UNIVERSITÉ DU QUÉBEC À MONTRÉAL

VALEUR AJOUTÉE DANS LE MODÈLE REGIONAL CANADIEN DU CLIMAT : COMPARAISON DE LA PRÉCIPITATION AUX ÉCHELLES DU MODÈLE GLOBAL CANADIEN DU CLIMAT

MÉMOIRE
PRÉSENTÉ COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN SCIENCES DE L'ATMOSPHÈRE

PAR
ALEJANDRO DI LUCA

JANVIER 2009

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ADDED VALUE IN THE CANADIAN REGIONAL CLIMATE MODEL: COMPARISON OF PRECIPITATION AT THE SCALES OF THE CANADIAN GLOBAL CLIMATE MODEL

THESIS
PRESENTED IN PARTIAL FULLFILMENT
OF THE REQUIREMENTS FOR THE MASTER'S DEGREE
IN ATMOSPHERIC SCIENCES

BY
ALEJANDRO DI LUCA

JANUARY 2009

REMERCIEMENTS

À mon directeur, René Laprise, pour ses précieuses contributions à ce travail et pour son encouragement constant tout au long de la maîtrise.

À mon codirecteur, Ramón de Elía, avec qui il est si agréable de travailler et qui m'a poussé à réfléchir au-delà des questions scientifiques. Ainsi que par ses généreux conseils qui m'ont aidé énormément au cours des dernières deux années.

À Ross Brown pour les intéressantes discussions au sujet des observations de précipitation. À Georges, Abderrahim et Mourad pour leur inestimable aide technique. Aux membres de l'équipe de simulation climatique d'Ouranos, en particulier à Anne, Michel et Dominique qui, entre autres choses, m'ont aidé avec les données du MRCC.

À tous les collègues et amis de l'UQAM et d'Ouranos, qui sont toujours prêts à donner un coup de main quand j'en ai eu besoin et qui ont égayé quelques soirées à l'extérieur du travail.

Au centre ESCER pour le soutien financier et au Consortium Ouranos pour l'espace physique où j'ai pu accomplir ma recherche.

À la famille et les amis, qui même à distance, sont toujours le support moral et la plus grande source d'inspiration.

Finalement, à mes colocataires: Anso, Caro, Guigui et Nico, qui sont en grande partie responsables de ces deux dernières années très heureuses.

TABLE DE MATIÈRES

LISTE DE FIGURES

Figure Page

LISTE DE TABLEAUX

Tableau Page

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AGCM	Atmospheric general circulation model
ALTA	Alberta (province of Canada)
AOGCM	Atmosphere-ocean general circulation model
ANSD	Absolute normalized seasonal deviation
BBE	Big Brother Experiment
BC	British Columbia (province of Canada)
CFL	Conditions aux frontières latérales
CGCM	version 3 of the Canadian centre for climate modelling and analysis coupled global climate model
CRCM	Canadian regional climate model
DD	Dry days
DJF	December-January-February
GCM	General circulation model
IPCC	Intergovernmental panel of climate change
JJA	June-July-August
LAM	Limited area model
LBC	Lateral boundary conditions
MAM	March-April-May
MAN	Manitoba (province of Canada)
MCG	Modèle de circulation générale
MCCG	Modèle couplé de circulation générale
MCGC	Modèle de circulation générale canadien
MINK	Région de Missouri, Iowa, Nebraska et Kansas
MRC	Modèle régional du climat
MRCC	Modèle régional Canadien du climat
NARCAAP	North American Regional Climate Change Assessment Program
NCEP	National Center for Environmental Prediction
NCAR	National Centers for Atmospheric Research
NSD	Normalized seasonal deviation

PR	Precipitation rate
PRUDENCE	Prediction of Regional Scenarios and Uncertanties for Defining European Climate Change Risks and Effects
QC	Quebec (province of Canada)
RCM	Regional climate model
SAS	Saskatchewan (province of Canada)
SON	September-October-November
UK	United Kingdom
WD	Wet days

LISTE DES SYMBOLES

B	distribution binomiale
f	fonction histogramme
Gs	échelles spatiales résolues seulement par le MCG
H	numéro de classifications dans l'histogramme linéaire
i	indice des stations ou des points de grille
j	indice des jours
K	numéro de classifications dans l'histogramme logarithmique
m	indice des mois
mod	données simulées
obs	données observées
P()	probabilité de…
r	indice des régions
R_{dd}	division entre les jours secs simulés et observés
Rss	échelles spatiales résolues seulement par le MRC
s	indice des saisons
S	mesure statistique définie par Perkins *et al.* (2007)
y	indice des années
Ω_h	taille de la classification (mm/jour)
Ω_k	taille de la classification de l'histogramme logarithmique (mm/jour)
Δx_G	intervalle de la grille du MCG (km)
Δx_R	intervalle de la grille du MRC (km)

RÉSUMÉ

La modélisation du climat à haute résolution est nécessaire pour une meilleure compréhension des impacts des changements climatiques. Les modèles régionaux du climat (MRC) constituent une des principales sources de ce type de données puisque les modèles de circulation générale (MCG) ne fonctionnent toujours pas à une résolution suffisante pour répondre à ces besoins.

Une fois que les MRC sont devenus des outils capables de générer des simulations physiquement réalistes, un effort important a été fait pour évaluer leur capacité de mise à l'échelle, en se concentrant principalement sur des variables moyennées temporellement. Cet effort ne s'est pas traduit par des améliorations sans équivoque par rapport aux simulations produites par les MCG.

L'objectif principal de cette étude est d'examiner l'existence de la valeur ajoutée dans les simulations du modèle régional Canadien du climat (MRCC) par rapport à celles du modèle de circulation général canadien (MCGC) utilisé comme pilote. Dans cette première étape, il a été nécessaire d'analyser les échelles temporelles et spatiales communes aux deux modèles, le MRCC et le MCGC. Une comparaison est effectuée en ramenant les données à haute résolution des stations météorologiques et du MRCC à la résolution du MCGC.

L'évaluation se base sur la comparaison des histogrammes d'intensités de précipitation et des 95^{e} centiles des distributions afin de caractériser les événements extrêmes. On estime le degré de chevauchement entre les distributions simulées et observées en utilisant la mesure S définie par Perkins *et al.* (2007). Cette dernière reflète principalement le comportement des intensités faibles et modérées.

Les résultats montrent que les statistiques quotidiennes des précipitations simulées par le MGCC et le MRCC sont généralement très similaires. En comparant les résultats des deux modèles, il n'existe aucune preuve de l'existence de la valeur ajoutée. En outre, pendant l'été, les données simulées par le modèle MCGC sont plus proches des observations que celles générées par le MRCC. Cette amélioration provient d'une meilleure simulation de la fréquence des jours secs. Pour les événements quotidiens les plus intenses, le MCGC produit aussi des résultats plus proches des valeurs observées que le MRCC. Ce dernier montre une sous-estimation constante de la fréquence d'occurrence des événements intenses. C'est aussi le cas dans les régions caractérisées par d'importants forçages de surface, où la différence entre les topographies des deux modèles pourrait avoir un impact.

Mots clés : MRCC, mise à l'échelle, précipitation, valeur ajoutée, histogrammes.

INTRODUCTION

Les outils principaux permettant d'étudier le climat futur sont les modèles couplés de circulation générale (MCCG). Ces modèles sont dérivés des lois physiques fondamentales et incluent, entre autres, des composants dynamiques décrivant des processus atmosphériques, océaniques, de la surface terrestre, ainsi que de la glace de mer. La dynamique est soumise à des approximations appropriées pour la grande échelle du système climatique (par exemple, l'approximation hydrostatique dans la composante atmosphérique) et la discrétisation des équations provoque une approximation supplémentaire. La multitude et la complexité des processus à résoudre, la longueur nécessaire des simulations pour l'étude du climat, ainsi que la nécessité d'effectuer des ensembles de simulations comprenant plusieurs membres afin d'obtenir des estimations statistiquement robustes imposent des contraintes de temps de calcul qui limitent l'intervalle de la grille sur laquelle les équations sont discrétisées. Présentement, les distances horizontales des grilles atmosphériques varient entre 125 et 400 km (Randall *et al.*, 2007), et elles sont insuffisantes pour reproduire la structure à petite échelle des variables climatiques. Par conséquent, des paramétrages doivent êtres utilisés pour représenter les procédés physiques non résolus. Ainsi, la confiance envers les modèles climatiques pour fournir des estimations quantitatives crédibles du climat est limitée aux échelles continentales et plus grandes (Randall *et al.*, 2007).

Dans ce contexte, une alternative pour obtenir des projections climatiques régionales futures est l'utilisation des modèles régionaux du climat (MRC) à haute résolution, en utilisant des conditions aux frontières latérales (CFL) des MCG à plus basse résolution (Dickinson *et al.*, 1989; Giorgi et Bates, 1989; Laprise, 2006).

La plus haute résolution horizontale des MRC implique deux grands avantages potentiels par rapport aux MCG : une discrétisation plus précise des équations qui permet une plus large gamme d'échelles spatiales explicitement résolues et, peut-être plus importante encore, une amélioration de la représentation des forçages de surface comme la topographie, les contrastes terre - mer, *etc.*

L'impact de l'augmentation de la résolution horizontale a été le sujet de plusieurs études dans les MCG (Boer et Lazare, 1988; Boville, 1991; Boyle, 1993) et dans les MRC (Marinucci et Giorgi, 1996; Castro *et al.*, 2005; Xue *et al.*, 2007), généralement en utilisant des résultats du même modèle mais pour différentes résolutions. Ces études montrent que les simulations à plus haute résolution ne produisent pas nécessairement des résultats plus proches des valeurs observées, mais que les performances dépendent fortement du comportement des paramétrages. En d'autres termes, l'augmentation de la résolution horizontale peut aggraver le comportement des paramétrages des processus de sous échelles et donc, nécessiter de faire appel à des nouveaux paramétrages ou à modifier ceux qui existent déjà. Par exemple, Marinucci et Giorgi (1996) ont étudié des données de précipitation simulés par un MRC à haute résolution et ont constaté que «the effects of physical forcings (e.g., a better representation of topography and coastlines) may be masked by the direct sensitivity of the model parameterizations to resolution itself, at least in the continental scale».

Les avantages provenant de l'utilisation de plus haute résolution des MRC sont influencés non seulement par la sensibilité des paramétrages à la résolution elle-même, mais aussi par la fiabilité technique de mise à l'échelle dynamique. Ici, la fiabilité de la mise à l'échelle dynamique est définie de manière similaire au *second principe* de Laprise *et al.* (2007) : la petite échelle générée par le MRC possèdes des amplitudes et statistiques climatiques qui seraient présents dans les données de pilotage si elles n'étaient pas limitées par la résolution. Cette affirmation a été étudiée en isolant les erreurs de la technique de pilotage sans prendre en compte ceux qui viennent des modèles particuliers ou des CFL, c'est-à-dire dans le contexte d'une *approche parfaite*. Tel qu'établi par Laprise *et al.* (2007), le *second principe* semble être valable dans certaines conditions particulières: aux latitudes moyennes, pour les niveaux inférieurs et pour des domaines suffisamment grands. La méthode la plus populaire des approches parfaites a été développée par Denis et al. (2002) et est désignée sous le nom de l'Expérience Grand Frère (EGF). Le protocole de l'EGF a été appliqué dans plusieurs contextes (voir Denis *et al.*, 2002, 2003; Antic *et al.*, 2005, de Elía *et al.*, 2002; Dimitrijevic et Laprise, 2005; Herceg, 2006; Koltzow *et al.*, 2008).

La fiabilité de la technique de mise à l'échelle dynamique prouve l'existence de *valeur ajoutée potentielle* dans des simulations du MRC et constitue une condition nécessaire à l'existence de *valeur ajoutée réelle*. Cette dernière doit être identifiée par l'étude des simulations du MRC dans des contextes plus réalistes que des approches parfaites, en établissant l'utilité de la petite échelle générée par le modèle. Certaines des difficultés qui apparaissent sur la détermination de la valeur ajoutée seront examinées au cours du mémoire.

Certaines études se sont penchées sur la présence de la valeur ajoutée dans des simulations des MRC en comparant avec un certain type de données observées. Giorgi *et al.* (1998) ont comparé la précipitation et la température de surface simulées par le modèle régional du NCAR RegCM et par le modèle mondial CSIRO avec des données observées (Legates et Willmott, 1990) dans la région du Missouri, de l'Iowa, du Nebraska et du Kansas (région MINK). La comparaison des moyennes saisonnières des précipitations dans la région MINK montre des différences importantes entre les résultats des deux modèles et les observations. L'existence de valeur ajoutée par le MRC n'est donc pas évidente. Le résultat le plus intéressant est la représentation des champs de précipitation pour la saison estivale. Bien que la région MINK ne soit pas caractérisée par une topographie complexe, la corrélation spatiale de la moyenne estivale entre les champs observés et simulés par le MRC est de 0,77 alors qu'elle est de –0,69 entre les observations et les résultats du MCG, ce qui montre une grande amélioration dans la distribution spatiale de la précipitation.

Durman *et al.* (2001) ont étudié la précipitation quotidienne simulée par deux versions du modèle HadCM2 : une version mondiale et une version à plus haute résolution et à aire limitée (HadRCM) qui utilise comme pilote la version mondiale. En comparant les distributions de fréquence d'intensité simulées et observées, ils trouvent que, pendant l'hiver, le MRC produit trop souvent des événements caractérisés par des taux de précipitation intenses. Au contraire, pendant l'été, le MRC produit une meilleure représentation de la distribution de précipitation par rapport au MCG, particulièrement due à une amélioration de la simulation des événements les plus intenses.

Feser (2006) utilise un filtre spatial pour séparer les résultats du MRC en deux gammes d'échelles spatiales : les échelles moyennes (entre 250 et 550 km) et les grandes échelles (plus de 700 km). Les résultats montrent que, lors de l'évaluation des variables à grande

échelle, comme la pression au niveau de la mer, les structures spatiales produites par le MRC ne sont pas très différentes de celles qui sont obtenues par les réanalyses du NCEP/NCAR. Au contraire, lors de l'évaluation de la température près de la surface dans les échelles moyennes, la performance du MRC dans la représentation des structures spatiales surpasse celles des réanalyses NCEP. L'auteur suggère que l'amélioration provient d'une meilleure représentation des propriétés de la surface dans le MRC.

L'objectif principal de ce projet est de contribuer à détecter objectivement la valeur ajoutée générée par un modèle régional du climat. Comme première étape dans ce but général, nous avons évalué les statistiques quotidiennes des précipitations simulées par un MRC et un MCG (utilisé comme pilote du MRC) à l'aide d'observations, pour plusieurs régions du Canada. L'analyse se concentre sur des échelles temporelles et spatiales représentées par les deux modèles, mondial et régional, mais où le modèle mondial devrait avoir peu d'habileté pour résoudre les processus dû au fait que les échelles sont près de sa limite de troncature (Laprise, 2003; Feser, 2006). La fine échelle temporelle et spatiale produite par le MRC n'a pas été explicitement évaluée parce que, comme nous allons le montrer plus tard, celle-ci suit une approche complètement différente. Certes, pour obtenir une évaluation plus globale, ce travail devra être complété par des études examinant d'autres aspects de la valeur ajoutée, y compris les caractéristiques spatiales et temporelles de fine échelle, d'autres variables, etc.

Le travail est présenté sous forme d'article rédigé en anglais dans le but de le soumettre à une revue scientifique. La première partie de l'article comprend une brève discussion de certaines questions importantes sur l'évaluation de la valeur ajoutée. Par la suite, les données observées et simulées utilisées dans les comparaisons et les différentes statistiques servant à évaluer les performances des modèles sont présentées. Ensuite, on présente l'analyse des résultats obtenus lors de l'évaluation du cycle annuel et la valeur quotidienne des précipitations pour conclure avec une discussion et un résumé des résultats obtenus.

CHAPITRE I

VALEUR AJOUTÉE DANS LE MRCC : COMPARAISON DE LA PRÉCIPITATION AUX ÉCHELLES DU MGC

Added value in the Canadian Regional Climate Model: comparison of precipitation at the scales of the Canadian Global Climate Model

by
Alejandro Di Luca [1], Ramón de Elía [2] and René Laprise [1]

Centre ESCER (Etude et Simulation du Climat à l'Échelle Régionale)

[1] Université du Québec à Montréal
[2] Ouranos Consortium

August 2008

Corresponding author's address:

Alejandro Di Luca
UQAM
Ouranos, 550 Sherbrooke St West, 19th floor, West Tower
Montréal (Québec), Canada, H3A 1B9
Tel: +1 (514) 282 6464 ext. 244
E-mail: diluca@sca.uqam.ca

Abstract

High-resolution climate information is currently in high demand in climate-change impact studies. Regional climate models (RCMs) constitute one of the main sources of this kind of datasets since present-day General Circulation Models (GCMs) do not run at a resolution sufficient to satisfy these needs.

Once RCMs were shown to be technically feasible, a large effort has since ensued to assess their capability as a climate downscaling tool, mostly concentrating on time-averaged fields. This effort has not resulted into unequivocal gains when compared to GCM simulations performed at much coarser resolution.

The primary aim of this study is to investigate the existence of added value in the Canadian RCM simulations compared to its driving Canadian coupled GCM simulations. As a first but necessary step, temporal and spatial scales that are common to both, the CGCM and the CRCM, are considered in the analysis. The comparison is performed by upscaling, at the CGCM level, data from the CRCM and from meteorological stations.

The assessment is based on the comparison of simulated and observed intensity frequency distributions of precipitation and on the computation of 95^{th} percentile of the distributions to characterize more extreme events. The S score defined in Perkins et al. (2007) is used to measure the overlap between simulated and observed distributions and reflects mainly the simulation of light-moderate precipitation rates.

Results show that the daily statistics of precipitation as simulated by the CGCM and by the CRCM are generally very similar and, when comparing both data, there is no evidence of the existence of added value in CRCM simulations. Moreover, in summer season, the CGCM shows a better agreement with observed data than the CRCM and this improvement comes from a better simulation of the frequency of observed dry days. In the case of more extremes daily values, the CGCM produce results closer to the observed values than the CRCM. The latter shows a consistent underestimation of the frequency of occurrence of heavier events. This is even the case in regions characterized by important surface forcings, where differences between model topographies may be expected to have an impact.

Key words: Regional Climate Model, driven data, added value, upscale precipitation, intensity frequency distributions.

1. Introduction

The primary and most comprehensive tools to study future climate are the Atmosphere-Ocean General Circulation Models (AOGCMs). These models are derived from fundamental physical laws and include dynamical components describing atmospheric, oceanic and land surface processes, as well as sea ice and other components. The dynamics is subject to approximations appropriate for the large-scale climate system, such as the hydrostatic approximation, and then further approximated through mathematical discretization. The large number and complexity of processes to be resolved, the long simulations needed for climate studies, and the need of ensemble members for better statistical estimates impose computational constraints that restrict the horizontal grid mesh used in the discretized equations. Present horizontal grid intervals of the atmospheric component, usually between 125 and 400 km (Randall et al., 2007), are insufficient to capture the fine-scale structure of climatic variables and parameterizations need to be used to represent the unresolved, subgrid-scale physical processes. As a result, confidence that climate models provide credible quantitative estimates of present-future climate is particularly true at continental scales and above (Randall et al., 2007).

In this context, an alternative to obtain future regional climate projections is the use of high-resolution Regional Climate Models (RCMs), nested at their lateral boundaries with low-resolution AOGCMs (Dickinson et al., 1989; Giorgi and Bates, 1989; Laprise, 2008). The enhanced horizontal resolution of the RCM implies two great advantages with respect to the AOGCM: a more accurate discretization of equations which permits a broader range of spatial scales explicitly resolved and, perhaps more important, an improvement in the representation of surface forcings such as topography, coastal regions, etc.

The impact of increasing resolution has been discussed by several authors, both for GCMs (Boer and Lazare, 1988; Boville, 1991; Boyle, 1993) and RCMs (Giorgi and Marinucci, 1996; Castro et al., 2005; Xue et al., 2007), generally using results from the same model running at different resolutions. These studies show that simulations at higher resolution do not necessarily produce results closer to the observed values and that the performances are strongly dependent on the behaviour of parameterizations. That is, the

increase in horizontal resolution may alter the appropriateness of parameterized subgrid-scale processes and hence call for new parameterizations or a retuning of the existing ones. For example, Giorgi and Marinucci (1996) studied the simulated precipitation by a RCM when increasing resolution and found that "the effects of physical forcings (e.g., a better representation of topography and coastlines) may be masked by the direct sensitivity of the model parameterizations to resolution itself, at least in the continental scale".

Benefits coming from the use of higher resolution RCMs are influenced not only by the sensitivity of parameterizations to resolution itself but also by the reliability of the one-way nesting technique. Here, the reliability of the dynamical downscaling technique is defined in a similar way as Tenet 2 in Laprise et al. (2008): the small scales generated (by RCMs) have the amplitudes and climate statistics that would be present in the driving data if it were not limited by resolution. This assertion was studied by isolating errors from the nesting technique without considering those coming from particular models or from lateral boundary conditions (LBCs), i.e. in the context of perfect prognosis approach. As stated by Laprise et al. (2008), Tenet 2 appears to be valid in some special conditions: mid-latitudes, low levels and for suitable large domain size. The most popular perfect prognosis approach was developed by Denis et al. (2002) and is referred as the "Big Brother Experiment" (BBE). The BBE framework has been applied in several contexts (see Denis et al., 2002, 2003; Antic et al., 2005; de Elía et al., 2002; Dimitrijevic and Laprise, 2005; Herceg, 2006; Koltzow et al., 2008).

The reliability of the one-way nesting technique, as proven in the context of the BBE, demonstrates the existence of "potential added value" in RCM simulations and constitutes a necessary condition to the existence of "real added value". This later added value must be identified through studying RCM simulations in more realistic frameworks than the perfect prognosis approach, establishing the usefulness of the small scale generated by the RCM. In other words, the study of added value in RCM simulations should also be carried out by including in the analysis those errors coming from the model itself and/or from LBC. The most obvious way to address this problem is by comparing RCM results with the observed climate.

Some studies have tried to identify the added value in RCM simulations by comparing with observations. Giorgi et al. (1998) compared the precipitation and surface temperature simulated by NCAR RegCM regional model and its driving CSIRO GCM in the Missouri, Iowa, Nebraska and Kansas (MINK) region. Comparisons of seasonal mean precipitation in the MINK region showed a large bias for both models, and there are no clear evidences of added value. The most interesting result is the representation of the spatial patterns during the summer season. Although the MINK region is not characterized by pronounced local topographic variability, the spatial correlation coefficients between observed and control run RCM for mean summer fields is 0.77 and between observed and control run GCM is –0.69, showing a great improvement in the spatial pattern.

Durman et al. (2001) studied the simulated daily precipitation by two versions of the HadCM2 model: a global (HadCM2) and a limited area version (HadRCM). The HadRCM is driven by the HadCM2 and the comparison is performed in two HadCM2 grid boxes that includes Scotland and south-east England. Comparisons between simulated and observed intensity frequency distributions show that, in winter season, the HadRCM has a large positive bias in the frequencies of heavier events and performs worse than the HadCM2. On the contrary, in summer season, the HadRCM greatly outperforms the HadCM2, particularly because of a better representation of the upper tail of the distribution.

Feser (2006) uses a spatial filter to separate the RCM results into two spatial-scale ranges: for medium scales (between 550 and 250 km) and for large scales (larger than about 700 km). Results show that, when evaluating a large-scale quantity such as the sea level pressure, spatial patterns produced by the RCM are similar to those given by the NCEP/NCAR reanalyses. On the other hand, when assessing near-surface temperature in medium spatial scales, the RCM outperforms the NCEP reanalyses in the representation of the seasonal mean spatial patterns. The author suggests that the improvement is coming from the better representation of physiographic data in the RCM.

The main objective of this project is to contribute to the effort of objectively detecting the added value generated by regional models. As a step in this direction, we have evaluated the statistics of daily precipitation as simulated by a RCM and a GCM (the same used to drive the RCM) using observations in several regions across Canada. The analysis is concentrated

in those scales that are represented by both, the global and the regional model, but where the global model is expected to have little skill due to the fact that scales are near its truncation limit (Laprise, 2003; Feser, 2006). No attempt is made to explicitly evaluate the fine temporal-spatial scales produced by the RCM because, as we will show later, this must includes a different approach. Certainly, to get a more complete assessment, this work will have to be complemented by studies examining other aspects of added value, including fine spatial and temporal scale features, others variables, etc.

The work is organized as follows. In section 2 a brief review of some important issues that arise in the evaluation of added value in realistic frameworks is introduced. In section 3, observed and simulated data used in the comparisons are presented. The different statistics used to evaluate the performance of models are introduced in section 4. Section 5 includes results obtained when evaluating the annual cycle and daily values of precipitation. The article ends with a discussion and a summary of the results obtained.

2. Issues in the evaluation of added value

In this section some topics that we consider important when evaluating added value in RCMs are briefly discussed. Most of them are inherently present when studies of added value are carried out (e.g., resolved scales in numerical models, the poor knowledge of climate statistics in regional scales, etc.), while others are particular issues associated with the present work (e.g., representativity of precipitation data).

2.1 Availability of observed data

A very important limitation when evaluating the performance of RCMs is the spatial and temporal resolution for which observed data are available (Christensen et al., 2007; Laprise et al., 2007). For example, high-resolution analyses with reliable information in fine spatial scales are necessary to assess spatial variability of atmospheric fields. Observed gridded dataset with grid spacing on the order of 10-50 km exists for some variables (precipitation, temperature, sea level pressure) and for some regions around the world. Surface observing networks with sufficiently high density of stations are limited to specific regions (generally near the more densely populated areas) and the produced gridded datasets have fine-scale information only in those areas (some parts of Europe and North America).

Given these limitations, the assessment of temporal variability of RCM-simulated climate is an interesting alternative since reliable estimations can be made in local-regional scales. However, these evaluations are also limited due to lack of data with adequate fine temporal information.

2.2 Point observations vs. model-simulated precipitation

When comparing simulated and observed data, an assumption about the "representativity" of the different types of data must generally be made. For example, weather station data are usually believed to constitute point estimations of precipitation. On the contrary, it is generally accepted that climate models produce area-averaged estimations of precipitation (Osborn and Hulme, 1997; Cherubini et al., 2002; Frei et al., 2003; McSweeney, 2007; Perkins et al., 2007; Chen and Knutson, 2008). The reasoning behind this agreement is that many of the processes that produce precipitation within a climate model are parameterized (not explicitly resolved) and parameterizations of precipitation are generally implicitly areal in their implementation (Skelly and Henderson-Sellers, 1996), as is the case of mass-flux-based moist convection parameterizations (Chen and Knutson, 2008). Also, parameterizations are "tuned" to reproduce time- and area-average statistics, not the details of the time series of observational records. Differences in the spatial scales of both estimations may produce very distinct statistics. Spatial averaging, in a similar way to temporal averaging, tends to smooth prominent characteristics of the original (point source) time series, decreasing the frequency of occurrence of extreme events in both tails of the distribution, such as producing fewer dry days and fewer heavy precipitation events. In other words, temporal variability in weather station data is expected to be higher than that of simulated data, and the lower the resolution of the model, the more accentuated this difference should be.

An equitable comparison between simulated and observed data is then possible when at least one of the two estimates is processed so that both data are thought to represent similar spatio-temporal scales. The process of converting a grid-box average into point estimation requires the use of a downscaling technique. The quality of the simulated data is then influenced by the performance of the downscaling technique hence adding a new source of error. The other alternative is to convert the point measurement into an area- average quantity, an operation that is usually referred to as upscaling. A reliable estimation of the spatially averaged precipitation will be possible with a suitable number of stations that are able to correctly account for spatial variations within the region. As stated by McSweeney (2007), "the

number of stations required ... depends on the grid box size and shape, station distribution over the area, station distribution over time, and the spatial variability in the region". For example, the spatial variability of precipitation is greater in regions characterized by complex topography so that a large number of stations, appropriately distributed, is required to well describe the area-averaged rainfall. The spatial scale of precipitation systems is associated with the atmospheric circulation and so, as a first approximation, the number of stations could depend also on the weather regime and the season considered.

2.3 Effective resolution in climate models

An important issue when dealing with results from numerical models is related to the actual resolution that outputs are supposed to represent. A number of authors have discussed differences between effective resolution and grid spacing, generally in terms of the minimum length scale that is resolved by the grid-point numerical models (Pielke, 1991; Durran, 2000; Walters, 2000) and spectral models (Laprise, 1992). For example, Walters (2000) has defined a numerical model's effective resolution as "the minimum wavelength the model can describe with some required level of accuracy". Effective resolution is suggested to be greater or equal than four grid intervals and the level of accuracy depends on the geometric relationship between the numerical grid and the true solution, but also on the discretization schemes used and the spatial arrangement of variables on the computational grid. As stated by Walters (2000), this definition of effective resolution "can be related to both the spatial variation of the structure at a given time step and the behaviour of the amplitude of the error as a function of time". That is, the definition of effective resolution in most studies is understood as the effective resolution of instantaneous fields. But statistics used in climate studies are generally computed from temporal mean fields (daily, monthly, etc.), not directly from instantaneous values, and little is known on how to deal with effective resolution in these temporal scales.

In practice, the assumption of an effective resolution of four grid intervals involve producing spatial averages in regions of four by four grid points of the model output data,

which prevent the identification of added value at the finest represented scales. Because of this, studies generally consider directly single grid-point results, for example in RCMs (Räisänen and Joelsson, 2001; Kunkel et al., 2002; Gutowski et al., 2003; Diffenbaugh et al., 2005; Buonomo et al., 2007; Boberg et al., 2008) or in GCMs (Wang and Zwiers, 1999; Kharin and Zwiers, 2000; Kharin et al., 2007; Perkins et al., 2007).

2.4 Resolved spatial scales: RCM vs GCM

Differences between RCM- and GCM-resolved spatial scales could be conceptually seen in a diagram adapted from Laprise (2003) presented in Fig. 1. In this diagram, four grid intervals are considered as the effective resolution of both climate models; the GCM with horizontal grid spacing of 400 km and the RCM with 50 km. The full blue arrow indicates spatial scales that are resolved by the GCM (between 1 600 and 30 000 km) and the red one those resolved by the RCM (between 200 and 5 000 km). The dotted blue arrow denotes those spatial scales that are only represented by the GCM (G_S, larger than 5 000 km) because of the limited-area domain of the RCM and dotted red arrow denotes scales only represented by the RCM (R_{SS}, smaller than 1 600 km) because its higher horizontal resolution. Between both regions, there is an interval of wavelengths (between 1 600 and 5 000 km) that are represented by both models and its denoted by a full black arrow in the diagram.

There is a general consensus that the added value is mainly associated with those spatial scales at which the coarse resolution driving re-analysis system or global simulation model has little or no skill (e.g., Laprise et al., 2002; Feser, 2006; Laprise, 2006; Castro et al., 2005). Here, little (or no) skill refers to the ability of the GCM to represent scales that are poorly (or not) resolved by the GCM because they are near (below) the truncation limit of the model. Thus, the interval of wavelength where added value is present is suggested to correspond to those spatial scales in the R_{SS} region.

The evaluation of added value in region R_{SS} could be separated in two parts: (1) for wavelengths between 400 and 1 600 km (designated with a dashed black square in Fig. 1), where data from both models coexists and; (2) for wavelength smaller than 400 km where

only data from RCM exists. Part 1 allows the evaluation of added value in a direct way by comparing results from the two models with observations to determine which one produces a better performance. The analysis of Part 2 is quite different because only data from the RCM is present and a direct comparison between the GCM and the RCM is not possible. As was stated in the introduction, this work will concentrate in the added value generated by the RCM in part 1 of the spectrum.

From what has been said above, it is understood that the search for added value should proceed for variables and climate statistics whose variance is important in spatial (and temporal) scales that are expected to be better resolved by the RCM. Precipitation, one of the most important variables in climate studies, displays a wide range of spatial scales. The advantage of using this variable can be seen by just comparing power spectra of precipitation with that of any predominantly large-scale variable, such as geopotential or sea level pressure. While the former shows that variability is important at all scales, the second shows that variance in small scales is several orders of magnitude lower than in large scales (Separovic et al., 2008). Precipitation is also a key variable because some of the most important societal impacts of climate change will probably result from changes in precipitation (Gutowski et al., 2003; Iorio et al., 2004, Trenberth et al., 2003).

Spatial scale dependence of any variable is also strongly dependent on the climate statistics used for the analysis. The analysis of time-averaged fields is not the ideal method to identify the benefits of increased resolution. The variance of time-averaged fields is always more concentrated into larger spatial scales than the original time-varying fields. For example, Boer and Shepherd (1983) and Boer (1994) studied the scale-dependence behaviour of the vertically integrated rotational kinetic energy field when decomposed in its time-mean and transient eddy component. Their results show the dominance of the mean structures for small wavenumbers and of the transient component for wavenumbers beyond about 10 (wavelengths of approximately 3 000 km). Whatever the intensity and location of particular weather events, time averaging will always smooth out the most outstanding features. As a consequence, time-averaged RCM fields do not look substantially different from those produced by much coarser global models. Some important exceptions, however, do exist,

especially with respect to features associated with strong fine-scale surface forcing (e.g. complex topography).

3. Data

3.1 Observed data: meteorological stations

Observed data used in this study were provided by the National Climate Data and Information Archive, operated and maintained by Environment Canada. Over 6000 meteorological stations with daily total precipitation rate in the period 1971-1990 are available for the analysis. Total precipitation includes all types of precipitation: rain, drizzle, freezing drizzle, freezing rain, hail and snow. Any snow quantity registered is melted and its liquid water amount is recorded in millimeters and added to amounts from other forms of precipitation. Daily total precipitation is recorded in mm with a precision that varies between 0.1 and 0.2 mm depending on the instrument used. Graduated to the nearest 0.2 mm, the Type-B rain gauge is presently used to measure rainfall at most of the stations (Metcalfe et al., 1997). Weighing-type precipitation gauges and tipping bucket gauges are also used on a number of automatic meteorological stations to measure precipitation rates at shorter time intervals with a precision of 0.1 mm (Metcalfe et al., 1997; see also http://www.climate.weatheroffice.ec.gc.ca).

Observed data is subject to several sources of errors and uncertainties. Systematic measurement bias arises from wind undercatch and wetting-evaporation loss in point measurements (Groisman and Easterling, 1994; Metcalfe et al., 1997). The first one, due to the wind deflection of hydrometeors, is on the order of 10 % in summer and could arise 50 % in winter (more important for snow precipitation) according to the results obtained by Sevruk (1982) in U.S. stations. Wetting on internal walls of the collector and evaporation from the container produce also an underestimation of the "ground truth" precipitation that could attain 10 %. In this study, no corrections for any on these problems are performed in any of the different types of rain gauges.

Another source of error in measurements of daily precipitation comes from observer bias, that is, the tendency for the observer to favor or avoid some precipitation values compared to others (Daly et al., 2007). As is shown in the Appendix A, these errors are very important in the definition of the "effective precision" of the instruments.

3.2 Coupled Global Climate Model: CGCM 3.1

The global model used in this study is the third generation of the Canadian Centre for Climate Modelling and Analysis Coupled Global Climate Model (CGCM3), hereafter referred as CGCM. The use of this model is two-fold. First, their fields of horizontal velocity, temperature, surface pressure, specific humidity and sea surface temperature are used to provide the LBC to the Canadian Regional Climate Model (hereafter called CRCM). Second, its simulated precipitation is used to determine the existence of added value in CRCM simulations.

The CGCM makes use of the same ocean component as that used in the earlier version, but it makes use of a substantially updated atmospheric component (AGCM). The ocean component is described in detail in Flato and Boer (2001) and Kim et al. (2002, 2003), and its sea-ice component is described in Flato and Hibler (1992). The third-generation AGCM (McFarlane et al., 2005; Scinocca et al., 2008) shares many basic features with the second-generation version (McFarlane et al., 1992): the spectral transform method is used to represent the horizontal spatial structure of the main prognostic variables while the vertical representation is in terms of finite elements defined for a hybrid vertical coordinate as described by Laprise and Girard (1990). The spectral representation currently used in the AGCM corresponds to a higher horizontal resolution than that used in the earlier version, being comprised of a 47-wave triangularly truncated (T47) spherical harmonic expansion. The vertical domain of this atmospheric component extends from the surface to the stratopause region (1 hPa, approximately 50 km above the surface) with a total of 32 layers.

In version 3, the penetrative mass-flux scheme of Zhang and McFarlane (1995) is used to model deep cumulus convection. This scheme is based on a bulk representation for an ensemble of cumulus clouds comprised of entraining updrafts and evaporative driven downdrafts. The gridded output of precipitation occurs on a 96 by 48 Gaussian grid (output data has a grid spacing of 3.75° in latitude and longitude).

It is important to note that the minimum interval from which cumulative precipitation is available from the CGCM corresponds to 24 hours, which prevents extending the analysis to

sub daily temporal scales. This could constitute a severe limitation to our attempt at identifying the added value of the CRCM.

3.3 Regional Climate Model: CRCM 4.2.0

The regional model used for this study is the version 4.2.0 of the Canadian Regional Climate Model initially described in Caya and Laprise (1999) but upgraded through the use of subgrid-scale physical parameterization package of the third-generation CGCM (see section 3.2), except for the Bechtold-Kain-Fritsch (BKF) deep and shallow convection parameterization. The BKF scheme is also a bulk mass flux scheme (Arakawa and Schubert, 1974) that follows the precepts of Kain and Fristch (1990) for closure assumption and cloud model but is slightly different in the formulation of the trigger function. For derivation details see Bechtold et al. (2001).

The CRCM simulations were performed with horizontal grid spacing of 45 km (on a polar-stereographic projection true at 60° N) over a North American domain covering Canada, United States and most of Mexico, with a total of 201 by 193 grid points (see Fig. 2). In the vertical, 29 unequally spaced Gal-Chen scaled-height levels were used (Gal-Chen and Somerville, 1975); the lowest thermodynamic level is about 25 m above the surface, and the computational rigid lid was located near 29 km. The use of semi-Lagrangian and semi-implicit marching schemes allows the use of a 15-min time step at this resolution.

The CRCM was driven at its lateral boundaries by the traditional nesting of Davies (1976) as well as in its interior with large-scale nudging (Riette and Caya, 2002). The CRCM uses a spectral nudging technique that follows closely the approach developed by von Storch et al. (2000) but it uses a scale decomposition based on the Discrete Cosine Transform (Denis et al., 2002).

Two CRCM simulations were considered in the present investigation differing only in the LBC used as nesting data. One simulation is driven by the CGCM and will be designated as CRCM (CGCM). The other simulation is nested by the National Centers for Environmental

Prediction (NCEP) - National Center for Atmospheric Research (NCAR) reanalyses (Kalnay et al., 1996) and will be designated as CRCM (NCEP).

Table 1 summarizes some important information on the three simulations used in the study. Column 1 gives the acronym of each simulation used in the text. Column 2 gives the name of the model and its source. Column 3 denotes the type of model, the horizontal grid spacing/triangular truncation and the number of vertical levels. Finally, column 4 indicates the data used to drive the regional model in each simulation.

4. Methodology

The methodology used to investigate the presence of added value in CRCM simulations is based on the assessment of CRCM performance when compared to its driving model (CGCM) and observed data.

With the aim of studying the added value under different atmospheric circulations, climate statistics are computed for annual values but also for the different seasons defined in the usual way, namely: March, April and May (MAM), June, July and August (JJA), September, October and November (SON), and December, January and February (DJF).

Evaluation of the effective resolution of the global model is investigated by assessing two different spatial scales. This is carried out by considering spatial-average precipitation in two size regions: regions corresponding to one CGCM grid point and regions including four (i.e., 2 by 2) CGCM grid points.

4.1. Regions and period of study

With the aim of investigating the response of climate models under diverse surface forcing conditions such as complex topography, land-surface variations and land-sea contrasts, different regions are considered in the analysis. Availability of observed data imposes some restrictions on the possible areas that could be evaluated. According to Osborn and Hulme (1997), a minimum of 10 to 15 stations are necessary to accurately estimate the variance and the rain-day frequency in the CSIRO-AGCM grid-box. This result was derived using rain gauge data from United Kingdom but it is acknowledged that the number of stations could change depending on the relative importance of the surface forcings (McSweeney, 2007). In this study we have used the same criteria as Osborn and Hulme (1997), but as we will see later, the number of stations used within the regions of interest far exceeds this threshold, giving further confidence to our results.

Fig. 3 shows a map of Canada with all available stations that measured daily total precipitation in at least some part of the period 1971-1990. Because of the relatively good spatial coverage, 5 different zones across Canada are selected for the analysis. A total of 10 regions are considered, 5 including one CGCM grid point and 5 including four CGCM grid points, indicated in the Fig. 3 with blue and red boxes respectively. Boundaries in latitude and longitude for all regions are presented in Table 2. Each zone is designated with the abbreviation of the province to which it belongs: BC for British Columbia, ALTA for Alberta, SAS for Saskatchewan, MAN for Manitoba and QC for Quebec. The name of each region is then completed by adding a digit that establishes the number of CGCM grid points including in the region (e.g., BC.1 is the region that belongs to British Columbia which size is one CGCM grid-box). Fig. 4 shows each region separately with the topography in colour filled contours as represented in the CRCM. BC region (BC.1 and BC.4) is characterized by complex topography from the Rocky Mountains and also may be influenced by land-sea interactions near the Vancouver area. ALTA region is also characterized by complex topography but in the leeward side of the Rocky Mountains. SAS, MAN and QC regions show a relatively simple topography, although SAS could be influenced by topography of the Rocky Mountains due to its proximity and the strong western flow characteristic of mid-latitudes.

It is important to note that the horizontal distribution of stations within each region is not uniform. A more homogeneous distribution is generally found in regions including only one CGCM grid point than in those including four CGCM grid points. Also, for both sizes of regions, MAN and QC zones seems to show more heterogeneous horizontal distribution than the others zones. These differences could have an impact in the estimation of the spatial-average precipitation rate.

Fig. 5 shows the number of stations in each region as a function of time in the period 1961-1990, and Table 3 indicates minimum and mean values of the number of stations within each region. Regions defined by a single CGCM grid-box (four CGCM grid-box) include a minimum of 70 (151) rain gauges per day during the period 1971-1990, exceeding by a factor 5 the minimum proposed by Osborn and Hulme (1997). It is interesting to note that all regions present an annual cycle in the number of stations, with a minimum during winter and

a maximum in summer. Differences between both seasons are on the order of 15 stations except for ALTA region where it reaches approximately 80 stations. Interannual variations are also important and some regions indicate a significant increase in the density of stations from 1961 to 1971 and, for this reason, the period 1971-1990 is selected with the aim of increasing the confidence of regional daily precipitation estimations.

Table 3 also shows the number of model grid points within each region. The number of CRCM grid points within a single CGCM grid box is variable because of the use of the polar-stereographic conformal projection in the CRCM.

4.2. Temporal and spatial scales of analysis

As discussed in section 2.2, a direct comparison between an RCM, a GCM and observed values can be properly done only when quantities are equivalent for the three sources of data. By equivalent, here it is meant that the evaluation is carried out at scales that are greater or equal than that of the coarser resolution model or observation system. In our case, the CGCM defines the minimum area at which to perform the comparison and this corresponds to one CGCM grid box. The choice of the coarser resolution as the unit of comparison forces us to transform high-resolution data into lower resolution. Assuming that climate models produce area-average estimations of precipitation (see section 2.2), upscaling RCM and observed results to the GCM level simply consists in computing the spatial-average of all grid-points and stations data within each GCM grid box. Similarly, the fact that the GCM cumulative precipitation data was archived at 24-hour intervals forces to carry the comparison at this time scale, thus discarding shorter time interval information.

For each source of data, let us denote with $PR=PR_{ijr}$ the mean precipitation rate of the j^{th} day for the i^{th} point within the region r of interest. In our case, $j \in [1,J]$ with J=7300 since we consider 20 years between 1971 and 1990 with 365 values each. The spatial average for each region r, in day j, is simply computed as:

$$\langle PR \rangle_{jr} = \frac{1}{I} \sum_{i=1}^{I} PR_{ijr} . \qquad (1)$$

where the subscript i represents stations or model grid points in each region r. Values of I for each data set and region are presented in Table 3. In the case of simulated data, the total number of grid points within each region is constant with time (e.g., I = constant). On the other hand, for observed data, all stations available each day are included in the estimations of the area-averaged precipitation rate, and the total number of stations depends on the completeness of the archive (i.e., $I = I(j)$). The minimum and mean values of $I(j)$ during the period 1971-1990 for each region are presented in Table 3.

For observed data, as was discussed in section 2.2, the accuracy of the estimation of the area-average is related to the number of stations, their spatial distribution, weather regime, etc. In the case of the CRCM data, the upscaling is done directly because the nature of data (area-average output) from the CGCM and of the CRCM is the same. We only have to make sure that in the calculation of the eq. (1), CRCM grid boxes overlap the CGCM region.

4.3. Statistical analysis tools

As was discussed in section 2.4, time-averaged variables may not be the ideal method to evaluate model performance. For this reason, an assessment of daily statistics derived from the calculation of intensity frequency distributions is performed in addition to the study of monthly mean values.

4.3.1. Monthly mean values

Following the notation defined above, monthly mean values for each time series are calculated in the usual form:

$$\langle PR \rangle_{mr} = \frac{1}{20} \sum_{y=1971}^{1990} \langle PR \rangle_{myr}, \quad (2)$$

where $\langle PR \rangle_{myr}$ correspond to the mean precipitation rate of the month $m \in [1,12]$ for the year $y \in [1971,1990]$ in the region r. In the same way we can calculate seasonal mean precipitation rate values ($\langle PR \rangle_{syr}$) by averaging over 3-month periods.

Simulated seasonal errors are calculated as departures from observed seasonal mean values and then normalized by the observed values, so that results from different regions, seasons and weather regimes could be compared in terms of relative differences. Normalized seasonal deviations (*NSD*) could then be expressed as

$$NSD_{sr}^{mod} = \frac{\langle PR \rangle_{sr}^{mod} - \langle PR \rangle_{sr}^{obs}}{\langle PR \rangle_{sr}^{obs}}, \quad (3)$$

To determine if *NSD* exhibits substantial differences for a particular region or weather regimes, the average of absolute values of *NSD* across seasons and regions respectively is calculated. Absolute values ($ANSD = |NSD|$) are used in the calculation of mean values to avoid compensations between regions and seasons with *NSD* of different signs.

4.3.2. Estimation of precipitation intensity distributions

Intensity distributions of simulated daily precipitation have been used in several studies (e.g., Durman et al., 2001; Gutowski et al., 2003; Perkins et al., 2007; Boberg et al., 2007). In all of these studies, the bins width of precipitation used to construct the histograms were kept constant. In this work, we have chosen variable bin sizes that vary logarithmically in order to account for the reduction on the number of events with increasing intensity. Histograms for each time series are constructed from the frequency of occurrence of events function defined in the following way:

$$f_r(k) = \left\{ \begin{array}{ll} DD_r = \dfrac{\text{n° of } \langle PR \rangle_{jr} \in [0\ ,\ 1]}{7300} & \text{for } k=0 \\ WD_r(k) = \dfrac{\text{n° of } \langle PR \rangle_{jr} \in [2^{k-1}\ ,\ 2^k]}{7300} & \text{for } k=1,\ 2,\ldots,\ 9 \end{array} \right\}, \qquad (4)$$

where *DD* correspond to our adopted definition of "dry days" that includes those events with precipitation rates less than 1 mm/day. The function *WD* represents the "wet days" and includes those events with precipitation rates greater or equal than 1 mm/day. Thresholds of the categories that define the histogram are determined by the parameter k. A more detailed explanation of the construction of histograms as well as their associated error bars is given in Appendix B.

With the aim of obtaining an objective comparison between precipitation distributions, a simple measure defined in Perkins et al. (2007) is used. This score measures the overlap between observed and simulated distributions and is given by:

$$S_r^{\text{mod}} = \sum_{k=0}^{9} \min\left(f_r^{\text{mod}}(k), f_r^{obs}(k)\right). \qquad (5)$$

S varies between 0 indicating that no overlap exists between the distributions, and 1, when both distributions are identical.

As stated by Lettenmaier (1995), at daily time scales, precipitation amounts do not have a continuous probability distribution. Instead, there is a discontinuity in the probability of zero and non-zero amounts of precipitation. The estimation of the probability of zero amounts is particularly difficult when working with gauge measurements. The limited precision of instruments induces a threshold in the definition of "zero" precipitation values. As discussed in appendix A, the precision of precipitation measurements is suggested to be 0.8 mm/day, with a threshold of equal value. These imply that a good way to characterize the zero precipitation events is through the frequency of events with precipitation rates smaller than 1.0 mm/day. This minimum category is conceptually different from the "zero" or "dry day" category usually used because it includes some small precipitation amounts. Notwithstanding

this discrepancy, we will still name the minimum category as "dry days" and the rest of the distribution (values greater or equal than 1.0 mm/day) as "wet days".

To complete the analysis of the performance of simulated frequency distributions accomplished through the *S* score, a separate study of dry and wet days is performed. In the next subsections, statistics used to evaluate these two different parts of the distribution are presented.

4.3.3. Analysis of dry days

To study the performance of the simulation of dry days, the ratio between simulated and observed values is computed as

$$R_{dd}^{\mathrm{mod}} = \frac{f^{\mathrm{mod}}(k=0)}{f^{obs}(k=0)} = \frac{DD^{\mathrm{mod}}}{DD^{obs}}. \qquad (6)$$

Defined in this way, $R_{dd} = 1$ when there is a prefect agreement between observed and simulated dry days, $R_{dd} > 1$ suggest an overestimation and $R_{dd} < 1$ an underestimation of the simulated frequency of dry days.

4.3.4. Analysis of heavier precipitation rate events

To study the performance of the simulations in the representation of heavier precipitation intensity events, values of 95th percentile are estimated. The 95-percentile is the minimum precipitation rate that exceeds the *95* % of the data. In this study, percentiles are computed from the wet day variable defined in equation (4).

5. Results

5.1. Monthly mean precipitation rate

5.1.1. The case with one CGCM grid point

Fig. 6 shows monthly mean precipitation rates observed and simulated by the CGCM and CRCM in the five regions defined by a single CGCM grid point (see Fig. 3 and Table 2 for boundary specifications). CRCM simulated precipitation is shown for the two simulations (see section 3.3 and Table 1) differing only in its LBC: one using the CGCM (CRCM (CGCM)) and the other using the NCEP/NCAR reanalyses (CRCM (NCEP)). Each region displays different climatic regimes evidenced by their individual annual cycle. BC.1 region, west of the Rocky Mountains and characterized by the presence of complex topography, shows a very wet winter with a relatively drier summer. ALTA.1, SAS.1 and MAN.1 regions, located east of the mountains, show moderately dry regimes with a maximum in summertime precipitation. QC.1 region shows a rather uniform monthly distribution of precipitation, with a maximum between summer and autumn seasons.

The observed annual cycle is generally well simulated in all regions by both models but differences between simulated and observed monthly values are noticeable. The most outstanding feature is a systematic over-estimation of monthly mean values, found in almost all regions and seasons. Fig. 7 shows normalized deviations, in all seasons and regions, for a) the CGCM, b) the CRCM driven by the CGCM and c) the CRCM driven by the NCEP/NCAR reanalyses. The overestimation is clear in results from both models with the exception of QC.1 region, particularly during winter. It is not surprising that ALTA.1 region shows the greatest normalized deviations in winter season because of its very low precipitation. For example, in DJF in ALTA.1, the difference between the observed and the CRCM (CGCM) simulated precipitation is less than 0.6 mm/day, which is not very large but of the same order that the observed precipitation.

With the aim of investigating whether deviations are related systematically to some particular model or season, normalized absolute deviations between observed and simulated

seasonal values, averaged across regions, are presented in Fig. 8a. Mean relative errors vary between 5 and 35 % and are of the same order of magnitude in all seasons, with minimum values in spring. Differences between seasonal values simulated by the CGCM and the CRCM (CGCM) are quite evident. However, the existence of added value is difficult to confirm since a systematic improvement of seasonal values is not found in the CRCM. For example, the inter-regional mean (normalized) deviations in SON vary from 35 % in the CGCM to less than 5 % in the CRCM (CGCM). In MAM on the other hand, the CRCM (CGCM) simulated precipitation produces a larger seasonal mean error. Annual errors of CRCM (CGCM) are smaller due to a large overestimation of precipitation in autumn season in CGCM simulations.

CRCM-simulated precipitation depends strongly on the driving data, showing important differences between seasonal mean values when using NCEP reanalyses or the CGCM as nesting data. It is interesting to note that seasonal values of precipitation simulated by the regional model when using the more realistic LBC from reanalyses does not systematically produce better simulation of the precipitation than when using LBC from the CGCM.

Fig. 8b shows normalized absolute deviations of simulated annual values for all models as a function of region. Differences between regions are sometimes important but generally of the same order as the inter-model differences. The exception is ALTA.1 that shows the largest errors (~ 30 %) with little dispersion between models. BC.1, SAS.1 and MAN.1 regions exhibit a better performance of the CRCM (for both LBC) to simulate annual values compared with those produced by the CGCM. On the other hand, QC.1 region seems to show a slightly better representation of annual precipitation in the CGCM compared to the CRCM.

5.1.2. The case with four CGCM grid points

Fig. 9 shows monthly mean precipitation rates observed and simulated for the five regions including four CGCM grid points (see Fig. 3 and Table 2 for boundary specifications). BC.4 and ALTA.4 regions, with important topographic features, tend to display slightly different climatic regimes compare to the corresponding smaller regions

shown in Fig. 6. For example, BC.4 region shows a decrease of the amplitude of the annual cycle when compared to BC.1, probably due to the influence of the different surface forcings affecting each sub region. The small region is located almost entirely west of the Rocky Mountains (see Fig. 4) and is greatly influenced by the topography-induced precipitation associated with the forced upward movement of large-scale west flow. Region BC.4 includes areas on the leeward side of the mountains with the precipitation shadowing effect of mountains ranges, producing a decrease in the spatial-average precipitation. The opposite effect (inclusion of areas with mountains forced precipitation) is probably causing the increment in spatial-average precipitation in ALTA.4 region when compared to ALTA.1. Regions SAS.4, MAN.4 and QC.4, with very little influence of the topography, show no significant differences compared to their respective smaller regions.

The performance of the CGCM and the CRCM in the five regions defined by four CGCM grid points are quite similar to that of a single CGCM grid boxes. Fig. 10 shows seasonal normalized deviations for a) the CGCM, b) the CRCM (CGCM) and c) CRCM (NCEP). Deviations are of the same order of magnitude as in the case of regions including one CGCM grid point but over-estimations are less important than in those regions, especially in the CRCM simulated precipitation.

As in the case of single CGCM grid box regions, there is a better simulation of annual precipitation rates in the CRCM (for both LBC) than in the CGCM, this difference coming particularly from larger errors in CGCM-simulated precipitation in SON and JJA (see Fig. 11a and Fig. 11b). BC.4 and SAS.4 regions indicate the largest improvements on annual values with normalized deviations changing from 25-30 % in the CGCM to 10-15 % in the CRCM.

5.2. Intensity frequency distributions of daily precipitation rate

5.2.1 The case with one CGCM grid point

As in the case of the annual cycles, each region displays distinct distributions of daily precipitation intensities. Fig. 12 shows observed and simulated daily precipitation intensity distribution in winter season for the five regions including a single CGCM grid point. The distribution for the very humid winter characteristic of BC.1 region shows a broad range of intensities where daily precipitation occurs. Regions ALTA.1, SAS.1 and MAN.1, located in the transition dry zone on the downstream slopes of the Rockies, show a narrow range of daily precipitation rates with more than 75 % of observed dry days. QC.1 region, with moderate precipitation in winter (mainly as snow), shows a wide range of intensities of precipitation with fewer heavy precipitation events than BC.1 region.

As was explained in section 4.2.2, the degree of overlapping between simulated and observed distributions can be quantified by using the S score (see equation (5)). Fig. 14 displays S values calculated for each season as a function of regions and models. In winter, differences in distributions across regions are well reproduced by both models but some difficulties arise in the representation of the lower and upper limits of distributions in regions QC.1 and ALTA.1, showing an underestimation of both, dry days and more extreme events. In SAS.1 and MAN.1 regions, S score is greater than 0.93 for the three simulations analyzed but it is important to stress that in these two regions, S describes mainly the simulation of non-precipitation events (observed dry days are more than 80 %). Finally, both models simulate very well the precipitation distribution in winter in QC.1, with a slightly over-estimation of dry days and underestimation of moderate precipitation rates (those between 4 and 16 mm/day). Table 4 summarizes seasonal results by showing the inter-region mean S score for the three simulations analyzed. In winter, values of the mean S score are 0.92 for the CGCM and 0.91 for the CRCM (for both LBC).

Based on these results, it is difficult to establish whether a model performed better than the other. While in regions SAS.1, MAN.1 and QC.1 both models performed similarly, some

differences arise in the others regions, particularly between CRCM simulations for different LBC.

The representation of precipitation intensity distribution in summer season (Fig. 13) seems to be a more difficult task for both models, and systematic differences are noted in all regions, with maybe the exception of ALTA.1 that maintains a similar behaviour to that of DJF. Mean values of S across regions (Table 4) are 0.85 for the CGCM, 0.76 for the CRCM (CGCM) and 0.82 for the CRCM (NCEP), showing not only a worse performance than in winter but also greater differences between the simulations analyzed.

Simulated daily precipitation by the CGCM consistently outperforms that from the CRCM (for both LBC) in almost all regions. Also interesting is that in all regions, the CRCM when driving by NCEP reanalyses performed better than the same model nested by the CGCM, suggesting that LBC are playing an important role in summer statistics. It is also difficult to establish whether any of the regions are more challenging but it seems that models perform poorly in BC.1 region.

In all regions and seasons but particularly in summer, both simulated distributions show the same pattern of deviations compared to observed ones: an underestimation of dry days associated with an overestimation of light precipitation rate events. To confirm that the source of differences between simulated and observed distributions comes from the different representation of dry days, Fig. 15 shows the ratio between simulated and observed dry days (R_{dd}) as a function of region for a) DJF, b) MAM, c) JJA and d) SON. The dependence of S with R_{dd} is evident: a good simulation of observed dry days ($R_{dd} \sim 1$) is associated with a high degree of overlapping ($S \sim 1$) and the opposite is found when the simulation of dry days is poor. This means that differences between simulated and observed distributions are dominated by differences in the representation of dry days. The greater differences in summer season are related to important underestimations of dry days, attaining a ratio R_{dd} of 0.4 in the CRCM (CGCM) in some regions.

Interestingly, in winter season, QC.1 region present ratios slightly larger than 1 for both models, indicating that simulated events with precipitation rate smaller than 1 mm/day are more frequent than in observations.

5.2.2 The case with four CGCM grid points

Similar results are obtained when regions including 4 CGCM grid points rather than one-grid box regions are analyzed (see Fig. 16 and Fig. 17). Values of the S score are of the same magnitude and higher skills are found in winter season compared to summer one. In winter, inter-regions mean S score is 0.92 for both the CGCM and the CRCM (CGCM) while in summer values are 0.81 for the CGCM and 0.74 for the CRCM (CGCM).

Similar to results from BC.1 region, BC.4 shows systematically the lowest values of S, for both models, when compared to the other regions. The difficulties of the models to represent the precipitation in this region may be related to the presence of complex topography.

5.3 Heavier precipitation rates

The 95th percentile is used to study the performance of models to reproduce those events with the highest rates of precipitation in each season and region. Fig. 18 shows the 95th percentile calculated for each simulation and season in regions including one CGCM grid point. As a first approximation, values of 95th percentile could be related with the total seasonal precipitation, showing higher values in JJA (DJF) for regions ALTA.1, SAS.1, MAN.1 and QC.1 (BC.1 region) because of its maximum summertime (wintertime) precipitation. However influences of other forcings, such as the presence of complex topography, could be important in the frequency and intensity of more extreme events. For example, in JJA, the highest 95th observed percentiles correspond to QC.1 (19.4 mm/day) and BC.1 (17.0 mm/day) regions, the latter with the minimum total summer precipitation between all regions. The relative greater importance of heavy precipitation rate events in the total summer precipitation of BC.1 region is probably related to the presence of complex topography. The precipitation generation type could influence the frequency and intensity of heavier events, in a similar way that the presence of topographical forcing. However, due to differences in the

total seasonal precipitation, it is difficult to determine whether the dominance of convection in summer produces distinct statistics of heavier precipitation events than the large-scale precipitation in winter. It seems that, at daily time scales, convection does not generate heavier precipitation events than large-scale precipitation.

Differences in the 95th percentile values across regions are generally well simulated by both models in all seasons. For example, in winter season, observed and simulated percentiles reach maximum values in BC.1, minimum values in MAN.1 and show a secondary maximum in QC.1 region. However, a quantitative analysis shows some important differences between simulated and observed percentiles as well as between the different simulations.

The most outstanding feature is the consistent underestimation of CRCM-simulated 95th percentile when compared to observed values. This underestimation is seen in all regions and seasons (more important in JJA and SON) whatever the LBC used to drive the CRCM. The CGCM produces generally a better agreement with observed 95th percentiles than the CRCM in all seasons, showing a slightly underestimation in BC.1 region and an overestimation in QC.1 region.

The influence of surface forcings in BC.1 is not well simulated by models that show an important underestimation of heavier events. The CRCM expected to better resolve topographical features and associated precipitation, shows similar results to the CGCM.

Differences between the CRCM simulation when driven by CGCM and NCEP/NCAR reanalyses are generally small and it is difficult to establish if one produces a better performance than the other.

Values of 95th percentiles for regions including 4 CGCM grid points (Fig. 19) are smaller than the one-grid box regions because of the larger area used in the average. However, the relative performance of climate models in reproducing observed percentiles is similar to those in one-CGCM grid box regions.

6. Summary and Discussion

The primary aim of this study was to investigate the existence of added value in CRCM simulations used to downscale CGCM-simulated fields. As a first but necessary step, temporal and spatial scales that are common to both, the CGCM and the CRCM, are considered in the analysis. The comparison is performed by upscaling, at the CGCM level, data from the CRCM and from meteorological stations.

Our evaluation is focused on daily precipitation as simulated by the CRCM and the CGCM using as reference daily observational time series. Temporal series of spatially-averaged observed precipitation are constructed from meteorological stations operated by Environment Canada. Because of their relatively good spatial coverage, five different zones across Canada are selected for the analysis, all exceeding by a factor five the minimum number of stations proposed by Osborn and Hulme (1997) when computing area-average estimations.

The assessment is based on the comparison of simulated and observed intensity frequency distributions of daily precipitation and on the computation of 95^{th} percentile of the distributions to characterize more extreme events. The S score defined in Perkins et al. (2007) is used to measure the overlap between simulated and observed daily distributions and reflects mainly the simulation of light-moderate precipitation rates. Considering different regions in the analysis allows the evaluation of added value as a function of surface forcings. In the same way, the dependence in weather regimes is studied by analysing seasonal statistics.

Results show that the daily statistics of precipitation as simulated by the CGCM and by the CRCM are generally very similar and, when comparing both data, there is no evidence of the existence of added value in CRCM simulations. For example, in winter season, the CGCM and the CRCM display similar skills to simulate the frequency and intensity of observed daily values by showing similar values of the S score, independently of region considered.

In summer season, both models have more difficulties than in winter season to reproduce the observed daily distribution, presenting smaller values of S than in wintertime. As suggested by the S score, the CGCM shows a better agreement with observed data than the CRCM and this improvement is coming from a better simulation of the frequency of observed dry days.

Representation of dry days is a recurrent problem in climate models (Trenberth et al., 2003), generally associated with an overestimation of the frequency of light intensity precipitation events (Dai et al., 1999; Paquin et al. 2002; Frei et al., 2003; Dai and Trenberth, 2004). All simulations presented in this study show an underestimation of dry days and an overestimation of light precipitation (underestimation/overestimation pattern), almost independently of the seasons and region considered. The only exception to this behaviour is winter season in Quebec region. Although both models underestimate the frequency of dry days, the CRCM seems to produce a pattern of underestimation/overestimation still more pronounced than the CGCM. A comparison of the simulated distributions of precipitation between the CRCM and others state-of-the-art regional climate models adds some evidence supporting this statement. Fig. 20 (taken from Christensen et al. (2008)) shows histograms of precipitation as simulated by 14 regional climate models in a domain that includes all Europe. The CRCM (identified as OURANOS_CRCM in Fig. 20) produces more light precipitation (and also less moderate-heavy precipitation) events than any other regional models considered in the evaluation. .

In the case of more extreme daily values, results show that the CGCM produces a better agreement with observed values than the CRCM. The latter shows a consistent underestimation of the frequency of occurrence of heavy precipitation events. This is even the case in regions characterized by important surface forcings, where differences between model topographies may be expected to have an impact.

More than a problem of relative performances of models, the absence of added value might be related to the failure of the assumptions from which added value could be expected. The hypothesis that support the assumption of existence of added value in our particular comparison was presented in section 1 and is given by:

Global model should have less skill at their smallest resolved scales due to the fact that these scales are near the truncation limit (Laprise, 2003; Feser, 2006).

As was discussed in section 2.3, this hypothesis has been studied generally for instantaneous fluctuating quantities but not for time-average quantities (e.g., daily and monthly values) and for discontinuous variables (e.g., precipitation). In this study, the hypothesis has been tested by comparing results obtained in regions including one and four CGCM grid points. The skill

of the global model seems to be similar in the two sizes of regions analyzed, suggesting that the model is "working well" in both spatial scales. In other words, because of the very good performance of the global model, independently of the size of regions considered, the hypothesis does not seem to be appropriate at least for the time scales here analyzed.

To avoid a misinterpretation of our results, it may be interesting to briefly discuss two particular features of the methodology. Both are linked to the partial character of the model evaluation:

- The approach employed in this study is based on the assessment of spatial and temporal scales of precipitation that are represented by the two models but which are near the truncation limit of the global model. Advantages of the RCM simulations due to its higher spatial-temporal resolution are not explicitly considered and, clearly, this is not the ideal way to highlight the benefits of the increased resolution in RCMs. For example, the minimum interval from which cumulative precipitation of 24 hours imposed by the CGCM could filter out some added value existing in smaller temporal scales. This characteristic of the study must be taken into account at the time of evaluating these results.
- The relative performance of some statistics of precipitation as simulated by both models with respect to observational data is used as a tool to detect added value. This approach has its downside. As stated by Oreskes et al. (1994), "If a model fails to reproduce observed data, then we know that the model is faulty in some way, but the reverse is never the case".

In other words, if a model A produces results that are closer to observed data than those of model B, these does not necessarily imply that model A is better than B. A better performance could be attained by compensation and not necessarily because of the right reasons. An example that attempts to illustrates this latter possibility is the LBC paradox shown in section 5: results produced by the CRCM when driven by the global model are sometimes closer to observed values than those using NCEP/NCAR reanalysis as LBC. Although NCEP/NCAR reanalyses are subject to errors, there is no doubt that they constitute a more reliable estimation of an evolving state of the atmosphere than GCM simulated data. So, how can it be possible to obtain "better" results when using worse LBC? Necessarily, the

errors in the LBC must be compensated by the CRCM, suggesting that the better results are coming from errors in the CRCM.

Analysing the added value generated by RCM is a complex issue. We have tried to contribute to this important subject in two different ways: a) by quantifying added value in a particular case and; b) by discussing some general issues that must be taken into account when conducting this kind of analysis. Among the several important questions that still remain open we can include:

- What is the effective resolution of climate models? What is the dependence of effective resolution on the time scale of the variable analyzed?
- RCMs present some advantages with respect to the coarser resolution GCMs such as a better representation of surface forcings, a greater range of processes resolved, etc. What is the relative importance of each advantage in producing added value?
- What is the dependence of added value on temporal and spatial scales?
- Is there any added value in large-scale results?

Finally, it could be very interesting to repeat the analysis with others models such as those involved in the European project PRUDENCE (Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects; http://prudence.dmi.dk/) or/and in the North American Regional Climate Change Assessment Program (NARCAAP; http://www.narccap.ucar.edu/; Mearns et al., 2005). The use of several models would help to determine whether some of the findings are inherent to the downscaling technique or related with a particular model.

Appendix A: Precision of observed data

As suggested by Osborn and Hulme (1997), the minimum precipitation measure (0.2-0.25 mm) could be used to separate periods with zero precipitation (e.g., dry day) and non-zero precipitation (e.g., wet day) at each station. Implicit in this definition of dry days is the assumption that precision in observed precipitation is the same as that of instruments. However, observer bias (by definition the tendency of the observer to favor or avoid some precipitation values compared to others) could be an important problem when defining the precision of observed data. This was shown by Daly et al. (2007) using daily precipitation measurements in the United States Cooperative Network (COOP). They found that there are two major types of observer biases: 1) underreporting amounts of less than 0.05 in, and 2) over-reporting of daily precipitation amounts evenly divisible by 5 and/or 10. They also suggest that underreporting of light intensities (less than 0.05 in) is associated with an unusual high frequency of zero amounts, possibly as a result of higher thresholds for inconsequential precipitation for some observer than others. Although observer biases discussed in Daly et al. (2007) refer to the COOP data measure system, characterized by measures in English units, some similar behaviour may be expected independently of the SI system used.

Fig. A1 shows observed precipitation intensity distributions of daily precipitation rate for the five regions indicated in blue in Fig. 3 (see Table 2 for boundaries specifications). Histograms are constructed with four different bin sizes: 0.1, 0.2, 0.4 and 0.8 mm per day. Constant bin sizes are used because only intensities less or equal 5 mm/day are shown in histograms and the number of events are relatively homogeneous at these rates. With the aim of investigating whether biases are widespread, histograms are constructed including all available data in each region between 1971 and 1990. The total number of values in each histogram depends on the number of stations in each region and on the completeness of each time series. Table A1 presents the mean and maximum number of stations in each region during the period 1971-1990 and the resulting total number of data. Also included is the equivalent number of years of a daily time series built assuming that all data are from the same site.

As stated by Daly et al. (2007), "the more precipitation events included, the smoother the appearance of the frequency histogram; at least 10 years of data are typically required to obtain a smooth histogram at an unbiased site, and to provide enough frequency counts in various precipitation bins to produce stable statistical results". In the present work we count with the equivalent of over 1400 yr of data, which suggests that the statistical robustness of the histogram is beyond question.

For the 0.1 mm/day bin size histogram (black line in Fig. A1), relative maxima associated with intensities of integer values (1, 2, etc. mm/day) are noticeable in all regions, suggesting an observer bias similar to type 2 of Daly et al. (2007). Another interesting signature is that there are no observations of events with an intensity of 0.1 mm/day in regions 2, 4 and 5, and negligible numbers in the other two regions, indicating a possible increase in zero thresholds for some observers, in accordance with Daly et al. (2007) results.

When using 0.2 mm/day as the interval value for constructing histograms, results are less irregular but still large departures from smoothness are observed. For example, when analyzing two stations with no visible observer bias (both passed all tests) and with very different precipitation regimes, Daly et al. (2007) find a similar behaviour in the distributions. Both stations exhibit a maximum frequency at 0.01 in (0.25 mm), with a relatively smooth decrease in frequency of occurrence as the daily precipitation amount increases (see Fig. A.2). Following this results and its apparent validity under different conditions, one should expect similar behaviour in the histograms presented in Fig. A1. However, all regions show that the number of events with precipitation rate in [0.2, 0.4] mm/day is smaller than the amount in [0.4, 0.6] mm/day. This is probably related to error of type 1 of Daly et al. (2007), presenting an overestimation of the dry days classification.

Accumulated results suggest that the use of the minimum precipitation measure (0.2 mm/day) as the definition of dry days could be an important source of error when comparing against simulated dry days. To avoid this potential error, a new "effective precision" will be considered. From Fig. A1, irregularity in the data seems to be important until a resolution of four times the precision is used and a bin width of 0.8 mm/day is chosen as the minimum interval from which observed precipitation is well resolved.

It is important to note that due to the variable bin sizes used to construct histograms (see Appendix B), this change in the precision of observed data only modifies the first classification (selection of dry days threshold). So, if we consider 0.8 mm/day as the resolution of observed daily precipitation, the definition of the minimum category correspond to the daily precipitation rate that is less or equal 0.8 mm/day.

When combining a number of stations that fall in a CGCM grid box, we define the dry day when the n-station average of each day is dry (e.g., the n-station mean is less than the threshold).

Appendix B: Construction of histograms and error sampling

Following von Storch and Zwiers (1999), to obtain a frequency histogram the interval of possible outcomes of daily precipitation rate is initially partitioned into h subsets Ω_h of 0.2 mm/day width, starting at 0.8 mm/day, and then divided by the total number of outcomes to obtain the frequency in each bin. Following the notation of section 4, the frequency histogram variable could be expressed as,

$$f(h) = \frac{\text{n° of } \langle PR \rangle_j \in \left[0.8 + h \cdot 0.2 \ , \ 0.8 + (h+1) \cdot 0.2\right]}{\text{total n° of } \langle PR \rangle_j} = \frac{m_h}{\mathrm{N}}, \qquad \text{(B.1)}$$

where h varies between 0 and 1750, m_h is the number of events in the classification h and N represent the total number of values in the temporal series. 1750 is chosen as the maximum h, giving a maximum precipitation rate of 350.8 mm/day, because simulated and observed values are always less or equal than 327 mm/day.

Due to decreasing daily precipitation events for increasing intensity rates, the resulting plot may behave irregular, depending on the choosing bin size and the total number of events. Decisions about the bin size also have consequences on the error made when estimating the frequency of occurrence of a particular event. This sampling error is inversely proportional to the number of events in the given category and, because of the general form of daily precipitation distribution, the error sampling is larger for more rare events when considering constant bin sizes as in eq. *B.1*. To account for the reduction on the occurrence of events when increasing intensity rates, a logarithmic scale is considered to determine the threshold used in each class. The histogram function at constant bin size is then used to construct a new frequency histogram where the interval of possible outcomes is partitioned into K subsets Ω_k. Thresholds of Ω_k are given by 1, 2, 4, 8, 16, 32, 64, 128, 256 and 512 mm/day and intervals for each classification is defined without including events which values are exactly the threshold. This function was expressed as *WD(k)* in eq. *(4)* and, when adding the dry days category, we obtain the histogram function *f(k)* (see eq. (4)). When plotting the histograms, frequencies are indicated in percentage using a logarithmic scale in base ten for ease of presentation.

In order to add error bars to the histograms of precipitation, confidence intervals could be constructed around the estimated frequency of each classification. These are intervals constructed to be wide enough to contain, with a specific probability, the population quantity corresponding to the sample statistic. To build the confidence intervals, we consider for each classification k, the variable m defined as the number of times that daily precipitation rate ($\langle PR \rangle_j$) is included in the interval width Ω_k during the period 1971-1990. In this way, the variable takes the value 1, when $\langle PR \rangle_j \in \Omega_k$, and 0 when $\langle PR \rangle_j \notin \Omega_k$. This random variable has a binomial distribution, where one or the other of two events occurs, and its distribution function B is given by,

$$B(m) = \left(\frac{N}{m}\right) \cdot f_k{}^m \cdot (1 - f_k)^{N-m}, \tag{B.3}$$

where N is the total number of events and f_k the frequency of occurrence of a given classification. The mean and variance of the binomial variable m are given by

$$\begin{aligned} \mu(m) &= N \cdot f_k \\ \sigma(m) &= N \cdot f_k \cdot (1 - f_k) \end{aligned} \tag{B.4}$$

If N is sufficiently large, the variable m will follow approximately the Gaussian distribution. Different criteria are used to determine when N is sufficiently large. Wilks (1995) suggest that, for a given m, N must satisfy the following criterion

$$0 < f_k \pm 2\sqrt{\frac{f_k(1 - f_k)}{N}} < 1 \tag{B.5}$$

Chalmer's criterion says that both $(1\text{-}f_k)N$ and f_kN must be greater than 5 and so there must be at least 5 occurrences in each interval for the Gaussian approximation to be valid.

So, considering that m is greater than 5 and that is normally distributed, the variable defined by

$$z = \frac{\sqrt{N}(m - \mu)}{\sigma}, \tag{B.6}$$

has a normal distribution with mean 0 and standard deviation equal 1. Because the real value of the variance is not known (we have an estimation S), confidence intervals are given by

$$P\left(m - t_U \frac{S}{\sqrt{N}} < \mu < m + t_U \frac{S}{\sqrt{N}}\right) = 1 - p' \tag{B.7}$$

where S is the estimated variance and t_U is the $p = 1 - \frac{p'}{2}$ th percentile of the probability distribution. For a 90 % confidence interval, p' is chosen to be 0.05 (two tails test) and $t_U = t_{(0.05)}$ with *(n-1)* degrees of freedom.

References

Antic, S., R. Laprise, B. Denis and R. de Elía, 2004: Testing the downscaling ability of a one-way nested regional climate model in regions of complex topography. *Clim. Dyn.*, 23, 473-493.

Arakawa, A., and W.H. Schubert, 1974. Interaction of a cumulus cloud ensemble with the large-scale environment: Part I. *J. Atmos. Sci.*, 31, 674-701.

Bechtold, P., E. Bazile, F. Guichard, P. Mascart and E. Richard, 2001: A Mass Flux Convection Scheme for Regional and Global Models. *Quart. J. Roy. Meteor. Soc.*, 127, 869-886.

Boberg, F., P. Berg, P. Thejll, and J.H. Christensen, 2007. Analysis of temporal changes in precipitation intensities using PRUDENCE data. Danish Climate Centre Report, 07-03, 43pp [Available from http://www.dmi.dk].

Boer, G.J., and T.G. Shepherd, 1983: Large-scale two-dimensional turbulence in the atmosphere. *J. Atmos. Sci.*, 40, 164-184.

Boer, G.J., and M. Lazare, 1988: Some results concerning the effect of horizontal resolution and gravity-wave drag on simulated climate. *J. Climate*, 1, 789–806.

Boer, G.J., 1994: Mean and transient spectral energy and enstrophy budgets. *J. Atmos. Sci.*, 51, 1765-1779.

Boville, B.A., 1991: Sensitivity of simulated climate to model resolution. *J. Climate*, 4, 469-485.

Boyle, J.S., 1993: Sensitivity to dynamical quantities to horizontal resolution for a climate simulation using the ECMWF (cycle 33) Model. *J. Climate*, 6, 796-815.

Castro, C.L., R.A. Pielke, Sr., and G. Leoncini, 2005: Dynamical Downscaling: An Assessment of Value Added Using a Regional Climate Model. *J. Geophys. Res.-Atmospheres*, 110, D05108, doi: 10.1029/2004JD004721.

Caya, D., and R. Laprise, 1999: A semi-implicit semi-Lagrangian regional climate model: The Canadian RCM. *Mon. Wea. Rev.*, 127 (3), 341-362.

Chen, C., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *J. Climate*, 21(7), 1605-1621.

Cherubini, T., A. Ghelli, and F. Lalaurette, 2002. Verification of Precipitation Forecasts over the Alpine Region Using a High-Density Observing Network. *J. Weather and Forecasting*, 17, 238-249.

Christensen, J.H., B. Hewitson, A. Busuioc, A. Chen, X. Gao, I. Held, R. Jones, R.K. Kolli, W.-T. Kwon, R. Laprise, V. Magaña Rueda, L. Mearns, C.G. Menéndez, J. Räisänen, A. Rinke, A. Sarr, and P. Whetton, 2007: Regional Climate Projections. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)].

Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press: 847-940 (http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter11.pdf).

Christensen, J.H., E. Sanchez, and E. Kjellström, 2008: Probability density distribution match of daily and monthly temperature and precipitation analysis (f_3). Deliver 3.2.2: RCM-specific weights based on their ability to simulate the present climate, calibrated for the ERA40-based simulations, project ENSEMBLE-based Predictions of Climate Changes and their Impacts.

Dai, A., F. Giorgi, and K.E. Trenberth, 1999: Observed and model simulated precipitation diurnal cycle over the contiguous United States. *J. Geophys. Res.*, 104, 6377-6402.

Dai, A., and K.E. Trenberth, 2004: The diurnal cycle and its depiction in the Community Climate System Model. *J. Climate*, 17, 930-951.

Daly, C., W.P. Gibson, G.H. Taylor, M.K. Doggett, and J.I. Smith, 2007. Observer Bias in Daily Precipitation Measurements at United States Cooperative Network Stations. *Bull. Amer. Meteor. Soc.*, 88, 6, 899-912.

Davies, H.C., 1976: A lateral boundary formulation for multi-level prediction models. *Quart. J. Meteor. Soc.*, 102, 405-418.

de Elía, R., R. Laprise, and B. Denis, 2002: Forecasting Skill Limits of Nested, Limited-Area Models: A Perfect Model Approach. *Mon. Wea. Rev.*, 130, 2006-2023.

Denis, B., J. Côté, and R. Laprise, 2002a: Spectral Decomposition of Two-Dimensional Atmospheric Fields on Limited-Area Domains Using the Discrete Cosine Transform (DCT). *Mon. Wea. Rev.*, 130, 1812-1829.

Denis, B., R. Laprise, D. Caya, and J. Cote, 2002b: Downscaling Ability of One-Way Regional Climate Models: The Big-Brother Experiment, *Clim. Dyn.*, 18, 627-646.

Dickinson, R.E., R.M. Errico, F. Giorgi, and G.T. Bates, 1989: A regional climate model for the western United States. *Clim. Change*, 15, 383-422.

Diffenbaugh, N.S., J.S. Pal, R.J. Trapp, and F. Giorgi, 2005: Fine-scale processes regulate the response of extreme events to global climate change," *Proc. Nat. Acad. Sci.*, 102, no. 44, 15774-15778.

Dimitrijevic, M., and R. Laprise, 2005: Validation of the nesting technique in a RCM and sensitivity tests to the resolution of the lateral boundary conditions during summer. *Clim. Dyn.*, 25, 555-580.

Durman, C.F., J.M. Gregory, D.C. Hassell, R.G. Jones, and J.M. Murphy, 2001. "A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates. *Quart. J. R. Meteor. Soc.*, 127, 1005-1015

Durran, D.R., 2000: Comments on "The Differentiation between Grid Spacing and Resolution and Their Application to Numerical Modeling". *Bull. Amer. Meteor. Soc.*, 81, 2478–2479.

Feser, F., and H. von Storch, 2005: Spatial two-dimensional discrete filters for limited area model evaluation purposes. *Mon. Wea. Rev.*, 133, 1774-1786.

Feser, F., 2006: Enhanced Detectability of Added Value in Limited-Area Model Results Separated into Different Spatial Scales. *Mon. Wea. Rev.*, 134, 2180–2190.

Flato, G.M., and G.J. Boer, 2001: Warming Asymmetry in Climate Change Simulations. *Geophys. Res. Lett.*, 28, 195-198.

Flato, G.M., and W.D. Hibler. III, 1992: Modelling Pack Ice as a Cavitating Fluid. *J. Phys. Oceanogr.*, 22, 626-651.

Frei, C., J.H. Christensen, M. Déqué, D. Jacob, R.G. Jones, and P.L. Vidale, 2003. Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *J. Geophys. Res.*, 108, n° D3, 4124, doi:10.1029/2002JD002287..

Giorgi, F., and G.T. Bates, 1989: The climatological skill of a regional model over complex terrain. *Mon. Wea. Rev.*, 117, 2325-2347.

Giorgi, F., and M.R. Marinucci, 1996: An investigation of the sensitivity of simulated precipitation to the model resolution and its implications for climate studies. *Mon. Wea. Rev.*, 124, 148-166.

Giorgi, F., L.O. Mearns, C. Shields, and L. McDaniel, 1998: Regional nested model simulations of present day and 2 CO2 climate over the Central Plains of the U.S. *Clim. Change*, 40, 457–493.

Giorgi, F., et al., 2001a: Regional climate information – Evaluation and projections. In: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., et al. (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, 583–638.

Gutowski, W.J., S.G. Decker, R.A. Donavon, Z. Pan, R.W. Arritt, and E.S. Takle, 2003: Temporal-spatial scales of observed and simulated precipitation in central U.S. climate. *J. Climate*, 16, 3841-3847.

Herceg, D., A.H. Sobel, L. Sun, and S.E. Zebiak, 2006: The big brother experiment and seasonal predictability in the NCEP regional spectral model. *Clim. Dyn.*, 26(4), 1-14. (DOI 10.1007/s00382-006-0130-z; http://dx.doi.org/10.1007/s00382-006-0130-z)

Iorio, J.P., P.B. Duffy, B. Govindasamy, S.L. Thompson, M. Khairoutdinov, and D.A. Randall, 2004: Effects of model resolution and subgrid-scale physics on the simulation of daily precipitation in the continental United States. *Clim. Dyn.*, 23, 243-258.

Kain, J.S., and J.M. Fritsch, 1990: A one-dimensional entraining/detraining plumes model an application in convective parameterization. *J. Atmos. Sci.*, 47, 2784-2802.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Kharin, V.V., and F.W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulation with a coupled atmosphere-ocean GCM. *J. Climate*, 13, 3760-3788.

Kim, S.-J., G.M. Flato, G.J. Boer, and N.A. McFarlane, 2002: A coupled climate model simulation of the Last Glacial Maximum, Part 1: transient multi-decadal response. *Clim. Dyn.*, 19, 515-537.

Kim, S.-J., G.M. Flato, and G.J. Boer, 2003: A coupled climate model simulation of the Last Glacial Maximum, Part 2: Approach to equilibrium. *Clim. Dyn.*, 20, 635-661.

Koltzow, M., T. Iversen, and J.E. Haugen, 2008: Extended Big-Brother experiments: the role of lateral boundary data quality and size of integration domain in regional climate modeling. *Tellus A*, 60, 398–410.

Kunkel, K.E., K. Andsager, X.Z. Liang, R.W. Arritt, E.S. Takle, W.J. Gutowski, and Z. Pan, 2002: Observations and Regional Climate Model Simulations of Heavy Precipitation Events and Seasonal Anomalies: A Comparison. *J. Hydrom.*, 3, 322–334.

Laprise, R., and C. Girard, 1990: A spectral general circulation model using a piecewise-constant finite element representation on a hybrid vertical coordinate system. *J. Climate*, 3, 32-52.

Laprise R., 1992: The resolution of global spectral models. *Bull. Amer. Meteor. Soc.*, 73, 1453-1454.

Laprise, R., R. Jones, B. Kirtman, H. von Storch, and W. Wergen, 2002: Atmospheric regional climate models (RCMs): A multiple purpose tool? Report of the « Joint WGNE/WGCM ad hoc Panel on Regional Climate Modelling », 19 p.

Laprise, R., 2003: Resolved scales and nonlinear interactions in limited-area models. *J. Atmos. Sci.*, 60 (5), 768-779.

Laprise, R., 2008: Regional climate modelling. *J. Comp. Phys.* 227, Special issue on «Predicting weather, climate and extreme events» (invited paper), 3641–3666.

Laprise, R., R. de Elía, D. Caya, S. Biner, Ph. Lucas-Picher, E.P. Diaconescu, M. Leduc, A. Alexandru, and L. Separovic, 2008: Challenging some tenets of Regional Climate Modelling. *Meteor. Atmos. Phys.* 100, Special Issue on Regional Climate Studies (invited paper), 3-22 (DOI 10.1007/s00703-008-0292-9).

Lettenmaier, D.P., 1995: "Stochastic Modeling of Precipitation with Applications for Climate Model Downscaling". Chapter 11 *in Analysis of Climate Variability*, H. von Storch and A. Navarra, eds., Springer-Verlag, Berlin, 197-212.

McSweeney, C.F., 2007: "Daily Rainfall Variability at Point and Areal Scales: Evaluating Simulations of Present and Future Climate". Thesis submitted for the degree of Doctor of Philosophy at the University of East Anglia, Norwich, 252 p.

McFarlane, N.A., G.J. Boer, J-P. Blanchet, and M. Lazare, 1992: The Canadian Climate Centre Second-Generation General Circulation Model and its Equilibrium State. *J. Climate*, 5, 1013-1044.

Mesinger, F., K. Brill, H. Chuang, G. DiMego, and E. Rogers, 2002: Limited area predictability: Can upscaling also take place? Research Activities in Atmospheric and Oceanic Modelling, Report n°. 32, WMO/TD – n°. 1105, 5.30-5.31.

Metcalfe, J. R., B. Routledge, and K. Devine, 1997: Rainfall measurement in Canada: changing observational methods and archive adjustment procedures. *J. Climate*, 10, 92-101.

Oreskes, N., K.S. Shrader-Frechette, and K. Belitz, 1994: Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science*, 263, 641–646.

Osborn, T.J., and M. Hulme, 1997: Development of a Relationship between Station and Grid-Box Rainday Frequencies for Climate Model Evaluation. *J. Climate*, 10(8), 1885.

Paquin, D., D. Caya, and R. Laprise, 2002: Treatment of moist convection in the Canadian Regional Climate Model. Ouranos, Équipe Simulations climatiques. Internal report n° 1, 30 p.

Perkins, S.E., A.J. Pitman, N.J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions. *J. Climate*, 20, 4356-4376, doi: 10.1175/JCLI4253.1.

Pielke, R. A., 1991: A recommended specific definition of "resolution". *Bull. Amer. Meteor. Soc.*, 72(12), 1914.

Räisänen, J., and R. Joelsson, 2001: Changes in average and extreme precipitation in two regional climate model experiments. *Tellus A*, 53, 547-566.

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi, and K.E. Taylor, 2007: Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Riette, S., and D. Caya, 2002: Sensitivity of short simulations to the various parameters in the new CRCM spectral nudging. *Res. Act. in Atmos. and Oceanic Modelling*, edited by H. Ritchie, WMO/TD - n° 1105, Report No. 32: 7.39-7.40.

Scinocca, J.F., N.A. McFarlane, M. Lazare, J. Li, and D. Plummer, 2008: The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. and Phys.*, submitted.

Separovic, L., R. de Elía, and R. Laprise, 2008: Reproducible and Irreproducible Components in Ensemble Simulations with a Regional Climate Model. *Mon. Wea. Rev.*, 136, 4942–4961.

Sevruk, B., 1982: Methods of correction for systematic errors in point precipitation measurements for operational use. *WMO Operational Hydrology Report,* n° 21. WMO- n°. 589, World Meteorological Organization, Geneva, Switzerland.

Skelly, W.C., and A. Henderson-Sellers, 1996: Grid box or grid point: What type of data do GCMs deliver to climate impacts researchers? *Int. J. Climatology*, 16, 1079–1086.

Sun, Y., S. Solomon, A. Dai, and R.W. Portmann, 2006: How often does it rain? *J. Climate*, 19, 916-934.

Trenberth, K.E., A. Dai, R.M. Rasmussen, and D.B. Parsons, 2003: The changing character of precipitation. *Bull. Amer. Meteor. Soc.*, 84, 1205-1217.

von Storch, H., H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, 128, 3664-3673

von Storch, H., and F.W. Zwiers, 1999: Statistical Analysis in Climate Research. *Cambridge University Press*, 484 p., ISBN 0521 450713.

Walters, M.K., 2000: Comments on "The Differentiation between Grid Spacing and Resolution and Their Application to Numerical Modeling". *Bull. Amer. Meteor. Soc.*, 81, 2475–2477.

Wang, X.L., and F.W. Zwiers, 1999: Interannual variability of precipitation in an ensemble of AMIP climate simulations conducted with CCC GCM2. *J. Climate*, 12, 1322-1335.

Zhang, G.J., and N.A. McFarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the CCC-GCM. *Atmos.-Ocean.*, 3, 407-446.

FIGURES

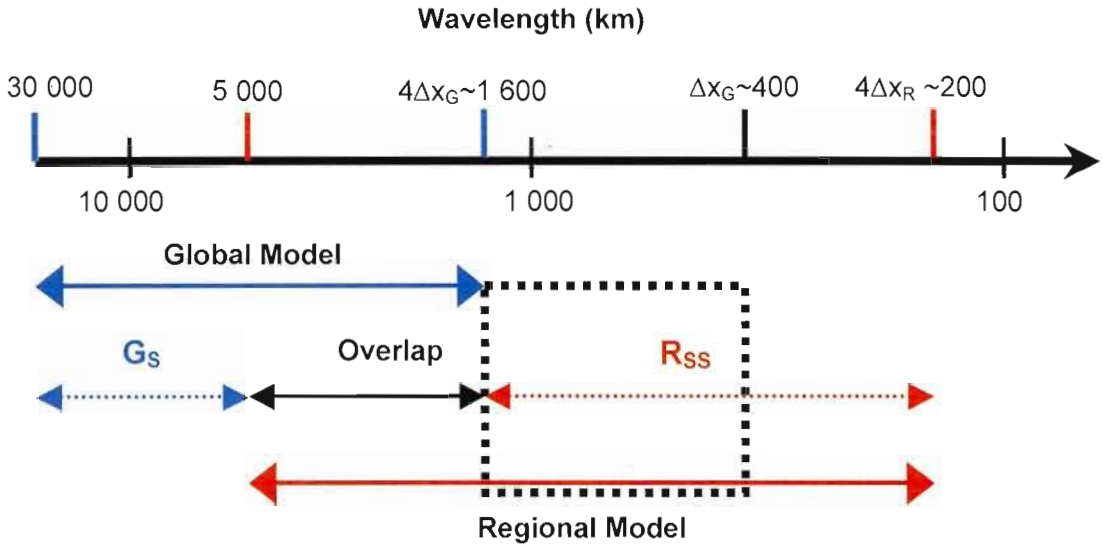

Fig. 1 Spatial scales resolved by a GCM (full blue arrow) and a RCM (full red arrow) with grid spacing of 400 km and 50 km respectively. Also indicated is the interval of wavelengths only resolved by the GCM (blue dotted arrow) and by the RCM (red dotted arrow). The black dashed square designates the spatial scales that will be considered in this study (see explanations in text).

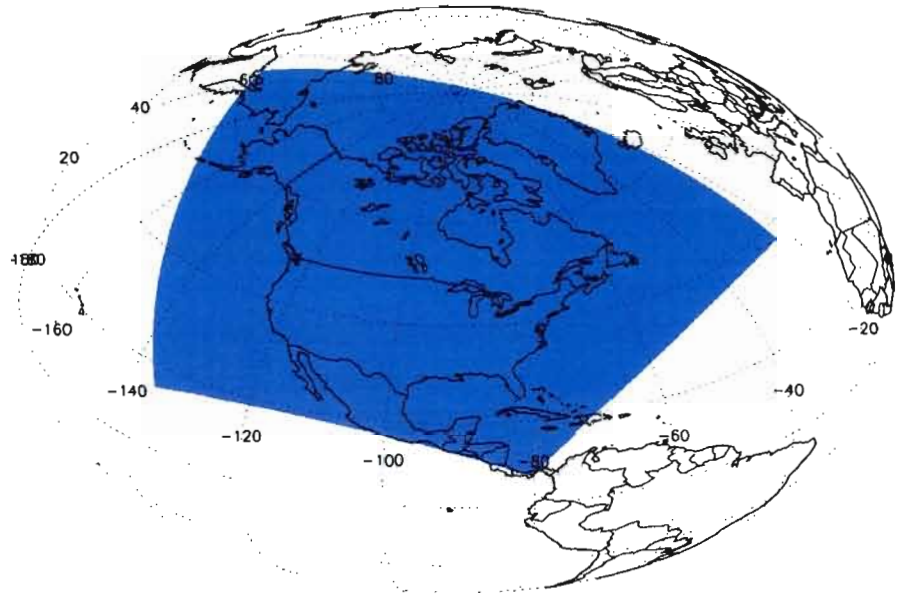

Fig. 2 Computational domain used in the CRCM (201 x 193 grid points) for the two simulations analyzed.

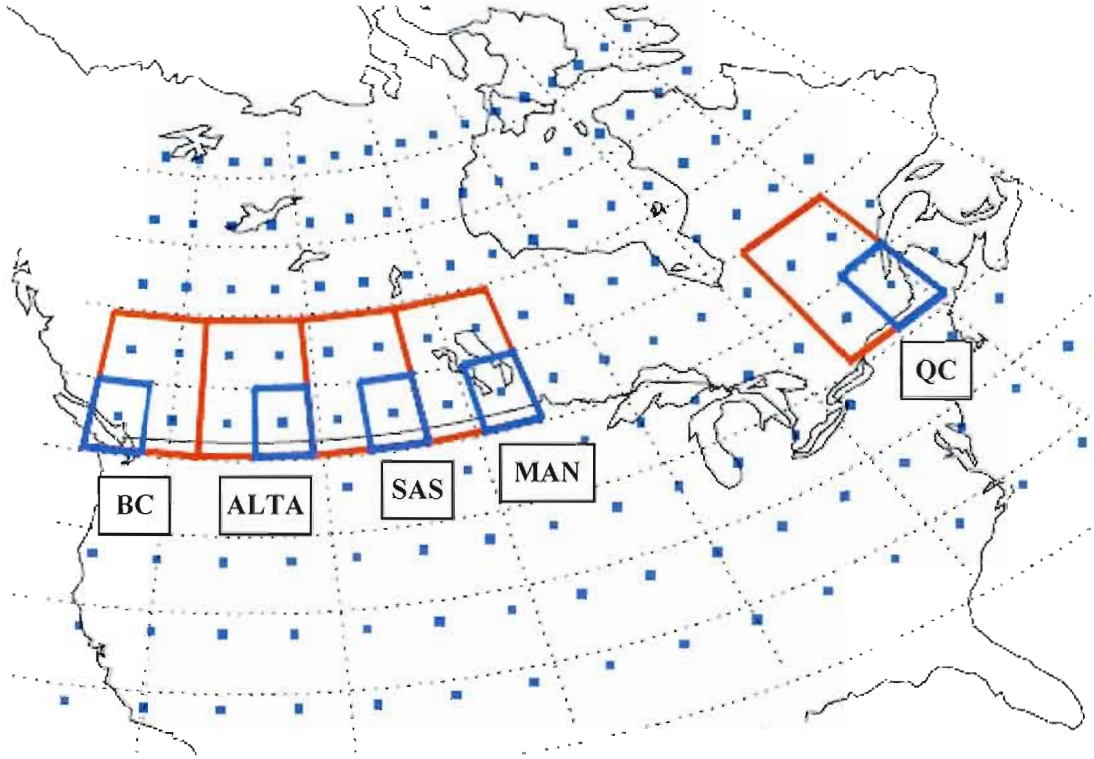

Fig. 3 Specification of areas of interest. Red and blue boxes indicate regions including four and one CGCM grid points respectively. Blue squares outlined the CGCM grid points available in North America.

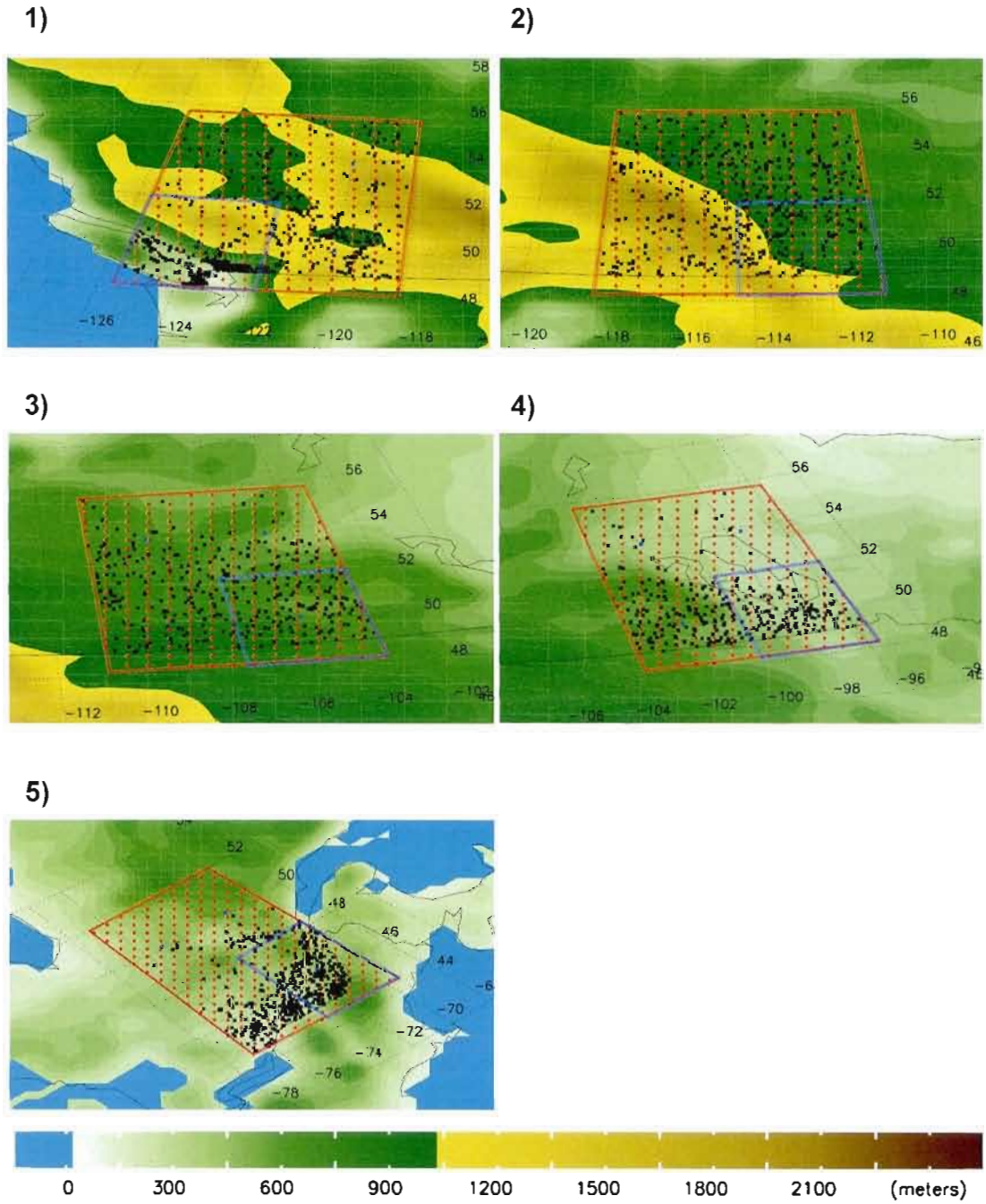

Fig. 4 Weather stations (black points), CRCM grid points (red points) and CGCM grid points (blue points) for regions indicated in Fig. 1: (1) BC, (2) ALTA, (3) SAS, (4) MAN and (5) QC. As in Fig. 1, the blue boxes denote those regions including one CGCM grid point and the red ones those regions that include four grid points. The topography field as represented in the CRCM is shown in color filled contours.

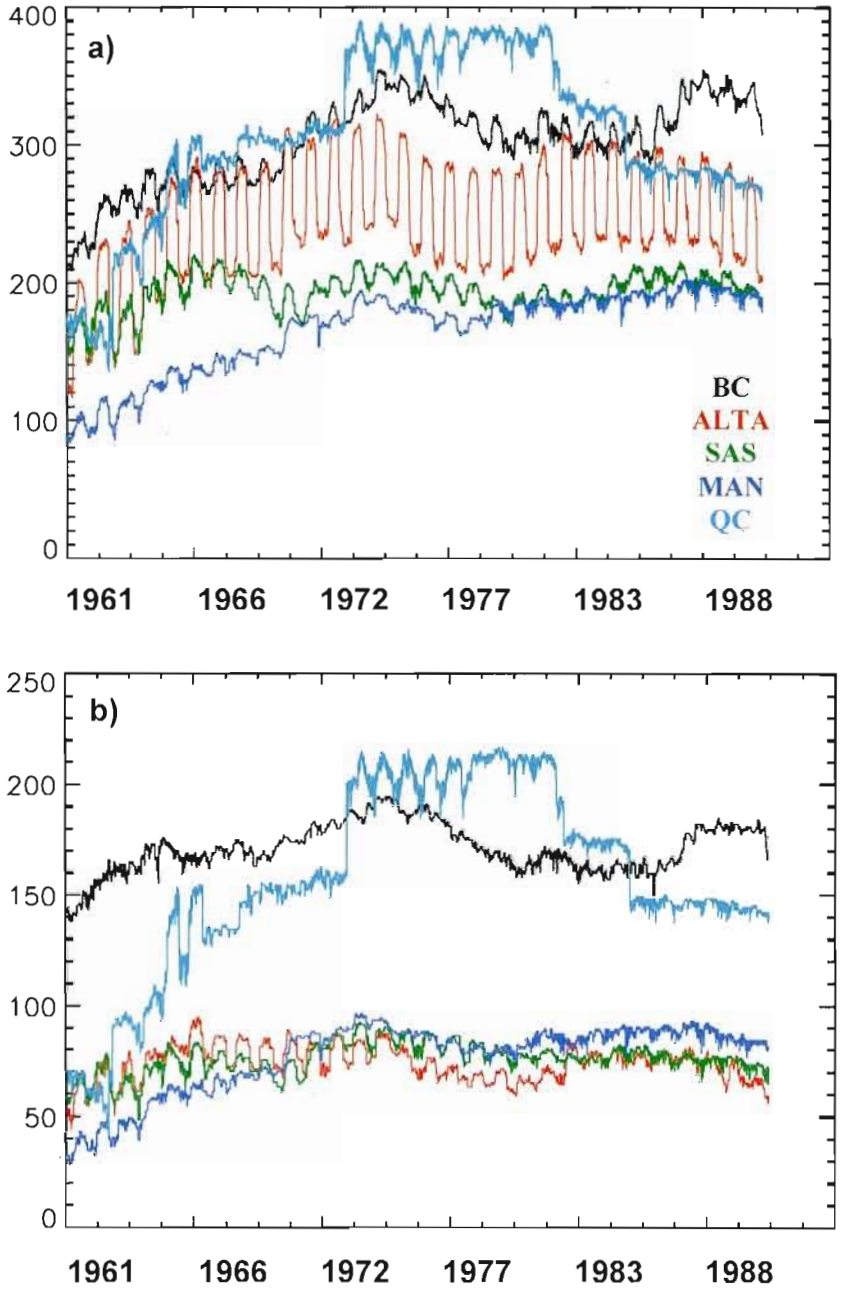

Fig. 5 Number of stations available in each region as a function of days for the period 1961 – 1990. Each region is indicated with a color. Fig. (a) shows regions including four CGCM grid points and (b) those regions including only one CGCM grid point.

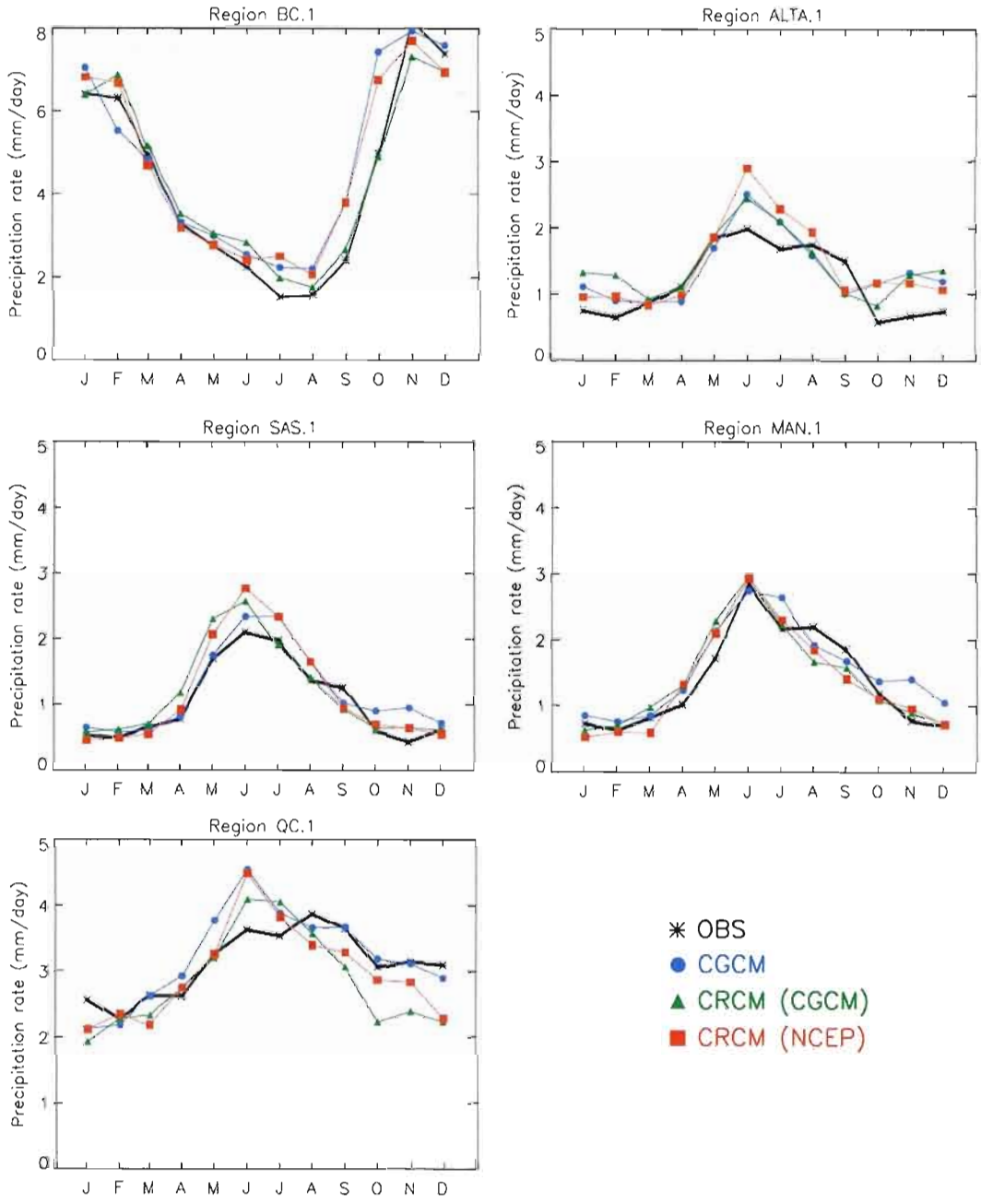

Fig. 6 Observed and simulated monthly mean precipitation rates (mm/day) for all regions including one-CGCM grid point.

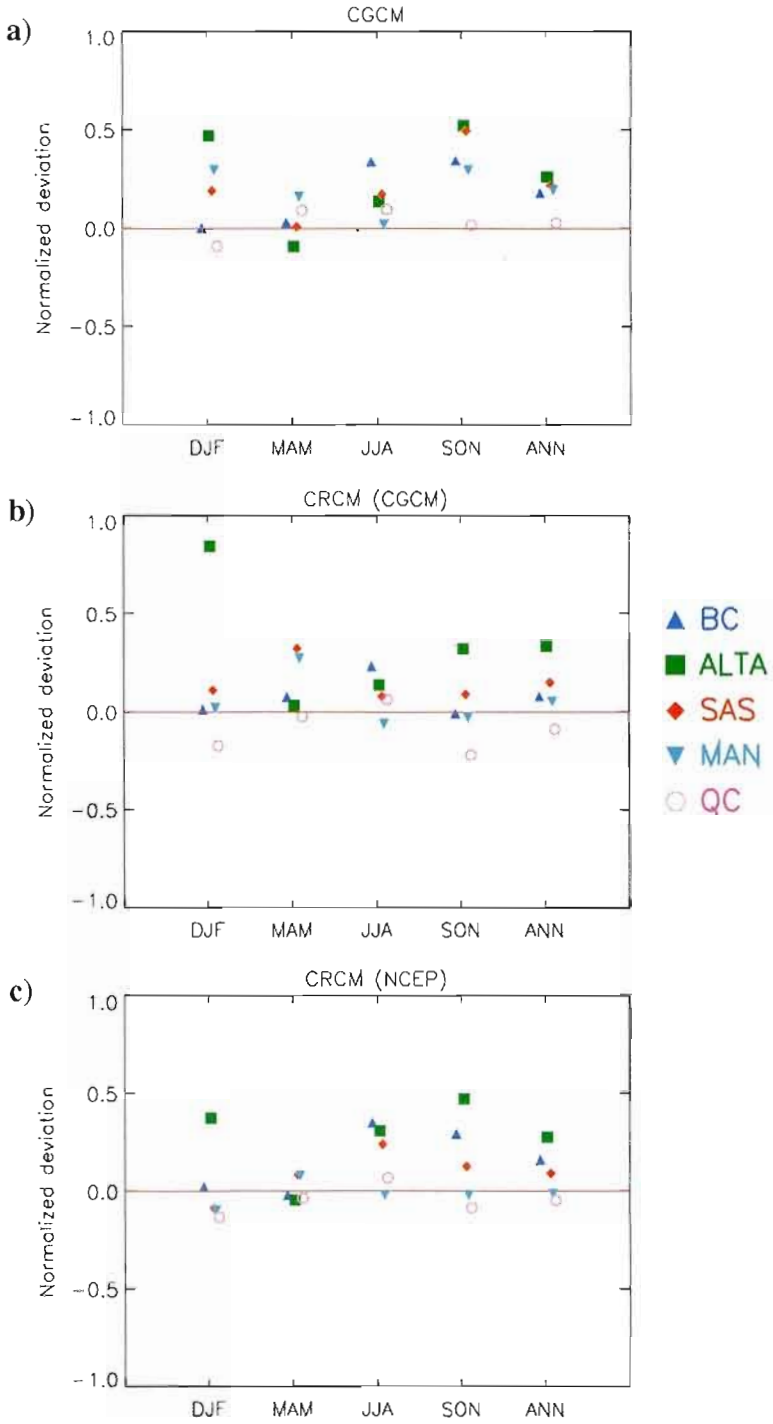

Fig. 7 Deviations between simulated and observed seasonal mean precipitation rate, normalized by observed values, for (a) CGCM, (b) CRCM when driven by the CGCM and (c) CRCM when driven by the NCEP/NCAR reanalysis. Different symbols represent each of the five regions analyzed.

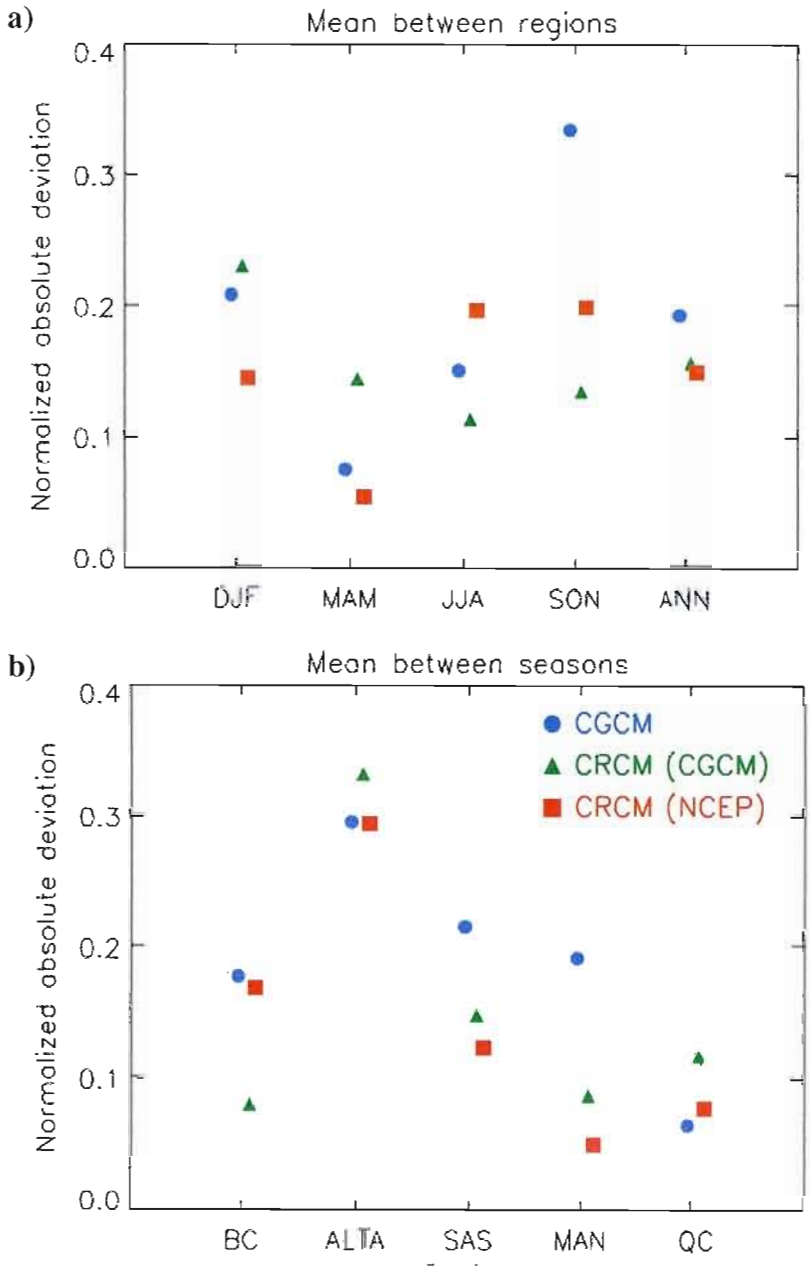

Fig. 8 Absolute deviations between simulated and observed seasonal mean precipitation rate, normalized by observed values, for (a) averaging between regions and (b) mean annual precipitation rate. Different symbols represent each of the five regions analyzed.

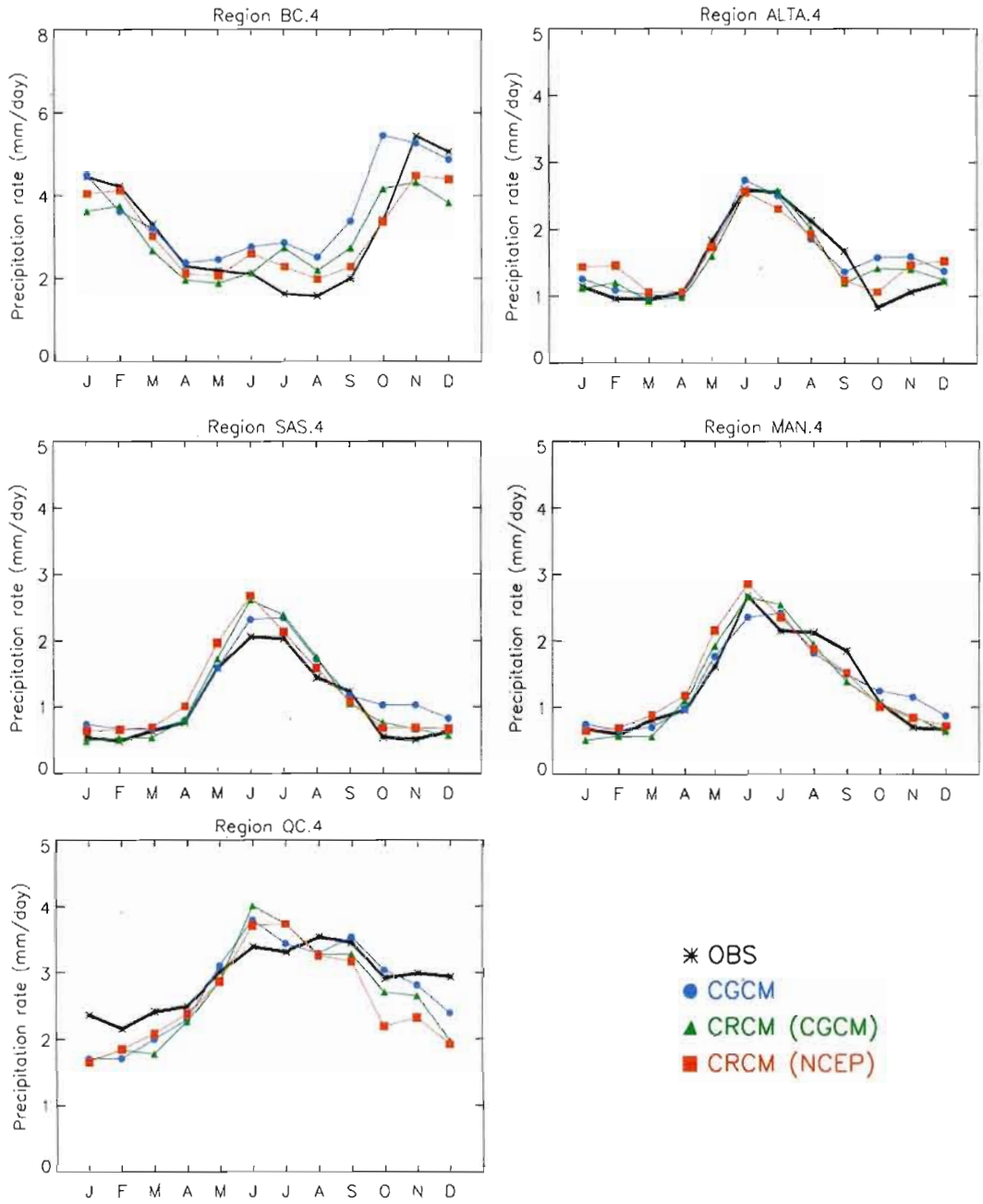

Fig. 9 As in Fig. 6 but for regions including four CGCM grid points.

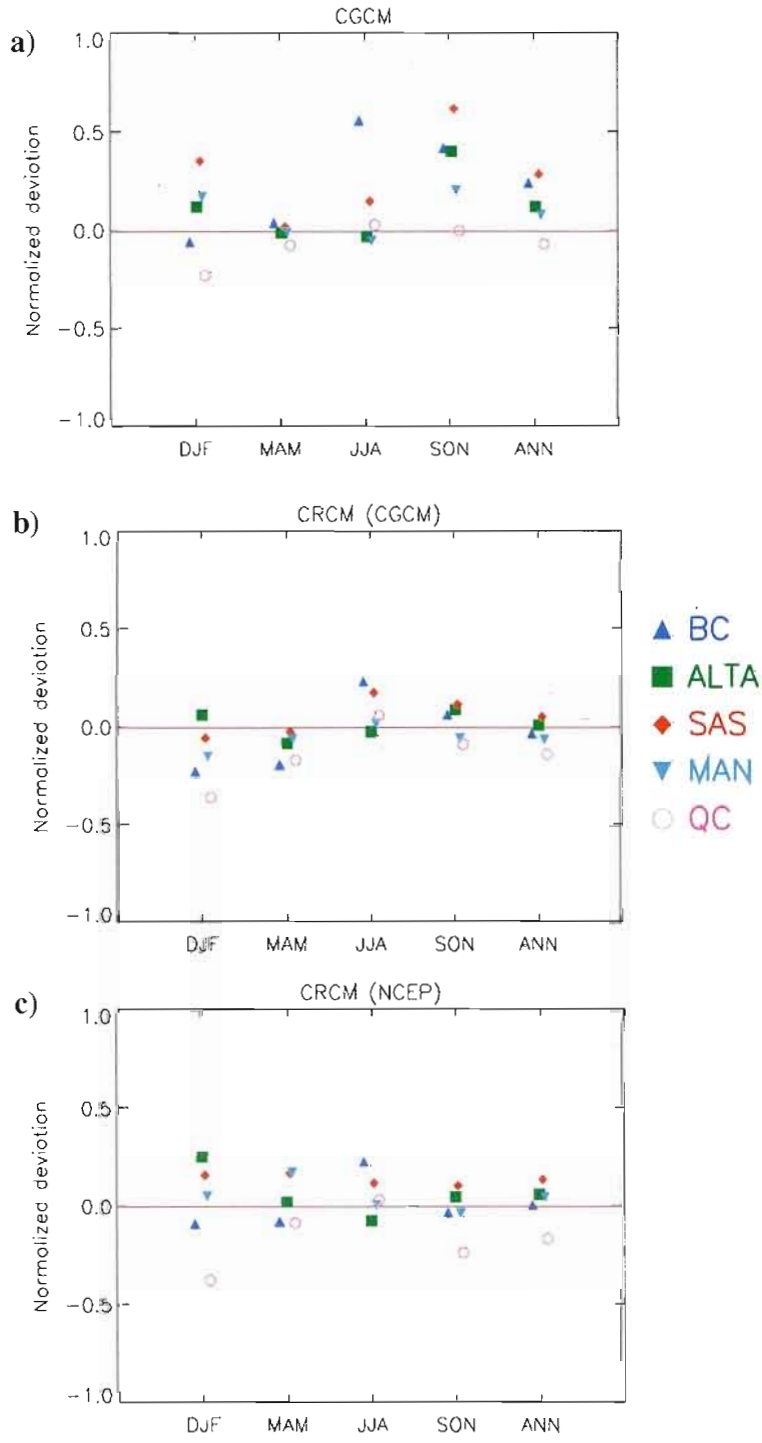

Fig. 10 As in Fig. 7 but for regions including four CGCM grid points.

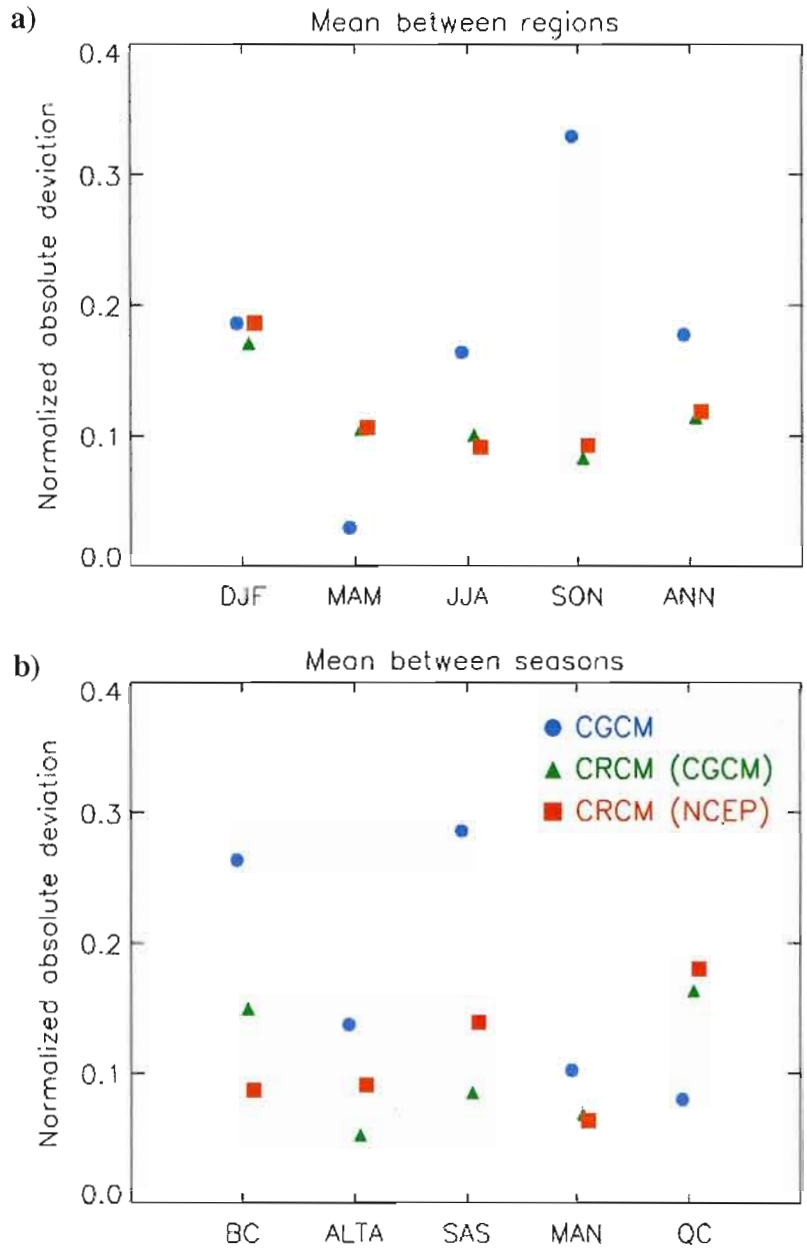

Fig. 11 As in Fig. 8 but for regions including four CGCM grid points.

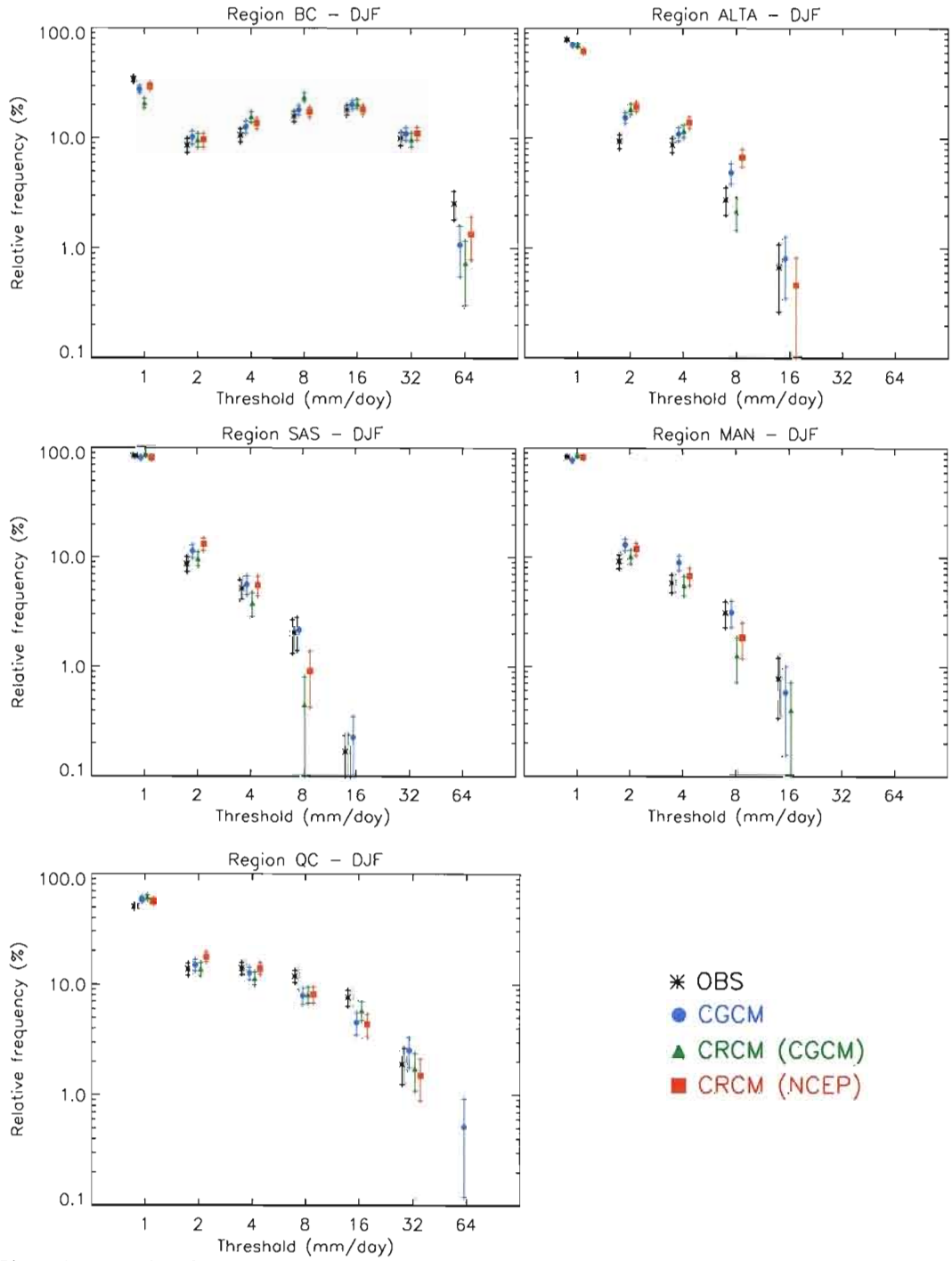

Fig. 12 Intensity frequency distributions of precipitation rate in DJF for the five regions studied. Observed data is in black, CGCM data in blue, the CRCM when driven by CGCM in green and when driven with NCEP/NCAR reanalysis in red. Only frequencies greater than 0,1 % are shown.

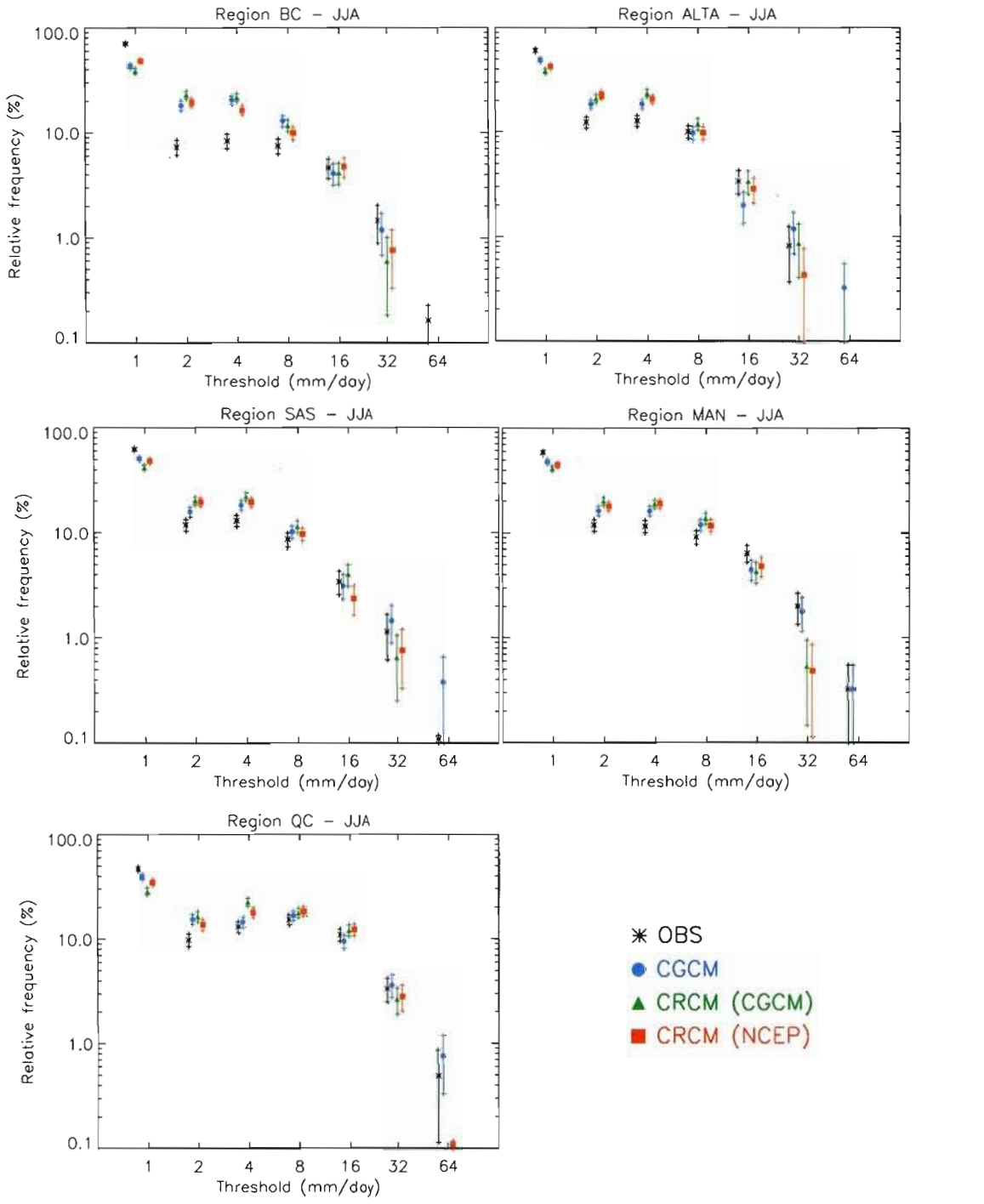

Fig. 13 As in Fig. 12 but for JJA.

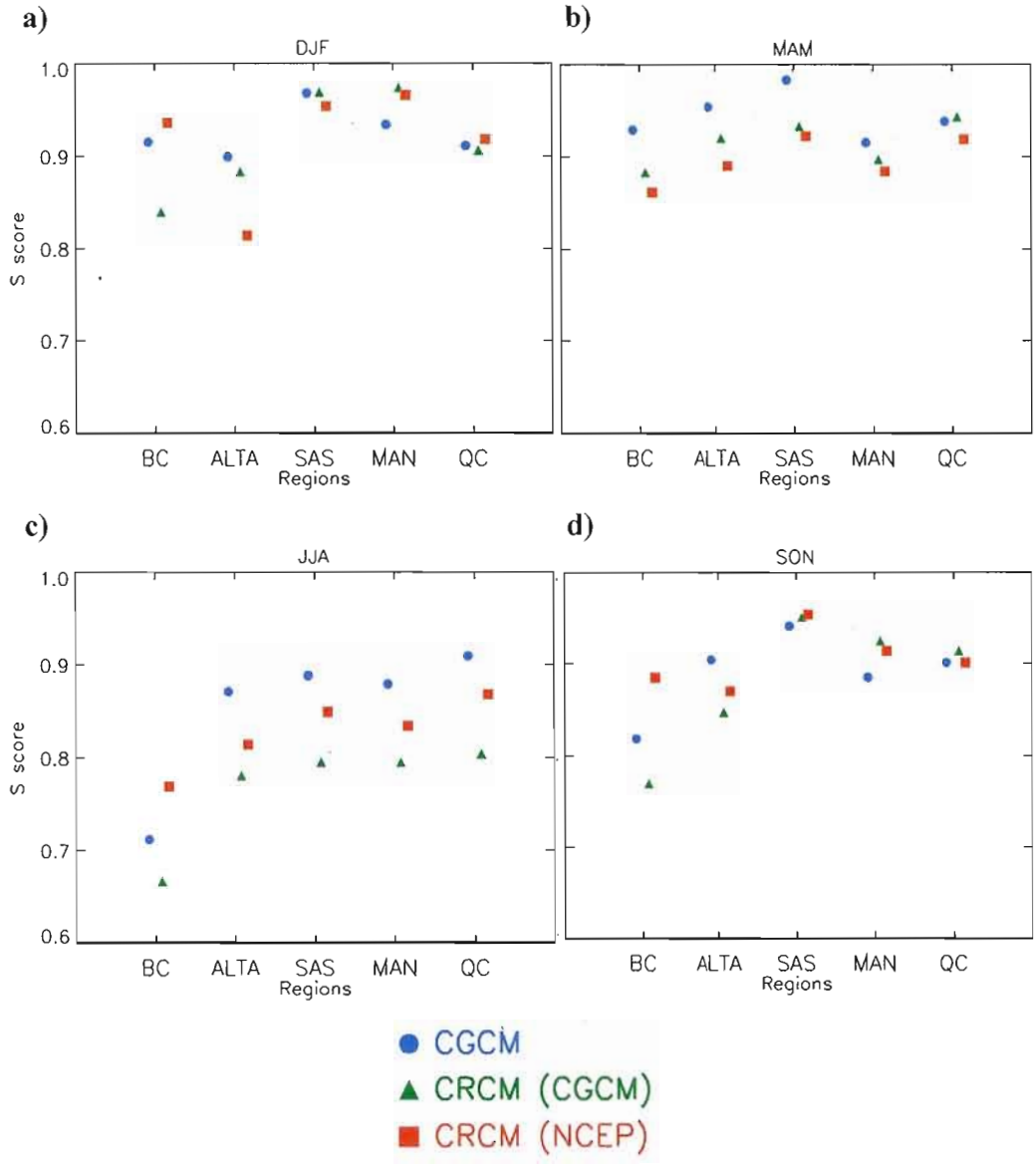

Fig. 14 S score (Perkins et al., 2007) calculated between the simulated and observed distribution as a function of regions for a) DJF, b) MAM, c) JJA and d) SON. Symbols denote the model used to produce data.

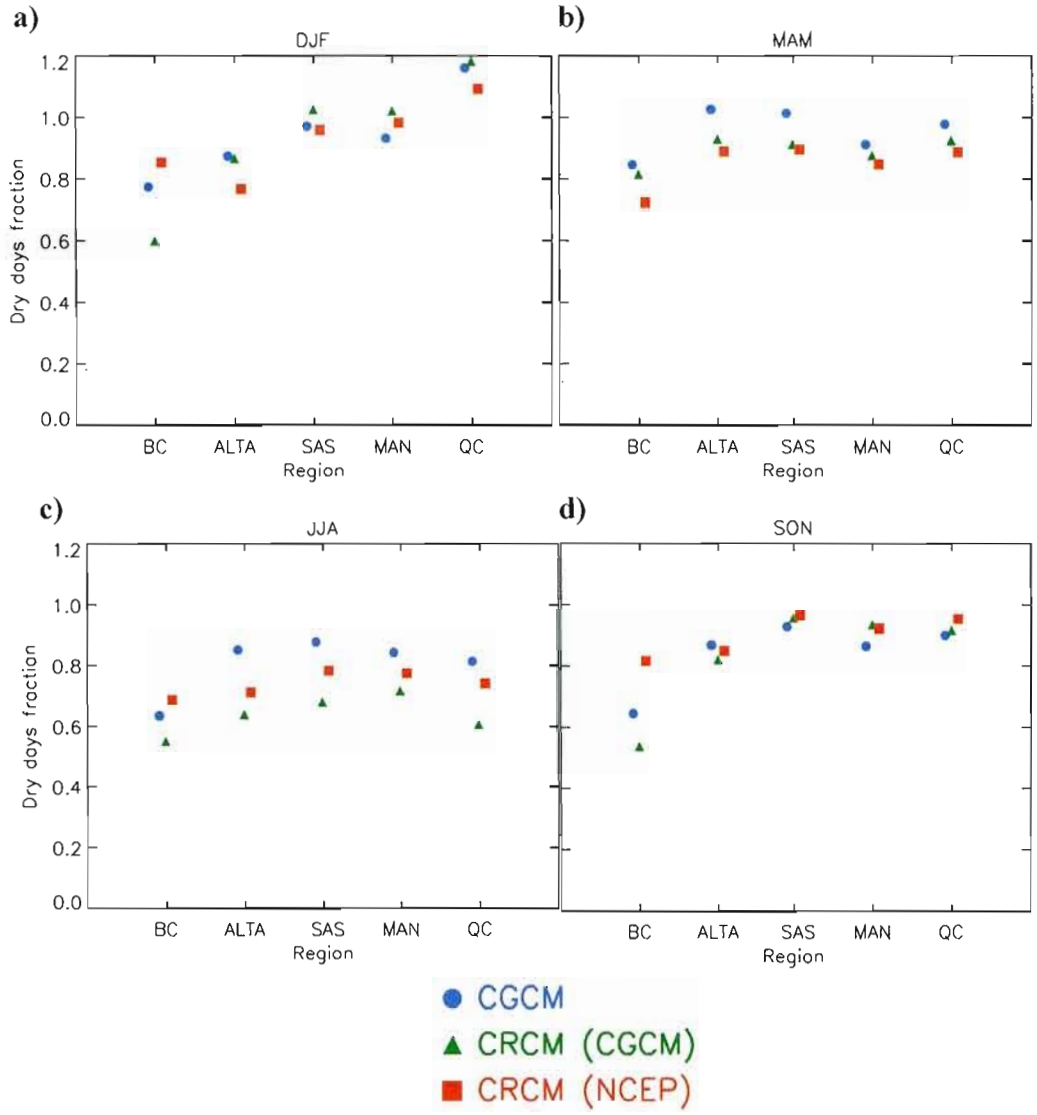

Fig. 15 Ratio between observed and simulated dry days as a function of regions for a) DJF, b) MAM, c) JJA and d) SON. Symbols denote the model used to produce data.

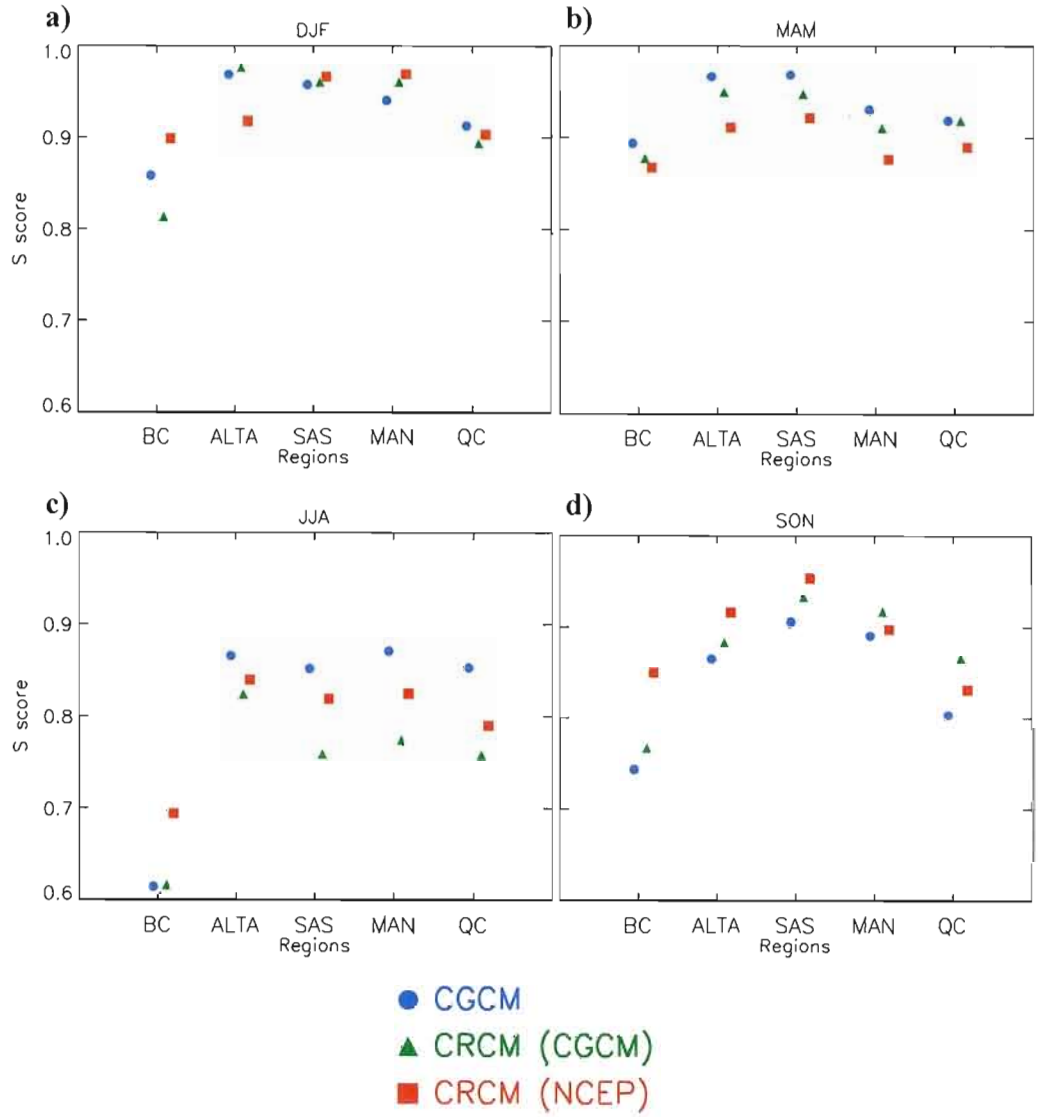

Fig. 16 As in figure 14 but for regions including four CGCM grid points.

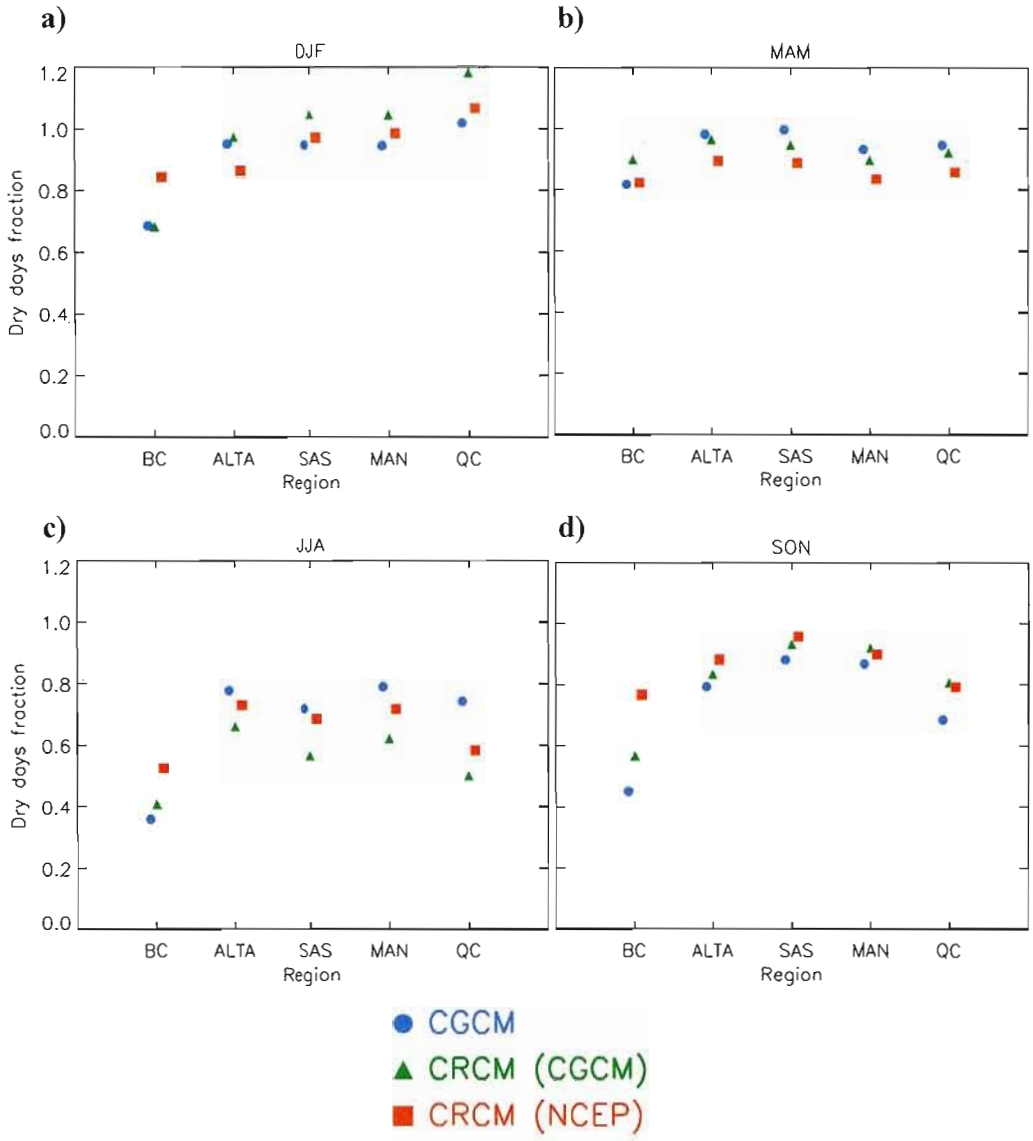

Fig. 17 As in Fig. 15 but for regions including four CGCM grid points.

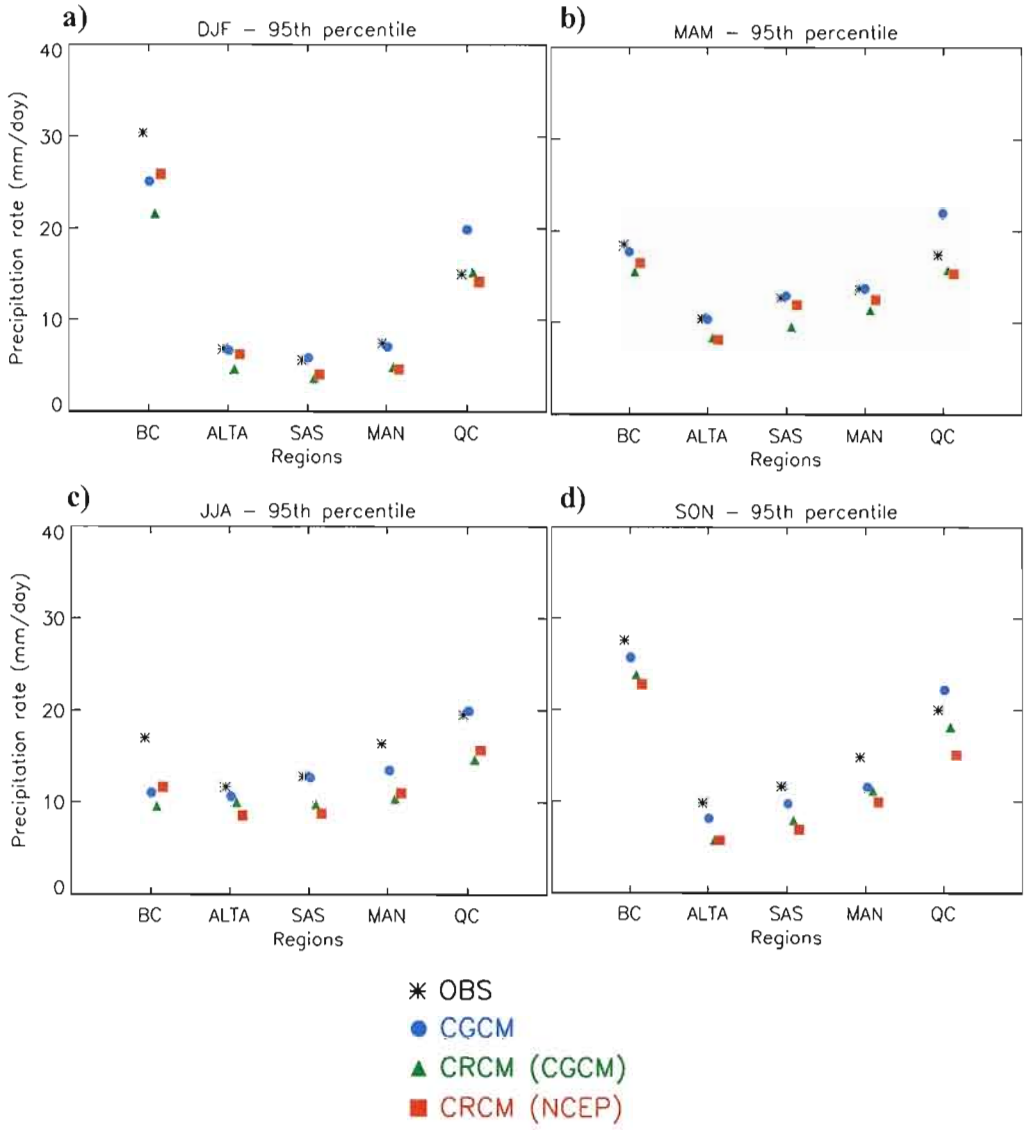

Fig. 18 Observed and simulated 95 % percentile calculated as a function of region for a) DJF, b) MAM, c) JJA and d) SON. Symbols denote the model used to produce data.

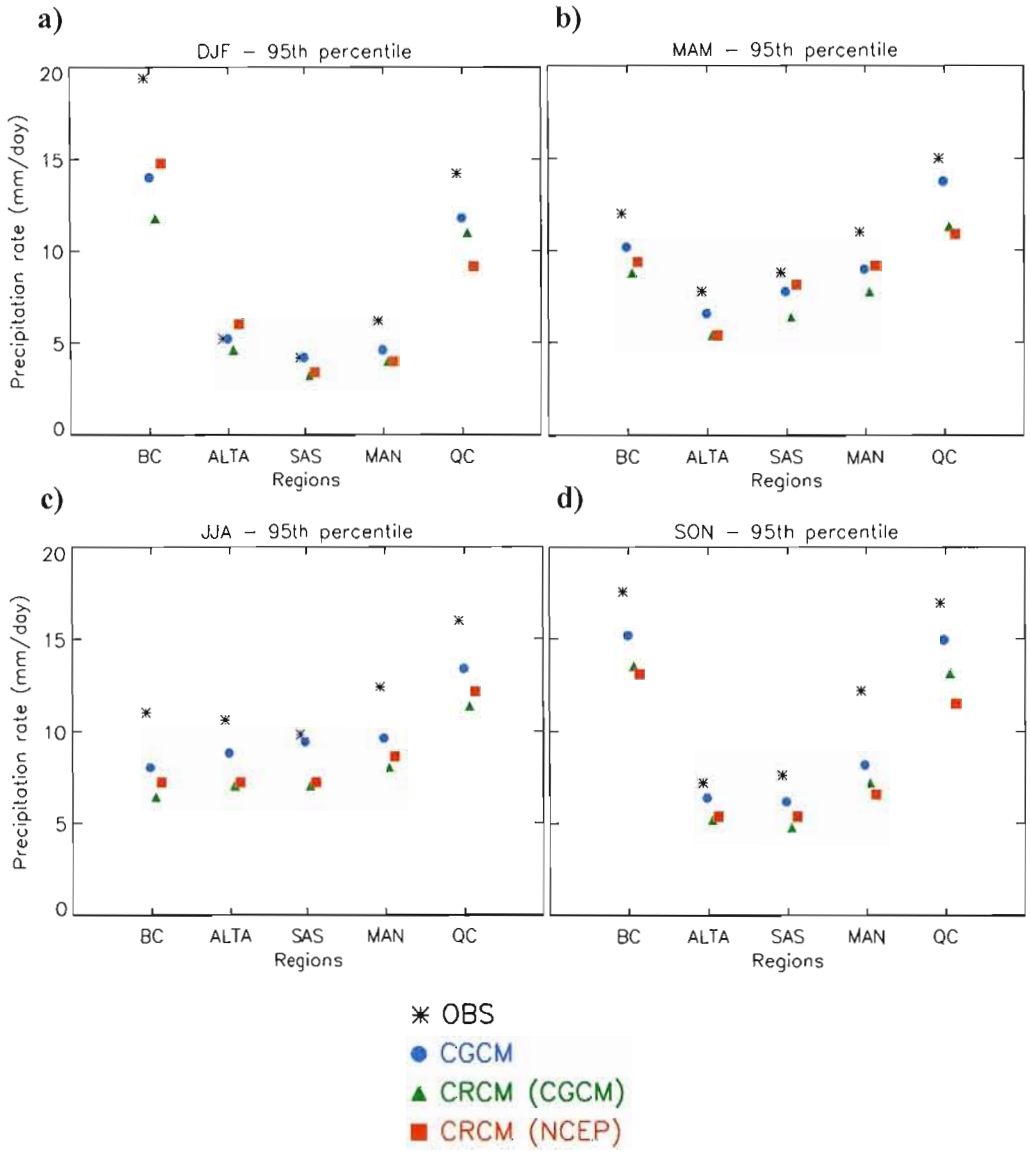

Fig. 19 As in Fig. 18 but for regions including four CGCM grid points.

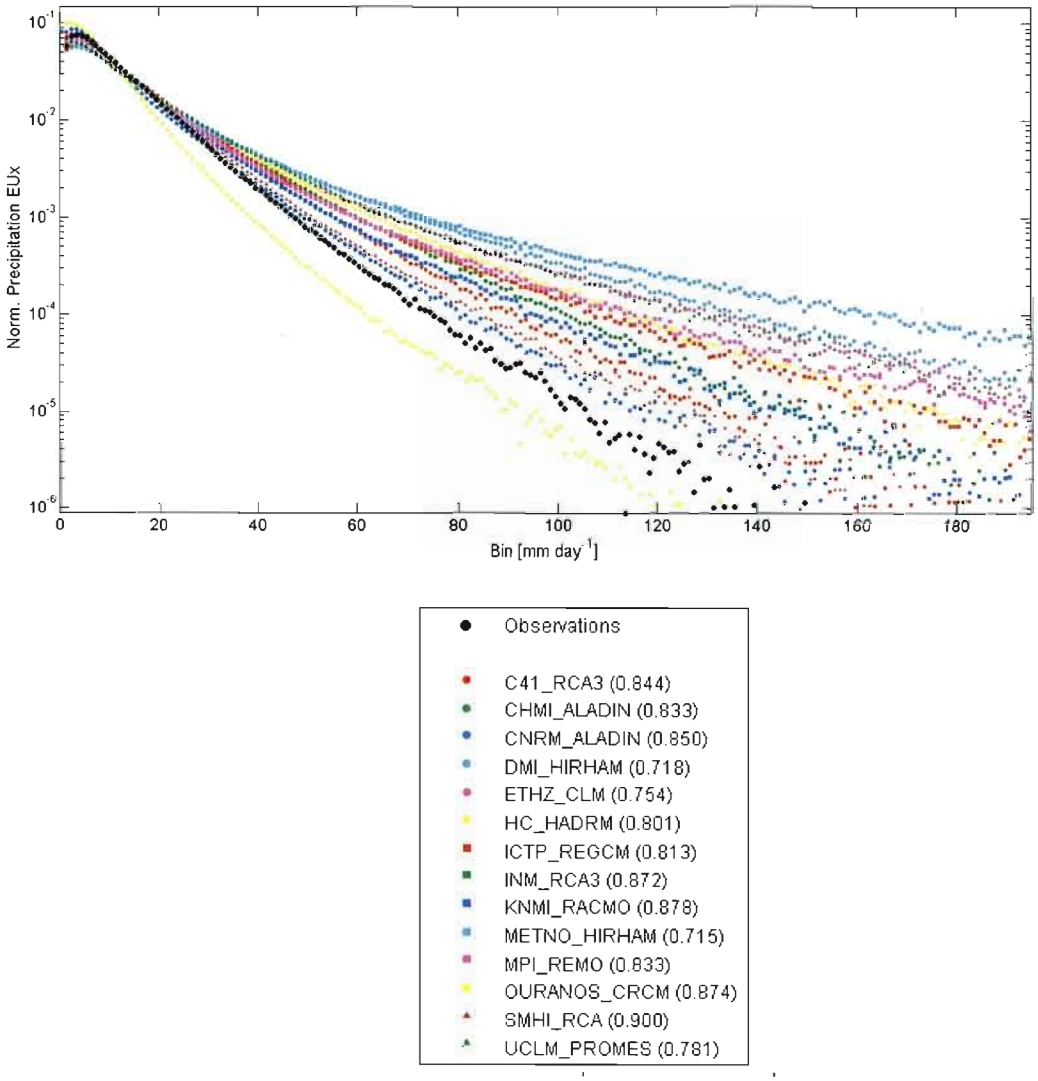

Fig. 20 Normalized precipitation distribution for the full European domain using all annual data over land only. Symbols denote the model used to produce data (Reproduced from Christensen et al. 2008).

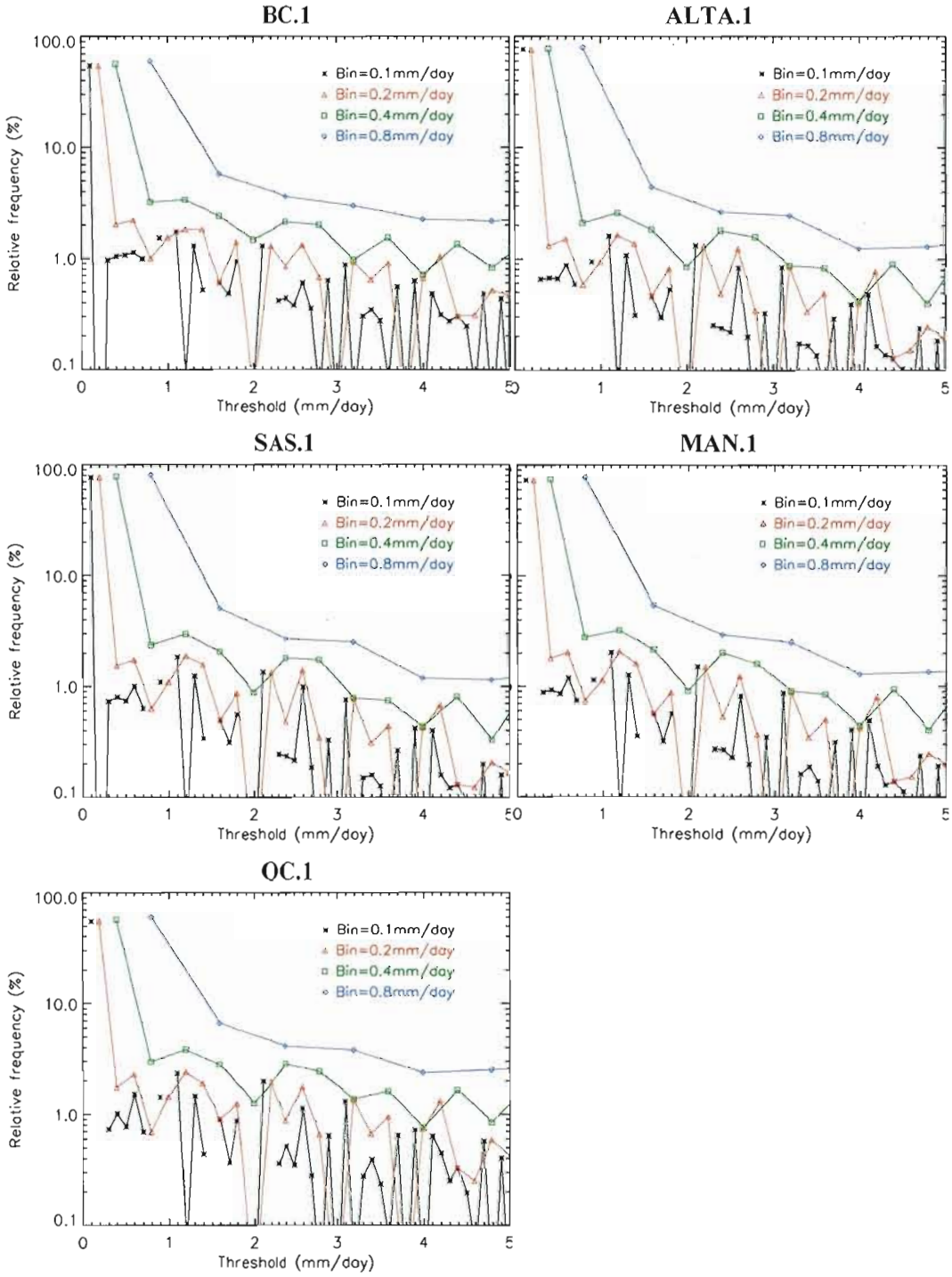

Fig. A.1 Observed intensity frequency distributions calculated on the basis of four different bin sizes (only intensities less than 5 mm/day are shown). Regions are indicated with blue lines in Fig. 1, corresponding to the size of one CGCM grid-box.

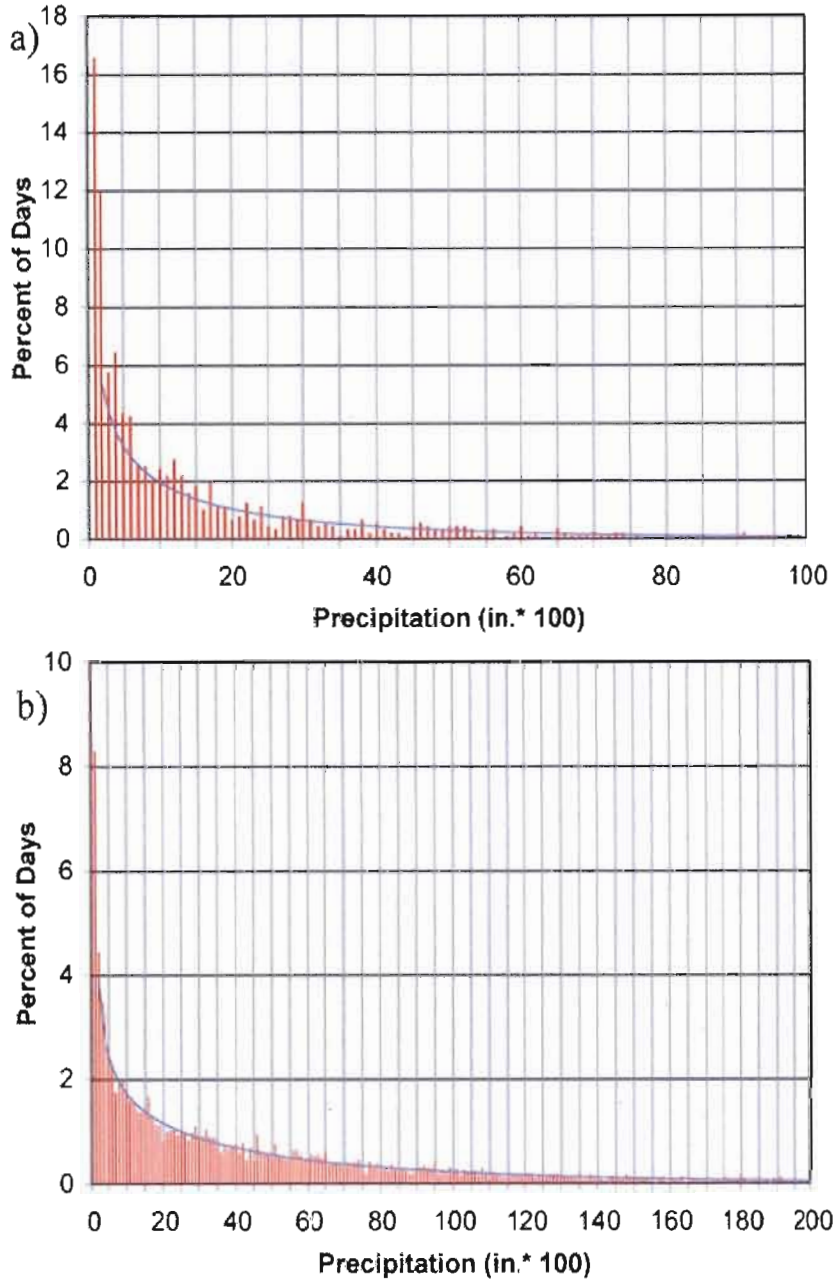

Fig. A.2 Percent frequency distribution of daily precipitation of at least 0.01 in. for the period 1971– 2000 at COOP stations: (a) Bishop, CA (040822), mean annual precipitation of 4.9 in. (125 mm); and (b) Quillayute, WA (456858), mean annual precipitation of 101.9 in. (2587 mm). Neither station exhibits appreciable observer bias. Solid curve is the fitted gamma function. (Reproduced from Daly et al. (2007) © 2007 American Meteorological Society).

TABLEAUX

Simulation name	Model label and Institution	Type Grid spacing /truncation N° of vertical levels	Lateral Boundary Conditions	Period analyzed
CGCM	CGCM version 3.1 Canadian Centre for Climate Modelling and Analysis	Global Model T47 L32	----------	1971 – 1990
CRCM (CGCM)	CRCM version 4.2.0 Ouranos Consortium	Regional Model 45 km at 60° N L29	CGCM version 3.1	
CRCM (NCEP)			NCEP/NCAR reanalyses	

Table 1. Simulations used in the study. Column 1 is the acronym of each simulation used in the text. Column 2 is the name of the model and its source. Column 3 denotes the type of model, the horizontal grid spacing/triangular truncation and the number of vertical levels and column 4 indicates the data used to drive the regional model in each simulation. Finally, column 5 gives the period of analysis of simulations.

Region name	Size (degree^2)	Latitude (°N)		Longitude (°W)	
		South limit	North limit	West limit	East limit
BC.1		48.22	51.97	125.62	121.87
ALTA.1		48.22	51.97	114.37	110.62
SAS.1	3.75 x 3.75	48.22	51.97	106.87	103.12
MAN.1		48.22	51.97	99.37	95.62
QC.1		44.52	48.27	73.12	69.37
BC.4		48.20	55.70	125.62	118.12
ALTA.4		48.20	55.70	118.12	110.62
SAS.4	7.5 x 7.5	48.20	55.70	110.62	103.12
MAN.4		48.20	55.70	103.12	95.62
QC.4		40.78	48.28	76.87	69.37

Table 2. Name, sizes (degrees^2) and boundaries in latitude and longitude for the regions examined in the study. The name of the region is composed by the abbreviated name of the province where it belongs and a digit that refers to the number of CGCM grid points within the region.

Region	CGCM	CRCM	Station data	
			Minimum	Mean
BC.1	1	61	150	165
ALTA.1	1	58	60	70
SAS.1	1	61	60	80
MAN.1	1	60	70	85
QC.1	1	66	150	170
BC.4	4	230	280	324
ALTA.4	4	233	203	257
SAS.4	4	231	154	198
MAN.4	4	231	151	185
QC.4	4	257	251	333

Table 3. Number of grid points of each model and number of weather stations within each region for the period 1971-1990. In the case of observed data, the minimum and mean value of stations are indicated.

Size of region (degree^2)	Season	CGCM	CRCM (CGCM)	CRCM (NCEP)
7.5 x 7.5	DJF	0.92	0.91	0.91
	MAM	0.94	0.91	0.89
	JJA	0.85	0.76	0.82
	SON	0.88	0.88	0.90
3.75 x 3.75	DJF	0.92	0.92	0.93
	MAM	0.93	0.91	0.89
	JJA	0.81	0.74	0.79
	SON	0.84	0.87	0.88

Table 4. Inter-region mean S score (Perkins et al., 2007) calculated between the simulated and observed distribution for each season and for each simulation. Results are presented for both sizes of regions. Minimum and maximum values in each simulation are denoting in blue and red respectively.

Region	N° of stations		Total n° values	Equivalent n° of years
	Mean	Maximum		
BC.1	165	322	1 204 500	3 300
ALTA.1	70	143	511 000	1 400
SAS.1	80	128	584 000	1 600
MAN.1	85	135	620 500	1 700
QC.1	170	259	1 241 000	3 400

Table A1 Mean and maximum number of stations in each region including one CGCM grid point during the period 1971-1990. Also presented are, the total number of data and the equivalent number of years of a daily time series built assuming that all data are from the same site.

CONCLUSION

L'objectif principal de cette étude était d'examiner l'existence de la valeur ajoutée dans des simulations du MRCC lorsque le MCGC est utilisé comme pilote. Des échelles temporelles et spatiales communes aux deux modèles MRCC et MCGC ont été considérées dans l'analyse. La comparaison a été effectuée en ramenant les données à haute résolution des stations météorologiques et du MRCC à la résolution du MCGC.

Notre évaluation a été focalisée sur la simulation de la précipitation selon le MRCC et le MCGC en utilisant comme référence des séries temporelles observées. Des observations journalières, moyennées spatialement, ont été construites à partir des stations météorologiques d'Environnement Canada. En raison du niveau relativement bon de la couverture spatiale, cinq (5) différentes régions qui traversent le Canada ont été choisies pour l'analyse. Dans toutes les régions, le nombre minimum de stations pour estimer la moyenne dans l'espace, est cinq (5) fois plus grand que la même proposée par Osborn et Hulme (1997).

L'évaluation a été basée sur la comparaison des histogrammes d'intensités de précipitation et des 95^{e} centiles des distributions afin de caractériser les événements les plus extrêmes. La mesure *S,* définie dans Perkins *et al.* (2007), a été utilisée pour estimer le degré de chevauchement entre les distributions simulées et observées. Cette dernière reflète principalement le comportement des intensités faibles et modérées.

La considération des différentes régions dans l'analyse permet d'évaluer la valeur ajoutée en fonction des forçages de surface. De la même manière, la dépendance de la valeur ajoutée dans les régimes de temps est étudiée suivant l'analyse des statistiques saisonnières.

Les résultats montrent que les statistiques quotidiennes des précipitations simulées par le MCGC et par le MRCC sont généralement très similaires et, en comparant les deux données, qu'il n'existe aucune preuve de l'existence de la valeur ajoutée. Par exemple, lors de la saison hivernale, les modèles MCGC et le MRCC montrent des performances semblables pour simuler la fréquence et l'intensité des valeurs observées, tout en donnant des valeurs du *S* très similaires, indépendamment de la région considérée.

Pendant la saison estivale, les deux modèles ont plus de difficultés à reproduire la distribution observée et présentent des valeurs plus petites de S qu'en hiver. Comme le montre la mesure S, le MCGC fournit un meilleur accord avec les données observées comparativement au MRCC. Cette amélioration vient d'une meilleure simulation de la fréquence des jours secs.

La représentation des jours secs est un problème récurrent dans les modèles climatiques (Trenberth *et al.*, 2003), généralement associé à une surestimation de la fréquence des événements de faibles intensités (Dai *et al.*, 1999; Paquin *et al.*, 2002; Frei *et al.*, 2003, Dai et Trenberth, 2004). Toutes les simulations présentées dans cette étude montrent une sous-estimation des jours secs et une surestimation des événements de précipitations faibles, presque indépendamment des saisons et de la région considérée. La seule exception à ce comportement est la saison d'hiver dans le sud du Québec (région QC.1). Bien que les deux modèles sous-estiment la fréquence des jours secs, le MRCC semble souffrir d'un biais de distribution encore plus prononcée que le MGCC. Une comparaison avec d'autres modèles régionaux du climat ajoute des évidences en faveur de cette supposition (voir, par exemple, Fig. 20 de Christensen *et al.* (2008)).

Dans le cas des événements journaliers les plus extrêmes, les résultats montrent que le MCGC produit un meilleur accord que le CRCM avec les valeurs observées. En effet, le CRCM montre une sous-estimation de la fréquence d'occurrence d'événements plus intenses. C'est aussi le cas dans les régions caractérisées par d'importants forçages de surface où les différences entre la topographie des modèles pourraient avoir un impact.

Plus qu'un problème des performances relatives des modèles, l'absence de valeur ajoutée pourrait être liée aux mauvaises hypothèses énoncées sur la valeur ajoutée dans les échelles étudiées. L'hypothèse soutenant l'existence de la valeur ajoutée dans notre comparaison a été présentée dans l'introduction et est énoncée comme suit :

> *Le modèle mondial a peu d'habileté pour résoudre les processus de fine échelle dû au fait que les échelles considérées dans l'analyse sont près de sa limite de troncature* (Laprise, 2003; Feser, 2006).

Cette hypothèse a été généralement étudiée dans le cas des variables instantanées, mais non lorsque la variable représente une moyenne temporelle (par exemple: pour des valeurs

journalières, mensuelles, etc.). Dans cette étude, l'hypothèse a été testée en comparant les résultats obtenus dans différents domaines : ceux définis par un seul point de grille et ceux comprenant quatre points de grille du MGCC. La performance du modèle mondial semble être similaire dans les deux domaines, ce qui suggère que le modèle «fonctionne bien» dans les deux échelles spatiales. En d'autres termes, en raison de la très bonne performance du modèle global, indépendamment de la taille des régions considérées, l'hypothèse ne semble pas être appropriée au moins pour les échelles de temps analysées ici (données journalières et mensuelles).

Pour éviter une mauvaise interprétation de nos résultats, il est intéressant d'examiner brièvement deux caractéristiques particulières de la méthodologie étant liées au caractère partiel de l'évaluation des modèles :

L'approche utilisée dans la présente étude est basée sur l'évaluation des échelles spatiales et temporelles des précipitations qui sont représentées par les deux modèles, mais proches de la limite de troncature du modèle global. Les avantages des simulations du MRC, en raison de sa plus haute résolution spatio-temporelle, ne sont pas explicitement pris en considération ce qui n'est pas la manière la plus efficace pour mettre en évidence les avantages d'une augmentation de la résolution dans un modèle MRC. Cette caractéristique de l'étude doit être prise en compte au moment de l'évaluation de ces résultats.

La performance relative de certaines statistiques de la précipitation simulée par les deux modèles par rapport aux données d'observation est utilisée comme outil pour détecter la valeur ajoutée. Cette approche a ses revers. Comme l'ont déclaré Oreskes *et al.* (1994), « If a model fails to reproduce observed data, then we know that the model is faulty in some way, but the reverse is never the case ». En d'autres termes, si un modèle A produit des résultats qui sont plus proches des données observées que ceux du modèle B, ceci n'implique pas nécessairement que le modèle A soit meilleur que le modèle B. Une meilleure performance peut être atteinte par des compensations des erreurs dans le modèle et pas nécessairement pour un meilleur représentation des processus physiques. Un exemple qu'essaie d'illustrer la suggestion précédente, est le paradoxe CFL: les résultats obtenus par le MRCC lorsque celui est piloté par le modèle global sont parfois plus proches des valeurs observées que lorsque celui-ci utilise les réanalyses NCEP/NCAR comme pilote. Bien que les réanalyses

NCEP/NCAR soient sujettes à des erreurs, il ne fait aucun doute qu'elles constituent une estimation plus fiable de l'évolution de l'état de l'atmosphère comparativement aux données simulées par le MCGC. Comment est-t-il donc possible d'obtenir de "meilleurs" résultats en utilisant des données de pilotage d'une qualité inférieure? Les erreurs dans les CFL doivent nécessairement être compensés par le MRCC. Ceci suggère que les meilleurs résultats proviennent d'erreurs dans les MRCC.

L'analyse de la valeur ajoutée générée par la MRC est un travail complexe. Nous avons essayé de contribuer à cette importante étude de deux manières différentes : a) par la quantification de la valeur ajoutée dans un cas particulier; et b) en débattant sur certaines questions générales qui doivent être prises en compte dans ce type d'analyse. Parmi les nombreuses questions importantes restant en suspens, nous pouvons inclure :

- Qu'est-ce que la résolution effective des modèles climatiques? Quelle est la dépendance de l'efficacité de résolution sur l'échelle de temps, de la variable analysée?
- Les MRC présentent certains avantages comparés aux modèles de circulation générale à plus basse résolution comme une meilleure représentation des forçages de surface, un plus large éventail de processus résolus, etc. Quelle est l'importance relative de chaque avantage dans la production de la valeur ajoutée?
- Qu'est-ce que la dépendance de la valeur ajoutée dans les échelles temporelles et spatiales?
- Y a t-il une valeur ajoutée à grande échelle dans les résultats?

Enfin, il peut être très intéressant de répéter l'analyse avec d'autres modèles tels que ceux impliqués dans le projet européen PRUDENCE (Prediction of Regional Scenarios and Uncertanties for Defining European Climate Change Risks and Effects) et/ou dans le programme d'évaluation des changements climatiques régionaux de l'Amérique du Nord (NARCAAP en anglais). L'utilisation de plusieurs modèles aiderait à déterminer si certaines des conclusions sont inhérentes à la technique dynamique de mise à l'échelle ou liées à un modèle particulier.

RÉFÉRENCES

Antic, S., R. Laprise, B. Denis, et R. de Elía. 2004. « Testing the downscaling ability of a one-way nested regional climate model in regions of complex topography ». *Clim. Dyn*, 23, 473-493.

Boer, G. J., et M. Lazare. 1988. « Some results concerning the effect of horizontal resolution and gravity-wave drag on simulated climate ». *J. Climate*, 1, 789–806.

Castro, C. L., R. A. Pielke, Sr., et G. Leoncini. 2005. « Dynamical Downscaling: An Assessment of Value Added Using a Regional Climate Model ». *J. Geophys. Res.*, 110, D05108, doi: 10.1029/2004JD004721.

Christensen J. H., E. Sanchez et E. Kjellström. 2008. « Probability density distribution match of daily and monthly temperature and precipitation analysis (f_3) ». Deliver 3.2.2:RCM-specific weights based on their ability to simulate the present climate, calibrated for the ERA40-based simulations, project ENSEMBLE-based Predictions of Climate Changes and their Impacts.

Dai, A., F. Giorgi, et K. E. Trenberth. 1999. « Observed and model simulated precipitation diurnal cycle over the contiguous United States ». *J. Geophys. Res.*, 104, 6377-6402.

Dai, A., et K. E. Trenberth. 2004. « The diurnal cycle and its depiction in the Community Climate System Model ». *J. Climate*, 17, 930-951.

de Elía, R., R. Laprise et B. Denis. 2002. « Forecasting Skill Limits of Nested, Limited-Area Models: A Perfect Model Approach ». *Mon. Wea. Rev.*, 130, 2006-2023.

Denis, B., J. Côté et R. Laprise. 2002a. « Spectral Decomposition of Two-Dimensional Atmospheric Fields on Limited-Area Domains Using the Discrete Cosine Transform (DCT) ». *Mon. Wea. Rev.*, 130, 1812-1829.

Denis, B., R. Laprise, D. Caya et J. Cote. 2002b. « Downscaling Ability of One-Way Regional Climate Models: The Big-Brother Experiment », *Clim. Dyn.*, 18, 627-646.

Dickinson, R. E., R. M. Errico, F. Giorgi et G. T. Bates. 1989. « A regional climate model for the western United States ». *Clim. Change*, 15, 383-422.

Dimitrijevic, M., et R. Laprise. 2005. « Validation of the nesting technique in a RCM and sensitivity tests to the resolution of the lateral boundary conditions during summer ». *Clim. Dyn.*, 25, 555-580.

Durman, C. F., J. M. Gregory, D. C. Hassell, R. G. Jones et J. M. Murphy. 2001. « A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates ». *Q. J. R. Meteorol. Soc.*, 127, 1005-1015.

Feser, F. 2006. « Enhanced Detectability of Added Value in Limited-Area Model Results Separated into Different Spatial Scales ». *Mon. Wea. Rev.*, 134, 2180–2190.

Frei, C., J. H. Christensen, M. Déqué, D. Jacob, R. G. Jones et P. L. Vidale. 2003. « Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps ». *J. Geophysic. Res.*, 108.

Giorgi, F., et G. T. Bates. 1989. « The climatological skill of a regional model over complex terrain ». *Mon. Wea. Rev.*, 117, 2325-2347.

Giorgi, F., et M. R. Marinucci. 1996. « An investigation of the sensitivity of simulated precipitation to the model resolution and its implications for climate studies ». *Mon. Wea. Rev.*, 124, 148-166.

Herceg, D., A. H. Sobel, L. Sun et S. E. Zebiak. 2006. « The big brother experiment and seasonal predictability in the NCEP regional spectral model ». *Clim. Dyn.*, 26(4), 1-14. (DOI 10.1007/s00382-006-0130-z; http://dx.doi.org/10.1007/s00382-006-0130-z)

Koltzow M., T. Iversen et J. E. Haugen. 2008. « Extended Big-Brother experiments: the role of lateral boundary data quality and size of integration domain in regional climate modeling ». *Tellus A*, 60, 398–410.

Laprise, R. 2003. « Resolved scales and nonlinear interactions in limited-area models ». *J. Atmos. Sci.*, 60 (5), 768-779.

Laprise, R. 2008. « Regional climate modelling ». J. Comp. Phys. 227, Special issue on «Predicting weather, climate and extreme events» (article par invitation), 3641–3666.

Laprise, R., R. de Elía, D. Caya, S. Biner, Ph. Lucas-Picher, E. P. Diaconescu, M. Leduc, A. Alexandru et L. Separovic. 2008. « Challenging some tenets of Regional Climate Modelling ». Meteor. Atmos. Phys. 100, Special Issue on Regional Climate Studies (article par invitation), 3-22 (DOI 10.1007/s00703-008-0292-9).

Oreskes N., K. S. Shrader-Frechette et K. Belitz. 1994. « Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences ». *Science*, 263, 641–646.

Osborn, T. J. et M. Hulme. 1997. « Development of a Relationship between Station and Grid-Box Rainday Frequencies for Climate Model Evaluation ». *J. Climate*, 10(8): 1885.

Paquin D., D. Caya et R. Laprise. 2002. « Treatment of moist convection in the Canadian Regional Climate Model ». Ouranos, Équipe Simulations climatiques. Rapport interne no 1, 30 p.

Perkins, S. E., A. J. Pitman, N. J. Holbrook, et J. McAneney. 2007. « Evaluation of the AR4 Climate Models' Simulated Daily Maximum Temperature, Minimum Temperature, and Precipitation over Australia Using Probability Density Functions ». *J. Climate*, 20, 4356-4376, doi: 10.1175/JCLI4253.1.

Randall, D. A., R. A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R. J. Stouffer, A. Sumi et K. E. Taylor, 2007: Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Trenberth, K. E., A. Dai, R. M. Rasmussen et D. B. Parsons. 2003. « The changing character of precipitation ». *Bull. Amer. Meteor. Soc.*, 84, 1205-1217.