

FINNISH METEOROLOGICAL INSTITUTE  
CONTRIBUTIONS

No. 106

PARAMETRIC UNCERTAINTY IN NUMERICAL WEATHER  
PREDICTION MODELS

Pirkka Ollinaho

Department of Physics  
Faculty of Science  
University of Helsinki  
Helsinki, Finland

ACADEMIC DISSERTATION in meteorology

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium Physicum D101 (Gustaf Hällströmin katu 2b) on April 10th, 2014, at 12 o'clock noon.

Finnish Meteorological Institute  
Helsinki, 2014

Author's Current Address: Finnish Meteorological Institute  
Climate Research  
P.O. BOX 503  
FI-00101 Helsinki, Finland  
e-mail [pirkka.ollinaho@fmi.fi](mailto:pirkka.ollinaho@fmi.fi)

Supervisors: Professor Heikki Järvinen, Ph. D.  
Department of Physics  
University of Helsinki, Helsinki, Finland

Researcher Marko Laine, Ph. D.  
Earth Observation  
Finnish Meteorological Institute, Helsinki, Finland

Pre-examiners: Assistant Professor Derek J. Posselt, Ph. D.  
Atmospheric, Oceanic, and Space Sciences  
University of Michigan, Ann Arbor, USA

Researcher Thorsten Mauritsen, Ph. D.  
Climate Dynamics  
Max-Planck-Institut für Meteorologie, Hamburg, Germany

Opponent: Professor Jón Egill Kristjánsson, Ph. D.  
Department of Geosciences, Meteorology and Oceanography Section  
University of Oslo, Oslo, Norway

Custos: Professor Heikki Järvinen, Ph. D.  
Department of Physics  
University of Helsinki, Helsinki, Finland

ISBN 978-951-697-823-2 (paperback)  
ISSN 0782-6117  
Unigrafia Oy  
Helsinki, 2014

ISBN 978-951-697-824-9 (pdf)  
<http://ethesis.helsinki.fi>  
Helsinki, 2014  
Helsingin yliopiston verkkojulkaisut



Published by Finnish Meteorological Institute

P.O. Box 503  
FIN-00101 Helsinki, FinlandSeries title, number and report code of publication  
Finnish Meteorological Institute  
Contributions 106, FMI-CONT-106Date  
March 2014

Author

Pirkka Ollinaho

Name of project

Commissioned by

Title

Parametric uncertainty in numerical weather prediction models

Abstract

Numerical Weather Prediction (NWP) models form the basis of weather forecasting. The accuracy of model forecasts can be enhanced by providing a more accurate initial state for the model, and by improving the model representation of relevant atmospheric processes. Modelling of subgrid-scale physical processes causes additional uncertainty in the forecasts since, for example, the rates at which parts of the physical processes occur are not exactly known. The efficiency of these sub-processes in the models is controlled via so called closure parameters. This thesis is motivated by a practical need to estimate the values of these closure parameters objectively, and to assess the uncertainties related to them.

In this thesis the Ensemble Prediction and Parameter Estimation System (EPPES) is utilised to determine the optimal closure parameter values, and to learn about their uncertainties. Closure parameters related to convective processes, formation of convective rain and stratiform clouds are studied in two atmospheric General Circulation Models (GCM): the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) and the ECMWF model HAMBURG version (ECHAM5).

The parameter estimation is conducted by launching ensembles of medium range forecasts with initial time parameter variations. The fit of each ensemble member to analyses is then evaluated with respect to a target criterion, and the likelihoods of the forecasts are discerned. The target criterion is first set to be 500 hPa level geopotential height Mean Squared Error (MSE) at forecast days three and ten. After the proof of concept experimentations, the use of total energy norm as the target criterion is explored. EPPES estimation with both likelihoods results in parameter values converging to more optimal values during a three-month sampling period. The improved forecast accuracy of the models with the new parameter values are verified through headline skill scores (Root Mean Square Error (RMSE) and Anomaly Correlation Coefficient (ACC)) of 500 hPa geopotential height and a scorecard consisting of multiple model fields.

The sampling process also provides information about parameter uncertainties. Three uses for the uncertainty data are highlighted: (i) parametrization deficiencies can be identified from large parameter uncertainties, (ii) parameter correlations can indicate a need for the coupling of parameters, and (iii) adding parameter variations into an ensemble prediction system (EPS) can be used to increase the ensemble spread.

The relationship between medium range forecasts and model climatology is explored, too. Closure parameter modification induced cloud cover changes at forecast day three carry over to the very long range forecasts as well. This link could be used to improve model climatology by enhancing the computationally cheaper medium range forecast skill of the model.

Publishing unit

Finnish Meteorological Institute, Climate Research Unit

Classification (UDC)

551.509.313

551.509.313.4

519.2.222

Keywords

numerical weather prediction, physical  
parametrization, parameter estimation,  
climate modelling

ISSN and series title

0782-6117 Finnish Meteorological Institute Contributions

ISBN

978-951-697-823-2 (paperback), 978-951-697-824-9 (pdf)

Language

English

Pages

123

Price

Sold by

Note

Finnish Meteorological Institute / Library

P.O. Box 503, FIN-00101 Helsinki, Finland



ILMATIETEEN LAITOS

Julkaisija

Ilmatieteen laitos

PL 503, 00101 Helsinki

Julkaisun sarja, numero ja raporttikoodi  
Finnish Meteorological Institute  
Contributions 106, FMI-CONT-106Julkaisu-aika  
Maaliskuu 2014

Tekijä

Pirkka Ollinaho

Projektin nimi

Toimeksiantaja

Nimike

**Numeeristen sääennustemallien parametrinen epävarmuus**

Tiivistelmä

Jokapäiväisten sääennusteiden pohjana ovat numeeristen sääennustemallien tuottamat ennusteet ilmakehän tulevasta tilasta. Mallien ennustetarkkuutta voidaan parantaa tarkentamalla mallille syötettävää ilmakehän alkutilaa tai mallintamalla ilmakehän ilmiöt realistisemmin. Hilaväliä pienempien ilmiöiden kuvaaminen malleissa tuottaa ennusteisiin oman epävarmuutensa, mm. koska näihin ilmiöihin liittyvien prosessien tehokkuutta ei tiedetä tarkasti. Malleissa näiden aliprosessien nopeutta säädellään ns. sulkuparametrien kautta. Tämän väitöskirjan tavoitteena on sulkuparametrien arvojen objektiivinen valinta sekä niihin liittyvien epävarmuuksien selvittäminen.

Tässä väitöskirjassa parametrien optimaalisten arvojen ja niiden epävarmuuksien estimointi suoritetaan EPES (Ensemble Prediction and Parameter Estimation System) -algoritilla. Konvektioon, konvektiiviseen sateeseen ja kerrospilvien muodostumiseen liittyviä parametreja tutkitaan kahdella globaalilla ilmakehämallilla: Euroopan keskipitkien sääennusteiden keskuksen (ECMWF) IFS (Integrated Forecasting System) -sääennustemallilla ja ECHAM5 (ECMWF model HAMburg version) -ilmastomallilla.

Parametrien estimointia varten niiden arvoja muunnellaan keskipitkien sääennusteiden ryväsennustejärjestelmässä. Jokaisen ryppään jäsenen ennustetta verrataan analyysikenttään ja ennusteen osuvuus mitataan ennalta määrätyllä kohdefunktiolla. Kohdefunktiona käytetään ensimmäiseksi 500 hPa painepinnan geopotentialkorkeuden MS (Mean Squared) -virhettä kolmen ja 10 päivän sääennusteissa ja EPES-algoritmin todetaan toimivan halutulla tavalla. Tämän jälkeen kohdefunktioksi vaihdetaan ilmakehän kokonaisenergianormi. Kolmen kuukauden otannoissa kummatkin käytetyt kohdefunktiot johtavat parametrien konvergoitumiseen optimoituihin arvoihin. Uusien parametrien todetaan parantavan ennusteita käyttäen validointimenetelminä 500 hPa painepinnan geopotentialkorkeudella RMSE (Root Mean Squared Error) ja ACC (Anomaly Correlation Coefficient) arvoja sekä laajoja mallien vertailutulokuita.

Estimoinnin aikana saadaan myös lisää tietoa parametreihin liittyvistä epävarmuuksista. Kolme käyttötarkoitusta nostetaan esiin: (i) suuret epävarmuudet parametreissa viittaavat puutteisiin parametrisaatioissa, (ii) voimakkaat parametrien korrelaatiot ilmaisevat tarpeesta parametrien yhdistämiseksi ja (iii) parametriveriaatioiden lisääminen ryväsennustejärjestelmään kasvattaa järjestelmän ryvähajontaa.

Viimeiseksi selvitetään yhteyttä keskipitkien ennusteiden ja mallin klimatologian välillä. Parametrien vaihtamisen aiheuttamien pilvisyyden muutosten rakenne kolmen päivän ennusteissa on havaittavissa myös mallin pitkäaikaisissa vuosittaisennusteissa. Näin ollen mallin klimatologiaa voisi parantaa myös tarkentamalla mallin ennustekykyä laskennallisesti halvemmissa keskipitkissä sääennusteissa.

Julkaisijayksikkö

Ilmastotutkimus

Luokitus (UDK)

551.509.313

551.509.313.4

519.2.222

ISSN ja avainnimike

0782-6117 Finnish Meteorological Institute Contributions

ISBN

978-951-697-823-2 (paperback), 978-951-697-824-9 (pdf)

Kieli

englanti

Myynti

Ilmatieteen laitos / Kirjasto

PL 503, 00101 Helsinki

Asiasanat

numeerinen sääennustusmalli, parametrisaatiot,  
parametri estimointi, ilmastomallitus

Sivumäärä

123

Hinta

Lisätietoja

Parametric Uncertainty in  
Numerical Weather Prediction Models



## PREFACE

The work presented in this thesis has been carried out at the Climate Research Unit of the Finnish Meteorological Institute (FMI) during the period 2010–2014. I had the pleasure of doing this thesis as part of a research consortium called Novel Advanced Mathematical and Statistical Methods for Understanding Climate (NOVAC). My work was funded by the Academy of Finland through the NOVAC project. The members of the consortium, led by Prof. Heikki Haario, provided invaluable help in deciding the direction of my research.

I wish to express my gratitude to all my co-authors, with your input and expertise the road to this point has been much easier to travel. To Peter Bechtold, Martin Leutbecher, Peter Bauer, Anton Beljaars and the rest of the talented and helpful people at ECMWF, thank you especially for the three months I got to spend in Reading working with you. I am grateful to Marko Laine, who developed the parameter estimation algorithm applied here. I certainly could not have achieved the results presented here without close and fruitful collaboration with Marko. The insightfulness of Antti Solonen has also helped a lot in figuring out the way to do things properly.

My supervisor Prof. Heikki Järvinen has been the guiding force behind this thesis and the research it contains. Heikki picked me to this position, and his ideas of the parameter estimation framework helped in finding the proper direction for my research. Heikki has also given me considerable amount of help with all the articles presented in this thesis. Simply put, this thesis would not have been possible without you. Thank you.

I would also like to thank all my friends for keeping a smile on my face throughout the years we have known each other. Special thanks to Ilona for giving valuable input at various stages of the writing process of the summary part for this thesis, to Outi for sacrificing her spare time to proofread parts of the summary, and to Tuuli for having patience to comment on various parts of the summary, as well as giving a major help with the Finnish abstract.

To Risto and Riitta, thanks for all the wonderful meals, those days have been much needed. To my brothers, it has been a privilege to have grown up with you. Especially my brother Ossi, when my morale and energy was low during the final push in writing the summary, your support was perhaps the single reason that kept me going. A big hug to you. And finally, Mom and Dad, you have always supported me and been there for me. Kiitos kaikesta.

Helsinki, March 27, 2014

Pirkka Ollinaho





# CONTENTS

LIST OF ACRONYMS	10
LIST OF ORIGINAL PUBLICATIONS	11
SUMMARIES OF THE ORIGINAL PUBLICATIONS AND AUTHOR'S CONTRIBUTION	12
1 INTRODUCTION	14
2 PARAMETRIC UNCERTAINTY	17
2.1 CLOSURE PROBLEM	17
2.1.1 Turbulence closure . . . . .	17
2.2 PARAMETRIZATION SCHEMES	19
2.2.1 Convection . . . . .	19
2.2.2 Cloud processes . . . . .	21
2.3 DETERMINING THE CLOSURE PARAMETERS	21
2.4 ON PARAMETER ESTIMATION	22
3 METHODOLOGY	24
3.1 MODELS	24
3.1.1 ECHAM5 . . . . .	24
3.1.2 IFS . . . . .	25
3.2 ENSEMBLE PREDICTION	25
3.3 PARAMETER ESTIMATION ALGORITHM	26
4 MAIN RESULTS	28
4.1 MODEL FORECAST SKILL OPTIMISATION	28
4.2 PARAMETRIC UNCERTAINTY AND EPS SPREAD GENERATION	34
4.3 FROM WEATHER TO CLIMATE	37
5 CONCLUSIONS AND FUTURE DIRECTIONS	42
REFERENCES	44

## LIST OF ACRONYMS

<b>ACC</b>	Anomaly Correlation Coefficient
<b>ECHAM</b>	ECMWF model HAMburg version
<b>ECMWF</b>	European Centre for Medium-Range Weather Forecasts
<b>EnKF</b>	Ensemble Kalman Filter
<b>ENS</b>	ECMWF Ensemble Prediction System
<b>EPES</b>	Ensemble Prediction and Parameter Estimation System
<b>EPS</b>	Ensemble Prediction System
<b>GCM</b>	General Circulation Model
<b>IFS</b>	Integrated Forecasting System
<b>MSE</b>	Mean Squared Error
<b>MCMC</b>	Markov Chain Monte Carlo
<b>NWP</b>	Numerical Weather Prediction
<b>PF</b>	Particle Filter
<b>RMSE</b>	Root Mean Squared Error
<b>TOA</b>	Top of the atmosphere
<b>UTC</b>	Coordinated Universal Time

## LIST OF PUBLICATIONS

- I Ollinaho, P., Laine, M., Solonen, A., Haario, H., and Järvinen, H., 2013. NWP model forecast skill optimization via closure parameter variations. *Q.J.R. Meteorol. Soc.*, **139**, 1520–1532, doi:10.1002/qj.2044.
- II Ollinaho, P., Bechtold, P., Leutbecher, M., Laine, M., Solonen, A., Haario, H., and Järvinen, H., 2013. Parameter variations in prediction skill optimization at ECMWF. *Nonlin. Processes Geophys.*, **20**, 1001–1010, doi:10.5194/npg-20-1001-2013.
- III Ollinaho, P., Järvinen, H., Bauer, P., Laine, M., Bechtold, P., Susiluoto, J., and Haario, H., 2013. Total energy norm in NWP closure parameter optimization. *Geosci. Model Dev. Discuss.*, **6**, 6717–6740, doi:10.5194/gmdd-6-6717-2013
- IV Solonen, A., Ollinaho, P., Laine, M., Haario, H., Tamminen, J. and Järvinen, H., 2012. Efficient MCMC for climate model parameter estimation: parallel adaptive chains and early rejection. *Bayesian Anal.* **7(3)**, 715–736. doi: 10.1214/12-BA724.

## SUMMARIES OF THE ORIGINAL PUBLICATIONS AND AUTHOR'S CONTRIBUTION

- I Ollinaho, P., Laine, M., Solonen, A., Haario, H., and Järvinen, H., 2012. NWP model forecast skill optimization via closure parameter variations. *Q.J.R. Meteorol. Soc.*, **139**, 1520–1532, doi:10.1002/qj.2044.

PAPER I demonstrates the application of the Ensemble and Parameter Estimation System (EPPES) in closure parameter estimation in a coarse resolution ECHAM5 climate model. Four closure parameters related to convection and cloud processes are optimised by targeting 500 hPa geopotential height errors at forecast days three and ten. The EPPES is able to find parameter values corresponding to improved model in the target criterion sense, proving that parameter estimation is feasible in an atmospheric General Circulation Model (GCM). In addition, the evolution of parameter covariances are studied. In an asymptotic test, where the sampling period is cycled through 10 times, noticeable parameter correlations appear. The climatology of the improved model is also tested in a 6-year simulation. The closure parameters improved with respect to 500 hPa geopotential height also enhance the top of the atmosphere radiative balance of the model. This indicates that the model medium range forecast skill and climatological behaviour might have a connection.

The author was responsible for all the model experiments and analysis of data, and participated in the writing.

- II Ollinaho, P., Bechtold, P., Leutbecher, M., Laine, M., Solonen, A., Haario, H., and Järvinen, H., 2013. Parameter variations in prediction skill optimization at ECMWF. *Nonlin. Processes Geophys.*, **20**, 1001–1010, doi:10.5194/npg-20-1001-2013.

PAPER II illustrates the application of the Ensemble and Parameter Estimation System (EPPES) in closure parameter estimation in the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Four closure parameters related to convection are optimised by targeting 500 hPa geopotential height errors at forecast days three and ten. In spite the increased resolution, high forecast skill of the IFS and inclusion of stochastic noise, the EPPES estimation is able to find parameter values corresponding to improved model in the target criterion sense. Parameter estimation is thus shown to be feasible even in a high forecast skill forecasting system. The additional estimation noise generated by stochastic physics is shown to slow down the convergence of parameter covariances, but the convergence to optimal parameter values is not affected by the noise. Probabilistic skill of the ECMWF Ensemble Prediction Sys-

tem (ENS) in lower-than-operational resolution is shown to benefit from the inclusion of parameter perturbations.

The author was responsible for all the model experiments and analysis of data, and for a major part of the writing process.

- III Ollinaho, P., Järvinen, H., Bauer, P., Laine, M., Bechtold, P., Susiluoto, J., and Haario, H., 2013. Total energy norm in NWP closure parameter optimization. *Geosci. Model Dev. Discuss.*, **6**, 6717–6740, doi:10.5194/gmdd-6-6717-2013

PAPER III explores the use of atmospheric total energy norm (EN) as a target criterion for parameter optimisation. The Ensemble and Parameter Estimation System (EPPES) is used to estimate four ECHAM5 closure parameters related to convection and cloud processes. The EPPES is able to find parameter values corresponding to improved model in the target criterion sense. Moreover, the EN improvements in the optimised model are the larger the longer the forecast range is. The improvements are shown to originate from more realistic tropical kinetic energy representation. At longer forecast ranges the improvements spread to higher latitudes, too. Finally, the optimised parameter values induce model wide improvements into the forecasts.

The author was responsible for all the model experiments and analysis of data, and for a major part of the writing.

- IV Solonen, A., Ollinaho, P., Laine, M., Haario, H., Tamminen, J. and Järvinen, H., 2012. Efficient MCMC for climate model parameter estimation: parallel adaptive chains and early rejection. *Bayesian Anal.* **7(3)**, 715736. doi: 10.1214/12-BA724.

PAPER IV introduces two concepts to improve the efficiency of Markov Chain Monte Carlo (MCMC) applications. First, a parallel MCMC algorithm allows for running MCMC applications in supercomputer environments effectively. Second, an early rejection method is presented and demonstrated. The early rejection method is applied in a MCMC-style closure parameter optimisation problem. The method is demonstrated to save 595 simulation years worth of CPU time in an experiment consisting of 3204 simulated years in total.

The author performed the atmospheric model runs and data analysis of these runs.

# 1 INTRODUCTION

Importance of weather information is constantly rising with growing human population and welfare. Agriculture, aviation, transport, and energy production are heavily dependent on the weather conditions, and thus correct weather forecasts can yield substantial financial benefits. Furthermore, accurate weather forecasts can prevent human injuries, or even save lives. Although extreme weather events are rare in Finland, even unusual weather conditions such as icy roads lead annually to approximately 47000 slipping accidents, resulting in about 420 million euros worth of losses to the society through costs in health care, lost workdays and social welfare (Ruuhela et al., 2005). On global scale these effects are much more pronounced. During 1975–2008 extreme weather events, such as hurricanes and floods, resulted in approximately 1.5 million deaths, and the direct costs alone to the global economy were in the order of 700 billion euros (Llosa and Zodrow, 2011). Clearly not all of these injuries, fatalities, and financial losses could be alleviated by accurate weather forecast alone. Nevertheless, the benefits of developing more accurate weather prediction models affect especially mitigation of injuries and losses of life, and aid in maintaining a reliable food and energy supply. The importance of the latter two is continuously increasing as the global temperatures rise, and the energy production shifts towards harvesting energy from the Sun, wind and waves.

Vilhelm Bjerknes postulated in 1904 that the future state of the atmosphere is, in principle, determined by its initial state and boundary conditions, and by a set of five partial differential equations (e.g. Kalnay, 2003): equation of motion (1.1), continuity equation (1.2), conservation of energy (1.3) and water mass (1.4), and the ideal gas law (1.5)

$$\frac{d\mathbf{v}}{dt} = -\frac{\nabla p}{\rho} - \nabla\phi + \mathbf{F} - 2\boldsymbol{\Omega} \times \mathbf{v} \quad (1.1)$$

$$\frac{d\rho}{dt} = -\rho\nabla \cdot \mathbf{v} \quad (1.2)$$

$$C_p \frac{dT}{dt} = \frac{1}{\rho} \frac{dp}{dt} + Q \quad (1.3)$$

$$\frac{dq}{dt} = E - C \quad (1.4)$$

$$p = \rho RT \quad (1.5)$$

The equations describe the time ( $t$ ) tendencies of three-dimensional wind field  $\mathbf{v}$ , pressure  $p$ , density  $\rho$ , temperature  $T$  and water vapour mixing ratio  $q$ .  $\mathbf{F}$  represents frictional forces,  $\boldsymbol{\Omega}$  the Earth's rotation, and  $\phi$  the gravitational potential. In addition, diabatic heating  $Q$ , evaporation  $E$ , and condensation  $C$  act as source and sink terms in the set of equations. Lastly,  $C_p$  and  $R$  are the specific heat at

constant pressure and the gas constant for air, respectively.

The set of equations, often referred to as primitive or governing equations, is a non-linear system without analytic solutions. Furthermore, the system is forced, dissipative and coupled. Lorenz (1963) argued that a system with these properties is entirely determined by its initial state. He continued that in such a system, even small differences in two initial conditions will eventually lead to two completely different solutions. Thus, an accurate representation of the initial state of the system is a fundamental part of a successful weather forecast. And consequently, uncertainties related to the atmospheric initial state constitute as an important part to a forecasting uncertainty. The initial state uncertainty has caught much attention throughout the history of numerical weather forecasting development. Although the amount of atmospheric observations has greatly increased due to proliferation of satellite data, the current atmospheric state cannot still be accurately observed. In order to improve the accuracy of the initial state, statistical methods are used to extract information of the observations. The most accurate state of the atmosphere, i.e. an analysis, is constructed through an optimal combination of observations and the model predictions. These so-called data-assimilation methods employ the short range model forecasts as a background field, which is then corrected with available observations.

When applying the primitive equations to numerical weather forecasting the equation system needs to be discretized. This introduces into the forecasting system a modelling uncertainty, or a modelling error, that adds to the initial state uncertainty. The modelling uncertainty originates from the numerical methods used to integrate the equations forward in time. The numerical results are approximations of the analytic solution. Furthermore, due to limitations in computational power, the numerical representation of the state is truncated, that is, the model domain is divided into grid boxes as illustrated in Fig. 1.1. For example, the currently most accurate global weather prediction model, the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF), operates on a grid that divides the globe into approximately  $16 \times 16$  km horizontal boxes. Thus, atmospheric phenomena occurring in scales smaller than a few grid spacing<sup>1</sup> are left unsolved by the numerical representation of the governing equations. In order to complement the lack of resolution, the net effect of the subgrid-scale phenomena are estimated with so-called parametrizations<sup>2</sup>, often referred to as "model physics". Erroneous or lacking representation of the

---

<sup>1</sup>The smallest scale phenomena that can be theoretically resolved in a finite difference scheme have a wavelength of two grid lengths. In practice at least four grid lengths are necessary due to stability considerations (see e.g. Grasso, 2000).

<sup>2</sup>The need for parametrizations can also be seen to follow from the discretization of the governing equations; the equations no longer predict the time evolution of the atmosphere precisely, and the equations are thus incomplete. The parametrized processes then complement the discretized equation set, and allow for accurate prediction of the future state of the atmosphere.

modelled physical processes then introduces into the forecasting system a parametric uncertainty, which constitutes a substantial part of the whole modelling uncertainty.

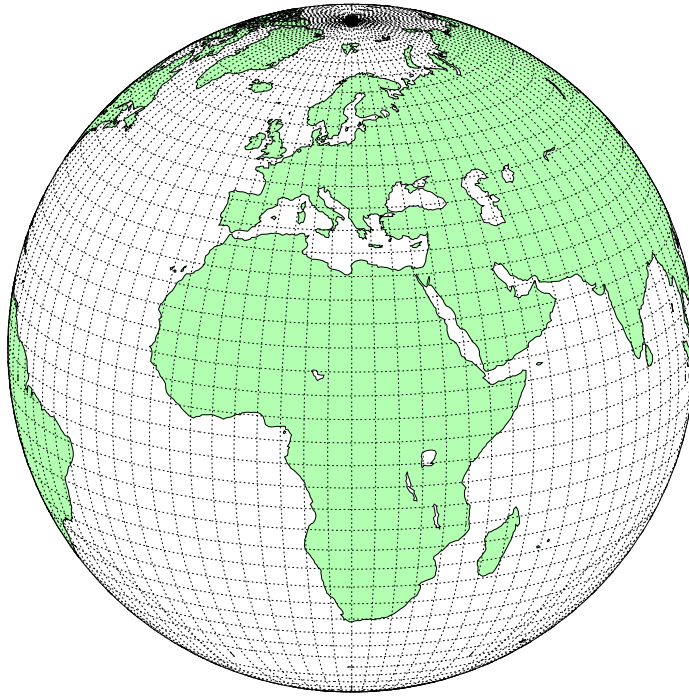


FIGURE 1.1. Earth divided into  $4^\circ \times 4^\circ$  (approximately  $440 \times 440$  km) grid boxes.

The main objective of this Thesis is to study and reduce the uncertainties and errors related to the parametrized processes. This research problem is addressed by (i) assessing if reducing model error is feasible through improving the parametrizations non-structurally, i.e. without changing the underlying numerical representation of the parametrizations, (ii) determining whether and how information about the parametric uncertainty could be utilised for improving the weather prediction models, and (iii) studying how the parametric uncertainty in medium range forecasts relates to the modelled long-term climatology.



## 2 PARAMETRIC UNCERTAINTY

The parametrizations characterise the subgrid-scale activity of physical processes, such as rain and cloud formation, radiative transfer in the atmosphere, and turbulence in the boundary layer to name a few. The parametrized processes might be erroneously described simply because the physical mechanisms behind the processes are not yet completely understood. Furthermore, some vital processes might even miss from the forecasting system. The parametrizations have thus many possible error sources, which lead to uncertainties in the forecasts. The forecast uncertainty caused by the imperfect parametrizations is here referred to as parametric uncertainty. Reducing the errors and the consequential parametric uncertainties is generally achieved by constructing more realistic parametrization schemes as the knowledge about the phenomena is increased and as additional computing power becomes available. This in general requires structural parametrization changes, i.e. revising the scientific reasoning accompanied by extensive code level rewriting in the parametrizations. However, the increase of realism could also be attainable through non-structural parametrization changes, motivated in the following.

### 2.1 CLOSURE PROBLEM

An important and often neglected part of the parametric uncertainty is the uncertainty related to closure parameters inside the parametrizations. The parametric representations of the sub-grid scale processes usually end up in a closure problem, where e.g. rate at which part of the physical process occurs cannot be fully described. Therefore, preset parameter values need to be defined, which then affect the outcome of the parametrization. For instance, convective precipitation formation in the IFS is controlled by the conversion rate from cloud water into rain, evaporation of precipitation, and the melting rate of snow. All three terms involve closure parameters that characterise the intensity of the individual processes. Thus, in order to achieve a realistic amount of precipitation in the model, the values of these parameters need to be set correctly. The closure parameters are thus an integral part of the parametrizations, and their values affect the accuracy of the model.

#### 2.1.1 *Turbulence closure*

In order to illustrate the closure problem, the turbulent movement of air is used as an example. A common turbulence representation splits the turbulent motions into a time mean part and a fluctuating part. The horizontal wind components can then be presented as  $u = \bar{u} + u'$  and  $v = \bar{v} + v'$ , where the overbars indicate time mean quantities and the primes perturbations from the mean. After

omitting advection and molecular viscosity terms, as well as treating the flow to be horizontally homogenous and incompressible, the time tendencies of the mean horizontal wind components can be expressed as

$$\frac{\partial \bar{u}}{\partial t} = -\frac{1}{\bar{\rho}} \frac{\partial \bar{p}}{\partial x} + f\bar{v} + \frac{\partial \overline{u'w'}}{\partial z} \quad (2.1)$$

$$\frac{\partial \bar{v}}{\partial t} = -\frac{1}{\bar{\rho}} \frac{\partial \bar{p}}{\partial y} + f\bar{u} + \frac{\partial \overline{v'w'}}{\partial z} \quad (2.2)$$

Here,  $u$  and  $v$  are the horizontal wind components, and  $w$  the vertical wind component. Furthermore,  $p$  is pressure,  $\rho$  the density of the column of air, and  $f$  the Coriolis-parameter. The overbars indicate time means and the primes perturbations.

In atmospheric models the mean values are available since they are part of the model state. However, there are two unknown terms in the equations,  $\overline{u'w'}$  and  $\overline{v'w'}$ . In order to calculate these terms explicitly, their tendencies need to be solved (Kalnay, 2003). Equations for these can be constructed by first deriving the tendency equations for the fluctuation parts, and then further solved as (see e.g. Stensrud, 2007)

$$\begin{aligned} \frac{\partial \overline{u'w'}}{\partial t} = & -\frac{1}{\bar{\rho}_0} \left( \frac{\partial \overline{p'w'}}{\partial x} + \frac{\partial \overline{p'u'}}{\partial z} - \overline{p' \frac{\partial u'}{\partial z'} \frac{\partial w'}{\partial x}} \right) + \overline{f'w'v'} \\ & - \overline{u'v'} \frac{\partial \bar{w}}{\partial y} - \overline{w'v'} \frac{\partial \bar{u}}{\partial y} - \frac{\partial \overline{u'v'w'}}{\partial y} \end{aligned} \quad (2.3)$$

$$\begin{aligned} \frac{\partial \overline{v'w'}}{\partial t} = & -\frac{1}{\bar{\rho}_0} \left( \frac{\partial \overline{p'w'}}{\partial y} + \frac{\partial \overline{p'v'}}{\partial z} - \overline{p' \frac{\partial v'}{\partial z'} \frac{\partial w'}{\partial y}} \right) + \overline{f'w'u'} \\ & - \overline{v'u'} \frac{\partial \bar{w}}{\partial x} - \overline{w'u'} \frac{\partial \bar{v}}{\partial x} - \frac{\partial \overline{v'u'w'}}{\partial x} \end{aligned} \quad (2.4)$$

However, to get values of  $\overline{u'w'}$  and  $\overline{v'w'}$ , additional correlation terms ( $\overline{p'w'}$ ,  $\overline{p'u'}$ ,  $\overline{p'v'}$  and  $\overline{u'v'}$ ) need to be calculated. Moreover, an unknown triple correlation  $\overline{u'v'w'}$  has appeared into the equations. Solving for this term then introduces a quadruple correlation term into the system, and the pattern continues when this term is solved for. This reappearance can be "closed" by e.g. describing the correlation terms as proportional to the vertical gradient of the mean part

$$\overline{u'w'} = -K \frac{\partial \bar{u}}{\partial z}, \overline{v'w'} = -K \frac{\partial \bar{v}}{\partial z}, \quad (2.5)$$

where the  $K$  refers to a constant, formally the vertical eddy viscosity coefficient (Savijärvi and Vihma, 2001). Thus, instead of solving directly for the correlated fluctuations, their contribution is now described through an approximation. Moreover, the value of this approximation is dependent on the value of the closure parameter  $K$ . The parametrization can also be constructed so that the value of  $K$

is set to be state dependent in the model, i.e. to vary according to the boundary layer conditions.

## 2.2 PARAMETRIZATION SCHEMES

When studying the impact of closure parameters on the model forecasts, it is practical to study a limited amount of closure parameters from a limited set of parametrizations. This way the physical impacts of the closure parameters might also be traceable, or even explainable, despite of the non-linear model response to the parameter changes. In order to make these choices it is important to understand how the parametrizations affect the forecasts; the physical processes vary in the extent that they affect the model forecasts, and some of the processes have a more localised effect and impact only some specific model fields. A correct representation of convection and cloud processes is crucial to achieve overall accurate forecasts. These two processes influence the short range forecasts of precipitation, humidity, winds, temperature, etc. Additionally, they have a substantial effect on the global energy balance in climatological runs and future simulations. Thus, PAPERS I, II, III and IV concentrate on studying closure parameters from these parametrization schemes. In the following, a concise overview of convection and cloud processes and their parametrization is presented. The closure parameters found in these parametrizations and considered in the parameter optimisation are presented in Chapters 3.1.1 and 3.1.2.

### 2.2.1 *Convection*

The turbulent motions of moist convection are inseparably connected to the large-scale dynamics of the tropical (20°S to 20°N) and extra-tropical tropospheres, and consists of deep cumulonimbus cloud systems and shallow, non-precipitating cumuli, as well as stratocumulus cloud sheets (Schneider and Sobel, 2007). The deep convective processes form the main heat engine of tropical circulation: the excess radiative forcing in the tropical regions is first redistributed in vertical through convective processes, and the vertical motions then drive global horizontal circulation patterns, such as the Hadley and Walker cells (Stensrud, 2007). Shallow convection on the other hand impacts global climate through modifying the surface radiation budget, and effects the structure of the planetary boundary layer (Randall et al., 1985). Finally, moist convection is the main driving force of the hydrological cycle.

For parametrization purposes it is convenient to separate the convective processes into deep and shallow convection (Stensrud, 2007); deep convective elements involve large part of the troposphere, while shallow convective elements cover vertically only a small part of the troposphere. Deep convection parametrization can be

conceptualised in many different ways and the schemes further divided into a number of basic types (Stensrud, 2007). So-called low-level control schemes (Mapes, 1997) assume that deep convection is determined by the physical processes controlling convective initiation (Schneider and Sobel, 2007). These schemes are concerned with how a parcel of air is able to overcome any convective inhibition present and activate its convective available potential energy (CAPE). This line of thought is incorporated in mass flux schemes, where the large-scale contribution of the convective processes is thought to comprise of an ensemble of updrafts and downdrafts representing the convective cloud elements within a grid box. These convective plumes then interact with the surrounding environmental air through entrainment and detrainment. The updraft and downdraft as well as their interactions with the environmental air in a single plume are illustrated in Fig.2.1. By calculating the bulk equations for the up- and downdrafts the contribution to the large scale budgets of heat, moisture and momentum can then be solved (Roeckner et al., 2003). In Tiedtke-style schemes (Tiedtke, 1989) entrainment in rising plumes is split into a term representing the turbulent mass exchange through the cloud edges, and a second term directly related to the large-scale moisture convergence, representing inflow into the cloud (Stensrud, 2007). Tiedtke schemes also solve shallow convection within the same line-of-thought.

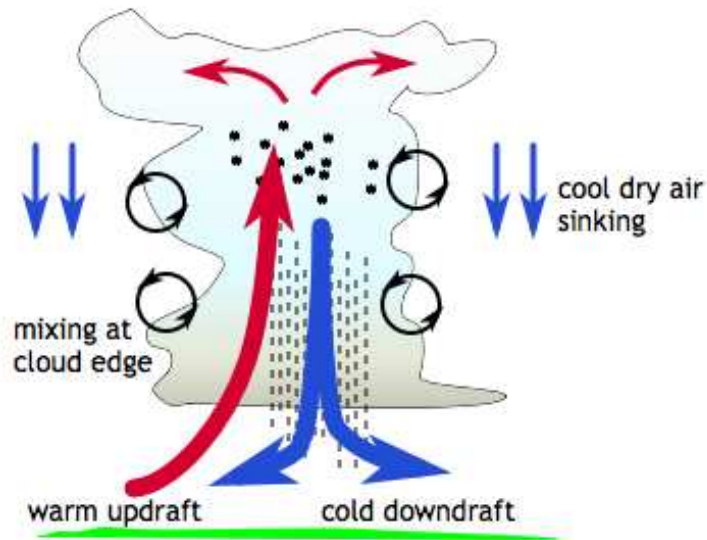


FIGURE 2.1. The main processes related to convection parametrization. Image courtesy of Center for Multiscale Modeling of Atmospheric Processes (CMMAP).

### 2.2.2 *Cloud processes*

Cloud cover greatly modifies the radiation received at the Earth’s surface, thereby affecting atmospheric circulation and climate (Stensrud, 2007). In addition to reflecting and absorbing shortwave (solar) radiation, clouds also absorb and emit longwave radiation. Thus the feedback of clouds to the Earth’s radiation budget is quite complex, e.g. extending low- and mid-level cloud cover in the tropics leads to surface cooling, owing to the increase of albedo, whereas a growing high-level cirrus cloud cover leads to surface warming, as a result of increase in longwave absorption and emission accompanied by a nearly unchanged solar radiation (Schneider and Sobel, 2007).

When considering cloud parametrization, it is essential to realise that clouds vary in both vertical and horizontal directions. In atmospheric models cloud cover is defined as the fraction of the cloud cover encompassing over the (selected) vertical layers. The parametrization schemes are divided into two main types (Stensrud, 2007): (i) Diagnostic cloud cover parametrizations diagnose cloud cover after each time step from the model state variables. (ii) In prognostic cloud cover parametrizations cloud cover, as well as cloud water, is added as a predicted model variable. This approach is more complicated and more expensive computationally. The prognostic parametrizations usually also involve cloud microphysics parametrization, which deal with the phase changes of water vapour as well as the interactions of cloud droplets and precipitating particles.

## 2.3 DETERMINING THE CLOSURE PARAMETERS

The closure parameters have been traditionally chosen based on the results of process studies, observational campaigns, or expert knowledge. Applying brute-force trial-and-error methods to search for the best fitting closure term is usually arduous (see e.g. Bender, 2008). One reason for this is the very non-linear model response to parameter variations, especially when multiple parameters are changed simultaneously. NWP models also operate at a high level of forecast skill, implying that the various multi-scale interactions and dynamic-physics feedbacks are tuned into harmony. Thus, re-tuning of these complex multi-scale modelling systems through manual closure parameter optimisation is a hard problem.

In recent years various statistical, algorithmic approaches have been studied in search for solutions to problems the size and complexity of NWP and climate models. These solutions have mostly been based on data-assimilation approaches, and thus focus on very short range forecast improvements. The most common of these approaches are based on the Kalman filter (Kalman, 1960) and particle filter (PF; Kivman, 2003; van Leeuwen, 2003). The Kalman filter applications often make use of the Ensemble Kalman Filter (EnKF; Evensen, 1994). The fundamental principle in these algorithms is to augment the state space with the

closure parameters; the parameters are considered to be a part of the model state together with the model variables, and calibrated concurrently with the model state during the analysis step. The EnKF approach in complex systems has been studied in relatively low-order models (e.g. Aksoy et al., 2006), as well as in coarse resolution climate model (Annan et al., 2005) and limited area models (LAM) of operational complexity (e.g. Aksoy et al., 1996; Hu et al., 2010). Though the EnKF approach shows promising results in simplified settings, moving to more realistic estimation cases reduces the parameter identifiability (Schirber et al., 2013). PF schemes have been studied in parameter optimisation in coarse resolution global climate model and LAM settings (e.g. Jackson et al., 2008; Yang et al., 2012).

Adaptive Markov Chain Monte Carlo (MCMC; Haario et al., 2006) has also recently been applied for parameter estimation with focus on model climatology; Järvinen et al. (2010) estimated ECHAM5 global GCM closure parameter posterior joint probability densities by targeting the top of the atmosphere (TOA) net radiation monthly errors. Although successful in optimising the climatological TOA net radiation, weak identifiability was observed in two of the estimated parameters, bringing up discussion on how to choose the target for the optimisation.

## 2.4 ON PARAMETER ESTIMATION

Atmospheric model systems are known to suffer from spinup and spindown problems, e.g. initial state imbalances in hydrological cycle causes the moist variables to display a tendency towards the model attractor (e.g. Betts et al., 2003; Trenberth and Guillemot, 1998). Parametric uncertainties are typically related to moist physical processes, and may thus be affected by this imbalance at very short forecast ranges. When moving to short-to-medium range forecasts, it is difficult to identify the individual contributions of the initial state and model errors to the forecast error; initial state errors are influenced by model errors too since the initial state generation contains the forecast model (Leutbecher and Palmer, 2008). Nevertheless, Savijärvi (1995) showed that very-short-range forecast error is dominated by the exponential growth of initial state errors, and that the linearly growing model errors influence the forecast error in longer model integrations. This would imply that (i) the effects of parameter variations are not substantial early in the forecast range, and (ii) at longer forecasts the parameter variation influence is stronger, but the indentifiability is masked by the non-linearity of the system. Forecast ranges between these extremes, where the parameter variations would already influence more strongly to the forecast error but atmospheric chaoticity does not yet dominate, would thus provide for an intriguing time window for parameter evaluation studies.

Järvinen et al. (2012) and Laine et al. (2012) introduced a methodology to apply this time window. The Ensemble Prediction and Parameter Estimation

System (EPPES) utilises an ensemble prediction system to make statistical inference about perturbed closure parameters. In EPPES, initial-time parameter variations are introduced on an ensemble of forecasts. Parametric uncertainty is then determined based on how likely the ensemble members are when evaluated against observations, with respect to a chosen target criterion. The evaluation can be performed at any forecast range covered by the forecasts, say at forecast day four. Therefore, the medium range forecast skill is targeted directly for optimisation. The algorithm is also in essence only monitoring the EPS, thus practically no-additional computation cost is added to the system in operational context.

### 3 METHODOLOGY

In the following, the atmospheric models used in the experiments, the concept of ensemble prediction, and the parameter estimation algorithm are introduced.

#### 3.1 MODELS

##### 3.1.1 *ECHAM5*

The ECMWF model HAMburg version 5 (ECHAM5; Roeckner et al., 2003, 2006) is a comprehensive GCM used for climate simulations. The ECHAM model has been developed by the Max-Planck-Institut für Meteorologie (MPI-M) in Hamburg since 1989. The dynamical core of ECHAM5 operates in spectral space, while the parametrized parts are calculated in grid point space. The vertical levels utilise pressure based terrain following sigma-coordinates. In PAPERS I, III and IV model version 5.4 is used with a coarse horizontal resolution of T42, i.e. triangular truncation at wave number 42, responding to a grid spacing of about 310 km. Correspondingly, 31 vertical levels are used, with model top at 10 hPa. A semi-implicit treatment is used in dynamics time-integration with a time-step of 20 min. The physical parametrizations are evoked at every time step, with the exception of radiative transfer, which is calculated once every two hours.

ECHAM5 convection parametrization is a Tiedtke style mass flux scheme with modifications for penetrative convection in line with Nordeng (1994). The stratiform cloud scheme consists of prognostic equations for water in vapour, liquid and ice phase, and a microphysical scheme after Lohmann and Roeckner (1996) with some revisions. A statistical cloud cover scheme involved in the stratiform cloud scheme was not used here, due to problems encountered in climate simulations performed with non-default parameter values. Table 3.1 outlines the three closure parameters from the convection scheme (CMFCTOP, CPRCON and ENTRSCV) and the stratiform cloud scheme one (CAULOC) that were used in PAPERS I, III and IV.

Table 3.1 ECHAM5 closure parameter subset used in model optimisation.

Parameter	Description
CAULOC	A parameter influencing the accretion of cloud droplets by precipitation (rain formation in stratiform clouds)
CMFCTOP	Relative cloud mass flux at the level above non-buoyancy (in cumulus mass flux scheme)
CPRCON	A coefficient for determining conversion from cloud water to rain (in convective clouds)
ENTRSCV	Entrainment rate for shallow convection



### 3.1.2 IFS

The forecast model of the Integrated Forecasting System (IFS) of the ECMWF is a global hydrostatic GCM. The model dynamics use a spectral, semi-implicit, and semi-Lagrangian two time-level dynamical solver. In PAPER II model version CY37R3<sup>1</sup> is used with horizontal resolution of T<sub>L</sub>159 (equalling to a grid spacing of about 125 km), 62 vertical levels and the model top at 5 hPa. Physical parametrizations are calculated at every time step (30 min), the exception being radiation, which is calculated once per three hours.

The physical parametrization of convection constitutes of a bulk mass flux scheme after Tiedtke (1989) with modifications according to Bechtold et al. (2008). The scheme is further divided into deep, mid-level, and shallow convection. The entrainment and detrainment formulation in the parametrization follows closely observations and output of cloud resolving models (de Rooy et al., 2013). The closure parameters considered in PAPER II relate to the entrainment and detrainment rates in deep convection, entrainment in shallow convection, and precipitation formation (Table 3.2).

Table 3.2 IFS closure parameter subset used in model optimisation.

Parameter	Description
ENTRORG	Entrainment rate for positively buoyant deep convection
ENTSHALP	Shallow entrainment defined as ENTSHALP $\times$ ENTRORG
DETRPEN	Detrainment rate for penetrative convection
CPRCON	Coefficient for determining conversion from cloud water to rain

## 3.2 ENSEMBLE PREDICTION

The atmosphere is chaotic by nature; tiny errors in the initial conditions will lead to different atmospheric states when the forecast is long enough. Lorenz (1969) estimated that due to the chaoticity the atmospheric predictability is about two weeks.

In probabilistic or ensemble forecasting, instead of having a single deterministic forecast of the atmospheric state, an ensemble of forecasts is used to predict what is the most probable state of the atmosphere, and how reliable this state is. The ensembles in an Ensemble Prediction System (EPS) are traditionally generated by perturbing the initial conditions around the analysis state, via e.g. introducing into the system optimally growing perturbations that depend on the dynamical state of the atmosphere (e.g. Buizza et al., 1993). Perturbing the initial conditions alone usually generates under-dispersive ensembles, i.e. the spread

<sup>1</sup>IFS documentation is available on-line at <http://www.ecmwf.int/research/ifsdocs>

of the ensemble does not cover the true atmospheric variability (see e.g. Slingo and Palmer, 2011). The missing spread due to modelling errors can be simulated by, e.g., perturbing the tendencies of the physical parametrizations (stochastic physics), or perturbing the physical parametrizations themselves. For this Thesis, parameter perturbations are particularly interesting.

### 3.3 PARAMETER ESTIMATION ALGORITHM

The Ensemble Prediction and Parameter Estimation System (EPPES) algorithm is described in detail in Laine et al. (2012), who also demonstrated its functionality using a stochastic version of the Lorenz-95 model (Lorenz, 1996; Wilks, 2005). In EPPES, it is assumed that for time window  $i$ , the optimal closure parameter  $\theta_i$  is a random realisation of a random variable, which follows a multivariate Gaussian distribution with a mean vector  $\mu$  of dimension  $p$  and a  $p \times p$  covariance matrix  $\Sigma$

$$\theta_i \sim N(\mu, \Sigma), i = 1, 2, \dots \quad (3.1)$$

The parameter estimation is thus formulated as a problem of estimating the unknown but static in time distribution parameters (or, hyper-parameters)  $\mu$  and  $\Sigma$ . These can be interpreted as follows: the distribution mean  $\mu$  represent the parameter values that work best on average considering all weather types, seasons, etc., and  $\Sigma$  reflects how much the optimal parameter values vary between time-windows due to evident modelling errors, such as inaccurate parametrization schemes.

At initial time, the distribution parameters ( $\mu$  and  $\Sigma$ ) are specified according to expert knowledge. Parameter bounds can also be issued to prevent the selection of unphysical or unwanted parameter values. The algorithm then draws a sample from the initial distribution. After running an ensemble of forecasts using these parameter values, the likelihood of each forecast is evaluated as a fit to observations. The likelihood is then used to weight each parameter vector and a re-sample is drawn from the weighted parameter sample, favouring parameters related to high likelihood (known as importance sampling). Lastly, the weighted sample is employed to update the hyper-parameters  $\mu$  and  $\Sigma$ . The updated distribution is then used in drawing a new sample for the next time-window, and the process is repeated. The algorithm steps are also given in detail in Chapter 2.3 of PAPER I.

For EPPES estimation purposes, an ECHAM5 EPS emulator was set up in PAPERS I and III by copying initial state perturbations generated by the operational ECMWF Ensemble Prediction System (ENS). The state variables of both models are the same (temperature, vorticity, divergence, logarithm of surface pressure, and specific humidity), thus the initial states can be conveniently used in

ECHAM5. However, since the perturbations have been constructed to be optimally growing in the ENS forecast model, they probably generate less spread in the EPS emulator. Model error in the emulator is represented by the initial-time parameter perturbations. The initial state and parameter perturbations generate approximately the same amount of ensemble spread in this system.

## 4 MAIN RESULTS

The main research objectives of the Thesis set in Chapter 1 are studied through the following questions:

- Q1 Can optimal parameter values be found algorithmically in a low resolution GCM of full complexity?
- Q2 Is parameter optimisation feasible in a system already at high level of forecast skill?
- Q3 How does the choice of target criterion affect the parameter estimation?
- Q4 Is there any useful information gained from studying the parameter covariance data?
- Q5 Do parameter perturbations affect the probabilistic skill of an EPS?
- Q6 Does medium range parameter optimisation have any relevance for model climatology?

Chapter 4.1 aims to answer the first three questions, the fourth and fifth are studied in Chapter 4.2, and the sixth one in Chapter 4.3.

### 4.1 MODEL FORECAST SKILL OPTIMISATION

To find answers to Q1 and Q2, a simple target criterion for the parameter estimation was appealing for demonstrative purposes. Therefore, in PAPERS I and II the target, or cost function, for EPPES sampling was chosen to be mean squared error (MSE) of 500 hPa level geopotential height at forecast days three and ten. In PAPER I, the EPPES evaluation was able to find a more skillful model in the target criterion sense (Fig. 5 of PAPER I). Furthermore, the optimised model was also more skillful outside the sampling period. These results indicate that the EPPES optimisation works in a problem the size and complexity of an atmospheric GCM. To further increase the complexity of the problem, in PAPER II the EPPES estimation with the same target criterion was experimented on with the ECMWF IFS. Additional noise was also introduced to the sampling through inclusion of stochastic physics. Like in the ECHAM5 experiment (Fig. 2 of PAPER I), the distribution mean values in the IFS experiment find the posterior values early on, but the distribution width does not converge well (Fig. 1 of PAPER II). And, it even grows in the ENTRORG and RPRCON distributions. In an experiment run without stochastic physics, the parameter distribution uncertainties did shrink. Therefore, the additional noise generated by the stochastic physics slows down the covariance information gain from parameter perturbations. However,

it does not prevent the distribution mean values from converging; the optimised model improves the root mean squared error (RMSE) scores with 95% confidence level up to forecast range nine and a half days (Fig. 3 of PAPER II). EPPES is thus able to find optimised parameter values in a high forecast skill NWP model. Even the added stochastic noise does not prove to be an obstacle for a successful parameter optimisation. Thus, summarising answer to Q1 and Q2 is that parameter estimation with the EPPES algorithm is feasible even in a problem the size and complexity of a true NWP model.

To answer Q3, the model response to the new parameter values needs to be studied more comprehensively. The optimised model in PAPER II shows a generally positive signal to the parameter changes (Fig. 4 of PAPER II), especially the tropics benefits considerably from the new parameters. However, there is a striking global degradation in the 100 hPa geopotential height RMSE. The degradation emphasises the selective nature of the 500 hPa geopotential height cost function; the target criterion is implicit about errors in mean temperature and humidity in the atmosphere below 500 hPa, and about processes which affect 500 hPa forecast errors. However, it is insensitive to e.g. geopotential height errors higher up. The choice of the target criterion for EPPES estimation is thus not trivial. It is not sufficient that the model is improved only w.r.t. the target criterion. It also has to force the model onto a forecast trajectory which imposes a model-domain-wide improvement in most resolved model variables. Also, the parameters to be estimated have to be sensitive to the target criteria and to trigger a noticeable change in the object value when parameter values are varied.

Motivated by this, PAPER III explores the use of atmospheric total energy norm (EN) as the target criterion. The energy norms in NWP context are mainly used in finding the fastest growing error structures to be used as initial state perturbations in EPS (e.g. Farrel, 1988; Palmer et al., 1994; Errico, 2000). In PAPER III the total energy norm is applied in the opposite sense, and a model is sought corresponding to the slowest possible forecast error growth in terms of the total energy norm. The energy norm is an integral quantity over the whole model atmosphere, and as such does not have preferences for any model variable, level, or geographical region. The discretised form used in PAPER III is

$$\begin{aligned} \Delta E = & \frac{1}{2} \sum_{p_0}^{p_1} \sum_A \left( (\Delta u)^2 + (\Delta v)^2 + \frac{c_p}{T_r} (\Delta T)^2 \right) dA dp \\ & + \frac{1}{2} R_d T_r p_r \sum_A (\Delta \ln p_{sfc})^2 dA. \end{aligned} \quad (4.1)$$

Here,  $u$  and  $v$  refer to the zonal and meridional wind components,  $T$  is temperature, and  $\ln p_{sfc}$  logarithmic surface pressure.  $\Delta$  indicates the difference between observed and forecasted atmospheric states. The ECMWF operational analyses

are used as observations.  $c_p$  is the specific heat at constant pressure,  $R_d$  gas constant of (dry) air,  $T_r$  a reference temperature (280K),  $p_r$  a reference surface pressure (1000 hPa),  $dA$  an areal element of the model grid, and  $dp$  the pressure difference between two pressure levels. In PAPER III  $dp = 1$  is used throughout the atmosphere in order to emphasize the surface pressure term. The first two terms in r.h.s. of Eq. 4.1 identify as kinetic energy, and the temperature and surface pressure terms as available potential energy (Lorenz, 1955, 1960).

In PAPER III the EPPES sampling is done with targeting 72-hour forecast EN errors. An identical test setup is used as in PAPER I, i.e. the ECHAM5 EPS-emulator is run with a 51-member ensemble twice daily (00 and 12 UTC) covering a time period from 1st January to 31st March 2011 (2011JFM). Thus, 9180 parameter subsets are tested in total. Fig. 4.1 illustrates the evolution of the four closure parameters during the three month sampling period. Mean distribution value (continuous line), and distribution width as two times standard deviation (dashed lines) representing the 95% probability range of the parameter uncertainty are shown. A vertical column of markers illustrates the ECHAM5 parameter values evaluated at the given date, darkness of the marker indicates the weight at re-sampling step. Similarly to the PAPER I experiment (Fig. 2 of PAPER I), CAULOC and CPRCON have a large initial shift to a new parameter region. Additionally, CMFCTOP and ENTRSCV evolve more conservatively in this experiment also.

To validate the the posterior distribution, the posterior mean values (Table 2 of PAPER III) are used to run the model. The forecasts from the default and the optimised model are then compared against ECMWF operational analyses. Validation is performed for (i) a dependent sample of 2011JFM, (ii) an independent sample of April 2011 (2011A), and (iii) an independent sample of January to March 2010 (2010JFM). In addition to the RMSE, the Anomaly Correlation Coefficient (ACC) is used. ACC score is sensitive to forecast patterns and insensitive to the model bias, thus it complements the bias sensitive RMSE. Fig. 4.2 illustrates the EN score validation results for the three validation samples. Mean score (continuous line) and the 95% confidence level of the mean (grey vertical bars) are given up to forecast day 10. Notation on both scores is such that positive (negative) values indicate where the optimised (default) model is performing better. The optimised model has improved day three EN scores at the 95% confidence level. Thus, EPPES has found a model which is improved in the target criterion sense. Moreover, the EN scores are improved statistically significantly at all forecast ranges. The improvements also carry to the independent samples, and the 2011A and 2010JFM samples are improved at the 95% confidence level for forecast ranges beyond two and five days, respectively. Next, validation of 500 hPa geopotential height is done for the three samples (Fig. 5 of PAPER III). The RMSE scores for all three periods show statistically significant improvements for

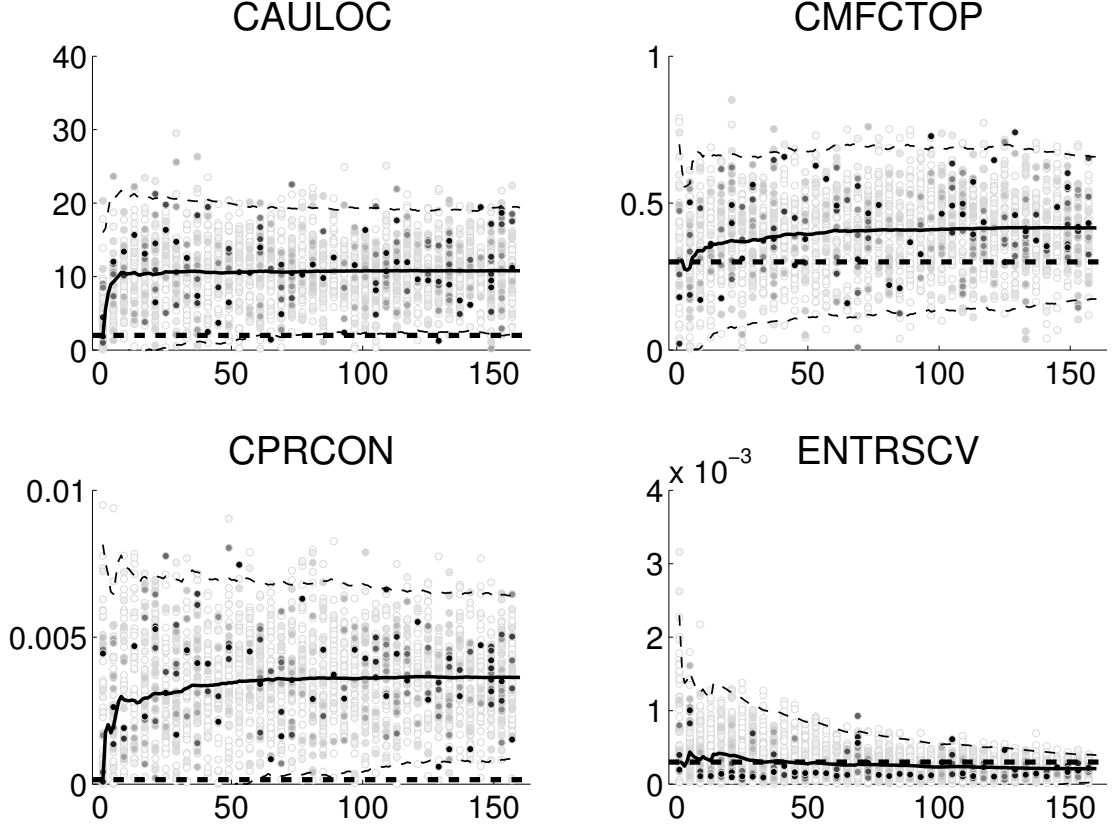


FIGURE 4.1. Time evolution of the parameter subset in 180 consecutive ensembles. A vertical column of markers represents parameter values of one ensemble. The marker shading corresponds to the weight in the distribution update. The parameter distribution mean value  $\mu$  (continuous line),  $\mu \pm 2 \times$  standard deviation (dashed lines) and default parameter values (thick dashed line) are also shown. For clarity, only every fourth ensemble is plotted. Figure from PAPER III.

all forecast ranges. The ACC scores in the 2011JFM sample are improved with 95% confidence level at forecast ranges 2.5 - 8 and 9.5 - 10 days. In the 2011A and 2010JFM samples the ACC scores are either improved or neutral, though the improvements do not hold at the 95% confidence level. All in all, the scores have improved more here than in the experiment conducted in PAPER I, where the 500 hPa geopotential height was specifically targeted for improvements .

To study where the EN improvements originate, Fig. 4.3 represents the zonally-averaged mean EN difference in the 2011JFM sample for forecast ranges of three, six, and 10 days (Figs. 4.3a, b, and c, respectively). Total energy norm (dark blue), and surface pressure (light blue), temperature (dark green) and kinetic energy (light green) terms are shown. Mean error (continuous line), and the 95% confidence interval of the mean (width of the coloured area) are also presented. At forecast day three, most of the total energy norm improvements

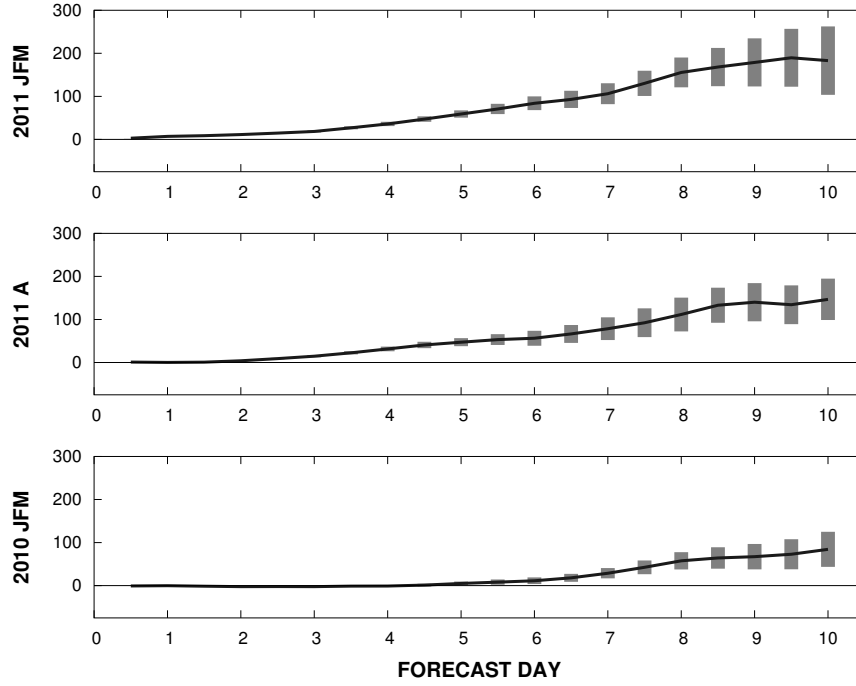


FIGURE 4.2. Energy norm differences between the default and optimised model. Top row: dependent sample (January to March 2011), middle row: independent sample of April 2011, bottom row: independent sample of January to March 2010. Mean difference (continuous line) and 95% confidence interval of the mean (grey bars). Figure from PAPER III.

are located in the tropics, and are dominated by the improvements in the kinetic energy terms. A favourable impact can also be seen in northern hemisphere (north of  $45^{\circ}\text{N}$ ). The only degradation is found in the southern hemisphere ( $25^{\circ}\text{S}$  to  $50^{\circ}\text{S}$ ). The oscillations of the surface pressure term are caused by orographically induced noise originating from the higher resolution analysis data. At longer forecast ranges, the tropical improvements are spread into the mid-latitudes. These mid-latitudinal improvements also grow in magnitude and are less dominated by the kinetic energy improvements the longer the forecasts are. By forecast day six, the largest improvements are found in the mid-latitudes. By forecast day 10 the surface pressure term improvements have exceeded those of the kinetic energy term, and mid-latitudinal improvements become even more pronounced. Note the different scaling in Figs. 4.3a, b, and c.

Figure 4.2 illustrated how the optimised model outperforms the default model the more the longer the forecast gets. This indicates that the optimisation procedure has reduced the model error, since the forecasts are initiated from the same conditions. Furthermore, Fig. 4.3 shows that the model error decrease



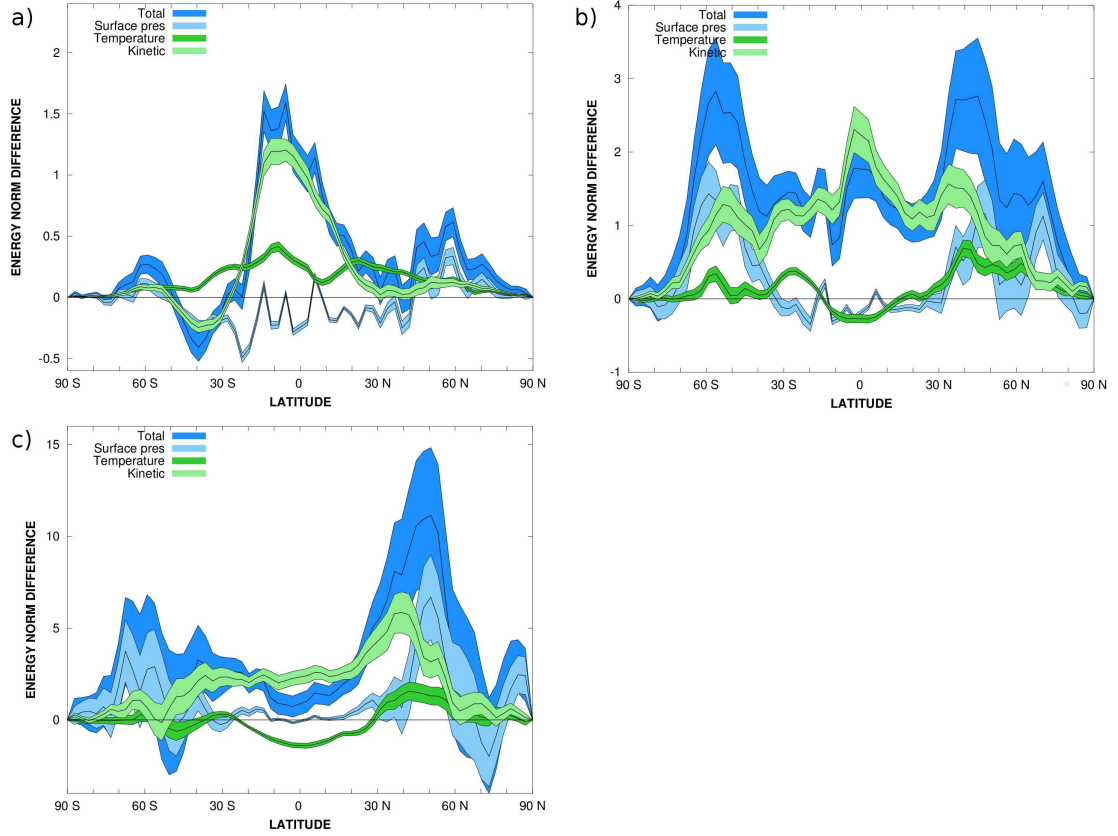


FIGURE 4.3. Zonally-averaged and areal-weighted energy norm differences between the default and optimised model from January to March 2011. a) Forecast day three, b) forecast day six, and c) forecast day 10. Total energy norm (dark blue), and surface pressure (light blue), temperature (dark green) and kinetic energy (light green) terms individually. Continuous black line indicates the mean error, and width of the coloured area represents the 95% confidence interval of the mean. Figure from PAPER III.

mainly affects evolution of kinetic energy in the tropics in the forecasts up to three days. These improvements are then spread by non-linear model dynamics into mid-latitudes. The kinetic energy term retains its dominating part in the total energy norm improvements up to forecast day six. This implies that the improved parametric processes continue to affect the tropical circulation and enhances the realism of the kinetic energy evolution in the tropics throughout the 10 day forecast range. Thus, the energy norm is a very promising target for the EPPES evaluation. The model-wide improvements achieved using the optimised parameter values (Fig. 5 of PAPER III) emphasise the point even further. The concluding answers to Q1–Q3 are then that even though parameter estimation is possible in a NWP model, the choice of target criterion for the estimation is crucial.

## 4.2 PARAMETRIC UNCERTAINTY AND EPS SPREAD GENERATION

Q4 and Q5 are studied through the EPPES produced covariance matrix  $\Sigma$ , which contains the in-between ensemble variability of the parameter values. In the experiments of PAPERS I and II weak parameter covariances begin to emerge even during a three month sampling period. In the PAPER I experiment, after the three months the parameter mean values have drifted away from the default values (Fig. 4.4). A slight tilt can also be observed in the ellipses, representing the parameter covariances, most noticeable between CMFCTOP and CAULOC, indicating correlation between the parameters. For stronger covariances to surface, the number of samples has to be increased; in PAPER I clear correlations are visible after the sampling is repeated 10 times for the same time period (Fig. 4 in PAPER I). The sample size does not necessarily have to be as large as this. Nevertheless, the large number of samples possibly required should not be an obstacle as such, since in EPPES the distribution mean  $\mu$  could even be frozen and only the covariance updated. Thus, one can collect covariance information around the default parameters, though the parameter values still need to be varied in the EPS. The covariance information can then be utilised in various ways, e.g. in detection of model deficiencies, coupling of parameters, and ensemble spread generation.

First, parametrization deficiencies can appear as immoderate parameter uncertainty and/or as weak parameter identifiability. EPPES systematically explores the identifiability of parameters, and can thus potentially discover the deficiencies. Caution is required though, since unidentifiability of a parameter might also be caused by an unsuitable target criterion: variation of a parameter might impose changes to model fields which have only a secondary or tertiary effect on the model fields observed by the target criterion. Thus, no real information about the parameter performance can be gained by monitoring the changes in the cost function. When estimating multiple parameters simultaneously, and a parameter in the set seems to identify poorly, it is crucial to understand whether this is caused by the insensitivity of the parameter to the target criterion or by a parametrization deficiency; target criterion changes triggered by the other parameters could overwhelm the changes caused by the weakly identifying parameter.

Second, strong parameter covariance arising in the estimation process calls attention to possibly coupling some parameters together. Klocke et al. (2011) coupled two of the parameters used in PAPERS I, III and IV (ENTRSCV and CMFCTOP) in their experiments, due to the opposite opposite effect the parameters have on TOA net radiation through their effect on low cloudiness. Interestingly, these two parameters also had a strong covariance in the extended sampling set performed in PAPER I.

Third, stochastic parameter perturbations can be used as a complementary EPS spread generation method, since they represent the model uncertainty. If

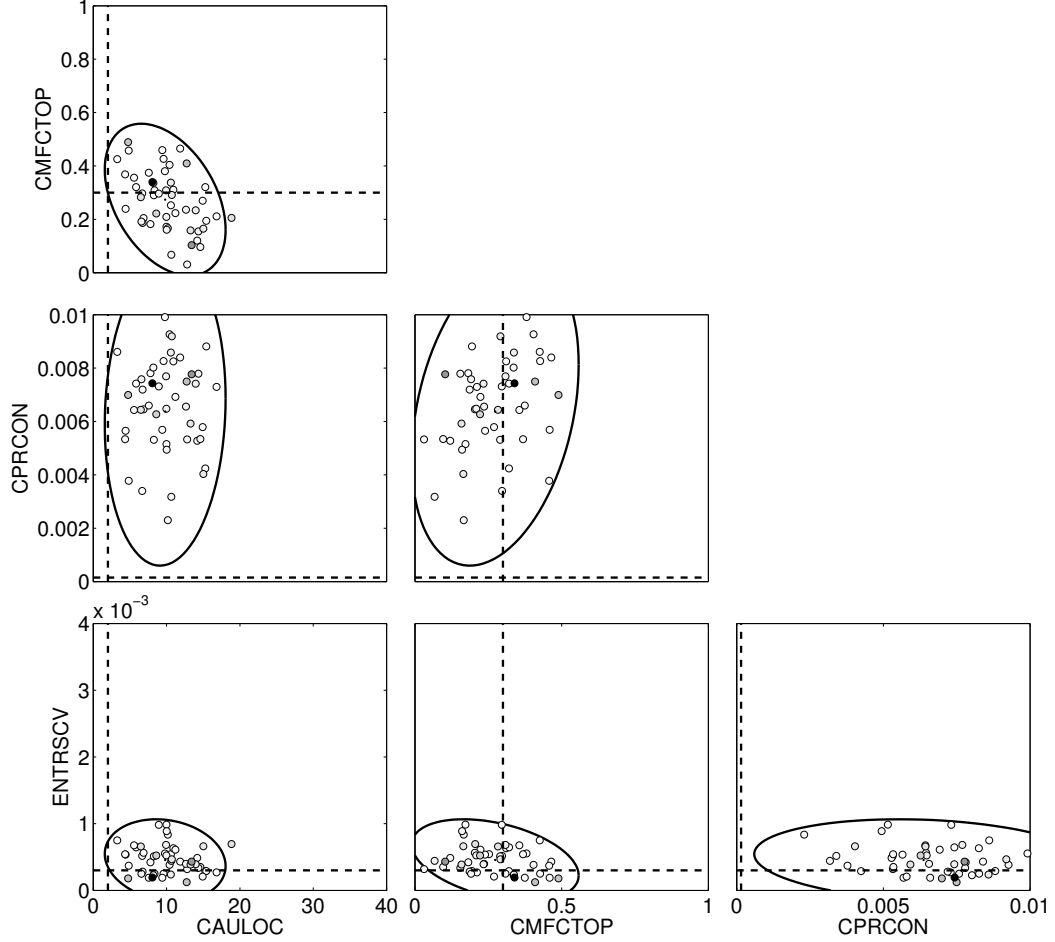


FIGURE 4.4. Pair-wise parameter covariances. Default parameter values (dashed lines) and the parameter covariances after 180 consecutive ensembles (ellipses) are shown. The small markers are the proposed parameter values at the step 180. Figure from PAPER I.

using parameter variations drawn from a uniform distribution, there is a risk of generating parameter sub-sets that correspond to sub-optimal, or even unphysical, models. These would then appear as outlying ensemble members, and deteriorate the skill of the whole ensemble. To potentially alleviate this risk, the covariance  $\Sigma$  already generated in an EPPES experiment could be used “offline”, and utilised to generate parameter sub-sets according to the covariance data. It is important to note that when EPPES provided covariance is used, the parameter values should be treated as stochastic, i.e. re-drawn for each time-window, since the covariance  $\Sigma$  represents the in-between time-windows variability of the optimal closure parameters.

The EPPES sampling itself also produces additional ensemble spread. During the EPPES sampling the ensemble spread generation is done “online”, and the parameter covariances evolve as explained before. Sampling from the initial

covariance, which in most cases is expert defined, can result in the aforementioned sub-optimal models. Therefore, in an operational system, it is crucial to execute the EPPES sampling conservatively, by e.g. starting from a tight initial covariance and/or inhibiting the distribution mean from taking any large steps. Alternatively, one could first run EPPES non-operationally and use the covariance matrix generated as the initial covariance for the operational system. Nevertheless, the initial covariance is updated to a more skillful one quickly during the first few distribution updates, and becomes more realistic the more updates have been performed. In PAPER II the spread generation was tested with the ENS by evaluating the probabilistic skill of the last 90 ensembles ran during the EPPES sampling. In this experiment the ENS was more under-dispersive than the operational system due to the lower-than-operational resolution. Nevertheless, the parameter variation experiment performed better than a reference experiment run without parameter variations (Fig 5 of PAPER II): the parameter perturbations improved the tropical Continuous Ranked Probability Skill Scores (CRPS) in all fields with exception of temperature around 200 hPa. The improvements originate from two possible sources: (i) the increased ensemble spread better matches the RMS error of the ensemble mean, and (ii) the average skill of the ensemble members has been improved as they use parameter values drawn from around the more skillful mean distribution  $\mu$ .

The ensemble spread increase generated by the parameter variations is not by any means additive to the spread generated by e.g. the initial state perturbations. The ECHAM5 EPS emulator used in PAPERS I and III was tested with using only either of the uncertainty sources. Even though both of the sources generate approximately equal amount of ensemble spread, having both uncertainty representations active increases the spread only slightly. Although not additive, the sources are complementary and generate an increased amount of ensemble spread in areas where the other uncertainty source generates only a small amount of it. This is nicely illustrated in Fig. 4.5, where the zonal mean energy norm averaged over 30 ensembles is shown for the ECHAM5 EPS emulator with a) only initial state perturbations active, b) only parameter perturbations active, and c) both sources of uncertainty active. Total energy (dark blue), and surface pressure (light blue), temperature (dark green) and kinetic energy (light green) terms are shown. The width of the coloured area represents  $\pm$  two standard deviations from the mean error (black lines). The complementary nature of the different uncertainty sources can be observed, for instance, in the total energy norm in the southern hemisphere; the parameter perturbations generate little spread in the southern hemisphere, whereas the initial state perturbations generate a lot of it. When both of the perturbations are active, the large spread originating from the initial state perturbations then also keeps the combined spread large. In contrast to this, both of the perturbations generate roughly the same amount of spread

in the northern hemisphere. Although their combined effect does increase the ensemble spread, the spread is not increased in an additive manner.

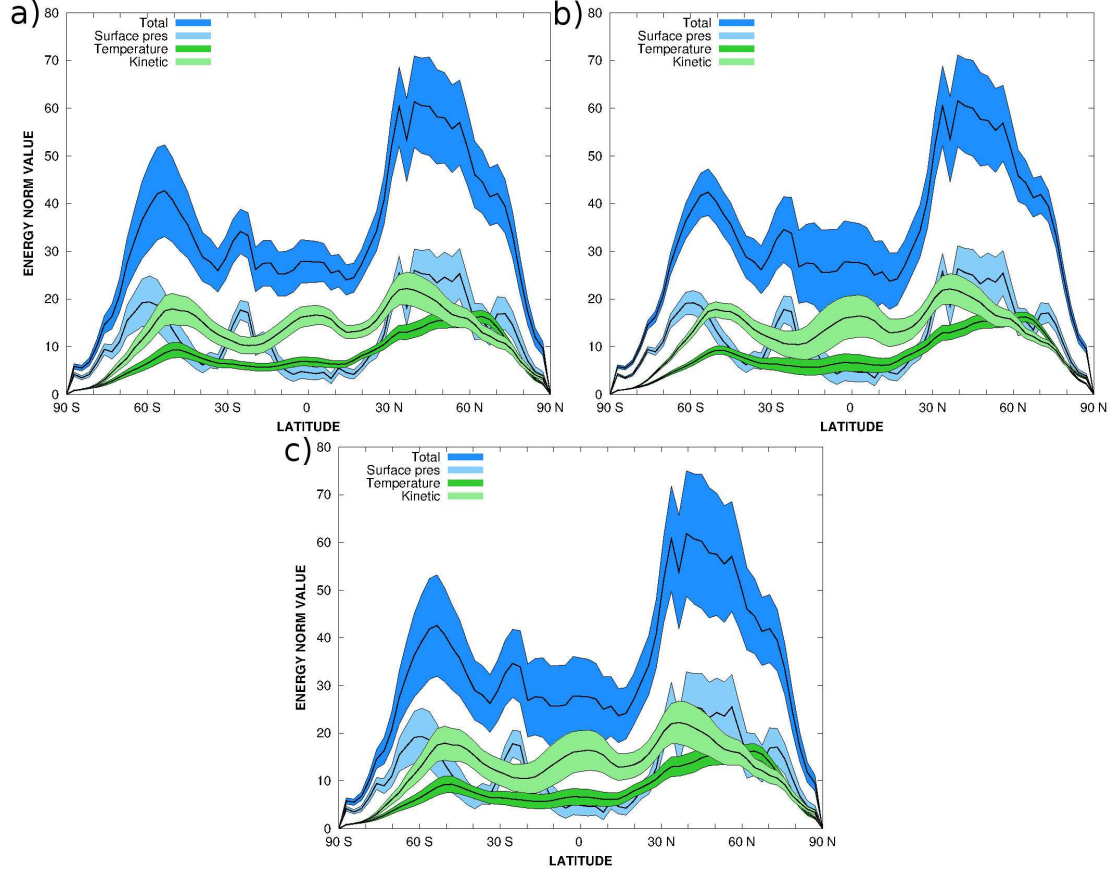


FIGURE 4.5. Ensemble spread of zonally-averaged and areal-weighted energy norm for 15 days (1st to 15th of January 2011) from 72-hour forecast. a) only initial state perturbations, b) only parameter perturbations, and c) both perturbations active. Total energy norm (dark blue), and individual terms; surface pressure (light blue), temperature (dark green) and kinetic energy (light green). Continuous black line indicates the mean model error. Width of the coloured area represents  $\pm$  two standard deviations from the mean.

To conclude, Q4 was studied by highlighting three uses for the parameter covariance data. The answer to Q5 is that the additional ensemble spread caused by parameter perturbations and the average skill increase of the EPS members affect positively the probabilistic skill of an EPS.

### 4.3 FROM WEATHER TO CLIMATE

In order to use climate models for long-term climate simulations, it is essential that the models maintain an observed radiative energy balance at TOA. This

generally involves arduous tuning of the model by hand (see e.g. Mauritsen et al., 2012). Thus, algorithmic tools would (i) help in making the process faster, and moreover (ii) produce a more realistic model objectively.

The idea of optimising the long term predictive skill of a GCM through adaptive MCMC was demonstrated in Järvinen et al. (2010) with ECHAM5. The targets for the optimisation were monthly and annual errors in TOA net radiative forcing. Järvinen et al. (2010) were able to improve the model in the target criterion sense. However, the optimised parameter values did not result in model-wide climatological improvements, i.e. the model was improved only in the target criterion aspect. Simulation of even one year with a climate model is time consuming, and in MCMC a large amount of data points is usually required (Järvinen et al. (2010) used 4500 one year simulations). In order to reduce the computational demand for such a process, PAPER IV introduced the concept of early rejection in MCMC: instead of running the model for one year and calculating the cost function at the end of the run, in early rejection the cost function is split into monthly slots. The cost function is then evaluated after each modelled month and conditionally rejected. In the ECHAM5 experiment of PAPER IV, the model resolution was also increased from the T21 used in Järvinen et al. (2010) to T42, thus computation time saving became even more important. In an experiment which consisted of 3204 simulated years, the early rejection method helped saving 595 years worth of simulations (equaling to approximately 695 hours of saved cpu time). Again, though ECHAM5 was improved in the radiative forcing sense, this improvement did not result in a modelwide improved climatology. Thus, though made possible by the algorithmic development, improving the model climatology through adaptive MCMC did not succeed because of poor choice of the likelihood function. The highly non-linear response of the model climatology to the parameter variations causes the target criterion selection to be an extremely hard problem in this context.

An alternative method to improve the model climatology is next explored in search for an answer to Q6. The hypothesis that improved medium range forecast skill would affect the very long range forecast skill of the model is motivated as follows: By decreasing the modelling uncertainty caused by parametrizations, the model is also altered to be more realistic. Therefore, when the same model is used for climatological runs, one could expect to see the effect of the improved realism of the model. This hypothesis was explored in PAPER I by using the default model and model optimised for medium range forecast skill for a 6-year climate simulation, and comparing the climatology of the models (Fig. 7 of PAPER I) with two observational data-sets, the Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled data-set (Loeb et al., 2009), and the International Satellite Cloud Climatology Project (ISCCP) D2 data (Rossow et al., 1996; Rossow and Duenas, 2004). The optimised model reduced the maximum

monthly mean net radiation biases by about  $10 \text{ Wm}^{-2}$ , and decreased the global annual-mean net radiation bias from default model value of  $1.84 \text{ Wm}^{-2}$  to  $-0.01 \text{ Wm}^{-2}$ . The radiation changes are due to the modified cloud fraction; the tropical region in the optimised model has less high clouds, and slightly more lower tropospheric clouds. The default ECHAM5 has too little cloudiness in the tropics according to the ISCCP data-set. Since the decrease of the high level cloudiness surpasses the increase of low level cloud cover, the optimised model deteriorates the cloud climatology even more. However, a realistic simulation of both clouds and TOA radiation fluxes in ECHAM5 is known to be challenging.

The relationship between medium range forecast improvements and climatology in ECHAM5 is further illustrated in Fig. 4.6. The pressure-latitude cross-sections of climatological averages (Fig. 4.6a,b) and of three day forecasts averaged over 141 cases in 2011JFM (Fig. 4.6c,d) are shown. Negative values in cloud fraction (Fig. 4.6a,c) indicate where the optimised model has less cloud than the default, and negative values in temperature (Fig. 4.6b,d) where the optimised model is colder than the default. The structure of the cloud cover changes in the medium range seem to carry remarkably well into the climatology range. Both the minimum in the cloud cover change around 200 hPa as well as the shape of the area of decrease bear similarity, but the latitudinal location has been switched to the other side of the equator. In addition, the decreased cloudiness in extra-tropics have similar structures in the climatology and medium range forecast. The increased cloud cover centred around 780 hPa in the climatology range can also be observed in the medium range, though it is less intense in the medium range. In the 10-day forecast range (Fig. 6 of PAPER I) these features are even clearer. The temperature response is very different, and more complex. The tropospheric warm anomalies between 800 and 500 hPa, and at the surface have vanished in the climatology run. The cold anomaly centred at 300 hPa has increased substantially, from  $-0.3^\circ\text{K}$  difference to  $-1.5^\circ\text{K}$  difference. There is also a large warming present in the stratosphere (centred around 75 hPa), and an extension from the cold anomaly down to 800 hPa. The 10-day forecast changes are almost identical to the 3-day forecasts, though in the 10-day range the 300 hPa cold anomaly has intensified somewhat and a warm anomaly has appeared around 150 hPa level. Since the cloud fraction and precipitation changes (not shown) in the climatological run are close to the changes in the 3-day forecast, we hypothesise that the cause-effect relationship follows the one presented in PAPER I. The absence of the warm anomalies are then caused by some secondary effects, surfacing only in the longer model run. Nevertheless, ECHAM5 equatorial troposphere is too warm and stratosphere too cold when compared to ERA40 reanalysis data (Upala et al., 2005). Thus, the climatological changes bring the model closer to the analysed climatology. The answer to Q6 is then that at least some of the changes in the model features present already at early medium range forecasts carry on

to the very long range, too. Thus, instead of using computationally demanding long model simulations in tuning the model climatology, a less consuming medium range forecast could be utilised to reach the goal.

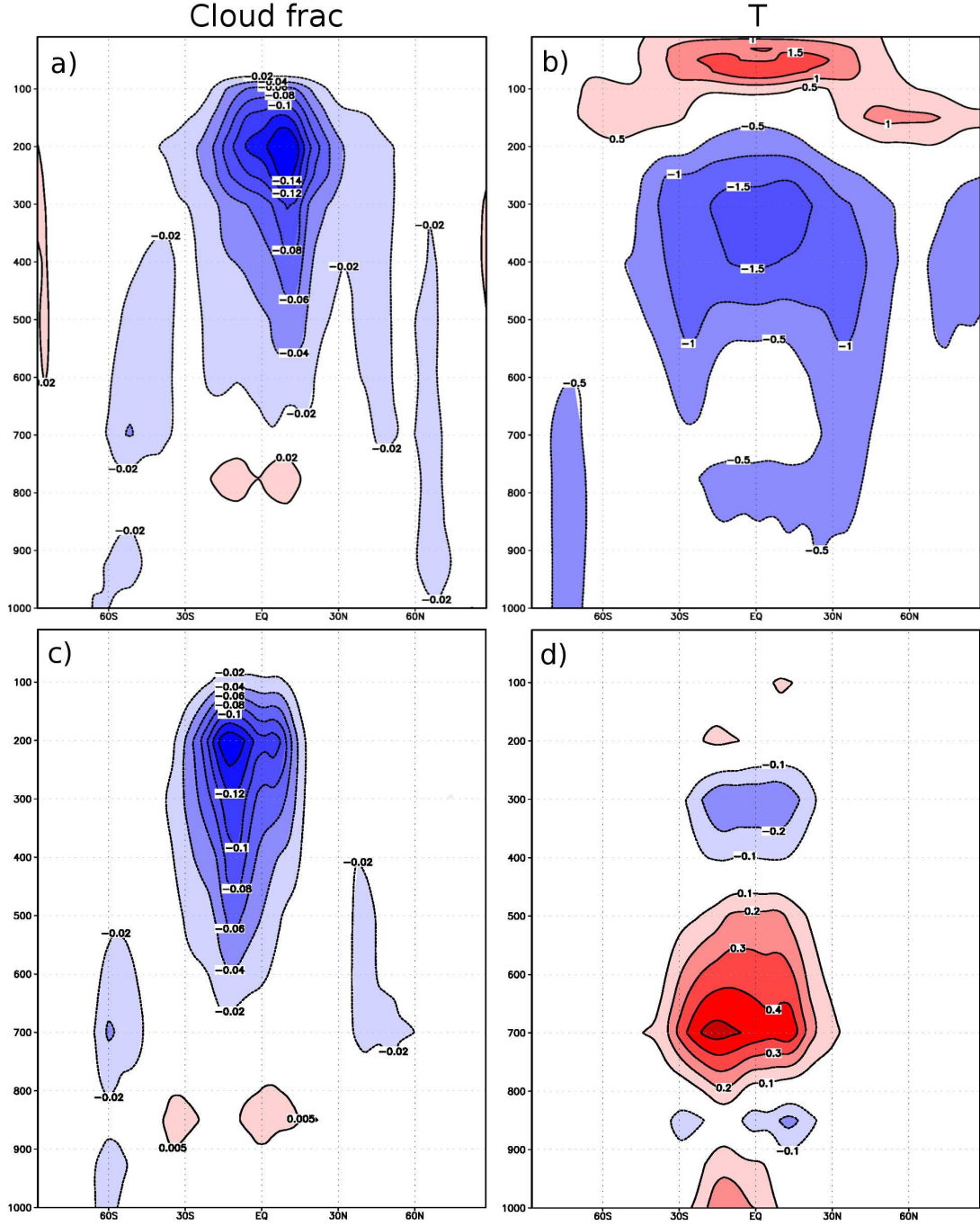


FIGURE 4.6. Pressure-latitude cross-section of the cloud fraction (left panel; non-dimensional values between 0 and 1) and temperature (right panel; unit K). Differences between the optimised and default models in 6-year climatological average (a,b), and in day three forecasts averaged over 141 forecasts (c,d).



In addition, the  $\Sigma$  information could also be used here; parameter perturbations drawn from  $\Sigma$  represent the parameteric model uncertainty, and, moreover, they are physically justified. Thus, by generating an ensemble of long climate simulations with these perturbed parameters the contribution of the model specific uncertainty can be assessed. If the same would be done with parameter values drawn randomly, there will likely be a number of poorly performing models included in the ensemble. These would then widen the uncertainty range of the simulations, and provide with an unrealistically large estimates of the model sensitivity.

## 5 CONCLUSIONS AND FUTURE DIRECTIONS

Applying the atmospheric governing equations in predicting the future state of the atmosphere requires discretization of the equations. To complement the discretized equation set, subgrid-scale physical processes have to be described in the model. The parametrization of these processes then strongly influence the accuracy of Numerical Weather Prediction (NWP) models. Imperfect representation of the processes introduces a parametric uncertainty into the forecasts. An important aspect in the uncertainties related to parametrizations comes from so-called closure problem; the closer the parametrizations go towards describing the phenomena in molecular level, the more lacking knowledge there is about the processes. Thus, at some point further modelling has to stop, and some closure parameter has to be set to e.g. define the rate at which a sub-process takes place, or to describe the efficiency of a sub-process. The closure parameters therefore influence the realism of the parametrizations and furthermore affect the forecast ability of the model. This Thesis has studied the forecast uncertainties related to the closure parameters from three aspects: (i) objectively estimating optimal values of the closure parameters, (ii) utilising the knowledge of closure parameter uncertainties for identifying problems within the parametrizations and constructing an improved Ensemble Prediction System (EPS), and (iii) showing how closure parameter changes in medium range forecasts relate to climatology of the model.

First, in order to study the closure parameter optimisation three research questions were posed:

- Q1 Can optimal parameter values be found algorithmically in a low resolution GCM of full complexity?
- Q2 Is parameter optimisation feasible in a system already at high level of forecast skill?
- Q3 How does the choice of target criterion affect the parameter estimation?

In search for answers to these questions, the Ensemble and Parameter Estimation System (EPPES; Järvinen et al., 2012; Laine et al., 2012) is used to evaluate closure parameters related to convection and cloud processes. The EPPES methodology is experimented with ECHAM5 climate model and the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Q1 and Q2 are studied through conducting parameter estimation by targeting 500 hPa geopotential height mean squared errors at forecast days three and ten. The EPPES estimation is able to find a closure parameter set corresponding to improved model in the target criterion sense with both used models. The study of Q3 shows that though improved in the target criterion sense, the models optimised for 500 hPa geopotential height have deficiencies in the upper level model fields. In order to find a target criterion leading to improved model

in most forecast fields, atmospheric total energy norm (EN) is then experimented with ECHAM5 by targeting EN errors at forecast day three. The estimation procedure is again able to optimise the closure parameter set with respect to the target criterion. Moreover, the EN error reduction is more pronounced the longer the forecasts are. This indicates that the optimisation has reduced the modelling error caused by the parametrizations. The EN improvements originate from more realistic kinetic energy representation in the tropics. At longer forecast ranges the improvement is also spread to higher latitudes by non-linear model dynamics. Furthermore, the optimised parameter set also improves the model with respect to most forecast quantities. Therefore, model closure parameter optimization seems to be a viable, and effective, way of reducing parameteric uncertainty without major structural changes inside the parametrizations. Although, the choice of target criterion has to be considered carefully prior to the estimation.

Second, the EPPES provided parameter uncertainties and covariances are studied in order to find answers to:

Q4 Is there any useful information gained from studying the parameter covariance data?

Q5 Do parameter perturbations affect the probabilistic skill of an EPS?

Answer to Q4 emphasis three possible uses: a) large parameter uncertainties could indicate deficiencies in the parametrizations, b) strong parameter correlations found would suggest need of coupling of parameters, and c) additional ensemble spread could be generated by introducing parameter variations into an EPS, and drawing parameter values from the EPPES generated parameter distributions. The EPPES sampling in itself also produces additional spread into an EPS. Answer to Q5 is found through experiments with the ECMWF Ensemble Prediction System (ENS), in which ensembles generated with EPPES estimation active were more skillful than default ensembles. This is due to increased ensemble spread and improved average skill of the ensemble members. Thus, in addition to finding optimal closure parameter values, the skill of an EPS benefits from utilising EPPES-style parameter perturbations.

Third, a hypothetical link between model medium range forecast skill and very long range forecast skill is studied through the following question:

Q6 Does medium range parameter optimisation have any relevance for model climatology?

The hypothesis is verified as the results indicate that model medium range and very long range improvements might be attainable simultaneously. The structural changes of cloud cover in medium range can be identified in the model climatology. In temperature fields the change structures do not carry to the very long range as well. Nevertheless, if universal in models, this connection could be used to improve

the very long range predictive skill of climate models by simply enhancing their medium range forecast skill.

The parameter evaluation through targeting EN errors is currently tested with the IFS. Similarly to the earlier experiments, the effects of stochastic noise will be verified. Inclusion of latent energy term in the EN remains still to be experimented on. The estimation in the experiments conducted in PAPERS I, II and III was done, and validated, only in a very seasonal samples; in PAPERS I and III during winter and early spring, and in PAPER II during summer. Thus, it would be of interest to study if the closure parameters have any annually cycling optimal values. Similarly, geographically dividing and optimising the closure parameters (see Wu et al., 2012) would also be an attractive topic of research. Lastly, though the initial results for the connection of medium range and climatology look promising, there is clearly need for more study on this. Particularly whether these structures can be observed in other model setups, and possibly seen in other model fields too, needs to be answered.

## REFERENCES

- Aksoy, A., Zhang, F., and Nielsen-Gammon, J.: Ensemble-based simultaneous state and parameter estimation in a two-dimensional sea-breeze model, *Mon. Wea. Rev.*, 134, 2951–2970, doi:<http://dx.doi.org/10.1175/MWR3224.1>, 1996.
- Aksoy, A., Zhang, F., and Nielsen-Gammon, J.: Ensemble-based simultaneous state and parameter estimation with MM5, *Geophys. Res. Lett.*, 33, L12 801, doi:10.1029/2006GL026186, 2006.
- Annan, J. D., Edwards, N. R., and Marsh, R.: Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter, *Ocean Modell.*, 8, 135–154, doi:<http://dx.doi.org/10.1016/j.ocemod.2003.12.004>, 2005.
- Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., Vitart, F., and Balsamo, G.: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales, *Quarterly Journal of the Royal Meteorological Society*, 134, 1337–1351, doi:10.1002/qj.289, 2008.
- Bender, F. A.-M.: A note on the effect of GCM tuning on climate sensitivity, *Environmental Research Letters*, 3, 014 001, doi:10.1088/1748-9326/3/1/014001, 2008.
- Betts, A. K., Ball, J. H., and Viterbo, P.: Evaluation of the ERA-40 surface water budget and surface temperature for the mackenzie river basin, *J. Hydrometeor.*, 4, 1194–1211, doi:[http://dx.doi.org/10.1175/1525-7541\(2003\)004<1194:EOTESW>2.0.CO;2](http://dx.doi.org/10.1175/1525-7541(2003)004<1194:EOTESW>2.0.CO;2), 2003.
- Buizza, R., Tribbia, J., Molteni, F., and Palmer, T. N.: Computation of optimal unstable structures for a numerical weather prediction model, *Tellus A*, 45, 388–407, doi:10.1034/j.1600-0870.1993.t01-4-00005.x, 1993.
- de Rooy, W. C., Bechtold, P., Fröhlich, K., Hohenegger, C., Jonker, H., Mironov, D., Pier Siebesma, A., Teixeira, J., and Yano, J.-I.: Entrainment and detrainment in cumulus convection: an overview, *Quarterly Journal of the Royal Meteorological Society*, 139, 1–19, doi:10.1002/qj.1959, 2013.
- Errico, R. M.: Interpretations of the total energy and rotational energy norms applied to determination of singular vectors, *Q.J.R. Meteorol. Soc.*, 126, 1581–1599, doi:10.1002/qj.49712656602, 2000.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research: Oceans*, 99, 10 143–10 162, doi:10.1029/94JC00572, 1994.

- Farrel, B.: Optimal excitation of neutral Rossby waves, *J. Atmos. Sci.*, 45, 163–172, doi:[http://dx.doi.org/10.1175/1520-0469\(1988\)045<0163:OEONRW>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1988)045<0163:OEONRW>2.0.CO;2), 1988.
- Grasso, L. D.: The differentiation between grid spacing and resolution and their application to numerical modeling, *Bull. Amer. Meteor. Soc.*, 81, 579580, 2000.
- Haario, H., Laine, M., Mira, A., and Saksman, E.: DRAM: Efficient adaptive MCMC, *Statistics and Computing*, 16, 339–354, doi:10.1007/s11222-006-9438-0, 2006.
- Hu, X.-M., Zhang, F., and Nielsen-Gammon, J. W.: Ensemble-based simultaneous state and parameter estimation for treatment of mesoscale model error: A real-data study, *Geophysical Research Letters*, 37, L08 802, doi:10.1029/2010GL043017, 2010.
- Jackson, C. S., Sen, M. K., Huerta, G., Deng, Y., and Bowman, K. P.: Error reduction and convergence in climate prediction, *J. Climate*, 21, 6698–6709, doi:<http://dx.doi.org/10.1175/2008JCLI2112.1>, 2008.
- Järvinen, H., Räisänen, P., Laine, M., Tamminen, J., Ilin, A., Oja, E., Solonen, A., and Haario, H.: Estimation of ECHAM5 climate model closure parameters with adaptive MCMC, *Atmos. Chem. Phys.*, 10, 9993–10 002, doi:10.5194/acp-10-9993-2010, 2010.
- Järvinen, H., Laine, M., Solonen, A., and Haario, H.: Ensemble prediction and parameter estimation system: the concept, *Q.J.R. Meteorol. Soc.*, 138, 281–288, doi:10.1002/qj.923, 2012.
- Kalman, R. E.: A new approach to linear filtering and prediction problems, *J. Basic Engineering*, 82D, 33–45, 1960.
- Kalnay, E.: *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 2003.
- Kivman, G.: Sequential parameter estimation for stochastic systems, *Nonlinear Processes Geophys.*, 10, 253–256, doi:10.5194/npg-10-253-2003, 2003.
- Klocke, D., Pincus, R., and Quaas, J.: On constraining estimates of climate sensitivity with present-day observations through model weighting, *J. Climate*, 24, 6092–6099, doi:<http://dx.doi.org/10.1175/2011JCLI4193.1>, 2011.
- Laine, M., Solonen, A., Haario, H., and Järvinen, H.: Ensemble prediction and parameter estimation system: the method, *Q.J.R. Meteorol. Soc.*, 138, 289–297, doi:10.1002/qj.922, 2012.

- Leutbecher, M. and Palmer, T. N.: Ensemble forecasting, *Journal of Computational Physics*, 227, 3515–3539, doi:10.1016/j.jcp.2007.02.014, 2008.
- Llosa, S. and Zodrow, I.: Disaster risk reduction legislation as a basis for effective adaptation, *Global Assessment Report on Disaster Risk Reduction: Revealing Risk, Redefining Development*, 2011.
- Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., Manalo-Smith, N., and Wong, T.: Toward optimal closure of the earth’s top-of-atmosphere radiation budget, *J. Climate*, 22, 748–766, doi: <http://dx.doi.org/10.1175/2008JCLI2637.1>, 2009.
- Lohmann, U. and Roeckner, E.: Design and performance of a new cloud microphysics scheme developed for the ECHAM general circulation model, *Climate Dynamics*, 12, 557–572, doi:10.1007/BF00207939, 1996.
- Lorenz, E. N.: Available potential energy and the maintenance of the general circulation, *Tellus*, 7, 157–167, doi:10.1111/j.2153-3490.1955.tb01148.x, 1955.
- Lorenz, E. N.: Energy and numerical weather prediction, *Tellus*, 12, 364–373, doi:10.1111/j.2153-3490.1960.tb01323.x, 1960.
- Lorenz, E. N.: Deterministic nonperiodic flow, *Journal of the atmospheric sciences*, 20, 130–141, doi:[http://dx.doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2), 1963.
- Lorenz, E. N.: The predictability of a flow which possesses many scales of motion, *Tellus*, 21, 289–307, doi:10.1111/j.2153-3490.1969.tb00444.x, 1969.
- Lorenz, E. N.: Predictability: A problem partly solved, in *Proceedings of seminar on predictability: Volume 1*. ECMWF, Reading, UK, 1, 1–18, 1996.
- Mapes, B. E.: Equilibrium vs. activation control of large scale variations of tropical deep convection, *NATO ASI Series C Mathematical and Physical Sciences*, 505, 321–358, 1997.
- Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H., and Tomassini, L.: Tuning the climate of a global model, *Journal of Advances in Modeling Earth Systems*, 4, M00A01, doi:10.1029/2012MS000154, 2012.
- Nordeng, T. E.: Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics, *Tech. rep.*, European Centre for Medium-Range Weather Forecasts, 1994.

- Palmer, T. N., Buizza, R., Molteni, F., Chen, Y.-Q., and Corti, S.: Singular vectors and the predictability of weather and climate, *Phil. Trans. R. Soc. Lond. A*, 348, 459–475, doi:10.1098/rsta.1994.0105, 1994.
- Randall, D. A., Abeles, J. A., and Corsetti, T. G.: Seasonal Simulations of the Planetary Boundary Layer and Boundary-Layer Stratocumulus Clouds with a General Circulation Model, *J. Atmos. Sci.*, 42, 641–676, doi:http://dx.doi.org/10.1175/1520-0469(1985)042<0641:SSOTPB>2.0.CO;2, 1985.
- Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5, Part I: Model description, Tech. Rep. 349, Max-Planck-Institut für Meteorologie, 2003.
- Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model, *Journal of Climate*, 19, 3771–3791, doi:http://dx.doi.org/10.1175/JCLI3824.1, 2006.
- Rossow, W. and Duenas, E.: The international satellite cloud climatology project (ISCCP) web site: An online resource for research, *Bulletin of the American Meteorological Society*, 85, 167–172, 2004.
- Rossow, W., Garder, L., Lu, P., and Walker, A.: International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud dataset, Tech. Rep. WMO/TD-737 349, World Meteorological Organization, 1996.
- Ruuhela, R., Ruotsalainen, J., Aschan, C., Torkki, M., and Kangas, M.: Preventing pedestrian slipping accidents with help of weather service, in: 17th International Congress of Biometeorology, ICB, pp. 330–332, 2005.
- Savijärvi, H.: Error Growth in a Large Numerical Forecast System, *Mon. Weather Rev.*, 123, 212–211, doi:10.1175/1520-0493(1995)123<0212:EGIALN>2.0.CO;2, 1995.
- Savijärvi, H. and Vihma, T.: *Rajakerroksen fysiikka I*, University of Helsinki, course literature, 2001.
- Schirber, S., Klocke, D., Pincus, R., Quaas, J., and Anderson, J.: Parameter estimation using data assimilation in an atmospheric general circulation model: From a perfect towards the real world, *J. Adv. Model Earth Syst.*, 5, 58–70, doi:10.1029/2012MS000167, 2013.



- Schneider, T. and Sobel, A. H., eds.: *The Global Circulation of the Atmosphere*, Princetown University Press, 2007.
- Slingo, J. and Palmer, T.: Uncertainty in weather and climate prediction, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369, 4751–4767, doi:10.1098/rsta.2011.0161, 2011.
- Stensrud, D. J.: *Parameterization Schemes: Keys to understanding Numerical Weather Prediction Models*, Cambridge University Press, 2007.
- Tiedtke, M.: A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, *Mon. Wea. Rev.*, 117, 1779–1800, doi: [http://dx.doi.org/10.1175/1520-0493\(1989\)117<1779:ACMFSF>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2), 1989.
- Trenberth, K. E. and Guillemot, C. J.: Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses, *Climate Dynamics*, 14, 213–231, doi:10.1007/s003820050219, 1998.
- Uppala, S. M., Kållberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., et al.: The ERA-40 reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.
- van Leeuwen, P.: A variance-minimizing filter for large-scale applications, *Mon. Wea. Rev.*, 131, 2071–2084, doi:[http://dx.doi.org/10.1175/1520-0493\(2003\)131<2071:AVFFLA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(2003)131<2071:AVFFLA>2.0.CO;2), 2003.
- Wilks, D. S.: Effects of stochastic parametrizations in the Lorenz '96 system, *Quarterly Journal of the Royal Meteorological Society*, 131, 389–407, doi: 10.1256/qj.04.03, 2005.
- Wu, X., Zhang, S., Liu, Z., Rosati, A., Delworth, T. L., and Liu, Y.: Impact of geographic-dependent parameter optimization on climate estimation and prediction: Simulation with an intermediate coupled model, *Monthly Weather Review*, 140, 3956–3971, doi:<http://dx.doi.org/10.1175/MWR-D-11-00298.1>, 2012.
- Yang, B., Qian, Y., Lin, G., Leung, R., and Zhang, Y.: Some issues in uncertainty quantification and parameter tuning: a case study of convective parameterization scheme in the WRF regional climate model, *Atmospheric Chemistry and Physics*, 12, 2409–2427, doi:10.5194/acp-12-2409-2012, 2012.

## **Finnish Meteorological Institute Contributions**

1. Joffre, Sylvain M., 1988. Parameterization and assessment of processes affecting the long-range transport of airborne pollutants over the sea. 49 p.
2. Solantie, Reijo, 1990. The climate of Finland in relation to its hydrology, ecology and culture. 130 p.
3. Joffre, Sylvain M. and Lindfors, Virpi, 1990. Observations of airborne pollutants over the Baltic Sea and assessment of their transport, chemistry and deposition. 41 p.
4. Lindfors, Virpi, Joffre, Sylvain M. and Damski, Juhani, 1991. Determination of the wet and dry deposition of sulphur and nitrogen compounds over the Baltic Sea using actual meteorological data. 111 p.
5. Pulkkinen, Tuija, 1992. Magnetic field modelling during dynamic magnetospheric processes. 150 p.
6. Lönnberg, Peter, 1992. Optimization of statistical interpolation. 157 p.
7. Viljanen, Ari, 1992. Geomagnetic induction in a one- or two-dimensional earth due to horizontal ionospheric currents. 136 p.
8. Taalas, Petteri, 1992. On the behaviour of tropospheric and stratospheric ozone in Northern Europe and in Antarctica 1987-90. 88 p.
9. Hongisto, Marke, 1992. A simulation model for the transport, transformation and deposition of oxidized nitrogen compounds in Finland — 1985 and 1988 simulation results. 114 p.
10. Taalas, Petteri, 1993. Factors affecting the behaviour of tropospheric and stratospheric ozone in the European Arctic and Antarctica. 138 s.
11. Mälkki, Anssi, 1993. Studies on linear and non-linear ion waves in the auroral acceleration region. 109 p.
12. Heino, Raino, 1994. Climate in Finland during the period of meteorological observations. 209 p.
13. Janhunen, Pekka, 1994. Numerical simulations of E-region irregularities and ionosphere-magnetosphere coupling. 122 p.
14. Hillamo, Risto E., 1994. Development of inertial impactor size spectroscopy for atmospheric aerosols. 148 p.
15. Pakkanen, Tuomo A., 1995. Size distribution measurements and chemical analysis of aerosol components. 157 p.

16. Kerminen, Veli-Matti, 1995. On the sulfuric acid-water particles via homogeneous nucleation in the lower troposphere. 101 p.
17. Kallio, Esa, 1996. Mars-solar wind interaction: Ion observations and their interpretation. 111 p.
18. Summanen, Tuula, 1996. Interplanetary Lyman alpha measurements as a tool to study solar wind properties. 114 p.
19. Rummukainen, Markku, 1996. Modeling stratospheric chemistry in a global three-dimensional chemical transport model, SCTM-1. Model development. 206 p.
20. Kauristie, Kirsti, 1997. Arc and oval scale studies of auroral precipitation and electrojets during magnetospheric substorms. 134 p.
21. Hongisto, Marke, 1998. Hilatar, A regional scale grid model for the transport of sulphur and nitrogen compounds. 152 p.
22. Lange, Antti A.I., 1999. Statistical calibration of observing systems. 134 p.
23. Pulkkinen, Pentti, 1998. Solar differential rotation and its generators: computational and statistical studies. 108 p.
24. Toivanen, Petri, 1998. Large-scale electromagnetic fields and particle drifts in time-dependent Earth's magnetosphere. 145 p.
25. Venäläinen, Ari, 1998. Aspects of the surface energy balance in the boreal zone. 111 p.
26. Virkkula, Aki, 1999. Field and laboratory studies on the physical and chemical properties of natural and anthropogenic tropospheric aerosol. 178 p.
27. Siili, Tero, 1999. Two-dimensional modelling of thermal terrain-induced mesoscale circulations in Mars' atmosphere. 160 p.
28. Paatero, Jussi, 2000. Deposition of Chernobyl-derived transuranium nuclides and short-lived radon-222 progeny in Finland. 128 p.
29. Jalkanen, Liisa, 2000. Atmospheric inorganic trace contaminants in Finland, especially in the Gulf of Finland area. 106 p.
30. Mäkinen, J. Teemu, T. 2001. SWAN Lyman alpha imager cometary hydrogen coma observations. 134 p.
31. Rinne, Janne, 2001. Application and development of surface layer flux techniques for measurements of volatile organic compound emissions from vegetation. 136 p.

32. Syrjäsoo, Mikko T., 2001. Auroral monitoring system: from all-sky camera system to automated image analysis. 155 p.
33. Karppinen, Ari, 2001. Meteorological pre-processing and atmospheric dispersion modelling of urban air quality and applications in the Helsinki metropolitan area. 94 p.
34. Hakola, Hannele, 2001. Biogenic volatile organic compound (VOC) emissions from boreal deciduous trees and their atmospheric chemistry. 125 p.
35. Merenti-Välimäki, Hanna-Leena, 2002. Study of automated present weather codes. 153 p.
36. Tanskanen, Eija I., 2002. Terrestrial substorms as a part of global energy flow. 138 p.
37. Nousiainen, Timo, 2002. Light scattering by nonspherical atmospheric particles. 180 p.
38. Härkönen, Jari, 2002. Regulatory dispersion modelling of traffic-originated pollution. 103 p.
39. Oikarinen, Liisa, 2002. Modeling and data inversion of atmospheric limb scattering measurements. 111 p.
40. Hongisto, Marke, 2003. Modelling of the transport of nitrogen and sulphur contaminants to the Baltic Sea Region. 188 p.
41. Palmroth, Minna, 2003. Solar wind – magnetosphere interaction as determined by observations and a global MHD simulation. 147 p.
42. Pulkkinen, Antti, 2003. Geomagnetic induction during highly disturbed space weather conditions: Studies of ground effects 164 p.
43. Tuomenvirta, Heikki, 2004. Reliable estimation of climatic variations in Finland. 158 p.
44. Ruoho-Airola, Tuija, 2004. Temporal and regional patterns of atmospheric components affecting acidification in Finland. 115 p.
45. Partamies, Noora, 2004. Meso-scale auroral physics from groundbased observations. 122 p.
46. Teinilä, Kimmo, 2004. Size resolved chemistry of particulate ionic compounds at high latitudes. 138 p.
47. Tamminen, Johanna, 2004. Adaptive Markov chain Monte Carlo algorithms with geophysical applications. 156 p.

48. Huttunen, Emilia, 2005. Interplanetary shocks, magnetic clouds, and magnetospheric storms. 142 p.
49. Sofieva, Viktoria, 2005. Inverse problems in stellar occultation. 110 p.
50. Harri, Ari-Matti, 2005. In situ observations of the atmospheres of terrestrial planetary bodies. 246 p.
51. Aurela, Mika, 2005. Carbon dioxide exchange in subarctic ecosystems measured by a micrometeorological technique. 132 p.
52. Damski, Juhani, 2005. A Chemistry-transport model simulation of the stratospheric ozone for 1980 to 2019. 147 p.
53. Tisler, Priit, 2006. Aspects of weather simulation by numerical process. 110 p.
54. Arola, Antti, 2006. On the factors affecting short- and long-term UV variability. 82 p.
55. Verronen, Pekka T., 2006. Ionosphere-atmosphere interaction during solar proton events. 146 p.
56. Hellén, Heidi, 2006. Sources and concentrations of volatile organic compounds in urban air. 134 p.
57. Pohjola, Mia, 2006. Evaluation and modelling of the spatial and temporal variability of particulate matter in urban areas. 143 p.
58. Sillanpää, Markus, 2006. Chemical and source characterisation of size-segregated urban air particulate matter. 184 p.
59. Niemelä, Sami, 2006. On the behaviour of some physical parameterization methods in high resolution numerical weather prediction models. 136 p.
60. Karpechko, Alexey, 2007. Dynamical processes in the stratosphere and upper troposphere and their influence on the distribution of trace gases in the polar atmosphere. 116 p.
61. Eresmaa, Reima, 2007. Exploiting ground-based measurements of Global Positioning System for numerical weather prediction. 95 p.
62. Seppälä, Annika, 2007. Observations of production and transport of NO<sub>x</sub> formed by energetic particle precipitation in the polar night atmosphere. 100 p.
63. Rontu, Laura, 2007. Studies on orographic effects in a numerical weather prediction model. 151 p.
64. Vajda, Andrea, 2007. Spatial variations of climate and the impact of disturbances on local climate and forest recovery in northern Finland. 139 p.

65. Laitinen, Tiera, 2007. Rekonnektio Maan magnetosfäärissä – Reconnection in Earth's magnetosphere. 226 s.
66. Vanhamäki, Heikki, 2007. Theoretical modeling of ionospheric electrodynamics including induction effects. 170 p.
67. Lindfors, Anders, 2007. Reconstruction of past UV radiation. 123 p.
68. Sillanpää, Ilkka, 2008. Hybrid modelling of Titan's interaction with the magnetosphere of Saturn. 200 p.
69. Laine, Marko, 2008. Adaptive MCMC methods with applications in environmental and geophysical models. 146 p.
70. Tanskanen, Aapo, 2008. Modeling of surface UV radiation using satellite data. 109 p.
71. Leskinen, Ari, 2008. Experimental studies on aerosol physical properties and transformation in environmental chambers. 116 p.
72. Tarvainen, Virpi, 2008. Development of biogenetic VOC emission inventories for the boreal forest. 137 p.
73. Lohila, Annalea, 2008. Carbon dioxide exchange on cultivated and afforested boreal peatlands. 110 p.
74. Saarikoski, Sanna, 2008. Chemical mass closure and source-specific composition of atmospheric particles. 182 p.
75. Pirazzini, Roberta, 2008. Factors controlling the surface energy budget over snow and ice. 141 p.
76. Salonen, Kirsti, 2008. Towards the use of radar winds in numerical weather prediction. 87 p.
77. Luojus, Kari, 2009. Remote sensing of snow-cover for the boreal forest zone using microwave radar. 178 p.
78. Juusola, Liisa, 2009. Observations of the solar wind-magnetosphere-ionosphere coupling. 167 p.
79. Waldén, Jari, 2009. Meteorology of gaseous air pollutants. 177 p.
80. Mäkelä, Jakke, 2009. Electromagnetic signatures of lightning near the HF frequency band. 152 p.
81. Thum, Tea, 2009. Modelling boreal forest CO<sub>2</sub> exchange and seasonality. 140 p.
82. Lallo, Marko, 2010. Hydrogen soil deposition and atmospheric variations in the boreal zone. 91 p.

83. Sandroos, Arto, 2010. Shock acceleration in the solar corona. 116 p.
84. Lappalainen, Hanna, 2010. Role of temperature in the biological activity of a boreal forest. 107 p.
85. Mielonen, Tero, 2010. Evaluation and application of passive and active optical remote sensing methods for the measurement of atmospheric aerosol properties. 125 p.
86. Lakkala, Kaisa, 2010. High quality polar UV measurements : scientific analyses and transfer of the irradiance scale. 156 p.
87. Järvinen, Riku, 2011. On ion escape from Venus. 150 p.
88. Saltikoff, Elena, 2011. On the use of weather radar for mesoscale applications in northern conditions. 120 p.
89. Timonen, Hilikka, 2011. Chemical characterization of urban background aerosol using online and filter methods. 176 p.
90. Hyvärinen, Otto, 2011. Categorical meteorological products : evaluation and analysis. 138 p.
91. Mäkelä, Antti, 2011. Thunderstorm climatology and lightning location applications in northern Europe. 158 p.
92. Hietala, Heli, 2012. Multi-spacecraft studies on space plasma shocks. 132 p.
93. Leinonen, Jussi, 2013. Thunderstorm climatology and lightning location applications in northern Europe.
94. Gregow, Hilppa, 2013. Impact of strong winds, heavy snow loads and soil frost conditions on the risks to forests in Northern Europe. 178 p.
95. Henriksson, Svante, 2013. Modeling key processes causing climate change and variability. 98 p.
96. Valkonen, Teresa, 2013. Surface influence on the marine and coastal Antarctic atmosphere. 114 p.
97. Saarnio, Karri, 2013. Chemical characterisation of fine particles from biomass burning. 180 p.
98. Honkonen, Ilja, 2013. Development, validation and application of numerical space environment models. 148 p.
99. Siljamo, Pilvi, 2013. Numerical modelling of birch pollen emissions and dispersion on regional and continental scales. 156 p.

100. Anekallu, Chandrasekhar, 2013. Energy conversion across the Earth's magnetopause observations. 166 p.
101. Riihelä, Aku, 2013. The surface albedo of the arctic from space optical imagers: retrieval and validation. 164 p.
102. Lindqvist, Hannakaisa, 2013. Atmospheric ice and dust : from morphological modeling to light scattering. 154 p.
103. Tuomi, Laura, 2014. (valmisteilla / in preparation)
104. Portin, Harri, 2014. Observations of aerosols, clouds and their interactions at Puijo measurement station. 133 p.
105. Carbone, Samara, 2014. Chemical characterization and source apportionment of submicron aerosol particles with aerosol mass spectrometers. 135 p.
- 106 Ollinaho, Pirkka, 2014. Parametric uncertainty in numerical weather prediction models. 123 p.