UNIVERSITÉ DU QUÉBEC À MONTRÉAL

EVALUATING THE CANADIAN REGIONAL CLIMATE MODEL THROUGH PROCESS TENDENCIES AND DATA ASSIMILATION

THESIS

PRESENTED

AS PARTIAL REQUIREMENT

FOR PHD DEGREE IN EARTH AND ATMOSPHERIC SCIENCES

BY

KAMEL CHIKHAR

JANUARY 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

ÉVALUATION DU MODÈLE RÉGIONAL CANADIEN DU CLIMAT À TRAVERS LES TENDANCES DES PROCESSUS ET L'ASSIMILATION DE DONNÉES

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN SCIENCES DE LA TERRE ET DE L'ATMOSPHÈRE

PAR

KAMEL CHIKHAR

JANVIER 2016

REMERCIEMENTS

Je tiens à remercier sincèrement mon directeur de thèse le Prof. Pierre Gauthier, professeur au Département des Sciences de la Terre et de l'Atmosphère, pour avoir accepté de diriger cette recherche en continuation à notre précédente collaboration dans le cadre de ma maîtrise. Cette recherche n'aurait jamais été possible sans son entière implication à travers le support scientifique, ses précieux conseils, son exemplaire disponibilité et sa grande générosité.

Je tiens également à remercier le Dr. Bernard Dugas pour ses précieux et incontournables conseils concernant l'utilisation des différents modèles numériques utilisés. Mes remerciements vont également à Madame Katja Winger, Monsieur Michel Valin et la Dre Ping Du pour leur inestimable assistance technique. Je remercie également les professeurs et collègues du Département des Sciences de la Terre et de l'Atmosphère de l'UQAM et du Centre sur l'étude et la simulation du climat à l'échelle régionale (ESCER) qui ont contribué de près ou de loin à l'aboutissement de ce travail.

Ce projet a été financé par le Ministère du Développement Économique, de l'Innovation et de l'Exportation (MDEIE) du Québec, Environnement Canada, le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), le Réseau canadien pour le climat régional et les processus météorologiques (en anglais *Canadian Network on Regional Climate and Weather Processes*, CNRCWP) ainsi que par des bourses d'études FARE de la Faculté des sciences de l'UQAM. Les ressources en calcul de haute performance ont été fournies par le consortium régional Calcul Québec de Calcul Canada sur la plateforme Guillimin.

Je terminerai en remerciant mon épouse Ounissa pour sa patience, son soutien et ses encouragements.

ÉNONCÉ D'ORIGINALITÉ ET DE CONTRIBUTION PERSONNELLE

Dans cette étude, le modèle régional canadien du climat (MRCC5) a été évalué en utilisant des outils généralement appliqués en prévision numérique du temps. La problématique de l'équilibre initial ainsi que l'impact des conditions aux frontières a été traitée dans un cadre à la frontière entre la prévision numérique du temps et la modélisation du climat.

Les aspects originaux présents dans cette étude sont :

- L'application de diagnostiques de prévision numérique du temps à un modèle régional de climat
- La réalisation de cycles d'assimilation de données avec un modèle à aire limitée, le MRCC, sur des périodes de un à deux mois
- L'étude de l'impact des conditions latérales sous une nouvelle perspective

Les résultats de l'étude visent à valider le modèle régional canadien du climat pour éventuellement produire des réanalyses régionales. En s'appuyant sur le MRCC5 et le système d'analyse variationnelle d'Environnement Canada, on bénéficie du travail considérable de validation de ces deux composantes.

Les résultats de cette recherche ont également permis de mettre en lumière le potentiel de nouvelles approches de validation du modèle régional canadien du climat permettant de mieux comprendre l'action des processus physiques et de les améliorer.

CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS

3D-Var	Three-dimensional variational data assimilation
3D-Var FGAT	Three-dimensional variational data assimilation with first guess at appropriate time
4D-Var	Four-dimensional variational data assimilation
4DEnvar	Four-dimensional ensemble-variational data assimilation
CanESM2	Second-generation Canadian Earth System Model
CGCM	Coupled Global Climate Model
CLASS	Canadian Land Surface Scheme
CORDEX	COordinated Regional Climate Downscaling EXperiment
CRCM5	Fifth-generation Canadian Regional Climate Model
ECMWF	European Centre for Medium-Range Weather Forecasts
EnVar	Ensemble-Variational data assimilation
ERA-Interim	European Centre for Medium-Range Weather Forecasts reanalyses data
ESCER	Étude et Simulation du Climat À l'Échelle Régionale
GEM	Global Environmental Multiscale
GEMCLIM	CRCM5 global version used in ESCER centre
GEM-LAM	Global Environmental Multiscale Limited Area Model
IFS	Integrated Forecast System
ITCZ	Intertropical Convergence Zone

LAM	Limited Area Model
MERRA	Modern-Era Retrospective Analysis for Research and Applications
MPI-ESM-LR	Max Planck Institute for Meteorology's Earth System Model (Low Resolution)
MSC	Meteorological Service of Canada
MSC3D	MSC analyses obtained by using 3D-Var data assimilation system
MSC4D	MSC analyses obtained by using 4D-Var data assimilation system
NARR	North American Regional Reanalysis
NCEP	National Centers for Environmental Prediction
NWP	Numerical Weather Prediction
QC-Var	Variational quality control
RCM	Regional Climate Model
SST	Sea Surface Temperature

RÉSUMÉ

L'objectif de cette thèse était d'utiliser une nouvelle approche basée sur celle proposée par Rodwell and Palmer (2007) (désignée dorénavant par RP07) pour valider le Modèle Régional Canadien du Climat (MRCC5). Cette appproche utilise les tendances temporelles initiales des processus physiques du modèle pour évaluer l'équilibre dynamique résultant de l'utilisation de différentes conditions initiales fournies soit par des réanalyses ou des simulations globales du climat. Ceci apporte des informations détaillées sur les interactions entre ces différents processus et complète les outils classiques de validation des modèles régionaux de climat. Pour un modèle de climat global, l'impact des conditions initiales s'estompe en cours d'intégration et n'a donc pas d'influence directe sur une simulation climatique. Par contre, un modèle régional de climat est constamment forcé à ses frontières soit par des réanalyses globales ou le résultat d'une simulation globale du climat produite à l'aide d'un modèle ayant des caractéristiques différentes. Cette thèse propose d'étudier en premier lieu la sensibilité d'un modèle régional du climat aux conditions initiales et aux frontières telle qu'évaluée par les diagnostiques sur les tendances initiales. Le MRCC5 (CRCM5 en anglais) est une version à aire limitée du modèle global GEM (Global Environmental Multi-échelle) d'Environnement Canada. Les tendances initiales ont d'abord été calculées pour la configuration globale de ce modèle pour un ensemble de courtes intégrations démarrées d'analyses issues d'un système d'assimilation utilisant le même modèle pour produire des analyses 3D-Var et 4D-Var d'Environnement Canada (EC). Le modèle est généralement bien équilibré mais une dégradation est notée au niveau de la zone de convergence inter-tropicale (ZCIT) pour le 4D-Var. Lorsque le modèle démarre d'une analyse produite par un autre modèle, dans notre cas les réanalyses ERA-Interim du Centre Européen de prévision météorologique à moyen terme (CEPMMT, ECMWF en anglais) à basse et haute résolution, une quasi-absence de la convection est notée dans les premiers instants de l'intégration lorsque les réanalyses à basse résolution sont utilisées. L'équilibre est meilleur avec les réanalyses à haute résolution. Il est également noté qu'une dégradation de la résolution verticale est plus dommageable que celle de la résolution horizontale. Pour le MRCC5, les conditions aux frontières utilisées pour le pilotage ont été, d'une part des analyses 4D-Var d'EC et d'autre part des réanalyses ERA-Interim à basse et haute résolution. Les diagnostics des tendances montrent un déséquilibre persistant au-delà de 15 jours dans le cas des réanalyses à faible résolution. Dans le

cas des réanalyses ERA-interim à haute résolution et les analyses 4D-Var d'Environnement Canada, le modèle affiche un bon équilibre initial mais une dégradation commence à apparaître après 5 jours environ en se propageant graduellement vers l'intérieur du domaine. Dans la deuxième partie, l'effet du pilotage sur le MRCC5 est évalué par sa réponse à différents types de conditions aux frontières produites par le modèle Canadien du climat (CanESM2), le modèle de l'Institut Max-Planck (MPI-ESM) ainsi que les réanalyses ERA-interim. Les diagnostics des tendances montrent un réchauffement excessif du modèle dans les basses couches et un refroidissement plus haut et ceci pour les différentes données de pilotage. Évaluées et moyennées sur une saison, les tendances de température ont révélé un refroidissement dans les basses couches lorsque le modèle est piloté par CanESM2 et un réchauffement lorsqu'il est forcé par MPI-ESM. Les tendances de l'humidité spécifique montrent également des effets différents selon le type de données de pilotage. Dans RP07, les tendances initiales sont obtenues dans le cadre d'un cycle d'assimilation de données dans lequel le modèle évalué est celui utilisé pour faire l'assimilation. La troisième partie de cette thèse présente l'évaluation d'un système régional d'assimilation de données basé sur le MRCC5 sur la région de l'Amérique du Nord. Le système d'assimilation variationnel d'Environnement Canada a été adapté pour produire des analyses aux 6-h pour les mois de Janvier et Juillet 2011. Bien que celles-ci concordent assez bien avec celles d'autres sources, les incréments d'analyse moyens, les écarts aux observations moyens (monitoring) et les tendances initiales ont révélé des anomalies près de la frontière nord du MRCC5 indiquant un problème avec le couplage aux frontières en présence de rapides variations dans la circulation transversale. La possibilité de faire de l'assimilation avec le MRCC5 ouvre la possibilité d'étudier différentes approches permettant d'aborder cette question et les diagnostics sur les tendances constituent un atout important pour élucider cette question.

Mots clés : Assimilation de données, équilibre dynamique, modélisation régionale du climat, processus physiques, modèles à aire limitée.

ABSTRACT

The objective of this thesis is the assessment of the fifth generation Canadian Regional Climate Model (CRCM5) using a new approach based on that proposed by Rodwell and Palmer (2007) (RP07, hereafter). In this approach, initial tendencies of resolved physical and dynamical processes are used to evaluate the dynamical balance when the model is initialized and driven by reanalyses of different sources or outputs from global climate simulations. This technique provides detailed information on interactions between these different processes and complements classical tools used in regional climate models validation. For global climate models, the impact of the initial conditions fades away progressively during the integration and consequently has no effect on the climate simulation. In the case of regional climate models, they are constantly forced at their lateral boundaries by global reanalyses or outputs from global climate models with different characteristics. This thesis aims to first examine sensitivity of regional climate models to initial and boundary conditions using the initial tendencies diagnostic.

The model used here is the fifth-generation Canadian Regional Climate Model (CRCM5), a limited-area version of the global model GEM (Global Environmental Multiscale) of Environment Canada. The initial tendencies are first computed for the global GEM model configuration by performing an ensemble of short integrations initialized by analyses from an assimilation system using the same model producing 3D-Var and 4D-Var analyses of Environment Canada. Results show that the model is fairly well balanced with a slight degradation in the case of 4D-Var analyses in the inter-tropical convergence zone (ITCZ). When the model is initialized by analyses based on a different model, in this case low and high resolution ERA-Interim reanalyses from the European Center for Medium-range Weather Forecasts (ECMWF), the results show a quasi-absent convection in the first steps of the integration when low resolution reanalyses are used. The reason could be a lack in humidity in lower levels preventing the triggering of convection. It is also noted that degrading the vertical resolution is more damaging than using a coarser horizontal resolution.

CRCM5 simulations have been conducted using driving data from 4D-Var analyses and ERA-Interim reanalyses at low and high resolution and the results showed persisting imbalances even after 15 days for the coarser reanalyses. When considered in the free interior zone, the model exhibits relatively good balance initially.

However, after 5 days of integration imbalances reappeared gradually, propagating into the interior of the model domain.

The second part looked at the effect of different driving data on the regional climate response by prospecting process tendencies. For this, the CRCM5 is assessed when driven by different boundary conditions. The latter are supplied by outputs from the second generation Canadian Earth System Model (CanESM2) and the Max Planck Institute for Meteorology's Earth System Model (MPI-ESM-LR) and also from ECMWF ERA-Interim reanalyses. Process tendencies diagnostics show a model excessive heating in lower levels and a cooling at higher levels for all different driving data. Temperature tendencies evaluated and averaged over a season revealed a cooling in lower levels when the model is driven by CanESM2 and a heating when driven by MPI-ESM2. Results from specific humidity tendencies also showed various impacts depending on the driving data used.

In the third part, results from regional data assimilation experiments using the CRCM5 over North America are presented. This system is adapted from the variational data assimilation system of Environment Canada to produce 6-h analyses for January and July 2011. These analyses are close to those obtained with other systems (e.g., ERA-Interim, MSC analyses). However, the mean analysis increments, mean departures from observations (monitoring) and initial tendencies revealed anomalies near the model Northern boundary due to problems in the driving procedure in the presence of rapid variations in transversal circulation. A regional data assimilation based on the CRCM5 makes it possible to evaluate different approaches to study this issue and, as proposed in RP07, initial tendencies diagnostics will be very useful for this.

Key Words : Data assimilation, dynamical balance, regional climate modeling, physical processes, limited-area models.

INTRODUCTION

Context

Climate simulations are obtained by using very complex climate models which take into account several physical processes to represent in the best possible way the evolution of the climate. The complexity of such coupled models is associated with its various components (e.g., atmosphere, ocean, biosphere, cryosphere, etc.) which leads to increasing sources of uncertainty in climate simulations. One of the main source of uncertainty are model errors especially their deficiencies in the treatment of physical processes. This is mainly due to their coarse resolution but also to lack of scientific understanding of these physical processes and how they interact with one another. Some studies treated the uncertainty related to model errors by using probabilistic approach in which an ensemble of model's versions is employed (Stainforth et al., 2005; Murphy et al., 2004) to produce an ensemble of climate simulations leading to probability density functions (PDFs). This approach is computationally expensive and showed some limitations on how to construct the ensemble members as pointed out by Rodwell and Palmer (2007)(RP07 hereafter). Indeed, they showed that the ensemble can include artificially unbalanced members, leading to an unrealistic appraisal of internal variability. Furthermore, this internal variability has been shown to be very sensitive to the parameterizations used in the model (Crétat and Pohl, 2012).

The origins of model error are still difficult to identify and represent a challenging issue. The most popular way to assess model errors is the confrontation to observations which also contain some error, including biases and variances. These metrics provide an estimate of model skill without further information or details related to eventual deficiencies in representing particular physical processes. Additional tools are needed to provide a more detailed assessment of model error.

Another issue in climate model validation is the high computational cost involved as the model to be assessed is often run for a long period and compared to the observed climate. In numerical weather prediction, diagnostic tools are used to evaluate the realism of fast acting physical processes and may be useful to evaluate the same processes in use in climate models. This could be less time consuming than having to run long climate simulations. These tools should be complementary to the traditional assessments procedures and can be seen as preliminary tests with a lower computational cost.

For regional climate models (RCMs), contrary to global climate models, the model error issue is complicated by the fact that RCMs are limited-area models (LAMs) that must be driven continuously by external data provided by coupled global climate models (CGCMs) or reanalyses. This driving or nesting can create inconsistencies in the blending zone that can penetrate in the interior of the RCM model domain. This is due to several factors such as the coarser spatial resolution in driving data and differences in models themselves especially the way they represent physical processes.

The total initial tendency diagnostic

During a model integration, at each time step, the atmospheric state is updated. This involves different physical and dynamical processes affecting the prognostic variables. RP07 proposed a diagnostic based on the physical processes tendencies computed in the first moments of the model integration, typically 6h. They showed that, when averaged over a large number of integrations, this measure is equivalent to the mean analysis increment brought in by the data assimilation. On the other hand, a mean analysis increment averaging to zero is an indication that the model used is unbiased with respect to observations and that the assimilation is well balanced. Given the equivalence to the mean analysis increment, one could use the initial tendency diagnostic to assess the dynamical equilibrium of the model. This approach can then easily be used, provided the model assessed is the same as the one used in the data assimilation system. The RP07 procedure consists in an ensemble of short integrations started from analyses obtained from an assimilation system using the same model. As we average over the first

moments of the integration, only the fast-acting physical processes are taken into account in the initial tendency computation. For a variable X, the mean total initial tendency can be expressed as,

$$\left\langle \left(\frac{\partial X}{\partial t}\right)_{Total} \right\rangle = \left\langle \left(\frac{\partial X}{\partial t}\right)_1 + \left(\frac{\partial X}{\partial t}\right)_2 + \ldots + \left(\frac{\partial X}{\partial t}\right)_N \right\rangle$$

where subscripts $1, 2, \ldots, N$ represent the different physical processes involved in the X variable update. The symbol $\langle \ldots \rangle$ denotes an average in time and space as the average can be done spatially over a particular region or globally.

Using an approach introduced by Klinker and Sardeshmukh (1992), RP07 used the initial tendency diagnostic to assess the European Center for Medium-range Weather forecasts (ECMWF) model, over a particular region, by modifying some parameters used in the convection parameterization scheme. They demonstrated the effect of those perturbed parameters on the model's dynamical balance and how this diagnostic points to inconsistencies in the model physics. In their study, RP07 modified the entrainment rate used in the convection scheme to regulate the moisture mixing between the convective plume and the environment. They found that this parameter setting can degrade significantly the dynamical balance. They also used this diagnostic to examine the improvement in the model balance when a new cloud scheme is incorporated as is often done in numerical weather prediction. They also made the point that this approach has a computational cost of only $\sim 5\%$ of that associated with the evaluation by performing several climate simulations. Rodwell and Jung (2008) used the same diagnostic on the ECMWF model when a new aerosol climatology is used and showed its impact on the radiation parameterization scheme and the subsequent impacts on other processes.

Regional data assimilation

RP07 have shown that the initial tendency diagnostic suitably averaged is a useful diagnostic to evaluate the dynamical balance of a model. A good balance is to be expected when the different physical and dynamical processes are consistent with one another. RP07 also showed that this diagnostic provides a detailed descrip-

tion on how the model treats every process and can help to identify which one is poorly or badly represented and how each process reacts to a change in the model. However, to use this diagnostic, the model evaluated has to be used within a data assimilation system providing the analyses used to initialize that model. Consequently, this is a good incentive to use a climate model within a data assimilation system to be able to apply this initial tendency diagnostic. Moreover, assessing a regional climate model using this diagnostic requires a regional data assimilation system.

Implementing a fully cycled regional data assimilation is a challenging task requiring a detailed validation of components associated with observations, and error statistics of observations and the short-term forecast of the climate model which would provide the *a priori* estimate referred to as the *background state*. Moreover, regional data assimilation needs to face the difficulties associated with the fact that the model is driven at its lateral boundaries. This raises issues in the way observations must be treated in the particular blending zone where discontinuities or strong gradients can occur.

A regional data assimilation system based on the fifth generation Canadian regional climate model (CRCM5) (Zadra et al., 2008), developed at the ESCER (Étude et Simulation du Climat à l'Échelle Régionale) centre, would certainly be very useful to the model development using the initial tendency diagnostic to its fullest. Moreover, the assimilation would provide a direct comparison to observations through the so-called innovations, defined as the departure between the background state and the observations. Furthermore, this regional data assimilation system can be a first step towards a reanalysis at a higher resolution, which would be useful for regional climate studies.

Previous work

Initial and lateral conditions impact on the model simulation represents an important aspect that needs to be considered with care. Wu et al. (2005) have shown that initial and lateral conditions have a significant impact on climate simulations. When used as initial conditions, analyses can produce imbalances in the first mo-

ments of the integration while the lateral conditions can induce biases. Analyses are known to produce spinup imbalances associated with spurious gravity waves (Gauthier and Thépaut, 2001). This issue is addressed using digital filters (Lynch, 1997) or normal mode initialization (Machenhauer, 1977). In a previous work, Chikhar (2011) studied imbalances associated with the initial conditions for the Global Environmental Multiscale model (GEM). Using 3D-Var and 4D-Var analyses from the Meteorological Service of Canada (MSC) (Gauthier et al., 1999, 2007) produced with a model very similar to the one being assessed, the results showed significant differences in the Tropics between 3D-Var and 4D-Var experiments. In the case of 4D-Var, the convective activity was found to be more intense particularly in the Intertropical Convergence Zone (ITCZ). This counter-intuitive result could be explained by the simplified physics in the MSC 4D-Var system which does not include convection. The impact of non-native or external analyses was also examined. Using ERA-Interim reanalyses from ECMWF, the model used in the data assimilation system is then different from the GEM model to be assessed. The initial tendency diagnosed revealed clearly the so-called 'initial shock' (Klocke and Rodwell, 2014) degrading considerably the model's balance due to inconsistencies between the two models. The initial tendency diagnostic showed that the vertical diffusion is very strong in the first moments of the integration whereas the convection process is nearly absent. The presence of a very stable lower layer as well as drier analyses (in the Tropics) could explain the non triggering of the convection.

The Canadian Regional Climate Model (CRCM5) dynamical equilibrium was also investigated. In this case, the model is supplied with different initial and boundary conditions. Results showed that the CRCM5 is better balanced when initialized and driven by MSC analyses, reflecting the fact that the CRCM5 is built from the limited-area configuration of the GEM model (same physics and similar resolution). It was also shown that the global model GEM is better balanced over the CRCM5 domain than the regional model. This result suggests that a higher resolution regional analysis could be better suited for the CRCM5. When initialized and driven by ERA-Interim reanalyses, the CRCM5 initial balance shows similar behavior as when it is driven by the global GEM analyses.

Objectives and methodology

The main objectives of this thesis is to investigate the potential benefits of using the tendency diagnostic as an assessment tool in the CRCM5 development, as well as applying new tools made available from a regional data assimilation system. This thesis is structured in the form of scientific papers presented as chapters.

In the first chapter, it is aimed to answer questions such as "How do initial and lateral conditions affect the CRCM5 dynamical balance in the first moments of the integration ?" and "To what extent initial conditions resolution impact the model dynamical balance ?". To address these questions, the initial tendency diagnostic is employed using different initial and driving data as MSC 4D-Var analyses and ERA-Interim reanalyses at low and high resolution. In the second chapter, the question to be answered was "How does the CRCM5 respond to various boundary conditions for longer integrations ?". To address this issue, 14 month simulations are performed using different driving data (outputs from CanESM2 and MPI-ESM models and ERA-Interim reanalyses). Analysis and comparison between the different simulations are done through the process tendencies computed and averaged over a season. Finally, the third chapter aims to present and evaluate a regional data assimilation system based on CRCM5 over North America, the objective being the validation of the model using the RP07 initial tendency diagnostic. This regional system is adapted from the Envar data assimilation system of Environment Canada and assimilates the same sets of observations. Fully cycled analyses are produced every 6-h over the months of January and July. Results are evaluated through the examination of mean analysis increments and data monitoring in addition to comparison to analyses from other sources.
The overall objective of this thesis is to provide new tools to assess CRCM5 through different perspectives. This will hopefully contribute to the model development.

CHAPTER I

IMPACT OF ANALYSES ON THE DYNAMICAL BALANCE OF GLOBAL AND LIMITED-AREA ATMOSPHERIC MODELS

This chapter is presented in the format of a scientific article. It is published in the Quarterly Journal of the Royal Meteorological Society. This article is entirely based on my work with the co-author involved in supervising all the tasks and text revision. The detailed reference is:

Chikhar K, Gauthier P (2014) Impact of analyses on the dynamical balance of global and limited-area atmospheric models. Q.J.R. Meteorol. Soc., 140: 2535–2545. doi: 10.1002/qj.2319

8

Impact of analyses on the dynamical balance of global and limited-area atmospheric models

Kamel Chikhar and Pierre Gauthier

Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale),
Département des Sciences de la Terre et de 1'Atmosphère
Université du Québec à Montréal (UQAM)
B.P. 8888, Succ. Centre-ville
Montréal (Québec) Canada H3C 3P8

Abstract

Dynamical imbalances can induce spurious variability which can be diagnosed from the physical tendencies observed in the first moments of short-term forecasts using as initial conditions analyses obtained from an assimilation system using this model. In this article this approach is taken to investigate differences in the balance obtained from 3D- and 4D-Var analyses, using the forecast–assimilation system of the Meteorological Service of Canada (MSC). The results indicate that the model is then in good balance globally but the 4D-Var analyses slightly upset the balance in the Tropics, thereby altering the characteristics of the Intertropical Convergence Zone (ITCZ). As the assimilation is driven by a particular model, the resulting analyses keep an imprint of the dynamics of that model and use of this analysis with another model may not be as well in balance due to the differences between the two models. To study this point, ERA-Interim 4D-Var reanalyses were used as initial conditions first at a lower horizontal and vertical resolution, and then at a resolution closer to that of the Global Environmental Multiscale (GEM) model. The higher-resolution reanalyses led to a better balance than that with a lower-resolution version of the ERA-Interim reanalyses. The coarser analyses create significant imbalances in the Canadian global model which persist for more than 5 days. In particular, it was noted that convection is nearly absent early on as if at a lower resolution, the ERA-interim analyses did not inject sufficient humidity to trigger convection. It was also noted that reducing the vertical resolution is more damaging than using a coarser horizontal resolution. In limited-area regional climate models, external analyses are used to define the boundary conditions and the Canadian Regional Climate Model (CRCM) was used to assess the impact of different ways to define the boundary conditions. The CRCM is a limited-area configuration of the GEM global model used in the 3D- and 4D-Var assimilation. Experiments were conducted in which the boundary conditions driving the CRCM are provided every 6 h as is usually done for the CRCM climate simulations. When using 4D-Var analyses and ERA-Interim reanalyses (coarse and full resolution) to define the boundary conditions, the results indicate that imbalances persist even after 15 days and are more significant for the coarser analyses. Moreover, even though the model exhibits relatively good balance initially, after 5 days imbalances

appear gradually in the interior of the regional model domain.

Key Words : data assimilation; dynamical balance; physical processes; limited-area models; numerical weather prediction; regional climate modeling

1.1 Introduction

Future climate projections are obtained by using increasingly complex models. These projections are associated with uncertainties related to different error sources (Stainforth et al., 2005; Murphy et al., 2004). An important source of uncertainty is the model error. Indeed, models contain errors of different kinds and these can lead to unrealistic simulations. A recognized source of error in the models lies in deficiencies in the representation of subgrid physical processes. The model errors are usually quantified by comparing the forecasts to observations. This technique provides an estimation of the model errors without a precise information on their origin. The latter is much harder to identify especially when multiple error factors combine. Complementary tools are then needed to assess models' performance and identify eventual errors more clearly and with more details. Rodwell and Palmer (2007), hereafter RP07, proposed a diagnostic tool based on the initial systematic tendencies. It provides useful information about the consistency of the model's physics as it relaxes to its own climatology in the first moments of the integration. However, it is well known that analyses can themselves create spin-up problems that can be addressed using normal mode initialization or digital filtering to remove spurious gravity waves. Gauthier and Thépaut (2001) showed that 4D-Var analyses provide initial conditions that are better balanced and do not require as much the application of other constraints to maintain this equilibrium.

The approach proposed by RP07 consists in carrying an ensemble of successive short time integrations from which physical tendencies are extracted. Each integration starts from an analysis produced by a data assimilation system using the same model. The meteorological variable (e.g. temperature or specific humidity) tendencies associated with each individual physical process are then determined and, when averaged in time and over a specific region, they yield systematic initial tendencies which give a reliable measure of the model's dynamical equilibrium. A well-balanced model would yield systematic tendencies close to zero, while large values would indicate imbalances and then inconsistencies in the representation of physical processes of the model. In this article, this diagnostic is used first to examine the Global Environmental Multiscale (GEM) model's dynamical balance when initialized by its own 3D-Var and 4D-Var analyses produced at the Meteo-

rological Service of Canada (MSC) (Gauthier et al., 1999, 2007). One could also be interested to assess the equilibrium of the GEM model when initialized by 'external' analyses. In the present case, the European Centre for Medium-range Weather Forecasts (ECMWF) reanalyses 'ERA-Interim' (Dee et al., 2011) were used and reflect the particular equilibrium of the model used in the assimilation system. The results indicate that better balance is obtained if spatial resolutions are close, particularly in the vertical. Finally, as regional climate models are driven through lateral boundary conditions defined by a different model, we wanted to assess the dynamical balance of the Canadian Regional Climate Model (CRCM) when initialized and driven by MSC 4D-Var analyses or ERA-Interim reanalyses. To assess how long it takes to recover from an initial imbalance, the model balance is examined when the models are integrated over a longer period (15 days).

The article is organized as follows. In section 1.2, the method used to compute the total initial tendency is presented. The models used, the data assimilation system and the experiments achieved are described in section 1.3. The different results obtained are presented and discussed in sections 1.4 to 1.7. Further discussion and some conclusions are given in section 1.8.

1.2 Diagnosing dynamical balance based on averages of physical tendencies

When initial conditions defined by an analysis are used, the model may exhibit a transient behaviour that can persist for some time. To avoid this, numerical weather prediction (NWP) models have resorted to nonlinear normal mode initialization or digital filters to filter out the spurious gravity waves associated with this initial imbalance (Machenhauer, 1977; Baer and Tribbia , 1977). Klinker and Sardeshmukh (1992) examined the initial physical tendencies of the model as a measure of the imbalance. Rodwell and Palmer (2007) then pointed out that the average of the physical tendencies over a period of 6 h corresponds, but with the opposite sign, to the analysis increment brought in by the assimilation. Therefore, since the analysis increments average to zero over a large number of analyses, one would then expect that, similarly, the total physical tendency would also average out to zero. It is important that, for this to hold, the model used should be the same as the one used to do the assimilation.

In the RP07 procedure, the first time step is excluded because of its distinctive nature with respect to the subsequent ones. The analysis being used as initial conditions, the first time step has a strong imprint from the analysis. Moreover, a cold start is often used in which some physical tendencies are set to zero, knowing that they will be quickly restored within a few time steps. Averaging over a period of 6 h is a better approach to capture variations associated particularly with the diurnal cycle, when the thermal tendencies are considered. In their study, RP07 highlighted that these initial imbalances reflect potential flaws of fast-acting physical processes which can create artificial variability. This information could be particularly helpful to assess the value of multimodel climate simulations often used to estimate the uncertainty associated with climate scenarios (Stainforth et al., 2005). Another interesting application is that presented in Rodwell and Jung (2008), who used the tendency diagnostic to evaluate the improvement obtained when modifications are made to the forecasting system.

In our study, this RP07 procedure will be used to compute our tendency diagnostic. A detailed description is given in the next subsection.

1.2.1 Total initial tendency

The atmosphere is simulated by models through a representation of a set of physical and dynamic processes, and prognostic variables are updated by all related processes at each integration time step. The resulting total temporal tendency of a certain variable consists of the combination of the individual tendencies from all processes. The diagnostic for temperature tendency used in this study is defined as

$$\frac{1}{M}\sum_{i=1}^{M}\dot{T}_i^{total} = \frac{1}{M}\sum_{i=1}^{M}\sum_{p=1}^{k}\dot{T}_i^{p} \tag{1.1}$$

where M is the total forecasts number, $\dot{T}_i^{total}$ is the total tendency, and $\dot{T}_i^{p}$, $p = 1, ..., k$, represent the individual tendencies of the k physical processes, namely advection, shallow and deep convection, radiation, large scale condensation and vertical diffusion. As shown in RP07, the average of the analysis increments obtained over an assimilation window τ corresponds approximately to the average of the temperature tendencies taken also over the period. If $\tau = n\Delta t$ is the length of the assimilation window, and considering M consecutive analyses at times t_i with $i = 1, \ldots, M$, it can be shown that the average of the analysis increments corresponds approximately to the average physical total tendency

$$\begin{aligned}\frac{1}{M}\sum_{i=1}^{M}\delta\mathbf{x}_a^{(i)} &\cong -\frac{1}{M}\sum_{i=1}^{M}\left(\sum_{j=1}^{n}(T_i^{(j)} - T_i^{(j-1)})\right) \\ &= -\frac{\Delta t}{M}\sum_{i=1}^{M}\left(\sum_{j=1}^{n}\dot{T}_i^{(j)}\right)\end{aligned} \tag{1.2}$$

where $T_i^{(j)} = T_i(t_i + j\Delta t)$. A detailed demonstration of this correspondence can be found in RP07. All physical parameterization schemes are applied at each time

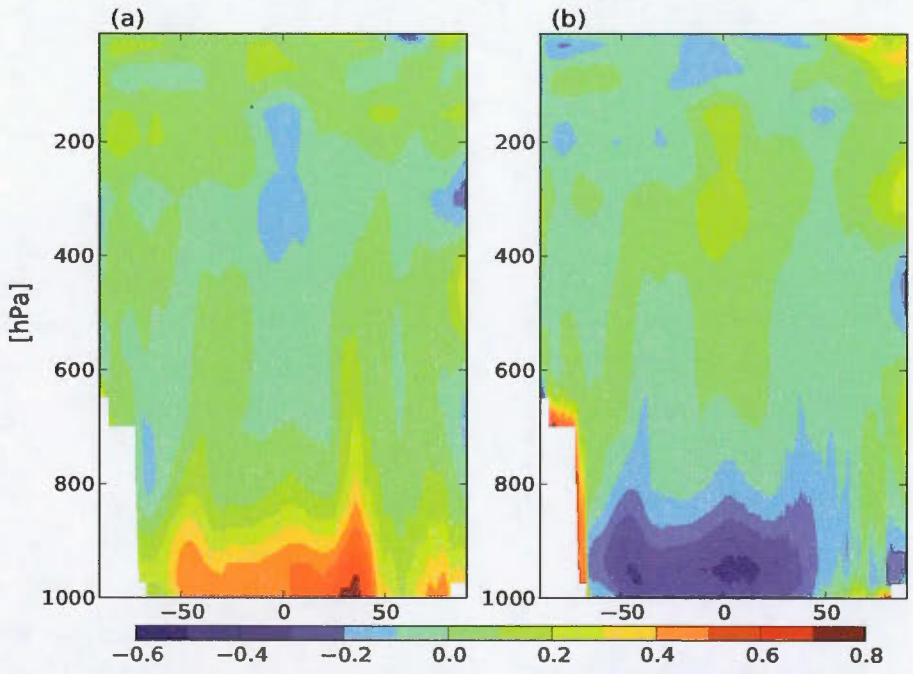

Figure 1.1 Zonal mean of (a) averaged temperature increments (K) and (b) averaged total physical tendency multiplied by the time step as expressed on the right-hand side of Eq. (1.2). The vertical coordinate is linear in pressure.

step so that temperature is updated every time step by all physical processes. The initial tendencies are averaged over a six hours period, the first time step being excluded as discussed previously. Tendencies are computed for each grid point and then averaged spatially over the whole globe as well as over specific regions like the Tropics, for example. These spatial averages are obtained by weighting every grid point by the size of the mesh. In order to evaluate the equivalence between the mean increment and the mean total physical tendency, an assimilation cycle for one month has been completed and the two sides of equation (1.2) have been computed for that period. The result is presented in figure 1.1 where this correspondence is clearly apparent.

1.3 Model, assimilation system and experiments

1.3.1 The Global Environmental Multiscale model

The Global Environmental Multiscale (GEM) model has been developed at the Meteorological Service of Canada (MSC) (Côté et al., 1998; Mailhot et al., 1998) and can be run in global mode with uniform or variable resolution, and also supports limited-area configurations for high-resolution regional forecasts. The global uniform configuration used in this study has a horizontal grid mesh of approximately 50 km at the Equator and uses a hybrid coordinate and has 80 vertical levels with a top at 0.1 hPa. The time step is 15 min. The model configuration considered here uses the following parametrization schemes : correlated-K solar and terrestrial radiations of Li and Barker (2005), Kain–Fritsch for deep convection (Kain and Fritsch, 1990, 1993), Kuo transient for shallow convection (Kuo, 1974), Sundqvist for large-scale condensation (Sundqvist, 1978), the vertical diffusion of Mailhot and Benoit (1982) and ISBA (Interactions between Soil–Biosphere–Atmosphere) for surface processes (Noilhan and Planton, 1989). A detailed description of the model physics can be found in Mailhot et al. (1998) and Bélair et al. (2009).

1.3.2 The Canadian Regional Climate model

The regional model used in this work is the CRCM5, a limited-area version of the GEM model (Zadra et al., 2008) that could be set over any area on the globe. In our study, the CRCM domain covers North America (Figure 1.2) with a 20 km horizontal grid mesh and 10 min time step. It has the same vertical discretization with the lid at 0.1 hPa and the same parametrization schemes as those of the global configuration described above. The CRCM lateral boundary conditions are supplied using a one-way nesting method (Davies, 1976; Yakimiw and Robert, 1990).

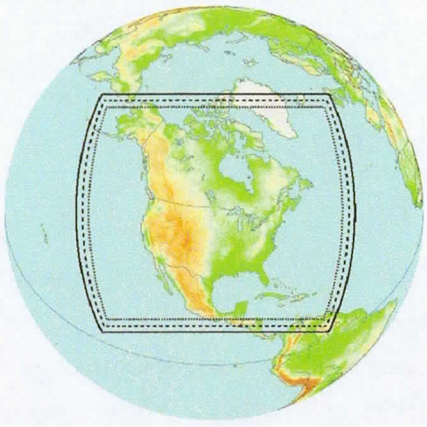

Figure 1.2 The CRCM domain.

1.3.3 Assimilation system

The three-dimensional variational data assimilation (3D-Var) system was implemented at MSC in 1997 and extended to four-dimensional variational data assimilation (4D-Var) in 2005 (Gauthier et al., 1999, 2007; Laroche et al., 2007). The variational formulation is based on the incremental approach (Courtier et al., 1994). The analysis increment is calculated at a lower resolution than the forecast model and is found by minimizing a cost function using the quasi-Newton algorithm (Gilbert and Lemaréchal, 1989). In 3D-Var, the background is obtained from a 6 h forecast and the analysis increment is determined at the centre of a 6 h assimilation window. Conventional data are assimilated over a 3 h window whereas radiances are assimilated over a 6 h window. In the 4D-Var system, the assimilation window is also 6 h and the background is now a trajectory obtained from a 9 h forecast. Data are assimilated over the whole window, thus increasing considerably their number compared to the 3D-Var system. A detailed description of these assimilation systems can be found in Gauthier et al. (1999, 2007). ERA-Interim reanalyses, used in this study, are produced at T255 ($\sim$ 79 km) horizontal grid mesh by a 4D-Var data assimilation system using a 12 h assimilation window. The forecast model used to produce the background is based on Integrated Forecast System (IFS) release Cy31r2 with T255 spectral horizontal resolution and 60 vertical levels with a top at 0.1 hPa. A detailed description of the ERA-Interim

reanalysis can be found in Dee et al. (2011).

1.3.4 Configuration of the experiments

Several experiments are realized in order to assess the GEM global model dynamical balance as well as that of the CRCM. In the case of CRCM, different types of initial and boundary conditions are used, the objective being the assessment of their eventual effect on model equilibrium. The initial and boundary conditions used in this study are MSC 3D-Var and 4D-Var analyses and ECMWF ERA-Interim reanalyses at low and high resolution. These datasets will be noted in this article as MSC3D, MSC4D, ERA-low and ERA-high respectively. In each experiment, several sets of integrations are performed covering the month of January 2009. Short and medium-range simulations are started every 6 h from 0000 UTC on 1 January. We consider that the total number of integrations is sufficient to resolve synoptic variability.

In a different set of experiments, longer-range simulations are completed with the GEM global model and the CRCM. As before, runs are started every 6 h from 0000 UTC on 1 January. The aim is to examine the temporal evolution of the dynamical balance when different initial and boundary conditions (for the CRCM) are used.

1.4 Impact of the assimilation method

Here, we examine the potential effect of the assimilation procedure (3D-Var and 4D-Var) on the model dynamical balance. We expect an improved dynamical equilibrium when the model is initialized by 4D-Var analyses, the latter being known to be better balanced than 3D-Var analyses (Gauthier and Thépaut, 2001; Laroche et al., 2007). In this experiment, GEM model integrations are initialized by MSC3D and MSC4D produced by MSC using a GEM model configuration very similar to that used here.

Figure 1.3 shows the resulting temperature tendency profiles averaged over the globe and over the Tropics. The different line colours identify the physical

processes acting on the temperature, and the thick black line represents the net tendency (i.e. the total of all individual process tendencies) In general, profiles are similar for the two cases (MSC3D and MSC4D). However, when initialized by MSC4D, the model is shown to be slightly more unbalanced compared to simulations using MSC3D, especially in the Tropics (Figure 1.3(d)). When averaged globally, the balance is established in the middle troposphere between the radiative cooling (radiation) and the heating due to latent heat release from convection and large-scale condensation. In the Tropics, the equilibrium is found mainly between cooling due to radiation and downward movement (vertical advection) and latent heat release from convection. In the lower levels, the profiles are very similar except small differences in radiation tendencies around 950 hPa and shallow convection around 900 hPa.

A closer look at profiles in the Tropics reveals significant differences in convection tendencies between MSC4D and MSC3D simulation. The convection activity is found to be stronger in the case of MSC4D, resulting in more heating in the atmospheric mid-levels.To examine more deeply this difference, the tendency due to convection is examined at 500 hPa (Figure 1.4). We can clearly see that the convective activity along the Intertropical Convergence Zone (ITCZ) is more intense in MSC4D simulations, especially over the oceans. This 'excess' in convective activity is responsible for the imbalance showed by temperature profiles (Figure 1.3(d)). This is a counter-intuitive result because 4D-var analyses are known to be better balanced than those produced using 3D-Var system. Specific humidity tendencies (not shown) show larger deficits in water vapour content and lead to the same conclusions, i.e. more convection in the ITCZ for MSC4D simulations. Another issue is the simplified physics used in the MSC 4D-Var system (which does not include convection) and this could also explain this difference. As shown in Figure 10 of Mahfouf and Rabier (2000), using simplified physics without convection results in an increase in the precipitation in the first moments of integration. Comparing the humidity analyses obtained with 3D-Var and 4D-Var (not shown) indicates that the latter are moister. The resulting precipitation leads to increased temperature tendency, particularly in the Tropics. This is consistent with the results shown in Figure 1.3(b, d).

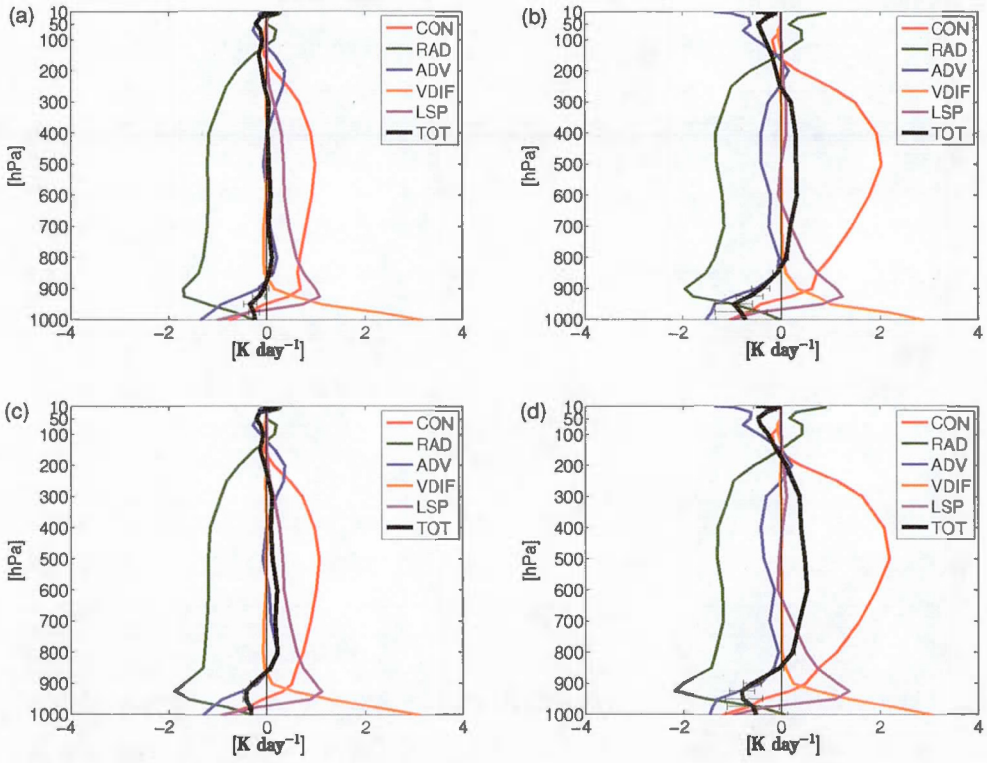

Figure 1.3 Temperature initial tendency profiles(K day^{-1}). Tendencies are computed for the first 6 h of the integration excluding the first time step. (a, b) show profiles obtained from simulations initialized by MSC3D and averaged (a) globally and (b) over the Tropics. (c, d) are as (a, b), but for integrations started from MSC4D. The different coloured lines correspond to the physical processes considered : radiation (green), advection (blue), convection (red), large-scale condensation (magenta) and vertical diffusion (orange) ; the black line shows the net tendency. The vertical coordinate is linear in pressure. Horizontal bars are 95% confidence intervals, showing very low values especially in the global case and above 800 hPa.

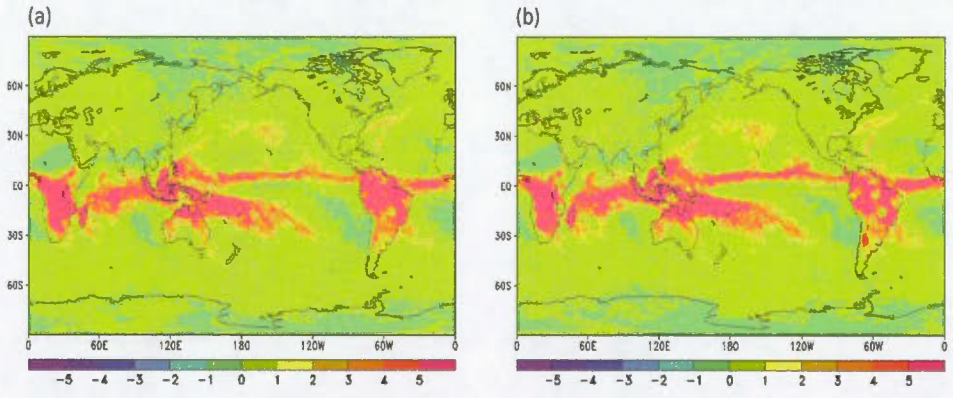

Figure 1.4 Mean temperature tendency due to convection (K day^{-1}) at 500 hPa from (a) MSC3D simulations and (b) MSC4D simulations.

The results presented in this section assessed the impact of analyses produced by assimilation systems driven by the same model as the one used to do the forecasts. In the next section, we examine the model balance when analyses are produced by a different model. The idea is to explore the interaction between a model and external analyses and compare the results obtained to those when the model is initialized from its own analyses.

1.5 Use of external analyses

In the previous experiment, the background state used in the assimilation producing the analyses is obtained from a model very similar to that we assessed, and this background state leaves an imprint on the analysis. The results obtained showed a fairly good dynamical balance. Now we investigate if this equilibrium is maintained when using external analyses based on a different model and assimilation system. An experiment identical to the previous one is conducted, except that the GEM model is now initialized by ECMWF ERA-Interim reanalyses. Initial conditions used in this experiment are the ERA-Interim atmospheric fields complemented by MSC surface analyses.

The ERA-Interim reanalyses were first used with a coarse mesh of 1.5° (around 150 km) and 37 levels in the vertical (ERA-low). Later, when they became available, the experiments were repeated using the reanalyses at higher resolution, that is 0.75° and 60 levels with a top at 0.1 hPa (ERA-high). This is to investigate not only the impact of external analyses on the model balance but also the impact of differences in horizontal and vertical resolution.

Systematic temperature tendencies profiles derived from this experiment are shown in Figure 1.5. Results indicate significant differences from what was obtained when using MSC3D or MSC4D. In this case, large imbalances are observed in profiles averaged globally as well as over the Tropics especially at low resolution where significant heating reaching 8 $K day^{-1}$ is present in the lowest levels induced by vertical diffusion and radiative heating. This indicates large vertical thermal gradients. In the above atmospheric levels, in both resolutions, we observe a cooling due mainly to a quasi-absence of deep convection activity. When comparing the zonal mean of specific humidity derived from the monthly temporal averages (not shown), we note that ERA-low is drier than the MSC4D in the tropical lower levels. This could contribute to prevent the triggering of deep convection. However, when full resolution reanalyses are used, a better dynamical balance is observed and the strong heating in the low levels disappears while the convective activity is more intense but still weak in the Tropics in particular. This result indicates the importance of the spatial resolution of the initial conditions in the horizontal as well as in the vertical. From this result, a question arises concerning the relative impact of each resolution (horizontal and vertical) on the model balance. This is confirmed by an experiment in which the initial conditions resolution of MSC4D was reduced horizontally and/or vertically. In the horizontal, the resolution went from 800 x 600 to 240 x 121 grid points, while in the vertical the original 80 vertical level analyses were interpolated to the 37 levels of the ERA-Interim at coarse resolution. The results based on the diagnostic of the initial tendencies on temperature are presented in Figure 1.6 and reveal that the model equilibrium is worse in both cases and that degrading the vertical resolution is more damaging than reducing the horizontal resolution. This result indicates that it is important to have a good correspondence in the resolution of the model and that of the initial conditions to maintain a good dynamical balance.

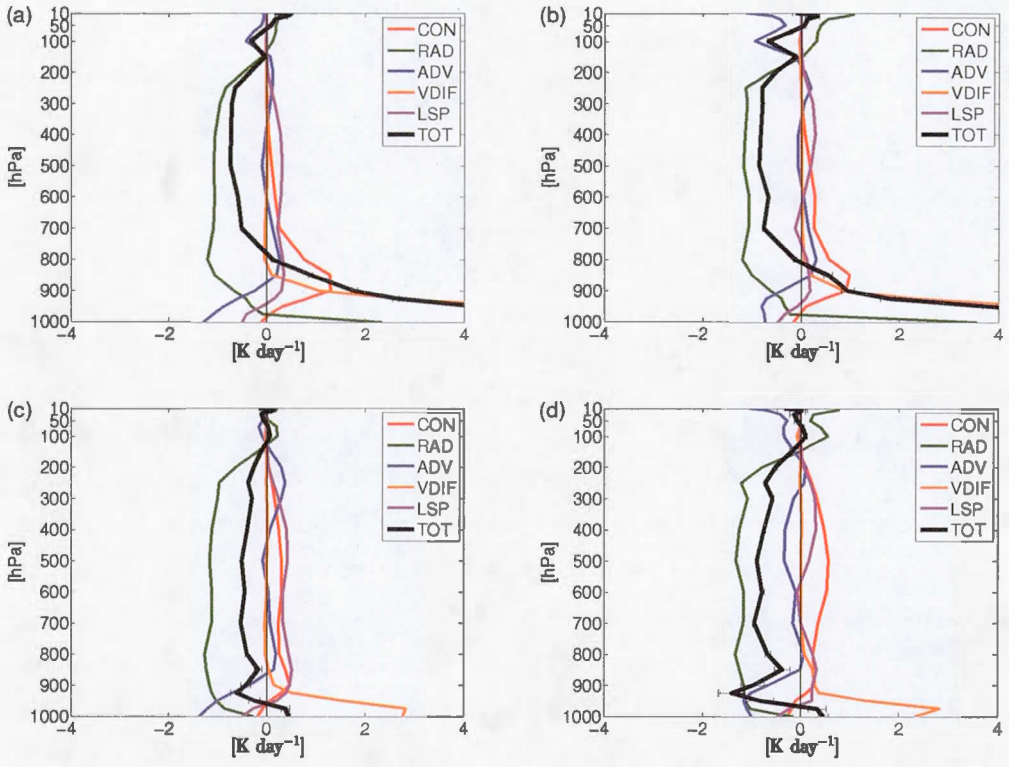

Figure 1.5 Temperature initial tendency profiles (K day^{-1}), computed for the first 6 h of the integration and obtained from simulations initialized by ERA-Interim reanalyses –(a, b) ERA-low and (c, d) ERA-high –and averaged (a, c) globally and (b, d) over the Tropics. Colour coding is as in Figure 1.3.

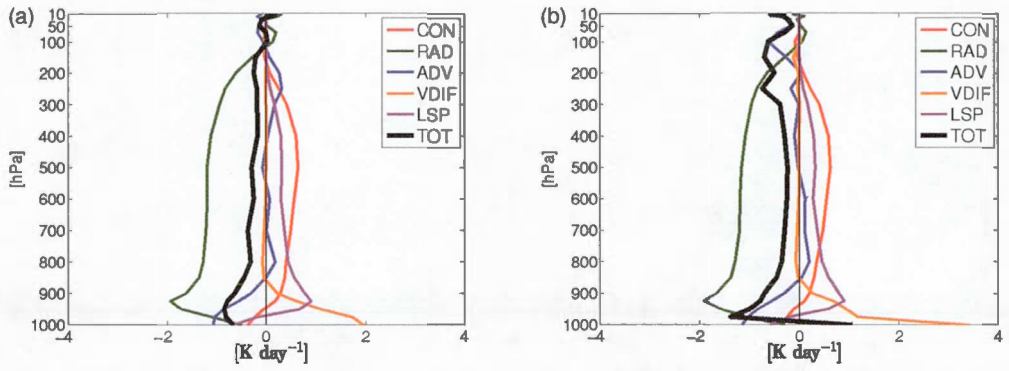

Figure 1.6 Temperature initial tendency profiles (K day^{-1}) computed for the first 6 h of integration with the initial conditions degraded (a) in horizontal resolution and (b) in vertical resolution. More details are given in the text.

As the boundary conditions of limited-area models (LAMs) are often prescribed every 6 h or less by forecasts from a global model which often has a coarser resolution, one must ask whether this can result in maintaining dynamical imbalances in the forecast of the LAM. This question will be investigated later using the CRCM limited-area regional climate model. Finally, how long will this imbalance persist before the model reaches its own equilibrium? This is addressed in the next section.

1.6 Time to dynamical equilibrium

In the previous experiments, the dynamical balance has been assessed for the first moments of the integration (typically the first 6 h). When imbalances are present, we are now interested to know how long it will take for the dynamical balance to be restored to the model's own climatology. Integrations were performed over longer periods, namely 15 days and, as before, the temperature tendency diagnostic is computed with a 6 h running mean over the duration of the simulation. Several sets of 15 days simulations starting from 1 January 2009 at 0000 UTC to 16 January at 1800 UTC have been performed (a total of 64 simulations in each set). As before, the GEM global model is used with different initial conditions, i.e. MSC4D, ERA-low and ERA-high.

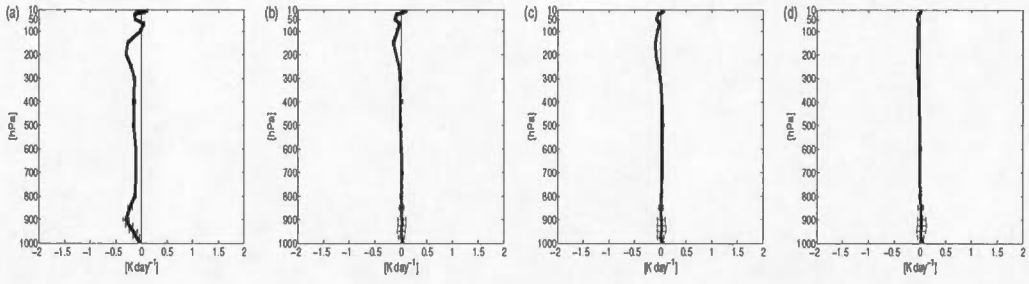

Figure 1.7 Temporal evolution of the mean total tendency profiles (K day^{-1}) after (a) 1, (b) 5, (c) 10 and (d) 15 days for the global GEM model. Simulations are initialized with MSC4D.

Temporal evolution of the total tendency is presented in Figures 1.7 and 1.8 when the GEM model is initialized by the three types of initial conditions. When MSC4D are used, the model is well balanced and seems to reach its own climatology after 15 days of integration. The model is better balanced in the middle of the atmosphere compared to the lower and upper levels where small imbalances persist. Simulations initialized with ERA-low exhibit a significant imbalance over the whole column where the model is warmer than its climatology. This result shows that the initial imbalances and the subsequent strong heating effects are still effective even after 15 days. Compared to the MSC4D simulations, the model initialized with ERA-low needs more time to reach its own equilibrium associated with the model's own climatology. However, the model is not completely balanced even after 15 days in both cases, especially at the upper levels where only radiation and advection processes are still active. Figure 1.8(a)–(d) show that the model's equilibrium is worse when initialized by ERA-low and a net warming appears in the first few days due to strong convective activity. This warming decreases in amplitude with time but remains even after 15 days integration.

However, ERA-high simulations yield a balance that is very similar to that obtained with MSC4D. MSC4D (Figure 1.7) and ERA-high simulations (Figure 1.8(e)–(h)) show rapid convergence towards the equilibrium state.

Finally, Figure 1.9 shows the vertically integrated total tendency as a function of time which is a good indicator of the model's imbalance over the atmosphere column; a well-balanced model should yield low values. The total tendency is

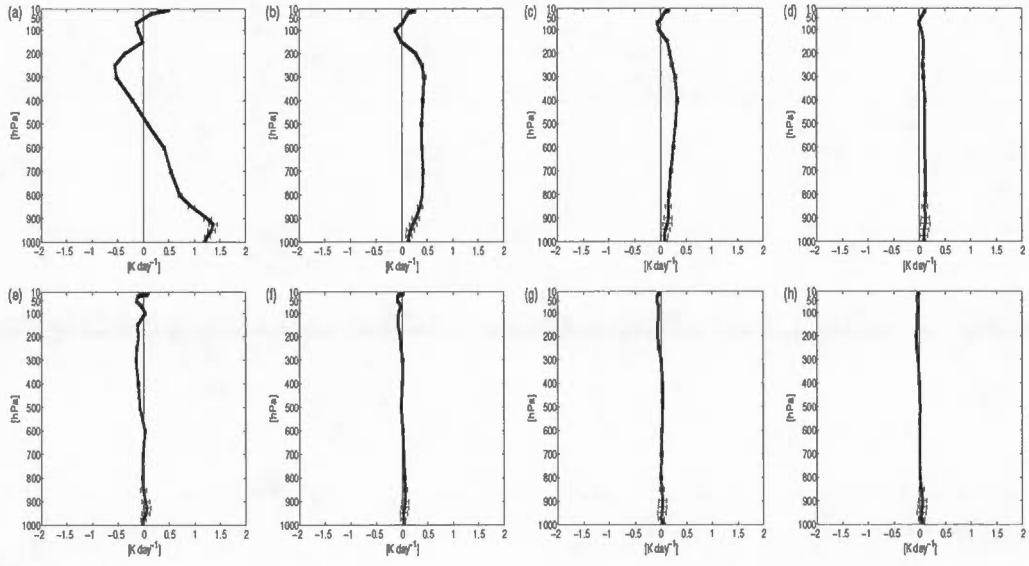

Figure 1.8 Temporal evolution of the mean total tendency profiles (K day^{-1}) after (a, e) 1, (b, f) 5, (c, g) 10 and (d, h) 15 days for the global GEM model, with simulations initialized with (a)–(d) ERA-low and (e)–(h) ERA-high.

mass-weighted vertically integrated over all model output levels from 1000 to 10 hPa. The absolute values were considered to prevent sign compensation between levels. The results show large differences from the three sets of simulations in the first days of integration during which a better balanced model is observed when initialized with MSC4D than with ERA-Interim simulations. The MSC4D simulations results indicate that the model's equilibrium state becomes nearly constant after roughly 2 days. With ERA-low, the large initial imbalance is clearly seen in the first days of integration but decreases with time especially in the first 2 days. However, the initial imbalance effect is too strong and requires several days of integration to be damped. The two curves converge reasonably well after 15 days but not completely. However, at full resolution the ERA-high simulations yield a better balance which compares well to that obtained with the MSC4D except in the first 18 h. However, afterwards, the two experiments have converged to the same level of balance. The two curves in Figure 1.9 converge after only 18 h of integration. This result confirms the advantage of using initial conditions with resolution close to the model's own. It is worth noticing that, from that point on,

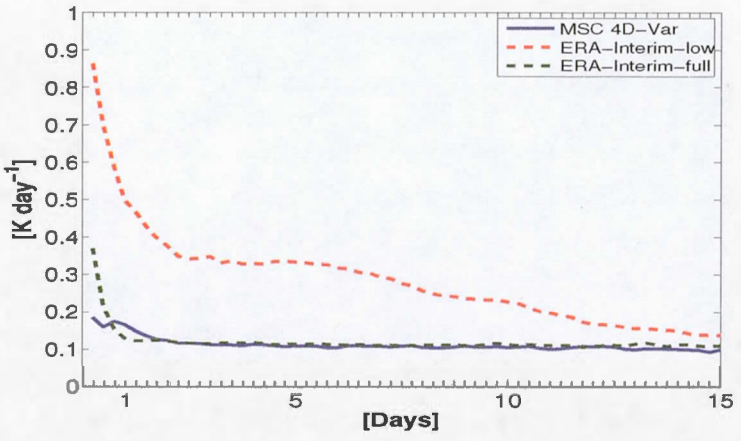

Figure 1.9 Vertically integrated absolute total temperature tendencies from GEM simulations initialized by MSC4D (blue line) and by ERA-low (dotted red line) and by ERA-high (dotted green line).

the balance has reached a constant level which does not vary in time. This is to be kept in mind for comparison with results to be presented in the next section with a limited-area model.

1.7 Assessing the balance of a regional climate model

In sections 1.4–1.6, the GEM dynamical balance was examined for its global configuration where only initial conditions are of interest. However, in the LAM configuration, the dynamical balance is not only influenced by the initial conditions but also through the lateral boundary conditions imposed at frequent intervals, typically 6 h. This is a significant difference for a regional climate model whose equilibrium defines the climatology to which it will tend. It therefore depends on the initializing data and the forcing information provided by either reanalyses or forecasts from another global climate model. In this section, the CRCM, a LAM configuration of the global model GEM, is used to study the resulting balance when a LAM is initialized and driven by different analyses. Based on the MSC4D, ERA-low and ERA-high as initial conditions, three sets of integrations were carried out using the CRCM and starting every 6 h from 1 to 30 January 2009 at 0000 UTC. The lateral boundary conditions are also extracted from the same ana-

lyses used in the experiment and supplied every 6 h. We note that these boundary conditions are linearly interpolated to each time step through the 6 h window.

The CRCM balance is measured by the RP07 diagnostic averaged over the complete domain of CRCM. These are presented in Figure 1.10 for the MSC4D and ERA-low/high experiments. In ERA-low, the strong heating associated with vertical diffusion observed in the global version is also present and is even larger, reaching 10 K day^{-1} (Figure 1.10(b)). A net cooling is dominant above 900 hPa throughout a large part of the troposphere due essentially to the absence of deep convection, even though this process is weak over North America in January. These results concur with those of the previous experiment, i.e. the model is better balanced when initialized and driven by MSC4D analyses. When ERA-high is used, the dynamical balance is much better, which confirms that a consistency in resolution is beneficial as was observed in the GEM (Figure 1.10(c)). Comparing Figure 1.10(a, c) indicates that MSC4D is more compatible with the CRCM dynamics than ERA-high which is not surprising given that the CRCM uses the same physical parametrizations as those of the GEM model used to produce the MSC4D analyses. As described previously, longer runs were also done with the CRCM model. In this case, the situation is different because of the combination of initial and boundary data effects on model equilibrium. Here we need to stress that only the boundary conditions are used, except at the beginning of the integration where initial conditions are provided. The motivation is that this is how a regional climate model (RCM) works, and it is bounded by input from an external source, be that analyses or global climate simulations. Since the initial condition effect on the model will vanish through the integration, this kind of experiment will help us detect the eventual impact of boundary conditions on the model balance. Moreover, it is known that in the boundary region or the blending zone, a significant damping is used to prevent contamination of the interior (Davies, 1976). As a consequence, most RCM results are examined only in the free zone, which is the interior domain excluding the blending zone.

In this study, the RP07 diagnostics are computed by firstly considering the whole domain (including the blending zone) and, secondly, considering only the interior 'free' region. The aim is to investigate the potential imbalance induced

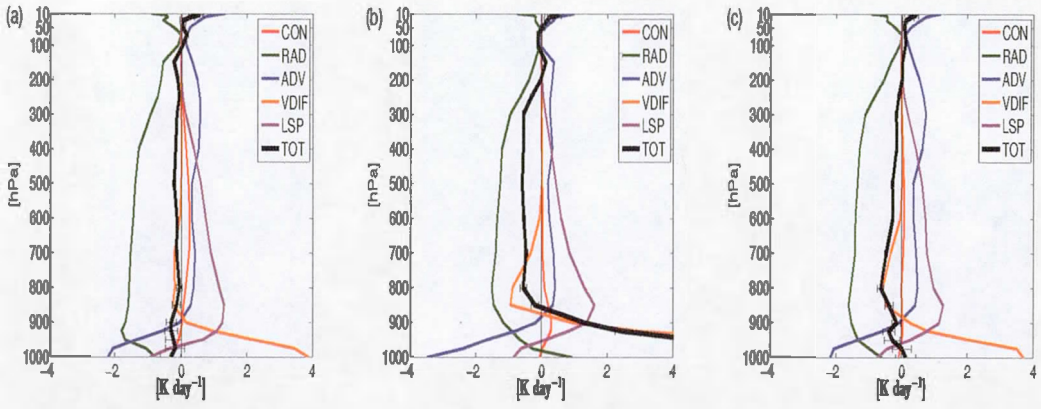

Figure 1.10 Temperature initial tendency profiles (K day^{-1}) obtained from CRCM simulations initialized and driven by (a) MSC4D, (b) ERA-low and (c) ERA-high analyses. Tendencies are computed for the first 6 h of the integration excluding the first time step and averaged over the complete domain including the nesting zone.

through the blending zone and to what extent this imbalance is introduced into the free interior zone. When tendencies are computed after 15 days of integration over the whole model domain, results show imbalances in the three types of simulations, with larger values observed in the case where the model is initialized and driven with ERA-low reanalyses (Figure 1.11). The large imbalances observed previously at the lower and upper model levels are also present here. When we compare these results with those obtained for the first 6 h (Figure 1.10), the model is less balanced, especially when it is initialized and driven by MSC4D analyses. This indicates the persistent effect of boundary conditions on the model equilibrium, suggesting that care is needed when looking at results over the whole domain.

In the interior domain, where the blending area is excluded, we obtain very different results (Figure 1.12). In its free zone, the model is shown to be less influenced by boundary conditions and it is much better balanced. These results show that the model, when considered in its free zone, reaches its equilibrium more rapidly, even in the case of ERA-low runs (Figure 1.12(b)). However, even in its free zone, the model exhibits some imbalances especially at the upper levels where the advection and radiation processes do not seem to balance each other. This

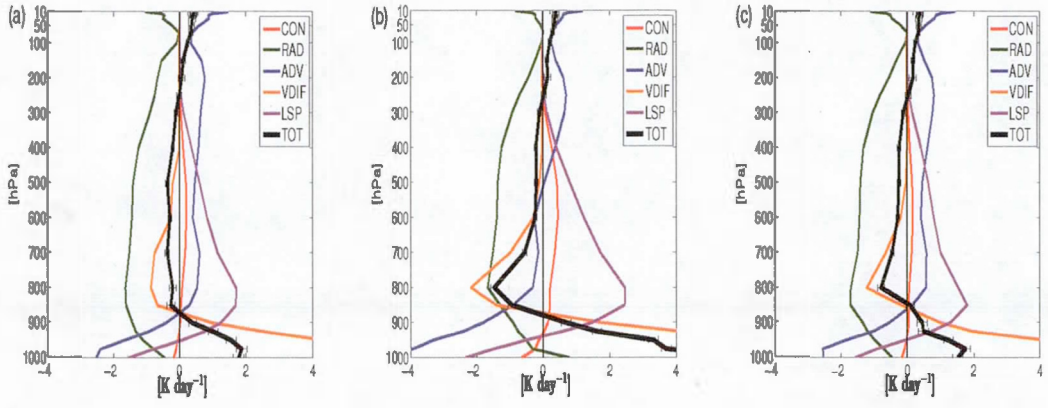

Figure 1.11 Temperature tendency profiles computed after 15 days integration over the whole CRCM domain including the blending zone, from runs using initial and boundary conditions from (a) MSC4D, (b) ERA-low and (c) ERA-high analyses.

imbalance is not seen in the initial temperature tendencies (Figure 1.10) nor in the global model case (Figure 1.7), reflecting consequently the important effect of boundary conditions and interactions between the forcing large-scale data and the LAM. This result also indicates that, even in the free zone, the model needs more time to establish its dynamical equilibrium towards its own climatology. A closer look at the total tendency temporal behaviour is presented in Figures 1.13 and 1.14 where we can clearly see the difference in model balance when the whole domain is considered, and when only the free zone is taken into account. Computations over the free zone show an improvement in the model balance during the first 5 days. Afterwards, we notice a degradation of this equilibrium over the whole atmospheric column, where we observe a net cooling in the mid-levels and a net warming above 250 hPa.

As before, we computed the vertically integrated total tendency for the CRCM model over the whole domain and over the free zone only, and results are presented in Figure 1.15. In the case where only the free zone is considered (Figure 1.15(b)), the simulations initialized and driven with ERA-low reanalyses show a rapid decrease during the first day of integration. The ERA-high curve is

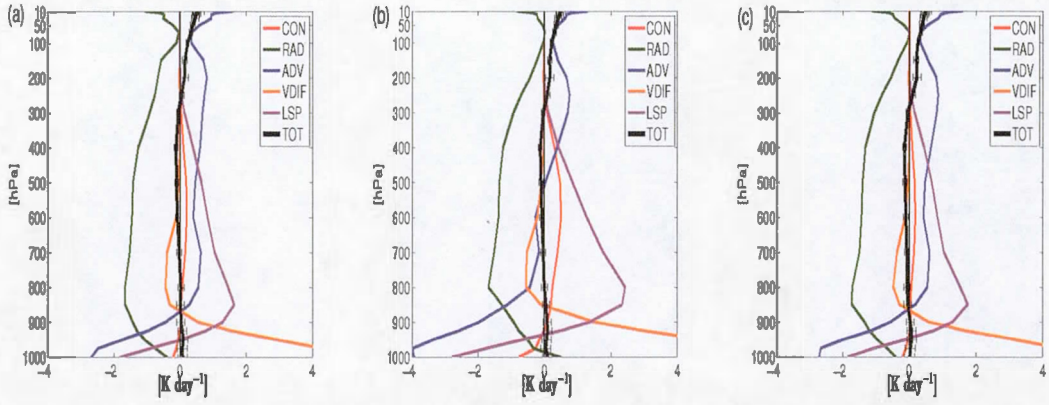

Figure 1.12 As Figure 1.11, but the tendency is averaged only over the CRCM free domain, excluding the blending zone.

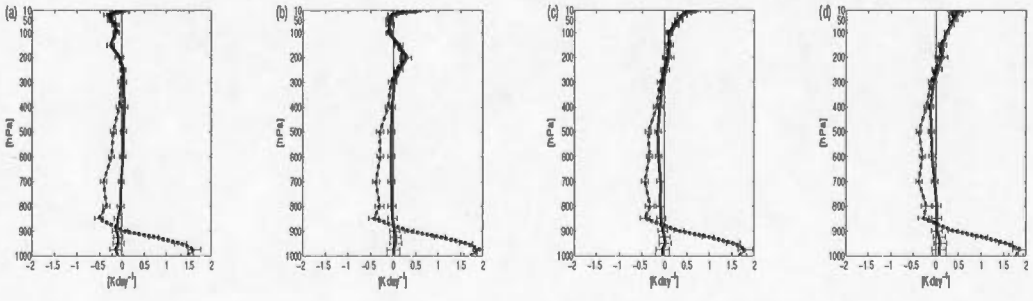

Figure 1.13 Temporal evolution of the mean total tendency profiles for the CRCM model after (a) 1, (b) 5, (c) 10 and (d) 15 days. Simulations are initialized and driven with MSC4D. The dotted line represents tendencies over the whole model domain and solid line those computed over the free zone.

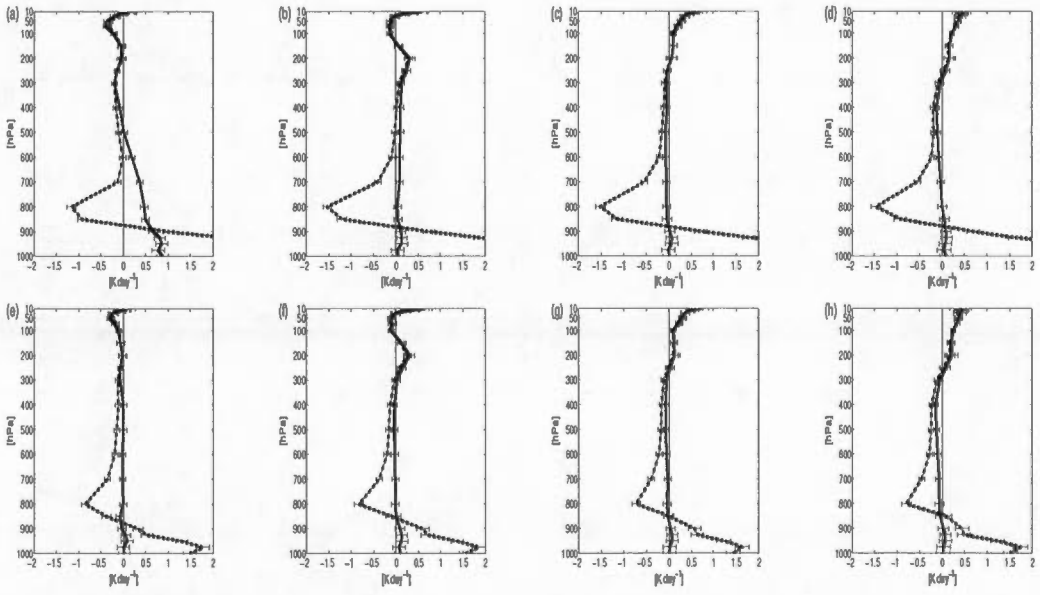

Figure 1.14 As Figure 1.13, but for simulations initialized and driven by ERA-Interim reanalyses (a)–(d) ERA-low and (e)–(h) ERA-high.

very close to that of MSC4D, showing the advantage of the higher resolution. The three curves exhibit a stable equilibrium between the first and the fourth days of integration. Afterwards, we note an increase in all curves (balance degradation), confirming what was seen in the net tendency profiles (Figures 1.13 and 1.14). Even though the MSC4D shows a good balance initially, Figure 1.15(b) indicates that imbalances start to reappear after 5 days of integration, indicating that the boundary forcing may create such imbalances after all.

1.8 Concluding remarks

In this study, an assessment has been made of the dynamical equilibrium of a model and its sensitivity to initial conditions and to boundary forcing. First, the global GEM model dynamical balance was looked at when initialized from its own 3D-Var (MSC3D experiment) and 4D-Var (MSC4D) analyses. The results showed significant differences, particularly with convection activity which was observed to be stronger for 4D-Var analyses in the ITCZ causing some imbalance in the model. This could be attributed to the absence of convection in the simplified

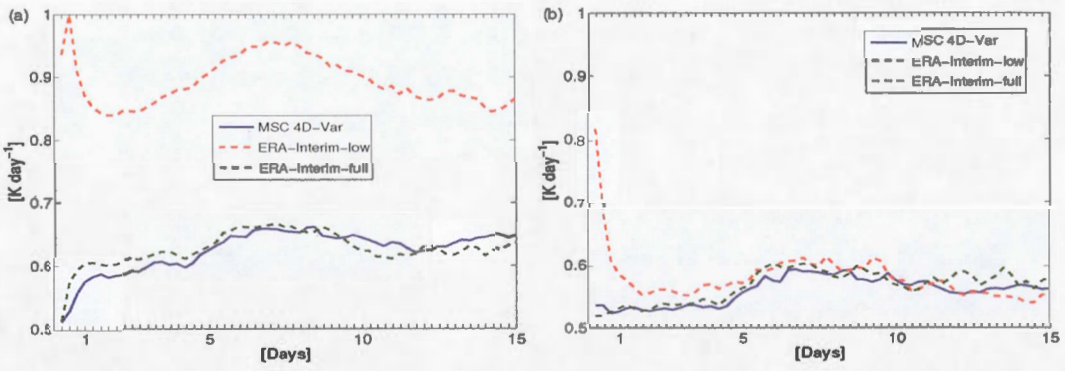

Figure 1.15 Vertically integrated absolute total tendencies of temperature for CRCM model computed (a) over the whole model domain including the blending zone, and (b) over the model free area only.

physics of the MSC 4D-Var (Gauthier et al., 2007).

When external analyses produced by an assimilation system driven by a different model are used, serious imbalances occur. When initialized with ERA-Interim analyses at low resolution (ERA-low), the global GEM model exhibits large imbalances especially in the lower levels due mostly to a vertical diffusion which is too strong and deep convection being virtually absent initially. Comparison of the two experiments (MSC4D and ERA-low) revealed that because ERA-low reanalyses are drier in the Tropics, there may not be enough humidity to trigger convection in the GEM global model. This can also contribute to imbalances in the lower troposphere. However, when used at their full resolution, ERA-Interim reanalyses (ERA-high) significantly reduce the imbalances compared to those seen at coarser resolution. Finally, comparison of the results of the MSC4D and ERA-high experiments indicate that, in the first 24 h, the model benefits from having its own analyses which yield a better initial equilibrium. However, this difference totally disappears after 48 h. This was obviously not the case for the ERA-low experiment for which convergence to equilibrium was not reached even after 15 days. Further experimentation in which the resolution of MSC4D analyses were degraded in the horizontal or in the vertical showed that reducing the vertical resolution has more impact on the model balance than reducing it in the horizontal.

The limited-area configuration of regional climate models requires that they are driven by imposed boundary conditions provided either by external global reanalyses or results from a global climate model. The sensitivity to boundary conditions has been examined by Wu et al. (2005), Laprise et al. (2013), Šeparovic et al. (2013) and others. The main objective of our study was to show if diagnosing the physical tendencies as proposed by RP07 could shed some light on this sensitivity. For this, the dynamical balance and the sensitivity to boundary conditions was studied for the CRCM LAM which, for the physical processes, is very close to that of the global GEM model used to produce the MSC4D analyses. First, the MSC4D experiment was repeated using these analyses as initial conditions and compared to the results obtained with the global model. The results show that the CRCM is in good equilibrium but a little less than that observed for the global model. Similar results were found when ERA-Interim reanalyses provided the initial and boundary conditions and were smaller when ERA-high reanalyses were used. This indicates, for a given model, the advantage of using initial and forcing data produced by a data assimilation system based on a model as similar as possible to that model.

Initial conditions are not a concern for climate simulations but boundary conditions are. The final set of experiments was to compare the sensitivity to boundary conditions only, which is the normal mode of integration of regional climate models. In this case, it was observed that imbalance observed in the blending zone where the boundary conditions are imposed is similar to what is observed in the experiments with the initial conditions. However, in the so-called free zone, the imbalances are much lower, as was seen in Figures 1.13 and 1.14, the latter showing that when ERA-low reanalyses are used for the boundary conditions, a significant imbalance is observed initially but seems to quickly disappear. However, a closer inspection shows that a slight imbalance persists even after 15 days. So, for the MSC4D, the balance is initially good but imbalances were noticed to gradually reappear after 5 days at which time this imbalance compares to that of the ERA-low and ERA-high analyses. The model still exhibits abnormal heating over the whole atmospheric column for the three cases for which it is found that longer runs are required to reach a good equilibrium. This will be further investigated in future work.

This work shows that the diagnostic of physical tendencies can provide very useful and detailed information about the model's dynamical balance. However, it remains to be seen if a link can be made between these diagnostics and estimation of internal variability as evaluated in regional climate studies. In future work, it is intended to use the same approach to assess the balance of the model with other analyses such as NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al., 2011) or NCEP's North American Regional Reanalysis (NARR) (Mesinger et al., 2006). Finally, the intention is also to produce regional analyses with the CRCM to be able to directly compare the model with respect to observations. The tendency diagnostics will be useful to assess the dynamical balance of the model when using its own analyses.

Acknowledgements

This research was funded by Québec's Ministère du Développement Economique, Innovation et Exportation (MDEIE), the Natural Sciences and Engineering Research Council of Canada (NSERC) and Environment Canada through its Grants and Contribution program. High performance computing resources were made available through a resource allocation granted by Calcul Canada on the Colosse and Guillimin platforms of the Calcul Québec regional consortium. The first author benefited also from a scholarship from the UQAM Science Faculty (FARE).

The authors would like to particularly thank Mrs. Katja Winger and Dr. Bernard Dugas for their help, numerous discussions and advice. Comments from Drs. Luc Fillion, Stéhane Laroche and Paul Vaillancourt from Environment Canada are gratefully acknowledged. The 3D-Var and 4D-Var analyses of Environment Canada were provided to us by Dr. Stéphane Laroche of the Data Assimilation and Satellite Meteorology research section of Environment Canada. This study benefited from the open access to ERA-Interim reanalyses from ECMWF.

CHAPTER II

ON THE EFFECT OF BOUNDARY CONDITIONS ON THE CANADIAN REGIONAL CLIMATE MODEL : USE OF PROCESS TENDENCIES

This chapter is presented in the format of a scientific article. It is published in the Climate Dynamics journal. This manuscript is entirely based on my work, with the co-author involved in results analysis and interpretation and also text revision.

The detailed reference is:
Chikhar K, Gauthier P (2015) On the effect of boundary conditions on the Canadian Regional Climate Model: use of process tendencies. Clim Dyn 45: 2515–2526 doi: 10.1007/s00382-015-2488-2

On the effect of boundary conditions on the Canadian Regional Climate Model : use of process tendencies

Kamel Chikhar and Pierre Gauthier

Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale),
Département des Sciences de la Terre et de l'Atmosphère
Université du Québec à Montréal (UQAM)
B.P. 8888, Succ. Centre-ville
Montréal (Québec) Canada H3C 3P8

Abstract

Climate simulations results can be very different when the regional climate model used is driven by different data. In this paper, the fifth-generation Canadian Regional Climate Model (CRCM5) response is assessed when driven by various boundary conditions. The latter are provided by outputs from the second-generation Canadian Earth System Model (CanESM2) and the Max Planck Institute for Meteorology's Earth System Model (MPI-ESM-LR) and also from ERA-Interim reanalysis. Physical and dynamical tendencies are analysed when the regional model is well spun up and is sufficiently affected by the lateral forcing data. The results indicate that the model is very sensitive to those imposed lateral conditions. Compared to observations, the CRCM5 exhibits excessive heating in the lower levels and cooling above when driven by the three driving data. It is also found that the two global models contribute to these anomalies but with different effects. Temperature tendencies revealed a cooling in lower layers when CRCM5 is driven by CanESM2 while a heating is noted when the model is forced by MPI-ESM-LR. Specific humidity tendencies also showed different effects depending on the driving data used.

Keywords : Regional climate modelling, Lateral driving data, CRCM5, Process tendencies

2.1 Introduction

Climate simulations are performed using increasingly complex models. The parameterized physical processes in addition to their interactions introduce errors in these climate simulations. Other factors such as model structure, vertical and horizontal resolution can contribute to these errors. Climate simulations are more complicated when using limited area models. Indeed, regional climate models (RCMs) are driven at their lateral boundary using data from Global Climate Models (GCMs) outputs or reanalyses. The lateral nesting introduces additional errors due to the coarser resolution in the driving data, the use of different parameterization schemes and also imbalances induced by the nesting technique (Warner et al., 1997). In addition, it has been shown that different driving data sources can introduce different results due to different errors (Šeparovic et al., 2013). In climate simulations, errors are mostly evaluated by comparing model outputs to observations and reanalyses (e.g., Šeparovic et al. (2013); Hernandez-Diaz et al. (2012); Martynov et al. (2013)). Evaluation of the bias, the root mean square error or anomaly correlations are often used to quantify models' skills and errors. Numerous studies focused on RCMs errors especially those related to initial and boundary conditions effect (e.g. Wu et al. (2005); Denis et al. (2003); Diaconescu et al. (2007)). However, the error sources are still difficult to identify and there is a need for additional or alternative tools to provide more information about the origin of these differences. Several studies have shown that the physical and dynamical tendencies of the model on temperature or humidity can provide valuable details on the model error when it is integrated in different configurations. Rodwell and Palmer (2007) showed that, starting from an analysis that defines the initial conditions, the initial imbalances detected through the physical and dynamical tendencies can reveal inconsistencies in the physical parameterizations associated with processes acting on fast time scales such as convection, radiation, or gravity wave drag. In numerical weather prediction (NWP), this is a method often used to assess the impact of changes brought to the physical parameterizations. An example is presented in Rodwell and Jung (2008) to evaluate a change in the impact of the treatment of aerosols in the ECMWF NWP model. In these studies the model used is the same as the one employed in the data assimilation system that

produced the analyses. Given that an analysis is a blend between observations and a short-term forecast model used as the background, Chikhar and Gauthier (2014) studied the impact on a given model of using different analyses. Their results indicate that differences between this model and the one used to produce the analysis can induce imbalances that are often very significant. Furthermore, these imbalances persist for some time but global models eventually reach an equilibrium representative of their own climate. Regional climate models on the other hand are constantly driven through lateral boundary conditions defined by either global reanalyses or climate simulations from a global climate model. Chikhar and Gauthier (2014) looked at the imbalances of the latest version of the Canadian Regional Climate Model (CRCM5 hereafter, Zadra et al. (2008)) when initialized and driven by global analyses produced by the global version of the model (Gauthier et al. 2007) or when initialized and driven by ERA-Interim reanalyses at low and high resolution. It was shown that using high resolution ERA-Interim reanalyses as initial and boundary conditions yields to a fairly good balance that compares well with that obtained from the global analyses. The object of this paper is to examine the physical and dynamical tendencies for a regional climate simulation after a long enough period of time for it to reach its own climatological equilibrium. The process tendencies are evaluated for the CRCM5 to investigate errors in this RCM simulations when driven by different data. The aim is to provide details on the model errors especially those related to the driving data and implicitly, the nesting technique. Three types of data are employed to drive the CRCM5 model : outputs from two Coupled Global Climate Models (CGCMs), the second-generation Canadian Earth System Model (CanESM2) and the Max Planck Institute for Meteorology's Earth System Model (MPI-ESM-LR) and also full resolution ERA-Interim reanalyses as those used in Chikhar and Gauthier (2014). The latter will be considered in this study as a reference for the boundary conditions. The document is organized as follows. In Sect. 2 , the methodology is introduced altogether with a description of the CRCM5 model and the design of the experiments. The results obtained are presented and analysed in Sects. 3–5. Concluding remarks are finally presented in Sect. 6.

2.2 Model and experiments description

2.2.1 Methodology

In this study, dynamical and physical tendencies are used to analyse the CRCM5 behaviour when driven by different data. The aim is to examine in more detail the dynamical and physical processes depicted by the model during its integration and to detect how they may be affected by the driving data at the lower and lateral boundaries.

We used three data sources to drive the CRCM5 model : outputs from two CGCMs (CanESM2 and MPI-ESM-LR) and ERA-Interim reanalysis. The first CGCM is CanESM2, the second generation Canadian Earth System Model. Its atmospheric component has a T63 spectral horizontal resolution with a linear transform grid of 2.8°, 35 levels in the vertical with a top at 1 hPa (Arora et al., 2011). The second CGCM is the Max-Planck-Institute Earth System Model (MPI-ESM-LR) having an atmospheric component (ECHAM6) operating at T63 with a quadratic transform grid of around 1.89° and 47 vertical levels with a top at 0.01 hPa (Stevens et al., 2013; Giorgetta et al., 2013). ERA-Interim reanalyses used in this study are at full resolution, namely at 0.75° horizontal resolution and 60 levels in the vertical with a top at 0.1 hPa (Dee et al., 2011).

The objective is to identify the impact of the boundary data on the model performance, by examining the process tendencies obtained from different simulations in which the model is driven by different data. The full resolution ERA-Interim reanalyses can be considered as a reliable representation of the atmosphere through the fit to all available observations. In addition, as pointed out by Chikhar and Gauthier (2014), in its first few days integration, the CRCM5 shows fairly good dynamical balance when initialized and driven by full resolution ERA-Interim reanalyses. The latter will then provide our reference boundary conditions and the resulting simulation will be compared to the results obtained from simulations where driving data are specified from outputs from two CGCMs. As pointed out by Laprise et al. (2013) and Šeparovic et al. (2013), a comparison of the dynamical and physical process tendencies obtained from a reanalysis-driven

simulation to 'observed' tendencies can shed some light on the 'structural bias' (SB) of the regional climate model, i.e, deficiencies in the model itself. On the other hand, the comparison between 'observed' tendencies and those obtained from a CGCM-driven simulation will provide details on the 'total bias' (TB) including 'structural bias' and a bias induced by the boundary conditions (BC). We can express this bias relation as

$$TB = SB + BC \tag{2.1}$$

Consequently, the bias induced by boundary conditions can be identified by subtracting the structural bias from the total bias

$$BC = TB - SB \tag{2.2}$$

The total bias from CanESM2 and MPI-ESM-LR driven simulations will be referred to as CRCM-Can-TB and CRCM-MPI-TB respectively. Similarly, biases due to the driving data are respectively denoted as CRCM-Can-BC and CRCM-MPI-BC. The 'observed' total tendencies are extracted from ERA-Interim reanalyses as these data provide a three dimensional representation of the meteorological variables and thus can be used to compare to model outputs. The CRCM5 structural bias will be denoted as CRCM-SB and corresponds, as explained previously to the total bias when the model is driven by ERA-Interim reanalyses. In this study, the emphasis will be the examination of biases on dynamics and physics tendencies computed for two variables, namely temperature and specific humidity. These tendencies are computed by taking into account all individual processes that combine to update the variable at each time step, the total tendency being their sum, i.e.

$$\dot{T}_i^{total} = \sum_{p=1}^{k} \dot{T}_i^{p} \tag{2.3}$$

where $\dot{T}_i^{total}$ is the total tendency, and $\dot{T}_i^{p}$, $p = 1, ..., k$, represent the individual tendency of the p-th process. In the case of temperature, processes considered are dynamics (i.e., horizontal and vertical advection along with horizontal diffusion), shallow and deep convection, solar and long wave radiation, large scale condensation and vertical diffusion. For specific humidity, the same processes are considered except radiation.

The regional climate model used (CRCM5) and the different simulations performed are described in the next two subsections.

2.2.2 CRCM5 description

The regional climate model used here is the fifth-generation Canadian Regional Climate Model (CRCM5) (Zadra et al., 2008). It is based on a limited-area configuration of the Global Environment Multiscale model (GEM) developed at the Meteorological Service of Canada (Côté et al., 1998). The CRCM5 operates on a rotated latitude longitude Arakawa C grid (Arakawa and Lamb, 1977) and includes a terrain-following vertical coordinate (Laprise, 1992). To supply CRCM5 with boundary data, a one way nesting technique is used (Davies, 1976; Robert and Yakimiw, 1986). It consists of a 10 grid points external halo zone for the semi-Lagrangian interpolation followed by a 10 grid points sponge zone where the prognostic variables are gradually relaxed to the driving data. The CRCM5 configuration considered here uses the following parameterization schemes : Kain-Fritsch for deep convection (Kain and Fritsch, 1990, 1993), Kuo Transient for shallow convection (Kuo, 1974), correlated-K for solar and terrestrial radiation (Li and Barker, 2005), Sundqvist for large scale condensation (Sundqvist, 1978), Mailhot and Benoit for vertical diffusion (Mailhot and Benoit, 1982) and CLASS for surface processes (Verseghy, 2000, 2008). A more detailed CRCM5 description

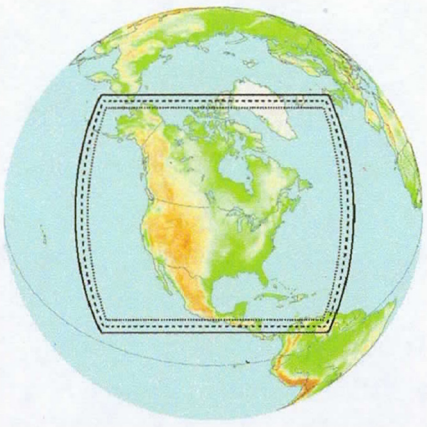

Figure 2.1 CRCM5 domain. The inner *light dotted line* delimits the model free zone

can be found in Martynov et al. (2013) and Hernandez-Diaz et al. (2012).

2.2.3 Configuration of experiments

The CRCM5 domain covers North America (Fig. 2.1) which is similar to that used in the COordinated Regional Climate Downscaling EXperiment (CORDEX) simulations (Šeparovic et al., 2013). The horizontal resolution is around 20 km ($0.22° \times 0.22°$) with a time step of 10 min. In the vertical, the model uses 56 levels with a top at 10 hPa.

As mentioned before, three CRCM5 simulations were performed, each using different data as boundary conditions. To be able to recompute the process tendencies for all the physical and dynamical processes, each of these simulations were restarted over a period of 14 months and beginning on 1st January 2008 at 00GMT. The restart conditions come from a sufficiently long period climate simulation using CRCM5 with the same configuration when driven by ERA-Interim reanalyses. In this way, we can ensure that CRCM5 is already well balanced with its own climatology while driven by the reference boundary data. In our simulations, the boundary conditions are updated every 6 h and interpolated in time to every time step within the 6-h interval. The physical and dynamical tenden-

cies are examined for the boreal winter DJF 2008–2009 season. This procedure is meant to ensure that the model is completely spun up and the boundary conditions effect is considered significant. Throughout the DJF season, the temperature and specific humidity mean process tendencies are computed every consecutive 6 h from 1st December at 06GMT to 1st March at 00GMT. The temporal mean process tendencies are then obtained for each process along with its evolution throughout the DJF season. In order to be examined as profiles over the whole atmospheric column, the tendencies have been vertically interpolated from model levels to pressure levels and then spatially averaged over the model free domain, i.e. the blending zone is excluded in computations. Spatial averages are obtained by weighting every grid point by its horizontal area of influence and masking all sub-surface grid points. The simulations will be denoted CRCM-Can, CRCM-MPI and CRCM-ERA when CRCM5 is driven by data from Can-ESM2, MPI-ESM-LR and ERA-Interim reanalysis respectively.

2.3 Impact of the boundary conditions on the temperature tendency

Our analysis will focus first on total tendencies to get an overview of the dynamical balance of the model. Then, examining individual process tendencies will provide more details about the impact of the different boundary conditions when driven by CanESM2, MPI-ESM-LR and ERA-Interim reanalysis.

2.3.1 Total tendency

Over a season, one would expect a cooling trend as the model goes from the beginning of December to March and averaging the temperature tendencies over that period should reflect this for the three RCM simulations forced from different boundary conditions. Mean total tendencies were computed as explained in the previous section and the profiles from each simulation are presented in Fig. 2.2. Figure 2.2a–d shows respectively the mean of 6-h total tendencies for these three simulations together with the observed temperature tendency computed from the ERA-interim reanalyses averaged over the same domain and period. The tendency

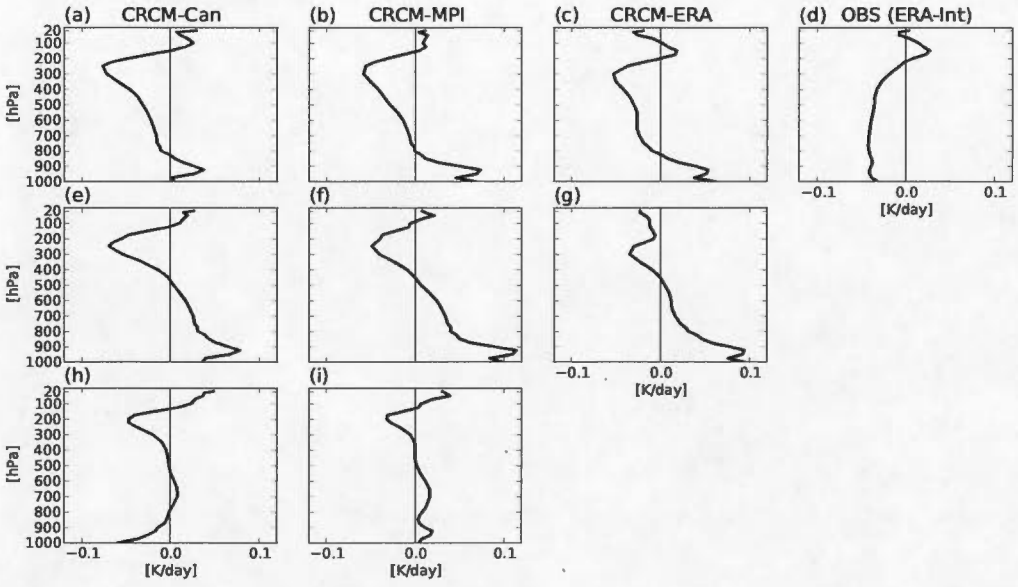

Figure 2.2 Total temperature tendency profiles in K/day computed for each simulation. Total DJF tendencies for CRCM-Can, CRCM-MPI and CRCM-ERA are given in (a), (b) and (c) respectively. The observed (from reanalyses) tendency is shown in (d). Total tendency biases CRCM-Can-TB and CRCM-MPI-TB are presented in (e) and (f) respectively whereas g indicates the structural bias CRCM-SB. Finally, the lateral effect biases CRCM-Can-BC and CRCM-MPI-BC are shown in h and i respectively. Vertical coordinate is linear in pressure

profiles indicate a model heating in levels below 800 hPa for CRCM-Can, CRCM-MPI and CRCM-ERA while a cooling for the three simulations occurs over a large part of the column with a different behaviour above $\sim$ 200 hPa in each profile. On the other hand, the observed tendency (Fig. 2.2d) indicates a cooling over the major part of the atmosphere from surface to $\sim$ 200 hPa whereas a warming prevails above that level. The comparison between profiles highlights some differences that are more consistent in lower and upper layers.

To highlight the differences between the three simulations, the observed tendency has been subtracted from the total tendencies for each of the three simulations to obtain the total bias in the tendency. Thus, as explained previously,

CRCM-Can-TB and CRCM-MPI-TB are computed by subtracting the observed tendency from the total tendency from CRCM-Can and CRCM-MPI respectively. This is presented in Fig. 2.2e, f which show a positive total bias over the lower part of the column (below $\sim$ 500 hPa) leading to an excessive model heating in these layers. In contrast, an abnormal cooling is observed between 500 and $\sim$ 100 hPa, especially for CRCM-Can-TB and CRCM-MPI-TB with a peak at around 250 hPa.

Considering the ERA-interim reanalyses as the best representation of the state of the atmosphere, the mean tendency total bias observed in the CRCM-ERA experiment (Fig. 2.2g) can be attributed to model deficiencies and therefore correspond to a structural bias (CRCM-SB). It shows a slightly different result with lower bias expressed in less warming in lower layers and also less cooling in the upper ones. The lateral boundary effect is measured as the difference between the total bias obtained by CRCM-Can or CRCM-MPI and that obtained from CRCM-ERA, i.e. the structural bias CRCM-SB. The resulting biases CRCM-Can-BC and CRCM-MPI-BC are shown in Fig. 2.2 h, i respectively. They clearly show that the driving effect is different for the two CGCMs. In the case of CRCM-Can, a negative bias (abnormal cooling) in surface layers is present whereas a light positive bias (heating) is observed for the CRCM-MPI experiment. Moreover, a heating is shown in the mid layers (850–500 hPa) for both cases but is more pronounced for CRCM-MPI. Above 500 hPa, the two profiles are similar with an excessive cooling peak at $\sim$ 200 hPa.

Finally, the driving data temperature total tendencies computed for the DJF season and over the CRCM domain are presented in Fig. 2.3. This figure reveals differences in the driving CGCMs data tendencies when compared to those from reanalyses (thin dotted line). In mid and lower atmosphere (below $\sim$ 400 hPa), it is shown that the cooling from both CGCMs is smaller than that observed in the reanalysis. From this level up to $\sim$ 100 hPa, we notice a different behaviour with more cooling for the two CGCMs. These differences could be essentially attributed to different physical parametrization schemes used in each model but also to differences in vertical and horizontal resolution.

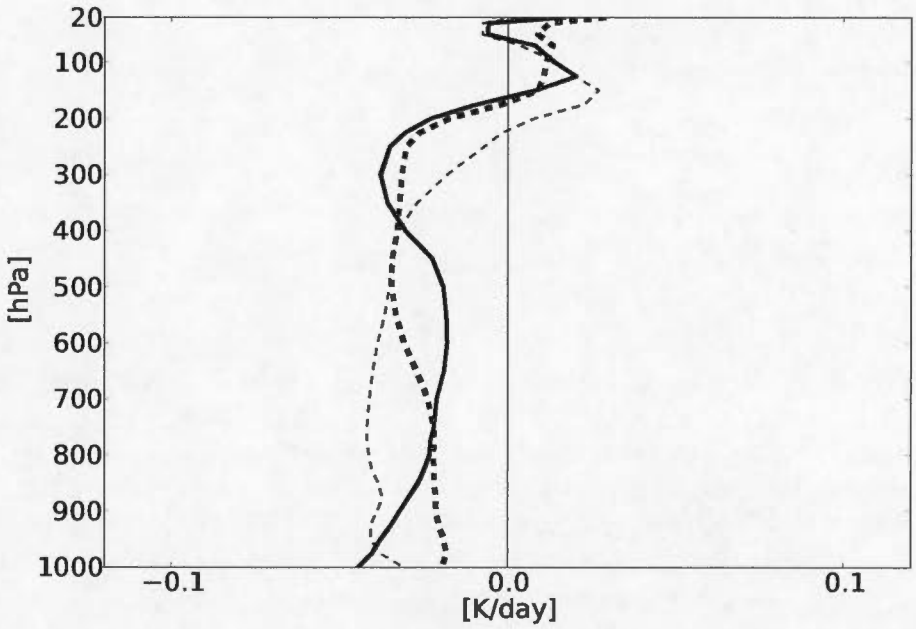

Figure 2.3 DJF temperature total tendency in K/day computed over the CRCM domain for CanESM2 (solid line) and MPI-ESM (thick dotted line). ERA-Interim temperature total tendency is plotted in thin dotted line

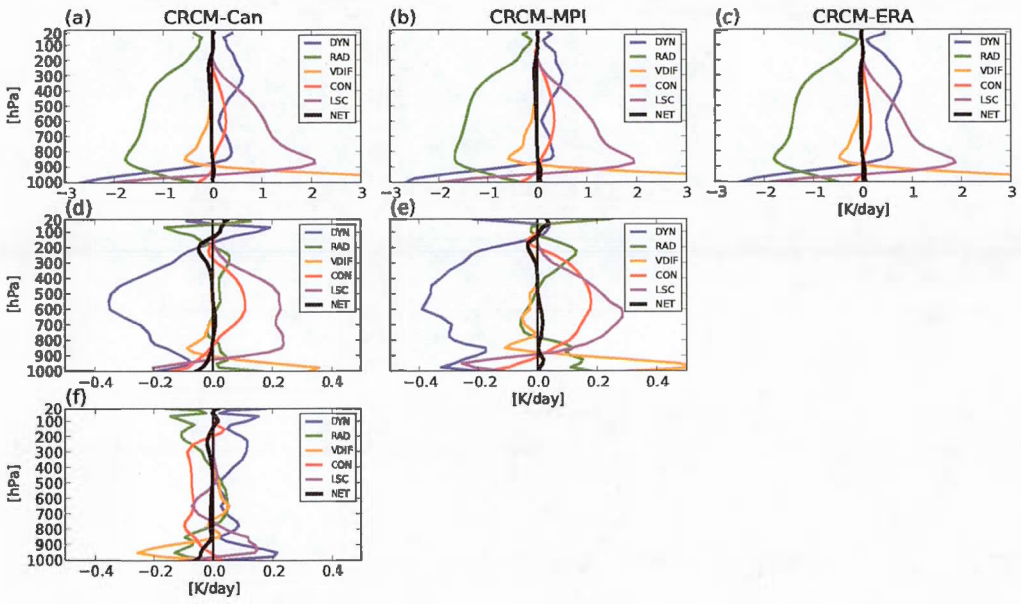

Figure 2.4 Individual process temperature tendencies in K/day computed for each simulation, CRCM-Can (a), CRCM-MPI (b) and CRCM-ERA (c). Lateral boundary effect biases CRCM-Can-BC and CRCM-MPI-BC are shown in (d) and (e) respectively. Difference between CRCM-Can-BC and CRCM-MPI-BC is depicted in (f). *Colored lines* represent radiation (*green*), vertical diffusion(*brown*), convection (*red*), large scale condensation (*magenta*) and dynamics (*blue*). The *thick black line* is the net tendency, i.e. the sum of all process tendencies. Note the change in x-axis scale in (d–f)

2.3.2 Contributions from the individual physical processes

The tendencies obtained from the various simulations are further examined by looking at individual tendencies from each physical and dynamical process. Profiles obtained from the three experiments are presented in Fig. 2.4 where the different colors represent dynamics (blue), the total of shallow and deep convection (red), the sum of solar and long wave radiation (green), large scale condensation (magenta) and vertical diffusion (brown). The thick black line is the net tendency obtained by summing all process tendencies.

The process tendencies (Fig. 2.4a–c) reveal that convection activity is stronger in CRCM-Can and CRCM-MPI than that in the CRCM-ERA simulation. This suggests that more favourable conditions are able to trigger convection more easily in CRCM-Can and CRCM-MPI. Figure 2.4a–c also shows a more active large scale condensation and less dynamical heating. Near the surface, CRCM-MPI shows that heating due to vertical diffusion outweighs the cooling due to dynamics, radiation and also evaporation from large scale condensation and convection schemes. The resulting effect is a weak warming near the surface. In the CRCM-Can simulation, a similar weak warming occurs but with less magnitude because of the lower heating associated with vertical diffusion. The cooling peak at $\sim$ 200 hPa is mainly due to less heating from dynamics. CRCM-Can-BC and CRCM-MPI-BC profiles are shown in Fig. 2.4d, e where a stronger convection activity is seen particularly in the case of CRCM-MPI. In the mid-troposphere, the positive biases from convection and large scale condensation exceed the negative one associated with dynamics leading to a net positive bias. In the lowest levels, in the case of CRCM-MPI-BC (Fig. 2.4e), positive bias from vertical diffusion and radiation exceeds the negative one observed from dynamics, convection and large scale condensation resulting in a net positive bias. The opposite is noted in the CRCM-Can-BC case leading to a positive bias.

The differences between CRCM-Can-BC and CRCM-MPI-BC (Fig. 2.4f) indicate a clearly stronger convection activity in the CRCM-MPI experiment because it offers better conditions for its triggering. Moreover, dynamics is responsible for a larger net cooling in the upper levels as shown in Fig. 2.4d, e where a clear negative bias is observed. This tends to destabilize the atmosphere leading to more favourable conditions for convection. Figure 2.5 shows the time series of the contribution of convection to temperature tendency. It shows that there is stronger convection activity for CRCM-Can and CRCM-MPI (Fig. 2.5a, b) compared to that observed in CRCM-ERA (Fig. 2.5c) confirming what we observed in the process profiles (Fig. 2.4). This convection activity is even larger in CRCM-MPI compared to CRCM-Can (Fig. 2.5d) particularly during the first half of the season. The impact of the driving data on convection can be more effective through the lower boundary conditions, i.e the sea surface temperature (SST). Figure 2.6 shows the mean DJF SST computed for the two CGCMs over the CRCM domain

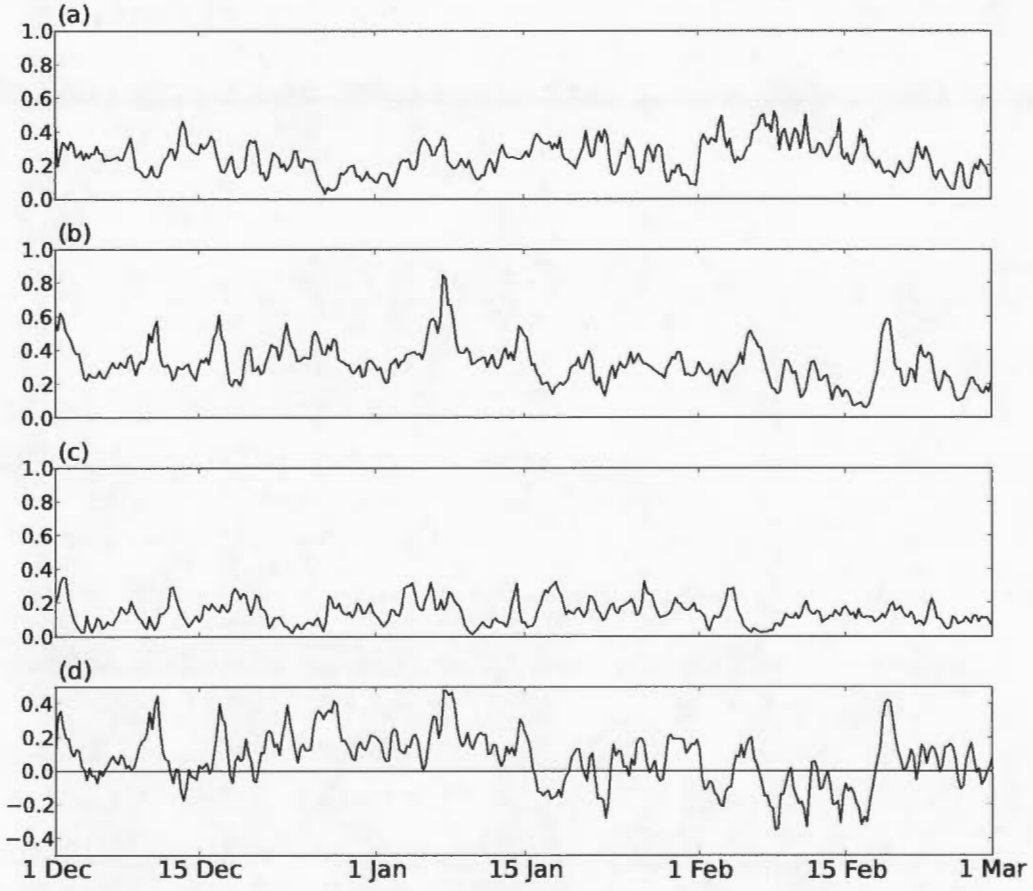

Figure 2.5 Temperature tendency due to convection in K/day computed at level 500 hPa for CRCM-Can (a), CRCM-MPI (b) and CRCM-ERA (c). Difference between CRCM-MPI and CRCM-Can is depicted in (d)

indicating warmer waters in the case of MPI-ESM model particularly in southern regions which is more favourable to initiate and strengthen convection. This is consistent with what was observed in Figs. 2.4 and 2.5.

The biases in the tendencies at 200 hPa (Fig. 2.7) indicate that there is cooling over the western and the northern parts of the domain in the case of CRCM-Can and CRCM-MPI. This can be attributed to a cold temperature advection through the northern boundary towards the inside of the model's domain. This strong cooling is not seen in the case of CRCM-ERA and this can be attributed to different air mass temperature for the three cases. Indeed, Fig. 2.8 shows the 200 hPa mean temperature and geopotential height for the three sets of driving data (CanESM2, MPI-ESM-LR and ERA-Interim). The shaded colors represent temperature whereas the contours are for geopotential height. The two CGCMs are clearly colder than ERA-interim especially MPI-ESM with abnormally low temperatures over the northern polar region. The stronger geopotential height gradient indicates stronger wind speeds for the two GCMs that enhance the cold air advection into the CRCM interior domain. This assumption is confirmed by the lateral bias due to dynamics shown in Fig. 2.9 where a cold bias is observed in the western part of the model domain. This cooling is more intense in the CRCM-MPI case and also extends to the northern region. Differences are also noted between these two simulations over the south-west part of the domain.

To summarize, the results indicate that convection is particularly sensitive to the boundary conditions. This is likely to have an impact on humidity as well and this is looked at in the next section.

2.4 Specific humidity

As was done for temperature, the specific humidity total tendencies are averaged over the DJF period for each simulation and compared with the 'observed' one obtained from ERA-Interim. Total tendencies, total biases and lateral boundary effect biases are shown in Fig. 2.10. The observed tendency (Fig. 2.10d) indicates a drying in the atmosphere (up to $\sim$ 400 hPa) throughout the DJF season as are the tendencies from the simulations but not at the surface (Fig. 2.10a–c). However,

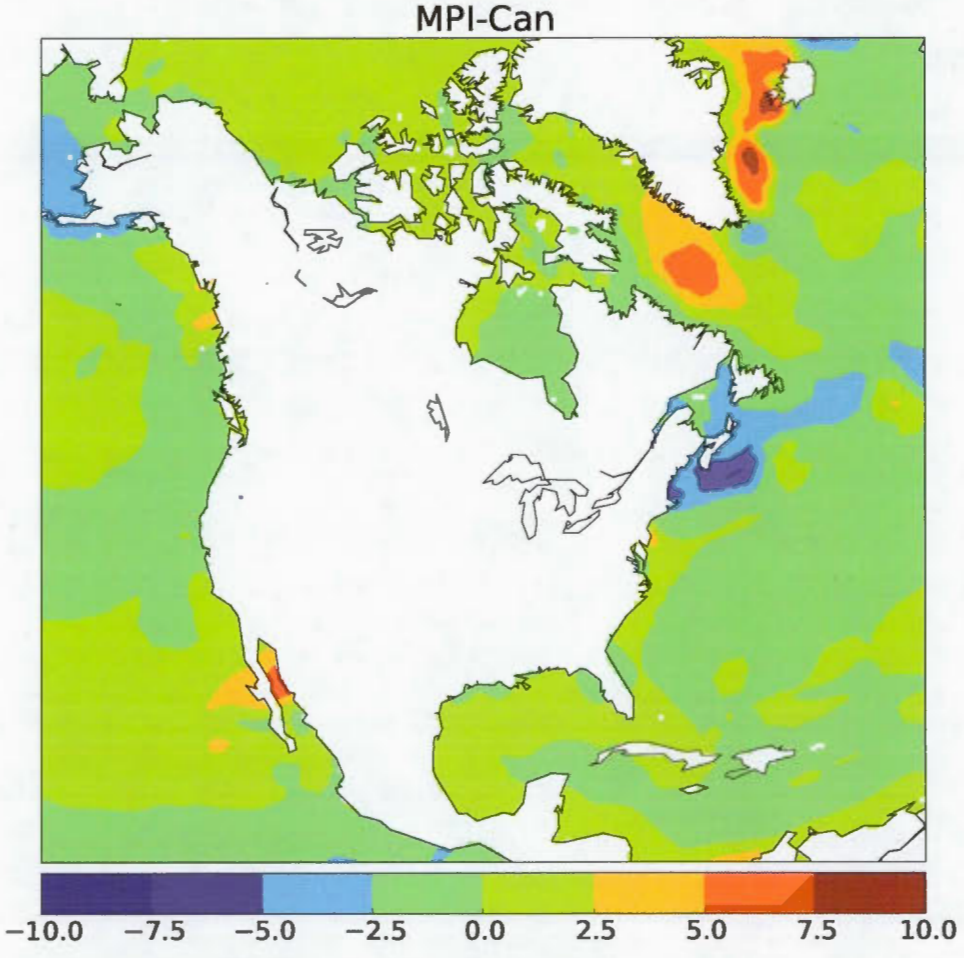

Figure 2.6 Difference in mean DJF sea surface temperature (in °K) between that of the MPI-ESM and that of the CanESM2 averaged over the domain of the CRCM

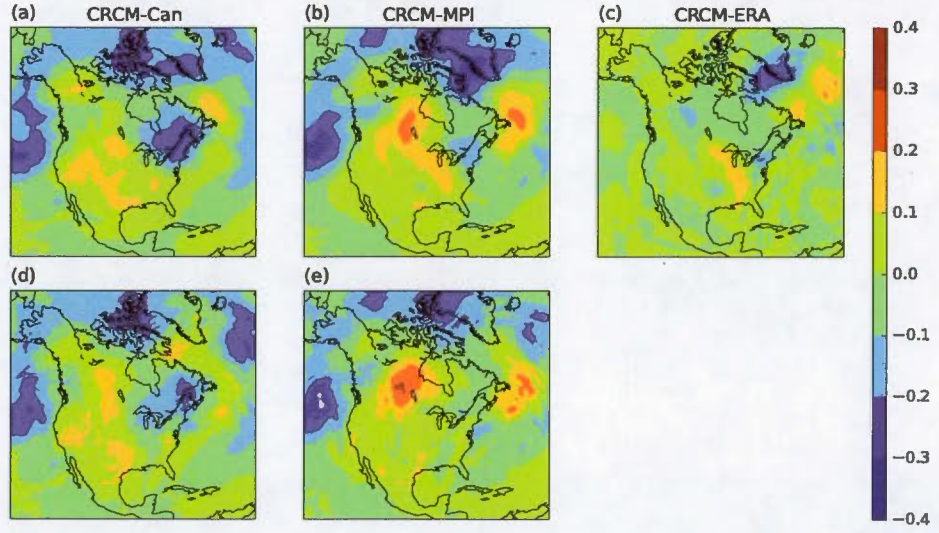

Figure 2.7 Temperature tendencies biases in K/day computed at 200 hPa for CRCM-Can (a), CRCM-MPI (b) and CRCM-ERA (c). Lateral boundary effect biases CRCM-Can-BC and CRCM-MPI-BC are shown in d and e respectively

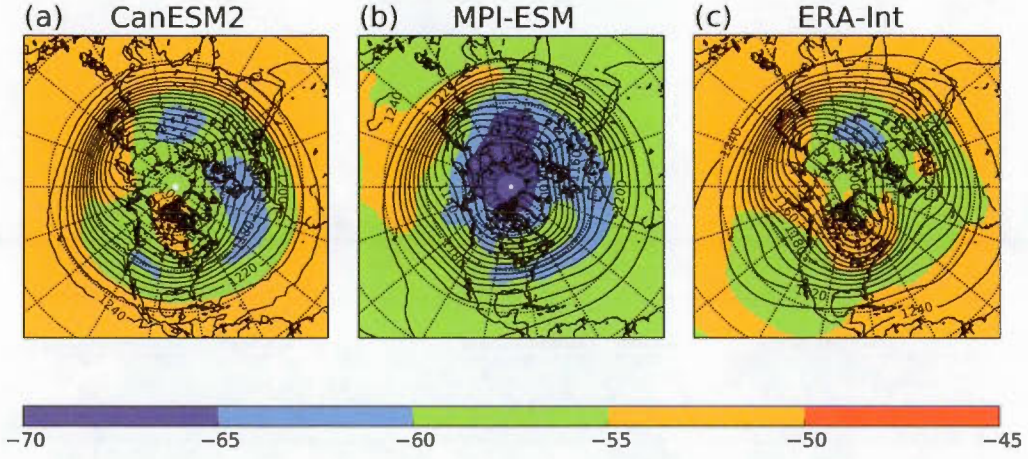

Figure 2.8 DJF mean temperature (°C) and geopotential height (gpm) at 200 hPa computed for CanESM (a), MPI-ESM-LR (b) and ERA-Int (c). Temperature is in *shaded colors* while contours represent geopotential height. The latter, plotted every hundred meters, is presented as indicative of the mean general circulation

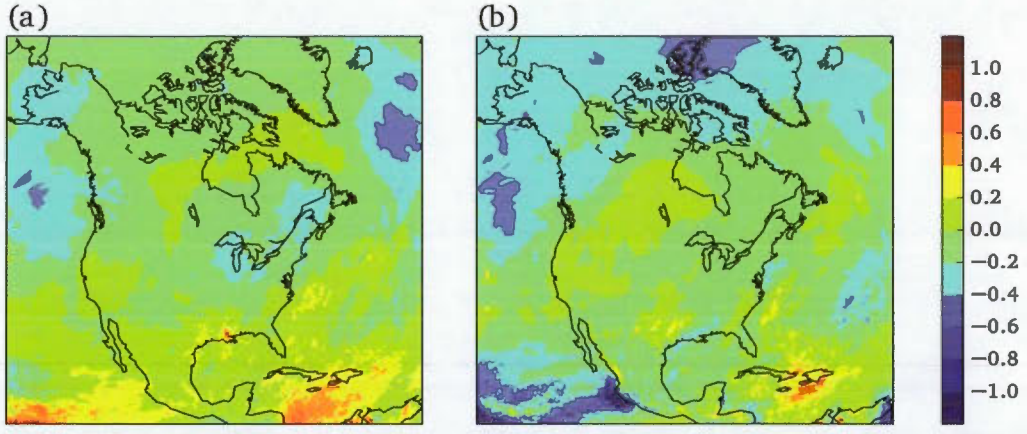

Figure 2.9 Temperature tendencies lateral biases in K/day computed for dynamics process at 200 hPa for CRCM-Can (a) and CRCM-MPI (b)

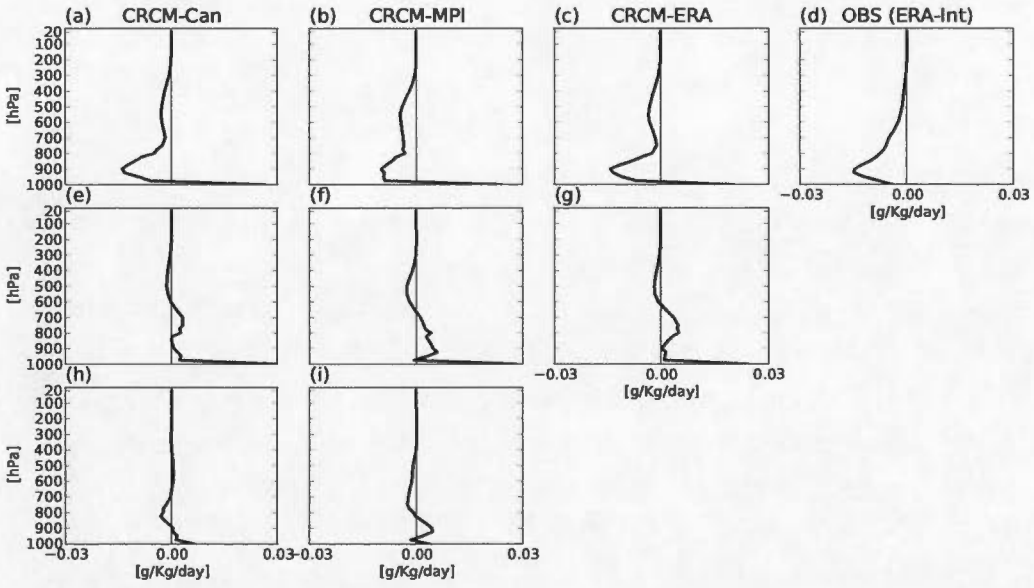

Figure 2.10 Same as Fig. 2.2 but for specific humidity

the total bias profiles (Fig. 2.10e, f) reveal differences compared to observations. In fact, a clear abnormal moistening in lower layers (up to $\sim$ 700 hPa) is observed in both CRCM-Can and CRCM-MPI. This excessive moistening is stronger in the first surface layer (between 1000 and 975 hPa). An abnormal drying (negative bias) is present at levels above 700 hPa particularly in the CRCM-MPI case. The structural bias profile (Fig. 2.10g) exhibits a large moistening in surface layer and low levels and a net drying above $\sim$ 650 hPa. The differences between CRCM-Can-BC and CRCM-MPI-BC biases are attributed to the different boundary data (Fig. 2.10h, i) : they indicate a moistening impact near the surface up to $\sim$ 900 hPa and a drying above with higher values observed in the CRCM-MPI simulation. However, the surface moistening is not as large as that observed in total biases. This suggests that this intense positive bias is mainly structural and can be attributed to the regional model itself through its parameterization schemes. This could be an excessive moisture transport from the surface by vertical diffusion or higher evaporation from large scale condensation.

The total tendency is made of different contributions associated with the individual processes. Figure 2.11 shows the process tendencies and biases associated with the net tendency profiles (Fig. 2.10). The lateral boundary effect biases (Fig. 2.11d, e) provide interesting information. Indeed, one can notice a higher vertical diffusion (brown line) in the CGCM-driven simulations with a stronger peak observed at $\sim$ 850 hPa in CRCM-MPI case (Fig. 2.11e). This means that more humidity is brought from the surface to these levels leading to a stronger shallow convection activity in the boundary layer (red line). In addition, more moistening is observed above 850 hPa in the CGCM-driven simulations due to dynamics (blue line). This indicates that more humidity is penetrating into the domain of the model through the lateral boundaries but also from the surface. This humidity injection leads to an excessive large scale condensation. Humidity from the surface is carried higher up through the vertical diffusion and then contributes greatly to intensify the deep convection as observed in mid-levels in lateral bias profiles of the CRCM-MPI simulation (Fig. 2.11e). Comparing the two CGCM-driven simulations, the CRCM-MPI shows more vertical diffusion resulting in more convection and large scale condensation especially below 700 hPa.

2.5 Tendencies climatology

An important aspect of this study is that the analysis examines only one winter season and the tendencies were averaged over the whole period and over the whole region (North America). Given the large extent of the domain, the synoptic scale variability averages out so that one can expect a net tendency that reflects the seasonal change in temperature and humidity. To confirm that these results are representative of what could be expected in other seasons, the ERA-Interim DJF tendency climatology has been evaluated over a period of 30 years (1979–2009). This climatology along with its variability is presented in Fig. 12. The natural variability expressed by the standard deviation (shaded area in Fig. 2.12) is relatively weak for temperature tendency (Fig. 2.12a). It is however larger above $\sim$ 200 hPa. As expected, the specific humidity tendency shows more variability (Fig. 2.12b).

When compared to this natural variability, the total tendency biases com-

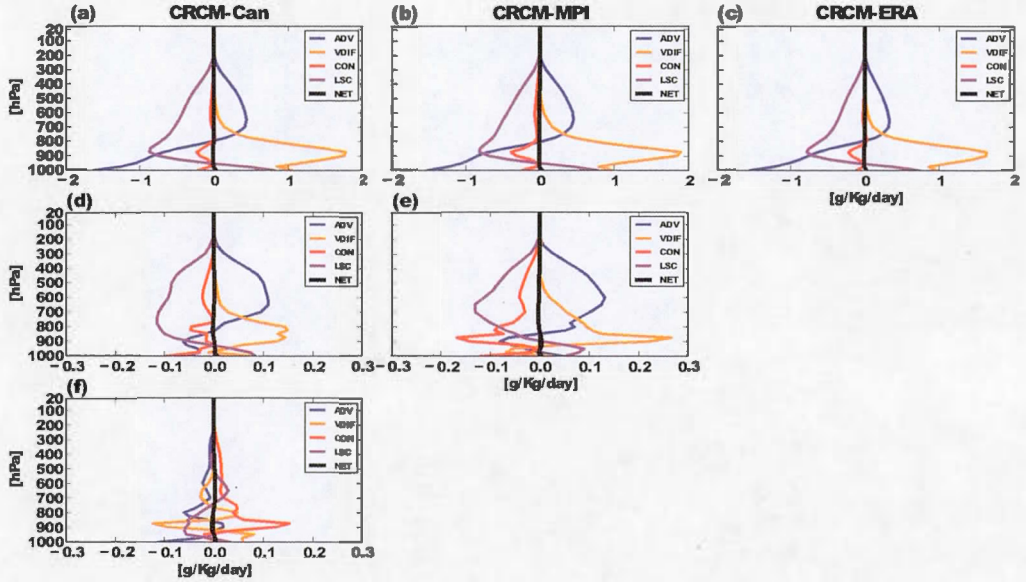

Figure 2.11 Same as Fig.2.4 but for specific humidity tendencies in g/Kg/day. Color coding is the same as in Fig. 2.4

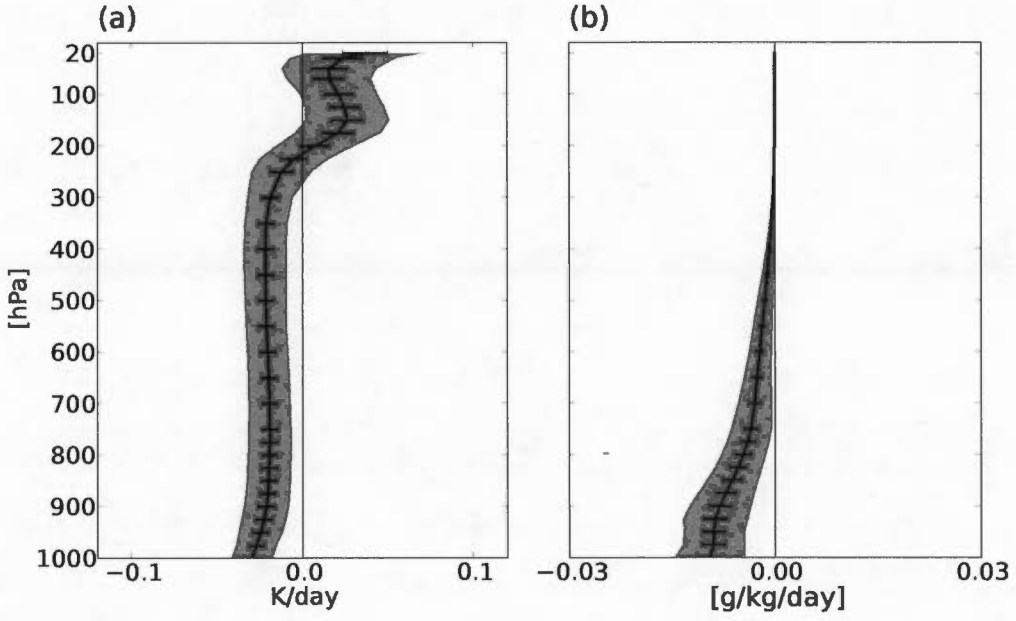

Figure 2.12 ERA-Interim DJF tendency climatology for temperature in K/day (a) and specific humidity in g/kg/day (b). *Horizontal segments* denote 95% confidence intervals. *Shaded area* represents the tendency standard deviation around the mean

puted for temperature and specific humidity exhibit larger values. This suggests that these biases can be attributed to deficiencies in the model simulations. Furthermore, the structural biases are found to be predominant relative to those attributed to boundary conditions. One could then argue that a CRCM simulation would lead to similar structural bias in any other analogous DJF season. The lateral boundary biases are also larger than natural variability demonstrating the notable driving data impact. However, it is important to remember that there could be noticeable differences from one winter to another, as was seen in Fig. 2.8 showing evident differences in temperature for the three experiments. However, these differences would not sensibly impact the results presented here for the total tendencies averaged over the domain.

2.6 Concluding remarks

The impact of prescribed boundary conditions on a regional climate model, i.e. the CRCM5, has been studied using three different data sources to drive the model, namely, outputs from two CGCMs (CanESM2 and MPI-ESM) and ERA-Interim reanalysis considered here as the reference. Following Rodwell and Palmer (2007), total and individual process tendencies averaged in time and space were computed for the three simulations. The tendencies analysed are computed for the DJF season after a 14 months integration starting from the initial conditions provided by the state of a 30-year climate simulation of the CRCM5 when driven by ERA-reanalyses. In this way, the model can be considered well spun up and the boundary data effect significant. The trend for temperature computed from the reanalysis is removed from that obtained for the CanESM2 and MPI-ESM to obtain the total bias. The bias due to lateral boundary conditions is then obtained by subtracting the CGCM-driven simulation tendencies from those obtained by the ERA-interim driven one.

We examined tendencies for two prognostic variables, temperature and specific humidity. Results showed different CRCM5 behaviour through the DJF season in every simulation. While the observed temperature tendency indicates a cooling in almost the whole column, CRCM5 exhibits a warming in lower layers and a stronger cooling above. The structural bias reveals that the abnormal warming seen in lower layers is mainly due to deficiencies in the regional model itself. Nevertheless, the two CGCMs are found to affect differently the CRCM5 especially in lowest layers where MPI-ESM tends to warm the model while CanESM2 induces a cooling. In higher layers above 500 hPa, the two CGCMs tend to cool CRCM5 with a peak around 200 hPa. When looking at specific humidity, results showed an excessive moistening in the lower layers for all simulations with a peak near the surface. At higher levels, a small drying is observed for the three simulations.

Individual process tendencies revealed additional details explaining the previous findings. For example, we noted a stronger convective activity in CRCM-Can and CRCM-MPI as shown in temperature and specific humidity tendencies for this process. This is explained by favourable conditions to trigger convection, having

more humidity available in the lower levels. It was also shown that, compared to CanESM2, MPI-ESM driving leads to more convection due warmer marine surfaces and more water vapour available in low levels supplied via vertical diffusion. The cold bias peak around 200 hPa is clearer when looking at dynamics process tendencies and can be attributed , at least partially, to colder CGCMs conditions for that DJF season. In surface layers, MPI-ESM is shown to heat the CRCM5 due to a strong vertical diffusion essentially. This is not the case in CRCM-Can leading to a model cooling when driven by CanESM2 data.

This study has shown that the CRCM5 is sensitive to the imposed driving data provided by CGCMs simulations. Consequently, climate predictions from regional climate models can be very different depending on the supplied driving data. We have to mention that only one DJF season has been examined in this work and extending the results for DJF for a 30 years simulations is required to produce reliable conclusions on the results such as those presented in Figs. 2.7 and 2.9. Nevertheless, this study showed that the process tendencies can bring different, additional and useful information regarding the RCMs sensitivity to driving data.

Acknowledgements

The authors would like to thank Mrs. Katja Winger and Dr. Bernard Dugas for their support through numerous discussions and advice. Comments from Dr. Oumarou Nikiema and Prof. René Laprise on early results helped us to better understand the implications of our results. This research was funded by the *Canadian Network for Regional Climate and Weather Processes* (CNRCWP) funded through the *Natural Sciences and Engineering Research Council of Canada* (NSERC) *Climate Change and Atmosphere Research programme* (CCAR), the Québec's *Ministère du Développement Economique, de l'Innovation et de l'Exportation* (MDEIE), the Grants and Contribution program of *Environment Canada* and NSERC's Discovery Grant program. High performance computing resources were provided by *Compute Canada* on the Guillimin platform of the Calcul Québec regional consortium. The first author benefited from a scholarship form the Faculty of Science of the Université du Québec à Montréal.

CHAPTER III

ASSESSMENT OF REGIONAL CLIMATE MODELS THROUGH DATA ASSIMILATION

This chapter is presented in the format of a scientific article. It is submitted to the Quarterly Journal of the Royal Meteorological Society. This manuscript is entirely based on my work, with the co-author involved in results analysis and interpretation and also text revision.

Assessment of regional climate models through data assimilation

Kamel Chikhar and Pierre Gauthier

Centre ESCER (Étude et Simulation du Climat à l'Échelle Régionale),
Department of Earth and atmospheric sciences
Université du Québec à Montréal (UQAM)
B.P. 8888, Succ. Centre-ville
Montréal (Québec) Canada H3C 3P8

Abstract

Regional and global climate models are usually validated by comparison to derived observations or reanalyses. The purpose of this paper is to build and evaluate a regional data assimilation system based on the fifth-generation Canadian Regional Climate Model (CRCM5) over North America. This system uses the operational variational data assimilation system of the Meteorological Service of Canada (MSC) and assimilates the same sets of observations. The motivation of this work is to validate the CRCM5 using the information made available by data assimilation and also through the initial tendency diagnostics proposed by Rodwell and Palmer (2007). This fully cycled regional assimilation system based on a limited-area model produced analyses over the months of January and July 2011. The results show that the main features of the analyses compare well with either global or regional analyses. However, the mean analysis increment could be associated with problems with the lateral coupling with the driving data provided by global analyses. This is detrimental to the forecasts particularly in the Northern part of the domain during the winter period. The initial tendencies indicate that this is indeed due to advection near the Northern boundary. These results are nevertheless very encouraging as the limited-area model had never been fully cycled before. This opens up the possibility of doing analyses with this regional climate model for the benefit of in-depth diagnostics of several aspects of the model associated with fast-acting physical processes.

Keywords : regional climate model, regional data assimilation, data monitoring.

3.1 Introduction

When used in data assimilation, a model is constantly compared to observations. Based on statistical estimation principles, a short-term model forecast from the previous analysis, is drawn towards the observations through the assimilation process that builds a correction to the background state, the analysis increment, which should in principle be unbiased. Therefore, a systematic correction is indicative of a bias associated with error in either the observations or the background state itself. On the other hand, Rodwell and Palmer (2007) (RP2007, hereafter) pointed out that a bias in the analysis increment corresponds to an opposite systematic physical tendency observed in the first instants of the 6-h forecast. A diagnostic based on tendencies therefore provides useful information to diagnose the fast-acting processes of the model. It is expected that the total tendency should average to zero as would the analysis increment. This suggests that data assimilation could be valuable even for climate models as a diagnostic approach to test, for example, different configurations to prevent the emergence of spurious internal variability associated with unbalanced physics in the model. This has been the underlying motivation of the work presented in this paper and two previous papers (Chikhar and Gauthier, 2014, 2015) .

In Chikhar and Gauthier (2014), the Global Environmental Multiscale (GEM) model and the Canadian regional climate model (CRCM5) were evaluated by analysing their initial dynamical balance based on the initial tendency diagnostic of RP2007. It is important to mention that the CRCM5 is very close to the GEM limited-area model (GEM-LAM) used for regional forecasts and analyses at the Meteorological Service of Canada (MSC). As explained in Chikhar and Gauthier (2014), this approach consisted in analyzing the mean tendency over a long enough period for a prognostic variable by considering all the dynamical and physical processes involved in updating this variable during the model integration. This study used existing analyses, namely analyses from MSC (Gauthier et al., 2007; Laroche et al., 2007) and those of ERA-interim (Dee et al., 2011), with the objective of studying the sensitivity to the initial and boundary conditions model used to initialize the two models. Bear in mind that the MSC global analyses are driven by a forecast model very similar to the model being assessed. They showed that

using different analyses as initial conditions has a significant impact on the initial tendencies. On the other hand, Chikhar and Gauthier (2015) showed that the physical and dynamical tendencies could detect a sensitivity of a regional climate model to changes in the lateral boundary conditions. This work was motivated by the fact that contrary to global climate simulations, regional climate models are constantly driven at their boundaries by analyses produced by different models with quite different physical parameterizations and spatial resolutions (vertical and horizontal).

These studies and others (e.g., Rodwell and Jung (2008)) revealed the usefulness of the tendencies diagnostics in model development and evaluation. This suggested to use the CRCM5 within a data assimilation system as is routinely done in numerical weather prediction. This is the essential motivation of the work presented here. Taking advantage of the fact that the CRCM5 is very similar to the limited-area regional model of MSC, GEM-LAM, used to produce regional analyses (Caron et al., 2015), it was technically possible to use the CRCM5 instead of the GEM-LAM in data assimilation and therefore, benefit from the immense work done to validate the system for the large volume of assimilated data. Validating a model in this context is a long process and this approach avoided many difficulties that arise when building a data assimilation system from scratch (e.g., quality control of the observations, detailed study of each observation operator, tuning of the error statistics). The objective of the study was to see first if the system could be cycled in time without drifting. Moreover, it was envisioned that the analysis increments and the physical tendencies on top of the data monitoring could help us pinpoint the source of problems if any.

In collaboration with Environment Canada (EC), the MSC regional variational data assimilation system (Buehner et al., 2015; Caron et al., 2015) was implemented on university computers supported by Compute Canada. This permitted the assimilation of all observations currently used at MSC. On the other hand, producing regional analyses in a fully cycled data assimilation system gives rise to difficulties related essentially to the way the regional model is driven through the lateral boundaries. This is because the lateral driving can induce important contrasts between the driving data and the regional model forecast due to several

causes such as differences in spatial resolution or different schemes used in parameterizing the physics in the driving and driven models. The resulting short range forecast provides the background state to the assimilation system but could be very different from reality and even unrealistic in the vicinity of the nesting zone. Consequently, very large deviations can occur between the model and observations which could lead the variational quality control to give very little weight to these observations or even reject them.

In many regional assimilation systems, the regional model used is driven by its global version or global reanalyses (e.g., Fillion et al., 2010; Mesinger et al., 2006). In our study, the lateral driving issue will be addressed by using, first, the ERA-Interim reanalyses to test their compatibility with the regional model. In a second step, global analyses based on GEMCLIM model, the CRCM5 global version, are used to drive the regional model. Some preliminary assimilation results for January and July 2011 will be presented.

In section 3.2, the experimental framework will be presented describing the assimilation system as well as of the model used. A preliminary evaluation of the regional assimilation is presented in section 3.3. A deeper investigation is presented in sections 3.4 to 3.6. An overall verification of this regional data assimilation system is presented in section 3.7 and, finally, concluding remarks are given in section 3.8.

3.2 The regional assimilation system design and experimental framework

The objective is to use the CRCM5 in data assimilation with a view of assessing its behavior as a short-term forecast model putting a particular emphasis on the balance between the different physical processes acting on temperature and also, humidity. The configuration of the CRCM5 used is the same as that described in Chikhar and Gauthier (2014). Namely, it uses a hybrid coordinate with 80 vertical levels up to 0.1 hPa. The horizontal resolution is also the same (around 20 km) with a timestep of 10 minutes. The background provided by the CRCM5 are forecasts valid at every time step for a 6-h period centered on the time of the

analysis. The domain used in the experimentation (see Fig.3.1) is the same as that used in the COordinated Regional Climate Downscaling EXperiment (CORDEX) over North America (Šeparovic et al., 2013).

The assimilation system is based on the ensemble-variational data assimilation (4DEnVar) recently implemented at MSC for deterministic weather prediction (Buehner et al., 2015; Caron et al., 2015). The 4DEnvar is an incremental variational assimilation system based on the use of hybrid background-error covariances with a weighted average of the static covariances B_{nmc} formulation, used in the previous 4D-Var and 3D-Var systems, and a flow-dependent 4-D ensemble covariances B_{ens} computed from ensembles, produced by the Ensemble Kalman Filter assimilation system also used at MSC (Houtekamer et al., 2014). In our experiments, only the static covariances B_{nmc} were used making the 4DEnVar a 3D-Var FGAT (First Guess at Appropriate Time). The minimisation employs the M1QN3 algorithm (Gilbert and Lemaréchal, 1989) with 70 iterations, the first 5 iterations being done without the variational quality control (QC-Var) (Gauthier et al., 2003). The assimilated observations located within the model domain include those from radiosondes, aircraft, surface land stations, buoys, ships, wind profilers, scatterometers, microwave and infrared satellite sounders and imagers, atmospheric motion vectors and satellite-based GPS radio occultation. The analysis increment is computed globally at a lower resolution ($\sim$ 100 km) and then interpolated to the CRCM5 higher resolution over the model domain.

The next section presents a preliminary evaluation of the regional assimilation system.

3.3 Evaluation of the regional assimilation system

To evaluate the regional assimilation system, assimilation cycles have been performed for the period spanning the months of January and July 2011. For each cycle, prior assimilation-forecast runs were done over a week to allow a spinup of the assimilation and the regional model. During these cycles, the CRCM5 model is driven 6 hourly at its lateral boundaries by ERA-Interim reanalyses.

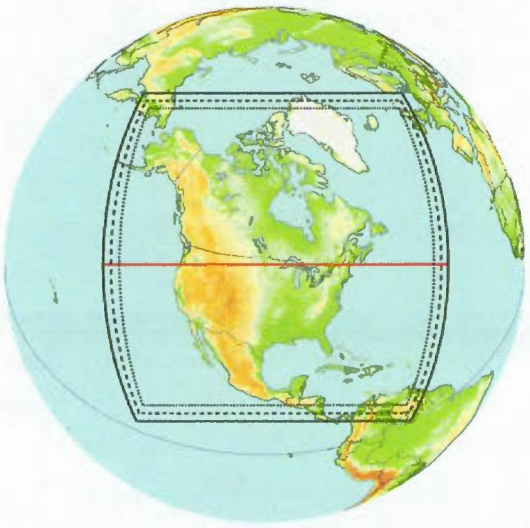

Figure 3.1 The domain used in the experimental regional assimilation system. The inner thin dotted lines indicate the free model zone limit while the area between the two dotted lines represents the blending region. The red line indicates the grid equator.

3.3.1 The mean analysis increments

The analysis increment is defined as the correction applied to the model forecast or background to obtain the analysis. Figure 3.2 shows the mean temperature analysis increment averaged over January and July 2011 at different vertical levels. It shows that, in July, the mean temperature analysis increment is small over a large part of the domain except at 100hPa level where slightly positive increments denote a relatively cold bias in the model. In January, large analysis increments are observed over the Northern part of the model's domain particularly over the Northern Canadian archipelago at 100hPa and 250 hPa levels as shown in Fig.3.2-a. In the lower part of the atmosphere however, the mean increments values are lower. The large differences between the model forecast and the analyses indicate that the observations are strongly correcting the background.

The relatively large negative values in the mean analysis increment over the Northern Canadian archipelago could be attributed to two possible causes. It may be because of some 'weaknesses' of the model causing a departure from 'reality'

defined by the observations in those regions which would indicate that the model is too 'warm' in the Northern part of the domain. Or it could be with the observations which may have a cold bias in that region. To complete the investigation, we now look at the monitoring of radiosondes particularly those located in the Northern part of the domain where large mean increments are observed. The background (6-h model forecast) is then compared with those observations.

3.3.2 Observation departures from background and analysis

In operational assimilation systems, observations go routinely through a quality control procedure including monitoring in which observations departures from the background state, here a 6-h forecast. These departures $[O-F]$ (or residuals) are the innovations and represent the difference between the model forecast and the observation in observation space. It is expressed as $\mathbf{d} = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$ where $\mathbf{y}$ is the observation, $\mathbf{x}_b$ the model forecast and H the so-called observation operator. The analysis residuals $[O-A]$, defined as $\mathbf{a} = \mathbf{y} - \mathbf{H}(\mathbf{x}_a)$ where $\mathbf{x}_a$ is the analysis, can also be looked at together with the innovations to assess how the analysis is close to observations or the background and to what extent the observations are correcting the background.

In the region where the anomalous mean analysis increment was observed over North of Canada in January (see Fig.3.2), there are a few radiosondes located as shown in Fig.3.3. Fig.3.4 shows the monitoring for all of them in January, with the innovations $[O-F]$ represented in blue and the $[O-A]$ in red. The innovations present large values indicating a background deviation from observations. The $[O-A]$ small values compared to innovations indicate that the assimilation draws the background towards the observations. At all levels, the mean innovations are negative meaning that the model is too warm compared to observations. This is particularly strong in the second half of the month and then stronger in the upper levels where the departure goes down to -8°C. The experiment was extended to the month of February and the results indicate that the mean innovations subside only to reappear after a few days later highlighting possible problems with the model. Moreover, temperatures observations from these sounding stations being of good quality, large departures from these observations are indicative of deficiencies in

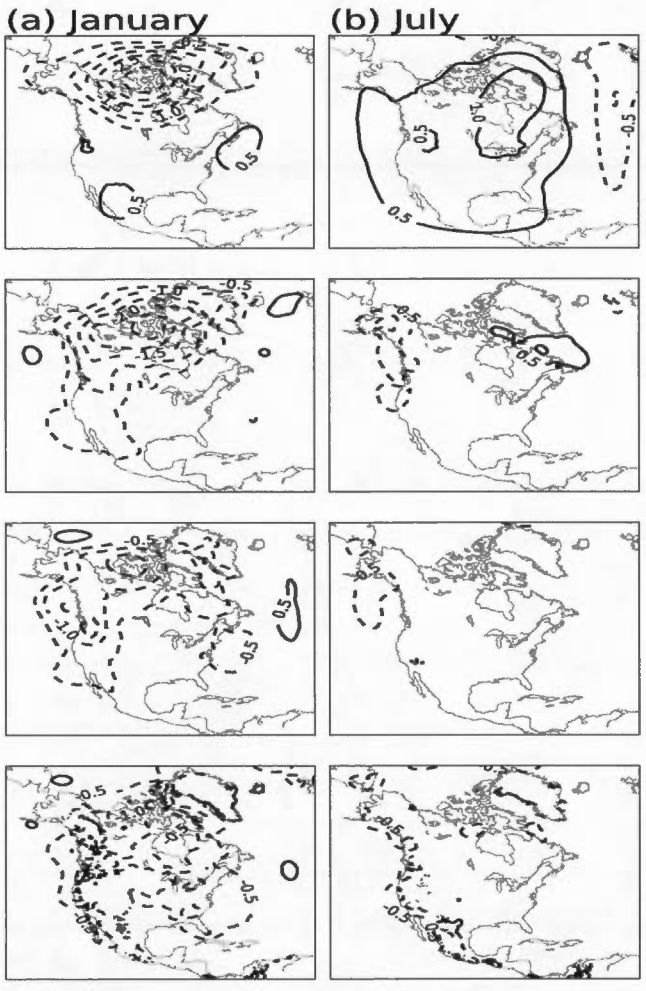

Figure 3.2 Temperature mean analysis increment in °C averaged over (a) January 2011 and (b) July 2011 for pressure levels 100hPa, 250hPa, 500hPa and 850hPa from top to bottom respectivey.

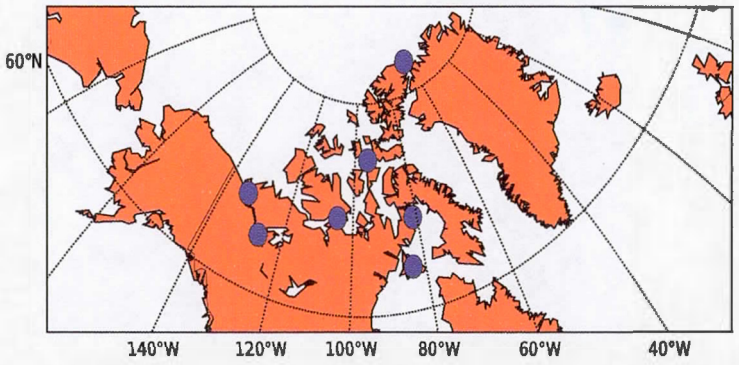

Figure 3.3 Sounding stations used in the data monitoring indicated by blue filled circles.

the model. In the summer case, (Fig.3.4-b) shows very small $[O-F]$ and $[O-A]$ indicating that the assimilation system is doing well and the model forecast is also close to observations except at 100hPa level where slightly positive $[O-F]$ are observed.

A new cycle has been performed over the month of January but using only radiosondes observations to rule out possible problems with the assimilation of satellite data over continents.

3.4 Assimilation using only radiosonde data

To better understand the origin of the problem observed in the Northern domain and confirm that it originates from model difficulties, it was decided to do another assimilation experiment for January in which only observations from radiosondes are used. These observations can be considered as a reference. It was suspected that satellite observations could be biased and contribute to the large analysis increments observed at higher levels. The mean analysis increment obtained from this cycle is shown in Fig.3.5 where it can be can noticed that the problem observed in the Northern part of the domain (Fig.3.2) has disappeared even at 100hPa. In this case, the mean increments are very low over the same region. The data monitoring on the same radiosondes as before (Fig.3.6) shows very large $[O-F]$ confirming that the large differences between the model and observations remain.

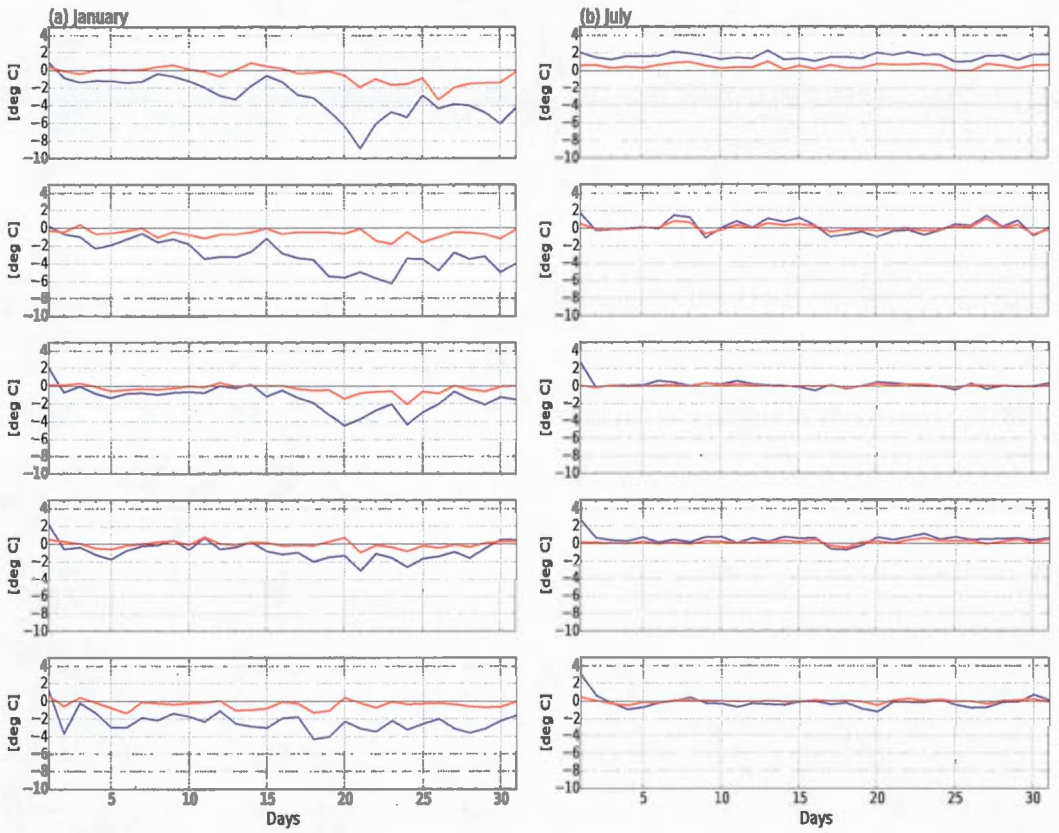

Figure 3.4 Mean $[O-F]$ (blue line) and $[O-A]$ (red line) for temperature soundings in the Northern Canada indicated in 3.3 over the month of January (left) and July (right) 2011 at levels 100hPa, 250hPa, 500hPa, 700hPa and 850hPa respectively from top to bottom. See text for $[O-F]$ and $[O-A]$ definitions.

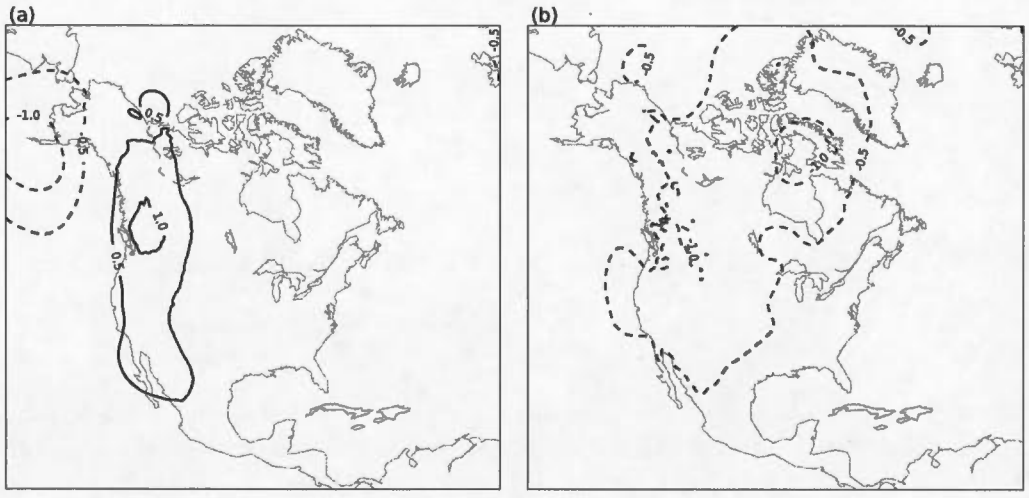

Figure 3.5 Mean analysis increment over the month of January 2011 at levels 100hPa (a) and 250hPa (b) from January 2011 assimilation cycle using only radiosondes observations.

In addition, the large $[O-A]$ is a clear indication that the observations have been rejected. In other words, the radiosonde observations located in the problematic region are not used to correct the model and are simply rejected by the QC-Var during the assimilation due to the large differences between these observations and the forecast.The combined results from the mean analysis increments and the data monitoring clearly demonstrate that the model forecasts are problematic in the northern part of the domain. On the other hand, rejecting good observations as those from radiosondes reveals that the variational quality control can be detrimental to the assimilation system in situations where the model forecast is far from what is observed.

The QC-Var assigns weights to observations based on the departure of the estimated analysis with respect to observations. When it is larger than the observation error, the QC-Var would assign a very low value to that observation. The mean departures observed in the previous experiment reach values of -8°C which is way larger than the prescribed observation error standard deviation of around 1°C used here. Moreover, as indicated in Gauthier et al. (2003), a flat Gaussian probability distribution is used in the variational system which would

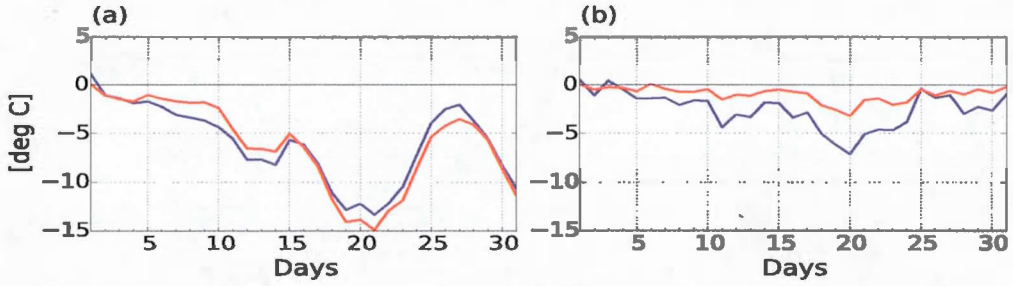

Figure 3.6 Mean $[O-F]$ (blue line) and $[O-A]$ (red line) computed for temperature from the same soundings considered in Fig.3.4 at levels 100hPa (a) and 250hPa (b) from January 2011 assimilation cycle using only radiosondes observations.

lead to a virtually null weight. In this way data far from the analysis are given very low weights and implicitly rejected. Variational quality control is designed to prevent eliminating good observations which initially disagree strongly with the background. In our regional analysis system, the minimization is performed with a limit of 70 iterations. The QC-Var is turned off during the first 5 iterations in which the preliminary analysis is built by considering all observations and move away from the background. It is then turned on during the rest of the minimization.

Fig.3.6 revealed some limitations in the QC-Var which rejected good observations specially at 100hPa level. As we trust the radiosonde observations, the same cycle has been restarted from the middle of January switching off the QC-Var during the remaining cycle period. Results presented in Fig.3.7 show smaller $[O-A]$ indicating that radiosondes temperatures are now correcting significantly the background even if the $[O-F]$ are still large. These results confirm that the QC-Var is rejecting incorrectly good quality data when these latter disagree strongly with the background. A possible reason is that 5 initial minimization iterations where the QC-Var is switched off are not sufficient to constrain the analysis towards the observations. Another reason could be the configuration of the QC-Var itself. To prospect the first issue, an experimental cycle using only radiosondes observations is performed where the QC-Var is switched off during

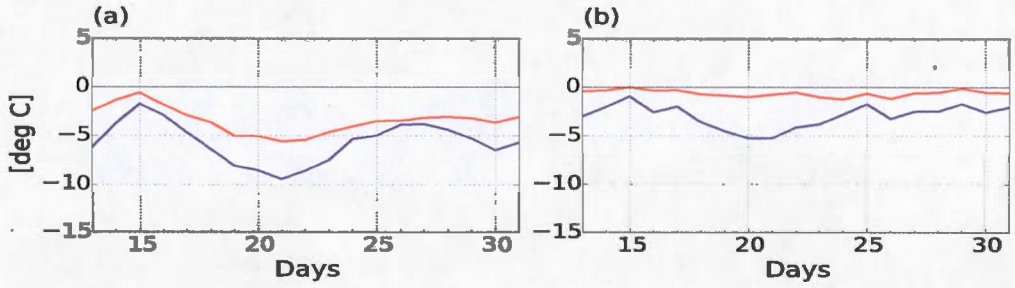

Figure 3.7 Same as Fig.3.6 but for assimilation cycle using only radiosondes observations and QC-Var turned off.

the first 30 iterations during the minimization. The resulting data monitoring (not shown) indicates slightly smaller analysis residuals and innovations reflecting that more observations have been used to constrain the analysis. However, these values remain larger than those obtained when the QC-Var is switched off completely indicating that the QC-Var can reject rapidly observations considered as outliers because of an excessive departure from the background. This experiment also shows that QC-Var being tightly associated with the quality of the background, its performance is seriously compromised in the case of poor quality first-guess. It is also noted that turning off the QC-Var brought back a mean analysis increment similar to what was observed in the experiment using all observations.

In the previous sections, it has been demonstrated that the large mean analysis increment observed in the Northern part of the domain is due to weaknesses in the model forecast that produces the background state. These problematic forecasts over that region need to be addressed and investigated more deeply to identify precisely the causes. In the next section, the initial systematic tendency diagnostic of RP2007 is used to examine the different processes of the model to better understand the source of the error.

3.5 Use of initial systematic tendency technique

The large mean analysis increments in the Northern domain indicate problems in the model forecasts. The initial systematic tendency technique of RP2007 is now used as described in Chikhar and Gauthier (2014). The mean initial tendencies profiles, averaged over the whole domain of the model (excluding the blending zone) extracted from the different processes updating the temperature variable are presented in Fig.3.8. These profiles reveal an abnormal heating from dynamics over most of the model levels leading to an excessive heating consistent with the large negative mean analysis increments noted previously. On the other hand, the 2-D mean initial tendency related to dynamics (Fig.3.9) shows a large warming at 100hPa and 250 hPa levels with maxima located over the same region as what was obtained when looking at the mean increments. This clearly indicates that the problems from the model forecast originate from dynamics, namely the advection process over the Northern domain.

On the other hand, very strong gradients can occur in the blending zone when inconsistencies are noted between the driving data and the regional model. An example is presented in Fig. 3.10 showing the geopotential field forecast where very tight isohypses contours denote too strong winds. This situation occurs several times during the January cycle because of mismatch between the driven and the driving model. The upstream point retrieved from the semi-Lagrangian scheme would lead to wrong temperatures values and consequently large temperature tendencies.

As mentioned previously, strong gradients can occur in the blending zone due to mismatch between the driving data and the regional model but also due to the nesting technique. In another experiment, the model domain has been extended further north and west to allow the model to better handle structures imposed through the boundaries. The mean analysis increments resulting from this experimental cycle for January 2011 (not shown) exhibit larger negative values and similar patterns as those obtained with the smaller domain due essentially to a lack of conventional observations as soundings which could correct the background in the polar region. Data monitoring applied to temperature from the same soun-

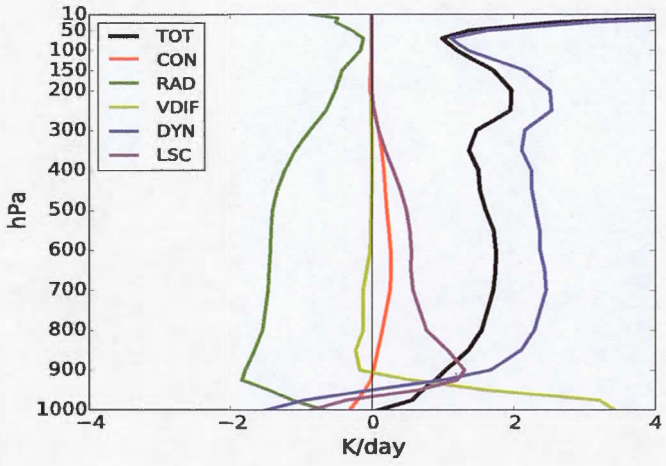

Figure 3.8 Mean temperature initial systematic tendency in K day^{-1} computed for the month of January 2011 and averaged over the free model zone. The different colors indicate the different processes involved, i.e. radiation (green), convection(red), large scale condensation (magenta), vertical diffusion (brown) and dynamics (blue). The black line is the net tendency.

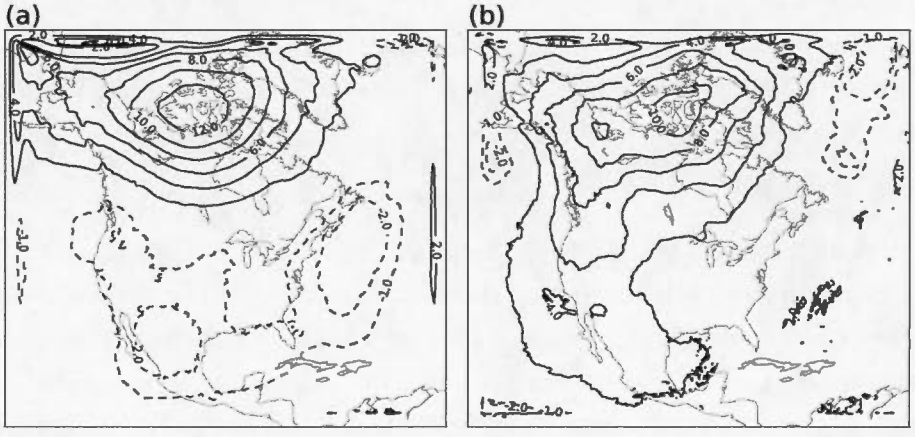

Figure 3.9 Mean temperature initial systematic tendency related to dynamics in K day^{-1} computed for the month of January 2011 at levels 100hPa (a) and 250hPa (b).

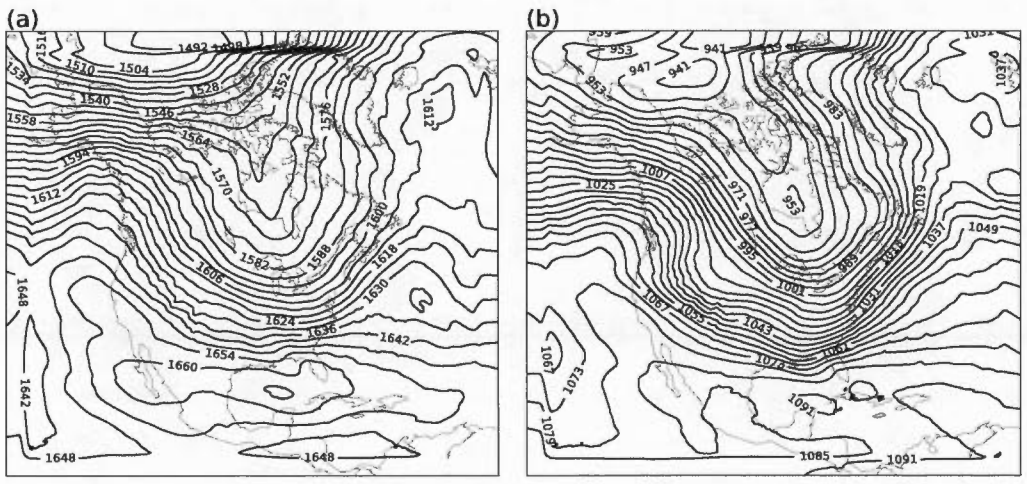

Figure 3.10 Geopotential height 6-h forecast valid on January 22 at 18 :00 GMT for levels 100hPa (a) and 250hPa (b).

dings as previously (not shown) reveals large departures between observations and the forecast suggesting that inconsistencies between the driving data and the regional model still exist over the northern domain and, as before, introduce a poor background in the assimilation system. Results from this experiment show that the blending across the Northern polar region can be tricky because of the special winter weather regime there.

Another experiment was also conducted in which the blending zone, where the solutions from the regional model are relaxed towards the driving data, has been increased from 10 grid points to 20 grid points. The results (not shown) are almost identical to those obtained with the original blending width. These results suggest that the blending zone width extension has no significant impact on reducing the potential large inconsistencies between the regional model and the driving from lateral boundaries. To go further in our analysis, a global version of the assimilation system has been used to eliminate the boundary driving issues. The results from this experiment and comparison to the regional case are discussed in the next section.

3.6 Comparison to global analyses

The previous experiments showed that the background states provided by CRCM5 forecasts led to important biases in the analysis increments that could be related to the process of imposing boundary conditions to the regional model based on the ERA-Interim driving data. To remove the impact of boundary conditions, another assimilation experiment has been performed in which global analyses were produced using a global version of CRCM5, referred to as GEMCLIM, to provide the background states to the assimilation. The horizontal resolution of GEMCLIM is $\sim 50km$, the number of vertical levels and the height of the top level are the same as the regional version. In this case, there are no lateral boundaries constraints and, consequently, this experiment will be taken as a reference to assess the impact of imposing lateral boundary conditions. Fig.3.11 shows the resulting mean increments over the CRCM5 regional domain which have a very small systematic correction. Moreover, Fig.3.12 shows the results of data monitoring for the same set of radiosondes used in Fig.3.6. The innovations $[O-F]$ as well as the analysis residuals $[O-A]$ are fairly small and do not have the huge bias seen before.

To go further in our comparison, the initial mean tendencies from the global model have been computed over the CRCM5 free zone and are presented in Fig.3.13. The process tendencies especially those from dynamics are much smaller from what was observed before (Fig.3.8). In this case, the mean total tendency as well as that from dynamics are much more realistic and similar to results obtained in Chikhar and Gauthier (2014) for example. In other words, the abnormal advection activity is not present, the dynamical balance is realistic suggesting that the model is correctly performing over the Northern region contrary to what has been noted in the regional case.

Since the global assimilation cycle uses a model that is very similar to the CRCM5 to provide the background states, these results provide good evidence that problems experienced in the regional case for January can be attributed to the lateral driving. Moreover, this lateral driving could be detrimental to the regional model in two ways. The first one is a mismatch between the lateral boundary forcing data and the regional model due to differences in the two models

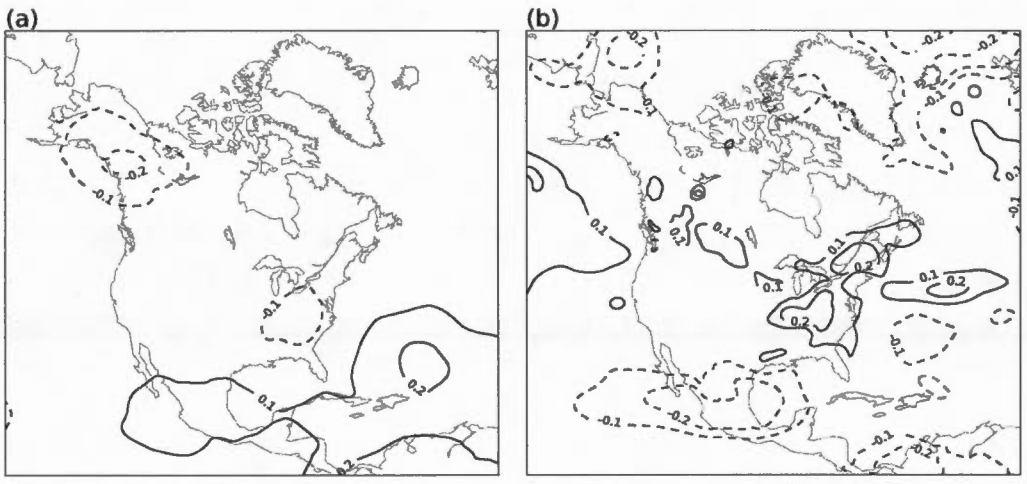

Figure 3.11 Mean analysis increment computed for January 2011 global cycle over the regional model domain for levels 100hPa (a) and 250 hPa (b)

or differences in resolution for example. The second could be, for example, the nesting procedure used to blend the global features from the driving data to the regional model.

The global analyses obtained with the global assimilation experiment, allow us to do a last experiment in which the assimilation was done using CRCM5 driven by our own global analyses deemed to be more consistent with the CRCM5. The resulting mean analysis increments are shown in Fig.3.14 and the results

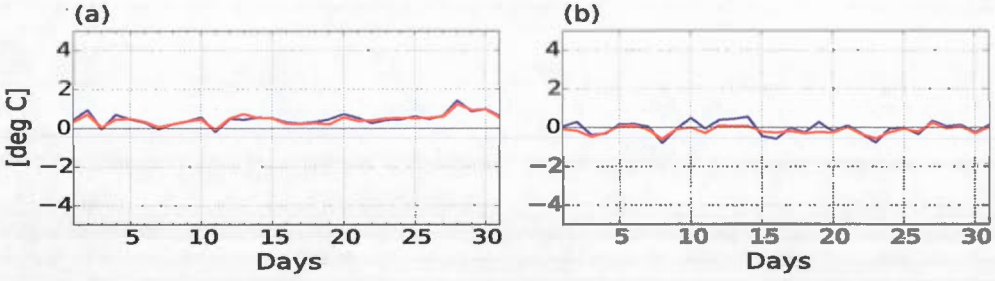

Figure 3.12 Same as Fig.3.6 but for the global cycle.

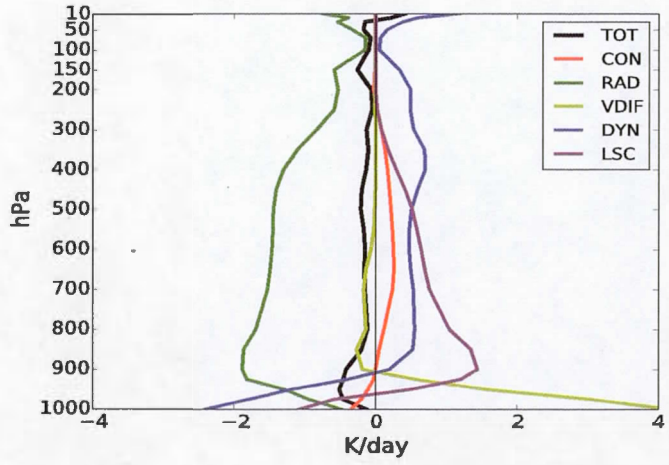

Figure 3.13 Mean initial tendencies in K day^{-1} from the global model averaged over the CRCM5 free zone.

are similar to what obtained in the case where CRCM5 driven by ERA-Interim reanalyses. This suggests that having driving data from the same model as the driven one does not prevent from mismatch and inconsistencies in the blending zone. These results also reinforce the idea that the nesting procedure in terms of nudging strength or frequency can be detrimental to the regional model forecast and consequently to the assimilation system in regions where observations are not available to correct the background.

3.7 Overall verification

Fig.3.2 and Fig.3.4 indicate that the problems with the regional analyses are more acute at upper levels than in the lower part of the atmosphere. As surface temperatures are of particular interest for regional climate studies, the quality of the CRCM5 regional analyses is examined by evaluating the bias and rms error for 2-m background temperature fits to observations for January and July 2011. Fig.3.15 indicates that the model presents a light negative bias (around 1 deg) in July while in January the model bias is more variable with higher day to day variations. The rms errors are also higher in January compared to those observed in July which exhibit a lower values and lower day to day variations.

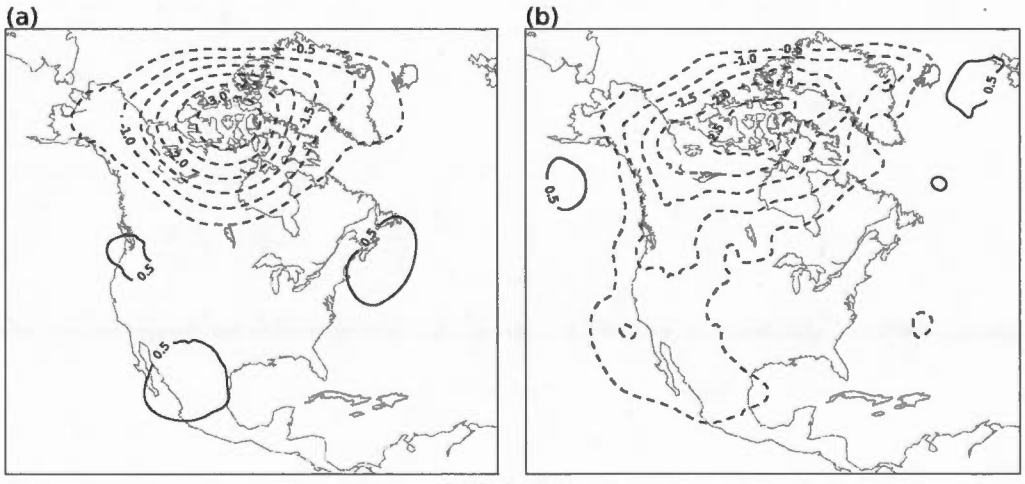

Figure 3.14 Mean analysis increment computed for January 2011 for levels 100hPa (a) and 250 hPa (b) from cycle where CRCM5 is driven by global analyses produced using the global cycle (see text for details).

The temperature profile from radiosondes are also compared to those from the background and analysis. This is shown in Fig.3.16 and 3.17 representing rms errors and bias for analysis and first-guess temperature fits respectively. As expected, the biases and rms errors are lower in analysis fits. We also observe that, on average, analyses are closer to observations in mid-atmospheric levels (from 500 hPa up to 250 hPa). When looking at background temperature fits to observations (Fig.3.16), the results are better in July with lower biases and rms errors. However, in January, higher values are noted in both bias and rms particularly at above 250 hPa. This can be explained by the large differences between the background and radiosondes observations over the northern part of the model domain as explained earlier in sections 3.3 and 3.4.

Although the statistics shown here are evaluated for only two months (January and July 2011), the results obtained are comparable to those in the NARR regional reanalyses (Mesinger et al., 2006) for the same two months even if the domain considered in their verification does not include the North of Canada.

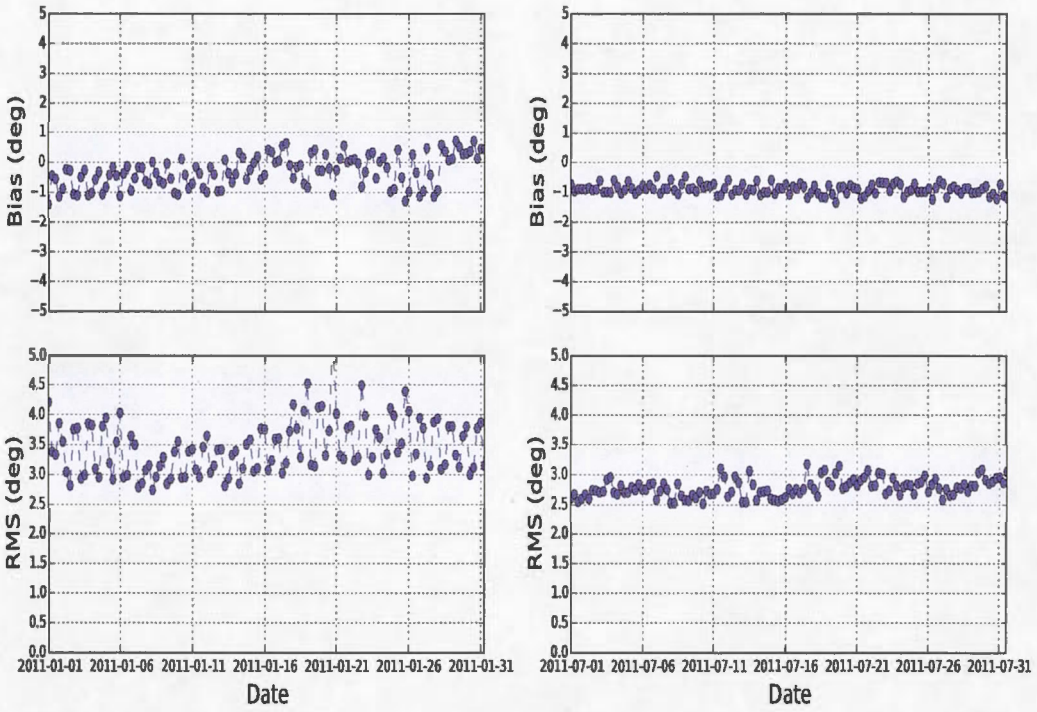

Figure 3.15 Bias (top) and RMS (bottom) for the background 2-m temperatures for January 2011 (left) and July 2011 (right).

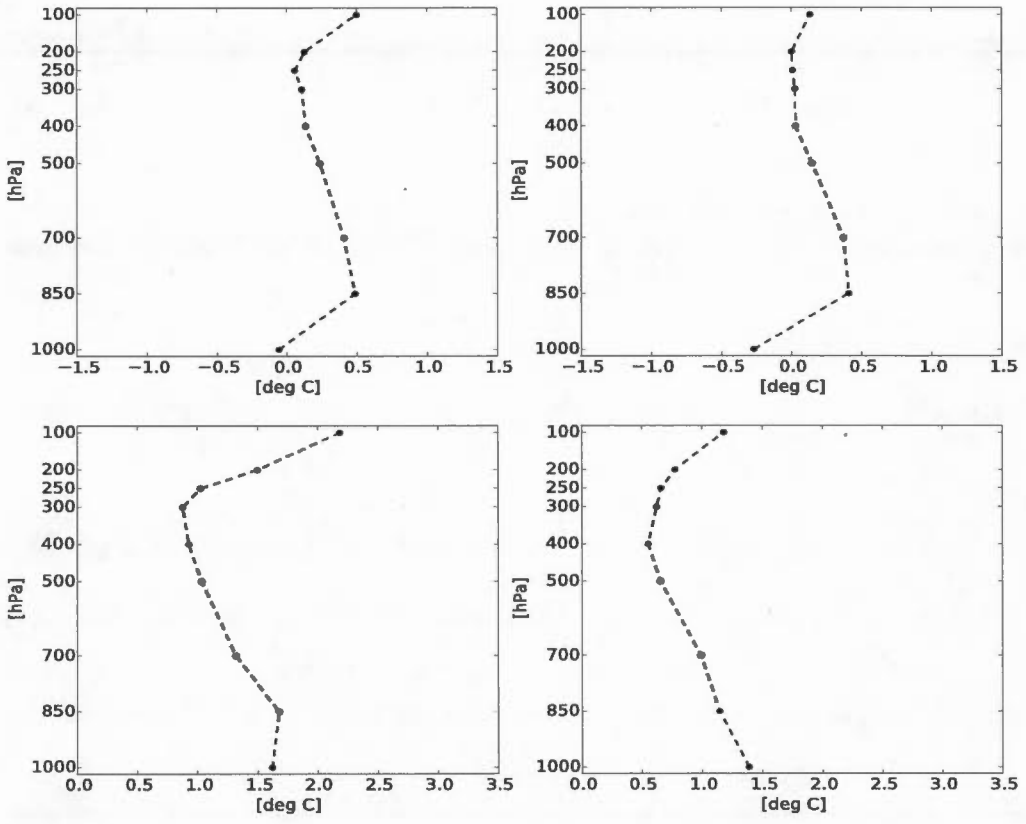

Figure 3.16 Bias (top) and RMS (bottom) for analysis temperature fits to radiosondes observations for January 2011 (left) and July 2011 (right).

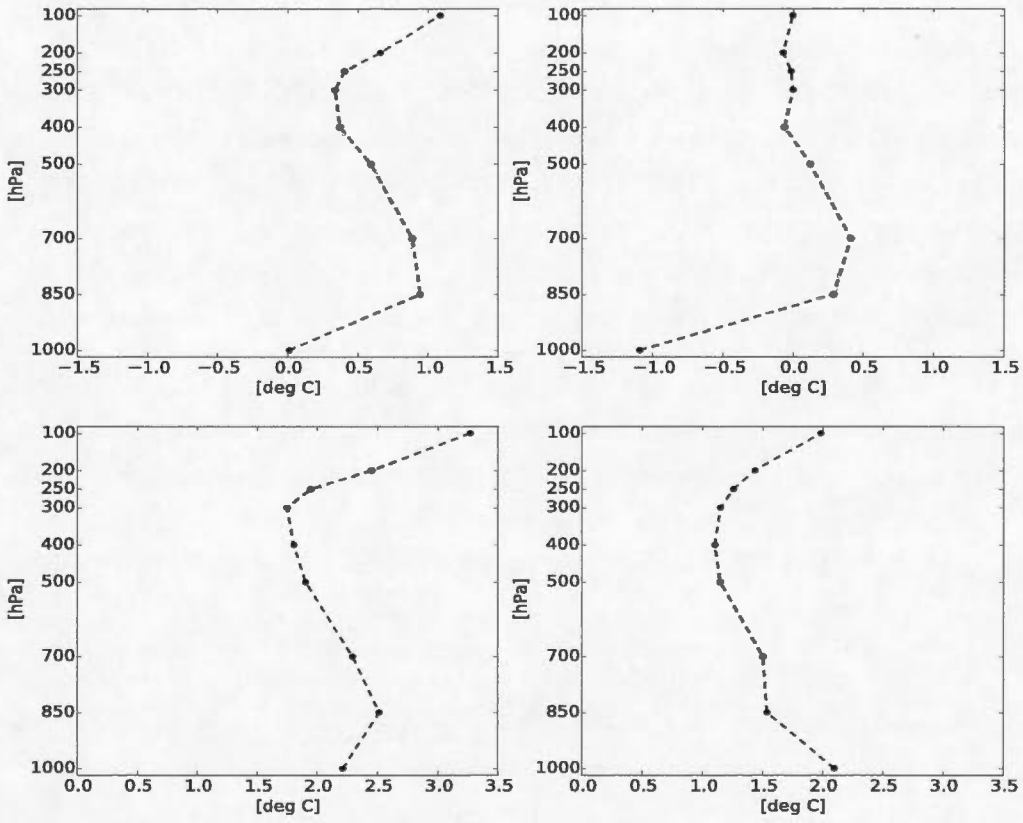

Figure 3.17 Bias (top) and RMS (bottom) for first-guess temperature fits to radiosondes observations for January 2011 (left) and July 2011 (right).

3.8 Concluding remarks

In this paper, a regional data assimilation based on the CRCM5 has been presented and evaluated. The main motivation of using CRCM5 forecasts as background states in this regional assimilation was to evaluate a regional climate model using the tools made available by data assimilation. Indeed, having an assimilation system based on the model to be evaluated is required to use the initial tendency technique of RP2007 who provides valuable information on the interactions between the different physical and dynamical processes present in the model. In addition, an assimilation system allows a direct comparison between the model forecast and observations which permits an evaluation of the observation departures of analyses and short-term forecasts that may reveal biases or unduly large variability.

In this study, a 3D-Var FGAT regional assimilation based on the CRCM5 has been implemented based on the Envar system of Environment Canada. This allowed to use a well tested system to assimilate the same observations. Several assimilation cycles have been completed over two months, January and July 2011 in a fully cycled way. In most of these cycles, the CRCM5 was driven by ERA-Interim reanalyses.

In the summer case in July, the first results showed small temperature mean analysis increments suggesting that analyses are close to the model forecast. However, in the winter case in January, the large mean negative increments at higher levels were observed in the Northern part of the domain associated with important departures from observations. The model was found to be too warm over these regions. In addition, a data monitoring of temperature observations with respect to soundings located in the Northern domain revealed small analysis residuals $[O - A]$ for the two months indicating that the assimilation system is doing well by correcting the background towards the observations. However, the innovations $[O - F]$ are unduly large in January. A January cycle using only radiosondes observations considered as reference data, identified problems in the model forecast over the problematic region ruling out that the problem could be associated with biased observations such as satellite data over land.

Using the systematic initial tendency technique of RP2007 revealed an abnormal advection activity in the northern model region. We notice here the usefulness of the initial tendency technique which allowed to identify with details the model 'failure' in resolving the advection process. This advection malfunction was shown to be related to very strong winds induced through the Northern lateral boundary due essentially to mismatch between ERA-Interim driving data and CRCM5. In an attempt to reduce this lateral driving issue, the domain was extended to the North and to the West. The results revealed no improvement in the system and the inconsistencies between the regional model and the driving data are still present.

To address this lateral driving issue, an assimilation cycle was done using the global CRCM5 version to produce global analyses. Results showed very low analysis increments over all of the CRCM5 domain and the initial tendency diagnostic applied to the global forecast over the CRCM5 domain revealed a normal advection activity. This experiment demonstrates that the model is performing well in the absence of the lateral driving constraints. In a final experiment, another cycle for January has been realised in which the CRCM5 was driven by global analyses produced in the previous global cycle using basically the same model in its global configuration. The resulting mean analysis increments are similar to the case where CRCM5 is driven by ERA-Interim reanalyses. Consequently, these combined results make it clear that the lateral driving procedure is responsible of the model difficulties occuring in January cycle.

This work has shown that a regional assimilation system can be very useful to evaluate a regional climate model by, first, enabling a rigorous use of the initial tendency diagnostic and, second, providing the possibility to compare the model to observations directly. Even though these were first experiments, the analyses in the lower part of the atmosphere seem reasonable and no significant drift was observed even over a period of two months (January-February). This is encouraging enough to envision reanalyses over a longer period. However, the results obtained also indicate that the coupling with lateral boundary conditions need to be revisited and doing it through assimilation experiments provide detailed diagnostics that would be useful to such an endeavor. This will be the object of future work.

Acknowledgements

The authors would like to thank Dr. Stéphane Laroche for his valuable comments and advice. We thank also Mr. Michel Valin, Drs. Bin He, Sylvain Heilliette and Ping Du for their assistance at implementing the assimilation system on Compute Québec's high-performance computing platform. Observation datasets used in the assimilation were provided by Environment Canada.

This research has been funded by the Grants and Contribution program of Environment Canada and a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant program. We gratefully acknowledge additional support from the Canadian Network for Regional Climate and Weather Processes (CNRCWP) funded through the NSERC's Climate Change and Atmosphere Research programme (CCAR). High-performance computing resources were provided by Compute Canada on the Guillimin platform of the Calcul Québec regional consortium.

CONCLUSION

In the first part of this work, the dynamical equilibrium of a global and limited-area model is studied in terms of its sensitivity to initial and lateral conditions. The methodology adopted to examine the dynamical equilibrium is to use the initial tendency diagnostic as proposed by Rodwell and Palmer (2007). This has been chosen because it provides detailed information and because of its relatively low computational cost. The models examined were the global GEM global model and CRCM5, its limited-area configuration. Different data sources were used as initial and lateral conditions, 3D-Var and 4D-Var analyses from the Meteorological Service of Canada (MSC) and also ERA-Interim reanalyses at low and high resolution.

For the GEM global model, it was shown that the initial dynamical balance is fairly good when initialized by MSC 3D-Var and 4D-Var analyses. However, results revealed some imbalances in the Tropics when the model is initialized by MSC 4D-Var analyses associated with a stronger convection activity especially in the ITCZ. This could be attributed to the missing convection process in the simplified physics used in the MSC 4D-Var data assimilation system (Gauthier et al., 2007). When ERA-Interim reanalyses are used at low resolution, the GEM global model exhibits large imbalances especially at lower levels. In fact, results show an abnormally intense vertical diffusion leading to excessive heating in lower levels. It is also noted that convection is nearly absent in the first moments of the integration. It was also found that analyses from ERA-Interim at low resolution are on average drier in the Tropics and, combined with the strong stable conditions in lower levels, could explain the lack of convection. This experiment also clearly showed the initial shock that occurs when the model and the initial conditions are incompatible in terms of origin or spatial resolution. When high resolution ERA-Interim reanalyses are used, these imbalances are reduced considerably and additional experimentation showed that degrading the vertical resolution is more

damaging than a degradation in the horizontal resolution. Restoring balance can take several days as observed with low resolution ERA-Interim experiment. However, results showed that the balance converges rapidly and becomes similar in the MSC 4D-Var and ERA-Interim high resolution reanalyses experiments.

In the case of the regional CRCM5 model, lateral boundary conditions are required in addition to the initial conditions. Results showed that CRCM5 is best balanced when using 4D-Var analyses as initial and lateral conditions as was the case with the global GEM model. Using ERA-Interim at low resolution leads to similar imbalances as in the global case whereas the model is better balanced with high resolution reanalyses. For longer integrations, the driving lateral data begin to influence more seriously the dynamical balance and their impact becomes more evident. The results from 15-days simulations show that the model remains unbalanced for all three cases. When considering the blending zone in the tendency diagnostic computation, the model is unbalanced without any indication of reaching a better equilibrium. On the other hand, when only the free zone is taken into account (the blending zone excluded), an improvement is noted in the first few days. However, further examination reveals imbalances reappearing after 5 days. The tendencies diagnostic provided good evidence of the sensitivity of the models' dynamical equilibrium to the initializing and driving data.

In the second part of this thesis, the impact of boundary conditions on a RCM was studied from the perspective of process tendencies. The regional climate model CRCM5 has been driven using three different data, i.e outputs from two CGCMs (CanESM2 and MPI-ESM) and also ERA-Interim reanalysis, the latter being considered as the reference driving data. Process tendencies were evaluated over a period of one season after a 14 months integration. All the simulations started from a state taken out of a multi-year climate simulation driven by ERA-Interim. This ensures a well spun up model and also a significant impact of the differences in boundary conditions. The total bias in temperature tendencies is computed by subtracting the 'observed' tendencies extracted from reanalyses from those obtained from CanESM2 and MPI-ESM driven experiments. Considering ERA-Interim reanalyses as the reference driving data, the bias due to boundary conditions is computed by subtracting ERA-interim driven simulation tendencies from those

obtained by CGCMs-driven ones.

The observed temperature tendency computed from ERA-Interim reanalyses shows a cooling in almost the whole column while the CRCM5 exhibits a warming in lower layers and a stronger cooling above for the three simulations. The structural bias reveals an excessive warming in lower layers suggesting deficiencies in CRCM5 itself. In addition, the bias due to boundary conditions reveals that MPI-ESM tends to warm the model in its lowest layers while CanESM2 causes a cooling. At higher levels, the two CGCMs are inducing a cooling with a peak at around 200 hPa. Results from specific humidity biases revealed an excessive moistening in the lower layers for the three simulations. Additional details are obtained by looking at the individual process tendencies. An example is the stronger convection in CRCM-MPI compared to CRCM-Can simulation. More humidity is also available in the case of CRCM-MPI leading to be more favourable conditions to trigger convection. It is also shown that the sea surface is warmer in the case of CRCM-MPI with more water vapour made available in the low levels through vertical diffusion. The heating noted in layers near the surface from the CRCM-MPI experiment can be attributed mainly to a strong vertical diffusion. This study showed that CRCM5 is sensitive to the imposed driving data provided by CGCMs outputs indicating that climate predictions from regional models can be impacted through the supplied driving data. It is also important to mention that the results examined come from one DJF season and it would be important to extend the analysis to more DJF seasons.

Assessing different model configurations of a RCM using physical tendencies requires the RCM to be part of the assimilation process. That is, the RCM being assessed should be the same as the one used to provide background states to the assimilation. In the third part of this thesis, a regional data assimilation based on the CRCM5 has been presented and evaluated. The assimilation system used here is based on the variational data assimilation system developed at Environment Canada implemented here as a 3D-Var FGAT regional assimilation based on the CRCM5 which is very similar to GEM-LAM model used in the operational regional data assimilation of Environment Canada. Using a well tested system also enabled to assimilate the same large volumes of observations. To evaluate this

regional assimilation system, several cycles have been performed over the months of January and July 2011 in a fully cycled way. That is, the analysis produced serves to produce the next model forecast providing the next background state to the assimilation. ERA-Interim reanalyses are used to force the model at the boundaries.

In July (the summer case), the resulting temperature mean analysis increments are small reflecting similarities between analyses and the model forecast. Different results were obtained in January, with large negative mean analysis increments observed over the Northern part of the domain. These important negative analysis increments is an indication of large departures between the model and observations and could reflect a positive bias in the model. A data monitoring on temperature observations from soundings over the Northern domain revealed small analysis residuals for the two months showing that the assimilation is correcting the model forecast towards the observations. On the other hand, the innovations are shown to be too large in January. Another January cycle using only the trusted radiosondes observations revealed problems in the forecasts over that problematic region and at the same time excluding the hypothesis that the problem could be associated with biased observations such as satellite data over land.

The systematic initial tendency diagnostic of RP07 identified an unusual advection in the North. This advection 'malfunction' was shown to be related to very strong winds induced through the Northern lateral boundary due mainly to discordance between ERA-Interim driving data and CRCM5. This was further examined by extending the model domain to the North and to the West without improvement indicating that mismatch between the regional model and the driving data still exist.

To further address this lateral driving issue, global analyses are produced based on the CRCM5 global version model. In this case, the analysis increments over the CRCM5 domain are very low. In addition, a normal advection activity is noted through the initial tendency diagnostic applied to the global forecast over the CRCM5 domain. Finally, another cycle for January has been performed in which the CRCM5 was driven by global analyses produced in the global cycle. The

resulting mean analysis increments are similar to those obtained when driving is from ERA-Interim reanalyses. Consequently, these combined results demonstrate clearly that the nesting procedure through the boundaries is responsible of the problems seen in the January cycle.

This work has shown the usefulness of evaluating a regional climate model through data assimilation. The regional assimilation system made possible a direct comparison between the model and observations and also the use of the initial tendency diagnostic providing details about the interactions between the numerous physical processes. This could certainly help to evaluate different model configurations and fine-tuning the model. These first experiments evaluated the current configuration of CRCM5 and revealed problems in the lateral coupling. On the other hand, comparison of our analyses to analyses from other sources showed acceptable results as no drift was observed when cycling over a period of two months (January-February). However, the results obtained indicate that the coupling with lateral boundary conditions needs to be improved if regional analyses with CRCM5 are to be produced.

In this research, the Canadian regional climate model (CRCM5) has been assessed from different perspectives. First, the model dynamical balance has been looked at through initial tendencies diagnostic revealing its sensitivity to initial and lateral conditions. This initial tendencies diagnostic strength has been demonstrated and its use in assessing new versions will certainly contribute to the model development. The research second part treated the lateral boundary driving impact on the model behavior. Even this aspect has been looked at in numerous studies, this research used a new approach and revealed that new information can be obtained to better understand the lateral driving issue. However, extending this study to multi-annual seasons will certainly bring more rigorous and reliable results. Another interesting extension to this study is the assessment of a model climate simulation through process tendencies. This will certainly provide some insights on the model response to a changing climate and will help better understand the climate change often expressed in temperature or precipitations variations. Finally, implementing and using a regional data assimilation system based on a regional climate model can be a powerful tool in validating different

model aspects. This regional data assimilation will certainly help improve the model through direct confrontation to observations and also through the use of initial tendencies diagnostic. Moreover, this regional data assimilation is based on the Environment Canada operational variational system and is consequently thoroughly validated. A well validated regional model together with a robust regional data assimilation system constitute a good starting point to envision production of regional reanalyses over long periods. This will be very useful to describe a coherent view of the observed climate against which the model can be compared and validated.

REFERENCES

Arakawa A, Lamb WR (1977) Computational design of the basic dynamical processes of the UCLA general circulation model. In: General circulation models of the atmosphere (A78-1066201-47). Academic Press Inc, New York, pp 173–265

Arora VK, Scinocca JF, Boer GJ, Christian JR, Denman KL, Flato GM, Kharin VV, Lee WG, Merryfield WJ (2011) Carbon emission limits required to satisfy future representative pathways of greenhouse gases. Geophys Res Lett 38: L05805

Baer F, Tribbia J (1977) On complete filtering of gravity modes through nonlinear initialization. Mon Weather Rev 105: 1536–1539.

Bélair S, Roch M, Leduc AM, Vaillancourt P, Laroche S, Mailhot J (2009) Medium range quantitative precipitation forecasts from Canada's new 33-km deterministic global operational system. Wea Forecasting 24: 690–708.

Buehner M, McTaggart-Cowan R, Beaulne A, Charette C, Garand L, Heilliette S, Lapalme E, Laroche S, Macpherson S, Morneau J, Zadra A (2015): Implementation of Deterministic Weather Forecasting Systems based on Ensemble-Variational Data Assimilation at Environment Canada. Part I: The Global System. Mon Weather Rev. doi:10.1175/MWR-D-14-00354.1, in press.

Caron JF, and Co-authors (2015) Implementation of Deterministic Weather Forecasting Systems based on Ensemble-Variational Data Assimilation at Environment Canada. Part II: The Regional System. Mon Weather Rev doi: 10.1175/MWR-D-14-00354.1

Chikhar K (2011) Évaluation de l'impact de la méthode d'assimilation utilisée sur la phase de démarrage d'un modèle atmosphériguqe. Mémoire de maîtrise en sciences de l'atmosphère, Université du Québec à Montréal, 77 pages. (available from UQAM at http://www.archipel.uqam.ca/id/eprint/4609).

Chikhar K, Gauthier P (2014) Impact of analyses on the dynamical balance of global and limited-area atmospheric models. Q J R Meteorol Soc, 140: 2535–2545. doi: 10.1002/qj.2319.

Chikhar K, Gauthier P (2015) On the effect of boundary conditions on the Canadian Regional Climate Model: use of process tendencies. Clim Dyn 45: 2515–2526. doi: 10.1007/s00382-015-2488-2.

Côté J, Gravel S, Méthot A, Patoine A, Roch M, and Staniforth A (1998) The operational CMC-MRB global environmental multiscale (GEM) model. Part I - Design Considerations and Formulation, Mon Weather Rev 126: 1373–1395.

Crétat J, Pohl B (2012) How Physical Parameterizations Can Modulate Internal Variability in a Regional Climate Model. J Atmos Sci 69: 714–724

Courtier P, Thépaut J-N, Hollingsworth A (1994) A strategy for operational implementation of 4D-Var using an incremental approach. Q J R Meteorol Soc 120: 1367–1387.

Davies HC (1976) A lateral boundary formulation for multi-level prediction models. Q J R Meteorol Soc 102: 405–418.

Dee DP, and Co-authors (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137: 553–597.

Denis D, Laprise R, Caya D (2003) Sensitivity of a regional climate model to the resolution of the lateral boundary conditions. Clim Dyn 20: 107–126

Diaconescu EP, Laprise R, Sushama L (2007) The impact of lateral boundary data errors on the simulated climate of a nested regional climate model. Clim Dyn 28: 333–350.

Fillion L, and Co-authors (2010) The Canadian regional data assimilation and forecasting system. Wea Forecasting 25: 1645–1669.

Gauthier P, Charette C, Fillion L, Koclas P, Laroche S (1999) Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. Atmos-Ocean 37: 103–156.

Gauthier P, Thépaut J-N (2001) Impact of the digital filter as a weak constraint in the preoperational 4D-Var assimilation system of Météo-France. Mon Weather Rev 129: 2089–2102.

Gauthier P, Chouinard C, Brasnett B (2003) Quality control: methodology and applications. In *Data Assimilation for the Earth System.* NATO Science Series. IV. Earth and Environmental Sciences, vol. 26, p.177-187.

Gauthier P, Tanguay M, Laroche S, Pellerin S, Morneau J (2007) Extension of 3D-Var to 4D-Var: Implementation of 4D-Var at the Meteorological Service of Canada. Mon Weather Rev 135: 2339–2354.

Gilbert J-C, Lemaréchal C (1989) Some numerical experiments with variable-storage quasi-Newton algorithms. Math Program 45: 407–435.

Giorgetta MA, and Co-authors (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, J Adv Model Earth Syst 5: 572–597, doi:10.1002/jame.20038.

Hernandez-Diaz L, Laprise R, Sushama L, Martynov A, Winger K, Dugas B (2012) Climate simulation over CORDEX Africa domain using the fifth-generation Canadian regional climate model (CRCM5). Clim Dyn doi:10.1007/s00382-012-1387-z

Houtekamer PL, Deng X, Mitchell HL, Baek SJ, Gagnon N. (2014) Higher Resolution in an Operational Ensemble Kalman Filter. Mon. Weather Rev. 142: 1143–1162, doi:10.1175/MWR-D-13-00138.1.

Kain JS, Fritsch JM (1990) A one-dimensional entraining/detraining plume model and its application in convective parameterization. J Atmos Sci 47: 2784–2802.

Kain JS, Fritsch JM (1993) Convective parameterization for mesoscale models: The Kain-Fritsch scheme. The representation of cumulus convection in numerical models. Meteor Monogr 24, Am Meteorol Soc, 165–170.

Klinker E, Sardeshmukh PD (1992) The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. J Atmos Sci 49: 608–627.

Klocke D, Rodwell MJ (2014) A comparison of two numerical weather prediction methods for diagnosing fast-physics errors in climate models. Q J R Meteorol Soc 140: 517–524

Kuo HL (1974) Further studies on the parameterization of the influence of cumulus convection on large-scale flow. J Atmos Sci 31: 1232–1240.

Laprise R (1992) The Euler equation of motion with hydrostatic pressure as independent coordinate. Mon Weather Rev 120:197–207

Laprise R, and Co-authors (2013) Climate projections over CORDEX Africa domain using the fifth-generation Canadian Regional Climate Model (CRCM5). Clim Dyn doi:10.1007/s00382-012-1651-2

Laroche S, Gauthier P, Tanguay M, Pellerin S, Morneau J (2007) Impact of the different components of 4D-Var in the global forecast system of the Meteorological Service of Canada. Mon Weather Rev 135: 2355–2364.

Li J, Barker HW, (2005) A radiation algorithm with correlated k-distribution. Part I: local thermal equilibrium. J Atmos Sci 62: 286–309

Lynch P (1997) The Dolph–Chebyshev Window: A Simple Optimal Filter. Mon Weather Rev, 125: 655–660.

Machenhauer B (1977) On the dynamics of gravity oscillations in a shallow water model with application to normal mode initialization. Contrib Atmos Phys 50: 253–271.

Mahfouf J-F, Rabier F (2000) The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. Q J R Meteorol Soc 126: 1171–1190.

Mailhot J, Benoit R (1982) A finite-element model of the atmospheric boundary layer suitable for use with numerical weather prediction models. J Atmos Sci 39: 2249–2266.

Mailhot J, Bélair S, Benoit R, Bilodeau B, Delage Y, Fillion L, Garand L, Girard C, Tremblay A (1998) Scientific Description of RPN Physics Library (Version 3.6). Recherche en Prévision Numérique, Service de l'environnement atmosphérique, Dorval (Québec).

Martynov A, Laprise R, Sushama L, Winger K, Separovic L, Dugas B (2013) Reanalysis-driven climate simulation over CORDEX North America domain using the Canadian Regional Climate Model, version 5: model performance evaluation. Clim Dyn doi:10.1007/s00382-013-1778-9

Mesinger F, DiMego G, Kalnay E, Mitchell K, et al. (2006) North American Regional Regional Reanalysis. Bull Am Meteorol Soc 87: 343–360.

Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430: 768–772.

Noilhan J, Planton S (1989) A simple parameterization of land surface processes for meteorological models. Mon Weather Rev 117: 536–549.

Rienecker M, and Co-authors (2011) MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J Clim 24: 3624–3648.

Robert A, Yakimiw E (1986) Identification and elimination of an inflow boundary computational solution in limited area model integrations. Atmos Ocean 24: 369–385.

Rodwell MJ, Palmer TN (2007) Using numerical weather prediction to assess climate models. Q J R Meteorol Soc 133: 129–146.

Rodwell MJ, Jung T (2008) Understanding the local and global impacts of model physics changes: An aerosol example. Q J R Meteorol Soc 134: 1479–1497.

Šeparović L, and Co-authors (2013) Present climate and climate change over North America as simulated by the fifth-generation Canadian Regional Climate Model (CRCM5). Clim Dyn, doi:10.1007/s00382-013-1737-5

Stainforth DA, and Co-authors (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. Nature 433: 403–406.

Stevens B, and Co-authors (2013) Atmospheric component of the MPI-M Earth System Model: ECHAM6. J Adv Model Earth Syst 5 (2):146–172

Sundqvist H (1978) A parameterization scheme for non-convective condensation including prediction of cloud water content. Q J R Meteorol Soc 104: 677–690.

Verseghy LD (2000) The Canadian land surface scheme (CLASS): its history and future. Atmos Ocean 38: 1–13

Verseghy LD (2008) The Canadian land surface scheme: technical documentation—version 3.4. Climate Research Division, Science and Technology Branch, Environment Canada.

Warner TT, Peterson RA, Treadon RE (1997) A Tutorial on Lateral Boundary Conditions as a Basic and Potentially Serious Limitation to Regional Numerical Weather Prediction. Bull Am Meteorol Soc 78: 2599–2617.

Wu W, Lynch AH, Rivers A (2005) Estimating the uncertainty in a regional climate model related to initial and lateral boundary conditions. J Clim 18: 917–933.

Yakimiw E, Robert A (1990) Validation experiments for a nested grid-point regional forecast model. Atmos-Ocean 28: 466–472.

Zadra A, Caya D, Côté J, Dugas B, Jones C, Laprise R, Winger K, Caron LP (2008) The next Canadian Regional Climate Model. Phys Can 64, 75–83.