between evaporation, runoff and the model-generated precipitation. Over central and southern Quebec, ERA-Interim shows *B* values between -30% and -10%, and a relatively closed water balance with *B* close to 0 over the rest of the province. NARR presents positive *B* values of about 20% in the far north latitudes, negative *B* values in the center, and reaches its lowest negative values in the south. The underestimations of NARR precipitation at the US-Canadian border and high runoff values close to the Manicouagan Lake illustrated in figure-A I-3 lead to significantly underestimated *B* values (from -70% to -90%).

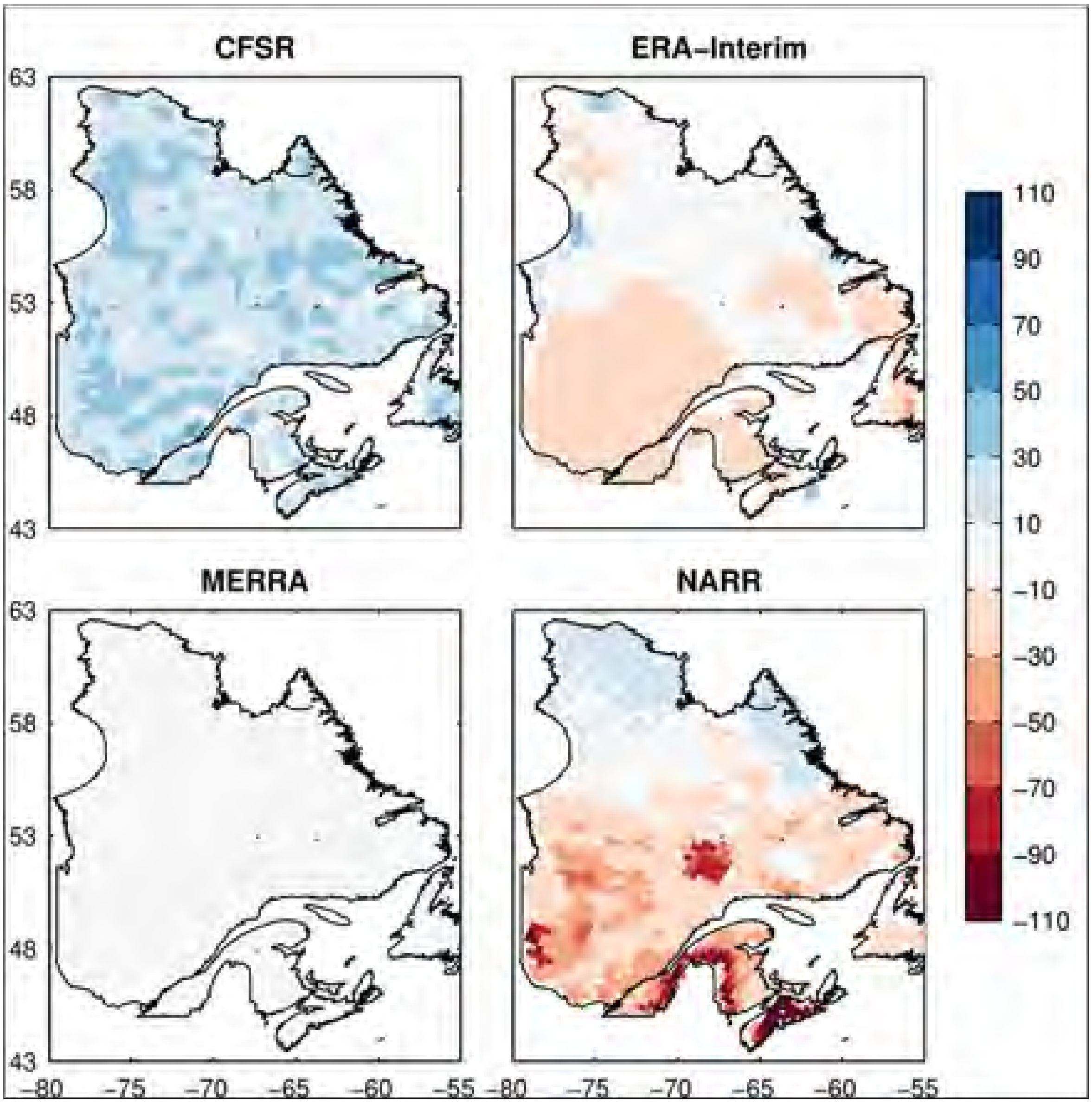

Figure-A I-4 Long-term mean relative water balances using *B* values (% precip) from CFSR, ERA-Interim, MERRA and NARR over the province of Quebec

I.3.2 Water cycle components over the climatic regions

Figure-A I-5 shows relative water balance *B*, runoff, evaporation and precipitation of the four reanalyses over the five climatic regions of Bukovsky, with the addition of the NRCAN dataset for the case of precipitation. Mean annual cycles are shown for runoff, evaporation and precipitation, whereas mean annual values are shown for the relative water balance *B* (one value per year), from 1979 to 2008.

As already seen in figure-A I-3, precipitation from CFSR is generally higher than that from the other four datasets over the five climatic regions, except in the summer months over the GL and NA regions (figure-A I-5). Precipitation from NARR is lower than for MERRA, ERA-Interim and CFSR, especially over the EB, GL and NA regions (0.5 to 1 mm day^{-1} below the values of the other reanalyses). Nevertheless, although precipitation values may differ greatly depending on the dataset, the temporal distributions are quite similar over the EB, ETA and ETU regions, which is in agreement with the precipitation from NRCAN. Over the GL and NA regions, precipitations from each dataset show discrepancies between another, mainly in the summer (JJA months).

Regarding evaporation, MERRA summer peak values are systematically higher than those from the other reanalyses, particularly over the EB, ETA and ETU regions, reaching respectively 4.3, 3.5 and 3 mm day^{-1}. Over the NA region, the evaporation values from MERRA are also very close to those from NARR, with summer peaks being 1.4 times higher than those from ERA-Interim and CFSR. Globally, the evaporation is questionable, as summer maximum values may double or triple from one reanalysis to another, depending on the climatic region. The general North-South gradients seen in figure-A I-3 for precipitation and evaporation are also reproduced in the mean annual cycles, as precipitation and evaporation generally tend to decrease from southern to northern regions, with precipitation from CFSR being an exception, as already pointed out, and precipitation over the GL region being lower than that over the NA region. Such gradients agree well with the known gradients in precipitation and evaporation across the province, also illustrated by the Atlas Of Canada (2014, 2015).

Runoff values from ERA-Interim are higher than those from MERRA, NARR and CFSR all year long for NA and GL regions, and during fall and/or winter seasons for EB, ETA and ETU regions (1 to 2 mm day^{-1} more than the other reanalyses). For instance, over the GL region, the ERA-Interim peak values during the spring are 2.5 times higher than those from the other three datasets. On the other hand, MERRA, NARR and CFSR agree well over this region. Likewise, MERRA, NARR and CFSR show similar runoff values over the NA region, whereas those for ERA-Interim are substantially higher. Moreover, MERRA presents higher maxima over ETA and ETU (4.3 and 3.6 mm day^{-1} respectively) compared with ERA-Interim, NARR and CSFR values. Generally, runoff values from the different reanalyses show more disagreement over the EB, ETA and ETU regions, as the low and high flow seasons show a large range of values.

The relative water balance values from the four reanalysis datasets cover a range from -50% to +50% of precipitation, except for the NARR dataset in the NA region, which will be discussed further. The high precipitation values and average values of evaporation and runoff from CFSR induce positive values of relative water balance *B* (about 25%), higher than for MERRA, ERA-Interim and NARR over the six climatic regions. In some years, the underestimation of precipitation from NARR leads to negative values of *B* over the five regions, especially over NA in the early 2000s, where *B* values are significantly low, below -150%. This may have been induced by the high evaporation values during summer, combined with the low precipitation of NARR in this region, which includes the US-Canadian border where precipitation from NARR is strongly underestimated (section I.3.1). Despite the highest evaporation maxima from MERRA, this reanalysis shows closed water balance with *B* values around 0% over the five climatic regions, as expected (section I.2.1). Regarding the ETA and ETU regions, ERA-Interim reveals a relatively closed water balance with *B* values close to 0%, compared with NARR, and particularly CFSR. However, the higher runoff values from ERA-Interim over the EB, GL and NA regions tend to produce a negative relative water balance (about -15%) for most years. Globally, the reanalyses do not agree in terms of relative water balance, while some of their water cycle components present some similarities. On the other hand, although MERRA was designed to close its water balance, it also reveals some overestimation of different water cycle components (with respect to the other reanalyses, and with respect to NRCAN observed precipitation mostly in ETA and ETU regions).

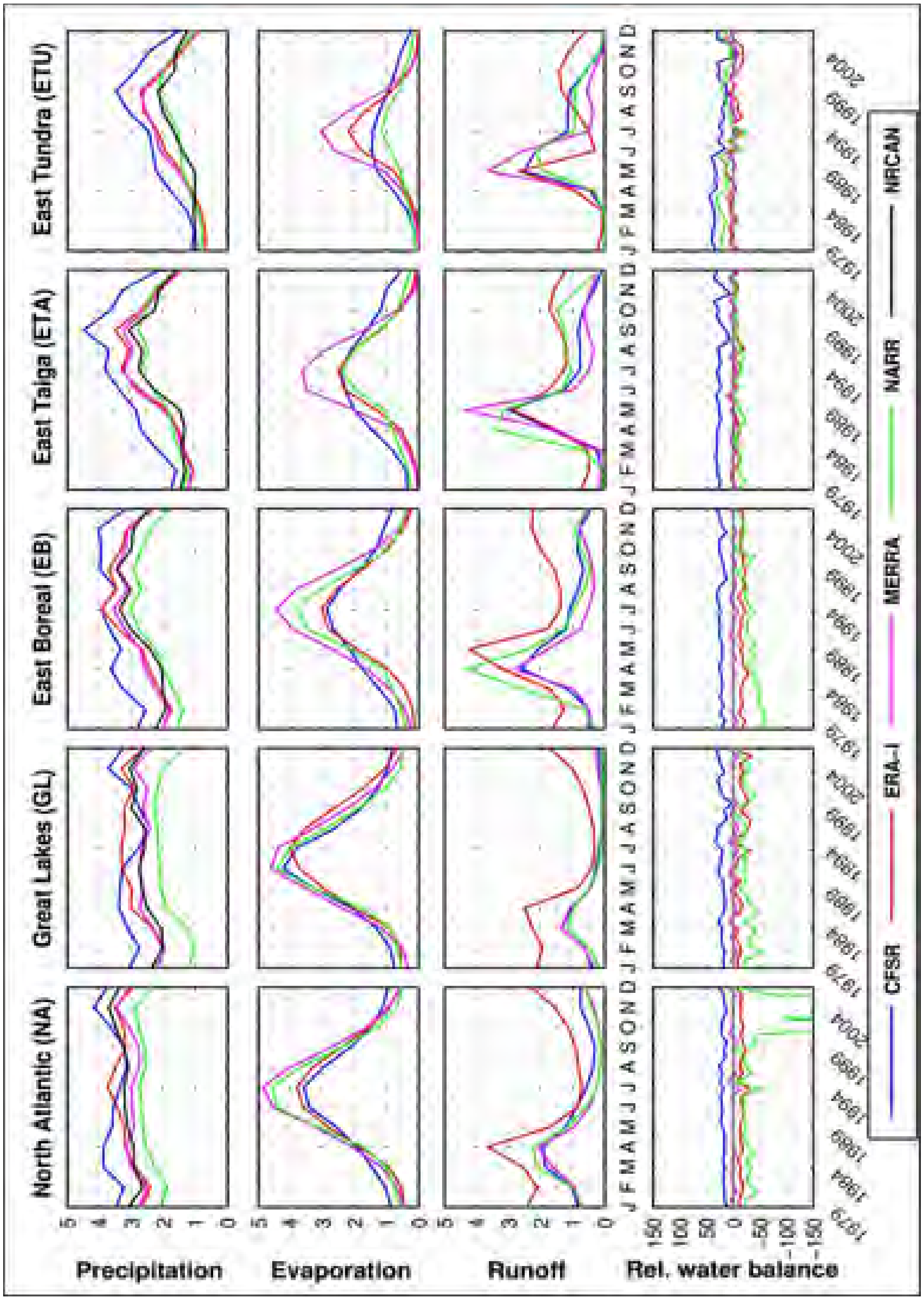

Figure-A I-5 Mean annual cycle of precipitation, evaporation and runoff (in mm day^{-1}), and annual mean relative water balance (in % of precip), from 1979 to 2008, of MERRA, ERA-Interim, NARR and CFSR, with the addition of NRCAN for the case of precipitation for the five climatic regions from Bukovsky (2011)

I.3.3 Water cycle components within river basins

Mean annual precipitation from the four reanalyses and NRCAN datasets were computed from 1979 to 2008 (figure-A I-6a) as spatial averages over the eleven selected river basins (RBs) (section I.2.3 and figure-A I-2b). As already seen in the previous sections, precipitation from CFSR is generally higher than that of the other datasets over the eleven river basins, including NRCAN. Among the four reanalyses, MERRA and ERA-Interim seem to be the less biased with respect to the NRCAN observations. On the other hand, NARR presents significant underestimations, as compared to ERA-Interim and MERRA, especially over RB 1 and RB 6 (1.5 to 2 mm day^{-1} less). Furthermore, a sudden change in precipitation from NARR around 2003 is clearly noticeable over RBs 1 to 6 (which are respectively located in the GL climatic region for RB 1, in the EB region for RB 2, 3 and 4, and in the NA region for RBs 5 and 6). Figure-A I-6b shows the distribution of the yearly mean annual precipitation from the four reanalyses and NRCAN datasets over the river basins. Globally median values from the five datasets differ more from one another over the southern river basins than over the northern ones. One the other hand, the dispersion of annual values from the five datasets is quite similar. The sudden change in precipitation from NARR around 2003 is also noticeable in figure-A I-6b over the affected river basins (outlier values).

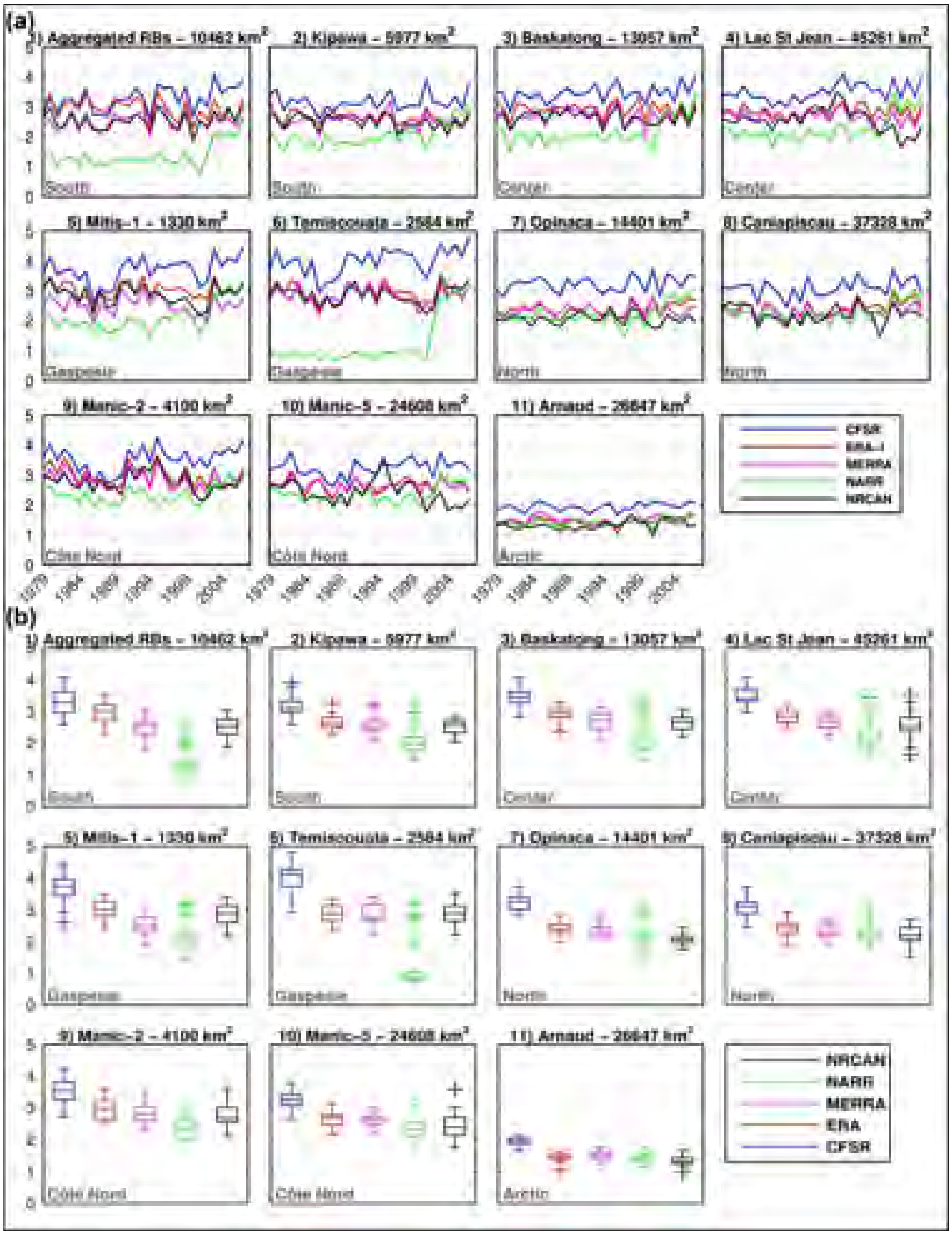

Figure-A I-6 (a) Mean annual precipitation (mm day^{-1}) from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the eleven selected river basins. For each river basin, the total area is indicated in square kilometers, and the hydrological region is indicated on the bottom left corner. (b) Distribution of the yearly mean annual precipitation (mm day^{-1}) from CFSR, ERA-Interim, MERRA, NARR and NRCAN, from 1979 to 2008 over the eleven selected river basins. On each box, the central mark is the median, the edges of the box are the 25th (p25) and 75th (p75) percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. Points are drawn as outliers if they are larger than p75 + 1.5(p75 – p25) or smaller than p25 – 1.5(p75 – p25)

Looking at figure-A I-7, the spatial distribution of the mean precipitation from NARR is quite different between the 1979-2002 and 2003-2008 periods. Indeed, the bias towards low precipitation values over the US-Canadian border between 1979 and 2002 vanished between 2003 and 2008. In southern Quebec, there is significantly more precipitation after 2003 than before. This could likely be related to the update of the NARR assimilation system in April 2003, and to the change in the number and nature of assimilated observations (Mesinger 2004).

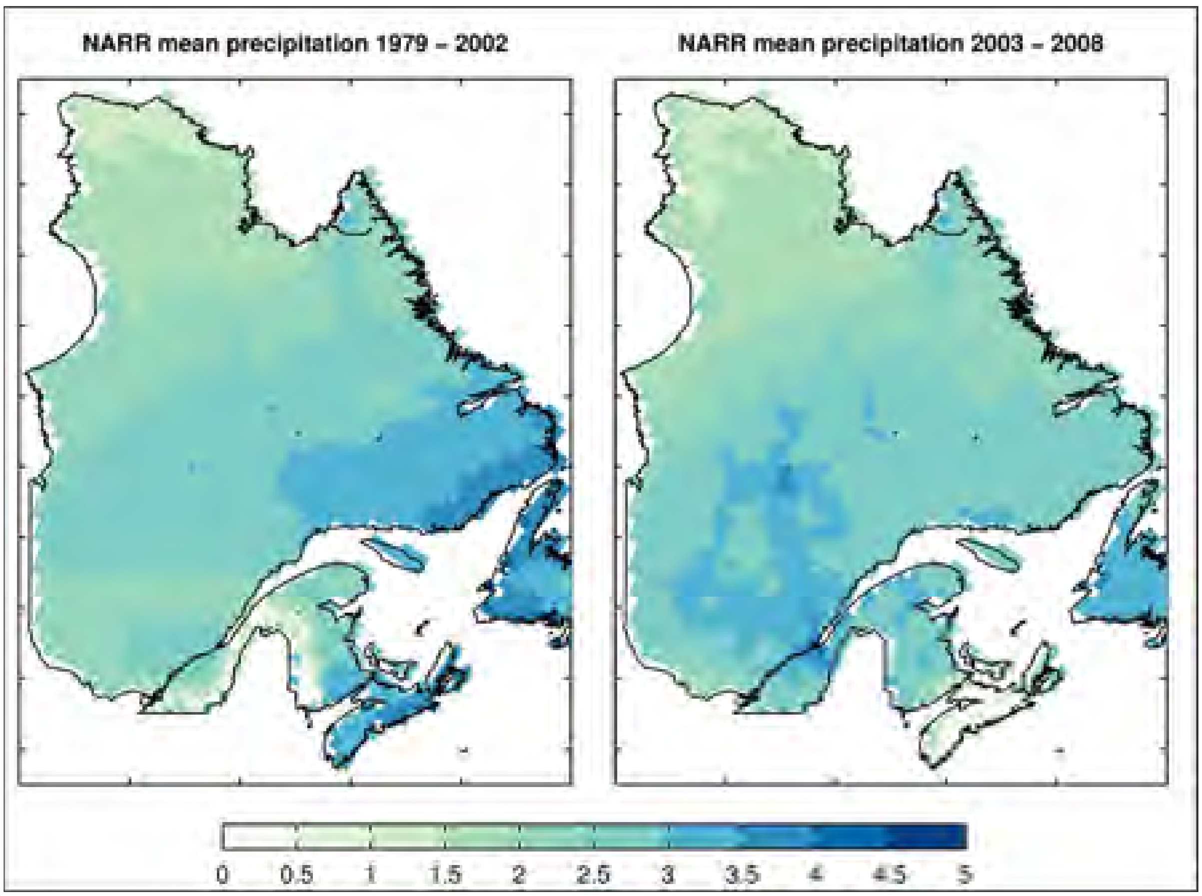

Figure-A I-7 Mean precipitation of NARR (mm day^{-1}) for the periods 1979-2002 (left) and 2003-2008 (right), over the province of Quebec

Table-A I-3 shows the correlation between daily time series of the spatial averaged precipitation from each of the four reanalyses and NRCAN over the selected river basins (section I.2.3). Despite the global overestimation of precipitation from CFSR, the correlations over all river basins are similar to those from MERRA and ERA-Interim. On the other hand, NARR presents low or inconsistent correlations, compared with MERRA, ERA-Interim and CFSR from -0.33 to 0.62. The negative value of NARR over RB 10 (-0.13) is not surprising when considering the low correlations of the other three datasets. However, the negative correlation value from NARR over RB 4 (-0.33) is probably related to the sudden shift in precipitation happening around 2003, discussed above (figure-A I-7). Nevertheless, these results must also be taken with caution since the quality of the precipitation from NRCAN is questionable in the northern part of the province, as the gridded data have been interpolated from very few weather stations (section I.2.2). For further analysis, it could be relevant to perform a direct comparison between reanalysis grid points and measurements from the closest weather station in order to avoid biases induced by the generation of gridded observation datasets. In any case, a correlation between daily time series around 0.8 or more should be acceptable to indicate that reanalyses represent the climate over the river basins of interest quite well, especially in southern Quebec, and therefore, may be useful for hydrological modelling purpose.

Table-A I-3 Correlation between daily time series of spatially averaged precipitation of CFSR, ERA-Interim, MERRA and NARR versus NRCAN over the 11 selected river basins

	Hydrologic region	**CFSR**	**ERA-Interim**	**MERRA**	**NARR**
(1) Aggregated RBs	South	0.78	0.78	0.81	0.62
(2) Kipawa	South	0.58	0.72	0.73	0.13
(3) Baskatong	Center	0.69	0.57	0.81	0.42
(4) Lac St-Jean	Center	0.34	0.45	0.62	-0.33
(5) Mitis-1	Gaspésie	0.75	0.80	0.79	0.41
(6) Temiscouata	Gaspésie	0.74	0.81	0.74	0.50
(7) Opinaca	North	0.36	0.26	0.16	0.09
(8) Caniapiscau	North	0.65	0.56	0.59	0.41
(9) Manic-2	Côte Nord	0.60	0.67	0.71	0.33
(10) Manic-5	Côte Nord	0.31	0.15	0.23	-0.13
(11) Arnaud	Arctic	0.14	0.13	0.04	0.37
Averaged correlation by reanalysis		0.54	0.53	0.57	0.26

Figure-A I-8 shows the distribution of the daily precipitation intensities from CFSR, ERA-Interim, MERRA, NARR and NRCAN, over the eleven river basins, for the period 1979-2008. Some of the main characteristics of the precipitation pattern over the province are well depicted in figure-A I-8. For instance, there are more dry days in summer than in winter, precipitation events are lower in winter, whereas the most extreme events generally happen during summer. On the other hand, reanalysis datasets underestimate the percentage of dry days compared with the NRCAN dataset. This remark is not valid for NARR in southern Quebec (RBs 1 to 6), as this dataset tends to underestimate the precipitation, compared with the four other datasets (figure-A I-5, figure-A I-6 and figure-A I-7). Moreover, CFSR shows the lowest percentage of dry days among the four reanalyses in winter. However, over the northern river basins in winter, NRCAN presents less dry days than the four reanalyses. One has to keep in mind that, in this part of the province, the density of weather stations (from which NRCAN has been generated) is very low (figure-A I-1b). In summer, NARR, which assimilates precipitation during the analysis process, shows quite similar distributions of precipitation, compared with NRCAN, except over RBs 1, 5, 6 and 9. Globally, among the four reanalyses, CFSR presents the highest percentage of the contribution of extreme events (16 to 128 mm day^{-1}) to the total amount of precipitation during winter and summer.

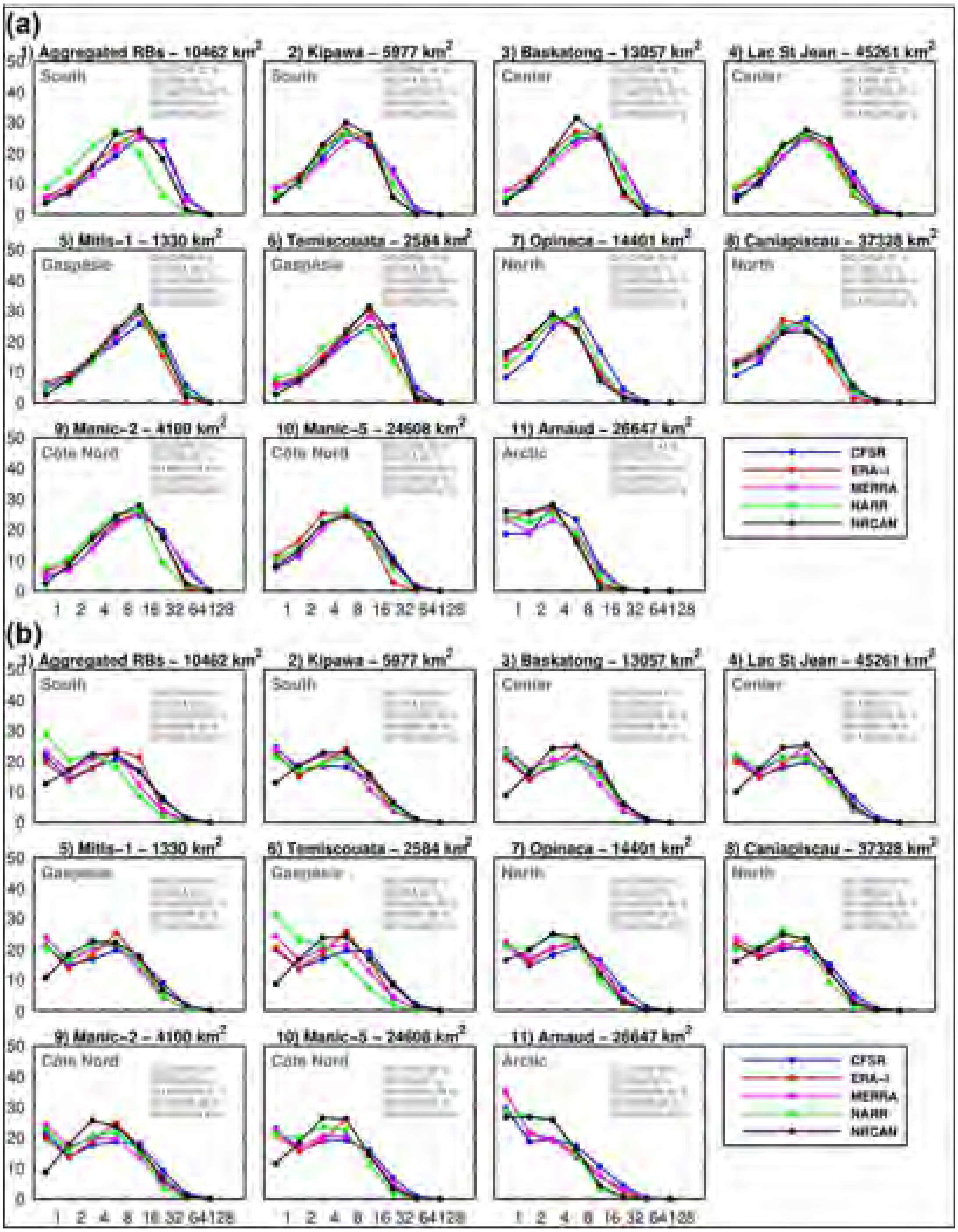

Figure-A I-8 Distribution of the daily precipitation intensities from CFSR, ERA-Interim, MERRA, NARR and NRCAN over the eleven river basins for the period 1979-2008 for (a) winter (DJF), and (b) summer (JJA). A threshold of 0.25 mm day^{-1} has been applied to dissociate dry days from wet days. The percentage of dry days (DD) is indicated into each subfigure for the five datasets. The y axis shows the contribution (%) of the daily precipitation of each bin to the total amount for the season of interest, while the x axis shows the precipitation intensity of each bin (mm day^{-1})

Figure-A I-9a shows the mean annual runoff from CFSR, ERA-Interim, MERRA, NARR and the runoff estimates from $(cQ)^2$ over the 11 river basins. The streamflow records for the 11 selected river basins do not completely cover the 1979-2008 period, so the runoff calculation periods differ among the river basins. Despite the high values of runoff from ERA-Interim over GL, NA, and EB climatic regions (figure-A I-3), this reanalysis shows runoff values similar to the runoff estimates at the watershed scale, with greater differences over RBs 3,4 and 9. On the other hand, CFSR and MERRA generally reveal comparable amounts of runoff water (from 0.5 to 1 mm $.day^{-1}$) which are systematically lower than those from ERA-Interim and the runoff estimates (about 1 to 1.5 mm day^{-1} less). Over RB 10, CFSR and NARR show high runoff values (3 and 5.5 mm day^{-1} respectively) on the period 1979 – 1984 (figure-A I-9a). This river basin encompasses the Manicouagan Lake, where high local values of runoff were also highlighted on Fig. 3 for the case of NARR and CFSR. As this result is common to these two reanalyses, one can assume that their common surface model Noah may be involved in the production of those high runoff values. Figure-A I-9b shows the distribution of the yearly mean annual runoff from the four reanalyses and the runoff estimates over the river basins. In agreement with the previous results, runoff median values of the four reanalyses differ from the one of NRCAN over all river basins, except for ERA-Interim. The dispersion of annual values from CFSR, ERA-Interim and MERRA is relatively similar to the one of NRCAN, except for RBs 1, 5, 6 and 10 for CFSR, RBs 1, 5 and 6 for MERRA and RBs 3, 5 and 8 for ERA-Interim. Regarding NARR, the dispersion of annual values is systematically different from the one of NRCAN, except for RB 11.

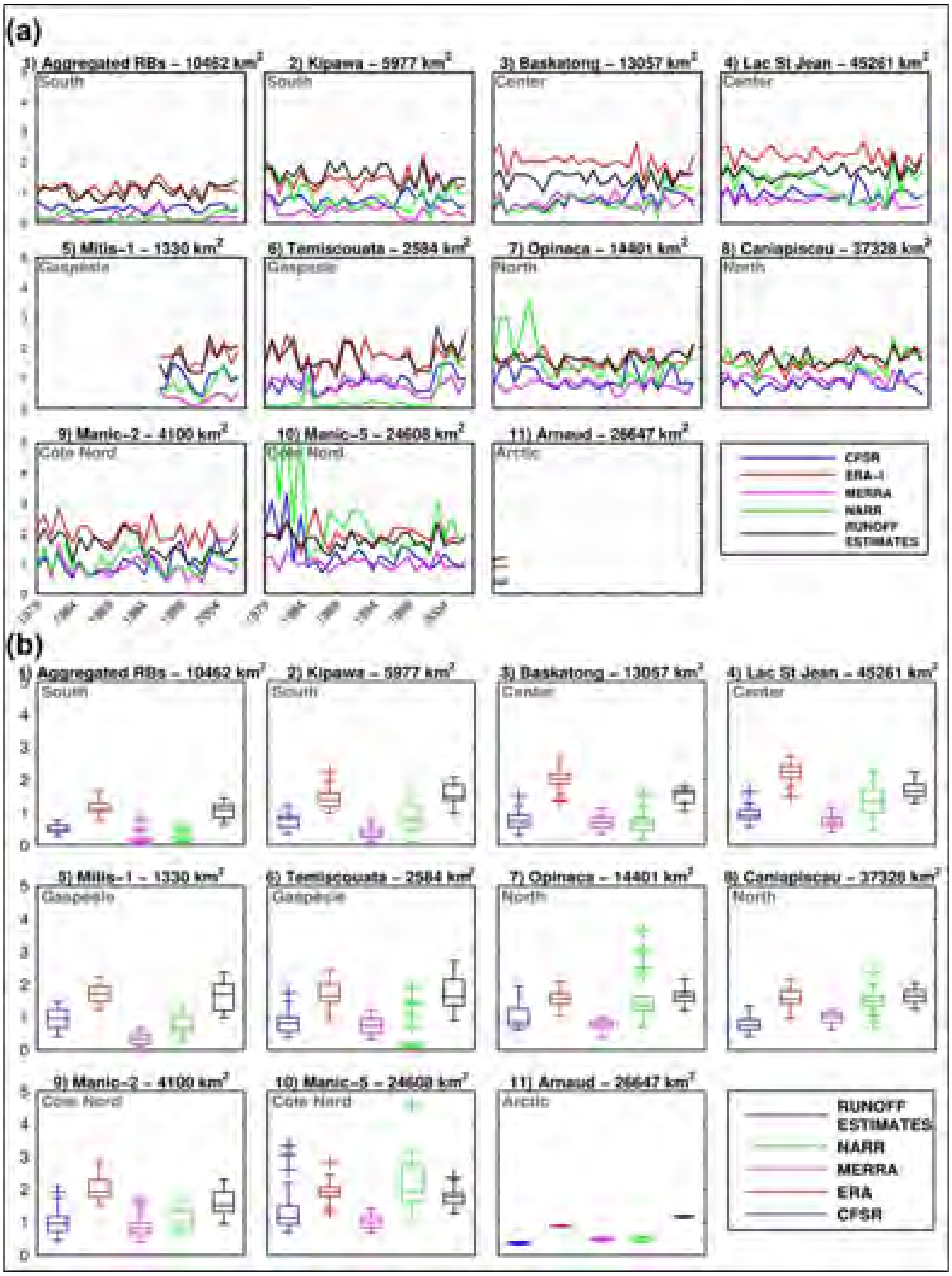

Figure-A I-9 (a) Mean annual runoff (mm day^{-1}) from CFSR, ERA-Interim, MERRA, NARR and runoff estimates, from 1979 to 2008 over the eleven selected river basins. Annual time series are limited to the available streamflow observations used to compute the runoff estimates. (b) Distribution of the yearly mean annual runoff (mm day^{-1}) from CFSR, ERA-Interim, MERRA, NARR and runoff estimates, from 1979 to 2008 over the eleven selected river basins. The characteristics of the box plots are described in the caption of figure-A I-6

Figure-A I-10a shows the mean annual *E/P* ratio from CFSR, ERA-Interim, MERRA, NARR and the observational estimates over the eleven river basins. As for the runoff results, the *E/P* calculation periods differ among the river basins. Considering the quality of the precipitation from NRCAN and of some $(cQ)^2$ streamflow time series, one can expect the observational estimates to sometimes be questionable (section I.2.2). Therefore, the observational estimates should not always be considered as the absolute truth, but rather, as another dataset for the purpose of analysis. Over southern and central Quebec, evaporation is expected to be about a half of precipitation, which means an *E/P* ratio around 0.5 (Fisheries and Canada 1978; Wang et al. 2014). Looking at figure-A I-10, the observational estimates ratio approaches 0.5 over RBs 1 and 3, and moderately underestimates 0.5 over RB 2. Looking more closely at RBs 4 and 10, the observational estimates ratio drops in the early 2000s (figure-A I-10a). This drop is also noticeable in figure-A I-6a, where, starting in 2004, the precipitation from NRCAN shows a slight downward trend, related to the quality issue of this precipitation dataset (discussed in section I.2.2) over Lac St-Jean and west of Côte Nord region, where RBs 4 and 10 are respectively located. For all river basins, MERRA and NARR show the highest *E/P* ratio values, especially NARR, which even exceeds 1 over the South, Center and Gaspésie river basins, which is not physically consistent. The high ratio values from NARR over RB 6 (and to a smaller extent, RBs 1 and 2) are directly related to the underestimation of precipitation close to the US-Canadian border, which extends through the river basins of Gaspésie (figure-A I-2b and figure-A I-7). Ratio values from MERRA tend to (and even sometimes exceed) 1 for the river basins in the South, Gaspésie and Center regions. This means that almost all the precipitation evaporates, which implies that there is no runoff in these regions for MERRA. This particularity is also illustrated in figure-A I-5 where MERRA shows the lowest runoff and the highest evaporation over the southern climatic regions in summer. Rienecker et al. (2011) highlighted deficiencies in MERRA that lead to the immediate evaporation of much of the rainfall, and consequently limit the surface runoff. Regarding ERA-Interim and CFSR, the *E/P* ratios are quite consistent with the expected value of 0.5 over the South and Center regions, except for some higher values in RB 1 between 1994 and 2008. Globally, the *E/P* ratios of the reanalyses show a more steady evolution during the 30 years over the North and Côte Nord river basins versus in the South, Gaspésie and Center river basins. However, the observational estimates ratios are most often lower than for the four reanalyses in the northern regions. In fact, the general limitations (discussed in section I.2.2) of the observed streamflow and the precipitation time series reduce the credibility of the evaporation observational estimates in the northern part of the province. Over RB 11 in the Arctic region, only three years of streamflow data were available to compute the *E/P* ratio. Nevertheless, the results show that the general behaviours of the reanalyses and the observational estimates appear to be similar to those over the other northern river basins. Figure-A I-10b shows the distribution of the yearly mean annual ratio from the four reanalyses and the observational estimates over the river basins. Dispersion of annual values from NRCAN is greater than the ones from CFSR, ERA-Interim and MERRA over the Center, Gaspésie, Côte Nord and North regions. In addition to the inconsistent values of NARR ratio, this reanalysis presents the largest dispersions among the five datasets over the South and Gaspésie regions.

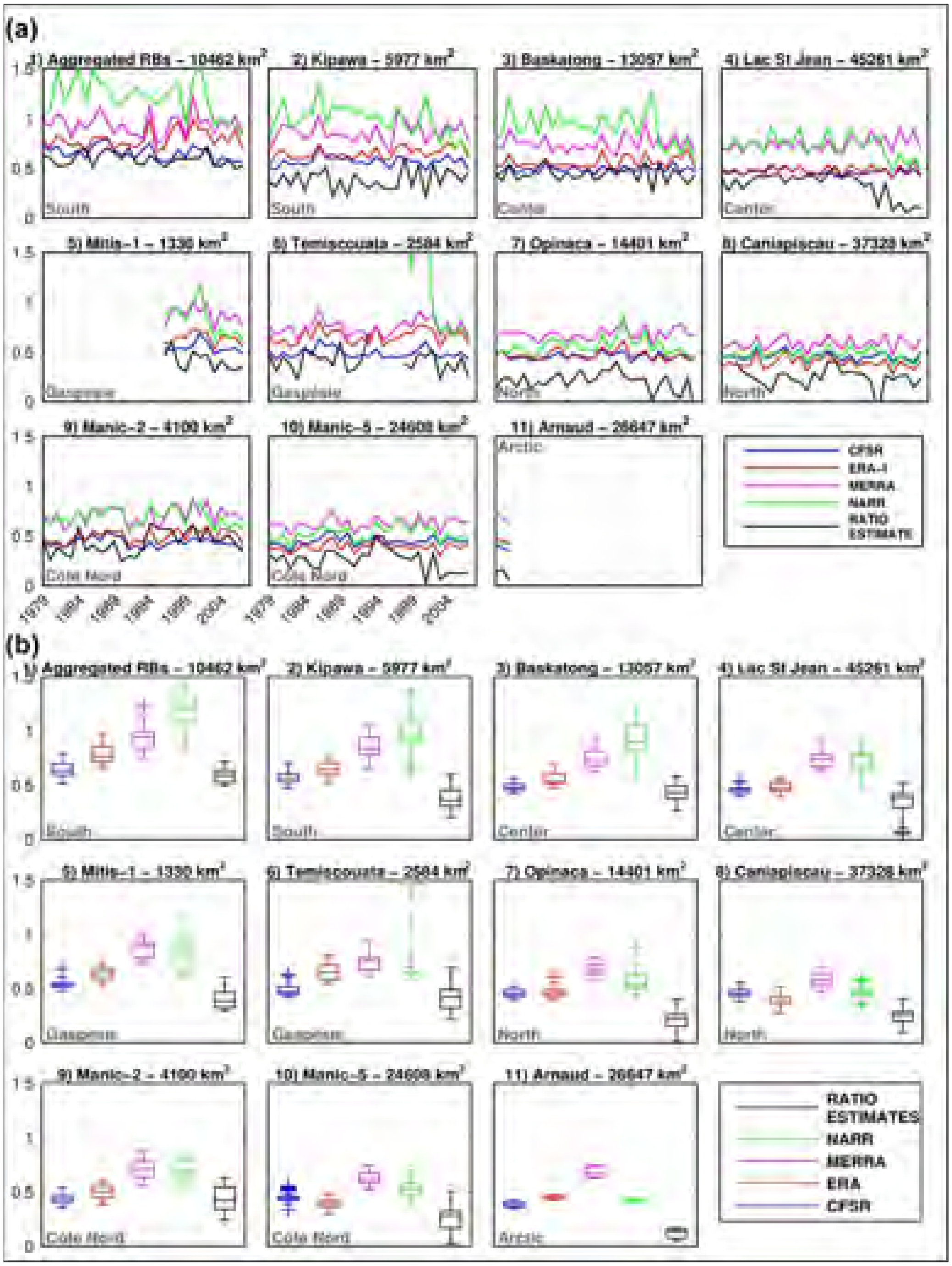

Figure-A I-10 Annual E/P ratio (evaporation divided by precipitation) from CFSR, ERA-Interim, MERRA, NARR and observational estimates, from 1979 to 2008 over the eleven selected river basins. Annual time series are limited to the available streamflow observations that were used to estimate E from observations. (b) Distribution of the yearly mean annual ratio (mm day^{-1}) from CFSR, ERA-Interim, MERRA, NARR and observational estimates, from 1979 to 2008 over the eleven selected river basins. The characteristics of the box plots are described in the caption of figure-A I-6

I.4 Concluding remarks

The scope of this study was to determine the reliability of the CFSR, MERRA, ERA-interim and NARR reanalyses in representing the terrestrial branch of the water cycle over the province of Quebec and to explore their potential in providing meteorological variables where few or no observations are available. The long-term mean of the water balance and the water cycle components were investigated for the four reanalyses. A first look at the water cycle components of the reanalyses showed that the water balance is not always maintained. Globally, the MERRA water balance is closed, as compared to the other three reanalyses, whereas NARR and ERA-Interim revealed negative water balances in central and southern Quebec, and CFSR revealed a positive water balance. The negative water balance of NARR is strongly influenced by inconsistencies in its components: for instance, underestimations of precipitation over the US-Canadian border and a large bias of runoff close to the Manicouagan Lake. Furthermore, the high precipitation from CFSR compared with those of MERRA, ERA-Interim and NARR, contributes to the positive values of the water balance over the entire province.

Precipitation, evaporation and runoff were also assessed on a multi-year monthly basis, over the Bukovsky climatic regions (Bukovsky 2011). For the four reanalyses, mean annual cycles of the water cycle components showed a similar temporal evolution. However, the amounts of water differed, especially the maxima of evaporation and runoff, and the precipitations from NARR and CFSR. A closer look at precipitation from the four reanalyses at the river basin scale confirmed that values from ERA-interim and MERRA agreed fairly well with one another and with those of NRCAN. On the other hand, CFSR and NARR almost systematically represented the highest and the lowest precipitation values, respectively, and consequently, do not appear to be particularly reliable over the province of Quebec, especially in the south.

Nevertheless, the temporal correlation of the daily precipitation from the reanalyses versus NRCAN over the river basins showed that precipitation from CFSR, ERA-Interim and

MERRA vary quite synchronously despite water amount disparities between them. Precipitation from NARR, which showed an overall poor daily correlation, should be taken with care, depending on the region of interest. Despite its highest runoff values among the reanalyses over southern climatic regions, ERA-Interim showed runoff values similar to runoff estimates over almost all studied river basins. Moreover, The *E/P* ratios revealed that ERA-Interim and CFSR succeeded relatively well in reproducing the distribution of precipitated water into evaporation and runoff, whereas NARR showed physically inconsistent results over southern Quebec. Regarding MERRA, although this dataset was designed to close the water balance, its surface model does not properly distribute an adequate amount of precipitated water into evaporation and runoff, since most of the precipitated water seems to evaporate before running off. Therefore, each of the four reanalyses has its own strengths and weaknesses. NARR, which has been shown to be quite reliable in the continental United States (Bukovsky and Karoly 2007; Mesinger 2004; Sheffield et al. 2012), appeared to be the least consistent dataset among the four reanalyses over eastern Canada. MERRA and ERA-Interim provide probably the most reliable precipitation over the province of Quebec, whereas evaporation remains questionable. The runoff values of ERA-Interim are the most consistent, compared with the observations, while runoff values from the other three reanalyses remain questionable.

No matter the case, for impact studies, care should be taken in using any of the reanalysis terrestrial water cycle components, for instance, in using evaporation and runoff as reference data in water resource management or as input or calibration data when performing hydrological modelling. In fact, it could be relevant to compare reanalysis datasets with weather station measurements (in the case of precipitation) to validate the ability of reanalyses to represent the weather, especially with the small number of weather stations available in remote regions (northern regions of Quebec). Such a task would involve accounting for the discrepancies between grid scales and point scales. A next step to this study will involve feeding hydrological models with reanalysis temperature and precipitation, and evaluating and comparing their potential for model calibration and validation. Moreover, scientific works are currently in progress concerning the use of hydrological variables from reanalyses, as the evaporation, as reference data in evaluating intermediate steps of simulation within the hydrological models, or introduced directly inside hydrological models in order to reduce the number of calibration parameters. Nevertheless, even though further investigations are necessary, these reanalyses have good potential to provide meteorological and hydrological information in remote areas of Canada. Moreover, some variables of these reanalyses succeed in revealing consistent values in regions where observational datasets are reliable.

I.5 Acknowledgements

We thank the National Centers for Environmental Prediction, the National Center for Atmospheric Research, the National Aeronautics and Space Administration and the European Centre for Medium-Range Weather Forecasts for developing and providing the reanalysis datasets. CFSR and NARR were obtained through the CISL Research Data Archive website at rda.ucar.edu/pub/cfsr.html and rda.ucar.edu/datasets/ds608.0, respectively. MERRA was downloaded from gmao.gsfc.nasa.gov/merra and ERA-Interim was downloaded from apps.ecmwf.int. We also thank Catherine Guay from the Institut de Recherche d'Hydro-Québec for providing the $(cQ)^2$ database and useful additional information about these data. We also would like to thank Blaise Gauvin St-Denis from Ouranos for providing valuable help in downloading and formatting reanalysis datasets, and finally, the organizations that funded this project, the Conseil de Recherches en Sciences Naturelles et en Génie du Canada, Hydro-Québec, Rio Tinto Alcan and Ontario Power Generation.

ANNEXE II

LISTE DES CONTRIBUTIONS SCIENTIFIQUES

1- Articles scientifiques – journaux avec jury (soumis / publiés)

Essou, G. R. C., Arsenault, R. et Brissette, F. P. (2016). Comparison of climate datasets for lumped hydrological modeling over the continental United States. *Journal of Hydrology,* 537, 334-345

Essou, G. R. C., Sabarly, F., Lucas-Picher, P., Brissette, F. et Poulin, A. (2016). Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling? *Journal of Hydrometeorology*, 17 (7).

Essou, G. R. C., Brissette, F. et Lucas-Picher, P. (2016). The use of reanalyses and gridded observations as weather input data for a hydrological model: comparison of performances of simulated river flows according to the density of weather stations. Journal of Hydrometeorology (Soumis en mars 2016).

Essou, G. R. C., Brissette, F. et Lucas-Picher, P. (2016). Impacts of combining reanalyses and weather station data on the accuracy of discharge modeling. Journal of Hydrology (Soumis en mars 2016).

Sabarly, F., **Essou, G. R. C.,** Lucas-Picher, P., Brissette, F. et Poulin, A. (2016). Use of Four Reanalysis Datasets to Assess the Terrestrial Branch of the Water Cycle over Quebec, Canada. *Journal of Hydrometeorology*, 17 (5), 1447-1466.

Essou, G. R. C. et Brissette, F. (2013). Climate Change Impacts on the Ouémé River, Benin, West Africa. *Journal of Earth Science and Climate Change*, 4, 161. doi:10.4172/2157-7617.1000161.

Chen, J., Chao, L., Sinha, E., Brissette, F. P., **Essou, G. R. C**. (2016). Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling. *Journal of Hydrology*. (Soumis en février 2016).

Arsenault, R., **Essou, G. R. C.** et Brissette, F. (2015). Improving hydrological model simulations using four gridded climate datasets in a multi-model averaging framework. *Journal of Hydrologic Engineering*. (Soumis en mai 2015).

2- Conférences internationales/nationales/provinciales – avec jury

Essou, G. R. C., Brissette, F., et Lucas-Picher, P. (2016). Comparaison des performances de simulations hydrologiques forcées par les réanalyses et les observations interpolées en fonction de la densité de stations météorologiques. 69e congrès de l'Association Canadienne des Ressources Hydriques : « L'eau à toutes les échelles : réduire la vulnérabilité et augmenter la résilience », Montréal, Canada, 25-27 mai 2016.

Essou, G. R. C., Sabarly, F., Lucas-Picher, P., Brissette, F. et Poulin, A. (2015). Peut-on utiliser les données de précipitation et de température provenant des réanalyses en modélisation hydrologique ? Colloque «la recherche hydrologique au Québec: État des lieux et perspectives», Montréal, Canada, 9-10 juin 2015.

Sabarly, F**., Essou, G. R. C.**, Poulin, A., Brissette, F. et Lucas-Picher, P. (2014). Réanalyses: étude des variables hydrologiques par région climatique en Amérique du Nord. 6e Symposium Scientifique d'Ouranos sur les changements climatiques, Québec, Canada, 4-5 décembre 2014.

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

Adam, J. C., and D. P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **108**.

Ajami, N. K., Q. Duan, X. Gao, and S. Sorooshian, 2006: Multimodel combination techniques for analysis of hydrological simulations: Application to distributed model intercomparison project results. *Journal of Hydrometeorology*, **7,** 755-768.

Alessandroni, M. G., and G. Remedia, 2002: The most severe floods of the Tiber River in Rome. *IAHS PUBLICATION***,** 129-132.

Ali, M., S. Ye, H.-y. Li, M. Huang, L. R. Leung, A. Fiori, and M. Sivapalan, 2014: Regionalization of subsurface stormflow parameters of hydrologic models: Up-scaling from physically based numerical simulations at hillslope scale. *Journal of Hydrology*, **519,** 683-698.

Alsdorf, D. E., and D. P. Lettenmaier, 2003: Tracking fresh water from space. *Science*, **301,** 1491-1494.

Andréassian, V., C. Perrin, C. Michel, I. Usart-Sanchez, and J. Lavabre, 2001: Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *Journal of Hydrology*, **250,** 206-223.

Andrieu, H., and J. D. Creutin, 1995: Identification of vertical profiles of radar reflectivity for hydrological applications using an inverse method. Part I: Formulation. *Journal of Applied Meteorology*, **34,** 225-239.

Archer, D., 1999: Practical application of historical flood information to flood estimation. *IAHS-AISH publication***,** 191-199.

Arsenault, R., and F. P. Brissette, 2014a: Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches. *Water Resources Research*, **50,** 6135-6153.

Arsenault, R., and F. Brissette, 2014b: Determining the optimal spatial distribution of weather station networks for hydrological modeling purposes using RCM datasets: An experimental approach. *Journal of Hydrometeorology*, **15,** 517-526.

Arsenault, R., G. R. C. Essou, and F. Brissette, 2015a: Improving hydrological model simulations using multiple gridded climate datasets in multi-model and multi-input averaging frameworks. *Journal of Hydrologic Engineering*, **Submitted**.

Arsenault, R., A. Poulin, P. Côté, and F. Brissette, 2014: Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration. *J. Hydrol., Eng*, **19,** 1374–1384.

Arsenault, R., R. Bazile, C. Ouellet-Dallaire, and F. Brissette, 2015b: CANOPEX : A canadian hydrometeorological watershed database. *Hydrological Processes*, **Submitted**.

Arsenault, R., F. Brissette, J.-S. Malo, M. Minville, and R. Leconte, 2013: Structural and non-structural climate change adaptation strategies for the Péribonka water resource system. *Water resources management*, **27,** 2075-2087.

Arsenault, R., P. Gatien, B. Renaud, F. Brissette, and J.-L. Martel, 2015c: A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. *Journal of Hydrology*, **529,** 754-767.

Artan, G., H. Gadain, J. L. Smith, K. Asante, C. J. Bandaragoda, and J. P. Verdin, 2007: Adequacy of satellite derived rainfall data for stream flow modeling. *Natural Hazards*, **43,** 167-185.

Avila, F. B., S. Dong, K. P. Menang, J. Rajczak, M. Renom, M. G. Donat, and L. V. Alexander, 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,** 6-16.

Bailey, W. G., T. R. Oke, and W. Rouse, 1997: *Surface Climates of Canada.* Vol. 4, McGill-Queen's Press-MQUP.

Baillargeon, S., J. Pouliot, L. Rivest, V. Fortin, and J. Fitzback, 2004: Interpolation statistique multivariable de données de précipitations dans un cadre de modélisation hydrologique. *les actes du colloque national Géomatique 2004 de l'Association canadienne des sciences géomatiques*.

Barrett, E. C., 1970: The estimation of monthly rainfall from satellite data. *Monthly weather review*, **98,** 322-327.

Bastola, S., and D. François, 2012: Temporal extension of meteorological records for hydrological modelling of Lake Chad Basin (Africa) using satellite rainfall data and reanalysis datasets. *Meteorological Applications*, **19,** 54-70.

Bengtsson, L., and J. Shukla, 1988: Integration of space and in situ observations to study global climate change. *Bulletin of the American Meteorological Society*, **69,** 1130-1143.

Benito, G., A. Díez-Herrero, and M. F. de Villalta, 2003: Magnitude and frequency of flooding in the Tagus basin (Central Spain) over the last millennium. *Climatic Change*, **58,** 171-192.

Berger, J.-F., P. Blanchemanche, C. Reynès, and P. Sabatier, 2010: Dynamiques fluviales en basse vallée du Vidourle au cours des six derniers siècles. Confrontation des données pédosédimentaires à haute résolution temporelle à l'analyse fréquentielle des crues historiques. *Quaternaire. Revue de l'Association française pour l'étude du Quaternaire*, **21,** 27-41.

Betts, A., and M. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and arctic air-mass data sets. *Quarterly Journal of the Royal Meteorological Society*, **112,** 693-709.

Biemans, H., R. Hutjes, P. Kabat, B. Strengers, D. Gerten, and S. Rost, 2009: Effects of precipitation uncertainty on discharge calculations for main river basins. *Journal of Hydrometeorology*, **10,** 1011-1025.

Bitew, M. M., M. Gebremichael, L. T. Ghebremichael, and Y. A. Bayissa, 2012: Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. *Journal of Hydrometeorology*, **13,** 338-350.

Bosilovich, M., 2008: NASA's modern era retrospective-analysis for research and applications: Integrating Earth observations. *Earthzine.[Online]. Retrieved on*, **26**.

Bosilovich, M. G., 2013: Regional climate and variability of NASA MERRA and recent reanalyses: US summertime precipitation and temperature. *Journal of Applied Meteorology and Climatology*, **52,** 1939-1951.

Bosilovich, M. G., F. R. Robertson, and J. Chen, 2011: Global energy and water budgets in MERRA. *Journal of Climate*, **24,** 5721-5739.

Bosilovich, M. G., J. Chen, F. R. Robertson, and R. F. Adler, 2008: Evaluation of global precipitation in reanalyses. *Journal of applied meteorology and climatology*, **47,** 2279-2299.

Bosilovich, M. G., D. Mocko, J. O. Roads, and A. Ruane, 2009: A multimodel analysis for the Coordinated Enhanced Observing Period (CEOP). *Journal of Hydrometeorology*, **10,** 912-934.

Brázdil, R., and M. Bukáček, 2000: Chronology of floods in the catchment area of the river Morava (the Czech Republic) since the 16th century.

Brázdil, R., Z. W. Kundzewicz, and G. Benito, 2006: Historical hydrology for studying flood risk in Europe. *Hydrological Sciences Journal*, **51,** 739-764.

Bukovsky, M. S., cited 2011: Masks for the Bukovsky regionalization of North America. [Available online at http://www.narccap.ucar.edu/contrib/bukovsky/.]

Bukovsky, M. S., and D. J. Karoly, 2007: A brief evaluation of precipitation from the North American Regional Reanalysis. *Journal of Hydrometeorology*, **8,** 837-846.

Buytaert, W., R. Celleri, P. Willems, B. De Bievre, and G. Wyseure, 2006: Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of hydrology*, **329,** 413-421.

Cavadias, G., and G. Morin, 1986: The combination of simulated discharges of hydrological models. *Hydrology Research*, **17,** 21-32.

Chaplot, V., A. Saleh, and D. Jaynes, 2005: Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO 3–N loads at the watershed level. *Journal of Hydrology*, **312,** 223-234.

Chen, D., L. Gong, C. Y. XU, and S. Halldin, 2007: A HIGH-RESOLUTION, GRIDDED DATASET FOR MONTHLY TEMPERATURE NORMALS (1971-2000) IN SWEDEN. *Geografiska Annaler: Series A, Physical Geography*, **89,** 249-261.

Chen, J., F. P. Brissette, and R. Leconte, 2011a: Uncertainty of downscaling method in quantifying the impact of climate change on hydrology. *Journal of Hydrology*, **401,** 190-202.

Chen, J., F. P. Brissette, and R. Leconte, 2012: Downscaling of weather generator parameters to quantify hydrological impacts of climate change. *Climate Research*, **51,** 185.

Chen, J., F. P. Brissette, A. Poulin, and R. Leconte, 2011b: Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resources Research*, **47**.

Cherry, J. E., L. Tremblay, M. Stieglitz, G. Gong, and S. Déry, 2007: Development of the pan-Arctic snowfall reconstruction: New land-based solid precipitation estimates for 1940-99. *Journal of Hydrometeorology*, **8,** 1243-1263.

Chiew, F., J. Teng, J. Vaze, and D. Kirono, 2009: Influence of global climate model selection on runoff impact assessment. *Journal of Hydrology*, **379,** 172-180.

Choi, H., P. Rasmussen, and V. Fortin, 2013: Evaluation and Comparison of Historical Gridded Data Sets of Precipitation for Canada. *AGU Fall Meeting Abstracts*, 1481.

Choi, W., S. J. Kim, P. F. Rasmussen, and A. R. Moore, 2009: Use of the North American Regional Reanalysis for hydrological modelling in Manitoba. *Canadian Water Resources Journal*, **34,** 17-36.

Chow, V. T., D. R. Maidment, and L. W. Mays, 1988: *Applied hydrology.*

Christensen, O. B., J. H. Christensen, B. Machenhauer, and M. Botzet, 1998: Very high-resolution regional climate simulations over Scandinavia-Present climate. *Journal of Climate*, **11,** 3204-3229.

Chvíla, B., B. Sevruk, and M. Ondráš, 2005: Intercomparison measurements of recording precipitation gauges in Slovakia. *WMO/CIMO Technical Conference*.

Cœur, D., 2003: La maîtrise des inondations dans la plaine de Grenoble (XVIIe-XXe siècle): enjeux techniques, politiques et urbains, Grenoble 2.

Cœur, D., M. Lang, and A. Paquier, 2002: L'historien, l'hydraulicien et l'hydrologue et la connaissance des inondations. *La Houille Blanche***,** 61-66.

Coeur, D., and M. Lang, 2000: L'information historique des inondations: l'histoire ne donne-t-elle que des leçons? *La Houille Blanche***,** 79-84.

Cole, S. J., and R. J. Moore, 2009: Distributed hydrological modelling using weather radar in gauged and ungauged basins. *Advances in water resources*, **32,** 1107-1120.

Coulibaly, P., J. Samuel, A. Pietroniro, and D. Harvey, 2013: Evaluation of Canadian National Hydrometric Network density based on WMO 2008 standards. *Canadian Water Resources Journal*, **38,** 159-167.

Cressman, G. P., 1959: An operational objective analysis system. *Monthly Weather Review*, **87,** 367-374.

Creutin, J., and C. Obled, 1982: Objective analyses and mapping techniques for rainfall fields: an objective comparison. *Water resources research*, **18,** 413-431.

Curtis, S., A. Salahuddin, R. F. Adler, G. J. Huffman, G. Gu, and Y. Hong, 2007: Precipitation extremes estimated by GPCP and TRMM: ENSO relationships. *Journal of Hydrometeorology*, **8,** 678-689.

D'souza, G., E. Barrett, and C. Power, 1990: Satellite rainfall estimation techniques using visible and infrared imagery. *Remote Sensing Reviews*, **4,** 379-414.

Daly, C., 2006: Guidelines for assessing the suitability of spatial climate data sets. *International journal of climatology*, **26,** 707-721.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of applied meteorology*, **33,** 140-158.

Daly, C., G. Taylor, and W. Gibson, 1997: The PRISM approach to mapping precipitation and temperature. *Proc., 10th AMS Conf. on Applied Climatology*, 20-23.

Dee, D., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137,** 553-597.

Degré, A., S. Ly, and C. Charles, 2013: Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. *Base*.

Delrieu, G., H. Andrieu, and J. D. Creutin, 2000: Quantification of path-integrated attenuation for X-and C-band weather radar systems operating in Mediterranean heavy rainfall. *Journal of Applied Meteorology*, **39,** 840-850.

Diederich, M., A. Ryzhkov, C. Simmer, P. Zhang, and S. Trömel, 2015: Use of specific attenuation for rainfall measurement at X-band radar wavelengths. Part II: Rainfall estimates and comparison with rain gauges. *Journal of Hydrometeorology*, **16,** 503-516.

Diks, C. G., and J. A. Vrugt, 2010: Comparison of point forecast accuracy of model averaging methods in hydrologic applications. *Stochastic Environmental Research and Risk Assessment*, **24,** 809-820.

Dirks, K., J. Hay, C. Stow, and D. Harris, 1998: High-resolution studies of rainfall on Norfolk Island: Part II: Interpolation of rainfall data. *Journal of Hydrology*, **208,** 187-193.

Donat, M. G., J. Sillmann, S. Wild, L. V. Alexander, T. Lippmann, and F. W. Zwiers, 2014: Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets*. *Journal of Climate*, **27,** 5019-5035.

Dribault, Y., 2012: Caractérisation de la dynamique saisonnière de l'hydrologie des tourbières minérotrophes du Moyen-Nord québecois, à l'aide de l'imagerie satellitaire multispectrale à très haute résolution spatiale, Université du Québec.

Duan, Q., and Coauthors, 2006: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, **320,** 3-17.

Duncan, M., B. Austin, F. Fabry, and G. Austin, 1993: The effect of gauge sampling density on the accuracy of streamflow prediction for rural catchments. *Journal of Hydrology*, **142,** 445-476.

Eleuch, S., A. Carsteanu, K. Bâ, R. Magagi, K. Goïta, and C. Diaz, 2010: Validation and use of rainfall radar data to simulate water flows in the Rio Escondido basin. *Stochastic Environmental research and risk assessment*, **24,** 559-565.

Elsner, M. M., S. Gangopadhyay, T. Pruitt, L. Brekke, N. Mizukami, and M. Clark, 2014: How does the Choice of Distributed Meteorological Data Affect Hydrologic Model Calibration and Streamflow Simulations? *Journal of Hydrometeorology*.

Engman, E. T., 1996: Remote sensing applications to hydrology: future impact. *Hydrological sciences journal*, **41,** 637-647.

Engman, E. T., 1999: Remote sensing in hydrology. *Assessment of Non-Point Source Pollution in the Vadose Zone***,** 165-177.

Engman, E. T., and R. J. Gurney, 1991: *Remote sensing in hydrology.* Chapman and Hall Ltd.

Essou, G., F. Brissette, and P. Lucas-Picher, 2016a: The use of reanalyses and gridded observations as weather input data for a hydrological model: comparison of performances of simulated river flows according to the density of weather stations. *Journal of Hydrometeorology*, **Submitted**.

Essou, G. R., F. Sabarly, P. Lucas-Picher, F. Brissette, and A. Poulin, 2016b: Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling? *Journal of Hydrometeorology*, **17**.

Faurès, J.-M., D. Goodrich, D. A. Woolhiser, and S. Sorooshian, 1995: Impact of small-scale spatial rainfall variability on runoff modeling. *Journal of Hydrology*, **173,** 309-326.

Fekete, B. M., C. J. Vörösmarty, J. O. Roads, and C. J. Willmott, 2004: Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate*, **17,** 294-304.

Fisheries, and E. Canada, cited 1978: Water balance - derived precipitation and evapotranspiration. [Available online at http://geogratis.gc.ca/api/en/nrcan-rncan/ess-sst/910100c0-4f8c-5ae8-ae87-69bf230e43cf.html.]

Forman, B. A., E. R. Vivoni, and S. A. Margulis, 2008: Evaluation of ensemble-based distributed hydrologic model response with disaggregated precipitation products. *Water Resources Research*, **44**.

Fortin, J.-P., R. Turcotte, S. Massicotte, R. Moussa, J. Fitzback, and J.-P. Villeneuve, 2001: Distributed watershed model compatible with remote sensing and GIS data. I: Description of model. *Journal of Hydrologic Engineering*, **6,** 91-99.

Fortin, V., 2000: Le modèle météo-apport HSAMI: historique, théorie et application. *Institut de recherche d'Hydro-Québec, Varennes*.

Fuka, D. R., M. T. Walter, C. MacAlister, A. T. Degaetano, T. S. Steenhuis, and Z. M. Easton, 2014: Using the Climate Forecast System Reanalysis as weather input data for watershed models. *Hydrological Processes*, **28,** 5613-5623.

Fuller, J. D., 2012: Alpine wind speed and blowing snow trend identification and analysis, Colorado State University.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Weather and Forecasting*, **13,** 377-395.

Gagnon, P., 2012: Désagrégation statistique de la précipitation mésoéchelle, Université du Québec, Institut national de la recherche scientifique.

Gallo, K., and G. Xian, 2014a: Application of spatially gridded temperature and land cover data sets for urban heat island analysis. *Urban Climate*.

Gallo, K., and G. Xian, 2014b: Application of spatially gridded temperature and land cover data sets for urban heat island analysis. *Urban Climate*, **8,** 1-10.

Garen, D. C., G. L. Johnson, and C. L. Hanson, 1994: Mean areal precipitation for daily hydrologic modeling in mountainous regions1. Wiley Online Library.

Gbambie, A. S. B., A. Poulin, and M.-A. Boucher, 2016: Added value of alternative information in interpolated precipitation datasets for hydrology, **Submitted**.

Gees, A., 1996: *Analyse historischer und seltener Hochwasser in der Schweiz: Bedeutung für das Bemessungshochwasser.* Geograph. Inst. d. Univ.

Gervais, M., J. R. Gyakum, E. Atallah, L. B. Tremblay, and R. B. Neale, 2014: How well are the distribution and extreme values of daily precipitation over North America represented in the Community Climate System Model? A comparison to reanalysis, satellite, and gridded station data. *Journal of Climate*, **27,** 5219-5239.

Gibson, J., P. Kallberg, S. Uppala, A. Hernandez, A. Nomura, and E. Serrano, 1997: ERA description. ECMWF Reanalysis Project Report Series 1, European Centre for Medium Range Weather Forecasts. *Reading, UK*, **66**.

Glade, T., and P. Albini, 2001: *The use of historical data in natural hazard assessments.* Vol. 17, Springer Science & Business Media.

Glaser, R., and H. Stangl, 2004: Climate and floods in Central Europe since AD 1000: data, methods, results and consequences. *Surveys in Geophysics*, **25,** 485-510.

Gleason, C. J., and L. C. Smith, 2014: Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *Proceedings of the National Academy of Sciences*, **111,** 4788-4791.

Goodison, B., H. Ferguson, and G. McKay, 1981: *Measurement and data analysis.* Vol. 776, Pergamon Press, Ontario.

Goodison, B., P. Louie, and D. Yang, 1998: WMO solid precipitation intercomparison. Final Report. *World Meteorol. Org., Instruments and Observing Methods Rep*, **67**.

Goovaerts, P., 2000: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of hydrology*, **228,** 113-129.

Gottschalck, J., J. Meng, M. Rodell, and P. Houser, 2005: Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. *Journal of Hydrometeorology*, **6,** 573-598.

Granger, C. W., and R. Ramanathan, 1984: Improved methods of combining forecasts. *Journal of Forecasting*, **3,** 197-204.

Grimes, D., E. Pardo-Iguzquiza, and R. Bonifacio, 1999: Optimal areal rainfall estimation using raingauges and satellite data. *Journal of hydrology*, **222,** 93-108.

Guidoboni, E., and E. Guidoboni, 1998: Human factors, extreme events and floods in the Lower Po Plain (Northern Italy) in the 16th century. *Environment and History***,** 279-308.

Gupta, H. V., S. Sorooshian, and P. O. Yapo, 1998: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, **34,** 751-763.

Hansen, N., and A. Ostermeier, 1996: Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*, IEEE, 312-317.

Hansen, N., and A. Ostermeier, 2001: Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, **9,** 159-195.

Harris, I., P. Jones, T. Osborn, and D. Lister, 2014: Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International Journal of Climatology*, **34,** 623-642.

Hartkamp, A. D., K. De Beurs, A. Stein, and J. W. White, 1999: *Interpolation techniques for climate variables.* CIMMYT Mexico, DF.

Hasenauer, H., K. Merganicova, R. Petritsch, S. A. Pietsch, and P. E. Thornton, 2003: Validating daily climate interpolations over complex terrain in Austria. *Agricultural and Forest Meteorology*, **119,** 87-107.

Hay, L., R. Viger, and G. McCABE, 1998: Precipitation interpolation in mountainous regions using multiple linear regression. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, **248,** 33-38.

Haylock, M., N. Hofstra, A. Klein Tank, E. Klok, P. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **113**.

Herrera, S., J. M. Gutiérrez, R. Ancell, M. Pons, M. Frías, and J. Fernández, 2012: Development and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain02). *International Journal of Climatology*, **32,** 74-85.

Higgins, R., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. *NCEP/Climate Prediction Center Atlas*, **7**.

Higgins, R., V. Kousky, V. Silva, E. Becker, and P. Xie, 2010: Intercomparison of daily precipitation statistics over the United States in observations and in NCEP reanalysis products. *Journal of Climate*, **23,** 4637-4650.

Higgins, R. W., J. E. Janowiak, and Y.-P. Yao, 1996: *A gridded hourly precipitation data base for the United States (1963-1993).* US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, 2005: Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology*, **25,** 1965-1978.

Hingray, B., C. Picouet, and A. Musy, 2009: *Hydrologie: Une science pour l'ingénieur.* Vol. 2, PPUR presses polytechniques.

Hofstra, N., M. New, and C. McSweeney, 2010: The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. *Climate dynamics*, **35,** 841-858.

Hopkinson, R. F., D. W. McKenney, E. J. Milewska, M. F. Hutchinson, P. Papadopol, and L. A. Vincent, 2011: Impact of aligning climatological day on gridding daily maximum-minimum temperature and precipitation over Canada. *Journal of Applied Meteorology and Climatology*, **50,** 1654-1665.

Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich, 1998: Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, **34**.

Huiyi, Z., and J. Shaofeng, 2010: Uncertainty in the spatial interpolation of rainfall data. *Progress in Geography*, **23,** 34-42.

Hunter, S. M., 1996: WSR-88D radar rainfall estimation: Capabilities, limitations and potential improvements. *Natl. Wea. Dig*, **20,** 26-38.

Hutchinson, M., 1995: Interpolating mean rainfall using thin plate smoothing splines. *International journal of geographical information systems*, **9,** 385-403.

Hutchinson, M., 2004: ANUSPLIN Version 4.3. Centre for Resource and Environmental Studies, Australian National University.

Hutchinson, M. F., D. W. McKenney, K. Lawrence, J. H. Pedlar, R. F. Hopkinson, E. Milewska, and P. Papadopol, 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,** 725-741.

Jacobeit, J., R. Glaser, J. Luterbacher, and H. Wanner, 2003: Links between flood events in central Europe since AD 1500 and large-scale atmospheric circulation modes. *Geophysical Research Letters*, **30**.

Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly Weather Review*, **122,** 927-945.

Janowiak, J. E., A. Gruber, C. Kondragunta, R. E. Livezey, and G. J. Huffman, 1998: A comparison of the NCEP-NCAR reanalysis precipitation and the GPCP rain gauge-satellite combined dataset with observational error considerations. *Journal of Climate*, **11,** 2960-2979.

Johnson, G. L., and C. L. Hanson, 1995: Topographic and atmospheric influences on precipitation variability over a mountainous watershed. *Journal of Applied Meteorology*, **34,** 68-87.

Joss, J., and U. Germann, 2000: Solutions and problems when applying qualitative and quantitative information from weather radar. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **25,** 837-841.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR reanalysis 40-year project. *Bull. Am. Meteorol. Soc*, **77,** 437-471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. Hnilo, M. Fiorino, and G. Potter, 2002: Ncep-doe amip-ii reanalysis (r-2). *Bulletin of the American Meteorological Society*, **83,** 1631-1643.

Kidd, C., D. R. Kniveton, M. C. Todd, and T. J. Bellerby, 2003: Satellite rainfall estimation using combined passive microwave and infrared algorithms. *Journal of Hydrometeorology*, **4,** 1088-1104.

Kistler, R., and Coauthors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological society*, **82,** 247-267.

Kite, G., and A. Pietroniro, 1996: Remote sensing applications in hydrological modelling. *Hydrological Sciences Journal*, **41,** 563-591.

Kleidon, A., and S. Schymanski, 2008: Thermodynamics and optimality of the water budget on land: a review. *Geophysical Research Letters*, **35**.

Konan, B., M. Slivitzky, P. Gagnon, and A. N. Rousseau, 2010: Validation of the Meteorological Outputs of the Canadian Regional Climate Model Using a Kriging Method: Application to Southern Quebec. *Canadian Water Resources Journal*, **35,** 259-280.

Kopp, T., and R. Kiess, 1996: The air force global weather central snow analysis model. *CONFERENCE ON WEATHER ANALYSIS AND FORECASTING*, American Meteorological Society, 220-222.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 2006: World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, **15,** 259-263.

Kouame, K. F., A. M. Kouassi, B. T. M. N'GUESSAN, J. M. Kouao, T. Lasm, and M. B. Saley, 2013: Analysis of trends in the rainfall-runoff relation in the context of climate change: case of the N'zo-Sassandra watershed (Western Côte d'Ivoire). *International Journal of Innovation and Applied Studies*, **2,** 92-103.

Krajewski, W., and J. Smith, 2002: Radar hydrology: rainfall estimation. *Advances in water resources*, **25,** 1387-1394.

Krajewski, W. F., V. Lakshmi, K. P. Georgakakos, and S. C. Jain, 1991: A Monte Carlo study of rainfall sampling effect on a distributed catchment model. *Water resources research*, **27,** 119-128.

Kralisch, S., P. Krause, M. Fink, C. Fischer, and W. Flügel, 2007: Component based environmental modelling using the JAMS framework. *MODSIM 2007 International Congress on Modelling and Simulation*, 812-818.

Lang, M., D. Cœur, C. Lallement, R. Naulet, and G. Boudou, 1998: HISTORISQUE-Guiers(utilisation de l'information historique pour une meilleure définition du risque d'inondation).

Langlois, A., and Coauthors, 2009: Simulation of snow water equivalent (SWE) using thermodynamic snow models in québec, canada. *Journal of Hydrometeorology*, **10,** 1447-1463.

Lauri, H., T. Räsänen, and M. Kummu, 2014: Using reanalysis and remotely sensed temperature and precipitation data for hydrological modeling in monsoon climate: Mekong river case study. *Journal of Hydrometeorology*, **15,** 1532-1545.

Lee, G. W., A. W. Seed, and I. Zawadzki, 2007: Modeling the variability of drop size distributions in space and time. *Journal of applied meteorology and climatology*, **46,** 742-756.

Lenters, J. D., M. T. Coe, and J. A. Foley, 2000: Surface water balance of the continental United States, 1963–1995: Regional evaluation of a terrestrial biosphere model and the NCEP/NCAR reanalysis. *Journal of Geophysical Research: Atmospheres*, **105,** 22393-22425.

Li, C., E. Sinha, D. E. Horton, N. S. Diffenbaugh, and A. M. Michalak, 2014: Joint bias correction of temperature and precipitation in climate model simulations. *Journal of Geophysical Research: Atmospheres*, **119,** 13,153-113,162.

Lindsay, R., M. Wensnahan, A. Schweiger, and J. Zhang, 2014: Evaluation of seven different atmospheric reanalysis products in the Arctic*. *Journal of Climate*, **27,** 2588-2606.

Llasat, M.-C., M. Barriendos, A. Barrera, and T. Rigo, 2005: Floods in Catalonia (NE Spain) since the 14th century. Climatological and meteorological aspects from historical documentary sources and old instrumental records. *Journal of hydrology*, **313,** 32-47.

Lopes, V. L., 1996: On the effect of uncertainty in spatial distribution of rainfall on catchment modelling. *Catena*, **28,** 107-119.

Lorenz, C., and H. Kunstmann, 2012: The hydrological cycle in three state-of-the-art reanalyses: intercomparison and performance analysis. *Journal of Hydrometeorology*, **13,** 1397-1420.

Mahdian, M., S. R. Bandarabady, R. Sokouti, and Y. N. Banis, 2009: Appraisal of the geostatistical methods to estimate monthly and annual temperature. *Journal of Applied Sciences*, **9,** 128-134.

Manzanas, R., L. Amekudzi, K. Preko, S. Herrera, and J. M. Gutiérrez, 2014: Precipitation variability and trends in Ghana: An intercomparison of observational and reanalysis products. *Climatic change*, **124,** 805-819.

Mareuil, A., R. Leconte, F. Brissette, and M. Minville, 2007: Impacts of climate change on the frequency and severity of floods in the Châteauguay River basin, Canada. *Canadian journal of civil engineering*, **34,** 1048-1060.

Marshall, G. J., 2003: Trends in the Southern Annular Mode from observations and reanalyses. *Journal of Climate*, **16,** 4134-4143.

Maurer, E., A. Wood, J. Adam, D. Lettenmaier, and B. Nijssen, 2002: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States*. *Journal of climate*, **15,** 3237-3251.

Maurer, E. P., G. M. O'Donnell, D. P. Lettenmaier, and J. O. Roads, 2001: Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model. *Journal of Geophysical Research: Atmospheres*, **106,** 17841-17862.

McEvoy, D. J., J. F. Mejia, and J. L. Huntington, 2014: Use of an Observation Network in the Great Basin to Evaluate Gridded Climate Data. *Journal of Hydrometeorology*, **15,** 1913-1931.

McKee, J. L., and A. D. Binns, 2015: A review of gauge–radar merging methods for quantitative precipitation estimation in hydrology. *Canadian Water Resources Journal/Revue canadienne des ressources hydriques***,** 1-18.

Meng, J., R. Yang, H. Wei, M. Ek, G. Gayno, P. Xie, and K. Mitchell, 2012: The land surface analysis in the NCEP Climate Forecast System Reanalysis. *Journal of Hydrometeorology*, **13,** 1621-1630.

Merz, B., and A. Bárdossy, 1998: Effects of spatial variability on the rainfall runoff process in a small loess catchment. *Journal of Hydrology*, **212,** 304-317.

Mesinger, F., 2004: Coauthors, 2004: NCEP North American Regional Reanalysis. *Preprints, 15th Symp. on Global Change and Climate Variations, Seattle, WA, Amer. Meteor. Soc. P.*

Mesinger, F., and Coauthors, 2006: North American regional reanalysis. *Bulletin of the American Meteorological Society*, **87,** 343-360.

Minville, M., F. Brissette, and R. Leconte, 2008: Uncertainty of the impact of climate change on the hydrology of a nordic watershed. *Journal of hydrology*, **358,** 70-83.

Minville, M., F. Brissette, and R. Leconte, 2009: Impacts and uncertainty of climate change on water resource management of the Peribonka river system (Canada). *Journal of Water Resources Planning and Management*, **136,** 376-385.

Minville, M., D. Cartier, C. Guay, L. A. Leclaire, C. Audet, S. Le Digabel, and J. Merleau, 2014: Improving process representation in conceptual hydrological model calibration using climate simulations. *Water Resources Research*, **50,** 5044-5073.

Mizukami, N., and M. B. Smith, 2012: Analysis of inconsistencies in multi-year gridded quantitative precipitation estimate over complex terrain and its impact on hydrologic modeling. *Journal of Hydrology*, **428,** 129-141.

Moriasi, D., J. Arnold, M. Van Liew, R. Bingner, R. Harmel, and T. Veith, 2007: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. Asabe*, **50,** 885-900.

Moulin, L., E. Gaume, and C. Obled, 2009: Uncertainties on mean areal precipitation: assessment and impact on streamflow simulations. *Hydrology and Earth System Sciences*, **13,** 99-114.

Mudelsee, M., M. Börngen, G. Tetzlaff, and U. Grünewald, 2004: Extreme floods in central Europe over the past 500 years: Role of cyclone pathway "Zugstrasse Vb". *Journal of Geophysical Research: Atmospheres (1984–2012)*, **109**.

Muñoz, E., C. Álvarez, M. Billib, J. L. Arumí, and D. Rivera, 2011: Comparison of gridded and measured rainfall data for basin scale hydrological studies. *Chilean Journal of Agricultural Research*, **71,** 459-468.

Murray, S., I. Watson, and I. Prentice, 2013: The Use of dynamic global vegetation models for simulating hydrology and the potential integration of satellite observations| Macquarie University ResearchOnline.

Nakama, L. Y., and J. C. Risley, 1993: Use of a rainfall-runoff model for simulating effects of forest management on streamflow in the east Fork Lobster Creek Basin, Oregon. Water resources investigation.

Nash, J., and J. Sutcliffe, 1970: River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology*, **10,** 282-290.

Naulet, R., 2002: Utilisation de l'information des crues historiques pour une meilleure prédétermination du risque d'innondation. Application au bassin de l'Ardèche à Vallon Pont-d'Arc et St-Martin d'Ardèche, Université du Québec.

Naulet, R., M. Lang, D. Coeur, and C. Gigon, 2001: Collaboration between historians and hydrologists on the Ardeche river (France). *The Use of Historical Data in Natural Hazard Assessments*, Springer, 113-129.

Naulet, R., M. Lang, T. B. Ouarda, D. Coeur, B. Bobée, A. Recking, and D. Moussay, 2005: Flood frequency analysis on the Ardeche river using French documentary sources from the last two centuries. *Journal of Hydrology*, **313,** 58-78.

Neiman, P. J., F. M. Ralph, B. J. Moore, and R. J. Zamora, 2014: The Regional Influence of an Intense Sierra Barrier Jet and Landfalling Atmospheric River on Orographic Precipitation in Northern California: A Case Study. *Journal of Hydrometeorology*.

New, M., M. Hulme, and P. Jones, 1999: Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology. *Journal of climate*, **12,** 829-856.

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. *Journal of Climate*, **13,** 2217-2238.

Ngo-Duc, T., J. Polcher, and K. Laval, 2005: A 53-year forcing data set for land surface models. *Journal of Geophysical Research: Atmospheres*, **110**.

Nigam, S., and A. Ruiz-Barradas, 2006: Seasonal hydroclimate variability over North America in global and regional reanalyses and AMIP simulations: Varied representation. *Journal of Climate*, **19,** 815-837.

Nystuen, J. A., 1999: Relative performance of automatic rain gauges under different rainfall conditions. *Journal of Atmospheric and Oceanic Technology*, **16,** 1025-1043.

Obled, C., J. Wendling, and K. Beven, 1994: The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of hydrology*, **159,** 305-333.

Onogi, K., and Coauthors, 2007: The JRA-25 Reanalysis. *Journal of the Meteorological Society of Japan*, **85,** 369-432.

Ouarda, T., P. Rasmussen, B. Bobée, and J. Bernier, 1998: Utilisation de l'information historique en analyse hydrologique fréquentielle. *Revue des sciences de l'eau/Journal of Water Science*, **11,** 41-49.

Pai, D., L. Sridhar, M. Rajeevan, O. Sreejith, N. Satbhai, and B. Mukhopadhyay, 2014: Development of a new high spatial resolution (0.25× 0.25) long period (1901–2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, **65,** 1-18.

Pandey, R. K., 2013: Étude des bassins fluviaux en Inde par télédétection, Université de Toulouse, Université Toulouse III-Paul Sabatier.

Pappenberger, F., and R. Buizza, 2009: The skill of ECMWF precipitation and temperature predictions in the Danube basin as forcings of hydrological models. *Weather and Forecasting*, **24,** 749-766.

Parent du Châtelet, J., P. Tabary, and P. Lamargue, 2005: Evolution du réseau radar opérationnel de Météo-France pour une meilleure estimation des lames d'eau. *Hydrologie continentale*, **49,** 1-4.

Payrastre, O., E. Gaume, and H. Andrieu, 2006: Apport du recueil de données historiques pour l'étude des crues extrêmes de petits cours d'eau. Etude du cas de quatre bassins versants affluents de l'Aude. *La Houille Blanche*, **6,** 79-86.

Payraudeau, S., M. Tournoud, and F. Cernesson, 2002: An adapted modelling approach for the nitrogen load management on a catchment scale. *DEVELOPMENTS IN WATER SCIENCE*, **47,** 1741-1748.

Pechlivanidis, I., B. Jackson, N. McIntyre, and H. Wheater, 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,** 193-214.

Peixoto, J. P., and A. H. Oort, 1992: Physics of climate.

Pellarin, T., G. Delrieu, G.-M. Saulnier, H. Andrieu, B. Vignal, and J.-D. Creutin, 2002: Hydrologic visibility of weather radar systems operating in mountainous regions: Case study for the Ardèche catchment (France). *Journal of Hydrometeorology*, **3,** 539-555.

Perry, M., and D. Hollis, 2005: The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology*, **25,** 1041-1054.

Pfisterer, C., 1998: Wetternachhersage, 500 Jahre Klimavariationen und Naturkatastrophen 1496–1995. *Verlag Paul Haupt, Bern*.

Pietroniro, A., and R. Leconte, 2000: A review of Canadian remote sensing applications in hydrology, 1995–1999. *Hydrological processes*, **14,** 1641-1666.

Pitman, A., and S. Perkins, 2009: Global and regional comparison of daily 2-m and 1000-hPa maximum and minimum temperatures in three global reanalyses. *Journal of Climate*, **22,** 4667-4681.

Poli, P., S. Healy, and D. Dee, 2010a: Assimilation of Global Positioning System radio occultation data in the ECMWF ERA–Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **136,** 1972-1990.

Poli, P., D. Dee, P. Berrisford, and J.-N. Thépaut, 2010b: Overview of satellite data assimilation in the ERA-Interim reanalysis. *Proceedings of 2010 EUMETSAT Meteorological Satellite Conference*, 1-8.

Poulin, A., F. Brissette, R. Leconte, R. Arsenault, and J.-S. Malo, 2011: Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. *Journal of hydrology*, **409,** 626-636.

Price, K., S. T. Purucker, S. R. Kraemer, J. E. Babendreier, and C. D. Knightes, 2014: Comparison of radar and gauge precipitation data in watershed models across varying spatial and temporal scales. *Hydrological Processes*, **28,** 3505-3520.

Puech, C., 2000: Utilisation de la télédétection et des modèles numériques de terrain pour la connaissance du fonctionnement des hydrosystèmes. *Habilitationa diriger des recherches, Grenoble University*.

Rakotomalala, R., 2008: Comparaison de populations et tests non paramétriques. *Université Lumière Lyon*, **2**.

Ramli, S., and W. Tahir, 2012: Radar hydrology: new ZR relationships over the Klang River Basin, Malaysia for monsoon season rainfall. *IAHS-AISH publication***,** 81-86.

Rango, A., 1994: *Applications of remote sensing by satellite, radar and other methods to hydrology.* World Meteorological Organization.

Rango, A., and A. I. Shalaby, 1998: Operational applications of remote sensing in hydrology: success, prospects and problems. *Hydrological Sciences Journal*, **43,** 947-968.

Rasmussen, R., M. Dixon, S. Vasiloff, F. Hage, S. Knight, J. Vivekanandan, and M. Xu, 2003: Snow nowcasting using a real-time correlation of radar reflectivity with snow gauge accumulation. *Journal of Applied Meteorology*, **42,** 20-36.

Renard, F., and J. Comby, 2006: Evaluation de techniques d'interpolation spatiale de la pluie en milieu urbain pour une meilleure gestion d'événements extrêmes: le cas du Grand Lyon. *La Houille Blanche*, 73-78.

Rhynsburger, D., 1973: Analytic Delineation of Thiessen Polygons*. *Geographical Analysis*, **5,** 133-144.

Riboust, P., and F. Brissette, 2015: Climate change impacts and uncertainties on spring flooding of Lake Champlain and the Richelieu River. *JAWRA Journal of the American Water Resources Association*, **51,** 776-793.

Ricketts, T. H., 1999: *Terrestrial ecoregions of North America: a conservation assessment.* Vol. 1, Island Press.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, **24,** 3624-3648.

Risley, J. C., 1994: Use of a precipitation-runoff model for simulating effects of forest management on streamflow in 11 small drainage basins, Oregon Coast Range.

Roads, J., S. C. Chen, M. Kanamitsu, and H. Juang, 1998: Vertical structure of humidity and temperature budget residuals over the Mississippi River basin. *Journal of Geophysical Research: Atmospheres*, **103,** 3741-3759.

Roads, J., and Coauthors, 2003: GCIP water and energy budget synthesis (WEBS). *Journal of Geophysical Research: Atmospheres*, **108**.

Rudolf, B., and U. Schneider, 2005: Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. *Proc. Second Workshop of the Int. Precipitation Working Group*, 231-247.

Ruelland, D., S. Ardoin-Bardin, G. Billen, and E. Servat, 2008: Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa. *Journal of Hydrology*, **361,** 96-117.

Ruiz-Barradas, A., and S. Nigam, 2006: Great Plains hydroclimate variability: The view from North American regional reanalysis. *Journal of climate*, **19,** 3004-3010.

Rusticucci, M., N. Zazulie, and G. B. Raga, 2014: Regional winter climate of the southern central Andes: Assessing the performance of ERA-Interim for climate studies. *Journal of Geophysical Research: Atmospheres*, **119,** 8568-8582.

Sabarly, F., G. Essou, P. Lucas-Picher, A. Poulin, and F. Brissette, 2016: Use of Four Reanalysis Datasets to Assess the Terrestrial Branch of the Water Cycle over Quebec, Canada. *Journal of Hydrometeorology*, **17,** 1447-1466.

Sagintayev, Z., and Coauthors, 2012: A remote sensing contribution to hydrologic modelling in arid and inaccessible watersheds, Pishin Lora basin, Pakistan. *Hydrological Processes*, **26,** 85-99.

Saha, S., and Coauthors, 2010: The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, **91,** 1015-1057.

Saha, S., and Coauthors, 2014: The NCEP climate forecast system version 2. *Journal of Climate*, **27,** 2185-2208.

Saltikoff, E., J. Koistinen, and H. Hohti, 2000: Experience of real time spatial adjustment of the ZR relation according to water phase of hydrometeors. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **25,** 1017-1020.

Sauvageot, H., 2000: Le radar polarimétrique, une nouvelle approche pour l'observation des champs de précipitations.

Schaake, J., S. Cong, and Q. Duan, 2006: The US MOPEX data set. *IAHS publication*, **307,** 9.

Schaake, J., Q. Duan, M. Smith, and V. Koren, 2000: Criteria to select basins for hydrologic model development and testing. *Preprints, 15thConference on Hydrology***,** 10-14.

Schmugge, T. J., W. P. Kustas, J. C. Ritchie, T. J. Jackson, and A. Rango, 2002: Remote sensing in hydrology. *Advances in water resources*, **25,** 1367-1385.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2014: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, **115,** 15-40.

Schubert, S. D., R. B. Rood, and J. Pfaendtner, 1993: An assimilated dataset for earth science applications. *Bulletin of the American meteorological Society*, **74,** 2331-2342.

Schultz, G. A., 1996: Remote sensing applications to hydrology: runoff. *Hydrological sciences journal*, **41,** 453-475.

Schultz, G. A., and E. T. Engman, 2000: *Remote sensing in hydrology and water management.* Springer Science & Business Media.

Seo, D.-J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *Journal of Hydrology*, **208,** 37-52.

Serreze, M. C., and C. M. Hurst, 2000: Representation of mean Arctic precipitation from NCEP-NCAR and ERA reanalyses. *Journal of Climate*, **13,** 182-201.

Servat, E., and A. Dezetter, 1991: Selection of calibration objective fonctions in the context of rainfall-ronoff modelling in a Sudanese savannah area. *Hydrological Sciences Journal*, **36,** 307-330.

Sevruk, B., 1996: Adjustment of tipping-bucket precipitation gauge measurements. *Atmospheric Research*, **42,** 237-246.

Shafran, P., J. Woollen, W. Ebisuzaki, W. Shi, Y. Fan, R. Grumbine, and M. Fennessy, 2004: Observational data used for assimilation in the NCEP North American Regional Reanalysis. *Preprints, 14th Conf. on Applied Climatology, Seattle, WA, Amer. Meteor. Soc*.

Shamseldin, A. Y., K. M. O'Connor, and G. Liang, 1997: Methods for combining the outputs of different rainfall-runoff models. *Journal of Hydrology*, **197,** 203-229.

Sheffield, J., G. Goteti, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, **19,** 3088-3111.

Sheffield, J., B. Livneh, and E. F. Wood, 2012: Representation of terrestrial hydrology and large-scale drought of the continental United States from the North American Regional Reanalysis. *Journal of Hydrometeorology*, **13,** 856-876.

Shepard, D. S., 1984: Computer mapping: The SYMAP interpolation algorithm. *Spatial Statistics and Models*, Springer, 133-145.

Shiu, C. J., S. C. Liu, C. Fu, A. Dai, and Y. Sun, 2012: How much do precipitation extremes change in a warming climate? *Geophysical Research Letters*, **39**.

Shrestha, M., G. Artan, S. Bajracharya, and R. Sharma, 2008: Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin. *Journal of Flood Risk Management*, **1,** 89-99.

Shrubsole, D., R. Kreutzwiser, B. Mitchell, T. Dickinson, and D. Joy, 1993: The history of flood damages in Ontario. *Canadian Water Resources Journal*, **18,** 133-143.

Sillmann, J., V. Kharin, F. Zwiers, X. Zhang, and D. Bronaugh, 2013a: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmospheres*, **118,** 2473-2493.

Sillmann, J., V. Kharin, X. Zhang, F. Zwiers, and D. Bronaugh, 2013b: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres*, **118,** 1716-1733.

Simmons, A., K. Willett, P. Jones, P. Thorne, and D. Dee, 2010: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **115**.

Singh, R., S. Archfield, and T. Wagener, 2014: Identifying dominant controls on hydrologic parameter transfer from gauged to ungauged catchments–A comparative hydrology approach. *Journal of Hydrology*, **517,** 985-996.

Singh, V. P., and D. A. Woolhiser, 2002: Mathematical modeling of watershed hydrology. *Journal of hydrologic engineering*, **7,** 270-292.

Skaugen, T., and J. Andersen, 2010: Simulated precipitation fields with variance-consistent interpolation. *Hydrological sciences journal*, **55,** 676-686.

St-Hilaire, A., T. B. Ouarda, M. Lachance, B. Bobée, J. Gaudet, and C. Gignac, 2003: Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study. *Hydrological Processes*, **17,** 3561-3580.

Steiner, M., J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden, 1999: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resources Research*, **35,** 2487-2503.

STREAMFLOW, I. M., 2009: Identifying hydrologic processes in agricultural watersheds using precipitation-runoff models.

Suarez, M. J., and Coauthors, 2008: The GEOS-5 Data Assimilation System-Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0.

Sun, Q., C. Miao, Q. Duan, D. Kong, A. Ye, Z. Di, and W. Gong, 2014: Would the 'real'observed dataset stand up? A critical examination of eight observed gridded climate datasets for China. *Environmental Research Letters*, **9,** 015001.

Sun, X., and A. P. Barros, 2010: An evaluation of the statistics of rainfall extremes in rain gauge observations, and satellite-based and reanalysis products using universal multifractals. *Journal of Hydrometeorology*, **11,** 388-404.

Sun, X., R. Mein, T. Keenan, and J. Elliott, 2000: Flood estimation using radar and raingauge data. *Journal of Hydrology*, **239,** 4-18.

Svoboda, A., P. Pekárová, P. Miklánek, and S. výbor pre hydrológiu Bratislava, 2000: *Flood hydrology of Danube between Devín and Nagymaros.* Ústav hydrológie SAV.

Tang, Q., H. Gao, H. Lu, and D. P. Lettenmaier, 2009: Remote sensing: hydrology. *Progress in Physical Geography*, **33,** 490-509.

Tapsoba, D., V. Fortin, F. Anctil, and M. Haché, 2005: Apport de la technique du krigeage avec dérive externe pour une cartographie raisonnée de l'équivalent en eau de la neige: Application aux bassins de la rivière Gatineau. *Canadian journal of civil engineering*, **32,** 289-297.

Taupin, J.-D., 2003: Précision dé l'estimation des précipitations au Sahel selon la densité du réseau d'observation pluviométrique. *Comptes Rendus Géoscience*, **335,** 215-225.

Taupin, J., A. Amani, and T. Lebel, 1998: Variabilité spatiale des pluies au Sahel: une question d'échelles. 1. Approche expérimentale. *IAHS PUBLICATION***,** 143-152.

Taylor, G., C. Daly, W. Gibson, and J. Sibul-Weisberg, 1997: Digital and map products produced using PRISM. *Proc., 10th AMS Conf. on Applied Climatology, Amer. Meteorological Soc., Reno, NV, Oct*, 20-24.

Teo, C.-K., and D. I. Grimes, 2007: Stochastic modelling of rainfall from satellite data. *Journal of hydrology*, **346,** 33-50.

Thiessen, A. H., 1911: Precipitation averages for large areas. *Monthly weather review*, **39,** 1082-1089.

Thorndycraft, V. R., and G. Benito, 2003: *Palaeofloods, Historical Data and Climatic Variability: Applications in Flood Risk Assessment: Proceedings of the PHEFRA International Workshop Held in Barcelona, 16-19th October, 2002.* CSIC-Centro de Ciencias Medioambientales.

Thornton, P., M. Thornton, B. Mayer, N. Wilhelmi, Y. Wei, and R. Cook, 2012: Daymet: Daily surface weather on a 1 km grid for North America, 1980-2008. *Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, T, N. doi*, **10**.

Thornton, P. E., S. W. Running, and M. A. White, 1997: Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, **190,** 214-251.

Tol, R. S., and A. Langen, 2000: A concise history of Dutch river floods. *Climatic change*, **46,** 357-369.

Tozer, C., A. Kiem, and D. Verdon-Kidd, 2012: On the uncertainties associated with using gridded rainfall data as a proxy for observed. *Hydrology and Earth System Sciences*, **16,** 1481-1499.

Tran, T., 2010: ESTIMATION DE L'ÉTAT HYDRIQUE DES SOLS EN AFRIQUE DE L'OUEST PAR TÉLÉDÉTECTION SPATIALE, Université de Grenoble.

Trenberth, K. E., and C. J. Guillemot, 1998: Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. *Climate Dynamics*, **14,** 213-231.

Trenberth, K. E., J. T. Fasullo, and J. Mackaro, 2011: Atmospheric moisture transports from ocean to land and global energy flows in reanalyses. *Journal of Climate*, **24,** 4907-4924.

Tridon, F., 2011: Mesure des précipitations à l'aide d'un radar en bande X non-cohérent à haute résolution et d'un radar en bande K à visée verticale. Application à l'étude de la variabilité des précipitations lors de la campagne COPS, Université Blaise Pascal-Clermont-Ferrand II.

Turco, M., A. Zollo, C. Ronchi, C. D. Luigi, and P. Mercogliano, 2013: Assessing gridded observations for daily precipitation extremes in the Alps with a focus on northwest Italy. *Natural Hazards and Earth System Science*, **13,** 1457-1468.

Turk, F. J., P. Arkin, M. R. Sapiano, and E. E. Ebert, 2008: Evaluating high-resolution precipitation products. *Bulletin of the American Meteorological Society*, **89,** 1911-1916.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131,** 2961-3012.

Upton, G., and A. Rahimi, 2003: On-line detection of errors in tipping-bucket raingauges. *Journal of Hydrology*, **278,** 197-212.

Van de Griend, A., and E. Engman, 1985: Partial area hydrology and remote sensing. *Journal of Hydrology*, **81,** 211-251.

Vente, C., 1964: Handbook of applied hydrology: a compendium of water-resources technology.

Vicente, G., J. Davenport, and R. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *International Journal of Remote Sensing*, **23,** 221-230.

Vischel, T., 2006: Impact de la variabilité pluviométrique de méso-échelle sur la réponse des systèmes hydrologiques sahéliens: modélisation, simulation et désagrégation, Grenoble, INPG.

Vose, R. S., S. Applequist, M. J. Menne, C. N. Williams, and P. Thorne, 2012: An intercomparison of temperature trends in the US Historical Climatology Network and recent atmospheric reanalyses. *Geophysical Research Letters*, **39**.

Vu, M., S. Raghavan, and S. Liong, 2012: SWAT use of gridded observations for simulating runoff–a Vietnam river basin study. *Hydrology and Earth System Sciences*, **16,** 2801-2811.

Walker, J. P., 1999: Estimating soil moisture profile dynamics from near-surface soil moisture measurements and standard meteorological data, the University of Newcastle.

Wang, S., D. W. McKenney, J. Shang, and J. Li, 2014: A national-scale assessment of long-term water budget closures for Canada's watersheds. *Journal of Geophysical Research: Atmospheres*, **119,** 8712-8725.

Wang, W., P. Xie, S.-H. Yoo, Y. Xue, A. Kumar, and X. Wu, 2011: An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Climate dynamics*, **37,** 1601-1620.

Warner, T. T., E. A. Brandes, J. Sun, D. N. Yates, and C. K. Mueller, 2000: Prediction of a flash flood in complex terrain. Part I: A comparison of rainfall estimates from radar, and very short range rainfall simulations from a dynamic model and an automated algorithmic system. *Journal of Applied Meteorology*, **39,** 797-814.

Weedon, G., and Coauthors, 2011: Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology*, **12,** 823-848.

Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo, 2014: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, **50,** 7505-7514.

Wei, J., H. Su, and Z.-L. Yang, 2015: Impact of moisture flux convergence and soil moisture on precipitation: a case study for the southern United States with implications for the globe. *Climate Dynamics*, 1-15.

Westrick, K. J., C. F. Mass, and B. A. Colle, 1999: The limitations of the WSR-88D radar network for quantitative precipitation measurement over the coastal western United States. *Bulletin of the American Meteorological Society*, **80,** 2289-2298.

Whitaker, J. S., G. P. Compo, and J.-N. Thépaut, 2009: A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Monthly Weather Review*, **137,** 1991-1999.

Wibig, J., A. Jaczewski, B. Brzóska, K. Konca-Kędzierska, and K. Pianko-Kluczyńska, 2014: How does the areal averaging influence the extremes? The context of gridded observation data sets. *Meteorologische Zeitschrift*, 181-187.

Widmann, M., and C. S. Bretherton, 2000: Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. *Journal of Climate*, **13,** 1936-1950.

Willett, K., A. Dolman, B. Hall, and P. Thorne, 2012: Global climate [in ‘‘State of the Climate in 2011’’]. *Bull. Amer. Meteor. Soc*, **93,** S7-S55.

Williams, A., and D. Archer, 2002: The use of historical flood information in the English Midlands to improve risk assessment. *Hydrological sciences journal*, **47,** 57-76.

Winkler, T., 1993: Standardized hydrometric data collection in Canada: the career development program for technician training. *Water international*, **18,** 217-224.

Woo, M. K., and R. Thorne, 2006: Snowmelt contribution to discharge from a large mountainous catchment in subarctic Canada. *Hydrological Processes*, **20,** 2129-2139.

Xiaoyang, L., M. Jietai, Z. Yuanjing, and L. Jiren, 2003: Runoff simulation using radar and rain gauge data. *Advances in Atmospheric Sciences*, **20,** 213-218.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society*, **78,** 2539-2558.

Xu, X., J. Li, and B. A. Tolson, 2014: Progress in integrating remote sensing data and hydrologic modeling. *Progress in Physical Geography***,** 0309133314536583.

Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 2012: APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, **93,** 1401-1415.

Ye, S., and Coauthors, 2014: Regionalization of subsurface stormflow parameters of hydrologic models: Derivation from regional analysis of streamflow recession curves. *Journal of Hydrology*, **519,** 670-682.

Yilmaz, K. K., T. S. Hogue, K.-l. Hsu, S. Sorooshian, H. V. Gupta, and T. Wagener, 2005: Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *Journal of Hydrometeorology*, **6,** 497-517.

Yin, H., M. G. Donat, L. V. Alexander, and Y. Sun, 2014: Multi-dataset comparison of gridded observed temperature and precipitation extremes over China. *International Journal of Climatology*.

Yu, M., X. Chen, L. Li, A. Bao, and M. J. de la Paix, 2011: Streamflow simulation by SWAT using different precipitation sources in large arid basins with scarce raingauges. *Water resources management*, **25,** 2669-2681.

Zhang, L., A. Kumar, and W. Wang, 2012: Influence of changes in observations on precipitation: A case study for the Climate Forecast System Reanalysis (CFSR). *Journal of Geophysical Research: Atmospheres (1984–2012)*, **117**.

Zhang, Q., H. Körnich, and K. Holmgren, 2013: How well do reanalyses represent the southern African precipitation? *Climate dynamics*, **40,** 951-962.

Zurita-Milla, R., H. Mehdipoor, S. Batarseh, T. Ault, and M. D. Schwartz, 2014: On the use of gridded daily temperature data to calculate the extended spring indices phenological models. *EGU General Assembly Conference Abstracts*, 13427.