

DISSERTATION

**IMPROVING NUMERICAL WEATHER PREDICTION: ERROR GROWTH AT
THE CONVECTIVE SCALE AND SPEED**

Submitted by

Giovanni Leoncini

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2009

COLORADO STATE UNIVERSITY

December 11, 2008

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY GIOVANNI LEONCINI ENTITLED **IMPROVING NUMERICAL WEATHER PREDICTION: ERROR GROWTH AT THE CONVECTIVE SCALE AND SPEED** BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Dr. William R. Cotton

Dr. Wayne H. Schubert

Dr. Richard E. Eykholt

Dr. Roger A. Pielke Sr., **Advisor**

Dr. Richard H. Johnson, **Department Head**

ABSTRACT OF DISSERTATION

IMPROVING NUMERICAL WEATHER PREDICTION: ERROR GROWTH AT THE CONVECTIVE SCALE AND SPEED

Despite the continuous growth of the available computational power, it is undoubtedly beneficial, for both the research and operational communities, to increase the efficiency of Numerical Weather Prediction. Because parameterizations often occupy a significant portion of the total execution time the first focus of this work is to provide a methodology to transform parameterizations into algorithms that provide the same output at a fraction of the computational cost (i.e., transfer schemes).

Several transfer schemes are developed for the Harrington radiation parameterization, in the clear sky case and implemented in the Regional Atmospheric Modeling System. The best one requires roughly 5% of the computational expense of the parent scheme. Accuracy is generally preserved and an analysis of the main meteorological fields after two days of simulations does not show significant differences. The differences for the 2 m temperature are larger than for the other fields, but still smaller than the differences introduced by a second common parameterization.

A second area where NWP is in need of improvements is convective-scale forecasting. The advantages of more accurate forecasting derive from the high societal impact of convective events, which can be severe and lead to loss of life and property.

Ensemble forecasting is an ideal tool to handle uncertainties in forecasts and the second aim of this study is to identify the processes that lead to error growth at the convective scale, for a case study over the United Kingdom using the Met Office Unified Model. The perturbation was applied to the potential temperature at a specific model level within the boundary layer, either sequentially (every 30 minutes) or at specific times. It was determined that acoustic waves are generated and can affect the background state. Vertical stability is also altered and occasionally lids can be set or removed. The unique boundary-layer scheme also contributes to error growth, by triggering different parameterizations as a response to the perturbation. Finally there are qualitative differences between high amplitude perturbations (1 K) and the smaller ones (0.01 and 0.1 K), but the root mean square error reaches similar values at saturation.

Giovanni Leoncini
Department of Atmospheric Science
Colorado State University
Fort Collins, Colorado 80523
Spring 2009

ACKNOWLEDGMENTS

First and foremost, I would like to thank Dr Roger Pielke Sr., my adviser, for the guidance and the independence I was granted during this long journey. I also am grateful my committee members, Drs. Cotton, Schubert, and Eykholt, for their advice and patience. Drs. Gabriel, Gray, and Plant deserve my gratitude for their insightful suggestions.

The Pielke Research Group has been a constant source of support and knowledge. Among those, Chris Castro, Adriana Beltran-Przekurat, and Dallas Staley deserve my gratitude more than others. Chris and Adriana for their friendship and for finding the patience and the time to answer all my RAMS related question; Dallas Staley for ensuring everyday the ideal conditions to work, not only practically. Last but not least, I would not have been able to carry out this work, without the support of my family and friends, but especially my wife, Maria, who made the burden much lighter.

Funding for this research was provided by the U.S. Department of Defense Center for Geoscience/Atmospheric Research Grant at Colorado State University (under cooperative agreement W911NF-06-2-001 and DAAD19-02-2-0005 with the Army Research Laboratory) and by the FREE program of the British National Environmental Research Council.

To my grandfather Alessandro

TABLE OF CONTENTS

Abstract
Acknowledgments

1. Introduction

2. From Model-Based Parameterizations to Lookup Tables: An EOF approach

2.1 Introduction
2.2 Methodology
2.2.a Overview
2.2.b The EOFs
2.2.c The Weighting Strategy
2.2.d Tests
2.3 Results
2.3.a Accuracy Compared with the Parent Scheme
2.3.b Computational Speed
2.4 Discussion
2.5 Conclusions

3. Error Growth at the Convective Scale

3.1 Introduction
3.2 Case Overview
3.2.a The Case
3.2.b Model and Model Set Up
3.2.c Control Run
3.3 Perturbation Strategy
3.3.a Overview
3.3.b Perturbation Structure
3.3.c Frequency of Perturbation
3.3.d Height of Perturbation
3.3.e Perturbation Experiments
3.4 Diagnostics
3.5 Results: Direct Effects
3.5.a Perturbation Effects on CAPE
3.5.b Vertical Adjustment to the Perturbation
3.5.c Boundary-Layer Changes

- 3.5.d Total Water Path and Cloud Distribution Changes
- 3.6 Results: Indirect Effects
 - 3.6.a Boundary-Layer Changes
 - 3.6.b Cloud Distribution Changes
 - 3.6.c RMSP
 - 3.6.d Intensity and Displacement Errors
- 3.7 Discussion
- 3.8 Summary and Conclusions

4. Conclusion and Future Work

CHAPTER 1

INTRODUCTION

The forecast skills of Numerical Weather Prediction models used by several weather services have been improving steadily. Part of this increase in verification scores is certainly due to the increased resolution allowed by the greater computational power now available. However, additional computational costs of model simulations are carefully weighted against the improvements they bring to the overall verification scores. Thus, the ability of running the same model at a lesser computational expense, without a decrease in accuracy, can greatly benefit NWP by freeing resources which can be devoted to improve weak features of the model itself. Parameterizations use a significant part of the overall computational cost of a model run. Majewski et al. (2002) reports that parameterizations can make use of 48% of the overall computational cost of a high resolution global model run, and the radiation parameterization alone takes up 57% of the total parameterization costs. Furthermore such values are not unusual for limited area models; tests carried out with the Regional Atmospheric Modeling System (RAMS) for this study show that the Harrington radiation scheme occupies 13% of the total computational time. Therefore in order to increase the computational efficiency of NWP is reasonable to turn the attention to parameterizations, and more specifically to radiation

parameterizations. The first aim of this work is develop a general methodology to transform parameterizations into transfer schemes. The core concept is to exploit their limited accuracy to build a fast algorithm that provides the same output at a fraction of the computational cost. Such an idea is described in Chapter 2, along with its implementation into RAMS for the Harrington radiation scheme.

It was mentioned previously that the resolution of the operational NWP model has been increasing and currently several weather centers are running their models operationally with grid spacings of $O(1\text{km})$. For example NCEP has been running ETA with 4 km grid spacing for 3 years (Wesimann et al. 2006) and similarly the Met Office with the Unified Model. Whilst the increased resolution resulted in a general increase in performance of the models, verification studies show the accurate forecast of timing and location of severe convective events as well as precipitation amounts still remains a daunting challenge. Unfortunately this happens also because even small errors (i.e., of the order of a few grid spacings) can significantly affect the impact of a storm, for example by flooding a different river catchment with adverse consequences on the efficacy of warnings and rescue efforts.

The high societal impact of severe convective events renders ensemble forecasting particularly attractive for the convective scale, since it provides a measure of the forecast uncertainty. At the same time, the availability of large computational resources has made multiple high resolution simulation possible at the regional scale. A recent study by Hohenegger and Schar (2007) underlines that most of the many perturbation methods, efficiently employed at the global and synoptic scale to build ensemble forecasting systems, are unlikely to be as successful at the convective scale. This is due to the strong nonlinearities which result in error doubling times of a few hours

instead of days. As a first step towards the design of an ensemble system the second aim of this work is to identify the physical processes that lead to error growth at the convective scale and to test a novel method to perturb the model state. This study is described in Chapter 3 and is carried out using the Unified Model of the British Met Office for a specific convective event over southern England. The case study consists of strongly upper-level forced event characterized by mesoscale topographic and surface features which resulted in widespread scattered convection and a squall line that travelled eastward during the morning hours.

CHAPTER 2

FROM MODEL-BASED PARAMETERIZATIONS TO LOOKUP TABLES: AN EOF APPROACH

2.1 Introduction

The goal of this study is to transform the Harrington radiation parameterization into a transfer scheme or lookup table, which provides essentially the same output (heating rate profile and short and longwave fluxes at the surface) at a fraction of the computational cost. The methodology put forth here does not introduce a new parameterization simply derived from the Harrington scheme, but shows that given a generic parameterization it is possible to build an algorithm, largely not based on the physics, that mimics the outcome of the parent parameterization. The core concept is to compute the Empirical Orthogonal Functions of all the input variables of the parent scheme, run the scheme on the EOFs, and express the output of a generic input sounding exploiting the input-output pairs associated to the EOFs. The weights are based on the difference between the input and EOFs water vapor mixing ratios. A detailed overview of the algorithm and the development of a few transfer schemes are also presented. Results show very good agreement ($r > 0.91$) between the different transfer schemes and the Harrington radiation parameterization with a very significant reduction in computational cost (at least 95%).

This study aims to provide an algorithm, or transfer scheme (TS) which can accurately reproduce the output of a generic parameterization, in this specific case the Harrington radiation scheme (Harrington 1997; Harrington et al. 1999), using a fraction of the computational power required by the parent parameterization. The background of this work has been laid out in Pielke et al. (2006) and is briefly summarized here.

Most Numerical Weather Prediction (NWP) models solve numerically the equations of motion, but also use parameterizations to account for subgrid-scale processes (e.g., turbulence), short- and longwave radiative flux divergence and other processes which cannot be explicitly simulated within the dynamical core which accounts for pressure gradient, Coriolis Effect, advection, and mass continuity. Land-surface interactions are also among the parameterized processes.

Because of computational constraints and limited physical knowledge, parameterizations always involve tunable coefficients (unlike the dynamical core) and are based on approximations. Furthermore most times they are strictly 1D, in the sense that they use values only from one grid column at a time.

Regardless of accuracy, parameterizations represent a considerable share of the total computational burden. Majewski et al. (2002) showed that for high resolution global NWP models, parameterizations account for 46.8% of the total computational cost while radiation parameterizations occupy 57.5% of that number. Similarly Chevallier et al. (1998) reports that the longwave radiation scheme accounts for 10 and 18% of the total computing time required, respectively, by the general circulation model at the European Center for Medium Range Weather Forecast and the climate model of the Laboratoire de Météorologie Dynamic. Tests conducted during this study show that the

Regional Atmospheric Modeling System (RAMS) has similar performances and the Harrington radiation parameterization occupies 13% of the total CPU time.

Therefore there would be great gain if a lookup table (LUT), or a Transfer Scheme (TS), could accurately reproduce the outputs of a parameterization at a fraction of the computational cost.

The main problem in developing a TS is that a simple brute force approach would require running the parent parameterization for all the possible input cases it can possibly be given during a simulation. The number of combinations of all the possible values, aside from being impractically large would grossly misrepresent reality since not every combination is physically possible. For example, we know that superadiabatic profiles of temperature generally can only occur immediately above a surface such as the ground or a cloud top. There are also larger-scale constraints such as the gradient wind balance at the synoptic scale in the mid- and high-latitudes. Moreover parameterizations are very often a limited representation of reality due to the set of simplifying assumptions. Therefore a two-stream radiation parameterization, such as the Harrington scheme, cannot be expected to reproduce the variability of the heating rates as a line-by-line code would do, let alone as nature does. On the contrary, parameterizations or “physics packages” must be thought of as the best possible compromise between the available computational resources and the accuracy desired. Because of the limited representation of reality that parameterizations offer, their outputs (i.e., physical variables) are likely to depend on and/or to be sensitive to a more limited number of physical parameters and quantities as compared to the behavior that the same physical variables have in nature. This study aims to show that it is possible to reproduce the behavior of parameterizations, not reality, at a fraction of the computational cost, using techniques and/or algorithms

that are not strictly related to the physics which parameterizations attempt to represent. The choice of the Harrington scheme is due mainly to two factors: its relatively high computational cost, and the limited number of input variables which render it more easily tractable.

Pielke et al. (2006) suggested ways to reduce the input space of a parameterization which is usually much larger than the output space, and in this study the Empirical Orthogonal Functions (EOFs) have been employed. The two main reasons for this choice are first of all its ability to identify key “patterns” among all the realistic inputs to the parent scheme, and second, the fact that every input can be expressed as a linear combination of such patterns, thereby greatly reducing the dimensionality of the problem. This is possible because the key patterns (i.e., the EOFs) are obtained through the eigenvalues and eigenvectors of the correlation matrix of the data. This property of the EOF analysis (or Principal Component Analysis) has been exploited widely for decreasing the number of estimators in statistical estimation (e.g., Davis 1976; Lorentz 1956, 1977). A second benefit of the EOF analysis is that the physical significance of the patterns themselves is ranked by the percentage of the variance of the data they explain, in a linear sense. This feature has also been exploited widely for several other applications: for example to study several aspects of the large-scale circulation such as the North American Monsoon (e.g., Castro et al. 2007), the Northern Hemisphere mean flow (Kravtsov et al. 2006), the Arctic Oscillation (Thompson and Wallace 2000), the predictability of seasonal means (Schubert et al. 2002), the relation between surface-level humidity and column-integrated water vapor (Liu et al. 1991), the impact of microphysics parameterization on a cloud property retrieval algorithm (Biggerstaff et al. 2006), and to

approximate the difference in the reflectance of the O₂ A band obtained from a multiscattering line-by-line code and a two stream representation (Natraj et al. 2005).

Furthermore, the EOF analysis has been used widely not only recently, but since Lorenz (1956) first brought it to the attention of the atmospheric science community, and its sampling errors are well known especially after the studies of North et al. (1982), Storch and Hannoschöck (1985,) and more recently Quadrelli et al. (2005). A detailed description of EOFs and their utility in the atmospheric sciences can be found in Wilks (2006).

The patterns identified by the EOF analysis are essential for this study since the core concept of the TS is to compute the EOFs of all the inputs to the Harrington scheme (HS), obtain the heating rates, and the other outputs associated with the EOFs, and then use those input-output pairs to approximate the HS outputs for every input.

While the words “*transfer scheme*” better characterize the algorithm development herein, traditionally similar algorithms, or parts of algorithms have been called “*lookup tables*”, and both nomenclatures are used interchangeably here.

The details of the algorithm are explained in Section 2.2, while Section 2.3 describes the results of the accuracy and speed tests, and finally the description and discussion of the results obtained by embedding the TS in RAMS are given in Section 2.4.

2.2 Methodology

2.2.a Overview

The EOFs of the HS input variables (surface albedo, upwelling longwave radiative flux at the surface and the vertical profiles of pressure, temperature, and water

vapor mixing ratio) are computed for a particular location and a specific time of the day under clear sky conditions. Then, the EOFs (or alternatively the regression of the data onto the principal components) are fed to the offline version of the HS, which, for each individual EOF, provides downwelling surface radiative fluxes and a vertical profile of radiative heating rates. Each EOF is associated with the day exactly at the center of the period over which they have been calculated. When input-output pairs are established, they are used to obtain a synthetic HS output for a generic input at that specific location and time of the day. This process is carried out at each model grid column, and no assumption is made concerning horizontal spatial variations. The same calculation is performed at eight different times of the day every three hours, starting at 0 UTC, and the LUT output at intermediate times is obtained by linear interpolation. Because the EOFs are calculated offline prior to the simulation, the only part of this process which must be carried out within a NWP model is the generation of the synthetic output. The TS output will be a weighted average of the output of each individual EOF. The weights are, in turn, determined by an arbitrarily defined “distance” between the EOFs and the input itself, at each grid point. This process is represented schematically in Fig. 2.1. It is also important to remember that the aim of this study is merely to achieve a reduction in speed, and no physical interpretation of the EOFs or of their significance will be attempted.

Different weighting strategies were tried offline using an independent month of RAMS output (Sept 2005) and the most successful one was then implemented into RAMS and compared against the parent scheme for a 2-day simulation during Sept 2005. A few key meteorological fields are then compared from simulations that used the parent

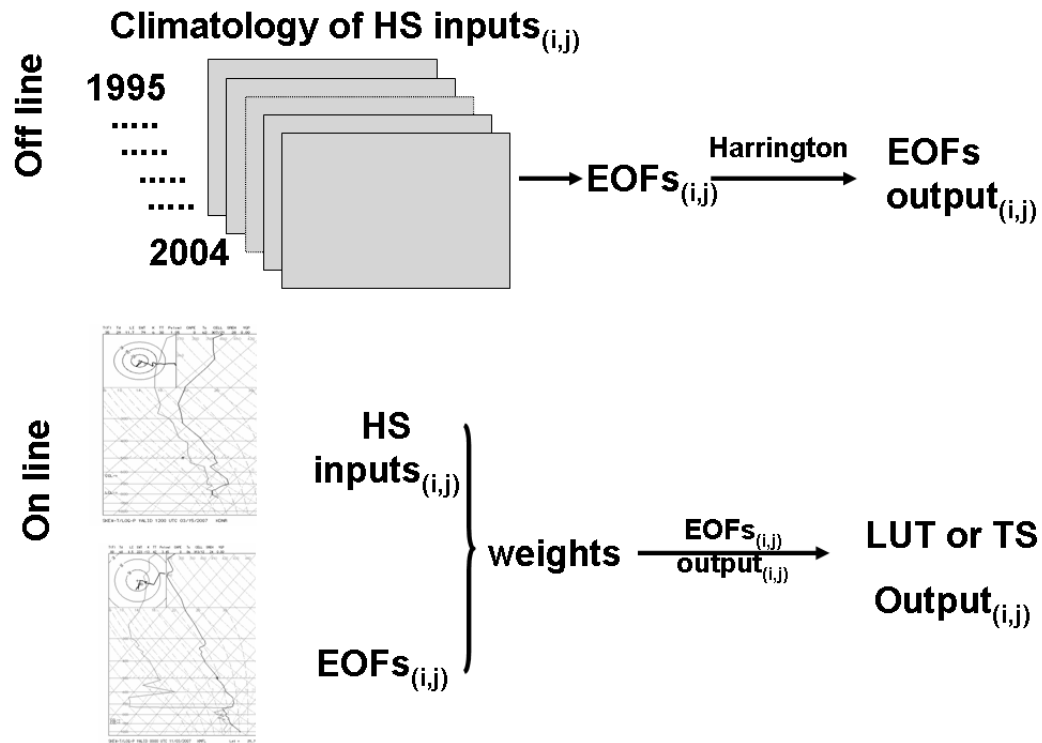


Figure 2.1 Schematic illustrating the process used to transform a parameterization in a transfer scheme. The subscripts i and j indicate grid dependent quantities.

scheme, the best TS, and the Chen-Cotton radiation parameterization (Chen and Cotton 1987).

The only difference between the version of the Harrington scheme applied on the EOF's and the version actually used in RAMS, is in how the air density is treated. The online (or in RAMS) version uses the background state density which is constant through each individual run, while the offline version computes the density from the temperature and pressure profiles which constitute part of the EOFs in order to have the thermodynamic quantities as balanced as possible. The differences between the two versions are one order of magnitude smaller than the error introduced by the LUT and thus they have been deemed negligible.

2.2.b The EOFs

Since the EOFs are meant to represent all the possible inputs which a clear sky atmosphere can provide to the HS, they were first computed using rawinsonde data for a particular station. Unfortunately the significance of the EOFs at each site varies with the different location histories and missing data periods. In order to avoid this inconvenience and to take full advantage of the temporal and spatial consistency that a NWP can offer, the EOFs were computed using the output of series of ad hoc simulations of the RAMS, which has already proved successful in this type of application as shown by Castro et al. (2007a, b).

The choices of the period of the year over which to compute the EOFs and how many years of data to use is a compromise mainly between two competing factors: on one hand the necessity of a small number of EOFs to have a faster algorithm, and on the other the necessity of having a sufficient number of EOFs to explain the variability of the input soundings. Considering also the computational constraints and the fact that with each EOF is associated a fixed day which is used to compute the zenith angle, along with hour and location, it was chosen to split the year into its 12 months. However, results from one month only are presented here, since the present study aims to show the feasibility of the LUT concept and the same implementation can be easily carried out for remaining periods. The effectiveness of the EOFs in building the TS is affected by the day of year associated to the EOFs, because if the corresponding zenith angle is very different from the zenith angle of the profiles input to the TS, the shortwave flux yielded by the TS will be unrealistic. In order to show that this is not detrimental to the TS, the month chosen for this study is September which belongs to a transition season, during

which the daily range of values of zenith angle at a grid point changes the most during the month.

As it has already been stated, this work does not aim to provide a physically meaningful interpretation of the EOFs, but it has been assumed that a more realistic set of EOFs would yield a more realistic set of associated heating rates and surfaces fluxes and therefore lead to a more accurate LUT. Pursuing this line of reasoning, several EOFs are retained, explaining up to 97.5% of the variance. Furthermore the model climatology data has not been filtered, nor averaged, and the EOFs have been rotated using the varimax method (Kaiser 1958) after scaling the eigenvectors, so that the each EOF has unity variance, but the principal components are mutually uncorrelated, and the EOFs are not orthogonal to each other, thus avoiding unphysical features introduced by this last mathematical constraints (Wilks 2006). This property of the rotated EOFs was also used in Arakawa (1993) which showed that the first few rotated EOFs of vertical profiles of heating rates are representative of typical cloud regimes.

To compute the EOFs for September, model-generated data from 10 Septembers (1995 through 2004) were used resulting in 300 temporal realizations of the sample because the EOFs are computed at eight hours of the day (from 00 to 21 UTC every three hours) using data at the same time of the day (1 sample per day, per 30 days, per 10 months), at each grid point of the RAMS domain. According to Quadrelli et al. (2005) this number is not enough to properly resolve all the eigenvectors and eigenvalues, especially for the high order EOFs. This issue, in principle, is even more important in this study, since the TS involves EOFs up to 97.5% of the total explained variance. Unfortunately the number of temporal realizations necessary to reduce the problem significantly for so many EOFs (up to 44) is too large for any practical use, and the

sample size was arbitrarily chosen. The RAMS simulations all share the same configuration which consists of one grid which covers the United States, east of the Rocky Mountains (Fig. 2.2). The radiation scheme was Harrington’s (Harrington 1997; Harrington et al. 1999), a two-stream parameterization, and was called every 20 minutes. Further details can be found in Section 2.3.

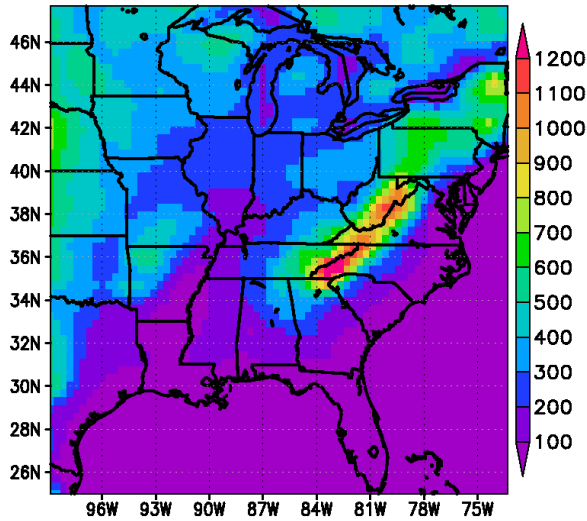


Figure 2.2: Domain and topography used for all simulations.

Although the ultimate measures of the TS value are its accuracy and computational cost, not the significance of the EOFs, they constitute the core of the TS and to investigate how it is affected by the EOFs six TSs have been implemented with different sets of EOFs (see Table 2.1). The control, “TS10RF00”, is based on rotated, unfiltered EOFs which are computed using the whole 10-year model climatology. The number of EOFs retained for each grid point, correspond to a total explained variance of 97.5%. Because the number of retained EOFs greatly affects the TS speed TS10RF00v75 and TS10RF00v50 used the same set of EOFs, but only up to 75 and 50% of the total explained variance, respectively. Time filtering the data results in fewer and more significant EOFs, so TS10RF11, TS10RF21, and TS10RF31 have been built

filtering the data with a Hamming low pass filter (Oppenheim et al. 1999) with windows of 11, 21, and 31 days. The same EOFs of TS10RF00, but non rotated, have been used for TS10NRF00, and TS9RF00 EOFs are rotated and unfiltered, as for TS10RF00, but based on a nine-year model climatology, from 1996 to 2004. The name of each TS is related to the characteristics of the EOFs: the first number (10 or 9) refers to the length of the climatology, and R (NR) stands for rotated (non- rotated) EOFs. The letter “F” and the following number indicate the filter window in days: 0 for unfiltered, 11 for eleven-day low pass filter, etc.

Table 2.1: Characteristics of the Different Transfer Schemes

TS Name	Num of years	Rotated EOFs	Filter width [day]	Variance retained [%]
TS10RF00	10	Yes	None	97.5
TS10RF11	10	Yes	11	97.5
TS10RF21	10	Yes	21	97.5
TS10RF31	10	Yes	31	97.5
TS10RF00v75	10	Yes	None	75
TS10RF00v50	10	Yes	None	50
TS10NRF00	10	No	None	97.5
TS9RF00	9	Yes	None	97.5

Finally, the EOFs and therefore the TSs are computed only for clear sky to reduce the number of variables involved in the calculations and have faster turnout times. A

model output column is defined as *clear sky* if its total cloud water mixing ratio is less than 10^{-8} kg kg⁻¹. The point is defined as *cloudy* otherwise. This value has been chosen because reducing it would not decrease the number of clear sky grid points as depicted in Fig. 2.3.

RAMS version 4.3 was used for all the numerical simulations in this study. Most of the choices made were driven by the necessity of using a standard NWP type of configuration, in order to emphasize the usefulness of the LUT for common applications without great computational costs. The basic configuration consists of one grid only of 68×67 points, with grid spacings of 40 km. The vertical grid has 33 atmospheric levels, and its spacing is stretched with a ratio of 1.12 starting from 50 m up to a maximum of 1500 m. The first level is roughly 24 m AGL while the model top is at 15800 m. Timestep is set to one minute. The North American Regional Reanalysis (Mesinger et al. 2006) was used to provide initial and lateral boundary conditions, and its total water mixing ratio, horizontal wind and potential temperature were nudged at the top model level using the Newtonian relaxation method. As mentioned earlier radiative fluxes were represented by the Harrington scheme and LEAF2 (Walko et al. 2000) parameterized the surface fluxes of heat and moisture. Diffusion was parameterized according to the anisotropic scheme of Smagorinsky (1963). The radiation parameterization (Harrington's) is called every 20 minutes. Finally the convection scheme was modified version of the Kain-Fritsch convection scheme (Kain and Fritsch 1993; Castro et al. 2002; Castro 2005). The only cloud process allowed was liquid condensation.

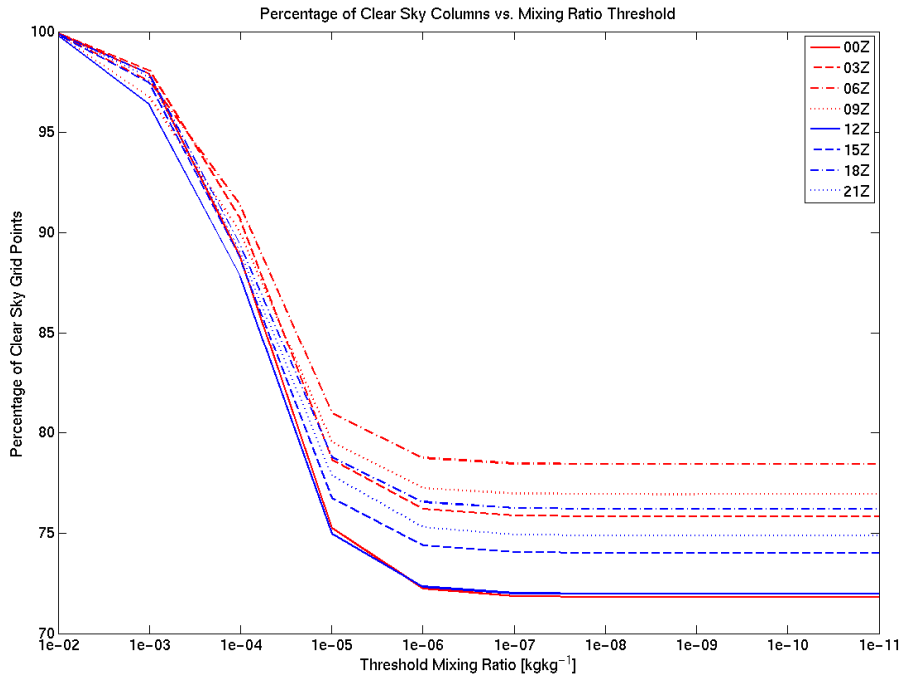


Figure 2.3: Percentage of clear sky grid points as function of the threshold mixing ratio.

2.2.c The Weighting Strategy

The output of the TS is a weighted average of the EOFs outputs, and the computation of the weights constitutes the core of the TS.

The orthogonality property of the unrotated EOFs can be exploited to compute the weights for each EOF, as it guarantees that, if the input sounding belongs to the same population that generated the EOFs, there exists one, and only one, linear combination of EOFs that is equal to the input sounding. Unfortunately despite its simplicity and elegance this method cannot effectively be applied, because not all the EOFs generated are orthogonal, even using double precision. This happens because, for the cases analyzed, over half of the EOFs (roughly 60) are effectively identical, resulting in a

nonorthogonal EOF matrix consistently with the above mentioned analysis of Quadrelli et al. (2005).

Further weighting strategies have been tested, but all were inversely related to the relative or absolute difference between the input sounding variables and the corresponding EOF variables. At first the weights were determined not only for each EOF, but also at each vertical level, but those strategies did not perform as well as the strategies that produced one weight per EOF. Most likely this is due to the implicit assumption that heating rates at different levels are independent from each other. The sum of the difference, as opposed to absolute value of the difference, also, has proven to be a poor measure of the similarity between the input sounding and the EOFs, because for a specific input sounding – EOF pair, mixing ratios larger for the sounding at a particular vertical level can be compensated by smaller values at different levels. Thus, two EOFs can have very similar weights even though one matches almost perfectly the vertical profile of water vapor mixing ratio and the other oscillates around the input sounding. The weights for the input sounding s , associated with each EOF e , which resulted in the best correlation and root mean square error (RMSE) between TS and HS outputs were determined by trial and error and are given by

$$w_s^e = \left(\sum_k |r_k^s - r_k^e| \right)^{-20}$$

where the subscript k indicates the vertical level. The high value of the exponent is necessary to cause the weights of the EOFs which are very different from the input sounding to go rapidly to zero. A smaller absolute value exponent would excessively weight the EOFs which are more distant from the input sounding, producing unrealistic heating rates and surface radiative fluxes. On the other hand, a larger absolute value of

the exponent fails to weight properly the EOFs which are closer to the input sounding, resulting again, in unrealistic values of heating rates and surface fluxes. Because of the very good code optimization at compilation, changing the exponent does not change the overall execution times. Exponentially decaying weights have also been tested, but the sensitivity of the weights to the value of the exponent is much stronger, making the tuning process more difficult and probably location dependent. The weights, which can be only zero or positive, are then normalized at each grid point (i.e., for each input profile), and are the same for all the HS output variables (heating rates, surface short- and longwave downwelling radiation fluxes).

From the tests performed using different weighting strategies a few conclusions can be drawn:

- a bulk weight per column results in improved correlations when compared against individual weights per grid point, confirming the integral character of radiative transfer.
- although all the input variables for the HS scheme (i.e., pressure, water vapor mixing ratio, and temperature profiles, upwelling longwave radiation, and albedo) have been used in the computation of the EOFs, mixing ratio alone works better than any other input variable or combination of variables, even when the combination includes the mixing ratio itself.
- if the weights computed as described are multiplied by the fractional explained variance and then normalized, the overall performance of the algorithm changes only slightly.

2.2.d Tests

To better describe the performances of the different LUTs implemented here, two kinds of accuracy tests were carried out. First, for each of the TSs, a two day simulation was carried out applying the HS at the cloudy points and computing both the TS and the HS at the clear sky points, but the TS output was driving the simulation. These simulations have been used to derive the error statistics of the LUTs compared with their parent parameterization, at the same grid points, fed the same input variables.

A second set of three simulations is then carried out, where the only difference is the radiation scheme, S1 uses HS, S2 uses Chen-Cotton, and S3 uses the best LUT and different meteorological fields (250 mb wind, 500 mb geopotential, vertical velocity at 500 mb, 2 m temperature, and 10 m wind) from the three simulations are compared. All the above mentioned simulations have the same set up used for the model climatology where the radiation scheme was called every 20 minutes and besides the different schemes for S2 and S3 the only relevant difference is the time period: from Sept 1st 0600 UTC to Sept 3rd 0600 UTC, 2005. This period was intentionally chosen to have the largest possible difference between the zenith angle of the input soundings and the EOFs one. As in the climatology case, the RAMS simulations are in agreement with the reanalysis (not shown).

2.3. Results

2.3.a Accuracy with Respect to the Parent Scheme

The determination of an acceptable error of the TS, compared with respect to its parent parameterization, depends on the use of the parameterization. For example, if the goal of the parameterization is to obtain diabatic heating rates in a numerical weather

prediction model, differences between the different approaches might not matter if the heating rate differences that matter in the prediction of weather in the model were less than about 0.1°C per hour. The ability of the different TSs to reproduce the behavior of HS is thoroughly evaluated here in this context.

The word “error” in this section, therefore, is used to refer strictly to the differences of the TSs relative to the HS, not against observation. Because of the lack of error analysis against observations for the HS, the uncertainty introduced by the use of the TS is evaluated by comparing meteorological fields obtained with the HS, the most accurate TS, and a second, widely used radiation parameterization (i.e., the Chen-Cotton scheme). A further discussion of the origin and significance of these uncertainties is given in Section 2.4.

The overall correlation coefficients, bias, root mean square error (RMSE), and error standard deviation for the heating rates, the long- and shortwave surface fluxes are presented in Table 2.2 for the eight TSs described in Section 2.2.b. All the different parameters changed little across the different TSs, indicating that the HS output is well replicated. For the heating rates the correlation coefficient (r) is always larger than 0.91, with a bias smaller than 0.009417 K h⁻¹, and a RMSE of approximately 0.024 K h⁻¹. The longwave flux has a higher correlation coefficient of at least 0.9592, with a negative bias that oscillates between -10.9737 and -12.1454 W m⁻², while the RMSE tends to be slightly larger (~14 W m⁻²). Although bias and RMSE for the shortwave fluxes are larger than for the longwave, the correlation coefficient is better for the shortwave fluxes (at least 0.9972). Relative errors (not shown) tend to be smaller and exhibit less variability over time, for the longwave than for the shortwave fluxes because the latter

Table 2.2: Correlation coefficient, bias, root mean square error, and error standard deviation for the 8 transfer schemes compared against the original Harrington scheme (see Table 2.1 for the details on the transfer schemes).

<i>Heating Rates</i>	TS10RF00	TS10RF11	TS10RF21	TS10RF31	TS10RF00v75	TS10RF00v50	TS10NRF00	TS9RF00
R	0.911906	0.911199	0.910942	0.910576	0.912823	0.911431	0.910756	0.912705
bias [K h ⁻¹]	0.009417	0.009336	0.009364	0.009372	0.009309	0.009374	0.009181	0.009310
rmse [K h ⁻¹]	0.024309	0.024337	0.024378	0.024420	0.024092	0.024268	0.024211	0.024118
std [K h ⁻¹]	0.022412	0.022475	0.022508	0.022550	0.022221	0.022385	0.022403	0.022249
<i>Longwave</i>								
<i>Flux</i>	TS10RF00	TS10RF11	TS10RF21	TS10RF31	TS10RF00v75	TS10RF00v50	TS10NRF00	TS9NRF00
R	0.9624	0.9618	0.9596	0.9596	0.9652	0.9642	0.9633	0.9645
bias [W m ⁻²]	-11.7819	-11.8744	-12.1454	-12.1235	-10.9737	-11.2017	-11.065	-11.4093
rmse [W m ⁻²]	14.6956	14.8387	15.1865	15.1596	13.8563	14.0971	14.1008	14.2383
std [W m ⁻²]	8.7835	8.8986	9.117	9.1012	8.4602	8.5585	8.7407	8.518
<i>Shortwave</i>								
<i>Flux</i>	TS10RF00	TS10RF11	TS10RF21	TS10RF31	TS10RF00v75	TS10RF00v50	TS10NRF00	TS9NRF00
r	0.9973	0.9978	0.998	0.9978	0.998	0.9972	0.9977	0.9979
bias [W m ⁻²]	-24.8747	-24.8067	-24.7471	-24.8	-24.9573	-24.8882	-24.9926	-24.7795
rmse [W m ⁻²]	45.8934	45.8114	45.7367	45.817	45.9621	45.9066	46.0698	45.7797
std [W m ⁻²]	38.5675	38.5139	38.4633	38.5248	38.596	38.5745	38.7015	38.4936

decreases (increases) in value before sunset (sunrise). TS10RF00v75 consistently gives better results, across all parameters for the three different outputs.

2.3.a.I Heating Rates

The time evolution of the heating rates correlation coefficients (Fig. 2.4a) shows a clear daily cycle for all LUTs, with a nighttime maximum and a daytime minimum. TS9RF00 and TS10NRF00 have the lowest minima during the day, confirming the value of the choices made for the EOF calculations. Also the correlation tends to be better close to the surface, although has a secondary maximum at the third to last model level and a secondary minimum at the second level from the ground (Fig. 2.4d).

The bias (Fig. 2.4b) has a similar behavior with negative daytime minima and positive nighttime maxima and virtually no spread among TSs. Moreover, during the day, and at times during the night as well, the lowest values of bias (in an absolute sense) correspond to the times when no time interpolation is done and only the EOF computed at the very same hour are used. This seems to indicate that a higher temporal frequency or a different interpolation scheme could further reduce the bias. The vertical profile of bias is consistent with the correlation analysis (Fig. 2.4e). The Root Mean Square Error (RMSE) behaves very similarly to the bias in time and space (not shown), but is roughly double in magnitude.

The average absolute error is shown in Figs. 2.4c and f, along with the 95th percentile of the distribution. The average absolute error is generally below 0.025 K h^{-1} and it tends to be larger during the night and smaller during the day. The 99th percentile of the absolute error distribution (not shown) has a more accentuated daily cycle between 0.04 and 0.11 K h^{-1} , for all the TS. Throughout the simulation all the TSs have similar

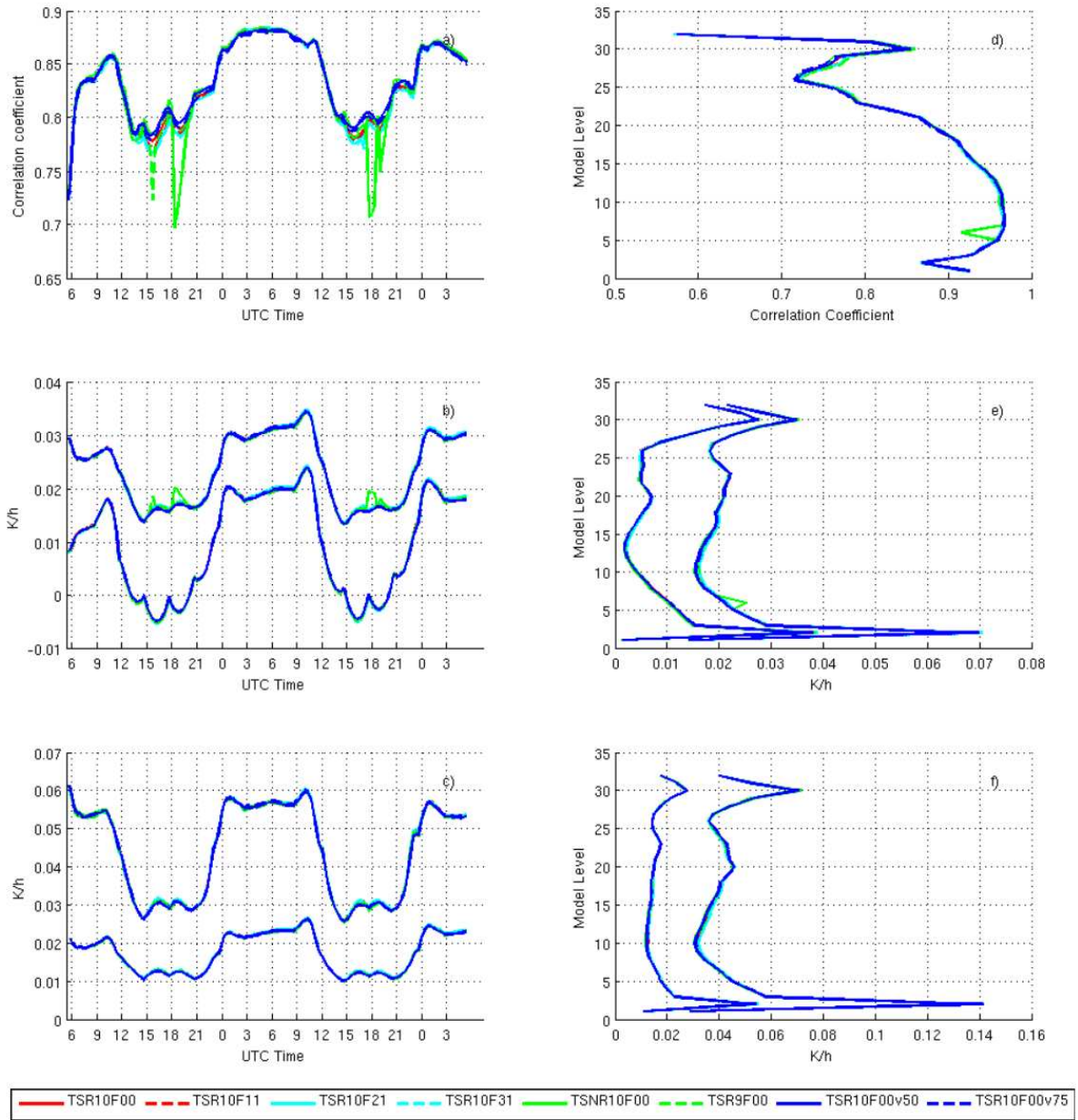


Figure 2.4: The left panels show the error statistics for the heating rates over the two days of simulation: a) correlation coefficient, b) bias (lower group of lines) and rmse (upper group), c) averaged absolute error (lower group) and 95th percentile of the absolute error distribution. The right panels show the same statistics per model level: d) Correlation coefficient, e) bias (left group of lines) and RMSE (right group), and f) averaged absolute error (left group) and 95th percentile (right group).

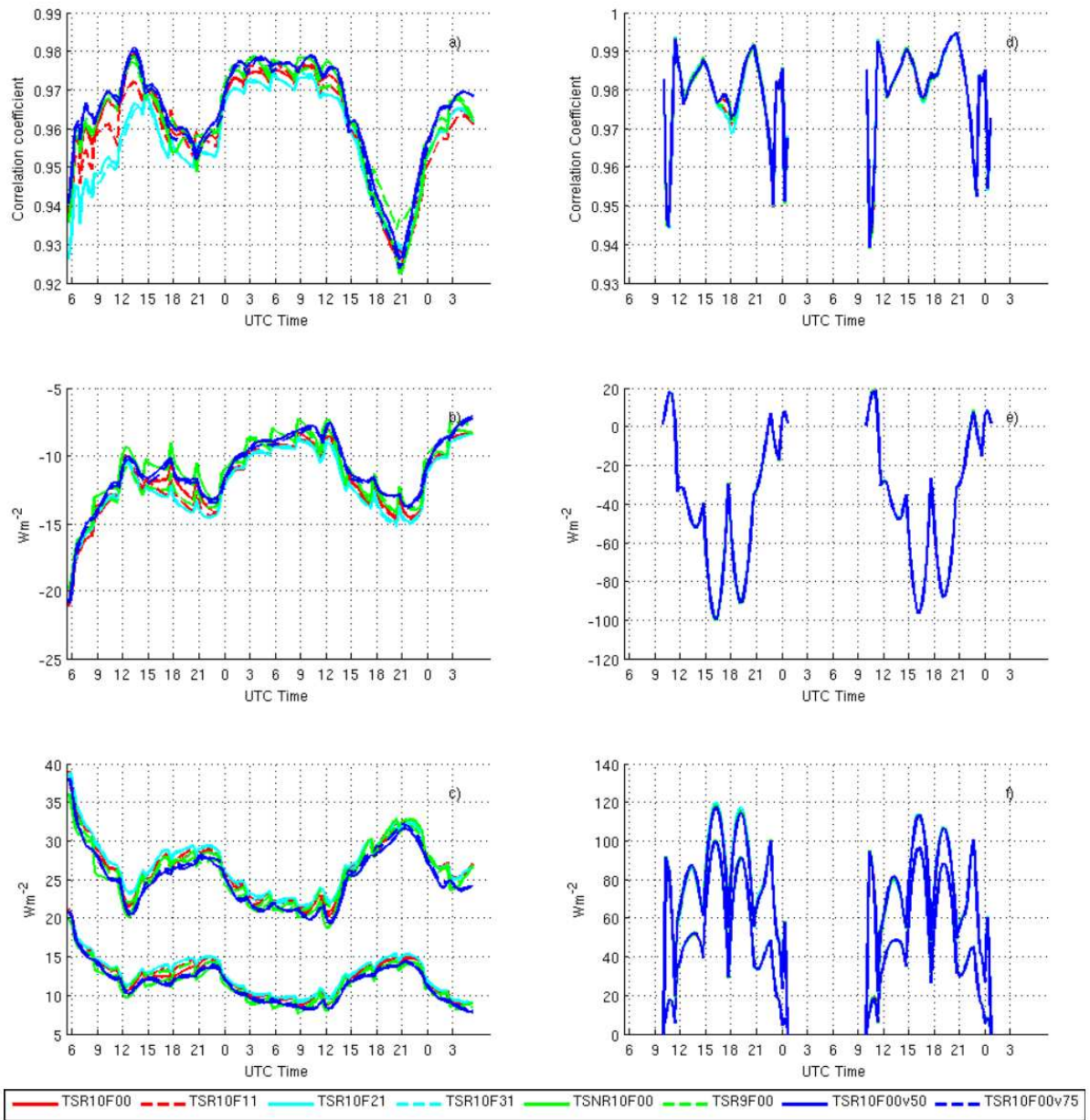


Figure 2.5: The first row of panels shows correlation coefficients, the second biases, and the third contains the averaged absolute error (lower group of lines) and rmse (upper group) for the longwave fluxes (left panels a, b, and c) and for the shortwave fluxes (right panels, d, e, and f).

absolute errors distributions such that 0.1% of the absolute errors are four times larger than the 95th percentile as for the absolute error distribution between the on line and off line versions of the parent parameterization itself. This last similarity in the error distributions leads to think that eliminating the small inconsistency between the two versions of the HS used in RAMS and the one utilized offline to obtain the heating rates for the EOFs would eliminate the largest errors, although their frequency is too small to improve significantly the other statistics.

2.3.a.II Longwave Fluxes

The minimum value of the correlation coefficient for the longwave fluxes throughout the two day experiment is very good (0.925-0.935) for all the TSs. A daily cycle is also present, as for the heating rates, but its amplitude is not nearly as constant (Fig. 2.5a). The two TSs which employ a smaller fraction of the variance (TS10RF00v50 and TS10RF00v75) tend to have higher correlation with the HS.

The bias also oscillates daily, roughly around -12 W m^{-2} , but stays always negative for every TS (Fig. 2.5b). Given the high correlation coefficient, the RMSE (not shown) is the mirror image of the bias.

Average and maximum absolute error have similar behaviors, (Fig. 2.5c) as they are relatively large at the beginning of the simulation, then they rapidly decrease to a relative minimum at 15 UTC, to increase until the afternoon of the first day. A similar cycle is repeated the second day.

Both the average absolute error and the bias show a pronounced sawtooth-like trend, with peaks corresponding to the hour of the EOF computation, although it is less evident for TS10RF00v75, especially for the averaged absolute error.

2.3.a.III Shortwave Fluxes

The time evolution of the correlation coefficient, shown in Fig. 2.5a does peak at the EOFs hours, and is lower at sunset and dawn, but its worse value is 0.94 and most of the time is above 0.98 for all TSs. There is virtually no spread among the different TSs, but again TS10RF00v75 proves better than the other TS because of its higher values around 18 UTC of September 1st.

The bias is generally negative and has relative minima between the EOFs hours (Figs. 2.5b and e). As for the longwave the RMSE is almost always equal in magnitude to that of the bias. Only before 12 UTC and after 00 UTC they both are positive.

The mean and the 95th percentile of the absolute error behave similarly to the bias (Fig. 2.5f).

2.3.a.IV Meteorological Fields

To show the overall effect of the new algorithm on some common meteorological fields the outputs of the simulations utilized for the above tests are shown in Figs. 2.6 through 2.10. All the fields are taken at the latest time of the simulations, two days, to ensure the maximum divergence between the three different set ups of the radiation scheme: the Harrington scheme, the best LUT (TSR10F00v75), and the Chen-Cotton scheme (herein CCS). The white areas indicate cloudy points at least for either of the Harrington or LUT simulations.

The differences are generally small, and have different characteristics. The 500 mb vertical velocity (Fig. 2.6) and the 10 m wind speed (Fig. 2.7) have average absolute errors which are about half of the corresponding errors of the CCS, while the maxima are more similar (see Table 3), indicating that the LUT tends to have small absolute errors,

with a few points with higher errors (i.e., the distribution has a very long tail). More precisely, for the two mentioned error distributions the difference between the 99th percentile and the maximum error is at least twofold the difference between the 99th and 50th percentile (not shown). The 2 m temperature (Fig. 2.8) has lower, but more widespread absolute errors: the average is 0.40 °C (Table 2.3), and the difference between the maximum and the 99th percentile is about two thirds of the 99th – 50th percentile difference (not shown). Furthermore the LUT error ranges from –2.77 to 1.89°C (Fig. 2.8 - top right panel), while the difference between the HS and the CCS goes from –8.02 to 0.86°C (Fig. 2.8 - lower right). The 500 mb geopotential height and the 250 mb wind speed have error distributions similar to those of the 10 m wind, but the maximum absolute errors, 1.19 m and 0.29 m s⁻¹, are negligible compared with the actual values of the two fields, and the top and middle left panels of Figs. 2.9 and 2.10 do not show any visible difference.

2.3.b Computational Speed

Computational speed has been tested profiling the four basic model configurations used thus far: RAMS standard version with HS, standard version with the CCS, RAMS with TSR10F00, the control TS, and TSR10F00v75, the most accurate TS. The profiling data consist of the number of times each subroutine, within RAMS, has been called, the total time has been spent on each subroutine (cost), and the time spent on each subroutine excluding the time due to call to or from other functions and subroutines (self time). This allows a detailed comparison of the execution times accounting not only for the subroutines which actually produce the desired outputs, but also for all those subroutines which preprocess data or carry out computations once per domain, instead of once per

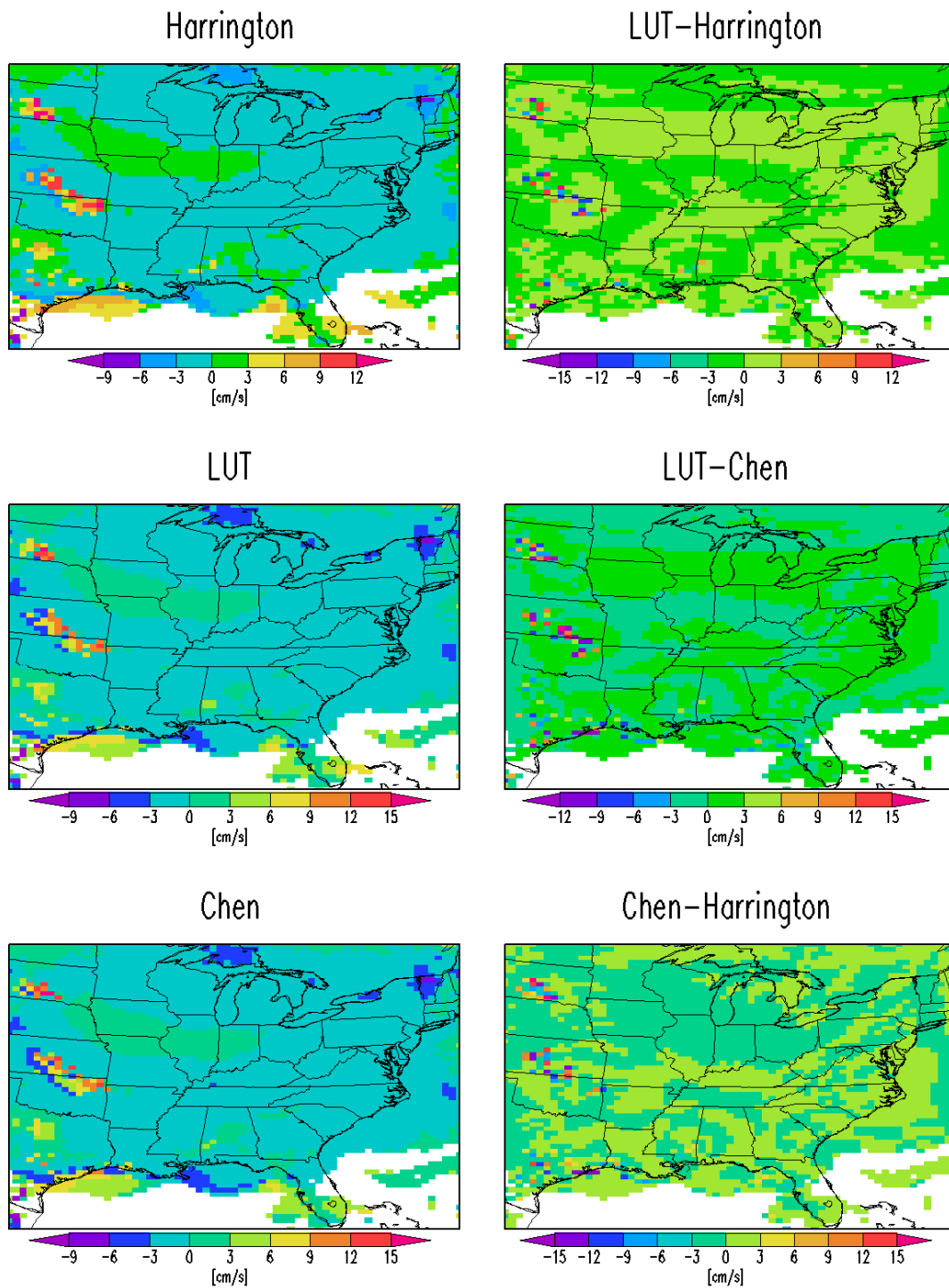


Figure 2.6: 500 mb vertical velocity [cm s^{-1}] at the end of the two day test simulation: top panels (left to right) correspond to runs with the Harrington scheme, Chen-Cotton, and the best LUT (TSR10F00v75). Lower panels show the difference between the best LUT and Harrington, the difference between the best LUT and Chen-Cotton, and finally the difference between the Chen-Cotton and Harrington scheme.

Table 2.3: Mean and maximum absolute error for the best transfer scheme and the Chen-Cotton parameterization, on Sept 3rd at 0600 UTC after 2 days of simulation.

	Mean Absolute Error LUT	Mean Absolute Error Chen-Cotton	Max Absolute Error LUT	Max Absolute Error Chen- Cotton
2 m Temperature [°C]	0.40	1.65	2.76	8.02
10 m Wind Speed [m s ⁻¹]	0.057	0.125	2.153	1.197
500 mb Geopotential Height [m]	0.084	0.161	1.19	1.22
500 mb Vertical Velocity [cm s ⁻¹]	0.29	0.36	16.5	16.5
250 mb Wind Speed [m s ⁻¹]	0.0134	0.0173	0.290	0.348

grid point. It was determined that the only non negligible operation, besides the schemes was the loading of the EOF data, necessary to the TSs. Every other preprocessing part of the computations, for all scheme are negligible compared to the core of the corresponding scheme. Table 2.4 shows the ratio of time spent on the different schemes to the time spent on the HS. For the LUTs a range is given because their execution time is comparable to the accuracy of the time measure. The execution time ratio for the CCS is 0.544 which means it is about twice as fast as the HS, but for the control TS the same ratio is 0.0732, implying a 93% reduction in the execution time. TSR10F00v75 performs even better and the reduction is 96%. The amount of additional data that the TSs require

to read at initialization depend on the number of grid points, and with this configuration is sizeable, 388 Mb for the control TS and 124 Mb for TSR10F00v75, but because of the great speed gain, assuming 70% of the domain is cloud free it takes 40 to 80 radiation calls to offset the initial overhead for the control TS (i.e., TSR10F00), and only 6 to 11 for TSR10F00v75. This number varies not only because of the uncertainty of the time measure of the TS, but also because it involves reading from a hard disk, whose timing depends on the size of data, if the disk is mounted over the network, and how many processes are accessing it.

The profiling data discussed above are accurate, but are also affected by the profiler itself which intrudes the original code in order to measure timing. In this specific case calling the LUT alternatively to the HS caused the parent subroutine to increase its self time even though the LUT was turned on from a configuration file, and the actual executable was not changed. Therefore the overall duration of the simulations was also measured without any profiler on a dedicated computer to minimize the overhead due profiling and the I/O times. The relative times are shown in the fourth column of Table 2.4, and are all inclusive. TSR10F00 and TSR10F00v75 result in a 10 and 11% reduction of the overall duration of the simulations, respectively. The use of CCS yields a similar reduction of 9% despite the longer time. The small difference between the CCS and the TS is due the additional overhead due to I/O and to the fact that the TS run only on 70 to 80% of the domain, and the HS takes care of the remaining 30 to 20%. Most likely also an increased self time of the parent subroutine still occurs, and can be eliminated when a “cloudy sky” version of the LUT will be ready and completely replace the HS.

Table 2.4: Time gains of the LUT over the HS. The second column is evaluated considering cloud free 70% of the domain.

Radiation Scheme	CPU Time	Num of calls to offset I/O	Simulations
	$\frac{\textit{Scheme}}{\textit{Harrington}}$		duration $\frac{\textit{Scheme}}{\textit{Harrington}}$
Chen Cotton	0.544	--	0.912
TSR10F00	0.0512 – 0.0732	40 – 80	0.907
TSR10F00v75	0.0129 – 0.0375	5.8 – 11	0.893

2.4. Discussion

Despite the very high correlation coefficients, the differences between the TSs and the HS for the surface fluxes are fairly large when compared to instrumental errors which are of the order of 10 W m^{-2} (Josey et al. 2003, Dong et al. 2006). Comparison with line-by-line codes often yield even better absolute errors (Fu and Liou 1992; Zhang et al. 2003), but fortunately for the longwave fluxes the absolute errors of the TSs constitute only a small fraction of the corresponding values of the HS: the averaged relative error over the two day test oscillates around 4%, and the 95th percentile is roughly 8% (not shown). The same relative errors are somewhat larger for the shortwave flux, as between 15 and 21 UTC the average relative error is at the most 12% and the 95th percentile peaks 15%. From dawn to 15 UTC and from 21 UTC to sunset the absolute errors are much lower than from 15 to 21 UTC (Fig. 2.5f), but the relative error is larger. The errors around dawn and sunset can be corrected by requiring that the interpolation of the EOFs at any point in time, for the shortwave fluxes happen only when both of the

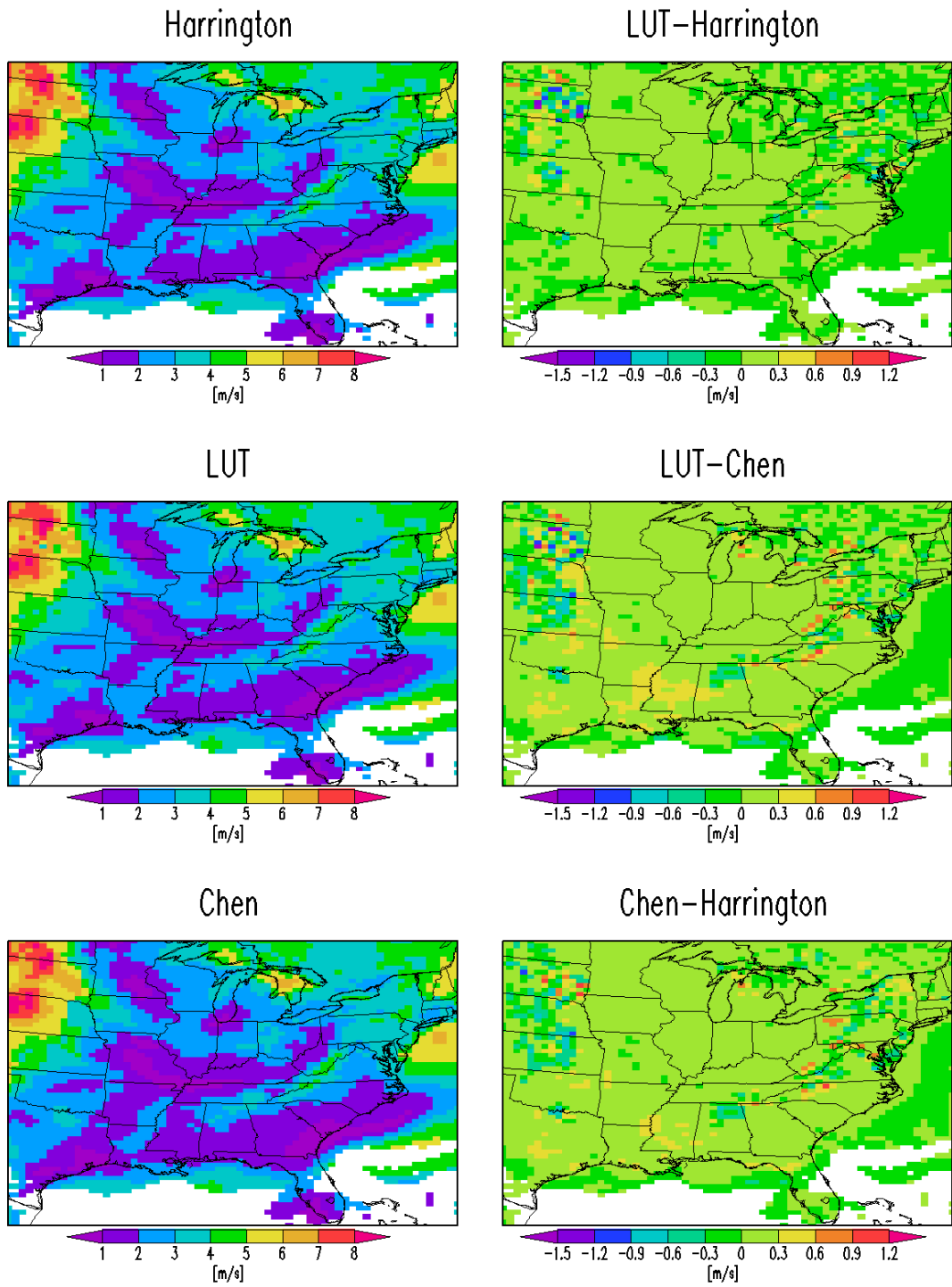


Figure 2.7 As Fig. 6 for the 10 m wind speed in $\text{m}^{\text{s}^{-1}}$

preceding and following EOF have a zenith angle large enough to trigger the computation of the Harrington scheme, which would be used otherwise. This approach has been attempted. While it slightly decreases the error around the hours of dawn and sunset it had little benefit for the overall accuracy performance of the algorithm, and thus, it has not been adopted.

It is noteworthy that all the errors of the shortwave fluxes have their relative maxima always at equal distance in time from the hours at which the EOFs were computed. This happens because the sine-like function that describes the daily behavior of the shortwave fluxes is not very well approximated by the linear interpolation of the relative low frequency (3 h) values obtained from the EOFs. This results in an underestimation of the shortwave fluxes, especially for the daily maximum. Most likely this behavior can be corrected by increasing the computation frequency of the EOFs at least during the daytime. Also a different weighting strategy can provide some benefit: by use of more EOFs, further away in time, a polynomial interpolation that keeps into account the current time of the day can ensure that the daily maximum of shortwave fluxes is not underestimated or is underestimated to a lesser extent.

Because of the poor knowledge of the clear sky errors of the HS, the accuracy of the TSs to reproduce the parent parameterization is compared against the parameterization of Fu and Liou (1992), Gabriel et al. (2000), and Zhang et al. (2003).

The above radiative transfer schemes predict heat fluxes which are later converted to heating rates via vertical divergence, so that the comparison between the heating rates would be affected by the different vertical grid spacing used for the tests. This is particularly important close to the surface where the vertical grid spacing used in this study (50 to 1500 m) is certainly smaller than the one regularly used for this kind of test

(700 to over 3000 m above the tropopause). Both Gabriel et al. (2000) and Zhang et al. (2003) radiative transfer schemes have errors of the order of 1 W m^{-2} , which at the first RAMS model level, translates into a 0.056 K h^{-1} error on the heating rate and compares well against the averaged absolute error of the TS (Fig. 2.4c and f). At higher altitudes where resolution is coarser and density smaller the same accuracy for the fluxes results in smaller heating rate errors, by about an order of magnitude. Here the TSs have larger errors (Fig. 2.4e and f). Fu and Liou (1992) used a constant vertical grid spacing of 1 km, and similar back of the envelope calculations using their errors on the heating rate are consistent with the above comparisons.

While the above error analysis provides a very good description of the strengths and weakness of the TSs, the more definitive proof that the uncertainties due to the imperfect reproduction of the HS output by the TS are acceptable lays in the effect on the meteorological fields. Figures 2.6 through 2.10 show that, at least after two days of simulation, the main meteorological fields are not significantly different from the simulations with the parent scheme, even at the higher altitudes, where the heating rates errors are relatively larger. The surface temperature differs more than other fields, but it is not unphysical and most of all, its variations from the HS are of the same order of magnitude of another common parameterization (Chen-Cotton), thus strengthening the suitability of the TS for climate simulations.

The computational gain over the HS is 95%, and makes the tradeoff error versus accuracy well worthy. The success of the TS is likely due to the clear sky condition, which implies weak multiple scattering, and low optical thickness. In this case, the up- and downwelling fluxes decouple and the multiple scattering problem reduces to Beer's law. The great computational gains are due to the fact that the use of HS output obtained

from the EOFs bypasses the on line calculation of the absorption coefficients whose computation is indeed the bottleneck of two-stream parameterizations such as the HS. Similarly, the selection rule of Gabriel et al. (2001) also exploits this simpler condition to reduce the number of radiative transfer calculations and thus the computational expense. Although this seems to indicate that a TS is likely to work less well when multiple scattering is significant, the study of Natraj et al. (2005), show the authors were able to accurately simulate the residual of the reflectance in the O₂ A band whose absorption coefficients were by a multiscattering line-by-line code. The number of lines used was selected through an EOF analysis with subsequent radiance calculations initialized by a two-stream model.

2.5. Conclusions

In this study a methodology to develop a LUT or TS from a parameterization has been presented. It is important to further clarify that this work does not to introduce a new parameterization, simply derived from a preexisting one, but reduces a parameterization to a TS, whose core concept is to compute the EOFs of the parent scheme input variables, under clear sky conditions, and run it on the EOFs. Then the TS output of a generic input is a weighted average of the EOFs output, where the weights are based on a form of distance between the input and each individual EOF. Several TS have been develop for the Harrington radiation scheme under clear sky conditions, by using different EOFs and their errors have been thoroughly analyzed, as well as their computational speed. The errors with respect to the parent parameterization, at times can be larger than what is commonly accepted as error for a radiation parameterization

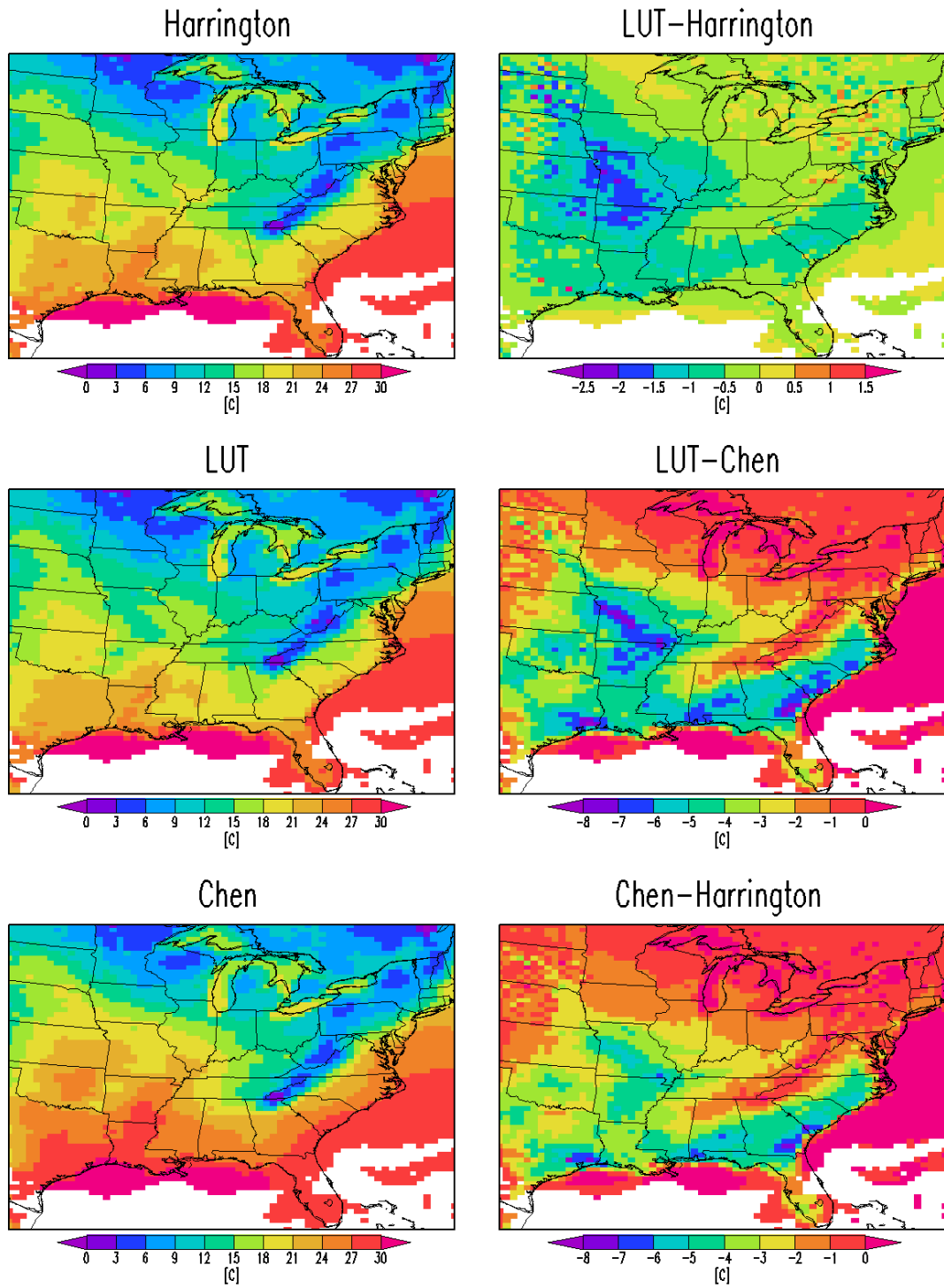


Figure 2.8: As Fig. 2.6 for 2 m temperature expressed in °C.

compared against a line-by-line code, but this kind of analysis has not been published for the HS, as well as for other mesoscale schemes, at least for the clear sky case. Therefore it is not possible to know with certainty the error introduced with the TS and it is suggested that a different weighting strategy is very likely to improve the shortwave flux errors. Furthermore once the best TS has been implemented into RAMS the meteorological fields after a two day simulation show a good agreement with the parent scheme and a comparison against the meteorological fields obtained by use of the Chen-Cotton scheme indicates that the uncertainties introduced by the TS, as compared with the HS are less significant than the ones due to the second scheme. Finally the calculations necessary for the TS are carried out at a fraction of the original cost.

While this study is limited to the Harrington radiation parameterization, it is reasonable to believe that the same methodology can be extended to a cloudy sky and applied to other parameterizations with similar results as first envisioned in Pielke et al. (2006).

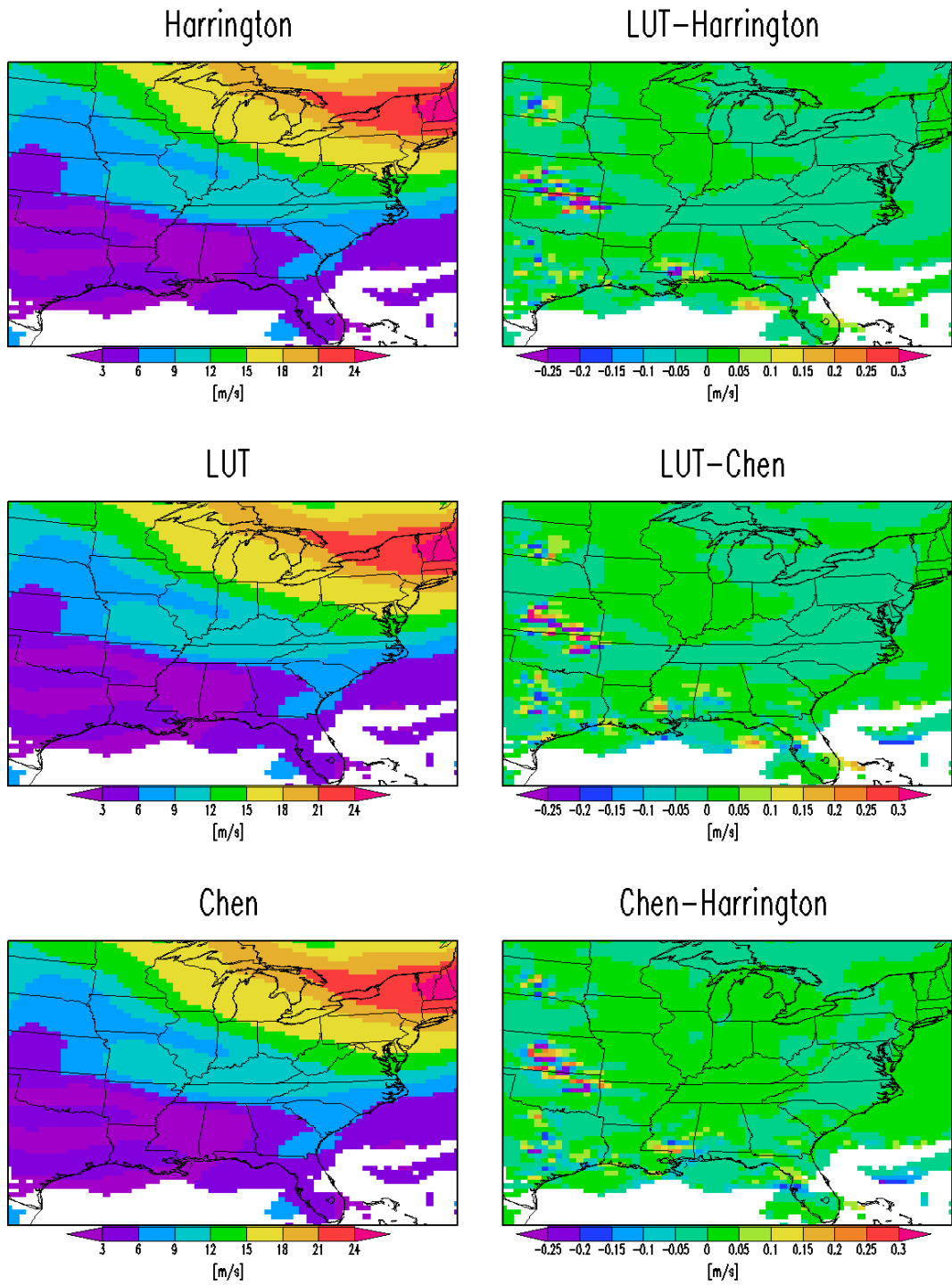


Figure 2.9: As Fig. 2.6 for the 500 mb geopotential height in m.

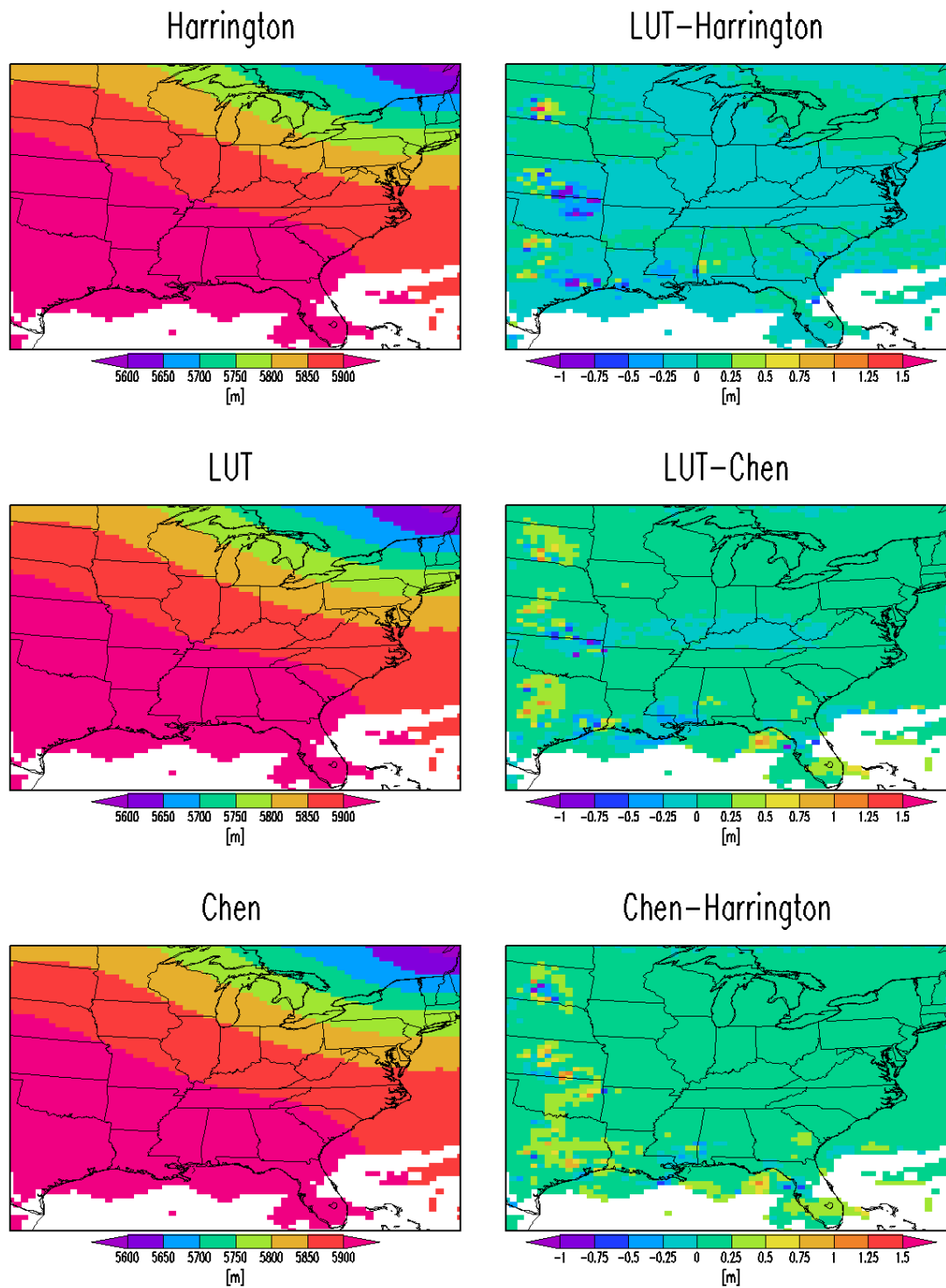


Figure 2.10: As Fig. 2.6 for the 250 mb wind speed in m s^{-1} .

CHAPTER 3

ERROR GROWTH AT THE CONVECTIVE SCALE

3.1 Introduction

Extreme rainfall from convective events is the leading causes of floods and flash floods over the summer months in the UK (Hand et al. 2004). Because of the high societal impact of the associated floods, accurately forecasting extreme convective events could benefit greatly the flood forecasting and specifically flash flood forecasting, as highlighted in Collier (2006).

Increased computational power has recently made numerical weather prediction possible over large domains, with grid spacings of the order of a few kilometers that resolve convection at least partially. While such grid spacings are not sufficient to resolve the individual convective elements (e.g., Bryan et al. 2003) they are generally able to describe convective phenomena more realistically than 10 or 12 km grid spacings (e.g., Weisman et al. 2008). For example the Met Office runs operationally at 4 km grid spacing over the entire United Kingdom with plans of upgrading to 1 km grid spacing on a similar domain by 2011. The National Center for Environmental Prediction has been running the WRF-ARW model at 4 km since 2003 (e.g., Weisman et al. 2008). However, in convection permitting simulations, with grid spacings of $O(1\text{km})$, the predictability of the atmosphere is remarkably different from that of the synoptic scale: the error growth rates are about 10 times larger and the tangent linear approximation breaks down very quickly, within hours instead of days as for the synoptic scale (Hohenegger and Schar 2007a). Such low predictability is most likely due to the high nonlinearities of the

atmosphere at such small scales. Microphysics, turbulence, radiation and flow dynamics are strongly coupled and amplify both model and observation uncertainties. This makes ensemble prediction systems particularly valuable because they provide a measure of confidence in the forecast, but at the same time it renders the large-scale methodologies for perturbation generation less likely to be effective (Hohenegger and Schar 2007a).

Despite these difficulties, the research into ensemble prediction systems at the convective scale is a developing field and Kong et al. (2006, 2007) represent the first attempt to design an ensemble prediction system for a full physics numerical model using operational initial conditions. More specifically they tested different methodologies over three nested grids with 24, 6, and 3 km grid spacing, applying the scaled-lagged average forecasting technique (Ebisuzaki and Kalnay 1991) to a tornadic storm. They found that the associated perturbations grew too slowly and produced little spread. However, the spread improves significantly when the perturbations are scaled by their amplitude, and not by the age of the forecast that generated them. Moreover they also point out that the 3 km ensemble members reflectivity patterns have greater spatial fidelity than the coarser grid spacing forecasts, but the conventional skills score (root mean square error, Brier score, etc.) do not always reflect such improvements. The use of radar data and data assimilation procedure were also tested, indicating that they certainly yield better forecasts, but the ensemble mean performed better than the deterministic control run even when its member did not assimilated any data.

Others (e.g., Zhang et al. 2003; Walser et al. 2004; Hohenegger et al. 2008a,b) have shown that ensembles of convection permitting simulations (with grid spacings of $O(1\text{km})$) can be used to investigate the predictability of specific events, perturbing the initial conditions or varying the lateral boundary ones. From all the above mentioned

studies which directly or indirectly addressed the feasibility of ensemble prediction systems at the convective scales, emerges that: a) moist convection and nonlinearities in general strongly favor rapid error growth with typical time scales of the order of the hour; b) the presence of moist convection alone does not necessarily imply low predictability because of a strong dependence upon the weather regime; c) lateral boundary conditions are also an important factor; and d) model uncertainties also affect predictability. Furthermore Hohenegger and Schar (2007b) determined that the fast and domain-wide error growth they observe in their simulations, is due to the propagation of small amplitude, fast acoustic waves and/or numerical noise, as well as large amplitude, slower gravity waves. Both phenomena associated with moist convective instability and the triggering potential of topography, contribute to error growth.

The goal of this study is twofold. First, to investigate the use of a novel technique to perturb the model state, in order to account for model uncertainties. Secondly, to determine the cause, or the causes, of the error growth at the convective scale for a convective event over the United Kingdom.

This chapter is structured as follows. Section 3.2 outlines the main features of the convective event, the model used and the control run. Section 3.3 describes the perturbation and the perturbation strategy in details, whilst Section 3.4 contains a description of the diagnostics used. The results are presented in Section 3.5 and discussed in Sections 3.6 and 3.7. Summary and conclusions are in Section 3.8.

3.2 Case Overview

3.2.a The Case

The Convective Storm Initiation Project (CSIP, Browning et al. 2007) was carried out during the period June-August 2005 (with a pilot study in the preceding year). The objective was to improve understanding of the mechanisms determining precisely when and where deep convective clouds initiate with an observational focus on Southern England. An overview of all 18 Intensive Observing Periods (IOPs) can be found in Browning and Morcrette (2006). IOP 18 was chosen for this study because the convection was primarily forced by a large-scale upper-level trough (suggesting predictability in the synoptic-scale forecast) but the evolution of the intense storms that formed and secondary convective initiation were driven by internal dynamics arising from cold downdraughts (suggesting that the details of the convective evolution will be sensitive to model perturbations).

The Met Office synoptic surface analysis at 1200 UTC for that day (August 25 2005) is shown in Fig. 3.1. The main features were well forecast (Clark and Lean, 2006) and are the cold front over the western edge of the European continent and the centre of the associated low pressure system to the north of the British Isles yielding westerly flow over the UK. Southern England lay below a tropopause fold running roughly along the southern coast of England. This led to widespread scattered convection not only over land but also over the surrounding seas. A squall line developed from a line of showers at 1015 UTC and formed a distinct arc by 1130 UTC with precursor cells formed at about 0815 UTC near the Bristol Channel. The squall line travelled east southeast to reach the East Coast of southern England at about 1400 UTC. This case was also characterized by secondary initiation, often linked to the numerous cold pools and gravity wave

propagation. A more comprehensive description of the synoptic and mesoscale observations can be found in Browning and Morcrette (2006) and in Clark et al. (submitted) which also includes a detailed analysis of the squall line.

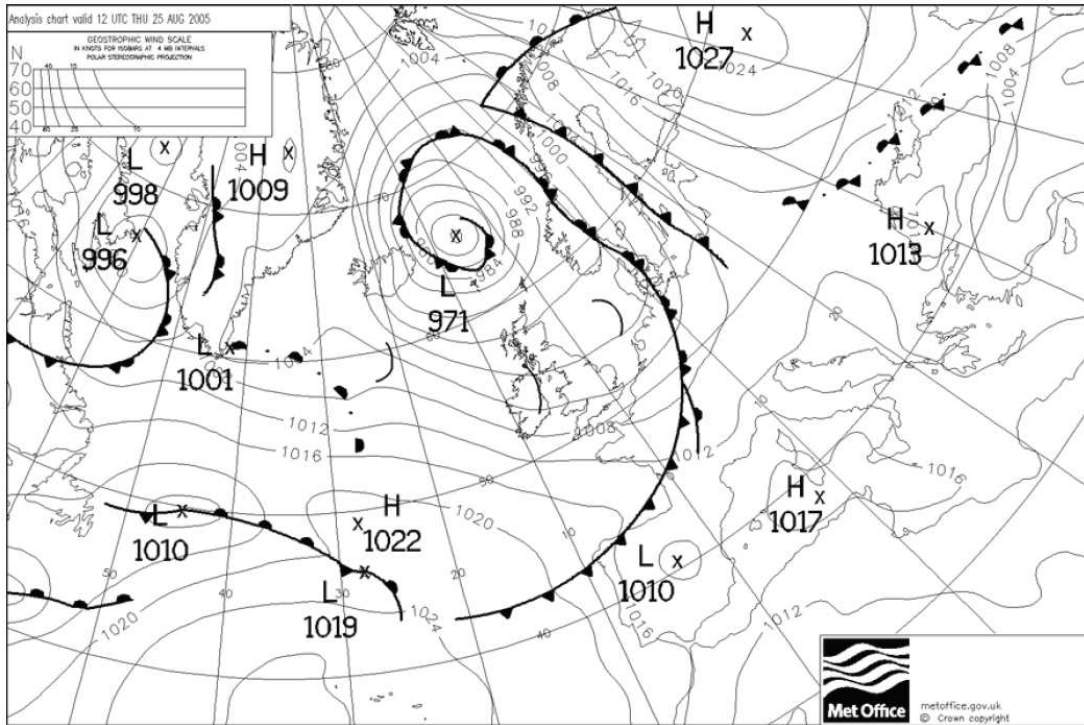


Figure 3.1: Met Office surface synoptic analysis for August 25, 2006.

3.2.b Model and Model Set Up

Version 6.1 of the Met Office Unified Model (UM) was used in this study. This model solves non-hydrostatic, deep-atmosphere dynamics using a semi-implicit, semi-Lagrangian numerical scheme (Cullen et al. 1997, Davies et al. 2005). The horizontal grid is rotated in latitude/longitude with Arakawa C staggering. The vertical coordinate system is terrain-following with a hybrid-height vertical coordinate and Charney-Phillips staggering. In this study, the model is run with 38 vertical levels and a horizontal grid

spacing of 4 km one-way nested within a domain with 12 km grid spacing. The model is currently run operationally at these resolutions and the higher resolution version was also quasi-operational during the CSIP campaign. The inner domain is centered over the UK, has 288×360 grid points and is the focus of the study. The analyses presented here are based on the slightly cropped domain shown in Fig. 3.6, which has been stripped of 25 grid points on each side to avoid any spin-up issues associated with the forced lateral boundaries. The simulations analyzed in this study were started from a Met Office mesoscale model (12 km grid spacing) analysis at 0100 UTC on 25 August 2005. The lateral boundary conditions for the 4 km grid spacing simulations were provided by a 12 km grid spacing simulation which in turn used lateral boundary conditions from the operational global model simulation. No data assimilation was used in the simulations performed for this study and as such the 12 km and control 4 km simulations performed here differ slightly from those described in Clark and Lean (2006) where the operational 12 km model simulation (including data assimilation) was used to provide the boundary conditions for the 4 km simulations. In this study the runs were started at 0100 UTC to allow the spin-up stage of the evolution to complete before sunrise (and to compensate to some degree for the lack of data assimilation).

The model makes use of a comprehensive set of parameterizations which includes the surface layer (Essery et al. 2001), radiation (Edwards and Slingo 1996) and the mixed-phase cloud microphysics (Wilson and Ballard 1999). The convection and boundary-layer parameterizations are key to this study and briefly described here. The convection scheme of Gregory and Rowntree (1990) is used for both the 12 and 4 km grid spacing simulations with a modification developed by Roberts (2003) applied at the higher resolution. The Gregory and Rowntree (1990) scheme has a trigger dependent on

the initial parcel buoyancy and a mass-flux determined by a specified timescale for adjustment of Convective Available Potential Energy (CAPE); the timescale is set here to the typical value of 30 minutes. The modification avoids the accumulation of high values of CAPE at the grid scale (which can lead to unphysical “grid-point storms”) by specifying the CAPE adjustment timescale as an increasing function of CAPE. This forces the model to explicitly resolve most of the deep convection; the parameterization scheme mainly deals with shallow convection. This modification was specifically designed for the 4km grid spacing configuration of the UM and has proven successful (Lean et al. 2005; Roberts and Lean 2008).

The boundary-layer parameterization identifies seven types of boundary layers: stable, stratocumulus over stable, well mixed, decoupled stratocumulus over cumulus, decoupled stratocumulus not over cumulus, cumulus capped, and shear-driven boundary layer. The first six of these are described in Lock et al. (2000); the shear driven boundary-layer type is a more recent addition. The categorization of each grid point into the different types depends on numerous parameters, but it is based on the adiabatic ascent of a parcel, rising from 10 m above the ground, and on its eventual descent from cloud top. To avoid overdue sensitivity to grid level noise, a threshold of 0.4 K is applied to the liquid-frozen water potential temperature before computing the ascent values. Once the boundary-layer type has been determined other parameterizations may be called (e.g., entrainment or convection) and/or specific turbulent viscosity coefficients are computed for the turbulent mixing parameterization.

3.2.c Control Run

Hindcast studies with the 4km grid spacing model starting at 0600 UTC were performed by the Met Office and found to capture the significant features of this IOP

(Clark and Lean 2006). A shower cluster triggered in approximately the right time and place with the first precipitation at about 0815 UTC preceded by a short convergence line along the coast (although unlike in the observations no cloud or precipitation occurred before this). By 1030 the model had a linear cluster of showers although this did not extend far enough to the south and propagated too slowly. The control (unperturbed) 4 km grid spacing run performed here behaves similarly to the Met Office performed run, in the sense that the main features of the event are captured but location and timing are slightly different. In our control run the squall line originated from a cluster of showers that formed around 630 UTC over the Bristol Channel and then moved inland intensifying at the right time and location. As for Clark and Lean (2006) by 1030 our simulation had a line of showers that did not extend far enough to the south, but it propagated at higher speed than observed. More generally, when compared against radar observations the convective precipitation features that encompass at least a few grid points are well captured and broadly consistent with the observations in space and time. As the typical horizontal extent of the storms diminishes in the later part of the afternoon the UM tends to underestimate both size and intensity, but the precipitation rates are still very realistic.

The domain average hourly accumulation for the “rainy” grid points (i.e., with an hourly accumulation of at least 1 mm) and the number of rainy grid points in the control simulation are shown in Fig. 3.2. The average accumulation peaks at 0700 UTC whereas the number of rainy grid points peak later at 1300 UTC. This shows the transition from intense but localized precipitation to weaker but more distributed stratiform precipitation. The decrease in the number of rainy points and the average accumulation is associated with the sharp decrease in CAPE throughout the domain.

The evolution of the main boundary-layer types in the control simulation is shown in Fig. 3.3. The dominant boundary-layer types are cumulus capped, stable, and well mixed. The evolution is characterized by two rapid transition periods, one between 0600 and 0730 UTC and one between 1700 and 1830 UTC. The first transition is the response to the increasing shortwave radiation at sunrise and is marked by a swift decrease in the

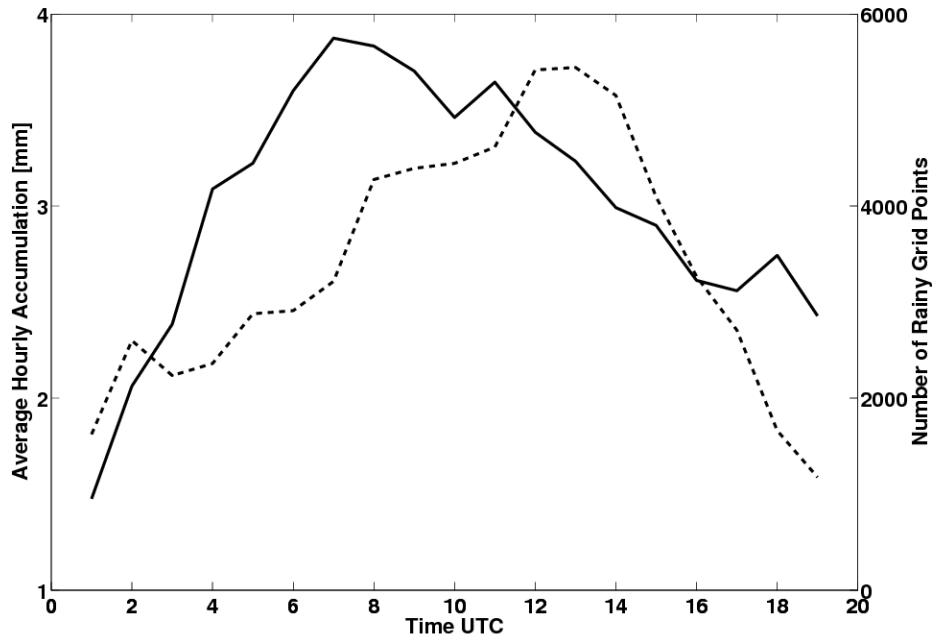


Figure 3.2 Hourly precipitation accumulation between the time shown and the following hour (solid line, left axis) and number of grid points with an hourly accumulation of at least 1 mm (dashed line, right axis). The total number of analyzed grid points is 73780.

percentage of grid points that have a stable boundary layer as they primarily switch to being categorized as cumulus-capped and well mixed. The second transition is the reverse of that occurring at sunrise and is the response to the diminishing shortwave radiation before sunset. The percentage of shear driven and decoupled stratocumulus not over cumulus grid points peak as intermediate states during both transitions. Overall, the percentage of cumulus-capped points increases from the outset of the simulation, peaks

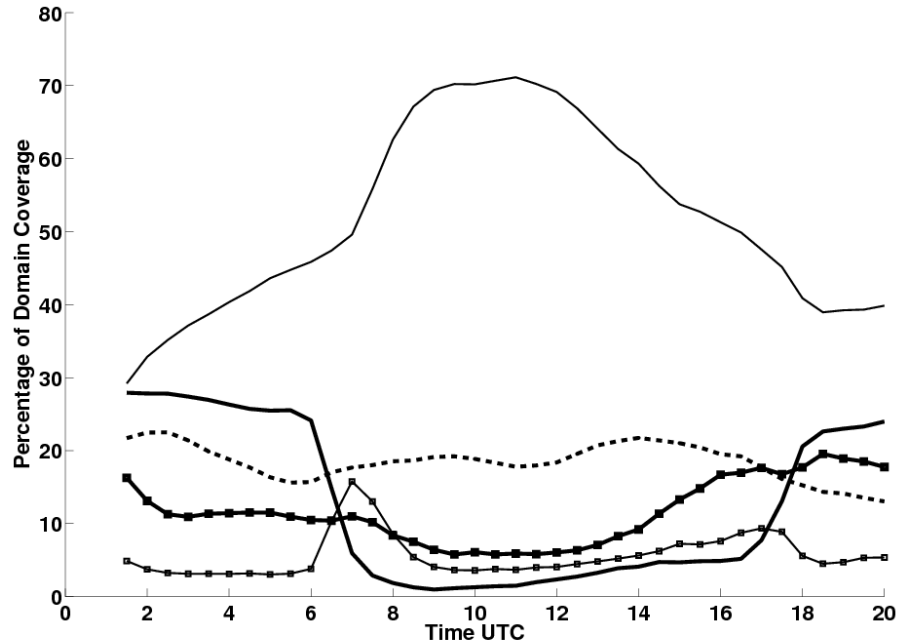


Figure 3.3 Percentages of domain covered by the seven boundary-layer types: Cumulus Capped (thin continuous line), Stable (thick continuous line), Well Mixed (thick dashed line), Shear Driven (thin line with squares) and all the stratus types (Stratus Over Stable + Decoupled Stratus + Decoupled Stratus Over Cumulus, thick line with squares).

between 9 and 12 UTC and decreases afterwards with the strongest time rate of change occurring during the morning transition period. The percentage of stable points decreases sharply during the first transition period, increases slowly after this, due to the formation of cold pools and then increases sharply again during the second transition period. The percentage of well-mixed points decreases slowly until the first transition period, varies somewhat during the day and then decreases further before and during the second transition period.

3.3 Perturbation Strategy

3.3.a Overview

The model state could of course be perturbed in numerous realistic ways. The aim of the present study is to determine those processes that are most effective in generating

realistic alternative simulations of scattered moist convection (i.e., of generating error growth). Convective triggering is a complex process (e.g., Bennett et al. 2006) that depends on the strength of boundary-layer thermals, wind convergence patterns, roughness gradients and other factors, but the essence is a means of lifting air parcels beyond their level of free convection.

A simple way of altering the triggering is to modify the buoyancy of a boundary-layer parcel by perturbing its potential temperature, moisture content or some combination of the two. Although boundary-layer moisture perturbations can be significant (particularly in altering the characteristics of triggered convective storms), temperature perturbations consistent with standard observational errors are a more important factor in determining triggering itself (e.g., Crook 1996; Fabry 2006). Here we perturb the potential temperature, which is a prognostic variable of the UM. We consider both repeated perturbations with a constant frequency (referred to as sequential-perturbation experiments) and single perturbations made at a specified time during the simulations (referred to as single-perturbation experiments). In the following subsections the perturbation structure, timing and height are justified and the experiments performed are detailed.

3.3.b Perturbation Structure

The random perturbation fields are constructed by the convolution of a random number field with a Gaussian kernel and applied at a specific model level, which is an average of 1280 m above ground (sensitivity tests to the height chosen are shown in Section 3.3.d). At each horizontal grid point, a random number is selected for the amplitude of a Gaussian distribution with standard deviation σ_{gauss} . Thus, for $\sigma_{gauss} > 0$ km the total perturbation at a grid point will be the sum of the amplitudes of

perturbations centered at all grid points in the domain. The sum of the perturbations constitutes the full two-dimensional perturbation field and it averages to zero over the entire domain. A set of random numbers generated with a specified seed and uniformly distributed between plus and minus unity is used to construct the perturbation field. This is then scaled to the desired peak amplitude before application. The scaling is determined from the standard deviation σ_{pert} of the perturbation field. Henceforth, we shall refer to $A=3\sigma_{pert}$ as the chosen perturbation amplitude.

Three perturbation amplitudes are considered, specifically $A=1$, 0.1 , and 0.01 K. The largest value was chosen to test nonlinearities and to directly affect storm development by significantly altering the buoyancy of the underlying air. It is also of the same order of magnitude of the largest values obtained by the 3DVAR data assimilation theta increments within the boundary layer that were used to generate the initial conditions for the simulations (Dixon personal communication, 2008). Furthermore Kong et al. (2007), tested three different methodologies for perturbing initial conditions and their increments were in excess of 1 K. The smallest value (0.01 K) provides an indicative bound on the potential predictability of a single convective-scale forecast: a practical system that can produce a model state to this level of accuracy is almost impossible to envisage. The 0.1 K perturbation amplitude is the most credible choice, being consistent with typical surface temperature measurement errors (e.g., Fabry 2006) and typical turbulent fluctuations in the convective boundary layer (e.g., Stull 1988, p358). Such perturbations are intended to be sufficient to change the location and timing of the triggering of moist convection.

Three values of σ_{gauss} are considered: 24, 8, and 0 km. These values correspond to typical length-scales in the full perturbation field of roughly $\sim 6\sigma_{gauss}$, as shown

below. A standard deviation of 24 km was chosen to provide a perturbation length-scale that is well resolved at the model grid spacing and larger than the typical lowest horizontal scale (80 km) of the 3DVAR data assimilation scheme used to produce the initial conditions for the outer grid. The choice of 8 km provides an intermediate scale between the well-resolved and the grid scale. The limiting case of a Gaussian for which the standard deviation tends to zero gives rise to spatially-uncorrelated grid-scale noise.

Figure 3.4 illustrates the effects of the perturbation field on the power spectrum for potential temperature at 1000 UTC, but the conclusions drawn here hold equivalently at any time of the day. The power spectra for the control (unperturbed) simulation and for those simulations for which the power spectra are distinguishable from the control simulation are shown in Fig. 3.4. The ratio of spectra between the perturbed and control potential temperature distributions shows a peak in added variance around a wavelength of $6\sigma_{gauss}$, or 144 and 48 km in Figs. 3.4b and c, respectively. For a perturbation amplitude of 1 K, the spectrum is significantly altered for wavelengths from $\sim\sigma_{gauss}$ to $10\sigma_{gauss}$, while changes are discernible with the 0.1 K amplitude and almost indistinguishable for 0.01 K. For spatially-uncorrelated perturbations (Fig. 3.4d), the relative magnitude of the contributions is much larger than for the correlated perturbation fields, since the perturbed spectrum field has much more power at near grid-scale wavelengths.

3.3 Frequency of Perturbation

Over a model domain that encompasses the entire UK (see Fig. 3.6), the phenomena that lead to the onset of convection and affect its development occur at different times of the day. Thus, perturbing the model state at a specified frequency

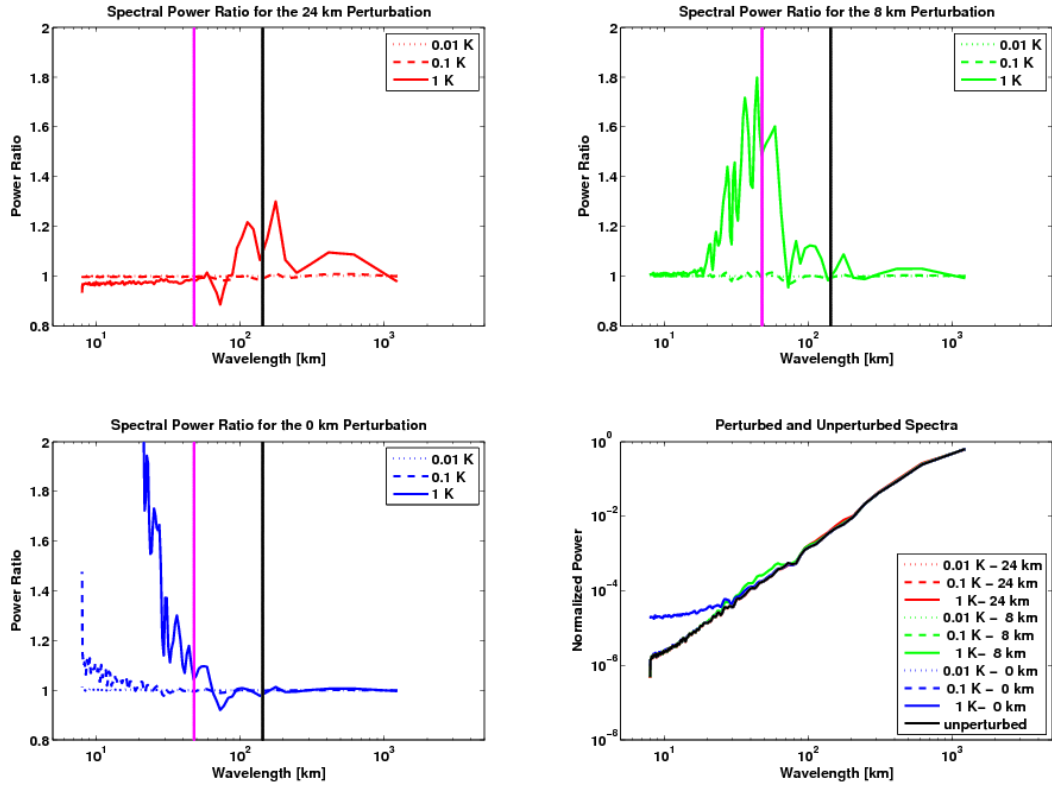


Figure 3.4 Normalized power spectrum of potential temperature, just before and just after a perturbation field is applied at 1000 UTC. In (a) the spectra are normalized such that an integral over the wavenumber produces unity. Note the log scale on the horizontal axis. In (b)–(d) the spectra are additionally normalized by the spectrum prior to perturbation. The amplitudes A and widths σ_{gauss} are indicated on the legends. The perturbation is applied at 1280 m (see Section 3.3.d).

(sequential-perturbation experiments) is a simple and effective way to ensure that a perturbation has been applied prior to all potentially sensitive times during the simulation. Perturbations applied at successive times during a simulation have no temporal correlation. A few single-perturbation experiments have also been performed to determine how the sensitivity of the simulation to perturbations changes during the day. Note that Grabowski et al. (2006) also perturbed the model state with random perturbations of temperature and moisture in the first kilometer of the boundary layer every 15 min., but their intent was to trigger instabilities to inter-compare different cloud resolving models.

In the sequential-perturbation simulations the first perturbation occurs one perturbation cycle after the start of the simulation. This is done to allow some time for the simulation to spin-up and produce a more balanced state from the interpolated lower-resolution initial conditions. The spin-up is thus considered to be a separate issue from the ongoing uncertainties that exist in the model state (see also Section 3.6.c).

The choice of the perturbation frequency represents a compromise between two considerations. On the one hand, a typical equilibrium timescale for a well-mixed boundary layer is of the order of 10 to 20 minutes (e.g., Nieuwstadt and Brost 1986; Stull, 1988, p450), and the boundary layer would not be able to fully adjust to each perturbation if perturbations were applied too frequently. On the other hand, infrequent perturbations could result in the absence of any perturbations during key transitions in the boundary-layer structure; for example, from stable to cumulus-capped after sunrise (Fig. 3.3). Any application frequencies intermediate between these two limits are hypothesized to be likely to lead to similar levels of model error growth.

The sensitivity of error growth to the perturbation frequency was investigated by considering two choices of frequency, 2 hr^{-1} and 1 hr^{-1} . Figure 3.5 shows the time evolution of the root-mean-square error (RMSE) of potential temperature on the perturbed model level (see Section 3.3.d) for a perturbation amplitude of 1 K and a perturbation standard deviation of 24 km (data at every model timestep - 100 s). Although only a four hour period is plotted for clarity the behavior is similar throughout the day. The reference field is taken from the end of a perturbation cycle and is reset on alternate cycles. Thus, each plot exhibits a pattern that is repeated every two perturbation cycles; error growth towards saturation can be seen during the first cycle, with a jump in RMSE signifying the start of a new cycle. This is followed by further growth until the reference

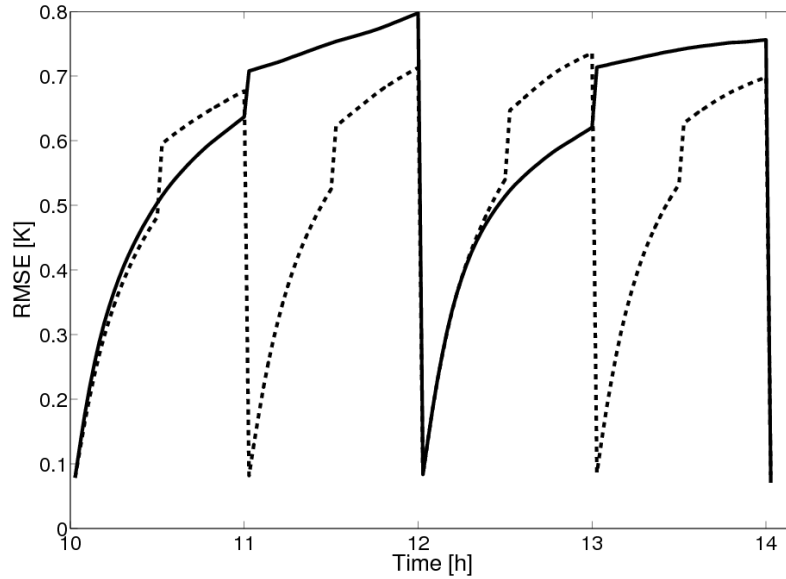


Figure 3.5 Time evolution of the RMSE of the potential temperature at the perturbed level. The reference field is taken from the end of a perturbation cycle and is reset on alternate cycles. Perturbations are applied at frequencies of 2 hr^{-1} (dashed line) and 1 hr^{-1} (solid line) with a perturbation amplitude of 1 K and a σ_{gauss} of 24 km.

state is reset. The jump is modest in comparison to the growth during the first cycle, indicating that error growth is mainly driven by the model's response to a perturbation, as opposed to the perturbation application itself. Somewhat larger jumps and more rapid growth are seen in the second cycle when perturbations are applied more frequently, with the result that the overall error growth over one hour of simulation is very similar, whether produced by a single perturbation, or two half-hourly perturbations.

A similar plot (not shown) for the perturbations of amplitude 0.01 K does not exhibit jumps on application of the second perturbation, consistent with the small changes produced in the power spectrum. Finally, we note that an analysis of the time evolution of the potential temperature power spectra with a 1 K perturbation amplitude indicates that while the effect of the perturbation at the perturbed scales decreases rapidly, after 30 minutes it remains perceptible. Hence, the perturbations do not dissipate entirely between applications.

3.3.d Height of Perturbation

The sensitivity of the results to the height of application of the perturbations was investigated by considering four choices of perturbation heights

1. Model level 8 (approximately 1280 m above the ground)
2. 500 m above the ground, as in Lean (2006);
3. from the surface up to 400 m with a vertical correlation of one throughout this depth (this is similar to Petch 2004, who used a depth of 200 m with stronger amplitude perturbations);
4. at the lifting condensation level.

No tests were performed applying perturbations above the boundary layer, since the aim is to perturb the triggering process (Section 3.3.a). Lean (2006) demonstrated that perturbations to potential temperature applied at 4500 m do not lead to significant error growth in idealized simulations of a cumulus ensemble. Experiments were performed with a perturbation amplitude of 1 K and a perturbation standard deviation of 24 km. The root-mean-square error for the hourly accumulation of total precipitation (for example) exhibited very little sensitivity to the height of the perturbations; hence the perturbation height was fixed at 1280 m in this study.

3.5 Perturbation Experiments

The experiments performed are summarized in Table 3.1. The perturbations were applied at the same height in all these experiments, namely model level 8, equivalent to a height of 1280 m. Nine *sequential-perturbation* experiments were performed with perturbations applied twice every hour and with varying perturbation amplitudes, A , and standard deviations, σ_{gauss} , as discussed in Section 3.3.b. These experiments are labeled

$\sigma r A s$, as a shorthand for $\sigma_{\text{gauss}}=r\text{km}$, $A=s\text{K}$. A trailing asterisk indicates that, for those experiments, a set of six simulations were performed differing only in the set of random numbers generated. We will refer to such sets of simulations as *ensembles*. They allow us to compare the spread that would be produced by varying the perturbation-field parameters with that arising from different realizations of the same perturbation process. Two ensemble members used different seeds to the first ensemble member for the random number generator and the remaining three ensemble members were generated from the first three ensemble members by multiplying the perturbation fields by a factor of -1, similarly to Done et al. (2008).

Eight *single-perturbation* experiments have also been performed. For these experiments the standard deviation σ_{gauss} was fixed at 24 km and two perturbation amplitudes ($A=0.01$ and 1 K) were considered. Four application times were tested: specifically in the initial conditions (0100 UTC) and at 0700, 0830, and 1000 UTC. These experiments are labeled in the form $t-s$. Here t indicates the application time in UTC (or else as IC for initial condition perturbations) and s is the perturbation amplitude A in degrees K. The first ensemble member in an ensemble experiment and all non-ensemble experiments used perturbation fields derived using the same random number sequence. As an example of the impact of these perturbations, Figure 3.6 shows the precipitating cloud fields at 1000 UTC for the unperturbed run and three perturbation experiments all with a perturbation amplitude of 1 K but with varying perturbation standard deviation. This snapshot shows that while the perturbations have changed the locations of individual clouds, on the regional scale the cloud distribution appears unaltered and is realistic.

Label	σ_{gauss} [km]	Amplitude, A [K]	Application Timing	Bias [$\times 10^{-2}$ mm]	Max 3h RMSP increment [mm]
Control	–	–	never	0.0	–
σ 24A1*	24	1	2 hr ¹	-6.37	2.07
σ 8A1	8	1	2 hr ¹	-6.47	2.55
σ 0A1	0	1	2 hr ¹	-3.92	2.13
σ 24A0.1*	24	0.1	2 hr ¹	-1.07	1.41
σ 8A0.1	8	0.1	2 hr ¹	0.67	1.12
σ 0A0.1	0	0.1	2 hr ¹	1.13	1.16
σ 24A0.01*	24	0.01	2 hr ¹	0.74	0.94
σ 8A0.01*	8	0.01	2 hr ¹	-0.63	1.00
σ 0A0.01*	0	0.01	2 hr ¹	-1.00	1.12
IC-1	24	1	IC	7.43	2.62
0700-1	24	1	0700	-0.08	1.52
0830-1	24	1	0830	0.28	1.47
1000-1	24	1	1000	-1.35	1.13
IC-0.01	24	0.01	IC	-0.49	0.84
0700-0.01	24	0.01	0700	1.11	0.95
0830-0.01	24	0.01	0830	-1.18	0.74
1000-0.01	24	0.01	1000	-0.09	0.74

Table 3.2: List of simulations performed and their characteristics. The simulation labeling is explained in the main text. Characteristics shown are the standard deviation σ_{gauss} , the perturbation amplitude A , its application timing, the bias in the domain-averaged precipitation accumulated during the simulation, and the maximum three-hourly increment of RMSP (as defined in Section 3.3.d). For the starred simulations the values reported refer to the first member of the ensemble.

3.4 Diagnostics

The diagnostics used in this study are presented here. Two types of diagnostics are considered: those associated with the *direct* impact of the perturbations - these reveal how the perturbations are 'felt' by the model (and may be somewhat model specific) - and those associated with the *indirect* impact of the perturbations i.e., error growth due to the model evolution. Four measures of the direct effect of the perturbations are examined: the effect on the convective instability of adding the perturbations (after one timestep),

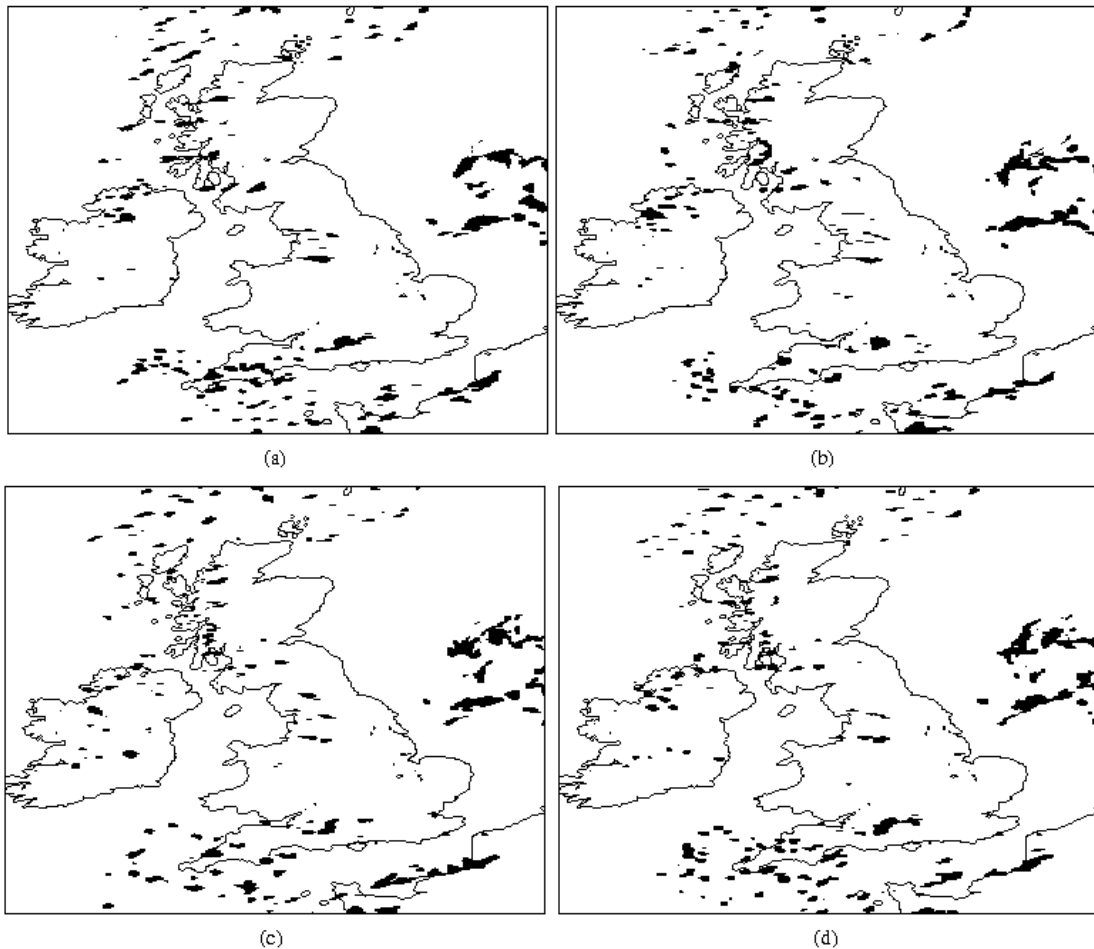


Figure 3.6: Precipitating clouds, as defined in Section 3.3.d, at 1000 UTC for (a) the control, (b) the σ_{24A1} , (c) the σ_{8A1} , and (d) the σ_{0A1} run.

the model adjustment to the perturbations, the direct effect on total cloud water and cloud distributions, and the direct effect on the diagnosed boundary-layer types. Three measures of the indirect effect of the perturbations are examined: the evolution of the effects on cloud distributions and diagnosed boundary-layer types and the root-mean-square error of the hourly-accumulated precipitation (RMSP).

The choice of diagnostics used to analyze the indirect effect of the perturbations is motivated by the focus of this study i.e., the error growth in convective-scale forecasts, and particularly the effect on precipitation. From a hydrological perspective, it is important to consider not simply the accumulation of precipitation, but also its location since a small change in storm location may potentially produce a significant change in the river catchment affected by the storm. The diagnostics are intended to be complementary, some provide an overview of the cloud and convective activity and some are sensitive to the timing and location of storms. Several recent papers provide methods to overcome these sensitivities (e.g., Casati et al. 2008, Ebert and McBride 2000, etc.). However they often require considerable computational resources and focus on one specific aspect (e.g., Roberts and Lean 2008, etc.), whereas the diagnostics used in this study aim to provide a more comprehensive view.

Convective instability is diagnosed using CAPE (convective available potential energy) and CIN (convective inhibition). The CAPE is defined here by the integrated buoyancy between the first model level (20 m above ground) and the level of neutral buoyancy (LNB), and so may be negative. Model adjustment mechanisms are determined by analyzing vertical profiles of horizontally-averaged pressure, vertical velocity and total cloud water content and their dependence on the sign of the potential-temperature perturbation.

Cloud distributions are affected both directly and indirectly by the applied perturbations: directly because a change of temperature will lead to a change in relative humidity and indirectly because the evolution of perturbed clouds and future cloud triggering will be affected. Two cloud definitions are considered: *precipitating clouds* and *non-cirrus cloud*. Precipitating clouds are defined based on grid points with rain rates of at least 1 mm h^{-1} (for consistency with RMSP). Note that tests have shown that the validity of the results presented in Section 3.6.b are not sensitive to a 10% variation of this threshold. Non-cirrus cloud is a more generic definition of cloudy air and is based on grid points with a vertically-integrated water path below 8 km of at least 0.05 kg m^{-2} . Non-cirrus clouds may or may not precipitate. The integration extends from the surface to the first model level above the maximum height of the squall line mentioned in Section 3.2.a in order to exclude higher clouds, namely cirrus, which are unlikely to be affected by the perturbations. An appropriate threshold for this definition may be somewhat case dependent, but sensitivity tests indicate that the relative behavior of such cloud statistics between simulations is unaffected by halving or doubling the threshold; however, the average cloud size and number are affected. Other definitions using liquid-water path or column-averaged total water density have also been tested, with similar results. “Clouds” are defined as connected clusters of cloudy grid points satisfying one of the above definitions. Grid boxes which share either an edge or a corner are considered to be connected. The size or area of the clouds is measured in grid boxes and cloud fields are evaluated every 30 minutes through the simulation. Including or excluding the spin-up phase does not significantly change any of the time-averaged cloud diagnostics presented.

The boundary-layer type determined by the UM is affected by the applied perturbations both directly and indirectly: directly because the boundary-layer type is

based upon the ascent of a parcel from 10 m above ground and, indirectly because the perturbation can induce changes which later cause the boundary-layer type to switch, as described by Lean (2006). The direct aspect will be addressed, based upon the percentages of the domain covered in the single perturbation experiments by the grid points that changed boundary-layer type with respect to the control run at perturbation time. While Lean (2006) analyzed the temporal evolution of the boundary-layer types within minutes of the perturbation, we evaluate the indirect effect by analyzing the time average of the domain cover for each of boundary-layer types. Such analysis provides useful information on how the sequential perturbations alter the boundary layer and the behavior of the model.

RMSP is a simple and widely used error norm (e.g., Molteni et al. 2001; Snyder and Zhang 2001):

$$RMSP = \sqrt{\frac{1}{N} \sum_{i=1}^N (ap_i - ac_i)^2}$$

where N is the number of grid points and ap_i and ac_i are the hourly-accumulated precipitation at grid point i in the perturbed and control runs respectively. The RMSP is computed relative to the control simulation and has contributions from those grid points which have a minimum hourly accumulation of 1 mm (hereafter “rainy” grid points) in either the control or the perturbed simulation. The 1 mm threshold is chosen to avoid undue sensitivity to grid points with very small accumulations. Tests indicate that varying the threshold from 0.1 to 10 mm does not alter any of the qualitative conclusions obtained from this diagnostic concerning the relative behavior of the simulations. Using the instantaneous rates, sampled every 30 min, also had no impact on the conclusions to be drawn. Such insensitivities may be due to the scattered nature of the convection within

the large domain. At any time in the simulation, there are multiple storms at different stages of their life cycles.

From both meteorological and hydrological perspectives it is important to know to what extent the perturbations tend to displace storms, alter their intensity, create new ones, or any combination of these three possibilities. A complete analysis of this issue could be provided only by keeping track of each storm at each timestep. However, some insight into such issues can be provided by decomposing the squared RMSP into three MSP (mean squared error in the hourly-accumulated precipitation) components, from three sets of contributing points: those that are rainy only in the control run, those that are rainy only in a given perturbed run, and those that are rainy in both simulations. The sets will be referred to as *CONTROL*, *PERTURBED*, and *COMMON* respectively. More specifically, the summed square error is computed for each set separately, and is normalized by the total number of rainy points yielding

$$\begin{aligned}
 TOTAL\ MSP &= \frac{1}{N} \sum_{j=1}^{N_{CONTROL}} ac_j^2_{MSP\ CONTROL} \\
 &+ \frac{1}{N} \sum_{k=1}^{N_{PERTURBED}} ap_k^2_{MSP\ PERTURBED} \\
 &+ \frac{1}{N} \sum_{l=1}^{N_{COMMON}} (ap_l - ac_l)^2_{MSP\ COMMON}
 \end{aligned} \tag{3.1}$$

where $N_{CONTROL}$, $N_{PERTURBED}$ and N_{COMMON} are the number of rainy control, perturbed and common points respectively (such that $N=N_{CONTROL}+N_{PERTURBED}+N_{COMMON}$) and j , k and l

are the grid points in the sets *CONTROL*, *PERTURBED* and *COMMON* respectively. Note ap_j and ac_k are zero by definition (hourly accumulation in the perturbed run at points that are only raining in the control run and hourly accumulation in the control run at points that are only raining in the perturbed run respectively). Note that an increase in the *CONTROL* contribution to the total MSP (for constant N and $N_{CONTROL}$) implies an increase in the average intensity of *CONTROL* precipitation; equivalently an increase in the *PERTURBED* contribution to the total MSP (for constant N and $N_{PERTURBED}$), implies an increase in the average intensity of *PERTURBED* precipitation. However an increase in the *COMMON* contribution to the total MSP (for constant N and N_{COMMON}) is due to an increased average difference in the precipitation intensity at the common points in the control and perturbed simulations.

3.5 Results: Direct Effects

3.5.a Perturbation Effects on CAPE

The overall changes of the CAPE due to the perturbation are very minor, so that the maximum bias for the strongest amplitudes is less than 0.5 J kg^{-1} in absolute value, at any time. The main contribution to such a small bias comes from a relatively small number of grid points (e.g., 0.6% of the grid points for the 1-0700 experiment), whose unperturbed ascent profile is always on the left of the temperature one. Thus they only have CIN and no level of neutral buoyancy (LNB) and for those points the CAPE is considered null. At some of these points a negative perturbation introduces a lid and sets the LNB at perturbation level, resulting in a negative contribution to the available energy that ranges from 0 to -60 J kg^{-1} . The opposite process also happens (i.e., a lid is removed), with changes in CAPE of opposite sign and slightly smaller values, and on a

smaller number of points. Thus, if a trigger is present, storms can be generated or suppressed at these points. Most of the points do not contribute to the overall bias because the perturbation level is above their LCL and the slight increase in CAPE associated with the negative perturbations is offset by the decrease due to the positive perturbations.

3.5.b Vertical Adjustment to the Perturbation

The first, although not the only, dynamical response to the heating consists of acoustic and Lamb waves which within minutes accomplish the expansion due to the heating (Chagnon and Bannon 2005). The sign of the response depends upon the sign of the perturbation, so when the average vertical pressure profile is computed using only the grid points associated with a positive perturbation (hereafter *positive points*), then a small positive bias in the pressure is visible. Such pressure bias between the control run and the 0700-1 simulation is shown, for the positive points, in Fig. 3.7. This bias decreases exponentially with height and the e-folding length for the tropospheric levels corresponds to the tropopause height as estimated from the two rawinsondes of Southern England (Camborne and Herstmonceux) taken at 1200 UTC that day. The pressure adjustment occurs within one timestep and involves the entire column as the stratospheric model levels are affected as well. Numerical calculations carried out for a linearized dry atmosphere using the domain average temperature vertical profile at 0700 UTC for the unperturbed run show that the perturbation of the pressure vertical profile is consistent with an acoustically adjusted state (Chagnon, personal communication 2008). Hohenegger and Schar (2007b) also observed acoustic waves as result of a temperature

perturbation similar in shape and amplitude, although in their case the perturbation was vertically decaying from the surface.

The magnitude of the pressure perturbation per se, is not likely to be of any meteorological concern. However, it has an associated Lamb wave, shown in Fig. 3.8(a). The vertical velocity bias over all the positive points is of the order of 3 mms^{-1} and of 10 mms^{-1} when the average is computed over the grid points whose perturbation is at least 1 K. Such small values of vertical velocity are unlikely to trigger new convection, but the modification to the unperturbed average vertical velocity is significant throughout the troposphere (Fig. 3.8b) and the maximum velocity in the mid-troposphere is increased by 38% for the 0700-1 simulation.

The changes in the vertical velocity associated with the Lamb wave for the 1 K simulations are significant, but their magnitude is fairly small and at the same time the Lamb wave travels at the speed of sound, resulting in very small displacements. However, the Lamb waves are generated throughout the entire domain, and although their intensity diminishes as they travel away from the source, they take roughly 30 minutes to cross the domain at the speed of sound. Moreover they affect the whole depth of the model atmosphere. It is then possible that they modify the environment enough to contribute to the error growth. It should also be pointed out that the importance of the Lamb and acoustic waves are generally not well represented since the coefficients of the off-centered advection scheme are designed to damp them (Davies et al. 2005) and because of the relatively long timestep (100 s) not apt to properly resolve the fast acoustic waves. However, they still propagate errors very quickly as shown by Hohenegger and Schar (2007b) for an individual Gaussian perturbation with an amplitude of slightly less than 1 K.

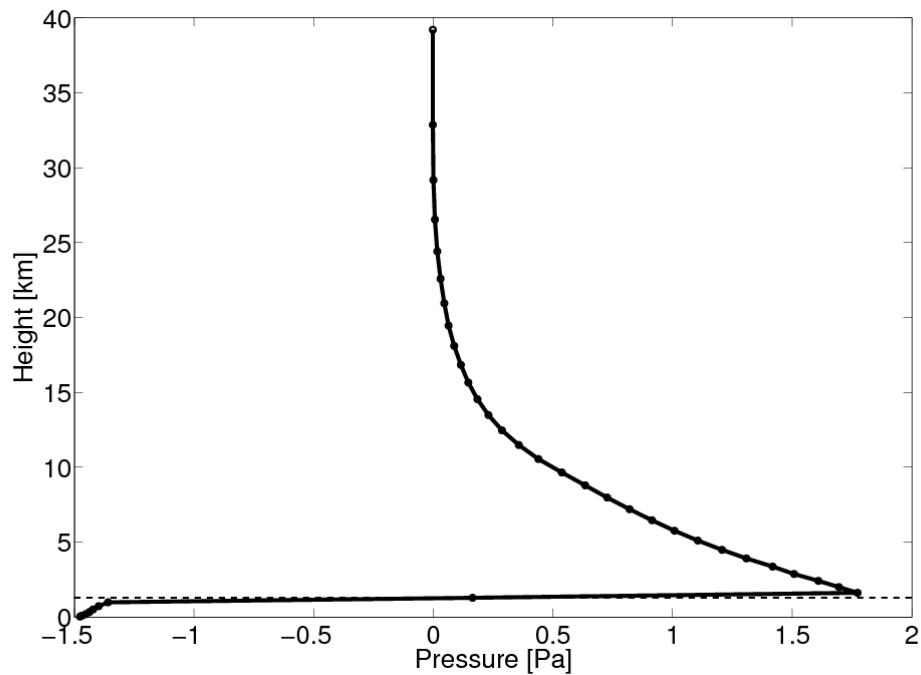


Figure 3.7: Vertical profile of the pressure bias between the 0700-1 and control runs, computed for the grid points with positive potential temperature perturbation at 0700 UTC, one timestep after the perturbation application. The dashed line marks the height at which the perturbations are applied.

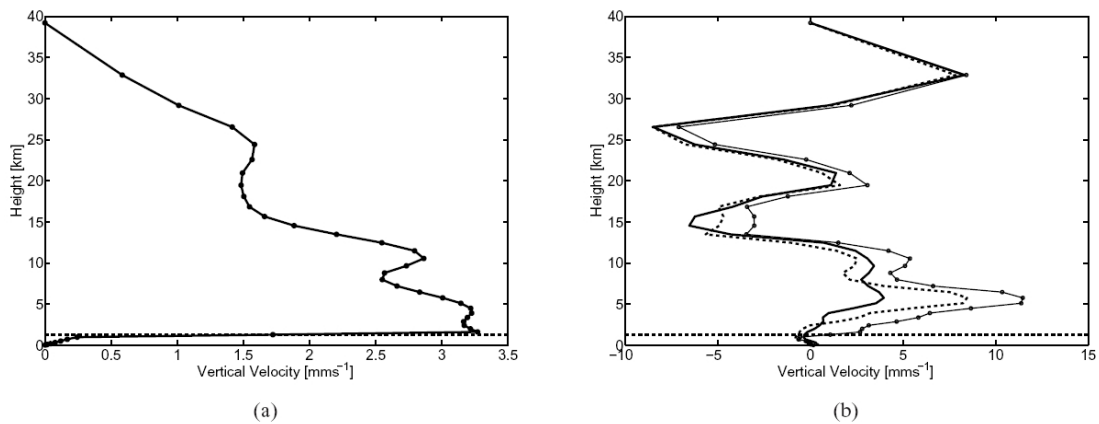


Figure 3.8: (a) Vertical velocity bias between the control and the 0700-1 runs over the grid points with positive perturbation. (b) Average vertical velocity over the entire domain for the control run (solid thick line), over the positive perturbation points for the control run (dashed line), and over the positive points for the 0700-1 run (solid thin line with circles). The dashed line in both panels marks the height at which the perturbations are applied.

The bias of the pressure and of the vertical velocity due to the potential-temperature perturbations scales linearly with the perturbation amplitude (i.e., are two orders of magnitude smaller for the equivalent experiment with 0.01 K amplitude) and the bias profiles over the negative points are symmetrical with the ones shown. The 1000-1 single perturbation run yields essentially identical results, indicating that the acoustic adjustment depends primarily on the perturbation and is not sensitive to the changes in the model atmosphere that occur between 0700 and 1000 UTC.

3.5.c Boundary-Layer Changes

The perturbations introduced can directly influence (i.e., immediately upon application of the perturbation) the determination of the boundary-layer types (as noted in Section 3.3.d). This effect has been calculated for six of the single perturbation runs, excluding the two runs for which the perturbations are made at the initial time. Table 3.2 shows that this effect is almost negligible for the 0.01 K amplitude runs as the maximum change in a boundary-layer type is 0.05%. This effect is more significant for the 1 K amplitude runs but the change is 2% at the most. However, it is important to point out that while the area covered by such changes is small compared to the entire domain, it can represent a significant fraction of the cover for a specific boundary layer (10% for stratus over stable in the 0700-1 run). Moreover, the number of grid points involved in such changes is fairly large as 2% of the domain corresponds to 1475 grid points and can contribute to error growth as shown by Lean (2006). Note that it is unlikely that the grid points which changed their boundary-layer type will immediately go back to the original one, since the even non-growing potential temperature perturbations persist for at least 30 min. (Section 3.3.c), albeit with decaying amplitude.

A thorough analysis of the changes in the boundary-layer type cover shows that generally such changes are directed towards numerous other types, with the exception of the stable boundary layer which only loses points and only loses them to the stratus over stable type. Thus the perturbation generates a stratus over these points.

Run ID	Well Mixed	Stable	Sc over Stable	Decoupled Sc not over Cu	Decoupled Sc over Cu	Cu capped	Shear dominated
Control at 0700	17.68	5.96	3.21	6.67	1.12	49.59	15.78
Control at 0830	18.66	1.27	0.26	5.71	1.55	67.13	5.41
Control at 1000	18.87	1.28	0.18	4.62	1.29	70.18	3.59
0700-1	0.92	0.31	0.31	0.98	0.20	1.12	0.81
0830-1	2.01	0.01	0.01	1.33	0.40	2.07	0.61
1000-1	2.21	0.04	0.04	1.81	0.30	2.07	0.47
0700-0.01	0.02	0.01	0.01	0.02	0.00	0.02	0.02
0830-0.01	0.04	0.00	0.00	0.03	0.01	0.03	0.01
1000-0.01	0.05	0.00	0.00	0.03	0.01	0.05	0.01

Table 3.3: The first three rows show the domain cover, expressed as percentage, for each boundary-layer type, at specific times within the control simulation. The subsequent rows show the percentage, for six single perturbation experiments, of the domain covered by the grid points that changed boundary-layer type with respect to the control run at perturbation time. Such percentages are based on the sum of the domain covers of the grid points that either before or after the perturbation were categorized to an individual boundary-layer type. This sum is then divided by two to avoid double counting, so that the sum of the percentages of the changes represents the domain cover of the total number of points that changed boundary-layer type.

3.5.d Total Water Path and Cloud Distribution Changes

The total water path (ice and liquid water) one timestep after the perturbation application is compared against the same quantity from the unperturbed run, at the same time. Changes above and below the perturbation level are negligible, but at perturbation level the evaporation (condensation) that follows the positive (negative) potential-temperature perturbations are significant. For the 0700-1 simulation the total water path, averaged over the positive points, decreases 18%, from 7.08×10^{-3} to 5.79×10^{-3} kgm^{-2}). Although the magnitude of the change diminishes by two order of magnitude for the equivalent smaller amplitude single perturbation experiment (0700-0.01) it is less consistent with a linear scaling than pressure and vertical velocity because of the non-linearity of the Clausius-Clapeyron equation. There is also a slight sensitivity to the time of the day as the changes described above are smaller in the single-perturbation experiments performed at 1000 UTC (e.g., 15% for the 1000-1 experiment).

The immediate repercussions of the total water path modifications on the cloud distributions are quite small, both on the number and average size at perturbation time for both cloud definitions. The number increases by 5.6% for the non-cirrus clouds of the 0700-0.01 experiment, consistently with the changes in the boundary-layer types, but this is an exception. Changes are generally below 1% for both cloud types and for all the single perturbation experiments. Changes in the cloud mean size are even less significant. It is worth mentioning that such changes are not necessarily consistent with the linear regime of the cloud distributions (Section 3.6.b since, for example, the 0700-1 perturbation causes an increase in both cloud mean size and number). This pattern is not surprising because, while the domain average perturbation is zero, the changes in the cloud condensate are obviously dependent upon its presence. Moreover an increase in

potential temperature causes more condensate to evaporate than the equivalent decrease causes it to condense.

3.6 Results: Indirect Effects

3.6.a Boundary-Layer Changes

The perturbations introduced can influence the determination of the boundary-layer types indirectly as well as directly (as shown in Section 3.5.c). Table 3.3 lists the time-averaged percentage for each of the seven types, and the changes to those values produced in the sequential-perturbation simulations.

For perturbation amplitudes of 0.01 and 0.1 K, the effects are very small (less than 0.2% of the control run value for all boundary-layer types), but for the larger perturbation amplitude of 1 K some more significant changes are found and roughly 10% of the domain has a different boundary-layer type from the control run. Inspection of the boundary-layer type percentages over the course of the simulations reveals that the changes brought about by the sequential perturbations are generally of the same sign and of a similar magnitude throughout the day (not shown). The main change is that the perturbations reduce the percentage of the cumulus capped boundary layer, and this is balanced chiefly by increases to the well mixed and the various stratocumulus boundary-layer percentages.

When the changes for the single perturbation experiments are compared against the sequential perturbation runs with the same metric, the changes observed in the sequential experiments are larger than the ones observed in the single perturbation runs, suggesting that during the course of the perturbed simulations the boundary-layer types change due to the indirect effects of the perturbation rather than to the direct ones.

3.6.b Cloud Distribution Changes

The time-averaged number of *non-cirrus* clouds in the model domain is plotted against their mean size in Fig. 3.9a. The control simulation has an average of around 125 such clouds with a mean size of 57 grid boxes (covering around 7% of the domain). The cloud number and size are clustered around these values for most of the perturbed simulations albeit with a tendency for slightly fewer larger clouds. The exceptions are for the sequential perturbation experiments and initial condition perturbation experiment with 1 K perturbation amplitude (σ_{24A1} , σ_{8A1} , σ_{0A1} , and IC-1). As discussed in Section 3.6.c these same simulations yield bias in the total domain rainfall. In these experiments, perturbations with smaller length scales give rise to smaller but more numerous clouds. These 1 K perturbation amplitude experiments appear to form a distinct subset of the experiments since regression analysis of the data from them and from all the other experiments separately both produce straight line fits with high correlations. This suggests that the non-cirrus cloud dynamics of the simulation is altered by these large amplitude perturbations. Similar comments apply for the precipitating clouds (Fig. 3.9b), which are of course smaller and less numerous covering on average just over 1% of the domain. The outlier for this diagnostic is the IC-1 simulation, with the data points for all other experiments being well fit by a straight line. This case has more precipitating clouds than in the control simulation, but of a similar size, a result which is consistent with its significant, positive precipitation bias (Table 3.1). As for the non-cirrus clouds, for a perturbation amplitude of 1 K the perturbation standard deviation affects the cloud sizes. Smaller standard deviations produce on average more and smaller clouds. Unlike for the non-cirrus clouds, the 1 K sequential perturbation and the IC-1 simulations have

the same linear response as all the other simulations, suggesting that the dynamics of precipitating clouds is not altered as significantly as for the non-cirrus clouds.

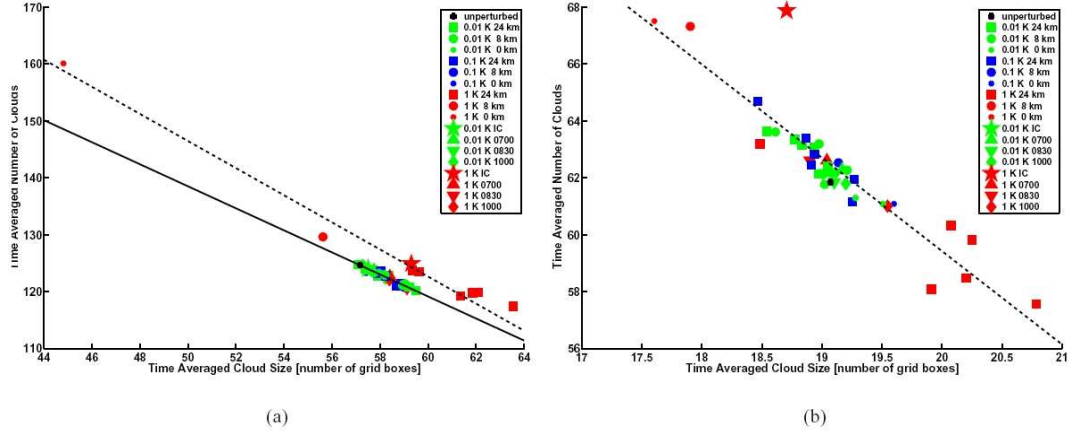


Figure 3.9: Mean cloud size and cloud number, time-averaged and plotted as a point for each simulation for (a) non-cirrus clouds and (b) precipitating clouds, where clouds are defined as in Section 3.4. Cloud size is measured in units of the model grid box. Cloud fields are evaluated every half hour starting 30 minutes into the simulation. The unperturbed simulation is denoted by a small black circle. Simulations with perturbation amplitudes of 0.01, 0.1 and 1K are denoted by green, blue and red symbols respectively. Sequential perturbation simulations with $\sigma_{\text{gauss}} = 24, 8$ and 0 km are denoted by squares, large circles and small circles respectively. Single perturbation simulations with perturbation application times of the initial time, 0700, 0830, and 1000 UTC are denoted by stars, upright triangles, inverted triangles and diamonds respectively. The dashed line in (a) is the least-squares fit for the sequential simulations with 1K perturbation amplitude, along with the IC-1 simulation. The solid line in (a) is the corresponding fit for all other simulations shown. In panel (b), the dashed line is the least-squares fit for all simulations other than IC-1.

3.6.c RMSP

Figure 3.10 shows the evolution of RMSP in the sequential perturbation runs. It is most responsive to the perturbation amplitude. When this is 1 K, the strongest error growth occurs after the second perturbation application (2 UTC). The RMSP peaks between 7 and 11 UTC before leveling off at about 3 mm. In contrast, for perturbations of amplitude 0.1 K, although the strongest growth is again seen at early times, there is no

peak in the RMSP evolution, and instead the error saturates at about 2.5 mm around 12 UTC. For the 0.01 K perturbations, the strongest growth is somewhat delayed, to 6 UTC. A clear saturation phase is not seen for these experiments within their duration although similar RMSP values to those achieved with the 0.1K perturbations are reached at the end of the simulations.

These results for the onset of strong growth suggest that small perturbations within the stable boundary layer have little effect upon any precipitation before sunrise. Only once the boundary layer starts to change its structure are the smaller perturbations capable of stimulating significant error growth. By contrast, the 1 K perturbations can be powerful enough to produce strong precipitation error growth almost from the outset. Figure 3.10b shows the RMSP evolution for three ensembles with varying perturbation amplitudes. Note that the range across the ensemble members increases with increasing perturbation amplitude. Systematic dependence on the perturbation standard deviation, σ_{gauss} , is not obvious in Fig. 3.10a, being comparable in size to the spread within the six-member ensembles. Thus, the horizontal-scale length of the perturbation does not significantly affect RMSP. However, some delay before the strongest growth can be seen for the experiments with perturbation standard deviations $\sigma_{gauss}=24$ km and the 1 K and 0.1 K amplitudes.

The strength of the strongest error growth in each simulation can be seen in

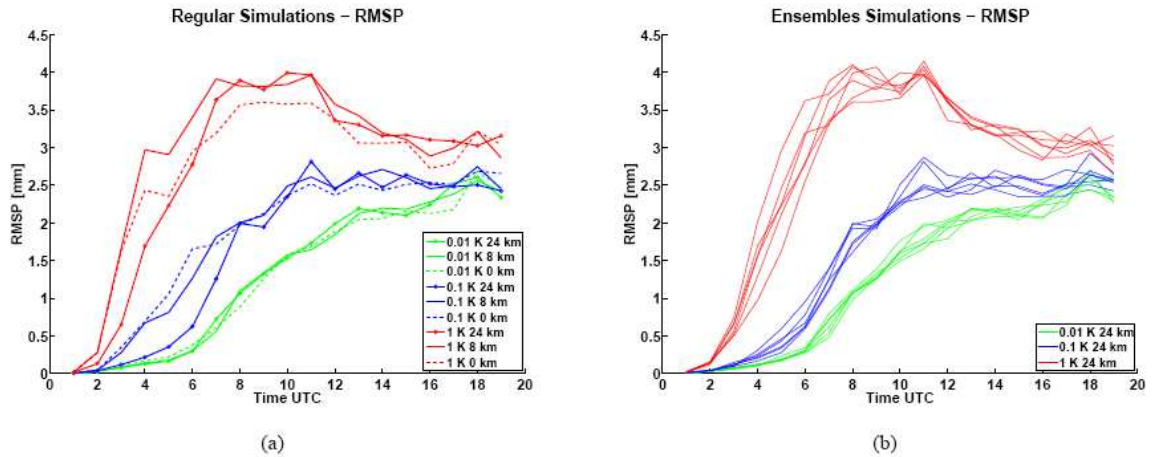


Figure 3.10: Evolution of the RMSD for (a) nine sequential-perturbation simulations with different perturbation amplitudes (0.01 K in green, 0.1 K in blue and 1 K in red) and standard deviations (24 km as solid lines with filled circles, 8 km as solid lines and 0 km as dashed lines) and (b) three ensembles with different perturbation amplitudes (0.01 K in green, 0.1 K in blue and 1 K in red), each with $\sigma_{gauss}=24$ km. All times refer to the beginning of the hour of accumulation.

Table 3.1, which gives the maximum three-hourly increments of RMSD. The RMSD is somewhat noisy and so it is convenient to apply a 1-2-1 filter to the increments; the relationship between error growth and perturbation amplitude is insensitive to the filtering. The maximum error growth is around twice as strong with the 1 K perturbations as with the 0.01 K perturbations. Furthermore when the same quantity is computed for the evolution of the ensembles plotted in Fig. 3.10b (not shown), the variations within each ensemble are smaller than the difference between the averages for each ensemble. Thus, at least for the $\sigma_{gauss}=24$ km simulations, the amplitude of the perturbation affects the RMSD more strongly than the random number sequence.

The domain-averaged precipitation accumulated over the full duration of the control run is 2.127 mm. The biases against that of the various perturbed simulations are given in Table 3.1. In most cases, the bias is at least two orders of magnitudes smaller

than the mean value, demonstrating that the total rainfall in this case study is primarily dictated by the large-scale convective forcing. While small-scale perturbations can alter the timing and location of particular storms, they do not affect the time-space averaged moisture budget.

The largest biases, of the order of a few percent, occur in the sequential perturbation experiments with 1 K perturbation amplitudes and the single perturbation experiment with the perturbation applied to the initial conditions (σ_{24A1} , σ_{8A1} , σ_{0A1} , and IC-1, Table 3.1). If the perturbation field is imposed on the initial conditions, the rainfall increases throughout the course of the simulation that follows, whereas in the sequential perturbation runs the total rainfall is reduced. The reduction occurs primarily between 5 and 15 UTC, somewhat offset by a positive bias later (not shown). These results highlight the point that the model is sensitive to strong perturbation amplitudes at early times, and also suggest that perturbations affecting the spin-up phase of the model can produce markedly different results, consistently with the changes in the cloud distribution (Section 3.6.b).

The RMSP of the single-perturbation simulations is shown in Fig. 3.11, along with that of the σ_{24A1} and $\sigma_{24A0.01}$ simulations for comparison. Generally the behavior for the single-perturbation simulations is similar to that for the sequential perturbations as RMSP grows from the time of the perturbation application and in some cases with the 1 K perturbation amplitude reach a clear saturation level. The IC-1 and IC-0.0.1 runs behave similarly in RMSP to their sequential-perturbation counterparts, σ_{24A1} and $\sigma_{24A0.01}$ (and the difference is within the spread of the ensemble generated by different the random number realizations (Fig. 3.10b). Note that, the different behavior of the IC-1 simulation to the σ_{24A1} simulation, as determined from the very different biases

in domain average accumulated precipitation (Table 3.1), is not revealed by the RMSP diagnostic. This indicates the importance of considering a range of diagnostics when assessing the impact of perturbations.

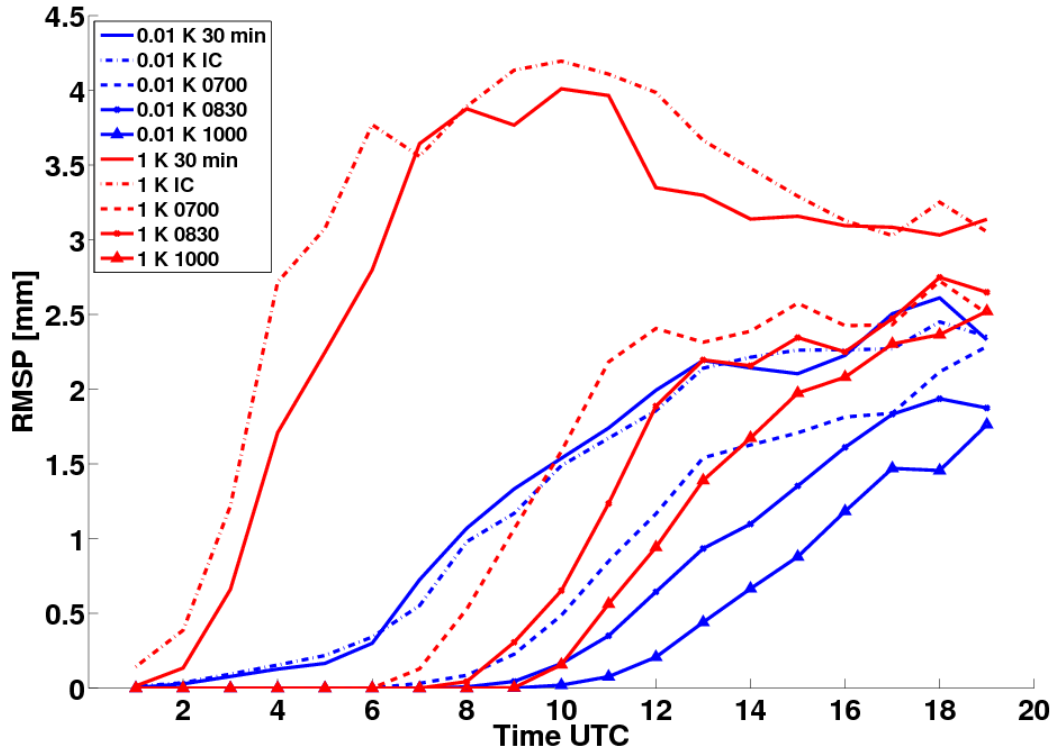


Figure 3.11: Evolution of the RMSP for the single perturbation experiments. The 0.01 K simulations have blue lines, while the 1 K have red lines. Continuous, unmarked lines represent the sequential simulations, while the dashed-dotted lines indicate the IC simulations. The dashed lines represent the 0700 UTC experiments, the solid lines with circles are for the 0830 UTC one and the 1000 UTC are represented by the solid lines with triangles.

In general the later the single perturbations are applied, the less likely the RMSP is to reach saturation and the less the RMSP is at the end of the simulation. The maximum growth rate of RMSP also reduces as the single perturbations are applied later in the day (Table 3.1). This shows that perturbing at early times is most effective in producing error growth.

3.6.d Intensity and Displacement Errors

In Section 3.4, a decomposition of the squared RMSP into the sum of three MSP components (*COMMON*, *CONTROL* and *PERTURBED*) was described (Equation 3.1). This decomposition for the sequential-perturbation simulations is given in Fig. 3.12, with MSP shown on the left and the fractions of each type of grid point (relative to the total number of rainy grid points) shown on the right. Note that the number of rainy points for the control simulation shown in Fig. 3.2 is the sum of the defined *COMMON* and *CONTROL* points here.

The points that are rainy in both the perturbed and control simulations (*COMMON* points) are considered first. The intensity of precipitation at *COMMON* points is altered in all of the perturbed simulations and most strongly for stronger perturbation amplitudes (Fig. 3.12a). This contribution dominates the total MSP at early times, the *COMMON* points having the fastest-growing MSP. Consistent with this observation, and with Fig. 3.10, the MSP for *COMMON* points grows more slowly and reaches a peak at later times for decreasing perturbation amplitudes. For a perturbation amplitude of 1 K, the fraction of *COMMON* points decreases from the outset of the simulations. From around 7 UTC most of the rainy grid points in the perturbation simulations differ from those in the control simulation (Fig. 3.12b). Thus, these perturbations are extremely effective from the outset at both displacing storms and altering the intensity of common storms. By contrast, at the same time in the simulations with weaker perturbation amplitudes the rain occurs in predominantly the same locations and at similar rates to the control simulation.

The points that are rainy only in the perturbed or control simulations are now considered. For the two smaller perturbation amplitudes these points exhibit similar

behavior. The simulations with 0.01 K perturbation amplitude start to generate points with a different rain status to the control simulation (i.e. raining in the perturbed run but not in the control or vice versa) around 1–2 h after such points are generated by 0.1 K perturbations. However, once produced the growth rates of the fraction of those points and the MSPs are similar, so that the same timing difference remains perceptible throughout the remainder of the simulations. With these perturbation amplitudes, the MSP contributions from *PERTURBED* and *CONTROL* rainy points to the total MSP are roughly equal.

By contrast, the simulations with 1 K perturbation amplitude have a different pattern of behavior for rainy but non-*COMMON* points. As seen in the 0.1 K and 0.01 K simulations, more storms are displaced earlier for a stronger perturbation amplitude. However, at early times the 1 K perturbations are more effective at triggering new storms than they are at suppressing storms seen in the control simulation. Thus, the fraction of points and MSP from *PERTURBED* points grows rapidly up to around 7 UTC (Figs. 3.12e and f). Beyond that time, the ability of the perturbations to trigger new storms, and the intensity of such storms, increases only slowly if at all. Interestingly also, the growth of the *CONTROL* fraction of points and MSP stalls at around the same time (Figs. 3.12e and f), indicating a reduced ability of the 1 K perturbations to alter storms. Thus, we can see that the period between 6 and 8 UTC is a critical one for the development of storms. It is during this time that perturbations of weaker amplitude first become effective at displacing storms. The 1 K perturbations meanwhile are extremely effective at producing additional storms, without greatly suppressing the triggering of storms in the control simulation (note the small MSP due to the *CONTROL* points prior to around 7 UTC).

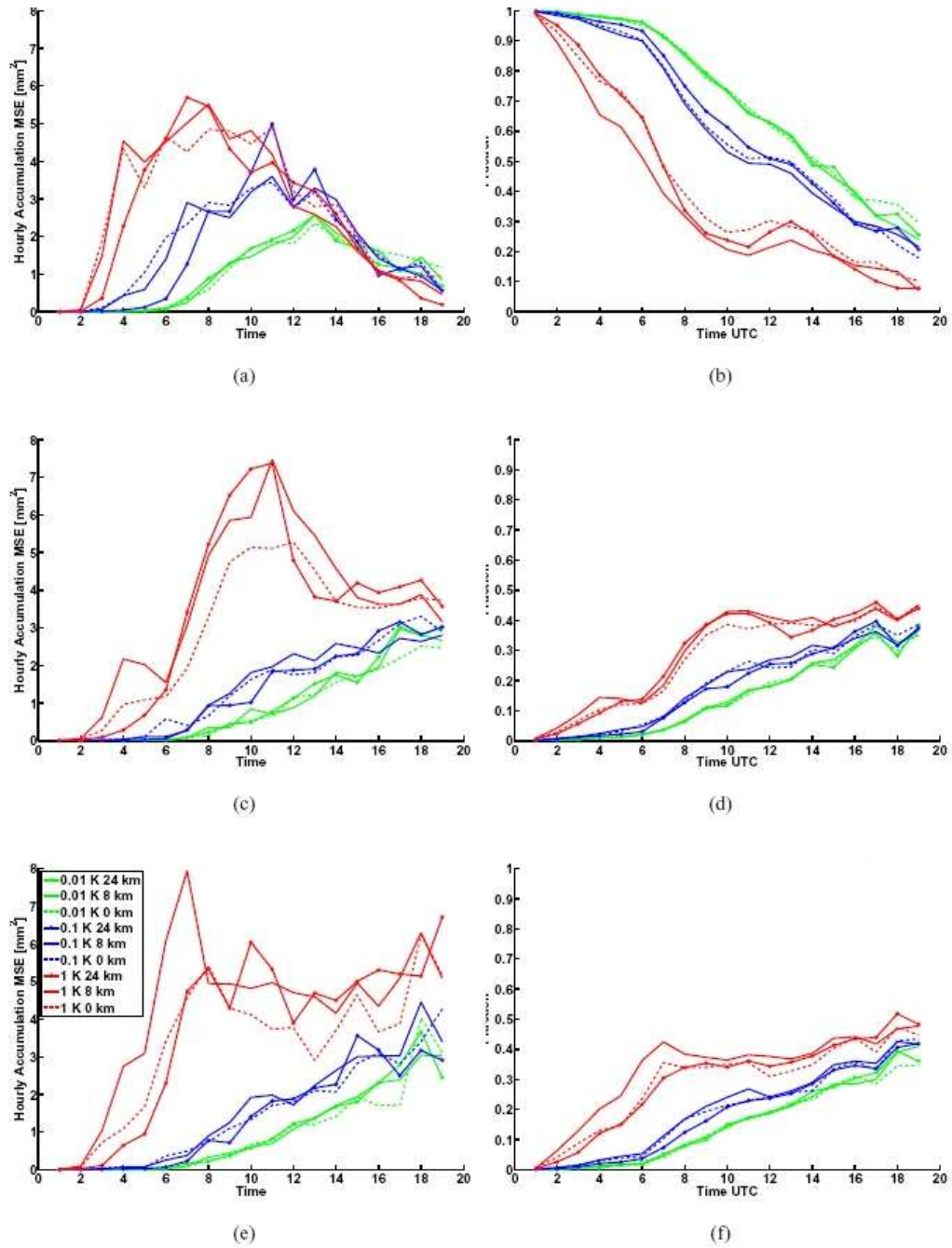


Figure 3.12: Panels in the left-hand column show the mean-square error for the rainy grid points which are classified as (a) *COMMON*, (c) *CONTROL* and (e) *PERTURBED*. At any time, the sum of the three, weighted by the domain fractions, gives the RMSPE, as shown in Fig. 10. Panels in the right-hand column show the fraction of the rainy part of the domain occupied by the (b) *COMMON*, (d) *CONTROL* and (f) *PERTURBED* points. In all panels, results are shown for nine sequential perturbation simulations with different perturbation amplitudes (0.01 K in green, 0.1 K in blue and 1 K in red) and standard deviations (24 km as solid lines with filled circles, 8 km as solid lines and 0 km as dashed lines).

Recalling the reduction in total rainfall at this time (Fig. 3.2), these results imply that the 1 K perturbations must be effective at reducing the strength of those storms that are in *COMMON* between the simulations. In essence, the strong perturbations produce more, but less intense, storms at this time. The storms present only in the control simulation are strongly affected by the 1 K perturbations during the late morning and early afternoon. From 1000 UTC their domain fraction remains constant or increases slightly, whereas the MSP decreases during the same period. Therefore, on average the intensity of the storms decreases. Comparing Figs. 3.12e and c we also note that the *PERTURBED* storms are stronger than the *CONTROL* in the afternoon and evening.

In general the MSP decomposition, is only slightly sensitive to the standard deviation of the perturbations σ_{gauss} , particularly so for the perturbation amplitude 0.01 K. However, for the larger amplitudes there are indications in the σ_{8A1} and $\sigma_{8A0.1}$ simulations that the 8 km standard deviation is consistently the most effective at displacing the storms.

3.7 Discussion

The key processes involved with the error growth have been analyzed. Some affect error growth by altering the background state of model (acoustic waves, CAPE, theta and total condensate changes) while changes in the boundary-layer types alter the model itself through by triggering different parameterizations and/or different coefficients for the parameterizations. Amongst the first group acoustic waves affect the background more widely since they propagate very quickly, although they are not very well represented. However they have been already observed to propagate errors by Hohenegger and Schar (2007b). They are also generated by convective storms (Nicholls

and Pielke 2000) and can further accelerate error growth if storms are displaced, similarly to gravity waves. Furthermore, because analysis increments can be larger than 1 K (e.g., Kong et al. 2007), if they are not acoustically balanced they can excite a wave response stronger than observed in this study.

It has been shown (Section 3.5.a) that each perturbation does not modify significantly the average CAPE of the domain, but at some points, for the 1 K perturbations lids can be removed or set, thus altering more effectively storm development and consequently error growth. Moreover, even in the absence of a trigger at these specific locations the vertical motions will be altered by the changes, albeit small, in the static stability.

The changes in the time-averaged boundary-layer covers introduced by the sequential perturbation (Table 3.3) involve more grid-points than the changes directly due to the perturbation as seen in the single perturbation runs (Table 3.2). This suggests that the boundary-layer changes of the sequential perturbation are due to the evolution of the perturbed model atmosphere (indirect effect of the perturbation) rather than the potential-temperature perturbation itself (direct effect).

While the changes in the boundary-layer types certainly affect error growth as suggested by Lean (2006), we did not attempt to quantify the relative importance of the different processes mentioned thus far, including the effect of theta itself and of total cloud water on the evolution of the model atmosphere.

No simulation was designed to specifically investigate the role of buoyancy (gravity) waves. It is well known that a potential-temperature perturbation induces gravity waves whose characteristics depend mainly on the vertical stability, on the duration and intensity of the heating, and on the size and aspect ratio of the heated region

(e.g., Chagnon and Bannon 2005). The examination of the sequentially perturbed potential temperature fields (output at every timestep) did not show any significant buoyancy wave activity at the perturbation level. At times, with no evident spatial pattern, oscillation of the potential-temperature followed the perturbation, but their frequency and amplitude vary greatly from grid point to grid point. Also many grid points do not have a clear oscillatory pattern and the perturbation decays smoothly to the control run values or simply the perturbed potential-temperature has tendency very similar to the control simulation. Furthermore no specific peaks were identified during the analysis of the potential-temperature spectra of the perturbed runs (Section 3.2). Finally, the small amplitude of the perturbation and the lack of correlation in time probably prevent resonance phenomena due to gravity waves, such as the one investigated by Robinson et al. (2008).

The absence of perturbation-generated buoyancy waves is probably due to numerous contributing factors. The horizontally heterogeneous shear and stability, which generate spatially incoherent responses, but also the very limited vertical extent and duration of the heating. For example, Chagnon and Bannon (2005) in their idealized study used a 1 K amplitude perturbation, but its vertical extent was at least 5 km. Robinson et al. (2008) to investigate the gravity wave response of surface hot spots used heating confined to the lowest model levels, but the maximum temperature perturbation reached 10 K spread over 1-2 h, whereas the depth of the perturbed layer in this study is 320 m and the perturbation is applied for one timestep only, which is 100 s. Furthermore, this specific value is not suitable to represent the slowest gravity wave modes whose period is $2\pi/N$ since N is of the order of 0.01 s. Such period can be reasonably expected and was observed in the idealized study of Chagnon and Bannon (2005). It also must be

mentioned that the long timestep of the model (100 s) filters out the short period gravity waves.

The above mentioned indirect effects brought about by the perturbations results in two main changes of the overall behavior of the simulations. First, the non-cirrus cloud distributions for the 1 K sequential perturbations are altered so that the linear regime which relates the mean cloud size to the mean cloud number changes significantly. This suggests that dynamics of such clouds is altered as the changes in mean size and cloud number are substantially larger than the ones due to the direct effects. We did not investigate how this change occurs. Secondly the perturbations, even the smaller ones, when applied every 30 minutes affect the storm location and intensity, as shown in Section 3.6.d. All the perturbations generate new storms and the MSE of the COMMON points peaks at later times as the perturbation decreases in amplitude. However, the 1 K perturbations behave differently as they start generating new storms earlier in the day, although they are less intense than the ones present only in the control run.

It has been shown that the 1 K perturbation experiments results in qualitatively different outcomes, not simply a more intense version of the smaller perturbation runs: the cloud distribution shifts to a different regime, the RMSP starts growing earlier and more vigorously and new storms are generated more effectively. Despite these important differences, both the sequential and single perturbation experiments (Section 3.6.c) reach similar RMSP over the course of the simulations (Fig. 3.10), even when only the initial conditions are perturbed. This could be specific to this case study which is strongly constrained by the large scale, but it seems to indicate that the nonlinearities of the system amplify the small perturbations more strongly. More accurate measures of spread

and error need to be implemented for a careful comparison with other studies similar studies such as Kong et al. (2006) or Hohenegger and Schar (2007b).

The time-scales of error growth are consistent with previous studies (e.g. Hohenegger and Schar 2007a), but at least in this analysis error growth rates depend on time of the day, amplitude and model balance. The dependence upon the time of the day is due to the transitions of the boundary layer: if the model state is perturbed before the morning transition from mainly stable to mainly mixed boundary-layer types then the perturbations will be amplified at higher rates than those generated after such transition. Secondly, the amplitude of the perturbation does affect such processes rendering it less dependent upon the time of the day since the strong growth rates of the RMSP for the 1 K sequential perturbation commence before the mentioned transition. However, it has been shown that the ability of the perturbation to affect new storms generation and change the intensity of the existing ones is different before and after such transition. Thirdly the model balance affects the error growth as the IC-1 experiments yields a very different forecast from the other 1 K single perturbation runs and from the sequential perturbations as well, indicating that the model during spin-up is more sensitive to the perturbation, which in turn can push the forecast on a different trajectory.

Finally the above mentioned sensitivity to the time of the day suggests that the turbulent plumes rising from the boundary layer find a slightly different environment and amplify the existing differences once moist convection starts. This interpretation is supported by the domain average mass flux at perturbation level (Fig. 3.13) which starts growing at 0600 UTC, peaks at 1300 UTC and then decreases afterwards. The early part of the increase in the mass flux is associated to transition in the boundary-layer types observed for the control run (Section 3.2.c) and also the 6 to 8 UTC period is critical for

the development of storms (Section 3.6.d). The decrease of the average mass flux from 1300 UTC onwards helps explaining the reduction of error growth rates for both the sequential perturbation and the single ones (Section 3.6.c) which happens roughly at the same time. Error growth can continue only at slower pace because, even though the average mass flux is diminishing, there are still storms in the domain and also the non-convective evolution of the flow leads to error growth. Within this framework, the larger error growth rates of the stronger amplitude perturbation runs are explained quite simply by the stronger initial differences which undergo nonlinear growth. Furthermore, the earlier the simulation is perturbed, the earlier the convective processes can act to generate error growth as seen in Section 3.6.c. This effect coupled with the smaller number of rainy points in the control run (Fig. 3.2) generates the early and swift error growth of the 1 K sequential perturbation.

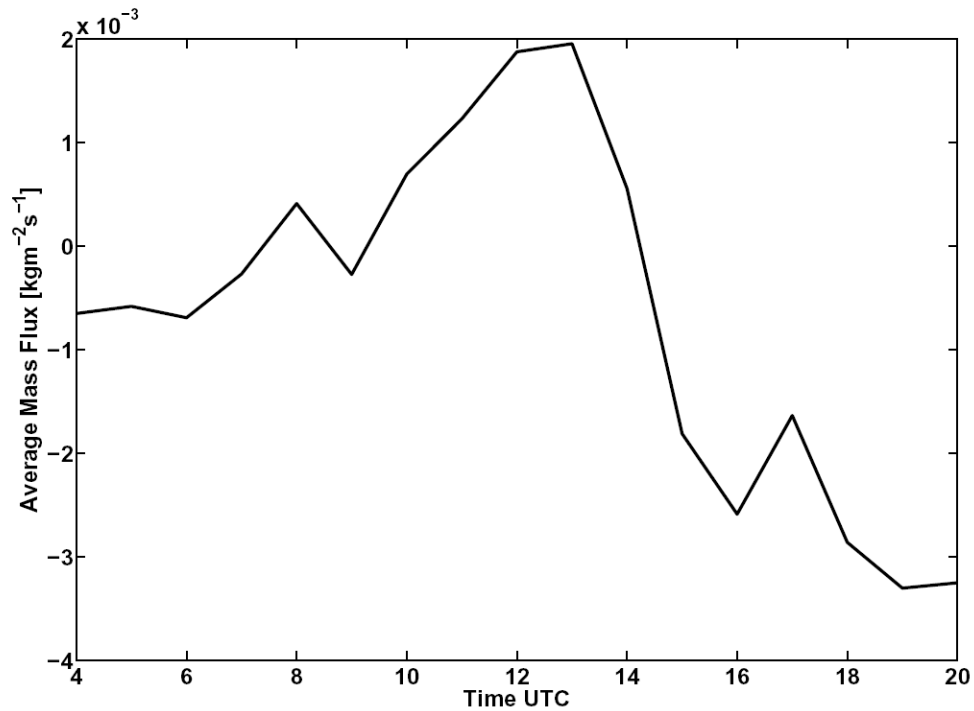


Figure 3.12: Horizontally averaged mass flux at the perturbation level. The values during the model spin-up are not shown.

3.8 Summary and Conclusions

The processes leading to convective-scale error growth and its sensitivity to the perturbation characteristics have been investigated for a CSIP case study. The case was chosen because it is strongly upper-level forced but with the detailed mesoscale/convective-scale evolution dependent on smaller-scale processes. The potential temperature was perturbed at a fixed model level within the boundary layer, most of the times above the LCL. Sensitivity studies show that RMSP is fairly insensitive to the height of the perturbation. The perturbation fields had varying horizontal length scales and amplitudes and both sequential perturbations (applied every 30 min throughout the run and uncorrelated in time) and single perturbations (at a specific time) were considered.

Diagnostics were carefully selected to elucidate both the direct effect (within one timestep) and indirect effect (as evolved by the model) of the perturbations on the model. In particular we have developed diagnostics to distinguish changes in precipitation intensity from changes in the location and distribution of clouds. This relates to the hydrological focus of the funding for this study (FREE).

The processes involved are the acoustic and Lamb waves, the changes in the boundary-layer type categorization, the changes in the CAPE, and in the total water condensate and theta. The acoustic waves affect quickly the whole domain, creating a different background. The changes of the potential temperature and in the total condensate act in a similar fashion, although they are limited to the perturbed level and its two neighbors. The perturbation affects only slightly the CAPE values of the individual grid points, but for the 1 K perturbations it creates or destroys lids, creating the conditions for changes in storm location and development favoring error growth. The

changes in the boundary-layer types affect error growth by directly altering the response of the boundary layer.

Overall, the amplitude of the perturbations (single and sequential) is the main driver of the error response, with the scale-length occupying a secondary role in modulating the response of the RMSP. However smaller-scale lengths for the 1 K sequential perturbations result in a larger number of clouds (both non-cirrus and precipitating ones) with a smaller average size. Storm displacement and generation, as defined in Section 3.6.d, is also affected more strongly by the 1 K sequential perturbation as they both displace storms and create new ones earlier in the day. Therefore the sequential perturbations result in qualitatively different simulations, although the RMSP is of the same magnitude as the smaller perturbations. This indicates, on the one hand, that the root mean square error is a fairly crude indicator of the error because it is not sensitive to important features, but on the other hand the nonlinearities of the atmosphere are such that the much smaller amplitudes are amplified more strongly. These features of the IC-1 experiment, where a 1 K perturbation was applied at the initial conditions, are even more markedly different from the remaining experiments and indicates that, at least for this specific event, the perturbation pushed the model on a different trajectory. Error growth has also been observed to be sensitive to the time of day, or better, to the state of the boundary layer since perturbations applied before the morning transition of the boundary layer are more effective at generating error growth. This, along with the different features of the sequential perturbation experiments, shows that the sequential perturbation is an effective method of generating error growth.

CHAPTER 4

CONCLUSIONS

In Chapter 2, the Harrington radiation parameterization was transformed into a transfer scheme for the clear sky. The transfer scheme requires a fraction of the computational cost of the parent parameterization. Accuracy is generally preserved, although at times the heating rates error is larger than what is normally accepted for a parameterization compared against a line-by-line code. However these larger errors occur near the tropopause where the model vertical resolution is coarsest and when heating rates are converted into heating flux via divergence they are of the same magnitude of those of Zhang et al. (2000) which used a 1 km vertical grid spacing for its radiative transfer scheme. Furthermore, after two days of simulations the RAMS meteorological fields (surface temperature, 500 mb geopotential height, etc.) obtained with the best transfer scheme yields results which are not significantly different from the parent scheme. More importantly they are physically realistic and a comparison of the same fields against the one obtained with the Chen-Cotton parameterization shows that the transfer scheme modifies the main meteorological fields less than the use of this second radiation parameterization.

Future work will focus on extending this methodology to a cloudy atmosphere. The first step is to use a modified version of the EOF analysis so that it does not result in negative cloud water content.

The processes that led to the error growth of the hourly-accumulated precipitation fields during a case study over southern England were analyzed in Chapter 3. The single perturbation experiments where the potential temperature perturbation is applied only once at specific times show that besides the expected changes in total cloud water content, three processes are involved in error generation: Lamb waves generated by the perturbation, for the stronger perturbation (1 K) small temperature inversions can be removed or set at the perturbation level, and also the categorization of the boundary layer is altered, affecting the boundary-layer parameterization. The first two processes affect the background state while the third alters the mixing coefficients and/or calls to other parameterizations (e.g., entrainment). The Lamb wave allows a quick propagation of the error throughout the whole domain, up to the stratospheric model levels, and in the case of the 1 K perturbation has a significant effect on the average vertical velocity. Hohenegger and Schar (2007) also observed acoustic waves as result of temperature perturbation. The changes in static stability can affect storm development and evolution, but they require the simultaneous presence of a trigger. The changes brought to the boundary-layer parameterization are similar to those observed, within the same Met Office model by Lean (2006). No attempt was made to quantify the influence of these processes on the error growth, but effects of the application of the perturbation every 30 minutes (sequential perturbation) yield in markedly different results depending mainly on perturbation intensity, but also its typical scale-length. The strongest perturbations (1 K) experiments have stronger error growth from the outset of the simulation and start generating new storm earlier in the day than the smaller amplitude counterparts (0.01 and 0.1 K). Also the cloud distributions are affected and the 1 K perturbations show an increase in the average number of clouds as the typical scale length decreases. The

smaller perturbations also have a linear relationship between average size and number of clouds, but with a significantly different slope.

The sequential perturbation method is overall successful at generating qualitatively different simulations, as shown above, but it also allows for of capture the sensitivity of the error growth to the perturbation to the time of the day. The perturbations applied before the morning transition of the boundary-layer types, cause a quicker error growth than those applied afterwards.

Future work will concentrate on three fundamental issues. First, to determine the relative importance of the different error sources relative to the error growth, as it is defined in this study. Secondly, it is important to test these conclusions on other case studies to test their significance towards the design of a convective-scale ensemble. Finally, while the hydrological focus of this study is very important it is also interesting to investigate the dynamical features induced by the sequential perturbation following up on the changes in the non-cirrus cloud distribution. This involves developing different measures of the error, possibly accounting for the location and timing errors of the convection

REFERENCES

- Arakawa A., 1993: Cumulus assumptions in the cumulus parameterization problem. *The representation of cumulus convection in numerical models, Meteor. Monogr.*, N. 46, Amer. Meteor. Soc., 1-16.
- Bennett, L. J., K. A. Browning, A. M. Blyth, D. J. Parker, and P. A. Clark, 2006: A review of the initiation of precipitating convection in the United Kingdom. *Q. J. R. Meteorol. Soc.*, **132**, 1001–1020.
- Biggerstaff, M. I., E. K. Seo, S. M. Hrstove-Veleva, K. Y. Kim, 2006: Impact of cloud model microphysics on passive microwave retrievals of cloud properties. Part I: model comparison using EOF analyses. *J. Appl. Meteor. Climatol.*, **45**, 930-954.
- Browning, K. A. and C. J. Morcrette, 2006: A summary of the Convective Storm Initiation Project Intensive Observation Periods. Tech. Rep. 474, MetOffice R&D. [Available on line at: http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html].
- Browning, K. A., et al., 2007: The Convective Storms Initiation Project. *Bull. Amer. Meteor. Soc.*, **88**, 1939–1955.
- Bryan, G. H., J. Wyngaard, and J. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Mon. Wea. Rev.*, **131**, 2394–2416.
- Casati, B., et al., 2008: Forecast verification: current status and future directions. *Meteorol. Appl.*, **15**, 3–18.
- Castro, C.L., W.Y.Y. Cheng, A.B. Beltrán, R.A. Pielke Sr., and W.R. Cotton, 2002: The incorporation of the Kain-Fritsch cumulus parameterization scheme in RAMS

with a terrain-adjusted trigger function. *Fifth RAMS Users and Related Applications Workshop*, Santorini, Greece, ATMET, Inc.

Castro, C.L., 2005: Investigation of the Summer Climate of North America: A Regional Atmospheric Modeling Study. Ph.D. Dissertation, Colorado State University, 223 pp.

Castro, C.L., R.A. Pielke Sr., and J. Adegoke, 2007a: Investigation of the summer climate of the contiguous U.S. and Mexico using the Regional Atmospheric Modeling System (RAMS). Part I: Model climatology (1950-2002). *J. Climate*, 20, 3844-3865 .

Castro, C.L., R.A. Pielke Sr., J. Adegoke, S.D. Schubert, and P.J. Pegion, 2007b: Investigation of the summer climate of the contiguous U.S. and Mexico using the Regional Atmospheric Modeling System (RAMS). Part II: Model climate variability. *J. Climate*, 20, 3866-3887.

Chagnon, J. M. and P. R. Bannon, 2005: Wave response during hydrostatic and geostrophic adjustment. Part I: transient dynamics. *J. Atmos. Sci.*, **62**, 1311–1329.

Chen C., and W. R. Cotton, 1987: The physics of the marine stratocumulus-capped mixed layer. *J. Atmos. Sci.*, **44**, 2951-2977.

Chevallier, F., F. Chérury, N. A. Scott and A. Chédin, 1998: A neural network approach for a fast and accurate computation of a longwave radiative budget. *J. Appl. Meteor.* **37**, 1385–1397.

Clark, P. A. and H. W. Lean, 2006: An overview of high resolution um performance for CSIP cases. Tech. Rep. 478, MetOffice R&D. [Available on line at: http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html].

- Collier, C. G., 2006: Flash flood forecasting: What are the limits of predictability? *Q. J. R. Meteorol. Soc.*, **133**, 3–23.
- Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785.
- Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the met office’s global and regional modelling of the atmosphere. *Q. J. R. Meteorol. Soc.*, **131**, 1759–1782, 10.1256/qj.04.101.
- Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **6**, 249–266.
- Done, J. M., G. C. Graig, S. L. Gray, and P. A. Clark, 2008: Case-to-case variability of predictability of deep convection in a mesoscale model. *Q. J. R. Meteorol. Soc.*, **submitted**.
- Dong, X., B. Xi, and P. Minnis, 2006: A climatology of midlatitude continental clouds from the ARM SCP central facility. Part II: Cloud fraction and surface radiative forcing. *J. Climate*, **19**, 1765–1783.
- Ebert, E. E. and J. L. McBride, 2000: Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrology*, **239**, 179–202.
- Ebisuzaki, W. and E. Kalnay, 1991: Ensemble experiments with a new lagged analysis forecasting scheme. research activities in atmospheric and oceanic modeling. Tech. Rep. 15, WMO.
- Edwards, J. and A. Slingo, 1996: Studies with a flexible new radiation code. Part I: Choosing a configuration for a large-scale model. *Q. J. R. Meteorol. Soc.*, **122**, 689–719.

- Essery, R., M. Best, and P. Cox, 2001: MOSES 2.2 Technical Documentation. Tech. Rep. 30, Hadley Centre. [Available on line at: <http://www.metoffice.gov.uk/research/hadleycentre/pubs/HCTN/index.html>].
- Fabry, F., 2006: The spatial variability of moisture in the boundary layer and its effect on convection initiation: Project-long characterization. *Mon. Wea. Rev.*, **134**, 79–91.
- Fu, Q., and K. N. Liou, 1992: On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. *J. Atmos. Sci.*, **49**, 2139–2156.
- Gabriel, P., G. L. Stephens, I. L. Wittmeyer, 2000: Adjoint perturbation and selection rule methods for solar broadband two-stream fluxes in multi-layer media. *J. Quant. Spectroscopy & Radiative Transfer*, **65**, 693–728.
- Gabriel, P. M., P. T. Partain, and G. L. Stephens, 2001: Transfer. Part II: Selection Rules. *J. Atmos. Sci.*, **58**, 3411–3423.
- Grabowski, W. W., et al., 2006: Daytime convective development over land: A model intercomparison based on LBA observations. *Quart. J. Roy. Meteor. Soc.*, **132**, 317–344, 10.1256/qj.04.147.
- Gregory, D. and P. R. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-depend closure. *Mon. Wea. Rev.*, **118**, 1483–1506.
- Hand, W. H., N. I. Fox, and C. G. Collier, 2004: A study of the twentieth-century extreme rainfall events in the United Kingdom with implications for forecasting. *Meteorol. Appl.*, **11**, 15–31.
- Harrington, J. Y., 1997: The effects of radiative and microphysical processes on simulated warm and transition season Arctic stratus. Colorado State University, 289 pp.

- Harrington, J. Y., M. P. Meyers, W. R., Cotton S. M Kreidenweis, 1999: Cloud resolving simulations of Arctic stratus. Part II: Transition season clouds. *Atmos. Res.*, **55**, 45-75.
- Hohenegger, C. and C. Schar, 2007a: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88**, 1783–1793.
- Hohenegger, C. and C. Schar, 2007b: Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.*, **64**, 4467–4478, 10.1175/2007JAS2143.1.
- Hohenegger, C., D. Luthi, and C. Schar, 2008a: Predictability mysteries in cloud-resolving models. *Mon. Wea. Rev.*, **134**, 2095–2107.
- Hohenegger, C., A. Walser, W. Langhans, and C. Schar, 2008b: Cloud-resolving ensemble simulations of the august 2005 alpine flood. *Q. J. R. Meteorol. Soc.*, **134**, 889–904, 10.1002/qj.252.
- Josey, S. A., R. W. Pascal, P. K. Taylor, and M. J. Yelland, 2003: A new formula for determining the atmospheric longwave flux at the ocean surface at mid-high latitudes, *J. Geophys. Res.*, **108 (C4)**, 3108, doi:10.1029/2002JC001418.
- Kain, J. S., and J.M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. *The Representation of Cumulus Convection in Numerical models*, 165-170.
- Kaiser, H. F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika.*, **23**, 187-200.
- Kong, F., K. K. Droegemeier, and N. L. Hickmon, 2006: Multiresolution ensemble forecasts of and observed tornadic thunderstorm system. Part I: Comparison of coarse- and fine-grid experiments. *Mon. Wea. Rev.*, **134**, 807–833.

- Kong, F., K. K. Droegemeier, and N. L. Hickmon, 2007: Multiresolution ensemble forecasts of and observed tornadic thunderstorm system. Part II: Storm scale experiments. *Mon. Wea. Rev.*, **135**, 759–782.
- Kravtsov, S., A. W. Robertson, and M. Ghil, 2006: Multiple regimes and low-frequency oscillations in the Northern Hemisphere's zonal-mean flow. *J. Atmos. Sci.*, **63**, 840-860.
- Lean, H. W., S. P. Ballard, P. A. Clark, M. A. Dixon, Z. Li, and N. M. Roberts, 2005: The summer 2004 reruns with the High Resolution Trial Model. Tech. Rep. 466, MetOffice R&D. [Available on line at: http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html].
- Lean, P. W., 2006: The predictability of convective storms over the ocean. Phd thesis, The University of Reading, Department of Meteorology, The University of Reading, Reading, UK, RG6 6BB.
- Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith, 2000: A new boundary layer mixing scheme. Part I: scheme description and single-column model tests. *Mon. Wea. Rev.*, **128**, 3187–3199.
- Liu, W. T., W. Tang, and P. P. Niiler, 1991: Humidity profiles over the ocean. *J. Climate*, **4**, 1023-1034.
- Lorenz, E.N., 1956: Empirical orthogonal function and statistical weather prediction. Science Report 1, Statistical Forecasting Project, Department of Meteorology, MIT (NTIS AD 110268), 49 pp.
- Lorentz, E.N., 1977: An experiment in nonlinear statistical weather forecasting. *Mon. Wea. Rev.*, **105**, 590-602.

- Majewski, D., D., Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner, 2002: The Operational Global Icosahedral-Hexagonal Gridpoint Model GME: Description and high-resolution tests. *Mon. Wea. Rev.* **130**, 319–338.
- Mahrer, Y., and R.A. Pielke, 1977: A numerical study of the airflow over irregular terrain. *Beitrage zur Physik der Atmosphere*, **50**, 98-113.
- Mesinger F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jović, J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343-360.
- Molteni, F., R. Buizza, C. Marisgli, A. Montani, F. Nerozzi, and T. Paccagnella, 2001: A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments. *Quart. J. Roy. Meteor. Soc.*, **121**, 2069–2094.
- Natraj, V., X. Jiang, R. Shia, X. Huang, J. S. Margolis, and Y. L. Yung, 2005: Application of principal component analysis to high spectral resolution radiative transfer: A case study of the O₂ A band. *J. Quant. Spectroscopy & Radiative Transfer*, **95**, 539-556.
- Nicholls, M. E. and R. A. Pielke, 2000: Thermally induced compression waves and gravity waves generated by convective storms. *J. Atmos. Sci.*, **57**, 3251–3271.
- Nieuwstadt, F. T. M. and R. A. Brost, 1986: The decay of convective turbulence. *J. Atmos. Sci.*, **43**, 532–546.
- North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, **110**, 699–706.

- Oppenheim, A.V., R.W. Schafer, and J. R. Buck, 1989: *Discrete-Time Signal Processing*, Prentice-Hall, 1999, 870 pp.
- Petch, J. C., 2004: The predictability of deep convection in cloud-resolving simulations over land. *Quart. J. Roy. Meteor. Soc.*, **130**, 3173–3287.
- Pielke Sr., R. A., T. Matsui, G. Leoncini, T. Nobis, U. Nair, E. Lu, J. Eastman, S. Kumar, C. Peters-Lidard, Y. Tian, and R. Walko, 2006: A new paradigm for parameterizations in numerical weather prediction and other atmospheric models. *National Weather Digest*, **30**, 93-99.
- Quadrelli, R., C. S. Bretherton, and J. M. Wallace, 2005: On Sampling Errors in Empirical Orthogonal Functions. *J. Climate*, **18**, 3704-3710.
- Roberts, N. M., 2003: Stage 2 report form the Storm-scale numerical modelling project. Tech. Rep. 407, MetOffice R&D. [Available online at: http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/index.html].
- Roberts, N. M. and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97.
- Robinson, F. J., S. C. Sherwood, and Y. Li, 2008: Resonant response of deep convection to surface hot spots. *J. Atmos. Sci.*, **65**, 276–286.
- Schubert, S.D., M.J. Suarez, P.J. Pegion, and M.A. Kistler, 2002: Predictability of Zonal Means during Boreal Summer. *J. Climate*, **15**, 420-434.
- Smagorinsky, J., 1963: General circulation experiments with the primitive equations. Part I, the basic experiment. *Mon. Wea. Rev.*, **91**, 99-164.

- Snyder, C. and F. Zhang, 2001: Assimilation of simulated doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677.
- Storch, H. V., and Hannoschöck, 1985: Statistical aspects of estimated principal vectors (EOFs) based on small sample sizes. *J. Appl. Meteor.*, **24**, 716–724.
- Stull, R. B., 1988: *An introduction to boundary layer meteorology*. 1st ed., Kluwer Academic Publishers, pp. 670.
- Thompson, D. W. J., and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: month-to month variability. *J Climate*, **13**, 1000-1016.
- Walko, R.L., L.E. Band, J. Baron, T.G.F. Kittel, R. Lammers, T.J. Lee, D.S. Ojima, R.A. Pielke, C. Taylor, C. Tague, C.J. Tremback, and P.L. Vidale, 2000: Coupled atmosphere-biophysics-hydrology models for environmental modeling. *J. Appl. Meteor.* **39**, 931-944.
- Walser, A., D. Luth, and C. Schar, 2004: Predictability of precipitation in a cloud-resolving model. *Mon. Wea. Rev.*, **132**, 560–577.
- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437.
- Wilks, D. S., 2006: *Statistical methods for the atmospheric sciences*. 2nd Edition, Academic Press, 648 pp.
- Wilson, D. R. and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. *Q. J. R. Meteorol. Soc.*, **125**, 1607–1636.

- Xu, K. M., and D. A. Randall, 1995: Impact of interactive radiative transfer on the macroscopic behavior of cumulus ensembles. Part I: radiation parameterization and sensitivity tests. *J. Atmos. Sci.*, **52**, 785-799.
- Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173–1185.
- Zhang, H., T. Nakajima, G. Shi, T. Suzuki, and R. Imasu, 2003: An optimal approach to overlapping bands with correlated k distribution method and its application to radiative calculations. *J. Geophys. Res.*, **10(D20)**, doi:10.1029/2002JD003358.