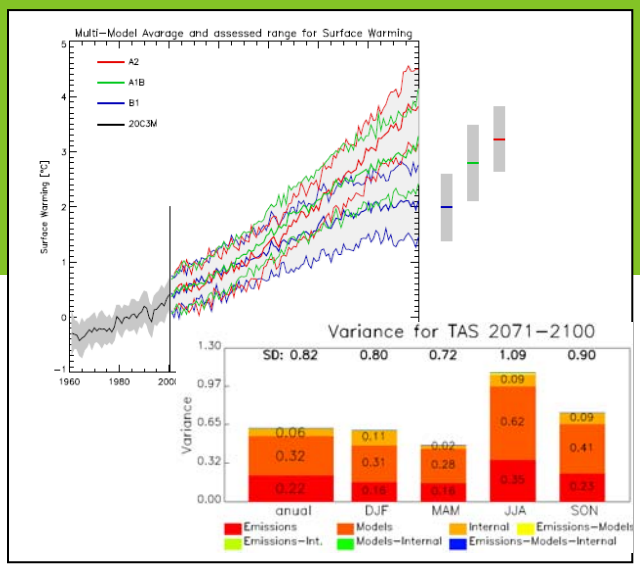
Wegener Center for Climate and Global Change
University of Graz



Scientifc Report No. 30-2009

Uncertainties in the Driving Data of Regional Climate Models in the Alpine Region

Andreas Franz Prein

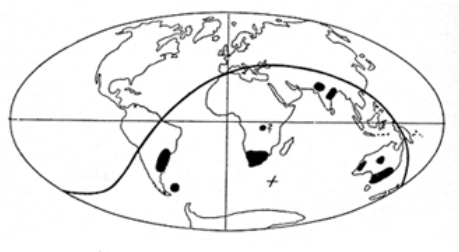
November 2009

kindly supported also by

Wegener Center
www.wegcenter.at

The **Wegener Center for Climate and Global Change** combines as an interdisciplinary, internationally oriented research center the competences of the University of Graz in the research area „Climate, Environmental and Global Change“. It brings together, in a dedicated building close to the University central campus, research teams and scientists from fields such as geo- and climate physics, meteorology, economics, geography, and regional sciences. At the same time close links exist and are further developed with many cooperation partners, both nationally and internationally. The research interests extend from monitoring, analysis, modeling and prediction of climate and environmental change via climate impact research to the analysis of the human dimensions of these changes, i.e., the role of humans in causing and being effected by climate and environmental change as well as in adaptation and mitigation. The director of the center, hosting about 35 researchers, is the geophysicist Gottfried Kirchengast, the lead partner and deputy director is the economist Karl Steininger. (more information at www.wegcenter.at)

The present report is the result of a Master thesis work completed in October 2009. Andreas Prein received partial financial support for this work from The Austrian Science Fund (FWF).



Alfred Wegener (1880-1930), after whom the Wegener Center is named, was founding holder of the University of Graz Geophysics Chair (1924-1930) and was in his work in the fields of geophysics, meteorology, and climatology a brilliant, interdisciplinary thinking and acting scientist and scholar, far ahead of his time with this style. The way of his ground-breaking research on continental drift is a shining role model — his sketch on the relationship of the continents based on traces of an ice age about 300 million years ago (left) as basis for the Wegener Center Logo is thus a continuous encouragement to explore equally innovative scientific ways: *ways emerge in that we go them* (Motto of the Wegener Center).

Wegener Center Verlag • Graz, Austria

ISBN 978-3-9502615-7-8

November 2009

Contact: Andreas Prein, M.Sc.
andreas.prein@uni-graz.at

Wegener Center for Climate and Global Change
University of Graz
Leechgasse 25
A-8010 Graz, Austria
www.wegcenter.at

Uncertainties in the Driving Data for Regional Climate Models in the Alpine region

Master Thesis

to obtain the degree of a Master of Natural Sciences
at the Department of Natural Sciences
University of Graz

Author
Bakk.rer.nat. Andreas Franz Prein
Wegener Center for Climate and Global Change
Matr.Nr.: 0311366

First Reviewer: Mag. Dr.rer.nat. Andreas Gobiet
Wegener Center for Climate and Global Change
Institute for Geophysics Astrophysics and Meteorology
University of Graz

Graz, September 22, 2009

Acknowledgment

I am grateful to my supervisor Mag. Dr.rer.nat Andreas Gobiet who gave me the opportunity to be a member of the Regional and Local Climate Modeling and Analysis Research Group at the Wegener Center in Graz. I want to thank him for his guidance, his invaluable support, his revisions and his useful suggestions during my research work.

Furthermore, I thank my colleagues at the Wegener Center for the warm reception in their group, their mental and physical support and for the many constructive discussions. My thankfulness also goes to the PCMDI (Program for Climate Model Diagnosis and Intercomparison), ECMWF (European Center for Medium-Range Weather Forecasts) and to the KNMI (Royal Netherlands Meteorological Institute) for the access to their databases.

Immeasurable gratitude must be expressed to my mother Mrs. Alfreda Prein and my brothers Mr. Michael and Mr. Wolfgang Prein, who supported me in every imaginable way during my whole studies. This thesis would have been impossible without them.

Finally I'd like to thank the Austrian Student Grand Authority for their financial aid during my studies and the University of Graz as the Wegener Center for the scholarship during the work on my diploma thesis.

Contents

Contents

Acronyms

Abstract **1**

Zusammenfassung **2**

1 Introduction **3**
1.1 Development and Methods of Climate Projection 3
1.1.1 Uncertainties in Climate Research 5
1.1.2 Types of Models 9
1.1.3 Regional Climate Models (RCMs) 11
1.2 GCM Driving Data for RCMs 13
1.2.1 The WCRP CMIP3 Multi Model Dataset Archive at PCMDI 14
1.2.2 GCM Driving Data Selection for RCM Simulations . . 14
1.2.3 Do Well Performing Models Provide Proper Information About the Future Climate? 16
1.2.4 Uncertainties in Driving Data for RCMs 17
1.3 Previous Work 18
1.3.1 The Performance of GCMs 18
1.3.2 Uncertainty Estimation 18

2 Materials and Methods **19**
2.1 GCM Data from CMIP3 Project 19
2.1.1 Considered Climate Variables 22
2.2 Reference Data 27
2.2.1 European Observation Dataset (E-OBS) 27
2.2.2 European Center for Medium-Range Weather Forecasts (ECMWF) 40 Years Reanalysis Dataset (ERA40) 27
2.3 Research Domains 29
2.4 Regridding 29
2.5 Measuring of Model Performance 30
2.5.1 Robustness of the MPI 32
2.5.2 Taylor Diagram 32
2.6 Box-Whisker Plots 34
2.7 Significance of the Climate Change Signal 35

2.7.1 Testing for Gaussian Distribution . . . 35
2.7.2 Test for the Homogeneity of Variances . . . 36
2.7.3 t-test for Homogeneous Variances . . . 36
2.7.4 t-test for Heterogeneous Variances . . . 37
2.8 Calculation of the Climate Sensitivity . . . 37
2.9 Spearman Rank Correlation . . . 38
2.10 Uncertainty Estimation . . . 38
2.10.1 Mean Variance of Uncertainty Sources . . . 39
2.10.2 Analysis of Variance (ANOVA) . . . 39
2.10.3 Reconstructing Missing Data . . . 44

3 Findings **47**
3.1 Model Evaluation for the Climate Period 1961 to 2000 . . . 47
3.1.1 Annual Cycle . . . 47
3.1.2 Spatial mean, max., min. Bias and RMS . . . 48
3.1.3 GCM Evaluation with Taylor Plots . . . 52
3.1.4 The Model Performance Index . . . 54
3.2 Analysis of the Climate Change Signal in the 21st Century . 62
3.2.1 Temperature at Surface (*tas*) . . . 62
3.2.2 Precipitation Flux (*pr*) . . . 66
3.2.3 Pressure at Sea Level (*psl*) . . . 69
3.2.4 Air Temperature (*ta*) . . . 69
3.2.5 Specific Humidity (*hus*) . . . 74
3.2.6 Geopotential Height (*zg*) . . . 74
3.2.7 North- and Estward Wind (*va*, *ua*) . . . 78
3.2.8 Vertical air Movement (*wap*) . . . 81
3.2.9 Overview of the GCM Climate Sensitivity . . . 84
3.3 Analysis of Correlations Between GCM Performance and their Climate Sensitivity . . . 85
3.4 Analysis of Uncertainties . . . 89
3.4.1 Uncertainties in Surface Temperature (*tas*) . . . 91
3.4.2 Uncertainties in Precipitation Flux (*pr*) . . . 91
3.4.3 Pressure at Sea Level (*psl*) . . . 95
3.4.4 Air Temperature (*ta*) . . . 97
3.4.5 Specific Humidity (*hus*) . . . 99
3.4.6 Geopotential Height (*zg*) . . . 101
3.4.7 Eastward Wind (*ua*) . . . 103
3.4.8 Northward Wind (*va*) . . . 105
3.4.9 Lagrangian Tendency of Air Pressure (*wap*) . . . 107

4 Discussion **109**

5 Conclusion **118**

Bibliography **119**

List of Figures **126**

List of Tables **132**

Acronyms

α	significance level
$\bar{o}_{\nu n}$	observed climatology for variable ν and grid point n
$\bar{s}_{\nu mn}$	simulated climatology for variable ν, model m and grid point n
ϵ_{ijk}	random variables
$\mu_{\cdot j\cdot}$	averaged expected values of the climate change signal for the emission scenarios
$\mu_{\cdot jk}$	averaged expected values of the climate change signal for the emission scenarios and internal runs
$\mu_{\cdot\cdot k}$	averaged expected values of the climate change signal for the internal runs
$\mu_{i\cdot k}$	averaged expected values of the climate change signal for the GCMs and internal runs
$\mu_{ij\cdot}$	averaged expected values of the climate change signal for the GCMs and emission scenarios
μ_{ijk}	averaged expected values of the climate change signal for the GCMs, emission scenarios and internal runs
$\mu_{i\cdot\cdot}$	averaged expected values of the climate change signal for the GCMs
μ_{ijk}	expected values of climate change signals
φ	latitude
ρ	density
ρ_{dA}	density of dry air
ρ_m	density of moist air
ρ_V	density of water vapor
$\sigma^2_{\nu n}$	interannual variance for variable ν and grid point n
σ_{ijk}	variance of climate change signals
σ_o	spatial variance of the observed variable
σ_s	spatial variance of the simulated variable
$\vec{v_H}$	horizontal wind vector
$\vec{v}$	wind vector
$\vec{x}$	random sample
B	number of samples for Bootstraping
df	degrees of freedom

E	choice of emission scenario with $j = 1, 2 \ldots, q$ scenarios
e	water vapor pressure
E_D	root mean square difference
E'	centered pattern RMS difference
$g(\varphi, z)$	gravity
g_N	standard gravity
hus	specific humidity
I	choice of internal run with $k = 1, 2, \ldots, r$ ensemble members
$I^2_{\nu m}$	performance index for variable ν and model m
I^2_m	model performance index (see also MPI)
M	choice of model with $i = 1, 2 \ldots, p$ models
M	molar mass
m	number of moles
N	choice of year with $l = 1, 2, \ldots, n$ years
n	number density of molecules
p	pressure
pr	precipitation flux
psl	air pressure at sea level
Q_{EI}	variance in the climate change signal caused by the combination of factors I and E
Q_{ME}	variance in the climate change signal caused by the combination of factors M and E
Q_{MEI}	variance in the climate change signal caused by the combination of factors M, E and I
Q_{MI}	variance in the climate change signal caused by the combination of factors M and I
Q_E	variance in the climate change signal caused by emission scenario uncertainty
Q_I	variance in the climate change signal caused by internal uncertainties
Q_M	variance in the climate change signal caused by GCM definition
Q_R	not explainable part of the variance
Q_T	total variance of the climate change signals
R	correlation coefficient
R_E	earth radius
R_G	gas constant
R_L	gas constant for dry air
r_s	rank correlation coefficient
R_V	gas constant for water vapor
s^2_E	standard deviation of the emission scenarios
s^2_I	standard deviation of the internal uncertainty
s^2_M	standard deviation of the model formulation

ta	air temperature
tas	near surface (2 m) air temperature
ua	eastward wind
V	volume
va	northward wind
w	vertical wind velocity
w_n	weights for area and mass averaging
wap	lagrangian tendency of air pressure
X_{ijk}	matrix, which contains the climatological mean values of the climate change signals
X_{ijkn}	matrix, which contains the annual values of the climate change signals
y_{ijkl}	total expected value μ_{ijk} plus the random variables ϵ_{ijk}
z	vertical height
zg	geopotential height
20C3M	climate of the 20th century
AGCM	Atmospheric Global Circulation Model
ANOVA	analysis of variance
AR4	fourth assessment report
CMIP3	phase 3 of the Coupled Model Intercomparison Project
DJF	December-January-February
EBM	energy balance model
E-OBS	European Observation
ECMWF	European Center for Medium-Range Weather Forecasts
ERA	European Reanalysis
GCM	General Circulation Model
IPCC	Intergovernmental Panel on Climate Change
JJA	June-July-August
KS-test	Kolmogorov-Smirnov test
MAM	March-April-May
MPI	model performance index
OGCM	Ocean General Circulation Model
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PRUDENCE	Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects
RCM	Regional Climate Model
REA	reliability ensemble averaging
RMS	root mean square
SON	September-October-November
WCRP	World Climate Research Program
WGCM	Working Group on Coupled Models

Abstract

The selection of high quality initial and boundary data (driving data) for Regional Climate Model (RCM) simulations is essential, because the model output quality is strongly dependent on the quality of the input data. For regional climate projections the driving data are usually supplied by General Circulation Models (GCMs).

In this study the data of 24 GCMs, used in the fourth assessment report (AR4) of the Intergovernmental Panel on Climate Change (IPCC), are analyzed to find out, which GCMs should provide the driving data for RCMs in the Alpine region. Further it is investigated how important the GCM formulation is compared to the emission scenario uncertainty and the internal uncertainty (different GCM initial conditions).

Therefore, nine parameters (three on the ground level and six in four different pressure levels) are evaluated against observed and reanalyis data in the climate period 1961 to 2000. The evaluation techniques include Taylor diagrams, annual cycle analyses, spatial bias maps and a model performance index (MPI). Concerning the forecast data the temporal development and the multi model spread of the parameters are plotted from 2000 to 2100 for the emission scenarios B1, A1B and A2. The climate change signals (periods 2021 to 2050 and 2071 to 2100 minus reference period 1971 to 2000) are analyzed for differences in the mean values and changes in the seasonal variances. Furthermore, correlations between the strength of the climate change signals and the model performance of GCMs are analyzed. Finally, two uncertainty estimation methods (one qualitatively and analysis of variance (ANOVA)) are used to calculate the composition of the variance in the climate change signals.

The results of this study should provide a guideline for the selection of driving data for RCM simulations in the Alpine region. They give an overview of the ability of state of the art GCMs to reproduce the contemporary climate in an orographically complex area and show the sources of uncertainty in the climate change signals of multiple climate variables.

Zusammenfassung

Die Auswahl von hochwertigen Anfangs- und Randbedingungen (Antriebsdaten) ist essenziell für Simulationen mit Regionalen Klimamodellen (RCMs), weil die Qualität der Modellergebnisse stark von jener der Eingangsdaten abhängt. Für regionale Klimaprognosen werden diese Daten in der Regel von allgemeine Zirkulationsmodellen (GCMs) bereitgestellt.

In dieser Studie werden die Daten jener 24 GCMs analysiert, die im vierten Sachstandsbericht (AR4) des Intergovernmental Panel on Climate Change (IPCC) verwendet wurden. Ziel ist es herauszufinden, welche GCMs verwendet werden sollen, um Antriebsdaten für RCMs im Alpenraum zu liefern (GCM Unsicherheit). Weiters wird evaluiert, wie wichtig diese Wahl im Verhältnis zur Emissionsszenario Unsicherheit und internen Modellunsicherheiten ist, welche von unterschiedlichen Randbedingungen herrührt.

Insgesamt werden neun Parameter (drei in Bodenhöhe und sechs in vier Druckniveaus) in der Klimaperiode 1961 bis 2000 mit Mess- und Reanalysedaten verglichen. Die Untersuchungsmethoden beinhalten Taylor Diagramme, Jahresganganalysen, Landkarten der räumlichen Fehler und die Berechnung eines Modellgüte Indexes (MPI). Die Prognosedaten werden in ihrem Verlauf von 1961 bis 2100 abgebildet und zusammen mit der einfachen Standardabweichung des GCM Ensembles für die Emissionsszenarien B1, A1B und A2 dargestellt. Die Klimasignale zweier Perioden (2021 bis 2051 und 2071 bis 2100 minus der Referenzperiode 1971 bis 2000) werden bezüglich ihrer Änderungen im Klimamittel bzw. saisoneller Variabilität untersucht. Des Weiteren wird nach Korrelationen zwischen der Stärke des Klimasignals und der Modellgüte von GCMs gesucht. Zuletzt werden noch zwei Unsicherheitsabschätzungsmethoden (eine Qualitative und Varianzanalyse (ANOVA)) angewannt, um die Zusammensetzungen der Varianzen der Klimasignale zu ermitteln.

Die Resultate dieser Studie sollen als Richtlinie für die Auswahl von Antriebsdaten für RCM Simulationen im Alpenraum dienen. Sie geben einen Überblick darüber, wie gut moderne GCMs den vorherrschenden Klimawandel in einem orographisch komplexen Gebiet nachbilden können und zeigen die Unsicherheitsquellen in den Klimasignalen verschiedenster Klimaparameter auf.

Chapter 1

Introduction

1.1 Development and Methods of Climate Projection

The word climate originates from the Greek expression "klinein" and means in its original content "bend". It is related to the curvature of the earth and the experience that the only possibility to observe a different region of the sky is a north to south movement on the ground. Nowadays the meaning of the word climate is very complex and so there are numerous ways to define it [Kraus, 2004]. One of the most used definitions is given by the Intergovernmental Panel on Climate Change (IPCC) [2007]:

> "Climate in a narrow sense is usually defined as the average weather, or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period of time ranging from months to thousands or millions of years. The classical period for averaging these variables is 30 years, as defined by the World Meteorological Organization. The relevant quantities are most often surface variables such as temperature, precipitation and wind. Climate in a wider sense is the state, including a statistical description, of the climate system."

Climate projection has like weather prediction a long history [Easterling and Stern, 1999]. Before the rise of numerical prediction methods people tried to estimate the weather and climate development by traditional methods of forecasting like divination, observed patterns (e.g., that a warm summer follows a cold winter in a specific region) or by precursors (animals or flowers that appear at special times). In the last century statistical forecasting methods have been developed, which predict parameters with either historical recordings of these parameters (univariate techniques) or with correlations to other quantities (multivariate techniques). The disadvantages of these methods are that they only work for short time predictions

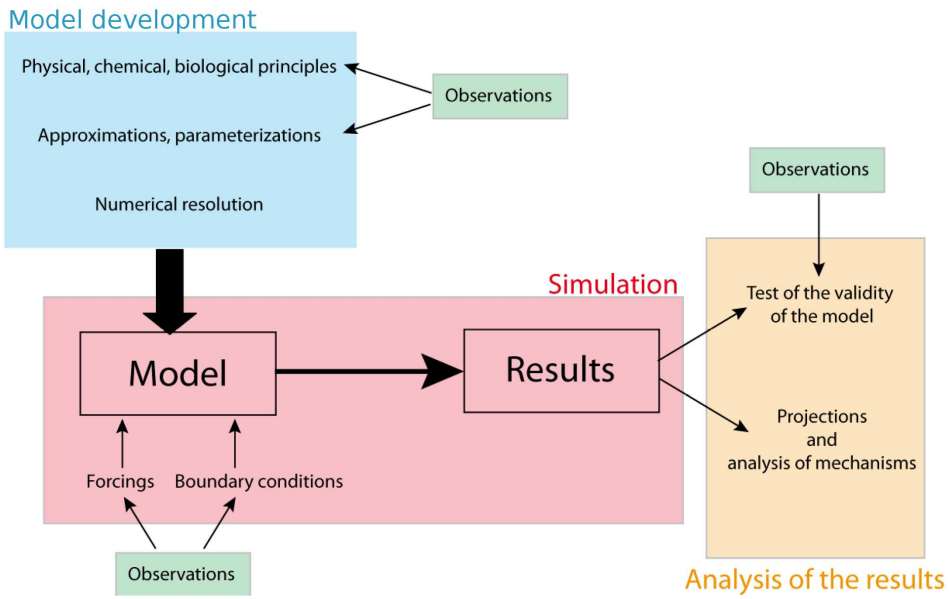


Figure 1.1: Overview of the general development of a climate Model [Goosse et al., 2008].

and cannot describe changes in the climate system, as they assume constant predictor - predictand relationships.

The first climate models were developed to explore the elementary dynamic of the climate system [Storch et al., 1999]. Their atmospheric model components rose from operational short time prognosis models known from weather forecasting. Since the awakening to the "climate problem" climate models got additional fields of importance. Nowadays they are the major source of information about anthropogenically caused climate change and its impact on nature and community. For Storch [1999] one major challenge for the research society is to determine which model information could be extracted and which not. In his opinion climate impact researchers refer much to high accuracy in time, space and precision to the simulated data than model construction can provide.

A modern climate model represents physical, chemical and biological processes, which are described and linked in a mathematically formalism [Goosse et al., 2008]. There are numerous types of models with different degrees of complexity starting from zero dimensional energy balance models (EBM) and ending with modern general circulation models (GCMs), which provide the information for the contemporary climate impact research. Figure 1.1 displays the schematic development of climate models. In every process step (model development, simulation and analysis of the results) observations play an elementary role.

The physical laws, which are the basis of each climate model are so complex that they have to be solved numerically. As a consequence, the output of GCMs is discrete in space and time. Some models have only global or zonal averages, whereas state of the art GCMs have a multidimensional grid with a spatial resolution below 200 km.

A summary of the history of GCM development is given in figure 1.2. More details can be found on the web page of the American Institute of Physics - Center of History of Physics at http://www.aip.org/history/sloan/gcm/.

1.1.1 Uncertainties in Climate Research

In the former section the basics of climate projection from the past to the present was described. This section deals with uncertainties, which are an important issue of climate research. A transparent and consistent treatment of uncertainties is one of the major goals of the IPCC [2007]. They divide uncertainties in three primary types:

- **Value uncertainties** occur when particular values are determined incompletely or if they are not appropriate to describe certain phenomenas. They are estimated using statistical techniques.
- **Structural uncertainties** arise from an incomplete understanding of a process. The estimation is limited to the judgment of the author and his/her confidence in the result.
- **Chaotic uncertainties** occur in systems which are not totally deterministic in nature.

Furthermore, there are two kinds of errors that arise from these three uncertainty types. The first are random errors, which decrease with increasing repetition of measurements (or repetition of model runs). The second are systematic errors, which are hard to detect and are not statistically quantifiable.

In climate research there is an uncertainty chain starting at the insecure future socio-economic development. That's why emission scenarios are needed, which describe the future anthropogenic greenhouse gas emissions. Additionally, the climate system is a chaotic system which can react very sensible to small changes. Therefor, modelers simulate ensembles of experiments, which only differ in their initial conditions. In this work this uncertainty source is called internal uncertainty. The next link in the chain raises because the climate system is very complex and GCMs are only a simplified representation of the reality. In each GCM different approaches are used to connect and solve the basic equations and to get a realistic reconstruction of the climate system, so that each GCM simulates a different future climate.

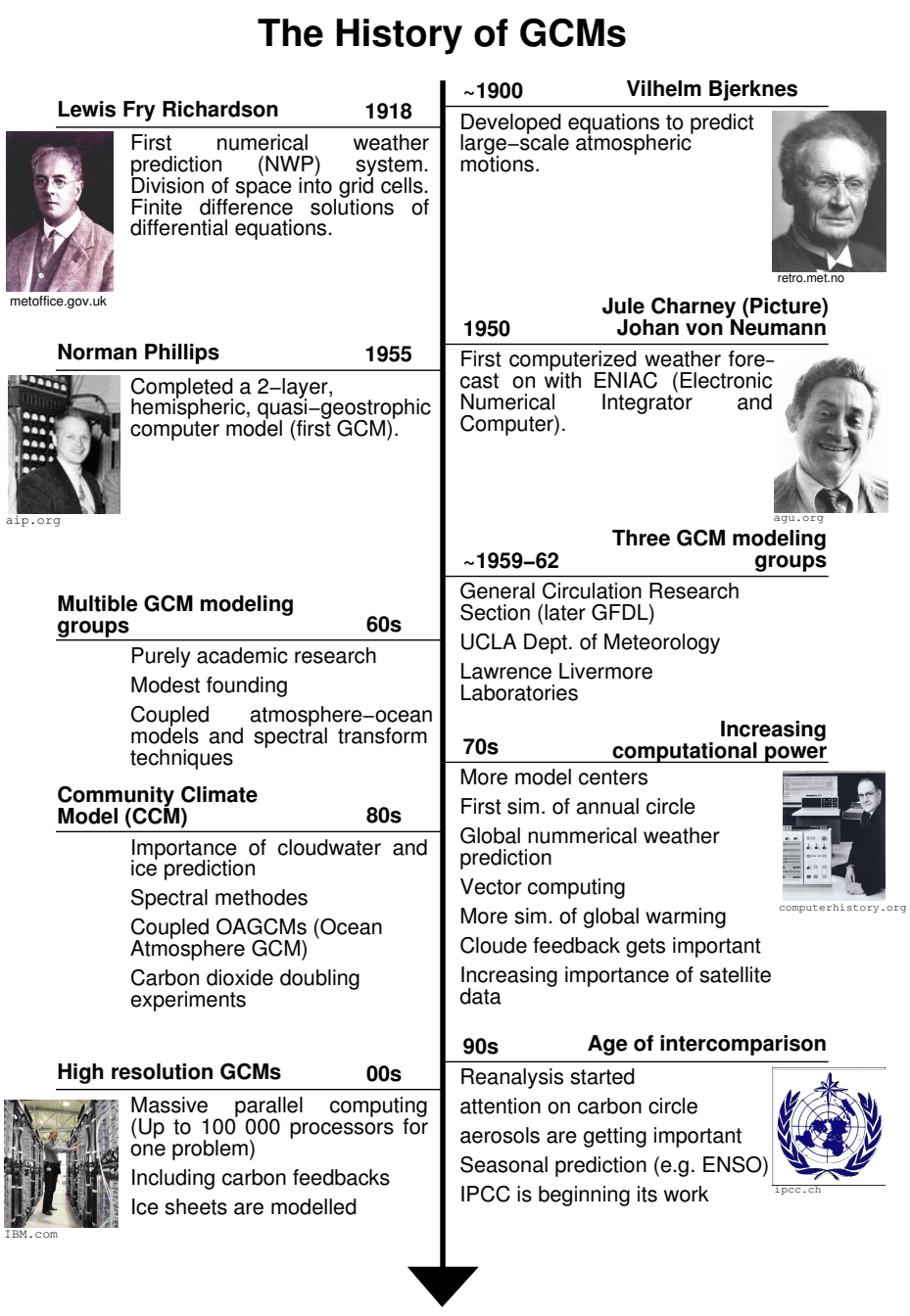


Figure 1.2: Short history of GCM development.

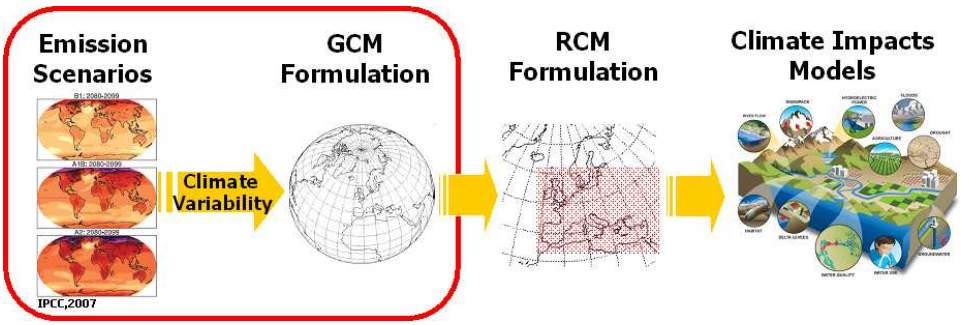


Figure 1.3: Uncertainty propagating from emission scenarios over internal, GCM and RCM uncertainties to climate impact model uncertainties. The red frame marks the part of the "uncertainty chain", which is the topic of this thesis.

This uncertainty source is called GCM formulation uncertainty. In this work the regional climate model (RCM) simulation uncertainty and the climate impact models uncertainty, which build the end of the uncertainty chain, are not a part of the analyses. The chain is displayed in figure 1.3 and the red box marks the parts which are discussed in this work.

Emission Scenario Uncertainty

It is highly significant that the recent climate change has an anthropogenic origin [IPCC, 2007]. Therefore it is understandable that climate projection is dependent on future human activities [Nakicenovic et al., 2007]. Since these activities are insecure the first link in the uncertainty chain, the emission scenario uncertainty, occurs (see figure 1.3). There are over 40 different emission scenarios specified in the IPCC Special Report on Emission Scenarios (SRES). They differ in future greenhouse gas emissions, future land use and other climate relevant driving forces. Overall the SRES scenarios can be grouped into four different scenario families A1, A2, B1 and B2 [Barker et al., 2007]. The focus of model intercomparison studies was mainly on the scenarios B1, A1B and A2.

- **B1:** In a globalized world the economic structure changes rapidly toward a services and technology orientated economy.
- **A1B:** Has the same population development as the B1 scenario, but with a technological change, which is balanced over all sources.
- **A2:** Shows a very heterogeneous world where the population is growing rapidly and the economic and technological development is slow.

There are no specifications concerning the likelihood of the scenarios in the SRES.

GCM Internal (Initial Condition) Uncertainty

The second important uncertainty source in figure 1.3 comes from the chaotic behavior of the climate system and is here called internal uncertainty. Many climate modeling groups performed ensembles of experiments by changing the initial conditions gradually to figure out how big the internal variability of the model is [PCMDI, 2009]. In a chaotic system a small change in the initial conditions can lead to huge changes in the result. In fact climate is a chaotic system and GCMs are showing chaotic behavior [Storch et al., 1999]. Nevertheless, by regarding climatological means, which are greater than 30 years and simulation times under 1000 years the differences between ensembles of GCM runs are found to be relatively small [Collins et al., 2001], [Rowell, 2006], [Déqué et al., 2007].

GCM Formulation Uncertainty

The the third uncertainty source in the chain from figure 1.3 is the uncertainty due to GCM formulation. Climate Models are of high complexity and need high performance computers for their simulations. At the moment there are more than a dozen research groups all over the world working on the formulation of GCMs [Reichler and Kim, 2008a]. Nevertheless, GCMs cannot perfectly reproduce our climate system due to incomplete theoretical understanding and assumptions which have to be made because of scaling problems (parameterizations).

Many approaches have been developed to evaluate climate models globally or over different regions of the world (e.g. [Gleckler et al., 2007], [Ulden and Oldenborgh, 2006], [Maxino et al., 2008], [Ki Min and Hensen, 2007]). One approach to define model performance in one single index was published by Reichler and Kim [2008]. Its theoretical background is described in section 2.5. They found, that the quality of GCMs increased steadily during the last years (see figure 1.4). However, evaluations of state of the art GCMs are showing a wide spread of model performance.

In section 1.2.3 it will be discussed if model performance metrics, based on the ability of GCMs to reproduce the current and past climate, can provide information about the accuracy of GCM results for the future climate change.

RCM Formulation Uncertainty

The third link in the uncertainty chain from figure 1.3 is the RCM formulation uncertainty. Similar as for GCMs there are numerous RCMs available which differ in their mathematical formulations and physical parameterizations. Furthermore, the simulation output of RCMs are also depend on properties like the domain setting or the resolution of the grid.

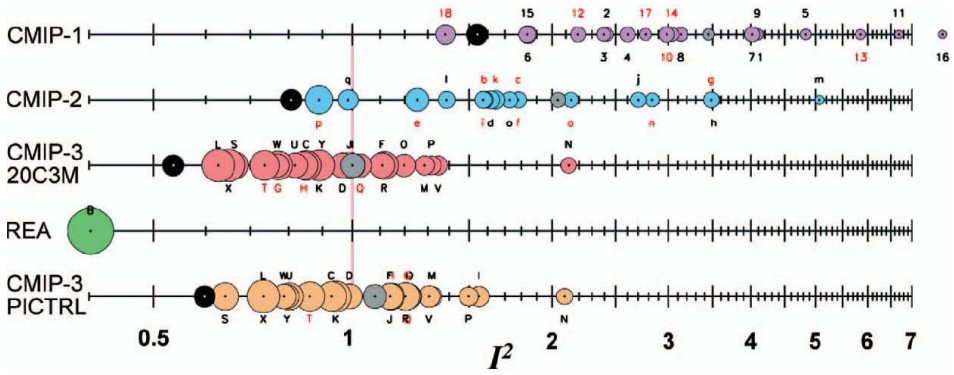


Figure 1.4: Model Performance Index I^2 for different model generations (rows) and different models (circles). The better the model the smaller the index. The diameter of the circles indicate the 95% confidence interval. Numbers and letters mark different models. Models marked with a red circle use flux correction. Grey circles show the average of all model performance indices (MPIs) in one model generation whereas black circles mark the MPI of the multi-model-mean of one model generation. The green circle REA belongs to the NCEP/NCAR reanalysis. In the last row the preindustrial control experiment (PICTRL) of the CMIP-3 project is displayed [Reichler and Kim, 2008a].

Climate Impact Model Uncertainty

Finally the high resolved data from RCMs are used in climate impact models to calculate the effects of climate change (e.g., changes in the water supply, heating and cooling of houses, agriculture, tourism and so on). The RCM data are initial and boundary conditions for impact models like hydrological, socio-economical, agricultural or environmental models.

1.1.2 Types of Models

In the former sections the concept of climate models as the different sources of uncertainty, which occur in climate projection, were discussed. Now the classification, structure and usage of different types of climate models will be described. One of the most common ways to classify climate models is to separate them into classes of complexity [Storch et al., 1999], [Goosse et al., 2008], [Trenberth, 1992]. On the one hand there are EBMs. In the simplest way variables are averaged over the whole globe and many processes are not included or parameterized. EBMs are mainly used to understand elementary processes in the climate system related to the global energy budget and the importance of interaction processes (e.g., temperature - albedo feedback). On the other hand are GCMs, which couple atmospheric AGCMs, ocean OGCMs and soil, vegetation, ice and chemistry models. These very complex models can provide three dimensional high resolved data in space and time. GCMs become more and more realistic with

increasing computational power and knowledge about components of the climate system. Figure 1.5 shows the development of GCMs with respect to the included components and processes since the mid 1970s.

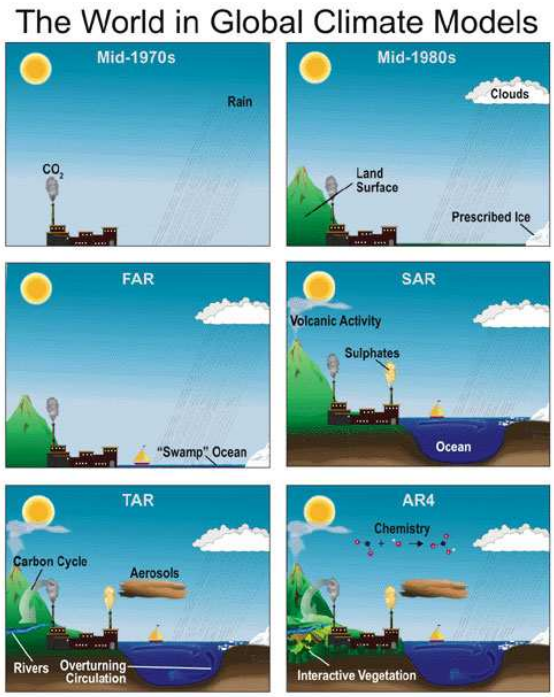


Figure 1.5: The development of components in GCMs since the mid 70th until the fourth IPCC Assessment reports [IPCC, 2007]. (FAR (1990): First Assessment Report -, SAR (1995): Second Assessment Report -, TAR (2001): Third Assessment Report -, AR4 (2007): Fourth Assessment report of the IPCC).

In the end complexity is only one side of the coin [Trenberth, 1992]. Even the most complex GCM is limited to its spatial and temporal resolution because of insufficient computational resources. Hence it's necessary to average over space and time and this causes a scale problem. Not only the orography of the surface must be approximated, but also numerous phenomena like clouds, which have smaller scales than grid boxes, have to be parameterized, whereas these parameterizations should be kept as physical as possible. For example, a parameterized cloud can be linked to variables that are resolved in the model, such as temperature and humidity. The degree of cloudiness within the area of a grid box is calculated from the correlation of temperature, humidity and cloudiness using empirical statistical methodologies. In climate model development the amount of model components are increasing while the size of model grids cells are shrinking. This leads to more orographical details, better regional simulations and the inclusion of more "small-scale" processes (see figure 1.6).

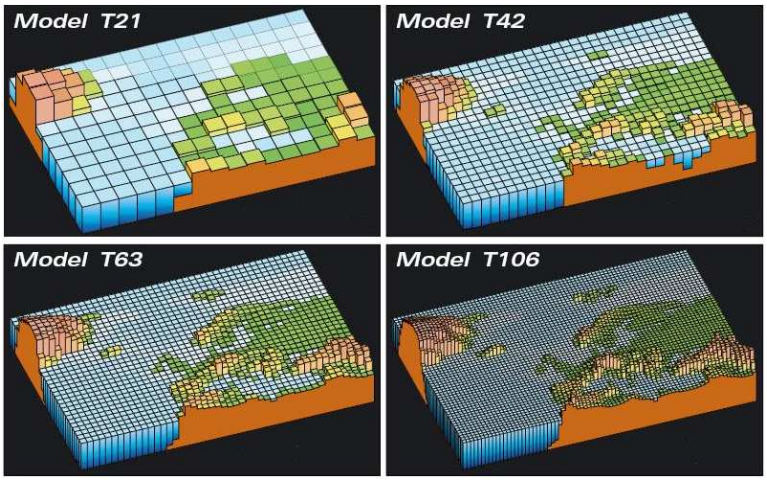


Figure 1.6: Europe for different model resolutions: T21 (ca. 500 km), T42 (ca. 250 km), T63 (ca. 180 km), T106 (ca. 100 km) [IPCC, 2007].

However, the increasing computational power and the increasing understanding of the climate system did not lead to a decrease of uncertainties. In 1979 the U.S. National Academy of Sciences wrote in the so called "Charney Report" that a doubling of CO_2 will lead to a temperature increase of 1.5 °C to 4.5 °C with a best estimate of 3 °C [Charney et al., 1979]. The current best estimation of the doubling CO_2 equilibrium temperature is given by the IPCC 2007 Fourth Assessment Report and is 2 °C to 4.5 °C with a best estimate of 3 °C.

Especially in regional and local climate projection as in the climate impact research the availability of highly resolved data is essential. Strong efforts are made to build GCMs with a horizontal resolution of up to 20 km [Noda et al., 2008], but within the next few model generations GCMs will not be able to provide these kind of data on a global grid. Therefore another solution was found in the form of RCMs.

1.1.3 Regional Climate Models (RCMs)

As mentioned in section 1.1.2 RCMs were developed in the need for high resolution climate information to understand the impacts of global climate change on small scales. Dickenson et al. [1989] and Giorgi and Bates [1989] were the creators of the first RCMs, which were mainly used as dynamic downscaling tools. In the early nineties Giorgi and Mearns [1991] wrote a review paper about possible techniques for regional climate simulations where they identified three basic methodologies:

1. **Empirical Approaches** using information from past climate

2. **Semi-empirical Approaches** are downscaling GCM fields statistically
3. **Modeling Approaches** are based on physical models like nested RCMs or variable resolution GCMs.

The focus of this study is the third methodology: RCMs nested in a GCM. The basic idea of a RCM is easy to understand: A limited area model, with detailed representation of physical processes and highly resolved orographic information, land use and complex topography can produce more realistic data with a higher information contend than global models[Wang et al., 2004]. It has been demonstrated that RCMs are able to improve simulations on a regional scale especially in regions with complex orography or coastlines. For Europe RCMs are in general able to reproduce the spatial distribution of temperature and precipitation better than global models, but still have errors like general warm biases in winter and summer and cold biases in the transient seasons as analyses of the PRUDENCE data by Jacob et al. [2007] showed.

Since RCMs can simulate only a limited part of the climate system they are dependent on input data at their lateral and upper boundaries [Storch et al., 1999]. Usually this input or "driving data" are taken from GCMs, measured, or reanalysis data, which provide the initial and time dependent lateral boundary conditions. The regional model is therefore embedded in the data field with a procedure called "nesting". The idea behind this concept is that the boundary data provides the global circulation and large scale forcing, whereas the regional model simulates the regional circulation on the sub grid scale of the driving data. To get a smoother transition between the boundary conditions and the regional model, RCMs often use exponentially weights and large buffer zones [Davies, 1976]. There is no general agreement about the size of a RCM domain. Large domains lead to a different large scale circulation between the RCM and the driving data (especially in regions where the boundary forcing is weak, e.g., in the tropics). Jones et al. [1995] suggest a small domain because in their opinion the synoptic circulation of the RCM should not differ too much from the driving fields, whereas other regional climate modelers like McGregor [1997] are using large domains and suggest that local area models should not only produce regional climate information, but also modify the large scale circulation with their high spatial information. Following McGregors idea it would be desirable that not only the global model influences the regional model, but also the regional model gives feedback to the global model. This technique is called "two way nesting" and is subject of current RCM development [Lorenz and Jacob, 2005]. The goal is to improve the results of the GCM without increasing the horizontal resolution on the entire globe.

One further way to improve the consistence of the driving data and the RCM data is a method called "spectral nudging" [Waldron et al., 1996]. The

idea behind this method is that small scale effects are the result of large scale circulations and small scale geographic information like orography, type of surface or land use [Storch et al., 2000]. Spectral nudging leads to a forcing that not only affects the RCM at the lateral boundaries, but also in the interior. To achieve this, large-scale nudging terms are added in the spectral domain. These terms have a major impact on large scale effects and are negligible for small scale processes. Furthermore there is a vertical gradient of the influence of the nudging terms so that lower atmospheric levels can adjust better to the topography of the surface. The conclusion of Storch et al. [2000] was that spectral nudging can help to retain the large scale circulation from the driving GCM in the RCM. However, it does not affect the ability of the RCM to produce realistic regional scale internal conditions suitable to its inherent spatial information.

By using GCM output as driving data for RCM simulations, GCM evaluation is very important because the ability of RCMs to reconstruct the observed regional climate depends strongly on the quality of the large scale boundary conditions [Wang et al., 2004]. GCM evaluation can help to select a GCM with small biases over the investigated domain, can provide information about uncertainties and can gain confidence in the RCM results. This topic will be discussed in more detail in section 1.2.2.

1.2 GCM Driving Data for RCMs

As described in section 1.1.3, RCMs need initial and lateral boundary conditions for their simulations. For future climate projections this information is provided by GCMs. There are various research groups, which are working on the development of GCMs. In the beginning of climate modeling the access to climate models and their data was limited to a small group of researchers [Meehl et al., 2007]. In the mid 1990s a committee of the World Climate Research Program (WCRP) called Working Group on Coupled Models (WGCM) coordinated the collection of GCM data in the framework of the first model intercomparison project. Afterward the Coupled Model Intercomparison Project (CMIP) was initialized by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) to make a subset of these data available for the research community outside the modeling groups. The following intercomparison projects CMIP2 and CMIP2+ suffered (similar as CMIP) under either the lack of model component data (atmosphere, land, ocean and sea ice) or the lack of experiments[1]. At the project start of the IPCC AR4 in 2003 a committee of experts decided to build up a better coordinated intercomparison project with more startup time for the modeling groups and sufficient time for the analyses of the data

[1] Until the CMIP2+ project only data of the control and the 1% CO_2 increase experiment were collected.

by the climate research community. The ambitious goal was to develop a state of the art collection of the human knowledge about climate change and variability based on model data. The PCMDI met the challenge and called it the third phase of the Climate Model Inetrcomparison Project (CMIP3), "... the largest international global coupled climate model experiment and multi model analyses effort ever attempted." [Meehl et al., 2007].

1.2.1 The WCRP CMIP3 Multi Model Dataset Archive at PCMDI

In the course of the development of the IPCC AR4 the PCMDI collected GCM outputs of the past, present and future climate from the leading modeling centers within the framework of the CMIP3 project [PCMDI, 2009]. The major part of the data were collected in the years 2005 and 2006 and the dataset is known as the "WCRP CMIP3 multi model dataset". Overall 17 modeling groups from 12 countries participated with 24 models. The data access is open for non commercial purposes after the registration and the agreeing to the terms of use at the PCMDI homepage[2]. As of February 2009 there were 36 terabyte of data stored, which were ready for download by 2220 registered users. There are parameters from every component of the GCM (atmosphere, sea-ice, land and ocean) in four temporal resolutions (every three hours, daily, monthly and annually) and time independent parameters. Furthermore, as shown in figure 1.7, there are multiple experiments available starting at twentieth and twenty-first century simulations going to emission scenario simulations and ending at climate change commitment experiments. The dataset contains approximately 140 different variables. A detailed list and precise definitions of the variables are available at the PCMDI homepage.

1.2.2 GCM Driving Data Selection for RCM Simulations

Each GCM, which is listed in figure 1.7 is able to provide initial and boundary conditions for RCM simulations. But the choice of the GCM, which generates the driving data, should be done carefully. Christensen [2007] points out the importance of the GCM formulation with respect to the numerical output of the RCM simulation. His conclusion is based on the results of the PRUDENCE project (Prediction of Regional scenarios and Uncertainties for Defining European Climate change risks and Effects) where two atmosphere-only general circulation models, four regional climate models and two emission scenarios were used to generate highly resolved climate simulations for Europe with the focus on uncertainties of the surface temperature and precipitation prediction. A similar result was achieved by Rowell [2006] who analyzed the PRUDENCE data for the United Kingdom.

[2]Access under: www-pcmdi.llnl.gov, September 22, 2009

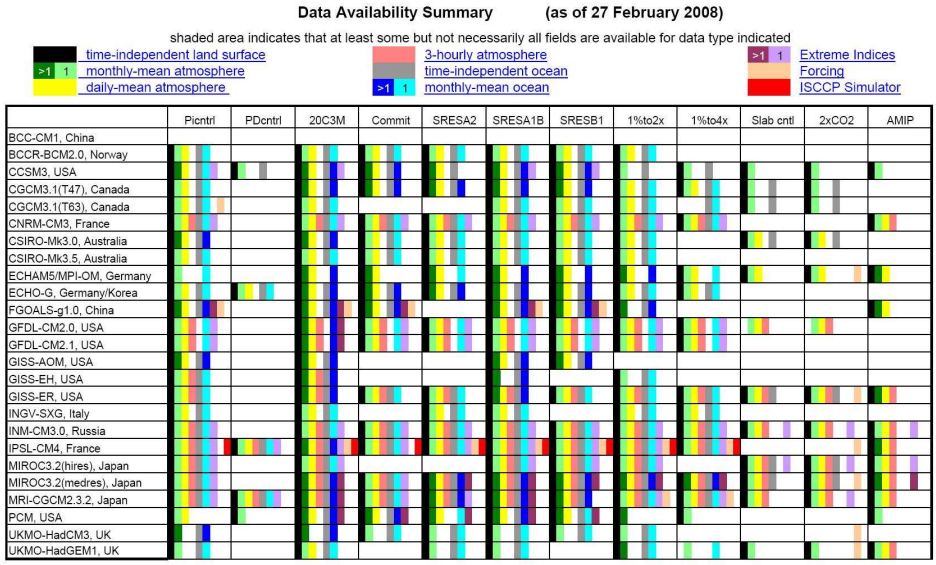


Figure 1.7: Summary of the data availability at the PCMDI. Colored boxes indicate that some variables (not necessarily all) are stored in the multi model dataset while shaded colors mean that there is an ensemble of runs for the corresponding model and experiment. The meaning of the colors is described in the legend [PCMDI, 2009].

His studies have shown that the dominant source of uncertainties in the climate change signal of the surface temperature and precipitation field is the structure and physic of the GCM. Déqué [2007] also made an uncertainty estimation based on the PRUDENCE project. His findings conform with the findings of Rowell: "The uncertainty introduced by the choice of the driving GCM is generally larger than the other three[3] sources." [Déqué et al., 2007]

All three studies in the paragraph above underline the importance of the GCM formulation for the output of RCM simulations. However, the achieved results suffer from some limitations. First of all only two GCMs namely HadAM3H from the Hadley Center and ECHAMP4/OPYC3 from the Max Planck Institute were used for the simulations. These two models are quite similar in their global climate change signal so that the real fraction of uncertainty due to GCM could be even larger [IPCC, 2001]. Furthermore only three experiments out of 19 where performed more than once, so that there are limitations concerning the validity of the variance caused by internal variability.

[3]These three sources are: the emission scenario, the RCM formulation and the internal variabilities represented by small changes in initial conditions.

1.2.3 Do Well Performing Models Provide Proper Information About the Future Climate?

In the former section the importance of the GCM driving data selection was illustrated, but the question how the selection process should look like is still open.

A frequently discussed topic in literature is, if future projections of GCMs, which show a good performance of the present day climate are more reliable than those of worse performing models. McAveny et al. [2001] stated that a good simulation of the current climate by GCMs is necessary, but not sufficient for considering their future climate projections as reliable.

Perkins and Pitman [2008] showed that there is a statistical significant difference in changes of maximum and minimum temperature, annual precipitation and rainy days between an ensemble of well performing GCMs and an ensemble of bad performing models over different regions in Australia. They suggested: "... that omitting weak models from regional scale estimates of future climate change help clarify the nature and scale of the projected impacts of global warming" [Perkins and Pitman, 2008].

Giorgi and Mearns [2002] introduced a "reliability ensemble averaging" (REA) method. On a sub-continental scale this method allows to calculate the averages, the uncertainty ranges and the reliabilities of future climate projections. There are two criterias that are taken into account by REA: first the model performance (ability of the model to perform present-day climate) and second the model convergence (the convergence of the future model projections of the ensemble). By rejecting GCMs with a bad performance or an outlying future projection Giorgi and Mearns [2002] succeeded in reducing the uncertainty range in the climate change signal and in the achieved reliability of simulated changes. A further result of this research was that: "In the set of experiments we analyzed, both for temperature and precipitation, there was no model that ubiquitously performed better or worse than all the others..." [Giorgi and Mearns, 2002].

However, the selection of good performing models and the rejection of bad ones is a controversial topic in climate research. For example Räisänen [2007] criticize the REA method because: "$\cdots$ the idea of rejecting or downweighting the climate-change projection from a specific model just because the climate change in this model differs a lot from that of the others involves a risk of circular reasoning." [Räisänen, 2007].

Often found in literature, and also shown in this study, is that the multi model mean of a GCM ensemble is able to outperform even the best model of the ensemble. Weigel et al. [2008] showed that this behavior is only achieved, if the models in the ensemble are not able to capture the full spread of uncertainty (that means, that they don't have the right ensemble spread at the right mean value). If this is true, even the addition of bad performing models to the ensemble can improve the multi model skill.

An other study, which contradicts the method of ensemble selection due to model performance in present climate was done by Pierc et al. [2009]. They showed that there are no significant differences in the results of selecting models by skill or selecting them randomly. They propose that enough realizations must be chosen for the ensemble building. In their case 14 GCMs were sufficient to obtain similar results as with the original ensemble. Concerning model skill they found that approximately five models are enough so that the model skill tends to converge. This means: "... that stable hindcasts (and forecasts) can be obtained by including a manageable small group of models." [Pierc et al., 2009].

1.2.4 Uncertainties in Driving Data for RCMs

In chapter 1.2.2 the importance of uncertainty analyses of GCM data concerning their use in regional climate model simulations is indicated. Overall there are three kinds of uncertainties, which can appear in the RCM driving data supplied by GCMs (see table 1.1) [Rowell, 2006].

Table 1.1: Summary of the sources of uncertainty [Rowell, 2006].

Abbreviation	Source of uncertainty	Data for uncertainty estimation
Internal	Large scale internal climate variations	multiple GCM ensembles with up to 9 members
Emissions	Projected anthropogenic emissions	GCMs forced by the SRES B1, A1B and A2 scenarios
GCM	GCM formulation	up to 24 GCMs from the CMIP3 dataset

In section 1.2.2 it is discussed that the boundary data for RCMs have a major impact on the result of the regional climate simulation. The basic questions are, which GCMs are suitable to provide driving data for RCMs in a particular region and how important this choice is compared with uncertainties from emission scenarios or internal variations due to initial conditions in the GCM runs. To answer these questions three methods are used:

1. The GCMs are evaluated with reference data so that the models can be ranked by performance.

2. The climate change signal of each model is compared with the multi model-mean to give an overview of the relative climate sensitivity of the models.

3. The uncertainties due to the effects listed in table 1.1 are estimated.

1.3 Previous Work

This section should give a rough overview of the previous work, which is the theoretical basis of this thesis.

1.3.1 The Performance of GCMs

In literature there are many approaches to define simple metrics, which are representative for the model performance (e.g. [Reichler and Kim, 2008a], [Gleckler et al., 2007], [Taylor, 2000], [WGNE, 2009], [Maxino et al., 2008]). A reliable model performance index (MPI) for climate models must contain a broad set of variables because climate models are used for a broad band of applications [Gleckler et al., 2007]. Furthermore it should consider variable phenomena in time and space on diurnal, inter seasonal, annual and longer periods. So far there is no single metric found to satisfy all these requirements.

Here the evaluation of GCMs is done by a model performance index, which was introduced by Reichler and Kim [2008]. It combines the spatial fit of multiple parameters for a particular climatology in one single index. The mathematical formulation and further details about this Index can be found in chapter 2.5. Additionally Taylor Plots [Taylor, 2000] are used to show the performance of the GCM in respect to the root mean square (RMS) error, the correlation coefficient and the standard deviation.

1.3.2 Uncertainty Estimation

Uncertainties are an important factor in climate projection and so they are a current issue in scientific research. Ensemble projects are a major tool to explore the uncertainty range and to estimate the insecurities of the result. For instance, one of the objectives of the the PRUDENCE project (see chapter 1.2.2 or http://prudence.dmi.dk/) was to estimate the uncertainties of regional climate projections over Europe.

One method to estimate the uncertainties was introduced by Rowell [2006] who analyzed the insecurities of the PRUDENCE data for the United Kingdom with an estimative and simple method.

A more comprehensive approach to estimate the uncertainty sources of RCM simulations was done by Déqué et al. [2007]. They used the PRUDENCE project ensemble to run an analysis of variance (ANOVA). In this study both approaches were used to estimate the uncertainty sources. The results of the methods will be compared to get a better estimation of the real sources of variance.

Chapter 2

Materials and Methods

2.1 GCM Data from CMIP3 Project

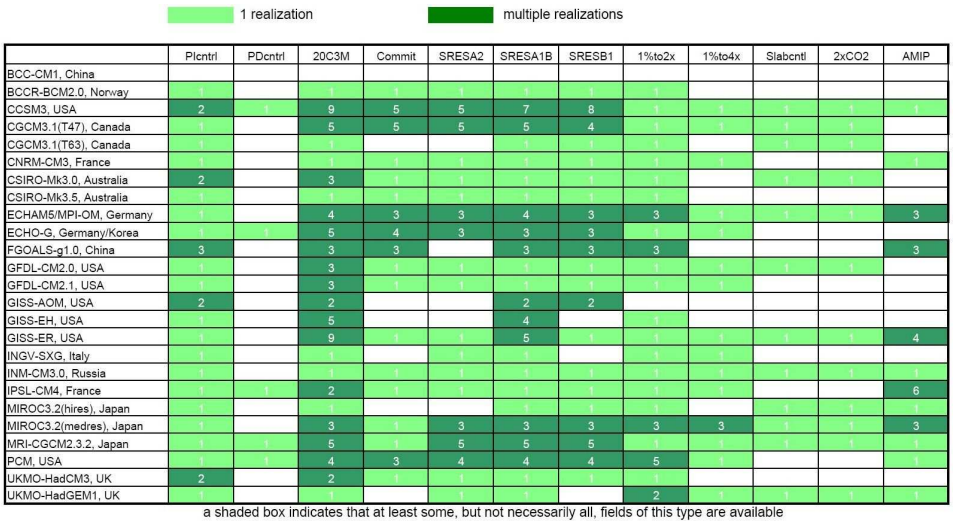
In this section details of the GCMs and certain parameters, which are used in this study, are described. A brief overview of the CMIP3 Multi Model Dataset was given in section 1.2.1. The tables 2.1 and 2.2 show the list of these GCMs, which are available in the PCMDI dataset.

Table 2.1: Model overview of the CMIP3 GCMs part I [PCMDI, 2009].

NR.	ORIGINATING GROUPS	COUNT.	MODEL	RES. (ATM.)	DETAILS	INTERNET ADRESS (September 22, 2009)
1	Beijing Climate Center	CHN	BCC-CM1	T63L16	1.875° lon x 1.875° lat	http://ncc.cma.gov.cn
2	Bjerknes Centre for Climate Research	NOR	BCCR-BCM2.0	T63L31 eq. T42	Triangular truncation T63 equivalent to T42 quadratic grid (2.8°)	www.bcm.uib.no
3	National Center for Atmospheric Research	USA	CCSM3	T85	Lateral resolution from 85-wavenumber of the dynamics. At the equator, the triangular spectral truncation spatial resolution is approximately 1.4°	www.ccsm.ucar.edu
4	Canadian Centre for Climate Modelling & Analysis	CAN	CGCM3.1(T47)	T47L31	Surface grid whose spatial resolution is roughly 3.75° lat/lon and L31	www.cccma.bc.ec.gc.ca
5	Canadian Centre for Climate Modelling & Analysis	CAN	CGCM3.1(T63)	T63L31	Surface grid whose spatial resolution is roughly 2.8° lat/lon L31	www.cccma.bc.ec.gc.ca
6	Meteo-France / Centre National de Recherches Météorologiques	FRA	CNRM-CM3	T63 qu. T42	Triangular truncation T63 with "linear" reduced Gaussian grid equivalent to T42 quadratic grid (2.8°)	www.cnrm.meteo.fr
7	CSIRO Atmospheric Research	AUS	CSIRO-Mk3.0	T63, L18		www.cmar.csiro.au/
8	CSIRO Atmospheric Research	AUS	CSIRO-Mk3.5	T63, L19		www.cmar.csiro.au/
9	Max Planck Institute for Meteorology	GER	ECHAM5/ MPI-OM	T63L31		www.mpimet.mpg.de
10	Met. Ins. University of Bonn, Met. Res. Ins. KMA	GER/ KOR	ECHO-G	T30L19		www.meteo.uni-bonn.de
11	LASG / Institute of Atmospheric Physics	CHN	FGOALS -g1.0	T42L26	The T42 gaussian grid (2.8° lon x 2.8° lat) is used between 65.58°N and 65.58°S, while a weighted even-area grid is used elsewhere (Wang et al 2004, Sci. China. Ser. A, 47,4-21)	http://web.lasg.ac.cn

Table 2.2: Model overview of the CMIP3 GCMs part II [PCMDI, 2009].

NR.	ORIGINATING GROUPS	COUNT.	MODEL	RES. (ATM.)	DETAILS	INTERNET ADRESS (September 22, 2009)
12	US Dept. of Commerce / NOAA / Geophysical	USA	GFDL-CM2.0	T42L24	2.5° lon x 2.0° lat, 24 levels	http://data1.gfdl.noaa.gov
13	Fluid Dynamics Laboratory	USA	GFDL-CM2.1	T42L24	2.5° lon, 2.0° lat, 24 levels	http://data1.gfdl.noaa.gov
14	NASA / Goddard Institute for Space Studies	USA	GISS-AOM	T30L12	4° lon, 3° lat, 12 levels	www.giss.nasa.gov
15	NASA / Goddard Institute for Space Studies	USA	GISS-EH	>T30	4° lat, 5° long	www.giss.nasa.gov
16	NASA / Goddard Institute for Space Studies	USA	GISS-ER	>T30	4° lat, 5° long	www.giss.nasa.gov
17	Instituto Nazionale di Geofisica e Vulcanologia	ITA	INGV-SXG	T106L19	Horizontal res: T106 (about 1.125° lat/lon). vertical res: L19	www.cmcc.it
18	Institute for Numerical Mathematics	RUS	INM-CM3.0	>T30L21	5° lon, 4° lat, 21 levels in vertical	http://www.sscc.ru/
19	Institut Pierre Simon Laplace	FRA	IPSL-CM4	>T42L19	Resolution: 2.5° lon, 3.75° lat (i.e. 96x71 grid points) 19 levels	http://dods.ipsl.jussieu.fr
20	Center for Climate System Research (The University	JPN	MIROC3.2(hr)	T106L56		www.ccsr.u-tokyo.ac.jp
21	of Tokyo) et.al.	JPN	MIROC3.2(mr)	T42L20		www.ccsr.u-tokyo.ac.jp
22	Meteorological Research Institute	JPN	MRI-CGCM2.3.2	T42	Approx. 2.8°	www.mri-jma.go.jp
23	National Center for Atmospheric Research	USA	PCM	T42	Lateral resolution from 42-wavenumber triangular spectral truncation of the dynamics. Equator resolution approx. 2.8°	www.cgd.ucar.edu/pcm
24	Hadley Centre for Climate Prediction and Research /	GBR	UKMO-HadCM3	>T42	2.75° latitude by 3.75° longitude; grid point model using an Arakawa B grid and hybrid vertical coordinates	www.metoffice.gov.uk
25	Met Office	GBR	UKMO-HadGEM1	T80N96L38	1.25° latitude by 1.875° 38 layers extending to over 39 km	http://www.metoffice.gov.uk



	PIcntrl	PDcntrl	20C3M	Commit	SRESA2	SRESA1B	SRESB1	1%to2x	1%to4x	Slabcntl	2xCO2	AMIP
BCC-CM1, China												
BCCR-BCM2.0, Norway	1		1	1	1	1	1	1				
CCSM3, USA	2	1	9	5	5	7	8	1	1	1	1	1
CGCM3.1(T47), Canada	1		5	5	5	5	4	1	1	1	1	
CGCM3.1(T63), Canada	1		1			1	1	1		1	1	
CNRM-CM3, France	1		1	1	1	1	1	1	1			1
CSIRO-Mk3.0, Australia	2		3	1	1	1	1	1		1	1	
CSIRO-Mk3.5, Australia	1		1	1	1	1	1	1				
ECHAM5/MPI-OM, Germany	1		4	3	3	4	3	3	1	1	1	3
ECHO-G, Germany/Korea	1	1	5	4	3	3	3	1	1			
FGOALS-g1.0, China	3		3	3		3	3	3				3
GFDL-CM2.0, USA	1		3	1	1	1	1	1	1	1	1	
GFDL-CM2.1, USA	1		3	1	1	1	1	1	1			
GISS-AOM, USA	2		2			2	2					
GISS-EH, USA	1		5			4		1				
GISS-ER, USA	1		9	1	1	5	1	1	1	1	1	4
INGV-SXG, Italy	1		1		1	1		1	1			
INM-CM3.0, Russia	1		1	1	1	1	1	1	1	1	1	1
IPSL-CM4, France	1	1	2	1	1	1	1	1	1			6
MIROC3.2(hires), Japan	1		1			1	1	1		1	1	1
MIROC3.2(medres), Japan	1		3	1	3	3	3	3	3	1	1	3
MRI-CGCM2.3.2, Japan	1	1	5	1	5	5	5	1	1	1	1	1
PCM, USA	1	1	4	3	4	4	4	5	1			1
UKMO-HadCM3, UK	2		2	1	1	1	1	1				
UKMO-HadGEM1, UK	1		1		1	1		2	1	1	1	1

a shaded box indicates that at least some, but not necessarily all, fields of this type are available

Figure 2.1: Monthly mean data available at the PCMDI data portal [PCMDI, 2009].

Figure 2.1 shows the availability of monthly mean data. For the evaluations in chapter 4 the data from the climate of the 20th century (20C3M) and the special emission scenarios B1, A1B and A2 were used.

2.1.1 Considered Climate Variables

The purpose of this study is to find adequate driving data for regional climate models in the Alpine region. So the set of parameters, which are used for the calculation of the MPI and the estimation of uncertainties, are the general used parameters (two meter surface air temperature, precipitation and sea level pressure) and parameters which are used as boundary conditions for RCM simulations. Table 2.3 shows the complete list of parameters and reference data (E-OBS (European Observation) and ERA40).

In the following paragraphs the parameters from table 2.3 are described more precisely, i.e., how they are defined, on what they are depending on and how they are interacting.

Air Temperature

The climate system is in good approximation an open system, which receives energy from the sun and re-emits it to the universe, while the mass of the system stays constant [Kraus, 2004]. Thereby, the air temperature ta (like pressure p, volume V or number density of molecules n) is a state variable of the climate. Usually state variables are not independent and can be connected with state equations. For an ideal gas (approximately the

Table 2.3: Climate variables, pressure level (for *tas*, *pr* and *psl* at ground level, others at 850, 700, 500 and 300 hPa) and corresponding reference data, which are used for the analyses.

	Variable	Unit	Pressure Level	Reference data
tas	near surface (2 m) air temperature	K	ground	E-OBS
pr	precipitation flux	kg/m^2s	ground	E-OBS
psl	air pressure at sea level	Pa	ground	ERA40
ta	air temperature	K	850, 700, 500, 300 hPa	ERA40
hus	specific humidity	g/kg		ERA40
zg	geopotential height	m		ERA40
ua	eastward wind	m/s		ERA40
va	northward wind	m/s		ERA40
wap	lagrangian tendency of air pressure	Pa/s		ERA40

lower atmosphere without condensed water vapor) the state equation can be written as:

$$p \cdot V = n \cdot \mathrm{R}_G \cdot ta = \frac{m}{M} \cdot \mathrm{R}_G \cdot ta \tag{2.1}$$

The number density of molecules n is hereby equal to m/M, with m is the number of moles and M is the molar mass. The parameter R_G is the ideal gas constant and has the value R_G=8.314472 J/mol·K. If ta is rising in formula 2.1 also p will rise when V/n stays constant.

An other important characteristic of temperature is its height dependence [Storch et al., 1999]. The energy in the lowest part of the atmosphere (troposphere) is mainly supplied by the surface in form of infrared radiation, sensible and latent heat flux. If an air parcel is lifted in higher layers of the troposphere it starts to expand and gets colder. As far as the air is not saturated with water vapor, the temperature gradient is approximately -10 K/km (adiabatic temperature gradient). After the air is saturated, condensation occurs, which leads to an energy input in the air parcel. This process leads to a decrease of the temperature gradient (saturated adiabatic temperature gradient), which is approximately -5 K/km to -8 K/km.

A more detailed description of the above discussed process can be found in [Kraus, 2004], [Liljequist and Cehak, 2006] and [Etling, 2008].

Precipitation Flux

Precipitation is a process where water or ice particles are build in the atmosphere and fall down at the surface [Kraus, 2004]. Thereby water vapor

starts to condensate or sublimate in the atmosphere and builds tiny water drops or ice crystals, which have a radius of approximately 1 μm to 100 μm. The water drops are so light that they can float in air. If these tiny particles are growing[1] to a radius $\geq 100\,\mu m$ they get a notable velocity to reach the surface and are called precipitation particles.

Precipitation is occurring in several areas of the troposphere:

- at fronts
- in convergence zones
- in convective zones
- in areas of static instability
- if air is forced to rise because of orography.

Sea Level Pressure

Air pressure is a value, which is proportional to the amount of air over a specific area [Liljequist and Cehak, 2006]. As mentioned in section 2.1.1, air pressure is a state variable of the climate system and is depending on the air temperature ta and density ρ via the ideal gas equation 2.1. The decrease of air pressure p with height z is described by the barometric formula:

$$\frac{\partial p}{\partial z} = -\frac{p}{\mathrm{R_L} \cdot ta} \cdot g \tag{2.2}$$

By integrating formula 2.2 it is possible to calculate the air pressure p at any altitude. In the case of sea level pressure, all pressure values are transferred to the mean sea level.

Specific Humidity

Specific humidity is defined via the following formulation [Kraus, 2004]:

$$\begin{aligned} hus = \frac{\rho_V}{\rho_m} &= \frac{\rho_V}{\rho_{dA} + \rho_V} = \frac{\frac{e}{R_V \cdot ta}}{\frac{p-e}{R_L \cdot ta} + \frac{e}{R_V \cdot ta}} \\ &= \frac{R_L}{R_V} \cdot \frac{e}{p - \left(1 - \frac{R_L}{R_V}\right) \cdot e} = 0.622 \cdot \frac{e}{p - 0.378e} \\ e &= \rho_V \cdot R_V \cdot ta \end{aligned} \tag{2.3}$$

[1] There are two growing processes: first the collision and coalescence with other particles (Langmuir-process) and second the diffusion of water vapor from liquid particles to solid particles in a supercooled mixed cloud (Bergeron-Findeisen-process).

In formula 2.3 ρ_V is the density of the water vapor and ρ_m is the density of the moist air, which itself is composed of the density of the dry fraction of the air ρ_{dA} plus ρ_V. By splitting the ρ values with the ideal gas equation (formula 2.1) e, the water vapor pressure, and the gas constants R_L for dry air and R_V for water vapor is needed. The unit of specific humidity is kg water vapor per kg of mixed air.

Geopotential Height

The geopotential height zg is used as vertical coordinate. It is defined as the geopotential[2] normalized by the standard gravity $g_N = 9,80665 m/s^2$ [Kraus, 2004]. If φ is equal to the latitude ($g(\varphi, z)$ is the gravity, R is the earth radius and z is the geometrical height) then zg can be written as:

$$zg = \frac{1}{g_N} \int_0^z g(\varphi, z) \approx \frac{1}{g_N} g(\varphi, z)$$

a good approximation for $g(\varphi, z)$ is:

$$g(\varphi, z) = g_N \cdot (1 - 0.0026373 \cos(2\varphi) + 0.0000059 \cos^2(2\varphi)) \cdot 1/(1 + z/R)^2$$

$$zg = (1 - 0.0026373 \cos(2\varphi) + 0.0000059 \cos^2(2\varphi)) \cdot \left(\frac{z}{1 + \frac{z}{R}} \right) \qquad (2.4)$$

Horizontal Air Velocity

By heating lower layers of the atmosphere the air expands, gets lighter and starts to rise. Because of this expansion the corresponding air column is increasing in height, which leads to a flow in higher layers [Storch et al., 1999]. In other areas where air is either cooled on the ground or not heated that efficient, it starts to sink because of its increasing density. This is the origin of cyclones (up floating air) and anticyclones (down floating air). The backflow on the ground closes the circulation. Differences in temperature are followed by differences in air pressure. This effect causes winds. Therefore not the absolute values of temperature and pressure are essential for the movements in the atmosphere, but their spatial gradients. The driving force is the force between cyclones and anticyclones (the so called “pressure gradient force”).

The two observed domains (see figure 2.4) are located in the westerlies. These winds are arising from the pressure gradient force between the high pressure in the horse latitudes (approximately at 30° north) and the low pressure in the mid latitudes (approximately at 60° north). The deflection of the south to the north directed force (forced to the right → westerly) occurs because of the rotation of the earth (coriolis force).

[2]The geopotential is given by the work that is needed to lift a unit mass (air package) from sea level to a certain altitude (potential energy of the unit mass) [Geer, 1996].

Vertical Air Velocity

Since climate models are using pascal rather than meters as unit for the vertical coordinate the vertical air velocity is given in Pa/s and is commonly labeled as *wap* (lagrangian tendency of air pressure). To convert *wap* into the vertical wind velocity w [m/s] the formula $wap = -\rho \cdot g \cdot w$ can be used in good approximation [Roger, 2006]. This means, that a negative *wap* value leads to an upraising vertical wind w. In the convention of the PCMDI *wap* is referred as positive down and so an increase of *wap* means directly an increase in vertical wind velocity.

The vertical wind is normally low (10^{-2} to 10^{-3} times the horizontal wind components) and hard to measure [Kraus, 2004]. A central formalism to calculate the vertical air movement is the continuity equation. It says that the density of the air within a certain volume can only change if there is a flow through the margins of the volume:

$$\frac{\partial \rho}{\partial t} = -\left(\frac{\partial(\rho \cdot ua)}{\partial x} + \frac{\partial(\rho \cdot va)}{\partial y} + \frac{\partial(\rho w)}{\partial z}\right) = -\mathrm{div}(\rho \vec{v})$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho \cdot \mathrm{div}\vec{v}$$

if the incompressibility condition ρ=const. (div$\vec{v}$=0) is valid than

$$0 = \frac{\partial ua}{\partial x} + \frac{\partial va}{\partial y} + \frac{\partial w}{\partial z}$$

the integration from the ground to the height z yields to

$$\begin{aligned}
\int_0^z \frac{\partial w}{\partial z}\mathrm{d}z &= -\int_0^z \left(\frac{\partial ua}{\partial x} + \frac{\partial va}{\partial y}\right)\mathrm{d}z \\
&= -\int_0^z \vec{v_H}\mathrm{d}z \\
w(z) - w_0 &= -\frac{1}{z}\int_0^z \vec{v_H}\mathrm{d}z \cdot z \\
w(z) &= -\overline{\mathrm{div}\vec{v}_H}^{z} \cdot z
\end{aligned} \tag{2.5}$$

The bar over the divergence of the horizontal wind means averaging over the height z. If the incompressibility condition is true, the continuity equation says that the vertical air velocity is increasing with increasing height z. The integral from the ground z_0, where the vertical wind is zero, to the height z yields the vertical wind velocity at z. It is the negative average of the horizontal wind divergence over the height z multiplied with the height z. The vertical wind w is positive (updraft) if the mean divergence of the horizontal wind is negative between the heights 0 and z.

The vertical air velocity w respectively *wap* plays a major role in cloud building and dissolving processes. A more comprehensive derivation of this mathematical formalism can for example be found in Kraus [2004].

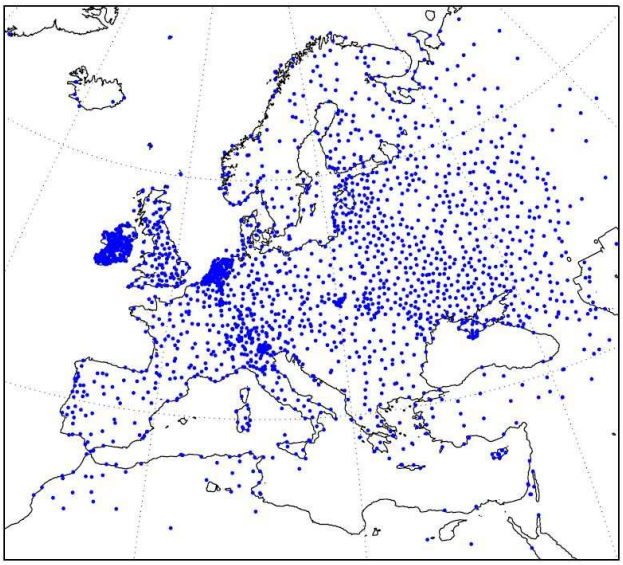


Figure 2.2: Station network of the E-OBS Dataset (August 2007) [Klok, 2007].

2.2 Reference Data

2.2.1 European Observation Dataset (E-OBS)

The E-OBS dataset contains a network of 2191 stations with daily observations and focuses on surface temperature and precipitation [Klonk and Klein Tank, 2008]. The density of the station network can be seen in figure 2.2 (state of August 2007). The average distance between two stations is 75 km. The time record period starts at 1950 and ends with 2006. The data are quality checked for plausibility and homogeneity. It was used as reference dataset for several projects including ENSEMBLES project (http://www.ensembles-eu.org/). In this work the E-OBS dataset is the reference dataset for the GCM evaluation of surface air temperature and precipitation in the period 1961 to 2000. The data is freely available for non-commercial research and can be downloaded from http://eca.knmi.nl/ (September 22, 2009).

2.2.2 European Center for Medium-Range Weather Forecasts (ECMWF) 40 Years Reanalysis Dataset (ERA40)

Many of the observation techniques, especially before satellite remote sensing techniques were developed, are point based (weather stations) or one dimensional data (radiosondes, ship measurements, flight data etc.). Since GCMs are working with grid based data, earth system models are used to recalculate the three dimensional state of the atmosphere at a specific point in time. The observed data provides initial, internal and boundary conditions within the model. This recalculated datasets are called reanalysis and are

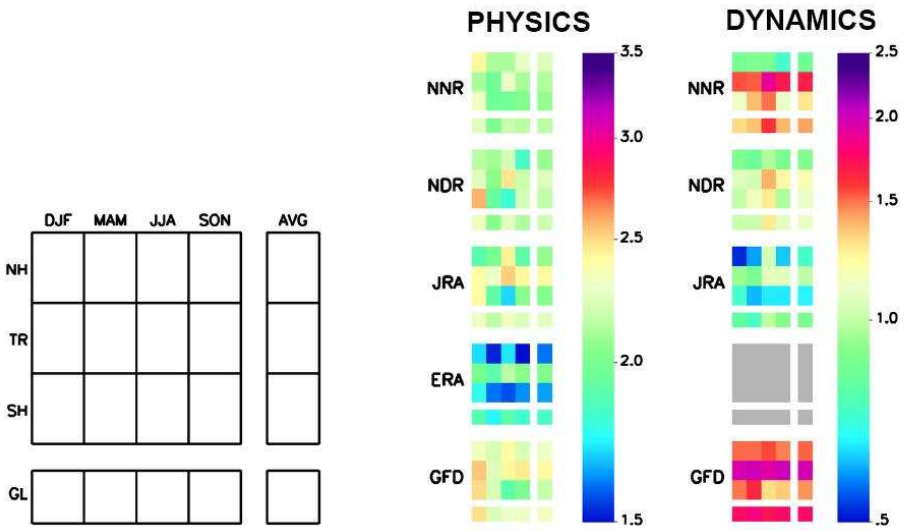


Figure 2.3: Normalized root mean square (RMS) errors of reanalysis/GCM data validated with the mean of multiple observations and averaged over all "physics" and "dynamics" quantities for different seasons (DJF - winter, MAM - spring, JJA - summer, SON - autumn, AVG - annual) and regions (NH - northern hemisphere, TR - tropes, SH - southern hemisphere, GL - global). (NCEP/NCAR (NNR), NCEP/DOE (NDR), ERA-40 (ERA), JRA-25 (JRA) and GFDL CM2.1 climate model (GFD)). Since the ERA 40 data were used as reference for the dynamic parameter analyze there are no results of the evaluation of these data (gray shaded boxes). [Reichler and Kim, 2008b].

provided by different research centers. Because reanalysis data is strongly based on statistics and model formulation, different reanalysis datasets show different results, especially in the pre satellite era and in regions with low observational density (e.g. southern hemisphere). There are different approaches to evaluate these reanalysis datasets (e.g., [Sterl, 2004; Reichler and Kim, 2008b]). For example, Reichler and Kim [2008b] evaluated four reanalysis datasets and one GCM with observation and European Reanalysis (ERA) data in the period 1979 to 1999 for 31 quantities. The calculations of normalized root mean square errors for different seasons, regions and observations showed that the quality of reanalysed data is strongly depending on the individual reanalysis model and the considered parameter. Reichler and Kim [2008b] stated that the ERA 40 reanalysis was able to reproduce the observed climate mean state better than other reanalysis products. Therefore, in this study the ERA40 data are used to evaluate GCM outputs like sea level pressure and the parameters in the free atmosphere (see section 2.5). An overview of the obtained results of the reanalysis dataset performance is given in figure 2.3.

The ERA40 dataset is provided by the ECMWF and covers the period

1957 to 2001 [Uppala et al., 2005]. The objective of the ERA40 project was to reconstruct the conditions in the atmosphere, land and surface for the mentioned period. The basis of the data reconstruction is the ECMWF data assimilation and forecasting system. The analyze is made on a T159L60 grid (1° lon, 1° lat) on a six hour time base [ECMWF, 2009]. For the reanalysis multiple sources of observation were merged, including satellite data since 1972 (see table 2.4).

The dataset can be downloaded directly from the ECMWF homepage (http://www.ecmwf.int/, September 22, 2009) on a daily and monthly basis with a spatial resolution of approximately 2.5° lat, 2.5° lon and 1.0° lat, 1.0° lon. In this work the higher resolved data were used. The data are provided without charges for research and education purposes only.

Table 2.4: Average daily counts of various types of observation supplied to the ERA40 data assimilation, for five selected periods [Uppala,2005].

Observation Type	1958-66	1967-72	1973-78	1979-90	1991-2001
SYNOP/SHIP	15313	26615	28187	33902	37049
Radiosondes	1821	2605	3341	2274	1456
Pilot balloons	679	164	1721	606	676
Aircraft	58	79	1544	4085	26341
Buoys	0	1	69	1462	3991
Satellite radiances	0	6	35069	131209	181214
Satellite winds	0	0	61	6598	45671
Scatterometer	0	0	0	0	7575
PAOBs	0	14	1031	297	277

2.3 Research Domains

The focus area of this study is the Alpine region. To estimate the effect of different domain settings (different boundary data) for RCM simulations, two domain sizes are investigated. As shown in 2.4 the large domain contains nearly entire Europe (with exception of Scandinavia), the Mediterranean Sea and parts of the Atlantic Ocean. The smaller one is embedded in the greater domain and contains the greater Alpine region.

2.4 Regridding

Both, references- and simulations data, are given on rectangular grids. The original grid size varies considerably from 0.5° (E-OBS) up to 5.0° (for the resolution of some GCMs (see table 2.1 and 2.2)). For an intercomparison of the data, all grid points have to be at exact the same position.

Figure 2.4: The two investigated domains (red: S=32°, N=60°, E=345°, W=35°), (green: S=38°, N=52°, E=0°, W=22°).

As standard a 0.5° latitude, longitude, rectangular grid was chosen. All datasets were regridded in two steps onto this "standard" grid.

In the first step the original data are regridded using linear interpolation to a grid with 0.1°. This step guarantees that the difference between the target domain and the refined grid is less than 0.05°. The interpolated data was cut out in the domains which are shown in figure 2.4.

In the second step the fine gridded data are sampled on the 0.5° standard grid. The average difference between the spatial mean values of the original data and the regridded data is under 0.1 percent.

2.5 Measuring of Model Performance

The CMIP3 Multi Model Dataset gives an overview of the state of the art general circulation models. With GCMs it is tried to reproduce the past climate processes as good as possible. It is questionable if models, which are able to reconstruct the past and present climate in a better way, are also able to give better estimates of the future climate than GCMs which perform bad (this topic was discussed in section 1.2.3). McAveny et al. [2001] found, that model quality is a necessary, but not sufficient criteria. Nevertheless, it is advisable to check the GCM performances in the past before looking at their forecasts. For this reason, in this study the data from the 24 GCMs from the CMIP3 database are evaluated against observations (E-OBS) and reanalysis data (ERA40) in the climatologically mean state from 1961 to 2001.

As discussed in section 1.3.1 there are different ways to evaluate the performance of GCMs. The MPI that was published by Reichler and Kim

[2008a] offers the possibility to examine the performance of GCMs concerning a broad range of climate variables.

The first step toward this MPI is the selection of a set of GCM output parameters. This selection should consider the application of the MPI result. The selected parameters in this study were described in section 2.1.1.

The second step toward the MPI is to calculate the normalized error variance e^2 for each model, each parameter and each pressure level by summing up the squared, weighted and normalized grid point differences between the simulated and the observational field (see formula 2.6).

$$e^2_{\nu m} = \sum_n (w_n(\overline{s}_{\nu mn} - \overline{o}_{\nu n})^2 / \sigma^2_{\nu n}) \tag{2.6}$$

Here $\overline{s}_{\nu mn}$ stands for the simulated climatology of the climate variable ν, the model m and the grid point n. Further, $\overline{o}_{\nu n}$ stands for the corresponding observation, while w_n are weights for area and mass averaging and $\sigma^2_{\nu n}$ represents the interannual variance of the corresponding observation. The division by $\sigma^2_{\nu n}$ is needed to homogenize the errors for all regions and variables.

In the next step the $e^2_{\nu m}$ values are normalized by the average values of all models ($\overline{e^2_{\nu m}}^m$) to ensure that the variables in the final MPI are equally weighted. The values $I^2_{\nu m}$ are called performance indices and are calculated for each parameter.

$$I^2_{\nu m} = e^2_{\nu m} / \overline{e^2_{\nu m}}^m \tag{2.7}$$

Finally the entire model performance index for one GCM is calculated by equally averaging over all performance indices of the climate variables ν (table 2.3).

$$I^2_m = \overline{I^2_{\nu m}}^\nu \tag{2.8}$$

The last step combines the performance of a GCM under consideration of multiple parameters to one single index. For the rare case that a parameter in the CMIP3 dataset was not available for a single GCM the missing values were filled with the average value of the remaining models. More background information about the MPI can be found in Reichler and Kim [2008a] and Murphy et al. [2004].

So far, (as mentioned in section 1.3) there is no single index which is able to evaluate all aspects of a GCM and also the MPI of Reichler and Kim [2008a] has its limitations. For example, the index only considers the mean state of climate and neglects higher moments like temporal variability or correlations which are essential for regional climate studies.

2.5.1 Robustness of the MPI

In section 2.5 the mathematical formulation of the model performance index is explained. The MPIs are single numbers, which are distributed around one. For a quantitative estimation of the robustness of the I^2 values a statistical method called bootstrapping is applied [Efron and Tibshirani, 1993].

The basis for this technique is that a random sample $\vec{x} = (x_1, x_2, \ldots, x_n)$ with an unknown probability distribution F has been observed. Here, this random sample is represented by 40 annually averages of a considered climate variable in the reference climatology 1961 to 2000. From this random sample, based on $\vec{x}$, a mathematical object (the model performance index) is calculated as described in section 2.5 so that $I^2 = t(F)$ ($t(F)$ is the real value of the MPI). Initially an estimate $\hat{I^2} = s(\vec{x})$ can be generated. The question is, how accurate this estimation of $\hat{I^2}$ is?

$$\hat{I^2} \rightarrow (x_1^*, x_2^*, \ldots, x_n^*) \tag{2.9}$$

The first step of the Bootstrap algorithm is to select B independent artificial subsamples, which are randomly drawn from $\vec{x}$ (with replacement). Each of the B samples consists of n elements.

In literature there is no clear answer how large B should be, but usually the number of bootstraps lies between 100 and 1000. For the calculations made in this thesis $B = 1000$.

In the second step the mathematical object is calculated for every artificial subsample b.

$$\hat{I^2}^*(b) = s(x_b^*)\ b = 1, 2, \ldots, B \tag{2.10}$$

In the third step the standard error $\hat{se}_F(\hat{I^2})$ is estimated by the standard deviation of the B samples.

$$\hat{se}_B = \left\{ \sum_{b=1}^{B} [\hat{I^2}^*(b) - \hat{I^2}^*(\cdot)]^2 / (B-1) \right\}^{1/2} \tag{2.11}$$

where $\hat{I^2}^*(\cdot) = \sum_{b=1}^{B} \hat{I^2}^*(b)/B$.

In general the standard error estimated with bootstrapping has a little bias [Efron and Tibshirani, 1993]. With increasing number of bootstraps the amount of the standard deviation is decreasing, so that small sample sizes B are generally overestimating the error.

2.5.2 Taylor Diagram

Karl E. Taylor [2000] developed a technique to display three different statistical values in one single diagram. These three values are:

1. **Correlation Coefficient** (R)
   The correlation coefficient is used to quantify similarities [Leonhart, 2004]. In this thesis the spatial correlations of the reference o and the simulated fields s in the climatology 1961 to 2000 were evaluated.

$$R = \frac{1/N \sum_{n=1}^{N} (o_n - \bar{o}) \cdot (s_n - \bar{s})}{\sigma_o \sigma_s} \tag{2.12}$$

   The correlation coefficient varies between plus and minus one. The maximum/minimum value are obtained when $\forall\, n,\ (o_n - \bar{o}) = \alpha(s_n - \bar{s})$, where α is a positive constant, if the data is positive correlated and α is a negative constant, if it is anti correlated.

2. **Root Mean Square (RMS) Difference** (E_D)
   The RMS of two fields is defined as:

$$E_D = \left[1/N \sum_{n=1}^{N} (o_n - s_n)^2 \right]^{1/2} . \tag{2.13}$$

   It is a combination of two parts. The first arises from the overall bias $\bar{E} = \bar{o} - \bar{s}$ and the second from the centered pattern RMS difference $E' = \left\{ 1/N \sum_{n=1}^{N} [(o_n - \bar{o}) - (s_n - \bar{s})]^2 \right\}^{1/2}$. The quadratic amount of these two quantities yield the full mean square difference $E_D^2 = \bar{E}^2 + E'^2$. The RMS gets lower for fields which have a high spatial match.

3. **Variances** (σ_o, σ_s)
   The variances of the two fields show the spatial variability and are calculated as follows:

$$\sigma_i = \left[1/N \sum_{n=1}^{N} (i_n - \bar{i_n})^2 \right]^{1/2} \qquad \text{for } i = s, o \tag{2.14}$$

All of the above described values provide specific statistical information and it is possible to display this information in one single diagram according to the following relationships:

$$\begin{aligned} E'^2 &= \sigma_o^2 + \sigma_s^2 - 2\sigma_o \sigma_s R, \\ &\text{and the Law of Cosines,} \\ c^2 &= a^2 + b^2 - 2ab\cos\phi. \end{aligned} \tag{2.15}$$

In the second equation in 2.15 the variables a, b and c are representing the length of the sides of a triangle and ϕ is the angle between a and b and the relationship of R, E', σ_o and σ_s is displayed in figure 2.5.

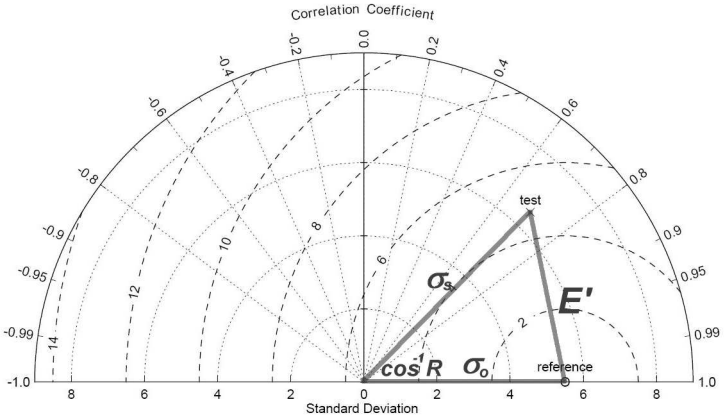


Figure 2.5: The azimuthal position of the simulated field (test) represents the correlation coefficient R. The distance between the origin and the test field is the variance of σ_s and the distance from the origin to the reference field is σ_o. The RMS E' is represented by the distance from the reference field to the simulated field [Taylor, 2000].

The Taylor plots which are used in this study have a standardized variance (σ_s/σ_o and $\sigma_o/\sigma_o = 1$), so that the reference point always has the coordinates [1/0].

2.6 Box-Whisker Plots

In this thesis box-whisker plots are used to illustrate the distribution of the climate change signals on an annual and seasonal time scale. They are combining values of the central moments with measures of dispersion in one graph [Leonhart, 2004]. The structure of a box with whiskers is shown in figure 2.6.

The line in the middle of the box shows the median value of the data. Around the median the box is drawn which carries 50 % of the data. The distribution of the data is symmetric if the median is in the middle of the box. A small box indicates a low variability whereas a large box shows a high one. The thin lines at both ends of the Boxes (whiskers) are showing the spread of the distribution without outliers. Their maximum length is 1.5 times the length of the box. Therefore, the statistical normal values of the distribution are lying within the fourfold box length. All data points, which are outside the whiskers are called extremes or outliers. More about box-whisker plots can be found in [Leonhart, 2004] or [Sachs and Hedderich, 2006].

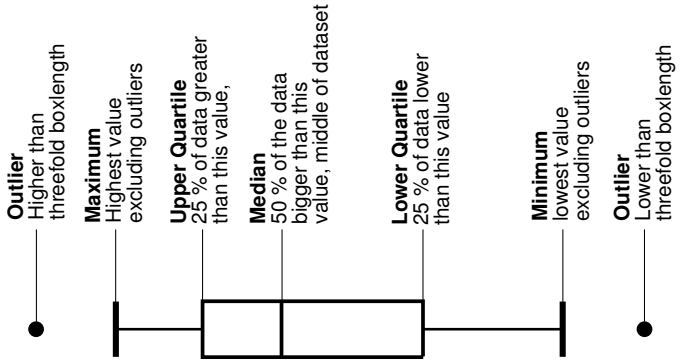


Figure 2.6: Theoretical structure of a box-whiscer plot.

2.7 Significance of the Climate Change Signal

The climate system is subject of interannual fluctuations, which are not a part of climate change signals in the long term. In this section a method is described how the significance of a climate change signal can be tested. The basic assumption is, that the variables of the two compared climatologies (reference 1971 to 2000 with future 2021 to 2050 and 2071 to 2100) are independent.

2.7.1 Testing for Gaussian Distribution

Since many statistical tests assume a normal distribution of the sample population the Kolmogorov-Smirnov test (KS-test) is used to check data for this property [Sachs and Hedderich, 2006]. The KS-test compares an empirical observed distribution with a known theoretical distribution (here with a normal distribution). The test value is the maximum difference between the empirical and the theoretical distribution.

The KS-test always tests the null hypothesis that the observed distribution $F(x_i)$ is matching a normal distribution $F(y_i)$. If the test value $\hat{D}$ gets bigger than the critical value (achieved from tables) there are significant deviations from the normal distribution and the tested variable can not be considered as normal distributed.

$$\hat{D} = \max_{1 \leq i \leq N} |F(y_i) - F(x_i)| \tag{2.16}$$

The evaluations in chapter 3 will show that the H0 hypotheses (the populations are normal distributed) always were acceptable on a test level of $\alpha = 0.05$. A corresponding table of the critical values and a more detailed derivation of the method can be found in [Sachs and Hedderich, 2006]

2.7.2 Test for the Homogeneity of Variances

Before a t-test for different meann vallues can be performed, the homogeneity of the variances has to be checked [Leonhart, 2004]. Homogeneity is given if the variance in both samples (reference and future climatology) is equal or varies only randomly. Therefore the F-test is used:

$$F = \frac{\sigma_1^2}{\sigma_2^2} \tag{2.17}$$

with the following degrees of freedom:

$$\begin{aligned} df_1 &= n_1 - 1 \\ df_2 &= n_2 - 2. \end{aligned} \tag{2.18}$$

This test defines σ_1^2 as the bigger variance so that F is always greater or equal one. The degrees of freedom, df_1 and df_2 from equation 2.18 are used to find the critical value of F in a table of F-distributions. The test is performed for the test level α=0.05. The two hypothesis are:

- H0: The variances are homogeneous

- H1: The variances are not homogeneous.

If the critical value F (obtained from tables in, e.g., [Sachs and Hedderich, 2006] or [Leonhart, 2004]) is larger than the calculated test value, the null hypothesis is retained.

Depending on the result of the variance homogeneity test the t-test for independent samples is performed for homogeneous or heterogeneous variances.

2.7.3 t-test for Homogeneous Variances

With this test the difference in the mean state of a climate variable in the reference (1971 to 2000) and future climatology (2021 to 2050 or 2071 to 2100) is analyzed. The precondition is that the two variances of the samples are homogeneous [Sachs and Hedderich, 2006].

Its definition is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n-1)\cdot s_1^2 + (n_2-1)\cdot s_2^2}{n_1+n_2-2} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{2.19}$$

with

$$df = n_1 + n_2 - 2.$$

If the calculated test value t is greater than the critical value t_c, the H0 hypothesis, which says that the mean values of the two samples are equal, is rejected and the H1 hypothesis, which says that the means are different, is accepted [Sachs and Hedderich, 2006]. The test levels are α=0.1, 0.05, 0.025 and 0.01.

2.7.4 t-test for Heterogeneous Variances

The t-test for heterogeneous variances is also used to calculate, if two mean values are equal or not, but in comparison to section 2.7.3 this test does not assume variance homogeneity.

The mathematical formulation of the test is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}$$

with

$$df = \frac{(n_1 - 1) \cdot (n_2 - 1)}{(n_2 - 1) \cdot c^2 + (n_1 - 1) \cdot (1 - c)^2} \tag{2.20}$$

and

$$c = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2}$$

Again, if the calculated test value t is greater than the critical value t_c the H0 hypothesis (the means of the two samples are equal) is rejected and the H1 hypothesis (the means are different) is accepted. The test levels are the same as in the case of homogeneous variances.

2.8 Calculation of the Climate Sensitivity

In the fourth assessment report of the IPCC the climate sensitivity is referred as: "... the equilibrium change in the annual mean global surface temperature following a doubling of the atmospheric equivalent carbon dioxide concentration." [IPCC, 2007]. In this study the word climate sensitivity is used as the response of different climate parameters due to 21st century greenhousegas forcing corresponding to the emission scenarios B1, A1B and A2. This responses are the differences between the reference climatology (1971 to 2000) and the two forecasted climatologies (2021 to 2050 and 2071 to 2100). This quantity is usually called climate change signal. Hereby, a bigger total difference between the two climatologies is showing a more climate sensitive GCM.

To obtain the mean climate sensitivity of a GCM (averaged over all evaluated parameters) the multi model mean climate change signals of the

parameters are calculated. Relative to these mean values the spread of the individual GCM signals are scaled between ± 100 % (the GCM with the smallest climate change signal has a value of -100 % and the GCM with the biggest +100 %). In the last step the relative climate change signals (in percent) are equally averaged over all parameters for every GCM to obtain the mean climate sensitivity of the models.

2.9 Spearman Rank Correlation

To test the hypothesis of section 1.2.3, if GCMs which are able to reproduce the past climate are more reliable for future forecasts, the correlations between the model performance indices and the climate change signals are tested. The hypothesis is, for example, that GCMs, with a better presentation of surface temperature in the past, have higher temperature signals in the future.

Since the climate change signals are based on an interval scale and the model performance indices are given on an ordinal scale, the Spearman rank correlation test is used [Leonhart, 2004]. For this test, the values of both indices are first transformed to a rank-order and then the correlations between the rank-orders are calculated. This test can be performed for parameters of any distribution. The differences between the ranks $d_i = \text{rank}_{xi} - \text{rank}_{yi}$ is the basis for the rank correlation:

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^{N} d_i^2}{N \cdot (N^2 - 1)} \tag{2.21}$$

Here N is the number of elements in x_i and y_i. The significance test is performed with the t-test:

$$t = \frac{r_s}{\sqrt{\frac{1-r_s^2}{N-2}}} \tag{2.22}$$

If t is bigger than the critical value t_c (with $df = N-2$ degrees of freedom) the null hypothesis (the two parameters are not correlated) is rejected.

2.10 Uncertainty Estimation

A central aspect of this thesis is the analysis of uncertainties in climate change signals. As mentioned in section 1.2.4 there are three sources of uncertainties (model-, emission- and internal (initial condition) uncertainties). In the sections 2.10.1 and 2.10.2 two techniques will be described how these uncertainties can be estimated.

2.10.1 Mean Variance of Uncertainty Sources

The method which is described below is taken from a paper of Rowell [2006]. Because some sample sizes are quite small, this method allows only a rough estimation of uncertainties. Nevertheless, it gives a good overview of the distribution of the data.

The standard deviation of the three uncertainty groups is calculated as followed:

$$
\begin{aligned}
s_E &= \sqrt{\frac{1}{p-1}\sum_{i=1}^{p}(X_{i\cdot\cdot} - X_{\cdots})^2} \text{ for } i = 0, 1, \cdots, p \\
s_M &= \sqrt{\frac{1}{q-1}\sum_{j=1}^{q}(X_{\cdot j\cdot} - X_{\cdots})^2} \text{ for } j = 0, 1, \cdots, q \\
s_I &= \sqrt{\frac{1}{r-1}\sum_{k=1}^{r}(X_{\cdot\cdot k} - X_{\cdots})^2} \text{ for } k = 0, 1, \cdots, r.
\end{aligned}
\tag{2.23}
$$

The indices E, M and I are corresponding to three uncertainty groups (emissions, models and internal). Each group has different group sizes p, q and r. The parameter X_{ijk} represents the matrix which contains the climate change signals (difference between 1971 to 2000 and 2021 to 2050 or 2071 to 2100) of a considered parameter of a GCM. X_{ijk} has a lot of missing values and is not equally filled. Models with simulations in all groups (three emission scenarios with up to nine runs, see figure 2.1) are therefore stronger weighted in the standard deviation calculation as GCMs with less simulations. The appearance of a dot in the index of X (e.g., $X_{i\cdot\cdot}$) means that the average over the concerning indices is calculated.

2.10.2 Analysis of Variance (ANOVA)

In this section the ANOVA is described, which can be used to estimate uncertainties occurring due to a set of indicator variables [Storch and Zwiers, 1999]. Here we analyze the dependence of a variable which is affected by three indicator variables (choice of model M, choice of emission scenario E and choice of model run I (initial condition uncertainty)). Each variable has different level sizes [Eckey et al., 2002]:

- **Models** M: $i = 1, 2, \cdots, p$; $p = 24$
- **Emissions** E: $j = 1, 2, 3$; $q = 3$
- **Internal** I: $k = 1, 2, \cdots, r$; $r = 9$
- **Investigation unit** N (years/seasons): $l = 1, 2, \cdots, n$; $n = 30$

Overall there are $p \cdot q \cdot r$ indicator levels or elements in the matrix y_{ijk} whereas each level contains n investigation units (the 30 annual mean values of the variable). The values y_{ijkn} are considered as a realization of a random variable (y), which is normally distributed in every sample of $p \cdot q \cdot r$. The population, which belongs to the groups $ijkl$ has an expected value of $\mu_{ijk\cdot}$ and a variance of $\sigma^2_{ijk\cdot}$ so that:

$$y_{ijkl} \approx N\left(\mu_{ijk\cdot}, \sigma^2_{ijk\cdot}\right) \text{ for all } i, j, k, l \tag{2.24}$$

The elementary approach has the form:

$$y_{ijkl} = \mu_{ijk\cdot} + \epsilon_{ijk\cdot} \text{ for all } i, j, k, l \tag{2.25}$$

where $\epsilon_{ijk\cdot}$ are random variables, which are independent and identical $N(0, \sigma^2_\epsilon)$ distributed.

The expected value of the ith group is:

$$\mu_{i\cdots} = \frac{1}{q \cdot r} \cdot \sum_{j=1}^{q} \sum_{k=1}^{r} \mu_{ijk\cdot\cdot} \tag{2.26}$$

The expected values of $\mu_{\cdot j\cdot}$ and $\mu_{\cdot\cdot k}$ are calculated identically. Further, the expected value of the groups, which are build by the factor levels i and j are:

$$\mu_{ij\cdot\cdot} = \frac{1}{r} \cdot \sum_{k=1}^{r} \mu_{ijk\cdot\cdot} \tag{2.27}$$

This definition can also be applied for the expected values of $\mu_{i\cdot k}$ and $\mu_{\cdot jk}$. Finally, the overall expected value is:

$$\mu_{\cdots\cdot} = \frac{1}{p \cdot q \cdot r} \cdot \sum_{i=1}^{p} \sum_{j=1}^{q} \sum_{k=1}^{r} \mu_{ijk\cdot\cdot} \tag{2.28}$$

The subject of investigation is, if the factors M, E and I as the factor combinations ME, MI, EI and EMI have a significant influence on the dependent variable.

Therefore, equation 2.25 is reformed so that the effect of the different levels gets visible.

$$\begin{aligned} y_{ijkl} - \mu_{\cdots\cdot} =& (\mu_{i\cdots} - \mu_{\cdots\cdot}) + (\mu_{\cdot j\cdot\cdot} - \mu_{\cdots\cdot}) + (\mu_{\cdot\cdot k\cdot} - \mu_{\cdots\cdot}) \\ &+ (\mu_{ij\cdot\cdot} - \mu_{i\cdots} - \mu_{\cdot j\cdot\cdot} + \mu_{\cdots\cdot}) + (\mu_{i\cdot k\cdot} - \mu_{i\cdots} - \mu_{\cdot\cdot k\cdot} + \mu_{\cdots\cdot}) \\ &+ (\mu_{\cdot jk\cdot} - \mu_{\cdot j\cdot\cdot} - \mu_{\cdot\cdot k\cdot} + \mu_{\cdots\cdot}) \\ &+ (\mu_{ijk\cdot} + \mu_{i\cdots} + \mu_{\cdot j\cdot\cdot} + \mu_{\cdot\cdot k\cdot} - \mu_{ij\cdot\cdot} - \mu_{i\cdot k\cdot} - \mu_{\cdot jk\cdot} - \mu_{\cdots\cdot}) \\ &+ \epsilon_{ijk\cdot} \end{aligned} \tag{2.29}$$

The effects, which arise from the single factors are written in the first line of equation 2.29, mixed effects of two factors are displayed in the second and third row and the effect of the combination of all three factors are represented through the expression in the fourth row.

For $\epsilon_{ijk\cdot}$ the following equation is true:

$$\epsilon_{ijk\cdot} = y_{ijkl} - \mu_{ijk\cdot} \tag{2.30}$$

The sum of squares decomposition of the variance is calculated by considering the different sample values.

$$Q_T = Q_M + Q_E + Q_I + Q_{ME} + Q_{MI} + Q_{EI} + Q_{MEI} + Q_R \tag{2.31}$$

Hereby the total variance Q_T is defined as:

$$\begin{aligned} Q_T &= \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n}(y_{ijkl} - y_{....})^2 \\ &= \underbrace{\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}^2}_{=Q_{0,T}} - \underbrace{\frac{1}{N}\cdot\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_0} \end{aligned} \tag{2.32}$$

where $N = p \cdot q \cdot r \cdot n$. The part of the sum of squares decomposition, which is explained by the factor M is:

$$\begin{aligned} Q_M &= \sum_{i=1}^{p} n \cdot (y_{i...} - y_{....})^2 \\ &= \underbrace{\sum_{i=1}^{p}\frac{1}{q\cdot r\cdot n}\cdot\left(\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M}} - \underbrace{\frac{1}{N}\cdot\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_0} \end{aligned} \tag{2.33}$$

The values for the factors E and I are calculated accordingly.

The factors ME are explained by the following part of the sum of squares decomposition.

$$
\begin{aligned}
Q_{ME} &= \sum_{i=1}^{p}\sum_{j=1}^{q} r \cdot n \cdot (y_{ij\cdot\cdot} - y_{i\cdots} - y_{\cdot j\cdot\cdot} + y_{\cdots\cdot})^2 \\
&= \underbrace{\sum_{i=1}^{p}\sum_{j=1}^{q} \frac{1}{r \cdot n} \cdot \left(\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M,E,ME}} - \underbrace{\frac{1}{n \cdot q \cdot r} \cdot \sum_{i=1}^{p}\left(\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M}} \\
&\quad - \underbrace{\frac{1}{n \cdot p \cdot r} \cdot \sum_{j=1}^{q}\left(\sum_{i=1}^{p}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,E}} + \underbrace{\frac{1}{N} \cdot \left(\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_0}
\end{aligned}
\tag{2.34}
$$

Again, the factors MI and EI are calculated simultaneously.
The interdependency factor of all three groups can be written as:

$$
\begin{aligned}
Q_{MEI} &= \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r} n \cdot (y_{ijk\cdot} + y_{i\cdots} + y_{\cdot j\cdot\cdot} + y_{\cdot\cdot k\cdot} - y_{ij\cdot\cdot} - y_{i\cdot k\cdot} - y_{\cdot jk\cdot} - y_{\cdots\cdot})^2 \\
&= \underbrace{\frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{q}\cdot\sum_{k=1}^{r}\left(\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M,E,I,ME,MI,EI,MEI}} + \underbrace{\frac{1}{n \cdot q \cdot r} \cdot \sum_{i=1}^{p}\left(\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M}} \\
&\quad + \underbrace{\frac{1}{n \cdot p \cdot r} \cdot \sum_{j=1}^{q}\left(\sum_{i=1}^{p}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,E}} + \underbrace{\frac{1}{n \cdot p \cdot q} \cdot \sum_{k=1}^{r}\left(\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,I}} \\
&\quad - \underbrace{\frac{1}{r \cdot n} \cdot \sum_{i=1}^{p}\sum_{j=1}^{q}\left(\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M,E,MI}} - \underbrace{\frac{1}{q \cdot n} \cdot \sum_{i=1}^{p}\sum_{k=1}^{r}\left(\sum_{j=1}^{q}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M,I,MI}} \\
&\quad - \underbrace{\frac{1}{p \cdot n} \cdot \sum_{j=1}^{q}\sum_{k=1}^{r}\left(\sum_{i=1}^{p}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,E,I,EI}} - \underbrace{\frac{1}{N} \cdot \left(\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_0}
\end{aligned}
\tag{2.35}
$$

The last missing element in 2.31 is the “not explainable” part Q_R.

$$
\begin{aligned}
Q_R &= \sum_{i=1}^{p}\sum_{j=1}^{q}\cdot\sum_{k=1}^{r}\sum_{l=1}^{n}\left(y_{ijkl}-y_{ijk\cdot}\right)^2 \\
&= \underbrace{\sum_{i=1}^{p}\sum_{j=1}^{q}\cdot\sum_{k=1}^{r}\sum_{l=1}^{n} y_{ijkl}^2}_{=Q_{0,T}} - \underbrace{\frac{1}{n}\cdot\sum_{i=1}^{p}\sum_{j=1}^{q}\cdot\sum_{k=1}^{r}\left(\sum_{l=1}^{n} y_{ijkl}\right)^2}_{=Q_{0,M,E,I,ME,MI,EI,MEI}}
\end{aligned}
\tag{2.36}
$$

Still missing are the degrees of freedom of the sum of squares decompositions, which can be decomposed as follows:

$$
df_T = df_M + df_E + df_I + df_{ME} + df_I + df_{MI} + df_I + df_{EI} + df_I + df_{MEI} + df_R, \tag{2.37}
$$

with:

$$
\begin{aligned}
df_T &= n\cdot p\cdot q\cdot r - 1 \\
df_M &= p - 1 \\
df_E &= q - 1 \\
df_I &= r - 1 \\
df_{ME} &= (p-1)\cdot(q-1) \\
df_{MI} &= (p-1)\cdot(r-1) \\
df_{EI} &= (q-1)\cdot(r-1) \\
df_{MEI} &= (p-1)\cdot(q-1)\cdot(r-1) \\
df_R &= n\cdot p\cdot q\cdot r - p\cdot q\cdot r.
\end{aligned}
\tag{2.38}
$$

The required test statistic can be defined as followed:

$$
\begin{aligned}
F_M &= \frac{Q_M}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{p-1} \\
F_E &= \frac{Q_E}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{q-1} \\
F_I &= \frac{Q_I}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{r-1} \\
F_{ME} &= \frac{Q_{MB}}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{(p-1) \cdot (q-1)} \\
F_{MI} &= \frac{Q_{MC}}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{(p-1) \cdot (r-1)} \\
F_{EI} &= \frac{Q_{EC}}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{(q-1) \cdot (r-1)} \\
F_{MEI} &= \frac{Q_{MEI}}{Q_R} \cdot \frac{n \cdot p \cdot q \cdot r - p \cdot q \cdot r}{(p-1) \cdot (q-1) \cdot (r-1)}
\end{aligned} \tag{2.39}
$$

The critical values are taken from a tabular, which for example can be found in Eckey [1999]. If one of the test values from formula 2.39 is greater then the corresponding critical value than it significantly effects the dependent variable.

The evaluations which will be shown in chapter 3 demonstrate that nearly all F values are significant with exception of the threefold combination term F_{MEI}.

2.10.3 Reconstructing Missing Data

In the previous chapter the ANOVA method was described and if there would be no missing data in y_{ijkl} (see figure 2.1) the eight values from equation 2.31 could be calculated straight forward. With gaps in the data matrix y_{ijkl} the variance Q_T can not longer be written as a sum [Déqué et al., 2007]. Further, this approach would lead to biases because, for example, the sum of experiments which were performed with the emission scenario A1B is not the same as it is in the scenarios A2 or B1.

The solution found by Déqué et al. [2007] (which is used in this study) is a reconstruction of the missing values so that the threefold interaction term Q_{MEI} in equation 2.31 is minimum.

An easy and short example should explain this method before it is described explicitly. In this simple assumption there are only two factors (e.g., GCMs in columns and scenarios in lines) for one variable and there is only one missing value x_8 in matrix X:

$$
\mathbf{X} = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & \mathbf{?} & x_9 \end{pmatrix} = \begin{pmatrix} 2.1 & 0.5 & 1.3 \\ 1.9 & 2.8 & 2.2 \\ 1.8 & \mathbf{?} & 2.3 \end{pmatrix} \tag{2.40}
$$

In this case the interaction terms are:

$$\begin{aligned} Q_M &= \sum_{i=1}^{3} (y_{i\cdot} - y_{\cdot\cdot})^2 \\ Q_E &= \sum_{j=1}^{3} (y_{\cdot j} - y_{\cdot\cdot})^2 \\ Q_{ME} &= \sum_{i=1}^{3} \sum_{j=1}^{3} (y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot})^2 \end{aligned} \tag{2.41}$$

The method assumes that the highest interaction term (here Q_{ME}) is the minimum so:

$$\begin{aligned} 0 &= y_{ij} - y_{i\cdot} - y_{\cdot j} + y_{\cdot\cdot} \\ y_{i,j} &= y_{i\cdot} + y_{\cdot j} - y_{\cdot\cdot} \\ y_{3,2} &= y_{3\cdot} + y_{\cdot 2} - y_{\cdot\cdot} = 2.05{+}1.65{-}1.86{=}1.84 \end{aligned} \tag{2.42}$$

So, the missing value for the model two and scenario three is approximately 1.84. This approach assumes that the effects of GCMs and emission scenarios are additively. For three factors the procedure is similar, but more complex. For each time slice $l = 1, \cdots, 30$ a three dimensional matrix with many missing values has to be filled up. The procedure can be summarized in five steps [Déqué et al., 2007]:

1. Calculating the mean $y_{\cdot\cdot\cdot}$.

2. Calculating the double mean $y_{i\cdot\cdot}$. This should be no problem since for every GCM there is at least one available experiment. The same calculation is done for $y_{\cdot j\cdot}$ and $y_{\cdot\cdot k}$.

3. Calculation of the single averages $y_{ij\cdot}$. Not all combinations of GCMs and emission scenarios can be calculated because there are missing values in the matrix. Therefore the method of minimizing the interaction terms is used. For the missing combinations of i and j the equation $y_{ij\cdot}$=$y_{i\cdot\cdot}$+$y_{\cdot j\cdot}$-$y_{\cdot\cdot\cdot}$ is applied. The same is done for $y_{i\cdot k}$ and $y_{\cdot jk}$.

4. The missing values of y_{ijk} are calculated using y_{ijk}=$y_{ij\cdot}$+$y_{i\cdot k}$+$y_{\cdot jk}$-$y_{i\cdot\cdot}$-$y_{\cdot j\cdot}$-$y_{\cdot\cdot k}$+$y_{\cdot\cdot\cdot}$. The terms for this equation are known from the previous steps.

5. Since the missing y_{ijk} should have been involved in the calculation process an iterative process is applied. This is possible because the equation system is linear. The terms from the right hand side of the

equation in step 4 are calculated from the result of y_{ijk} from the previous step. Only the missing terms in y_{ijk} are updated. If this routine is repeated 10 times, the changing in the y_{ijk} values is smaller than 0.1 %.

Chapter 3

Findings

In this chapter the methods of chapter 2 are applied to analyze:

- the ability of GCMs to reproduce the 1961 to 2000 climatology.
- the temporal development, significance and amplitude of the climate change signals.
- if well performing models are simulating a different future climate than worse performing GCMs.
- the amounts and sources of uncertainties of future climate projections.

3.1 Model Evaluation for the Climate Period 1961 to 2000

In the following sections at first the mean annual cycles of all parameters are displayed. Then statistical values like the mean, maximum and minimum bias and the RMS error of the spatial fits are evaluated. In Taylor diagrams the spatial variances, RMS and the spatial correlation coefficient are displayed. Furthermore the MPI of Reichler and Kim [2008a] is calculated and evaluated.

3.1.1 Annual Cycle

Figures 3.1 and 3.2 show the 40 year annual mean of the simulated data (thin colored lines) and the reference data (thick red line) for all considered parameters on the large domain. The thick black line marks the characteristics of the multi model mean.

In most cases the different GCMs scatter around the observation values (e.g. at *tas*, *hus* or *zg*), but there are also some systematic biases visible in the multi model field. For example, the winter precipitation is overestimated

and the sea level pressure is generally underestimated by the majority of GCMs.

For almost all evaluated variables there are some models with relative big biases. For instance, the iap_fgoals1_0_g model has an extreme cold bias in winter for *tas* and *ta* in 850 hPa or the geopotential height (*zg*) in the ipsl_cm4 GCM has a constant negative bias, which is more than two magnitudes lower than those of other models.

Although annual cycles don't give information about spatial properties like correlation or variability of the variables, still they are a good method to show mean biases and the seasonal variability of errors in GCM simulations. Furthermore, they displays the multi model spread of the CMIP3 dataset and detect GCMs with extreme biases. Nevertheless, it is possible that a GCM with nearly no mean spatial bias has areas with very high positive and negative biases, which cancel out by averaging. Therefore, further evaluation methods have to be applied.

The annual cycles of the variables in the small domain are similar to those of the large domain. Their general behavior is changing a little bit more to continental climate (stronger amplitudes in the annual cycles), but the scattering of the models stay similar and the outlying GCMs are staying the same. The corresponding plots are included in the digital appendage.

3.1.2 Spatial mean, max., min. Bias and RMS

In the former section the annual cycle of the GCMs was evaluated. On the following pages the focus is on spatial properties. Therefore, differences between model data and reference data are plotted on maps.

In figure 3.3 the spatial bias of surface temperature for two models and the multi model mean is displayed. The bigger map on the left shows the annual multi model mean minus the E-OBS reference dataset. The bias in the surface temperature field in areas where the ground level is rising (Alps, Pyrenees or Carpathians) is mainly caused by differences in the resolution of the GCM grids and the reference data grids[1]. This error is highest for surface air temperature because of its dependency on altitude. For the other parameters, especially those on fixed pressure levels, this error is generally small.

Concerning surface air temperature in figure 3.3 the mean bias of the multi model mean is 0.5 K and the maximum and minimum biases are 8.4 K and -4.8 K. The RMS error has an amount of 1.6 K. In the upper right picture in figure 3.3 the bias of the worst performing model for annually surface air temperature (the iap_fgoals1_0_g GCM) is displayed. It has a

[1]The orography is resolved finer in the reference data (E-OBS and ERA40) than in the GCMs dataset because E-OBS has a grid size of 25 km and ERA40 has approximately 100 km. Since there is no height correction of surface temperature in the regridding routine, there is a positive surface temperature bias in mountainous areas.

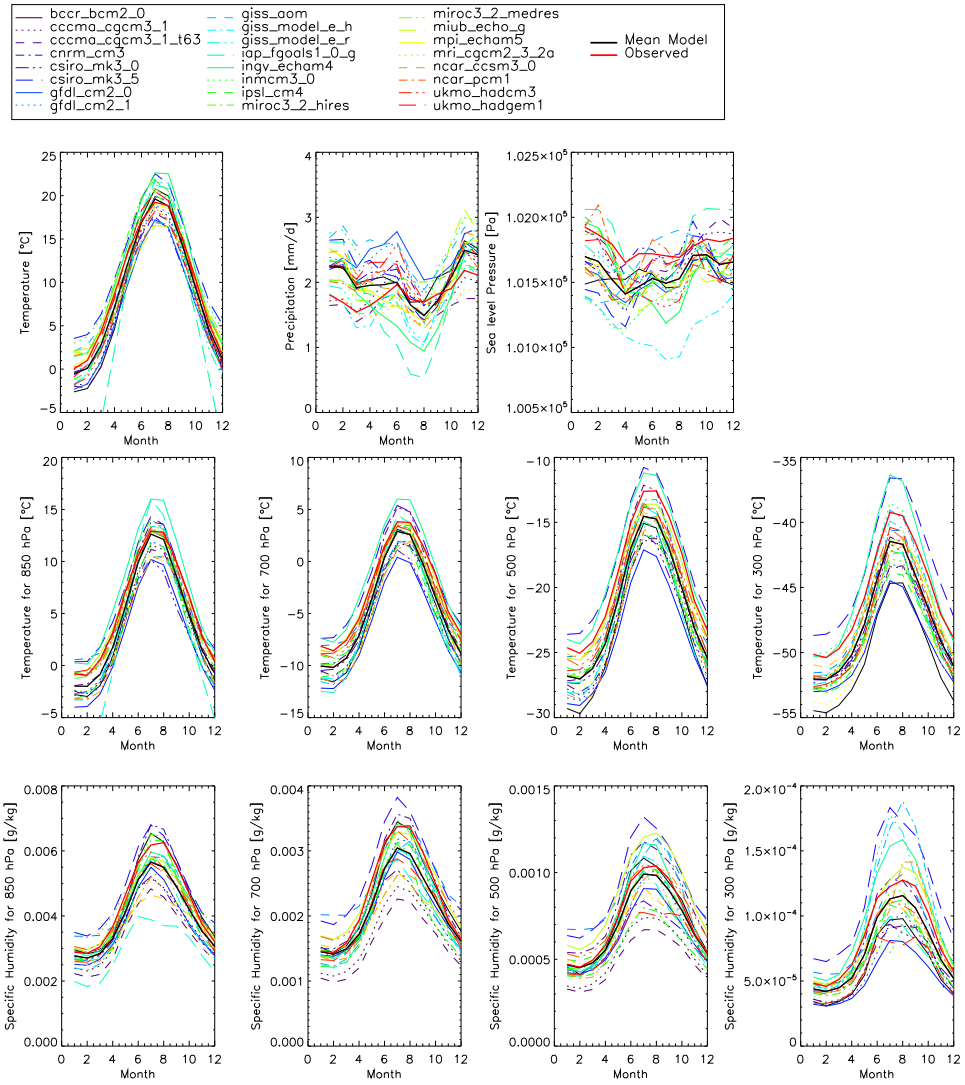


Figure 3.1: Spatial mean of the large domain annual cycle for *tas*, *pr*, *psl*, *ta* and *hus* in the 1961 to 2000 climatology. Thin collored lines are different GCMs while the thick red line is corresponding with the reference data.

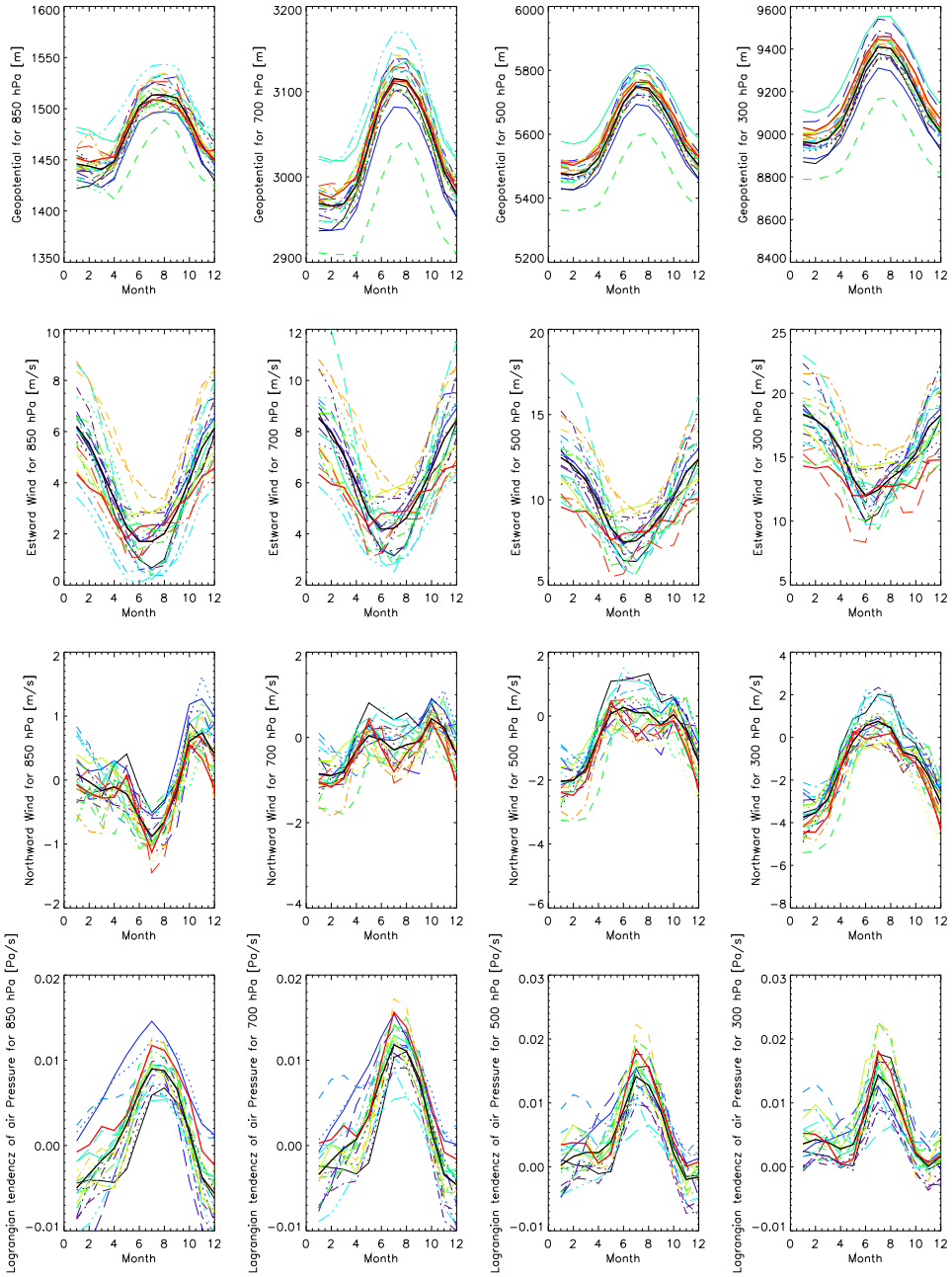


Figure 3.2: Same as figure 3.1 for *zg*, *ua*, *va*, and *wap*. The legend is the same as in figure 3.1.

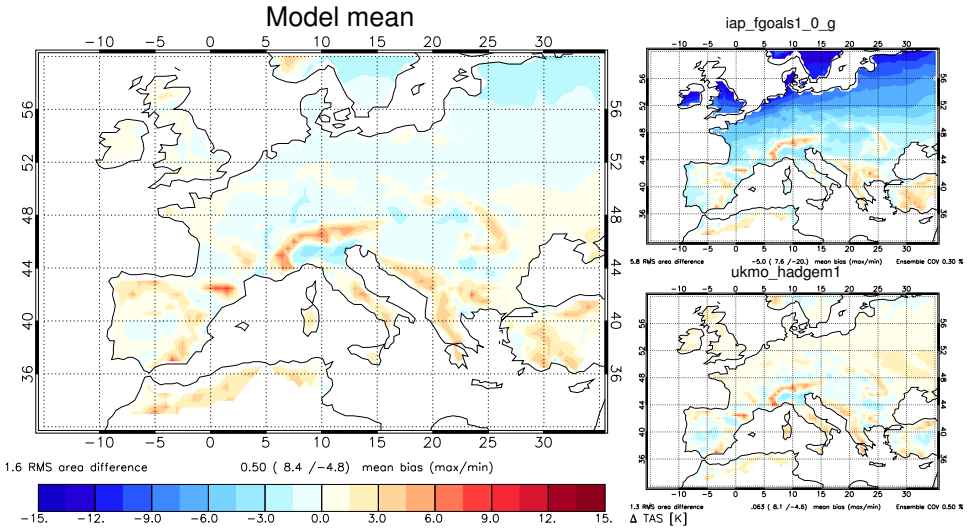


Figure 3.3: Spatial bias of the multi model mean (left), the iap_fgoals1_0_g (upper right) and the ukmo_hadgem1 (lower right) GCM minus the E-OBS reference dataset annually for *tas* in the 1961 to 2000 climatology.

cold bias in northern Europe with a maximum of up to -20 K and a RMS of 5.8 K. The last map (lower right corner) shows the best performing model (the hadgem3 GCM). Its RMS error is only 1.3 K and there is no strong bias outside of mountainous reagions.

The importance of seasonal evaluations can be seen in figure 3.4, where the errors of the precipitation flux of the multi model mean are shown. The season with the highest RMS (RMS=0.7 mm/d) is winter. There is a wet bias over large areas in Europe with maxima in France and northwest Spain. Very dry spots can be found in northern Portugal, Scotland and southern Norway. In the Alps the multi model winter precipitation is slightly underestimated.

In summer the overall RMS is smallest and there are no particularly wet or dry spots. However, one big difference to the other seasons is a dry area in the central and eastern alps. It is likely that this dry bias comes from difficulties in the convective precipitation parameterization. This is the prevalent form of precipitation during summer in this area [Cebon et al., 1998].

The spatial bias maps are drawn for every model, parameter, pressure level, domain and season. The corresponding plots are available in the digital appendix.

The advantage of clearness and visibility of spatial bias structures in maps has the disadvantage that it is hard to compare a couple of GCMs directly with each other. Therefore, the Taylor plot method with its high information density is used in the next chapter.

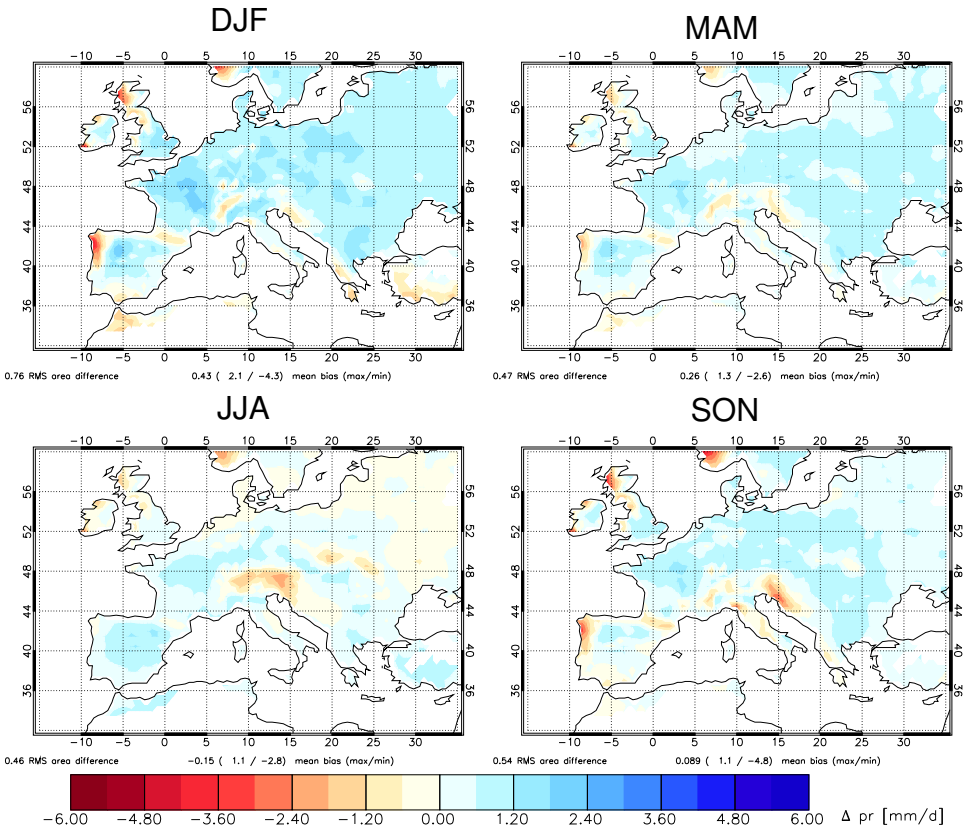


Figure 3.4: Seasonal bias for *pr* in the multi model mean field.

3.1.3 GCM Evaluation with Taylor Plots

Taylor plots are a tool to evaluate the correlation, variability and RMS error characteristics of spatial distributed data. An explanation of the theory behind a Taylor plot is given in section 2.5.2 or in the work of Taylor [2000].

In figure 3.5 a Taylor diagram for *tas*, *pr* and *psl* is depicted. The annual climatology is plotted for both domain sizes. Especially in the large domain (left side), three clusters corresponding to the three different parameters (different colors in the plot) can be seen.

The sea level pressure (yellow symbols) has nearly a perfect spatial correlation, a RMS close to zero and the simulated variability is very close to the spatial variability of the ERA40 reference data. The models are clustering and there are no outliers in both domains. The GCM performance in the large domain is slightly better than the performance in the small domain.

Differences between the two domains are especially visible in the simulated surface air temperature. In the large domain most GCMs have a spatial correlation coefficient (R) ≥ 0.90. In the small domain the average R is about 0.75. A major impact on the increase of the error on the small

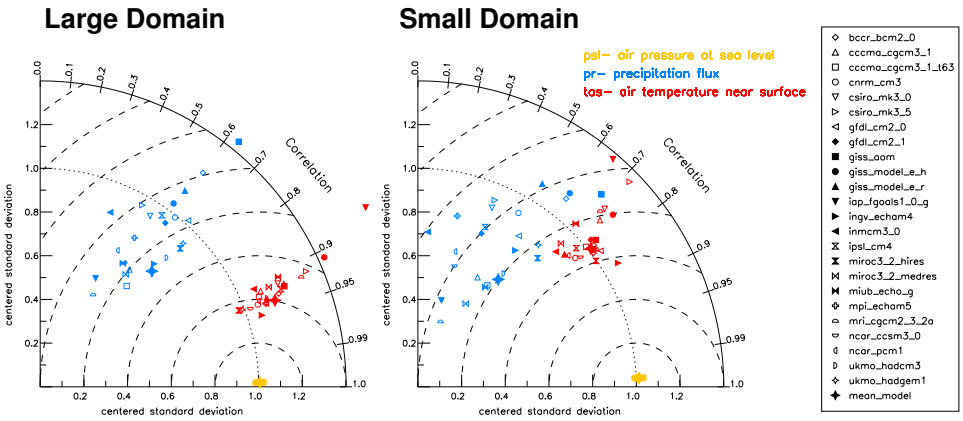


Figure 3.5: Taylor diagram for *tas* (red symbols), *pr* (blue symbols) and *psl* (yellow symbols) for the annually climatology in the large and small domain. Displayed are the spatial RMS error, the spatial correlation coefficient and the centered spatial variability.

domain is that the ratio of area with complex orography is higher than it is in the large one.

Like for *psl*, the multi model ensemble of *tas* is clustering very well but there are some outliers. Especially the iap_fgoals1_0_g model shows a weak spatial correlation and a far to high variability. An overestimated variability of the *tas* model data is a general problem in the large domain. This behavior improves in the small domain, where many GCMs perform close to the reference variability.

The spatial correlation of the precipitation field is hard to simulate and GCMs are performing very differently. The correlation coefficient lies between 0.70 to 0.38 in the large domain and 0.70 to 0.10 in the small domain. Most of the models are showing an underestimation of spatial variability and a very high RMS error. Nevertheless, there are some models like the ukmo_hadgem1 or the miroc3_2_hires GCM, which are performing well in both domains.

In figure 3.6 the parameters *ta*, *zg*, *hus*, *ua*, *va* and *wap* are displayed in Taylor diagrams for the four different pressure levels 850 hPa, 700 hPa, 500 hPa and 300 hPa. The parameters *ta* and *zg* are clustering well on every pressure level and there is no clear height dependency in GCM performance for these variables. Remarkable is the only outlier in the *ta* field at 850 hPa, which is (as on the ground level in figure 3.5) the iap_fgoals1_0_g GCM but its performance is improving in lower pressure levels. Specific humidity *hus* (yellow symbols) is captured well at 850 and 700 hPa. However the majority of simulations are getting worse with increasing altitude.

The quality of air motion parameter for *ua*, *va* and *wap* is very different in the GCM ensemble. Anyway, the spread is increasing with decreasing

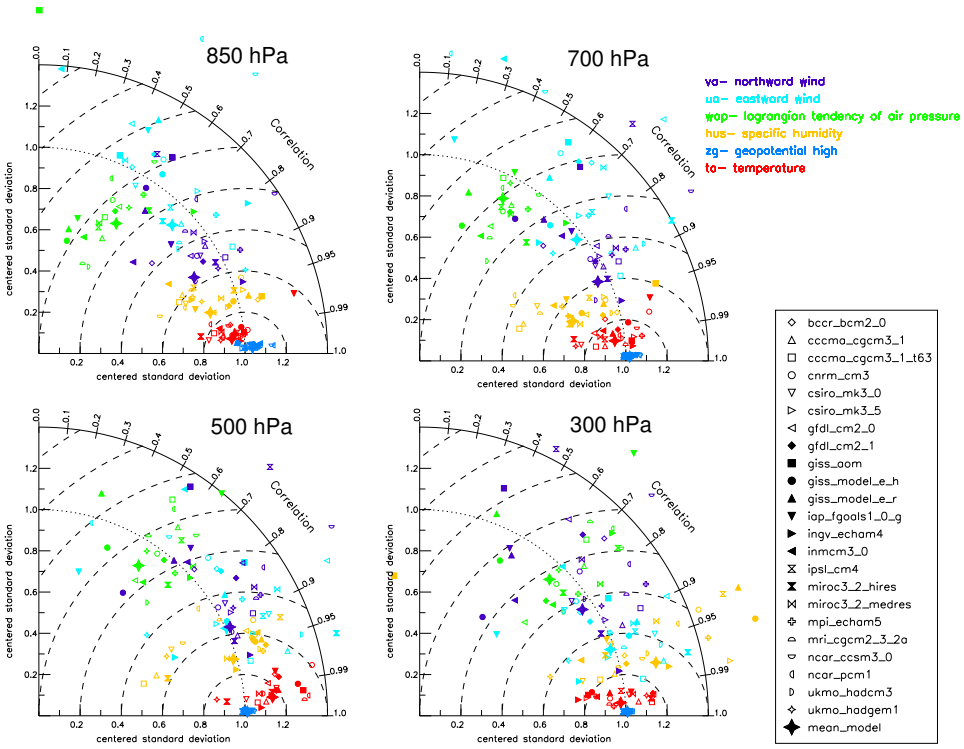


Figure 3.6: Taylor diagrams for *ta* (red), *hus* (yellow), *zg* (blue), *ua* (light blue), *va* (violet) and *wap* (green) in the pressure levels 850 hPa, 700 hPa, 500 hPa and 300 hPa for the annual climatology in the large domain.

air pressure. The northward (*va*) and eastward (*ua*) components of the wind vector are approximately equally good simulated. However, for *va*, the mean performance (big star) is decreasing with increasing height. This is exactly inverse to the eastward wind performance, which gets better in lower pressure levels.

The vertical air motion (parameter *wap*) is difficult to simulate. However, the quality of the reference data might also be worse as it is for other parameters because vertical air movement in the free atmosphere is hard to observe. Nevertheless, there are also models like the ingv_echam4 or the miroc3_2_hires GCM, which simulate the reference data well on different pressure levels.

Additionally to the displayed plots, Taylor plots for different seasons, domains and parameters are provided and can be found in the digital appendix.

3.1.4 The Model Performance Index

In the former section Taylor plots were shown, in which the performance of different GCMs and parameters can be displayed in a very compact and

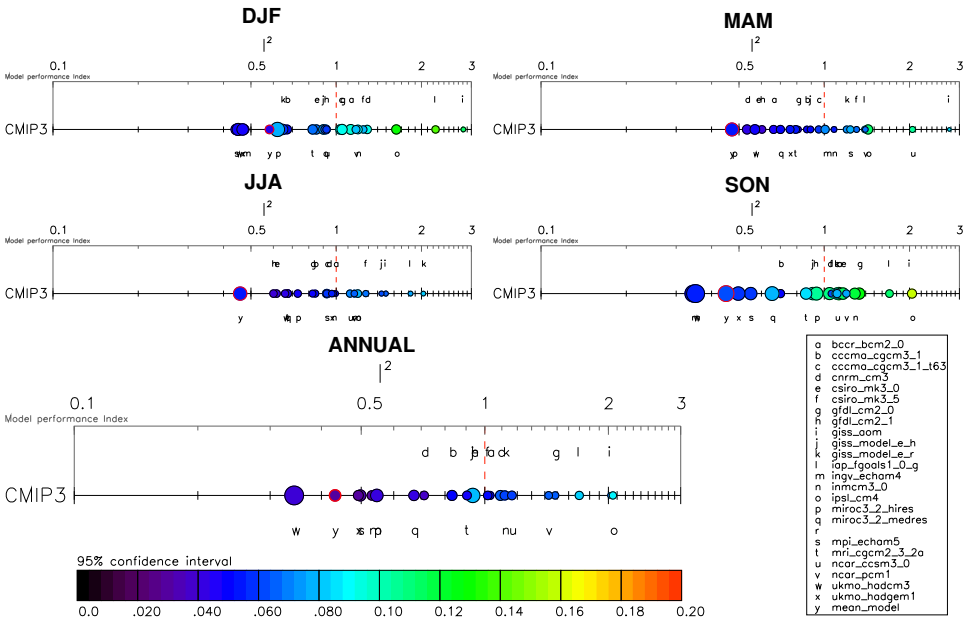


Figure 3.7: The MPI for the large domain, annually and for all seasons. Each circle is corresponding to a GCM from the legend. The colors and sizes of the circles are showing the uncertainty range (95 % confidence interval) of the single MPIs. The x-axis is scaled logarithmically so that differences in small MPIs are better visible. The circle with the red edging shows the multi model mean performance (model y).

comparable way. It is easy to search for well performing GCMs for particular parameters and seasons since there are plots available for different seasons, domains and parameters, but it is hard to find a clear answer to the question, which GCM performs best on average. This is essential for the usage of data from GCMs as driving data for regional climate models. The calculation of a MPI can give a guideline for this decision.

In this section the results from the GCM evaluation, using the model performance index of Reichler and Kim [2008a] are shown. The theoretical background was described in section 2.5.

In figure 3.7 the MPIs of the large domain annually and for all seasons are displayed. In the annual evaluation the Hadley Center model ukmo_hadcm3 is best performing with a MPI of 0.34. However, this model never performs best on a seasonal basis, but it ranks always under the top three GCMs. On the annual scale it is even better than the multi model mean.

On the second place, with an MPI of 0.49, is the other Hadley Center Model ukmo_hadgem1 closely followed by the mpi_echam5 model from the Max Planck Institute (MPI=0.50). The echam5 GCM has the best performance of all models in the winter season, but it has a substandard performance of 1.23 in spring.

The fourth and fifth best performance indices show the ingv_echam4

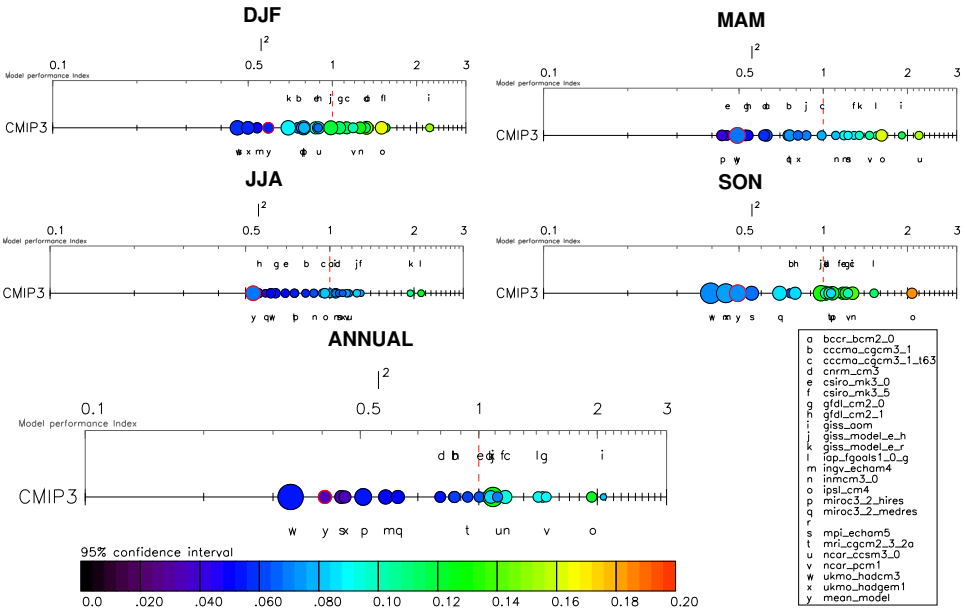


Figure 3.8: Same as figure 3.7, but for the small domain.

(MPI=0.52) and the miroc3_2_hires (MPI=0.53) GCM.

The same model performance calculations as for the large domain were made for the small one. In figure 3.8 the mean model performance indices for all seasons in the small domain are displayed. The MPI values are changing only slightly compared to those of the large domain and the ranking of the models are almost constant.

Figure 3.9 shows the performance indices for the parameters *ta*, *hus*, *zg*, *ua*, *va* and *wap* in the large domain at the 850 hPa level in the annual mean. Assuming that the weight factors w_n and the inter-annual variability $\sigma^2_{\nu n}$ in formula 2.6 are approximately constant, a rough estimation of the correspondent RMS error value of different performance indices is possible. Those values are written on an axis above the performance index axis. This method allows an estimation of how good models can simulate different parameters.

The evaluation of the parameters shows a wide spread of model performance. However, there is not a single model which is outperforming the others in all fields, but there are a few, which are performing above average universally. Of course, these are the GCMs which are well performing in the average over all parameters too. Nevertheless, it is important to take a look at the single parameter indices. For example the ipsl_cm4 model which is in the top five performing GCMs in the average, shows a substandard performance in the temperature and geopotential field.

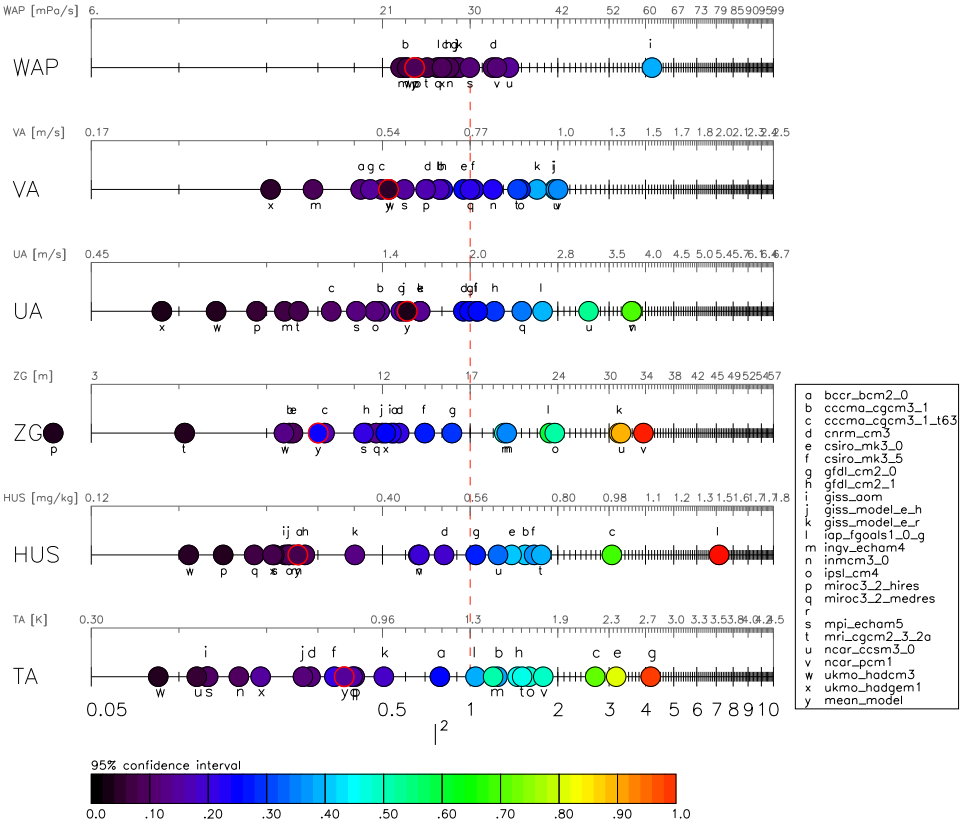


Figure 3.9: The performance indices for *ta*, *hus*, *zg*, *ua*, *va* and *wap* in the large domain, annually on 850 hPa. For reasons of visibility the radii of the circles are constant. The uncertainty of the values is displayed as filling color.

In figure 3.9 the likelihood ranges of the different MPIs are displayed with colors (the radii keept constant) where blue means a small range and red means a wide range of uncertainty. It is clearly visible, that the uncertainties of the indices are increasing with increasing MPI. If light colored circles are close together, it is statistically not possible to know which model is performing better.

This kind of plots are providing information on a parameter basis (annually and for different seasons) and are helpful to choose capable driving data for regional climate models and to interpret the results of RCM simulations.

Additional plots for different domains, seasons, parameters and pressure levels can be found in the digital appendix.

As discussed in section 2.5, the ranking of GCMs due to the model performance index is dependent on the regarded parameters, seasons and domains. This is displayed in figure 3.10, where all MPIs are displayed for both domains, annually and on seasonal basis. The rank of the GCMs on the x-axis are ordered corresponding to their MPI on the large domain annually

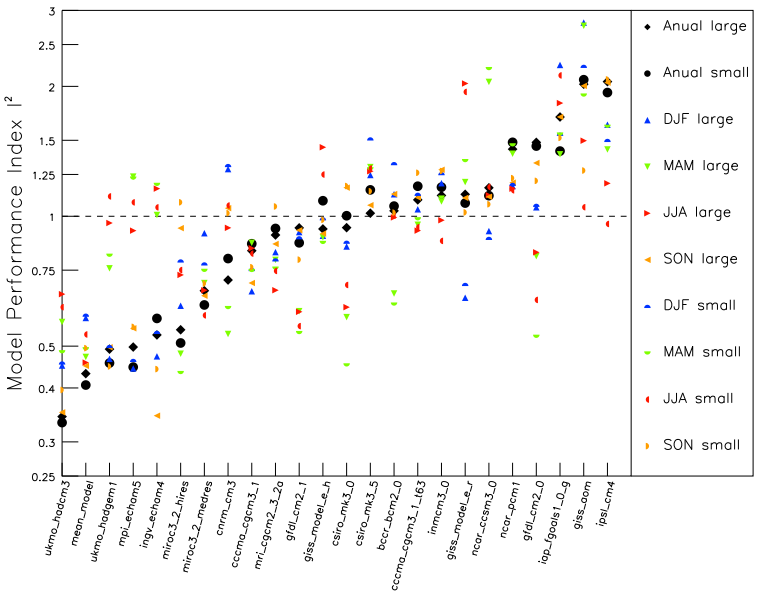


Figure 3.10: Overview of the MPIs for different GCMs, domains and seasons.

(black diamond symbols). On the y-axis the MPI is plotted logarithmically. Different seasons are drawn with different colors and different domains have different symbols corresponding to the legend on the right.

It is visible that equally colored symbols are close together and different colored are more spreaded. This means that the changing of season influences the MPI result stronger than the changing of the domain. The average variety of the MPIs of single GCMs is large. The mean model has the most compact MPI spectra and is universally well performing (its worst MPI is 0.6 in winter in the small domain). Also the ukmo_hadcm3 GCM has a low variance in its spatial and temporal performance indices and a very good average MPI.

Noticeable is the similar performance of the ukmo_hadgem1 and the mpi_echam5 GCM. Both are top performing annually in winter and in autumn, but in spring and summer they are performing only averagely. Even the ipsl_cm4 GCM, which has the worst performance on an annual time scale (on the large domain), has a better performance in summer on the small domain than the mpi_echam5 and the ucmo_hadgem1 model.

The distribution of the 27 parameters (different symbols) for the large domain annually is shown in figure 3.11. The averages over these parameters are marked with black diamond symbols in figure 3.10 and in figure 3.11 the black line passes these points. The y-axis is logarithmically scaled with a

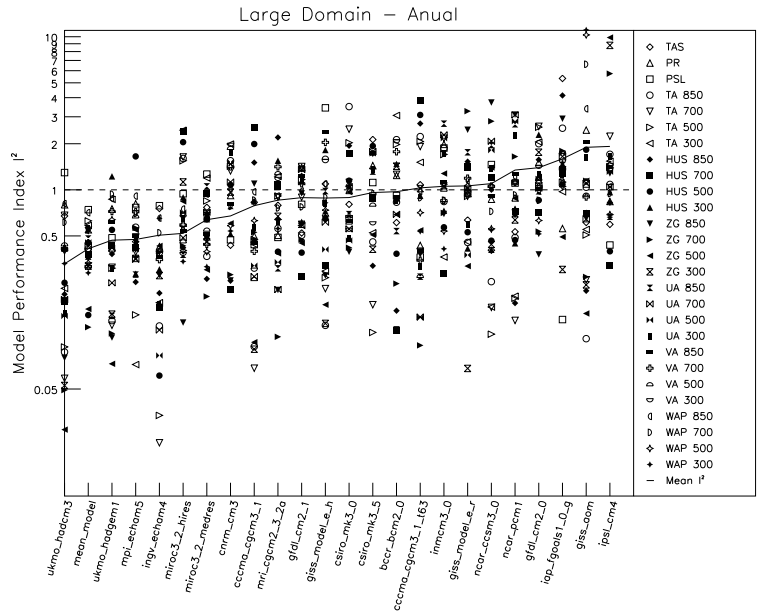


Figure 3.11: Performance indices for all GCMs and parameters on the large domain annually.

maximum performance index of 10.

Graphic 3.11 shows that a good average MPI of a certain GCM does not mean that all parameters in the model are well simulated. For instance, the sea level pressure in the ukmo_hadcm3 model has a performance index of 1.4, which is larger than the performance index of the ipsl_cm4. However, it should be kept in mind that these parameter based performance indices have a likelihood range (shown in figure 3.9 and discussed in section 2.5.1) and a significant distinction of two similar indices is often impossible.

For different domains and seasons additional plots (like figure 3.11) are attached in the digital appendix.

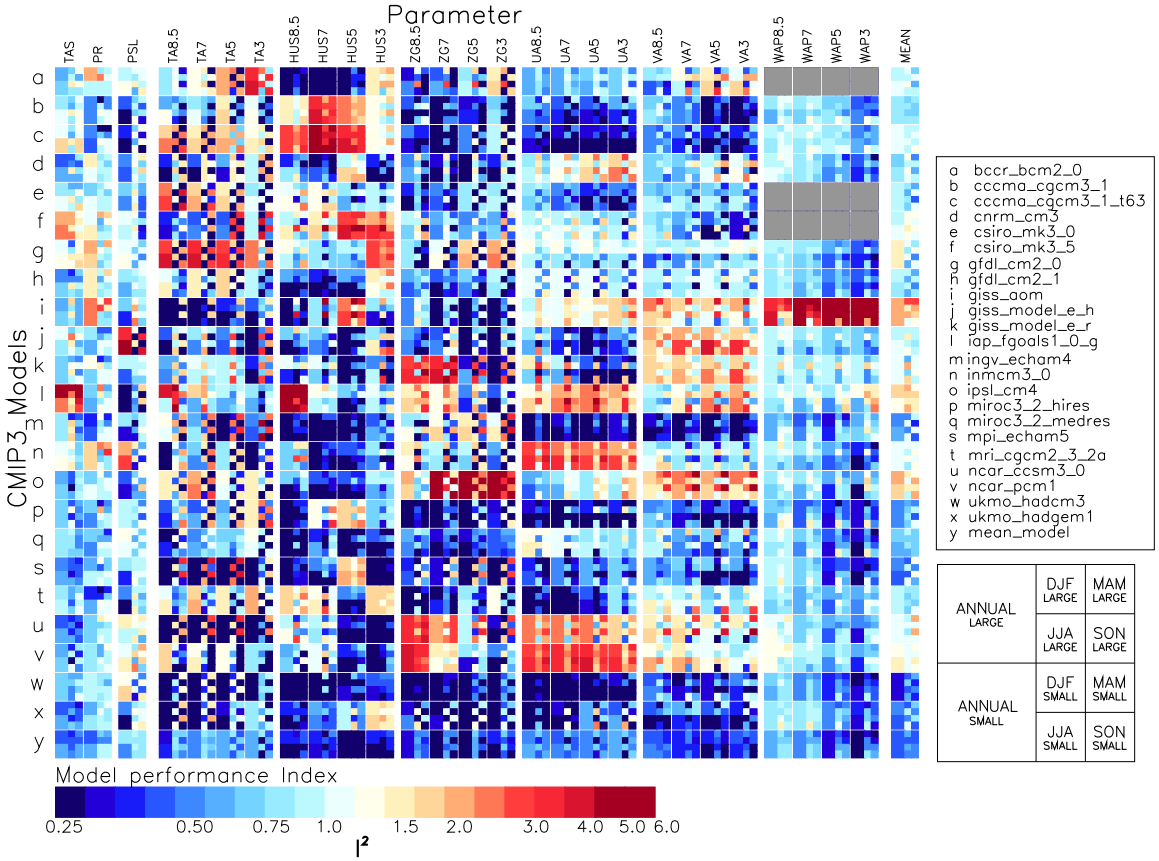


Figure 3.12: Overview of the model performance indices for all variables, domains and seasons. Different GCMs are plotted in the rows whereas different parameters are plotted in the columns. The outer right column shows the mean performance over all parameters. Blue colors are corresponding to good MPIs, whereas red colors represent bad performing models. The partition of the squares is explained in the lower right corner.

Even if figure 3.10 and 3.11 have a high information density it is still hard to get an good overview of the overall performance of GCMs in different domains, for different parameters and different seasons. Therefore, in figure 3.12 the MPIs of all GCMs, parameters, seasons and domains are plotted. The columns represent different parameters and the rows represent different GCMs (so, each square represents a parameter of a GCM). Each square is separated into 10 regions where the upper half of the square represents the large domain and the lower half of the square represents the small domain. The big sub squares on the left side show the performance indices of the annual climatology whereas the small sub squares on the right are corresponding to the four seasons starting on the left top with DJF and going to the right bottom over MAM and JJA to SON. A legend corresponding to the division of the single squares is drawn in the right bottom of figure 3.12.

The color of every sub square shows the performance index of a certain model-parameter-domain-season combination. Corresponding to the logarithmic color bar on the bottom of figure 3.12, a dark blue color means a good model performance whereas a red color means a bad performance.

The last column on the right shows the mean over all parameters. This is the MPI for each model, season and domain. The last row, which is marked with an y shows the performance of the mean model (multi model mean).

Looking at the model performance indices in the last right column, the five best performing models become visible because of their dark blue shade. The darkest shade of blue has the ukmo_hadcm3 model (line w). It has a superior MPI in every season and domain. Also on the parameter level, there are nearly no weaknesses visible. Only the sea level pressure and the air temperature *ta* in spring and summer show performance indices, which are greater than one.

As shown in figure 3.10 and discussed in section 1.2.3, the multi model mean (row y) has an outstanding performance. Four of the five best performing GCMs (ukmo_hadcm3 model, ukmo_hadgem1, mpi_echam5 and ingv_echam4) have their worst MPI results in summer and spring. Only the miroc3_2_hires does not show this feature.

Looking on the other side of the spectra, there is no single model, which performs bad in every observed field. For instance, the giss_aom GCM (line i) has a good performance in air temperature *ta* and in geopotential height *zg*.

Figure 3.12 gives a very good overview of the weaknesses and strengths of the different GCMs. In combination with Taylor plots, maps of biases and seasonal cycle analyses a widespread overview of the ability of CMIP3 GCMs to simulate the past climatologically mean state of the atmosphere in the Alpine region is given. These results are the basis for the analyses of future model projections for the 21st century, which is the focus of the next section.

3.2 Analysis of the Climate Change Signal in the 21st Century

With the knowledge of the skill of CMIP3 GCMs to simulate the past, a glimpse into the future is ventured now. In the following chapter the spatial mean temporal development of every single parameter will be examined and discussed for the emission scenario B1, A1B and A2. For the period of 1960 to 2000 the results of the 20C3M simulations are taken. After 2000 there are three datasets corresponding to the three emission scenarios. In the plots the uncertainties (simple standard deviation) due to different GCMs are gray shaded in the background and framed with the corresponding color of the emission scenario. The multi model mean for each emission scenario (average over all models) is drawn as a thick line. The gray bars on the right show the climatological mean and standard deviation of the 2071 to 2100 climate period. Offsets in the contact point between the 20C3M and the scenario simulations in the year 2000 can occur because of the different amounts of GCM runs in the different scenario experiments. All displayed lines are 30 year running means.

After these evaluations, a statistical check for significant changes in the climate mean state of the parameters between the reference period 1971 to 2000 and the two future climatologies 2021 to 2050 and 2071 to 2100 is done. Furthermore, the seasonal variances are checked to obtain an estimation of the future climate variability.

3.2.1 Temperature at Surface (*tas*)

In figure 3.13 the temporal development of *tas* in the large domain is displayed. In the top left picture the 20C3M and the emission scenario experiments are depicted. Until 2040 the three scenarios have a very similar trend and the A1B scenario shows the highest temperature increase until 2060. Even in 2100 it is statistically not possible to distinguish between the different emission scenarios because of the wide spread of standard deviation in the multi model field. At least the B1 and the A2 scenarios are showing statistically different results.

In the three other pictures of figure 3.13 the three emission scenarios with the course of each GCM are shown. Not every emission scenario has the same amount of GCMs (see also figure 2.1). The multi model standard deviation increases with time in all scenarios. This means that the forecast accuracy is decreasing in the long term.

Each of the 24 GCMs has a different strong climate sensitivity. For the two meter surface temperature, high climate sensitive models are, for instance, the miroc_hires & medres GCM and a low climate sensitive model is the csiro_mk3_0 GCM.

Overall, the mean surface temperature rise in the large domain between

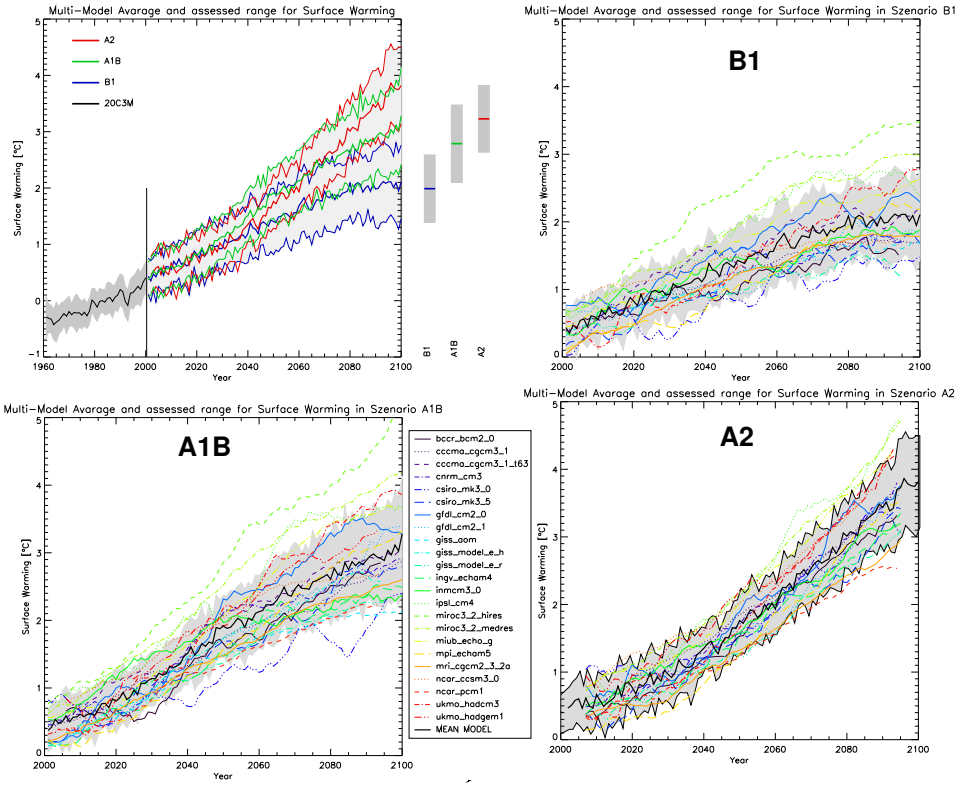


Figure 3.13: Temporal development of two meter surface temperature *tas* from 1960 to 2100.

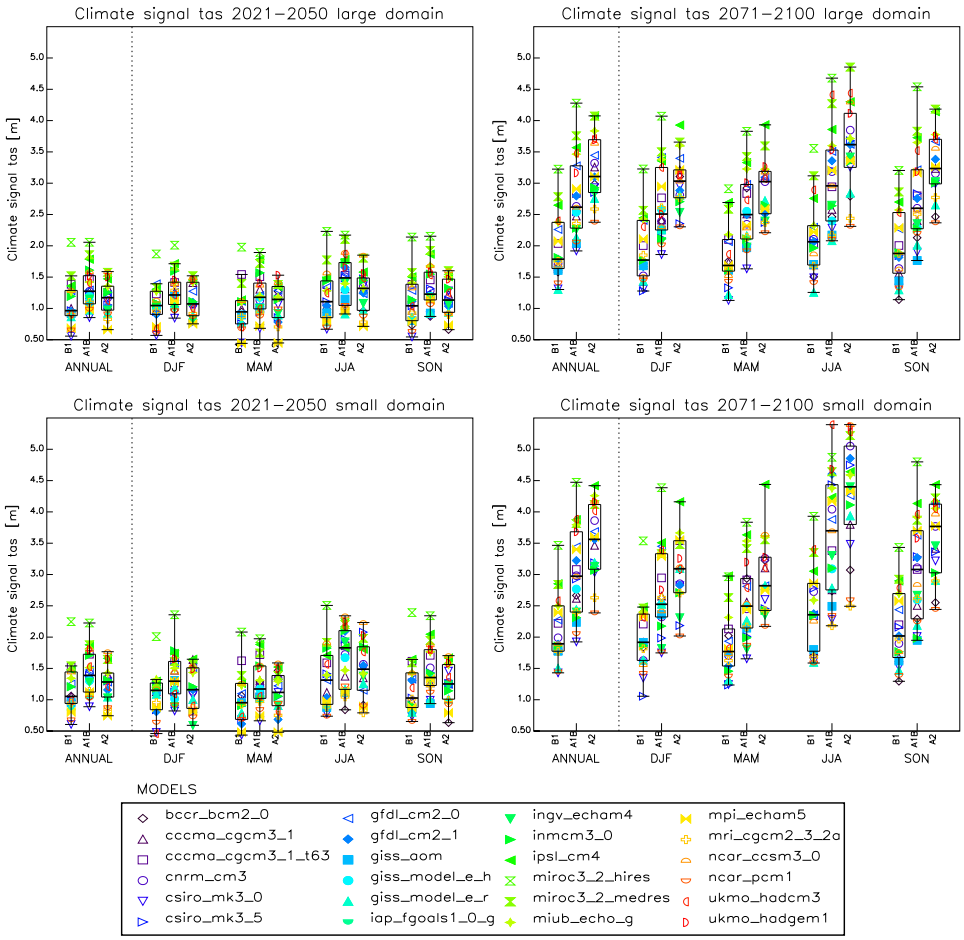


Figure 3.14: Box plots of surface temperature *tas* for all domains, climatologies, seasons and experiments.

1971 to 2000 and 2071 to 2100 is (2.0±0.6) K for B1, (2.8±0.7) K for A1B and (3.2±0.6) K for the A2 scenario. In the small domain the standard deviation and the ensemble mean temperatures are higher. In scenario B1 the rise in *tas* is (2.1±0.65) K, for A1B (3.1±0.8) K and for the A2 scenario (3.5±0.7) K.

Figure 3.14 shows the climate change signals for *tas* in both domains, both climatologies (ref. 1971 to 2000 and 2021 to 2050 and 2071 to 2100), for all experiments, annually and for all seasons in a box-whisker plot (see section 2.6). The climatology of 2021 to 2050 shows only low temperature rises and nearly no differences between the emission scenarios compared to the 2071 to2100 climatology. In the later climatology a clear diversification of the different experiments is visible, whereby the climate change signal increases with increasing greenhouse gas emissions. The most climate sensitive season for *tas* in both domains and in all experiments is summer.

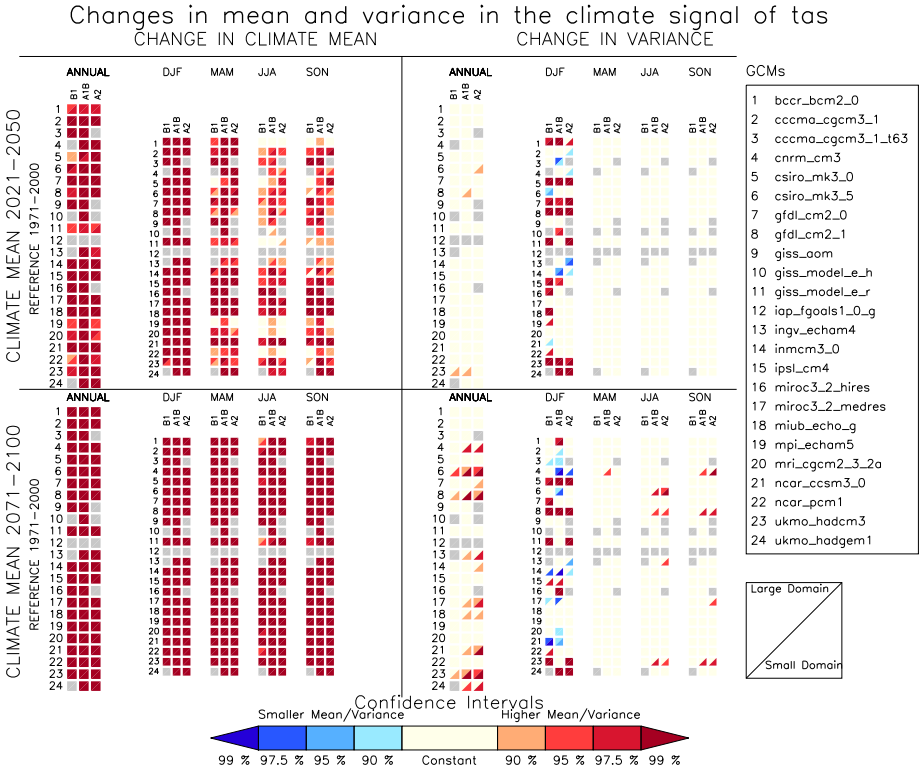


Figure 3.15: Significance analyses of the climate change signal (left side) and changes in inter-seasonal variability (right side) for *tas*. The upper side of the plot shows the climate change signals for 2021 to 2050 and the lower part those for 2071 to 2100. Different bars are showing the annual mean climatologies and different seasonal climatologies. Each bar is separated in boxes, where the three columns are presenting the three emission scenarios B1, A1B and A2 and the 24 rows show different GCMs. Each box is divided into two triangles where the upper one shows the significance of the climate change signal of the large domain and the lower one the significance of the small domain corresponding to the description on the right bottom. Red colors are indicating a significant higher mean or monthly variance in *tas* and blue colors are showing a significant decrease in mean or monthly variance of *tas*, respectively.

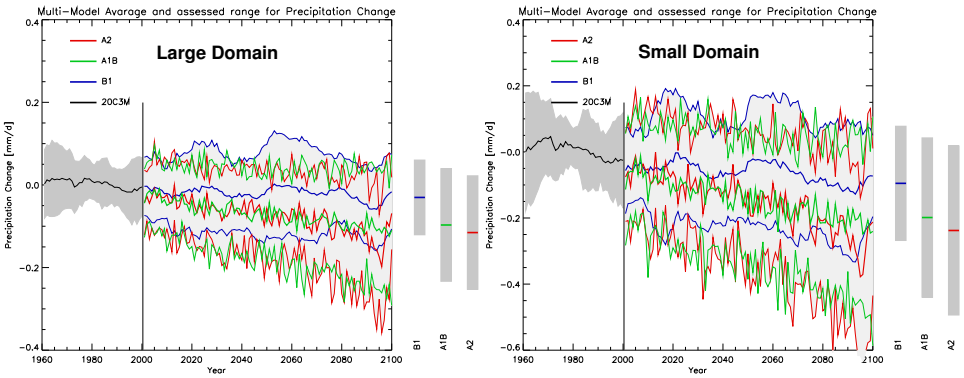


Figure 3.16: Temporal development of the spatial mean monthly precipitation in scenario B1, A1B and B2 on the large and small domain.

The significance of the climate change signal and the changes in the seasonal variance for *tas* are depicted in figure 3.15. Most GCMs show no changes in the seasonal *tas* variance between the two future climatologies and the reference period. Only in the winter season approximately half of the models are projecting significant changes, but they disagree in sign.

The mean changes in the *tas* climate change signal are showing a statistical significant increase in seasonal mean surface temperature. Even in the 2021 to 2050 climatology the changes in temperature are highly significant. Especially the winter period shows a consistent result for all GCMs and emission scenarios.

3.2.2 Precipitation Flux (*pr*)

Figure 3.16 shows spatial mean changes in precipitation for the three emission scenarios and both domains plotted as thirty year running mean. Compared to the surface temperature the changes of precipitation are more difficult to interpret. Especially the A1B and A2 scenarios show a decrease of precipitation in the multi model mean. In the B1 scenario no clear climate change signal is visible. However, the multi model spread is big and even increases in the small domain compared to the large one. Plots with the temporal development of the single GCMs in all emission scenarios are available in the digital appendix.

As visible from figure 3.17 it is not clear if the precipitation is increasing or decreasing in the 21st century. The annual median trend is negative in all scenarios. On the seasonal time basis, a positive median climate change signal in winter and a negative signal in the remaining seasons is detected. The highest climate sensitivity is visible in summer where some GCMs predict a strong drying.

There are no consistent changes in the *pr* mean visible in figure 3.18. However, there are many models with a significant decrease in precipitation,

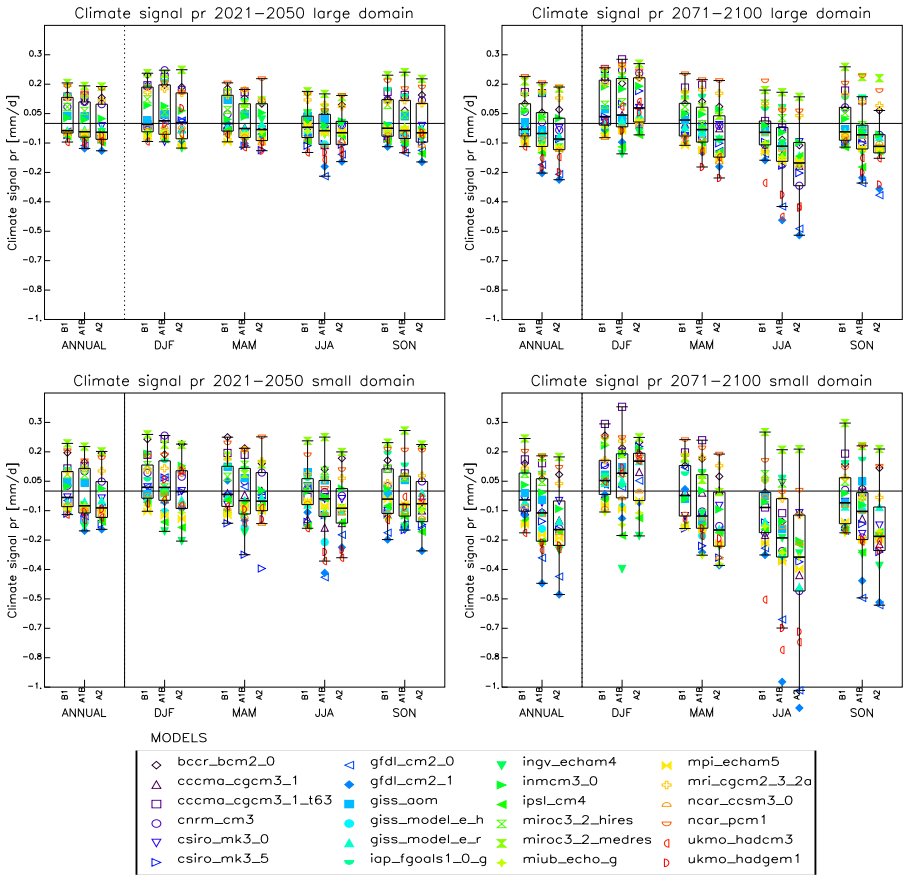


Figure 3.17: Same as figure 3.14 but for precipitation flux pr.

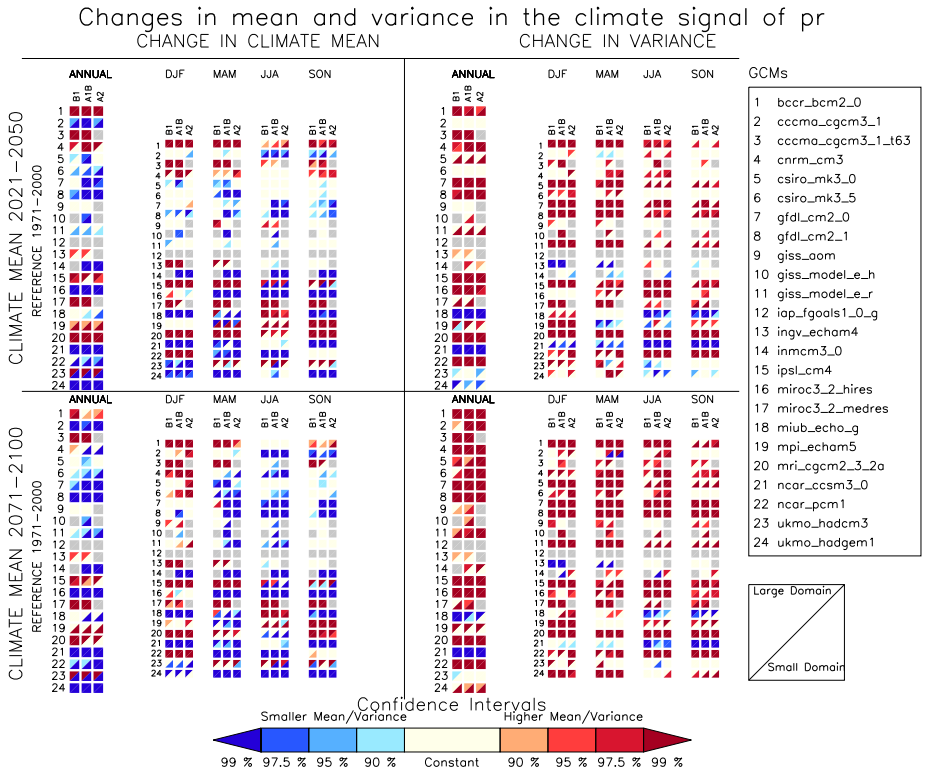


Figure 3.18: Same as figure 3.15 but for pr.

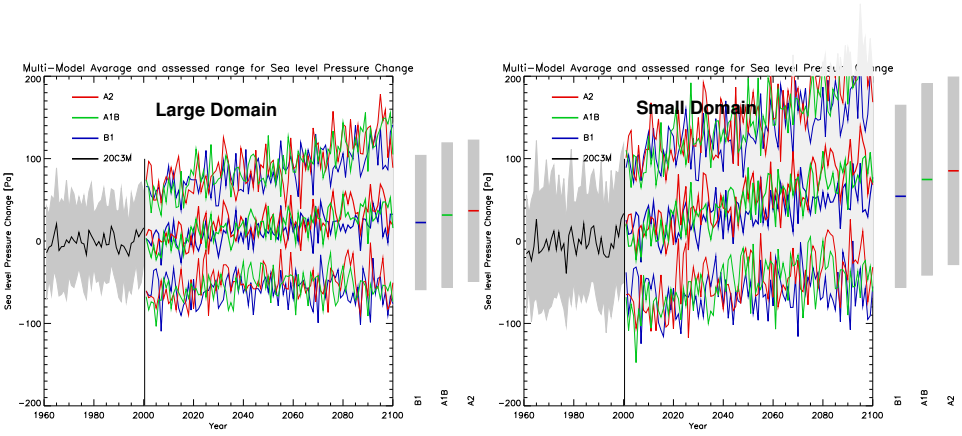


Figure 3.19: Temporal development of the spatial mean yearly sea level pressure in scenario B1, A1B and B2 in the large and small domain.

especially during the summer season at the end of the 21st century. The results of the variance homogeneity tests show a more uniform result. The outnumber of GCMs show an increase of seasonal variability for all seasons.

3.2.3 Pressure at Sea Level (*psl*)

There is a visible increase in the temporal development of psl until 2100, which is stronger in the small domain (see figure 3.19 and 3.20). The emission scenarios show nearly no differences.

As shown in figure 3.20 the GCMs predict a median increase in *psl* during winter, spring and autumn and a decrease in summer. This behavior is already evident in the first half of the 21st century in all emission scenarios. There is no season which could be picked as especially climate sensitive since the amplitudes of the climate change signals are approximately equally high.

Significance testing for differences in the mean *psl* climate change signal and for the homogeneity of variance indicates no clear results in the multi model field (see figure 3.21).

3.2.4 Air Temperature (*ta*)

In figure 3.22 the mean air temperature rise is drawn for 850 hPa and 300 hPa in the large domain. Compared to the surface air temperature, an earlier breakup point of the scenario runs is visible. Furthermore, the temperature rise is increasing in higher levels of the troposphere. For instance, the climate change signal (1971 to 2000 compared to 2071 to 2100) for the A2 scenario is at the 850 hPa level 3.3±0.6 K and at the 300 hPa level 4.4±0.8K. This means a shift in the laps rate in the troposphere due to

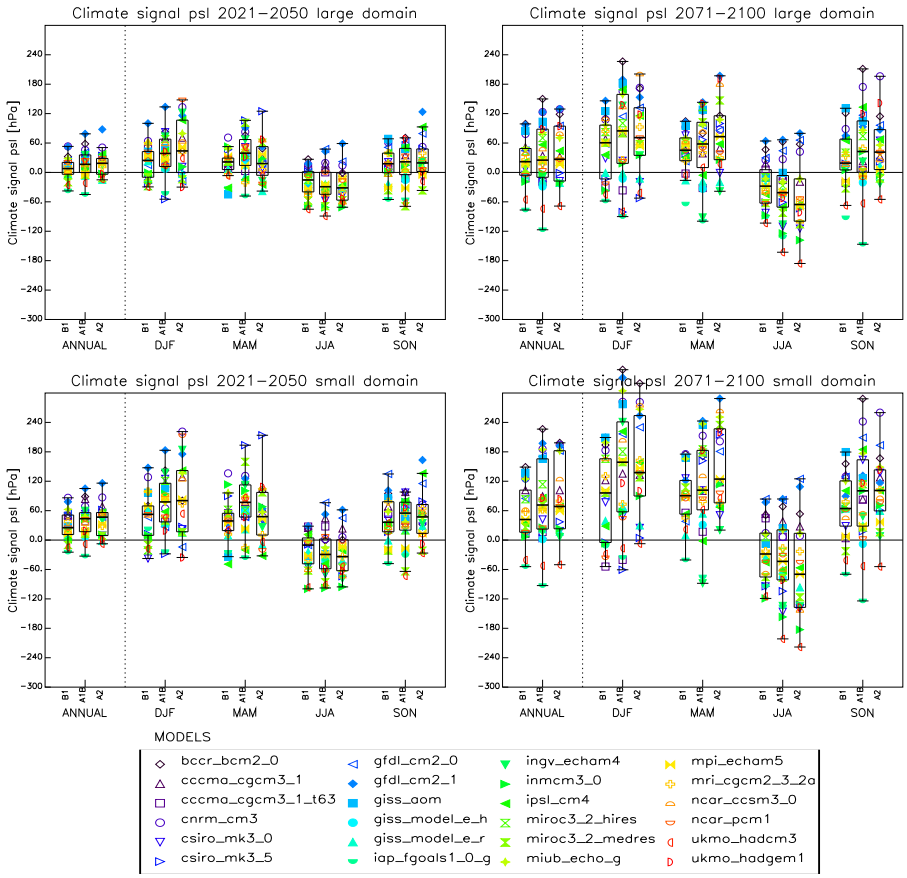


Figure 3.20: Same as figure 3.14 but for sea level pressure *psl*.

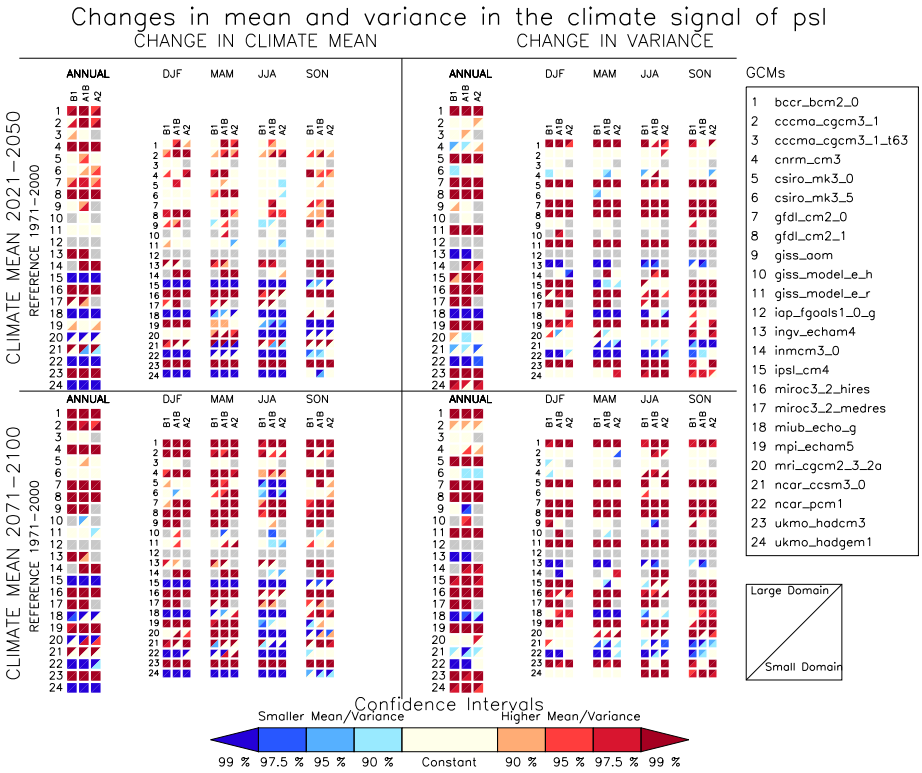


Figure 3.21: Same as figure 3.15 but for pressure at sea level.

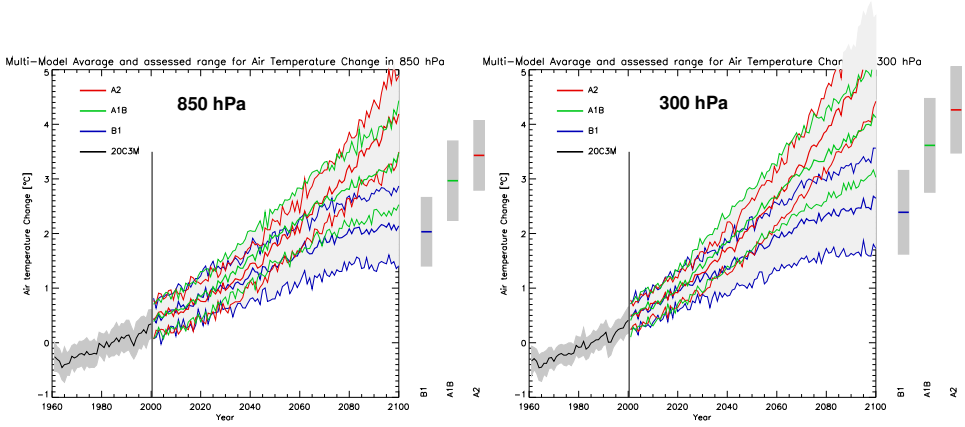


Figure 3.22: Mean air temperature (*ta*) rise at 850 and 300 hPa for the large domain.

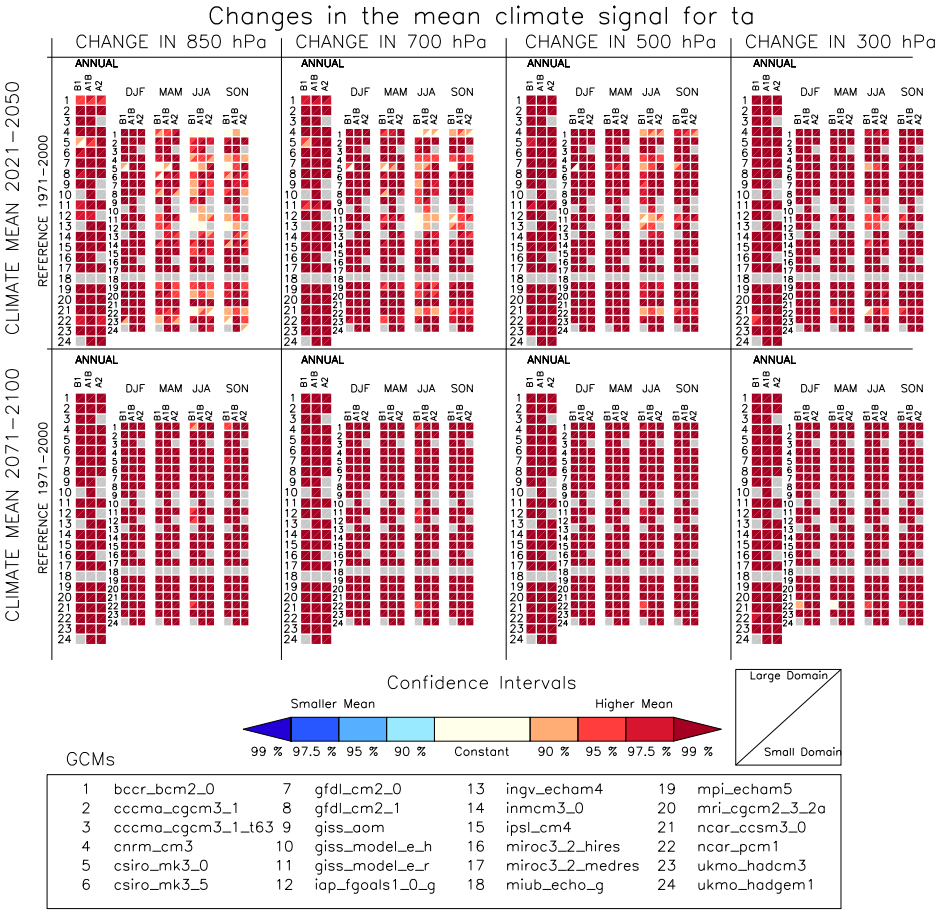


Figure 3.23: Significance in the change of mean air temperature at the four different pressure levels, the three different emission scenarios and the two different climatologies in the small and large domain.

lower vertical temperature gradients. In general this leads to a stabilization of the air layers.

The test for differences in the mean states of air temperature between the reference and the future climatologies shows a high significant temperature rise for all GCMs and all scenarios (see figure 3.23). The significance is increasing with altitude and time and is strongest in the winter season. In general, differences between the emission scenarios are small.

In figure 3.24 the significance of changes in the monthly variance is visible. The outnumber of models show a significant trend towards a stronger variability especially in the winter and spring season, which is increasing in high altitudes. However, in other seasons and at higher pressure levels there are nearly no significant changes.

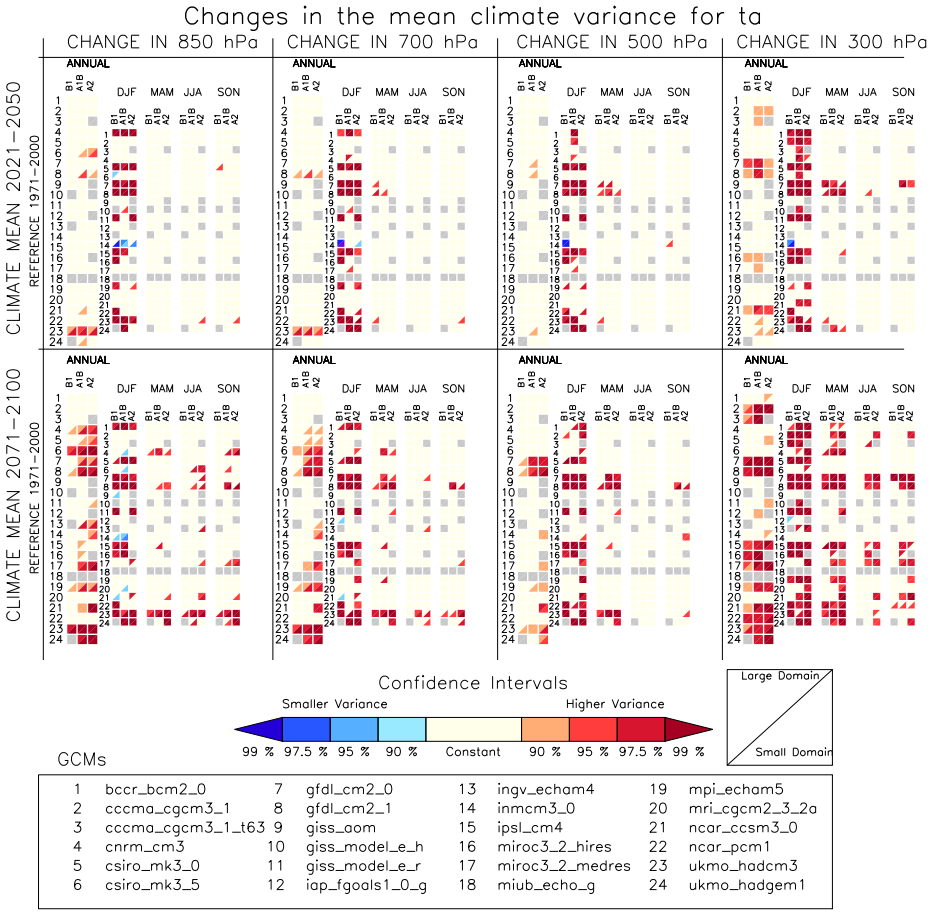


Figure 3.24: Same as figure 3.23 but for changes in monthly variance of *ta*.

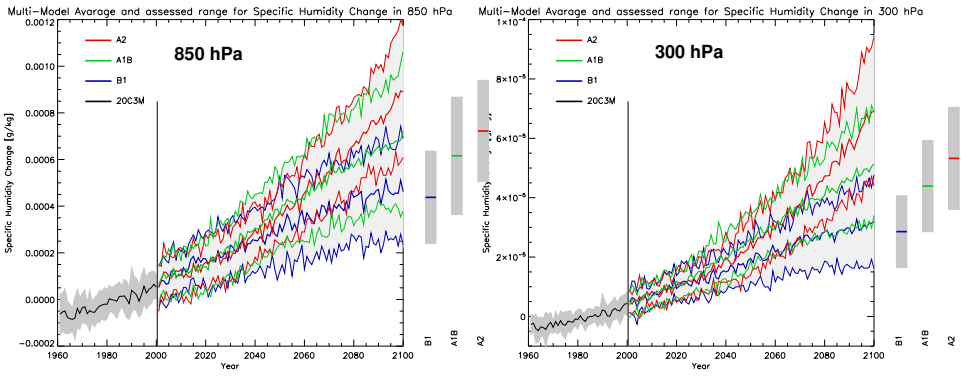


Figure 3.25: Changing of specific humidity in the spatial average of the small domain at the 850 hPa and 300 hPa level.

3.2.5 Specific Humidity (*hus*)

The changing in the mean specific humidity (hus) in the small domain is shown in figure 3.25 for the 850 hPa and 300 hPa level. There is a positive trend of *hus* in all pressure levels with an increasing amount at higher altitudes. The course of the curves look similar to those of the air temperature plots. This is related to the dependency of the specific humidity on air temperature due to formula 2.3. If the air temperature of a certain pressure level is rising, the specific humidity is going to rise as well.

The results for the large domain are looking similar to those for the small one, but they have a slightly lower range of uncertainty. The corresponding plots can be found in the digital appendix.

The box whisker plots, attached in the digital appendix, are showing that the summer is the season with the highest climate sensitivity for *hus*. This correlates well with the analyzed air temperature rise.

Also the significance in the changes of the mean specific humidity looks very similar to the changes in the mean *ta* field (see figure 3.23). The corresponding figure can be found in the digital appendix. Bigger differences between *hus* and *ta* occur in the changes of the seasonal variability. Figure 3.26 shows that there is a significant rise in the seasonal variability of *hus* with increasing altitude, which is stronger than for *ta*. For example, at the 300 hPa level, in the end of the 21st century the whole GCM ensemble is showing a highly significant increase for all emission scenarios.

3.2.6 Geopotential Height (*zg*)

The change in geopotential height during the 21st century in the large domain is displayed in figure 3.27. These plots are looking similar to those of *ta* and *hus*. The increase of the geopotential height at certain pressure levels is related to a vertical extension of the troposphere due to thermal expansion.

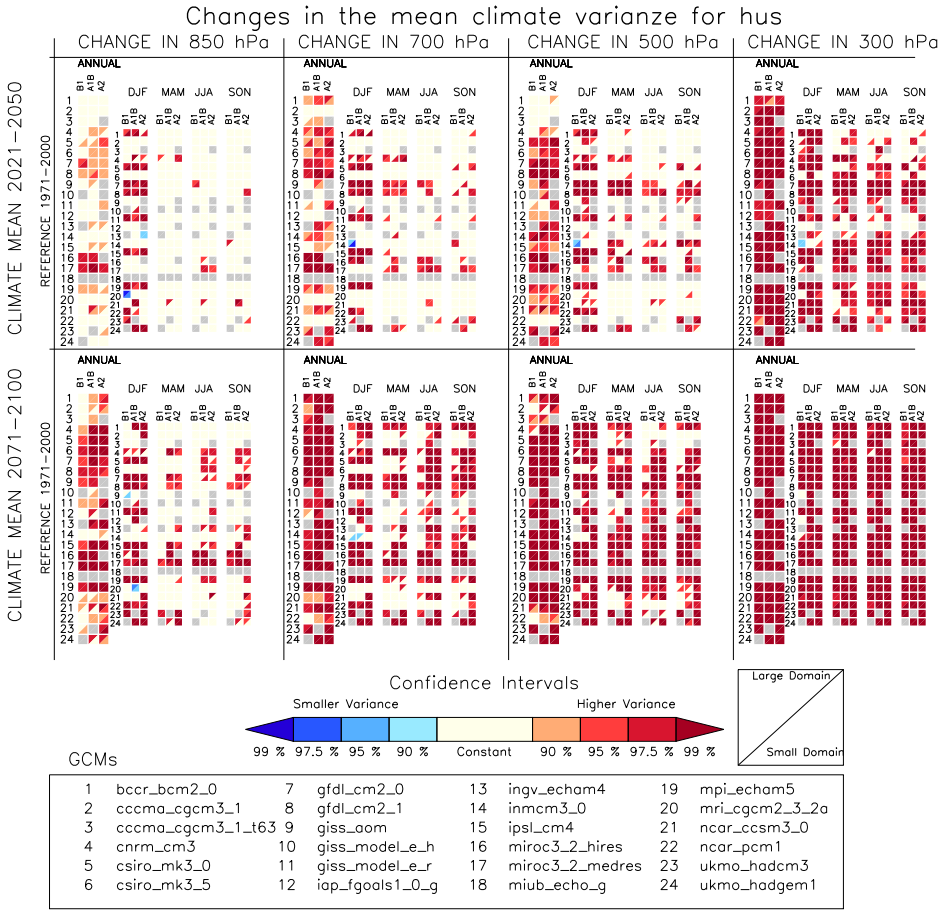


Figure 3.26: Same as figure 3.24 but for *hus*.

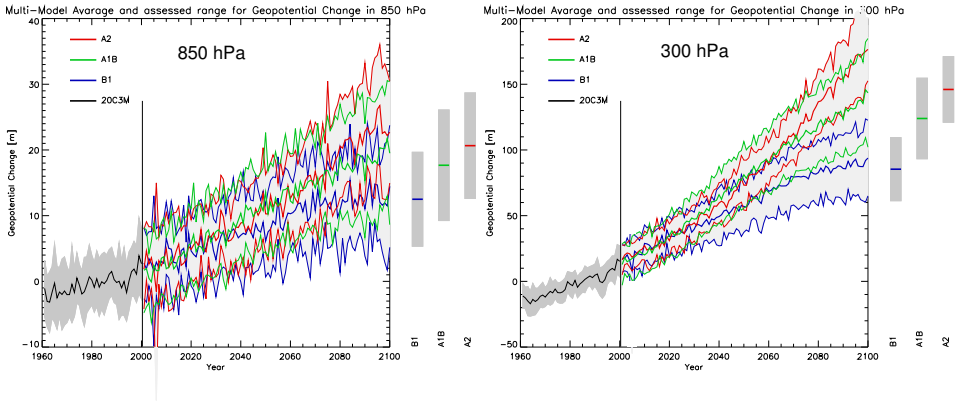


Figure 3.27: Temporal development of the mean geopotential height on the large domain.

The barometric formula describes this circumstance in mathematical terms:

$$p_z = p_0 \cdot \exp^{-\int_0^z \frac{dz}{H_z}}; \quad H_z = \frac{R_G \cdot (ta_0 + \Gamma_a \cdot z)}{\bar{M}_L \cdot g}$$

the integral in the exponent is:

$$\int_0^z dz \frac{1}{H_z} = \int_0^z dz \frac{\bar{M}_L \cdot g}{R_G \cdot (ta_0 + \Gamma_a \cdot z)} = \frac{1}{\Gamma_a} \cdot \ln\left(\frac{ta_0 + \Gamma_a \cdot z}{ta_0}\right)$$

inserted in the hydrostatic equation it is:

$$p_z = p_0 \cdot \exp^{-\frac{\bar{M}_L \cdot g}{R_G \cdot \Gamma_a} \cdot \ln\left(\frac{ta_0 + \Gamma_a \cdot z}{ta_0}\right)}$$

transformed to z, the equation yields:

$$z = \frac{\left(\frac{p_z}{p_0}\right)^{-\frac{R_G \cdot \Gamma_a}{\bar{M}_L \cdot g}} \cdot ta_0 - ta_0}{\Gamma_a} \tag{3.1}$$

In equations 3.1 z is corresponding to the geometric height, which is linked via equation 2.4 to the geopotential height. However, the difference between the geopotential and the geometric height in the first 30 km of the atmosphere is less thatn 0.5 % [Geer, 1996]. M_L is the molar mass of air (ca. $28.964 \cdot 10^{-3}$ kg/mol, g is the gravitational constant, R_G is the universal gas constant ($R_G = 8.3145$ J/K·mol), ta_0 is the temperature at the ground and Γ_a is the temperature gradient.

As mentioned above and visible in figure 3.22 the 21st century temperature rise increases with increasing altitude. So there is a shift in the

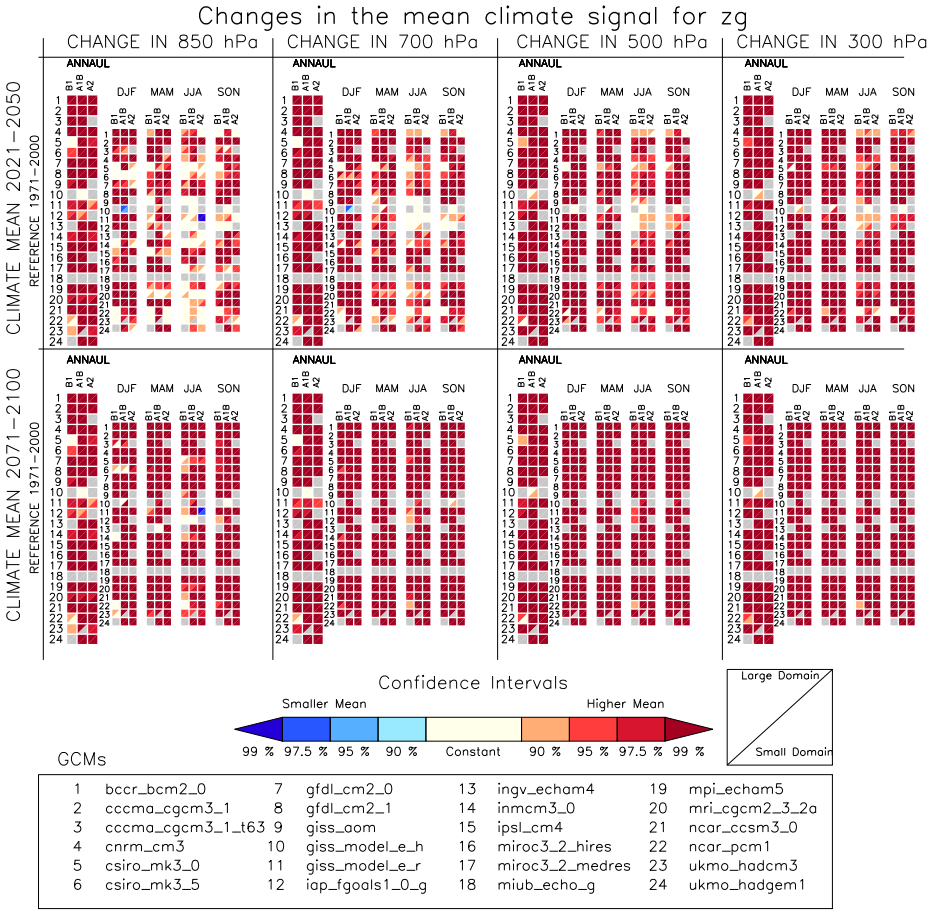


Figure 3.28: Same as figure 3.23 but for zg.

lapse rate of the troposphere due to lower vertical temperature gradients as already mentioned in section 3.4.4. Since the temperature gradient Γ_a is inversely proportional to the gepotential height (see equation 3.1) a decrease of Γ_a leads to an increase of z if the pressure levels p_z and p_0 are staying constant.

Even in the middle of the 21st century, the increasing of geometric height is very significant in the whole GCM ensemble with increasing consistency at higher altitudes (see figure 3.28). The changes in the seasonal variability of zg are very heterogeneous in the ensemble. Many GCMs are showing no changes. Only in the winter season the majority of models are showing an increase in variance. The corresponding plot can be found in the digital appendix.

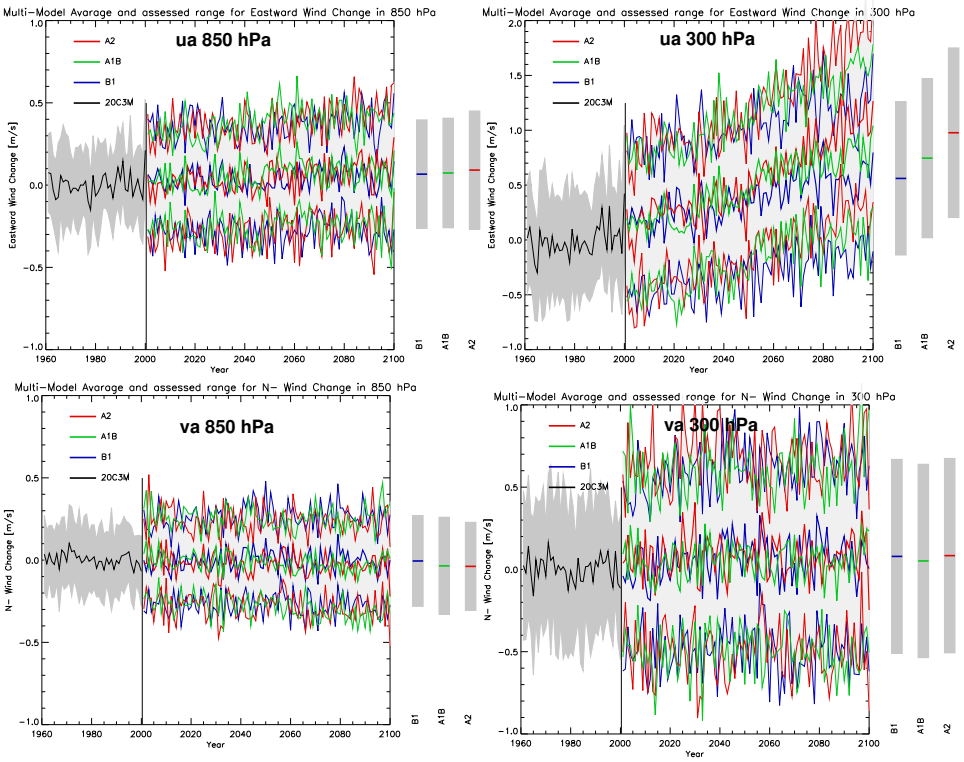


Figure 3.29: Temporal changing in the spatial mean state of eastward *ua* and northward *va* wind in the large domain at 850 hPa and 300 hPa.

3.2.7 North- and Estward Wind (*va*, *ua*)

Figure 3.29 shows the spatial mean development of the horizontal wind parameters eastward *ua* and northward *va* wind in the large domain for the 850 hPa and 300 hPa pressure level. The two pictures for 850 hPa are looking similar, but the northward wind component does not show any changing in the annual mean value with the altitude, whereas the mean eastward wind speed increases in all three scenarios with sinking pressure levels.

Even if there is no visible climate change signal on the annual time scale for the northward wind field, there are differences on the seasonal scale. On the right hand side of figure 3.30 the climate change signal for *va* is visible in a box whisker plot for both domains at the end of the 21st century. There is a significant increase in northward wind in winter and a slight decrease in the other three seasons so that the annual sum is almost zero for all scenarios. Furthermore, in the small domain not only the uncertainty range in the multi model field is increasing, but also the amplitude of the median climate change signals are rising.

On the left hand side of figure 3.30 the climate change signal for the eastward wind component is displayed. The most sensitive season is, as for

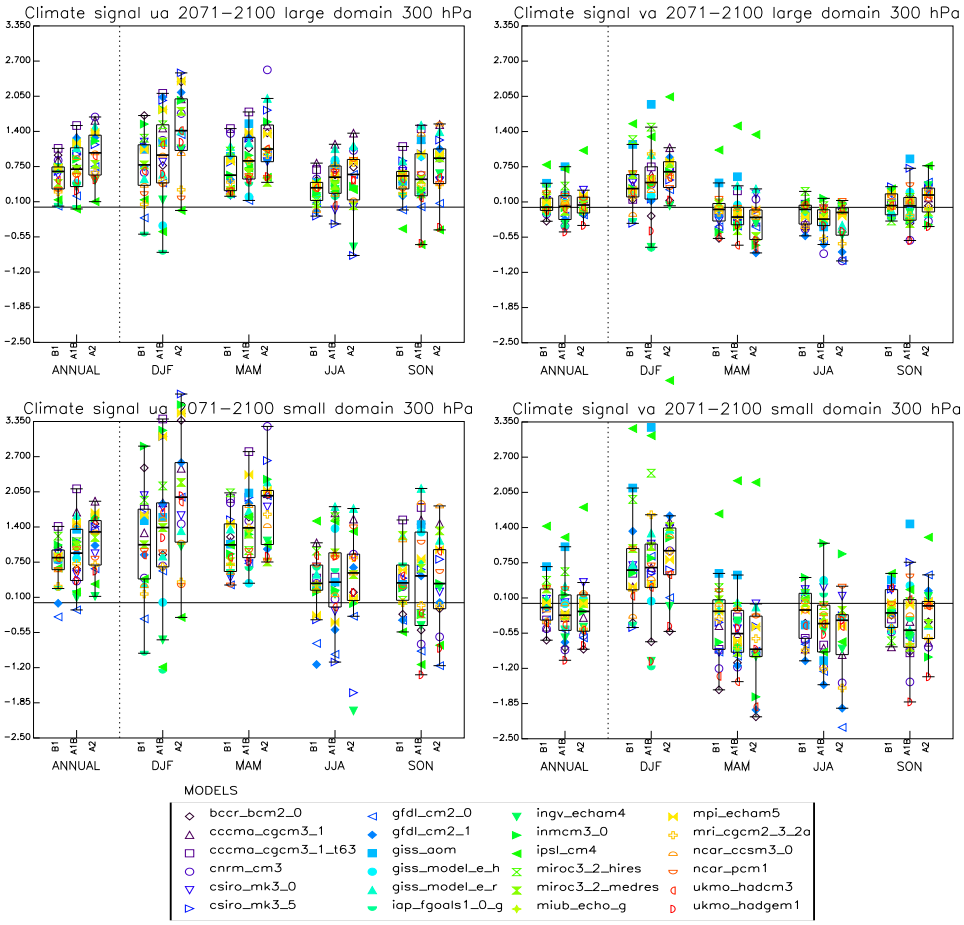


Figure 3.30: Box whisker plots for the eastward (left side) and northward wind velocity (right side) for all domains, seasons and scenarios at the 300 hPa pressure level.

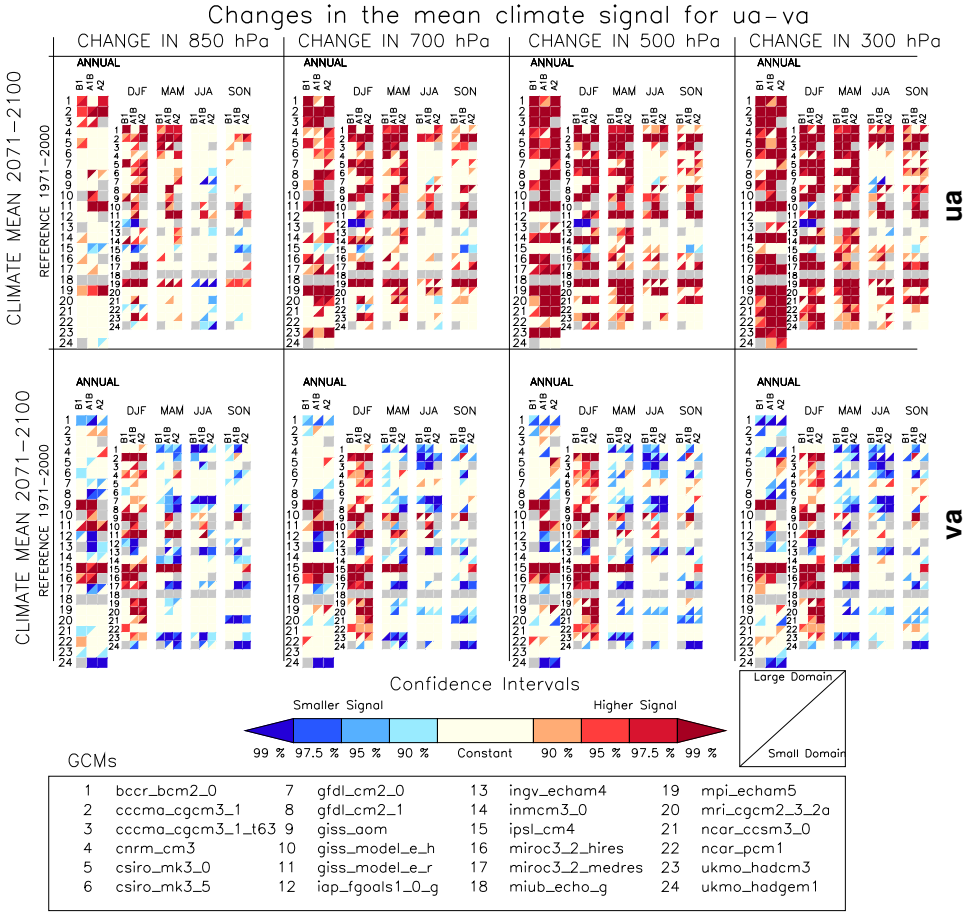


Figure 3.31: Test for significant changes in the mean horizontal wind fields *ua* (upper half) and *va* (lower half) in the end of the 21ct for all pressure levels, domains and scenarios.

va, the winter, but it is also the season with the highest uncertainty. In spring all models show an increasing eastward wind speed and in the annual average there is only one model with a decreasing trend.

At the higher three pressure levels the behavior of the horizontal wind velocity is similar to the 300 hPa level, but the amplitudes are decreasing at lower altitudes. The corresponding plots can be found in the digital appendix.

Figure 3.31 shows the significance of the changes in the horizontal wind fields *ua* and *va* for the climatology 2071 to 2100 compared with the reference period. In the annual analyses of the *va* field no significant trend is visible, whereas the seasonal time scale differences are visible. In winter the outnumber of models show a statistical increase in northward wind speed in all pressure levels. The ensemble signals in the other seasons are not so consistent.

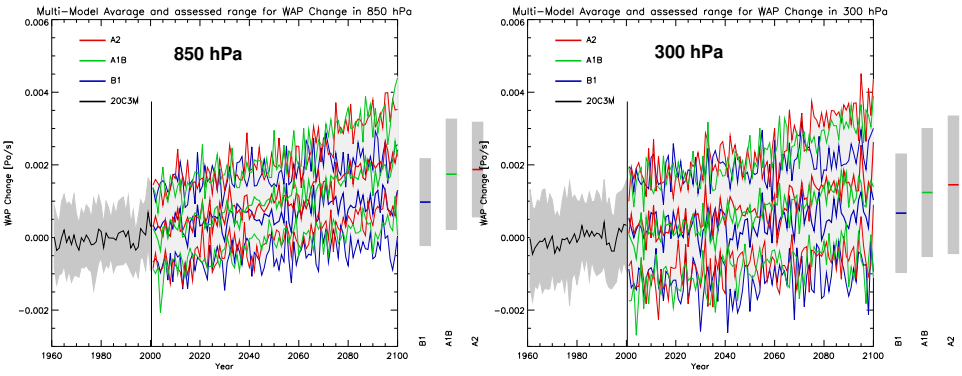


Figure 3.32: Temporal development of the lagrangian tendency of air pressure *wap* in the large domain.

In high altitudes a rising in the eastward wind speed, especially in winter and spring, is high significant. The significance disappears at the 850 hPa level particularly in the annual, JJA and SON analyses.

The changes in seasonal variances are very similar in both horizontal wind directions. Approximately half of the ensemble members are simulating a significant increase of *ua* and *va* at the end of the 21st century, independent of the pressure level, scenario or domain. The corresponding plots are visible in the digital appendix.

3.2.8 Vertical air Movement (*wap*)

The positive direction of wap (lagrangian tendency of air pressure) is downwards and $wap = -\rho \cdot g \cdot w$ (see section 2.1.1) which means that an increase in *wap* leads to an increase in vertical air velocity w. This means there is an increasing trend in wap visible in figure 3.32, and so the average vertical wind speed is rising. This trend is stronger in low altitudes. The behavior in both domains is very similar. The increasing of vertical air velocity is an indicator for more turbulences in the two research areas. Nevertheless, vertical air movements can appear on small scales and it is possible that regions with uprising winds canceling out with regions with downward winds and so in the mean there is no climate change signal visible. To exclude this behavior different analyze methods have to be applied.

By considering the information from the box whisker plot in figure 3.33 the seasonal structure of the climate change signal in *wap* can be seen. The multi model field shows no consistency for an increasing or decreasing vertical wind velocity. However, in most cases the median signal is positive (exceptions occur partly in winter and summer). The most climate sensitive periods are spring and autumn. Here a median increase in *wap* of up to 5 % is visible.

The significance analyses of the climate change signal for *wap* shows no

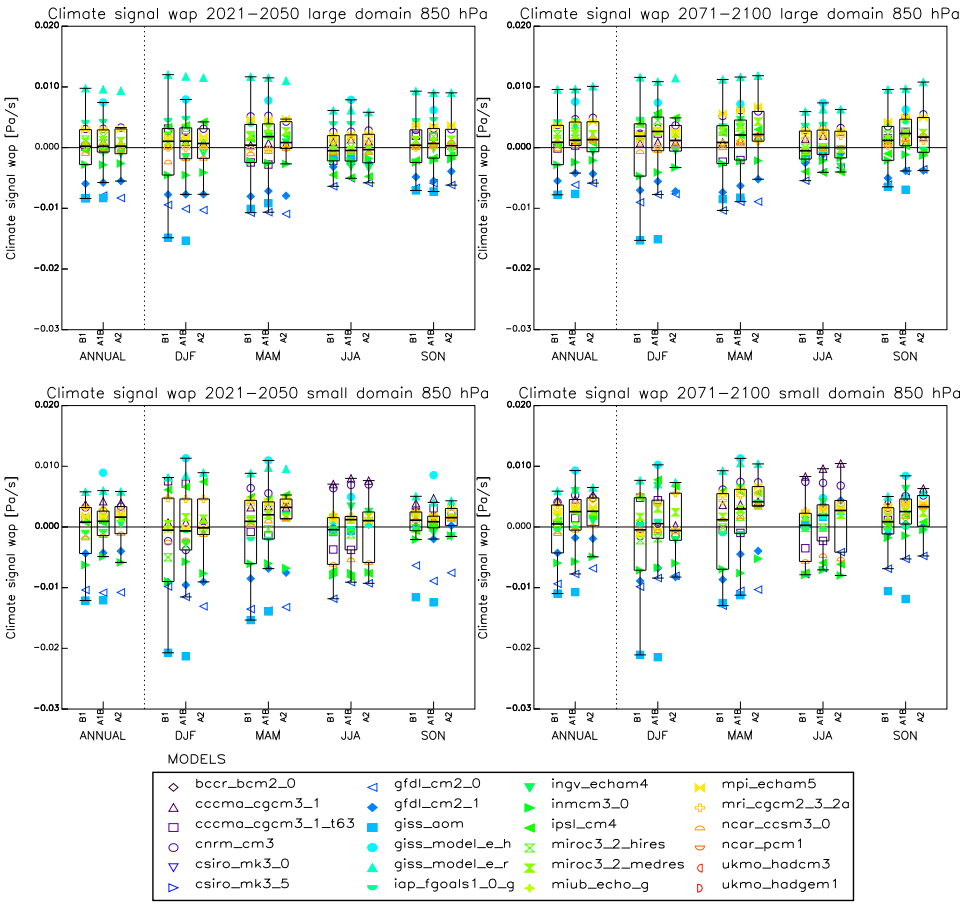


Figure 3.33: Box Whisker Plot of the climate change signal of *wap*.

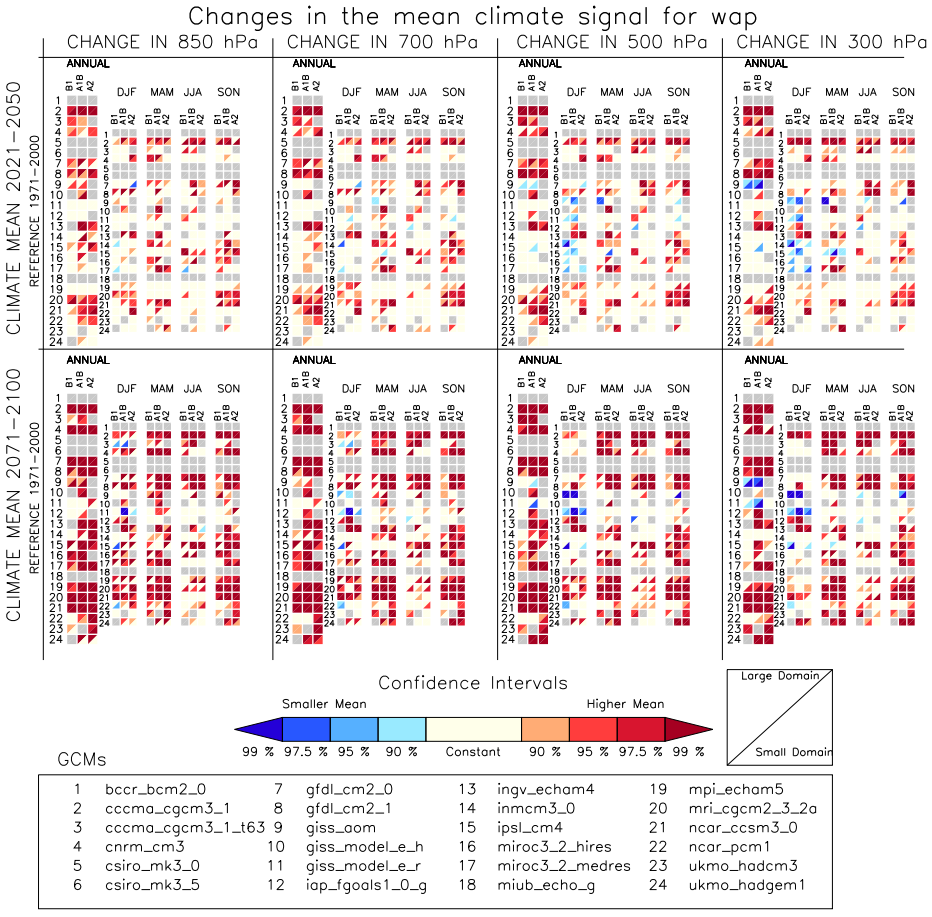


Figure 3.34: Same as figure 3.23 but for *wap*.

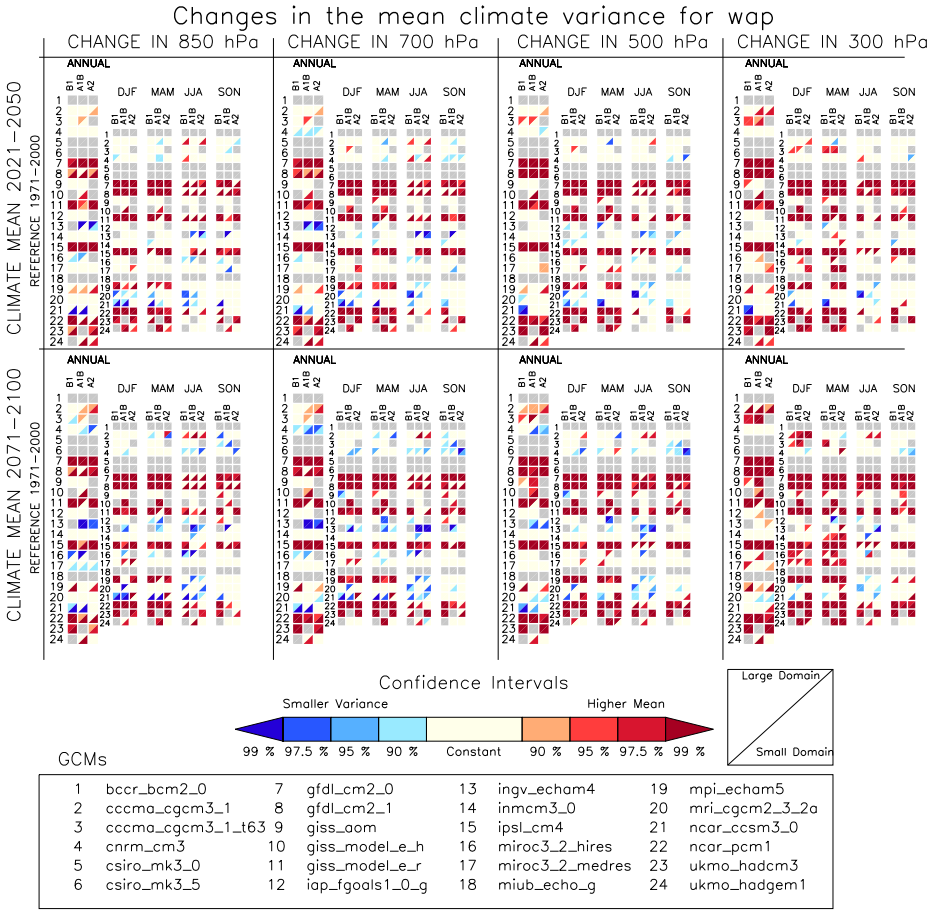


Figure 3.35: Same as figure 3.24 but for *wap*.

consistent results in the 2021 to 2050 climatology. However, at the end of the 21st century there is some evidence that the vertical air movement is increasing in all pressure levels, especially during spring and autumn. Nevertheless, there are also GCMs like the giss_model_e_r (number 11), which show no significant changes or even a decrease of *wap* over the Alpine region.

The ensemble is showing no clear result for changes in the monthly variance as figure 3.35 indicates. There are GCMs, like the gfdl_cm2_0, showing a significant increase of variance in all seasons and in both climatologies. But there are also some models showing no changes or decreases.

3.2.9 Overview of the GCM Climate Sensitivity

In figure 3.36 the relative difference (in percent) between the multi model mean climate change signals and the individual GCM climate change signals

are plotted for each parameter, pressure level, emission scenario and both domains on an annual basis. The GCM with the biggest difference determines the ±100 % interval. Red colors correspond to above-average climate change signals, blue colors to below-average climate change signals. The last column shows the mean climate sensitivity as an average over the signals of all parameters.

It is apparent that each model has at least one parameter with high climate sensitivity or insensitivity. That means, that the climate sensitivity of a GCM is dependent on the considered parameters. Nevertheless, it is possible to separate GCMs with their average sensitivity (last column in figure 3.36). Looking at the five best performing models from the model performance analyze:

- hadcm3 (MPI=0.34; rel. climat change signal=+7.8%)
- hadgem1 (MPI=0.49; rel. climat change signal=+5.4%)
- echam5 (MPI=0.50; rel. climat change signal=+20.8%)
- miroc3_2_hires (MPI=0.52; rel. climat change signal=+44.4%)
- ingv_echam4 (MPI=0.53; rel. climat change signal=-11.4%)

it gets visible that they have a wide bandwidth of climate sensitivity. The miroc3_2_hires GCM, for example, is the most sensitive GCM in the whole ensemble while the ingv_echam4 model shows a below average sensitivity.

A further result is that the sensitivity is very similar for all emission scenarios and that the results in both domains are looking similar too. However, variations occur between pressure levels (e.g., the sensitivity of the air temperature in the csiro_mk3_5 model increases with altitude) and can change temporally (e.g., the echam5 GCM which has a relative low climate sensitivity in the first half of the 21st century and an above average sensitivity at the end).

3.3 Analysis of Correlations Between GCM Performance and their Climate Sensitivity

From the model performance index analyses the capability of the models to reproduce the climate mean of the 1961 to 2000 period is known. Further, the analyze of the climate change signals showed the climate sensitivity of the GCMs. In this section links between these two quantities are analyzed.

In figure 3.37 and 3.38 the rank correlations between the performance indices and the climate change signals are displayed. The hypothesis behind this test was, that GCMs with good performance in a particular parameter are reacting similar on greenhouse gas emissions and have similar climate

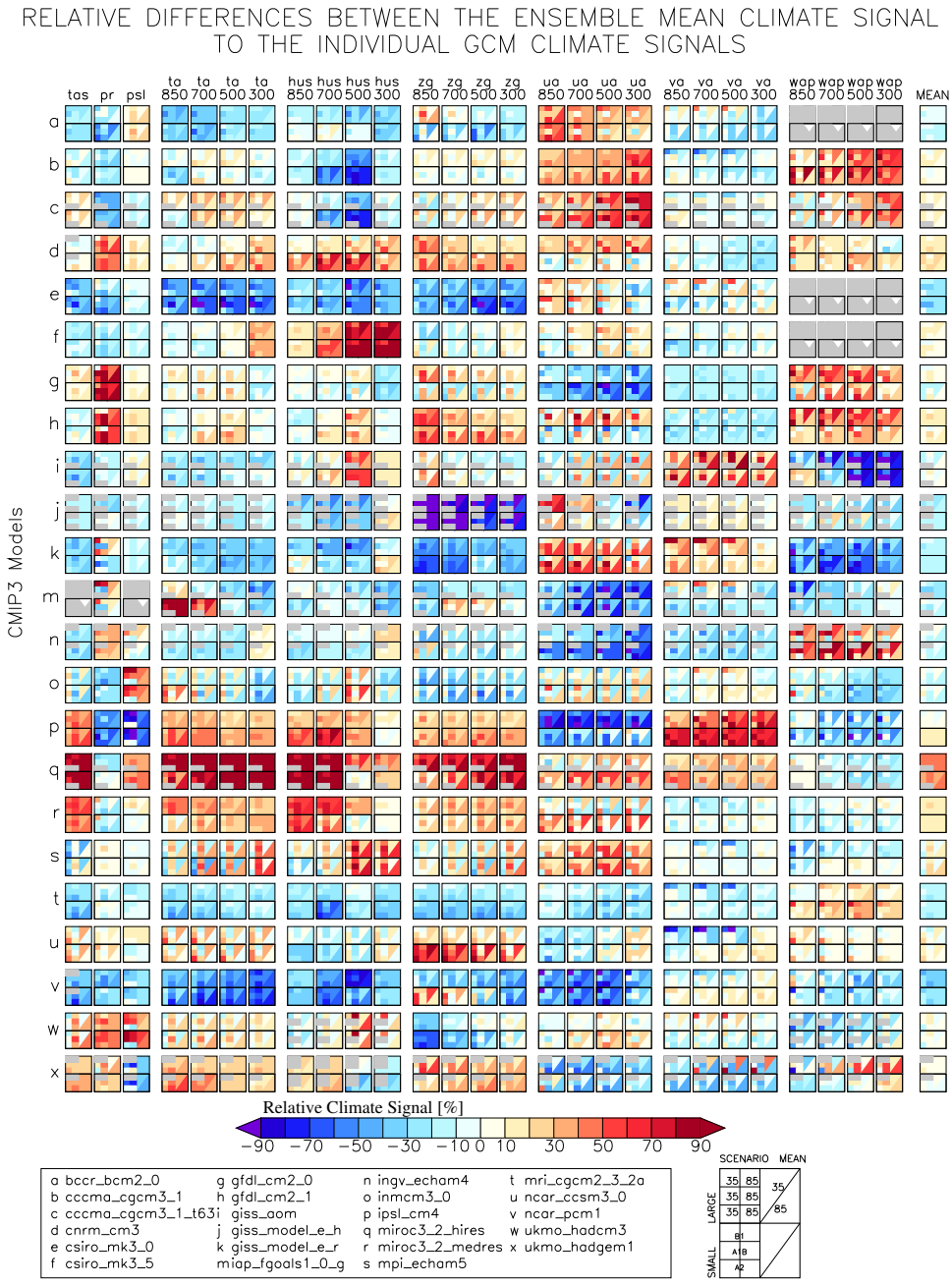


Figure 3.36: Relative differences in percent between ensemble-mean climate change signal and GCM climate change signals for all GCMs, parameters, emission scenarios and both domains and climatologies. Red colors indicate stronger climate change signals, blue colors are corresponding with weak climate change signals. Rows in the plot are different GCMs, columns display different parameters and pressure levels. The subdivision of single rectangles is explained in the lower right corner. The mean climate sensitivity is displayed in the last column (equally weighted average over all parameters).

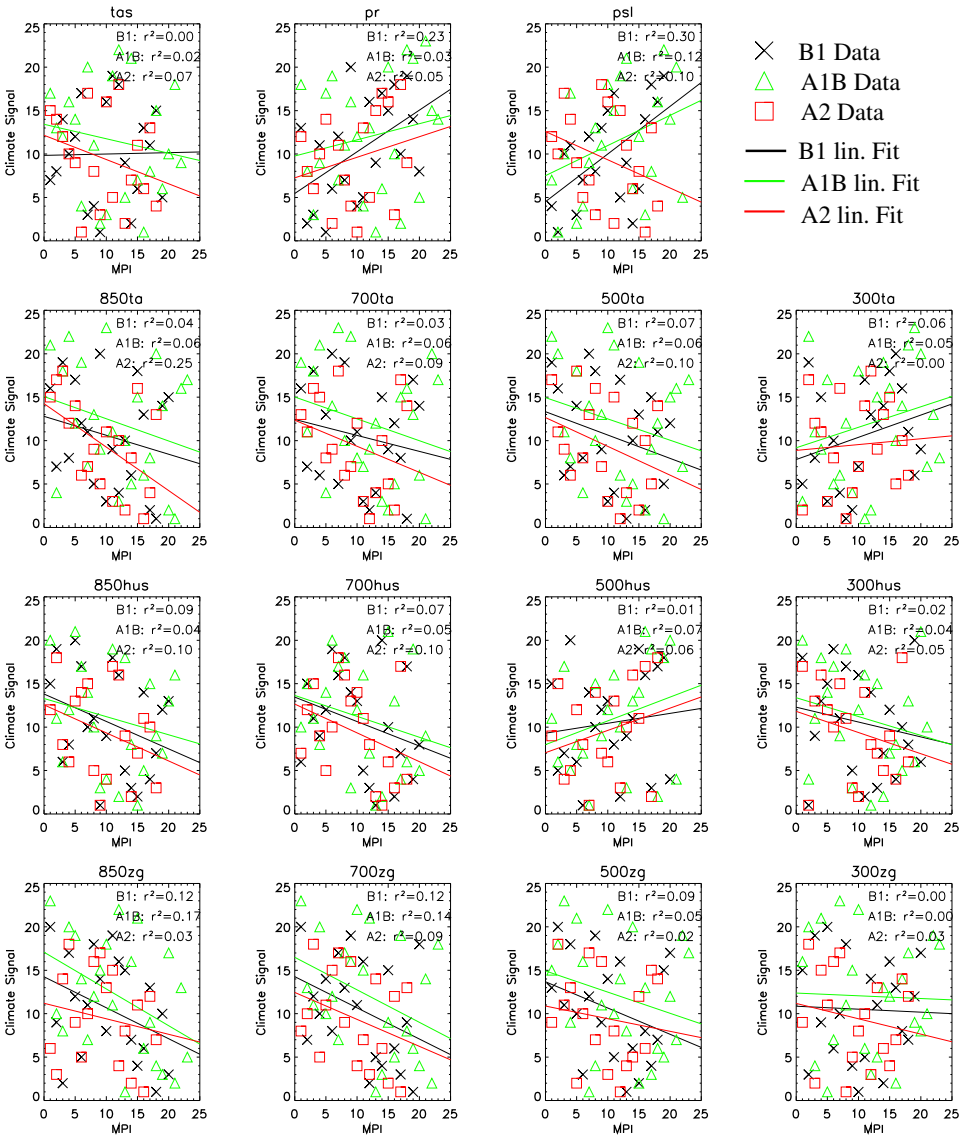


Figure 3.37: Analyze of the rank correlation between the performance indices of GCMs and their climate sensitivity for the parameters *tas*, pr, *psl*, *hus* and *zg*. The last three are analyzed at four pressure levels (850 hPa, 700 hPa, 500 hPa and 300 hPa). Different colors and symbols are showing different emission scenario experiments (black (cross) is B1, green (triangle) is A1B and red (box) is A2). The colored lines are representing the linear fit of the data. The coefficient of determination (r^2) is written on the right top of each subplot.

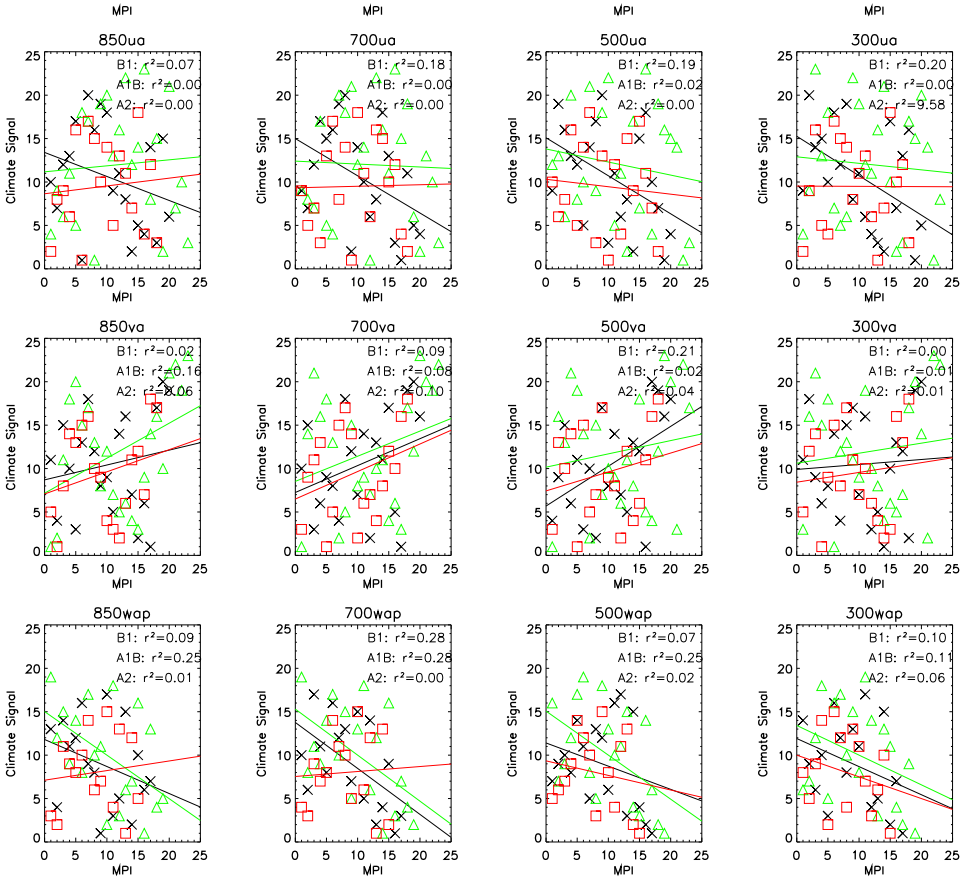


Figure 3.38: Same as figure 3.37 but for *ua*, *va* and *wap*.

sensitivities (e.g., GCMs, which have a good performance index in surface air temperature are simulating an higher temperature increase than worse performing models). The Spearman rank correlation test was chosen because the performance indices are ordinal scaled. The coefficient of determination r^2 is written on the right top of each plot for the three emission scenarios. Furthermore, the linear fit is shown as colored line in the plot.

The evaluation showed that there is no significant linear correlation between the MPIs and the strength of the climate change signals of the models. The highest coefficient of determination was found for *psl* in the scenario B1 with $r^2 = 0.30$.

As a second test for the connection between model performance and the amplitude of the climate change signals, the difference in the mean climate change signal between the best six models (concerning the MPI) and the worst six models was evaluated with a t-test for homogeneous or inhomogeneous variances. Out of 81 possible options 36 are showing a significant difference in the climate mean signal at least at the 0.90 test level. These significant parameter-emission scenario combinations are presented in table 3.1.

The best performing GCM ensemble always has a stronger climate change signal than the worst. Significant results occur especially in the temperature and the specific humidity field, where more than 80 % of the experiments are showing the above mentioned behavior. Furthermore, 50 % of the experiments concerning geopotential and one third of the eastward wind and lagrangian tendency of air pressure results are significant. It is interesting that for eastward wind only the B1 emission scenario experiments have significant stronger wind speeds in the best GCM ensemble. Finally, also the A1B experiment for surface temperature and the B1 experiment for sea level pressure have significant results.

The changes in variance are hard to interpret because there is no clear trend toward an in- or decrease between the best and the worst GCM ensembles.

3.4 Analysis of Uncertainties

In the former sections the results of the model performance analyses of the CMIP3 GCMs were presented. In the following pages the uncertainty of the climate change signal is in the focus of interest. The theoretical background of the uncertainty analysis is described in section 2.10.2. For this purpose, the results from two methods, one qualitatively [Rowell, 2006] and the ANOVA [Déqué et al., 2007], are compared. Each parameter is analyzed separately concerning:

- total uncertainty

Table 3.1: Parameters with significant differences in the climate change signal (reference: 1971 to 2000 and forecast: 2071 to 2100) between the six GCMs with the best MPIs and the six GCMs with the worst MPIs. The mean climate change signals of the ensembles are given as the difference from the multi model mean climate change signal.

		BEST		**WORST**		
PAR.	SCEN.	MEAN	VAR.	MEAN	VAR.	Sign. α
tas	A1B	0.51	0.50	-0.15	0.42	0.90
psl	B1	67.86	6293.92	-80.44	28711.40	0.90
850ta	A1B	0.67	0.45	-0.01	0.37	0.90
850ta	A2	0.38	0.18	-0.08	0.39	0.90
700ta	B1	0.43	0.34	-0.10	0.17	0.90
700ta	A1B	0.64	0.37	-0.12	0.33	0.95
500ta	B1	0.54	0.43	-0.26	0.23	0.95
500ta	A1B	0.69	0.56	-0.25	0.39	0.90
500ta	A2	0.34	0.14	-0.18	0.49	0.90
300ta	B1	0.66	0.44	-0.40	0.47	0.975
300ta	A1B	0.85	0.48	-0.36	0.50	0.975
300ta	A2	0.58	0.11	-0.38	0.73	0.975
850hus	B1	1.39E-04	6.36E-08	-4.47E-05	1.07E-08	0.90
850hus	A1B	1.91E-04	1.03E-07	-3.12E-05	2.07E-08	0.90
850hus	A2	1.36E-04	4.45E-08	-7.21E-05	2.74E-08	0.90
700hus	B1	8.42E-05	1.69E-08	-3.27E-05	8.27E-09	0.90
700hus	A1B	1.13E-04	3.05E-08	-1.87E-05	1.10E-08	0.90
700hus	A2	9.99E-05	8.74E-09	-6.27E-05	1.71E-08	0.975
500hus	A2	3.30E-05	1.20E-09	-2.13E-05	3.11E-09	0.95
300hus	B1	8.08E-06	9.95E-11	-4.62E-06	5.94E-11	0.975
300hus	A1B	9.13E-06	1.43E-10	-4.40E-06	6.65E-11	0.95
300hus	A2	7.73E-06	1.60E-10	-5.58E-06	7.40E-11	0.95
700zg	B1	4.93	51.26	-0.43	24.43	0.90
500zg	B1	9.77	169.68	-2.28	78.97	0.90
500zg	A1B	14.21	270.94	-1.11	173.93	0.90
300zg	B1	19.12	524.59	-7.51	284.64	0.95
300zg	A1B	26.73	761.74	-5.70	532.09	0.95
300zg	A2	13.94	219.38	-8.37	713.77	0.90
850ua	B1	0.03	0.00	-0.11	0.01	0.975
700ua	B1	0.06	0.01	-0.15	0.02	0.975
500ua	B1	0.10	0.02	-0.23	0.04	0.975
300ua	B1	0.14	0.03	-0.28	0.07	0.975
700wap	A1B	6.49E-04	1.10E-06	-3.69E-04	1.26E-06	0.90
500wap	B1	2.14E-04	1.96E-07	-3.79E-04	3.96E-07	0.90
500wap	A1B	5.17E-04	4.90E-07	-4.72E-04	1.64E-06	0.90
300wap	B1	1.96E-04	3.21E-07	-3.38E-04	3.78E-07	0.90

- specific uncertainty in percent
- composition of uncertainties
- the temporal dependency of the first three points (two climatologies)
- and the spatial dependency of the first three points (two domains).

3.4.1 Uncertainties in Surface Temperature (*tas*)

In figure 3.39 the uncertainty due to emission scenarios, GCM formulation and internal uncertainties (uncertainties in initial conditions) are estimated with the method described in section 2.10.1. As discussed, this method has shortcomings because of the unequally weighting of GCMs (not all GCMs have realizations in all emission scenarios and the range of single GCM runs lies between one and nine). Therefore, the method of section 2.10.3 is used to add the missing experiments artificially. The impact of this procedure on the outcome of the uncertainty estimation can be seen in the lower two graphs of figure 3.39. There is only a small change in the uncertainty distribution in the 2021 to 2050 climatology, whereas in the 2071 to 2100 climatology there is a shift of uncertainty to the internal variability of GCM runs. At the same time the emission scenario uncertainty is decreasing and the GCM formulation uncertainty stays constant.

Focusing on the differences between the small (lower two plots) and the large domain (upper two plots) in figure 3.40 it is apparent that there are only small differences in the percentage of variance in the left two plots. The main properties, like the seasonal characteristic and the ratio of different uncertainty groups, are similar in both domains. The two plots on the right show that the total amount of uncertainty is increasing in the small domain.

The major uncertainty source, as in figure 3.39, is the GCM formulation. Especially in the first half of the 21st century, 80 % of the total variance is covered by this group. At the end of the century also the emission scenarios are adding a lot of variance.

The amount of total variance is more than doubling between the two periodes. The total annual standard deviation is 0.33 K in the 2021 to 2050 climatology and 0.73 K in the 2071 to 2100 climatology. There are also big differences in the amount of variance between different seasons. The most uncertain period of the year is summer (JJA) whereas spring (MAM) has the lowest uncertainty.

3.4.2 Uncertainties in Precipitation Flux (*pr*)

In figure 3.41 the distribution of the climate change signal of precipitation flux (*pr*) is shown for the uncertainty groups. Again, as for *tas*, the major difference between the upper and the lower plots is an increase in the internal

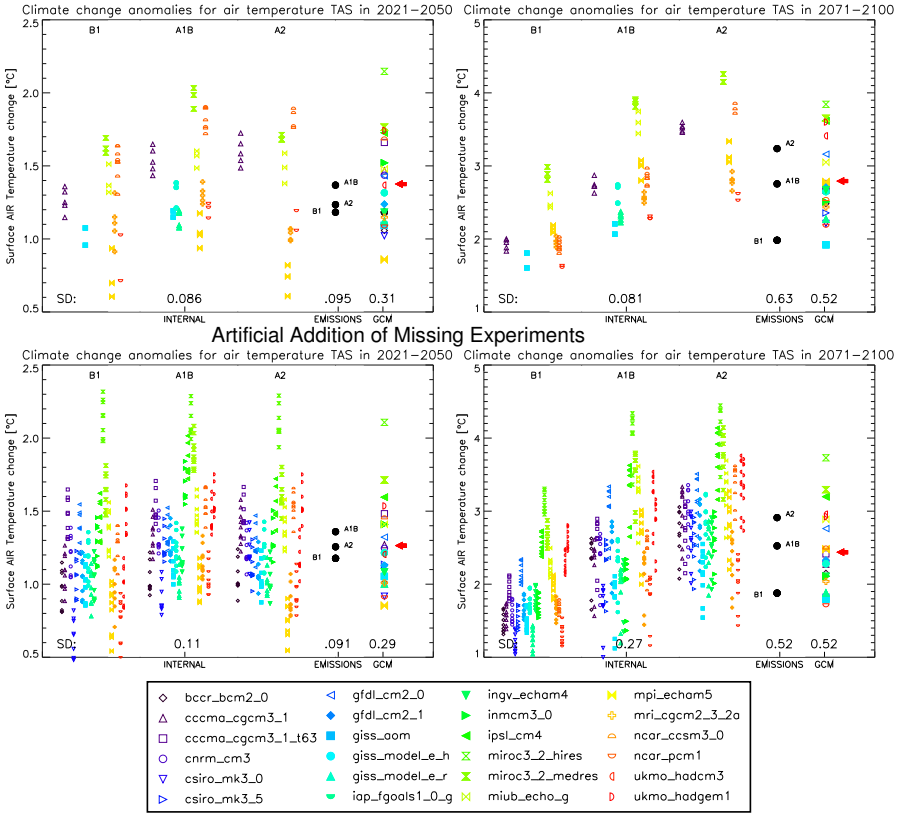


Figure 3.39: Differences in the uncertainty estimation for *tas* in both climatologies (2021 to 2050 and 2071 to 2100) for the original data matrix (upper two graphs) and the artificially filled data matrix (lower two graphs) in the large domain. The two graphs on the left hand side are showing the climate change signal for the 2021 to 2050 climatology whereas the plots on the right are showing the 2071 to 2100 climatology. The colored symbols are corresponding to the GCMs (see legend). The standard deviation (SD) of the three uncertainty groups are written on the bottom of each plot.

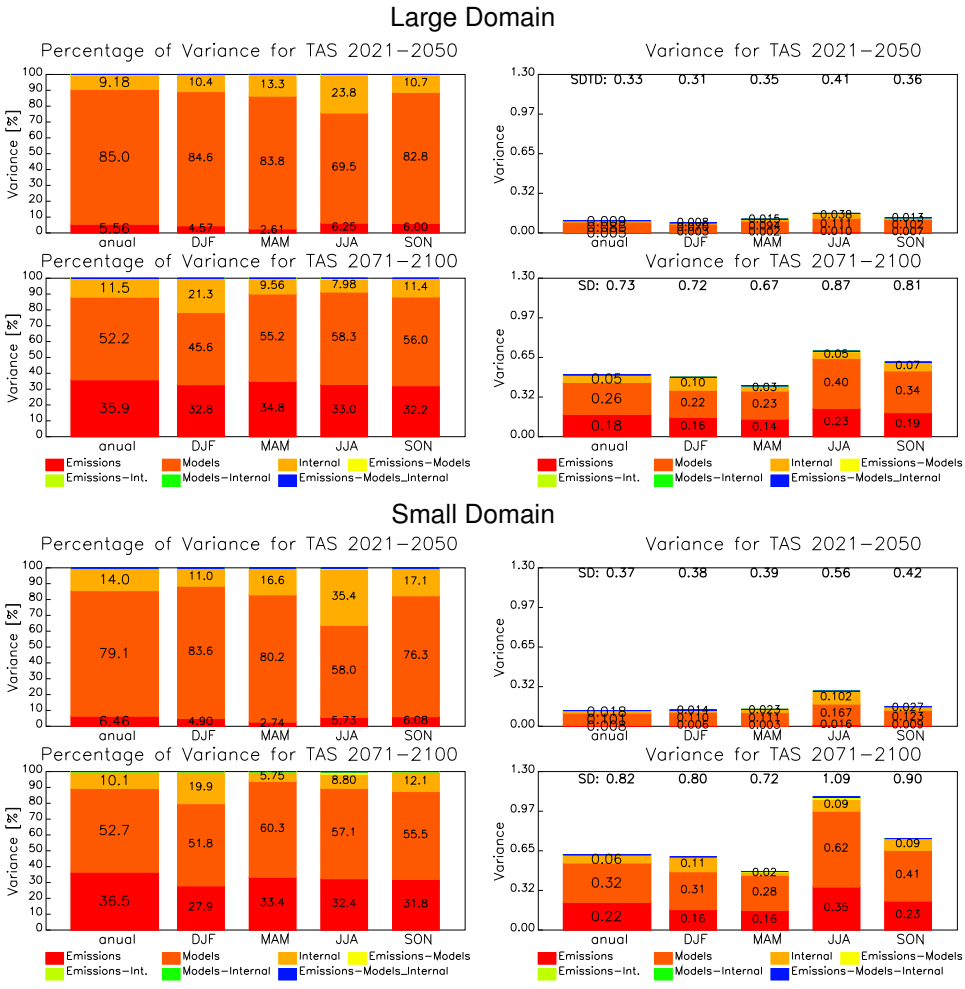


Figure 3.40: Temperature at surface ANOVA for both climatologies, domains and on annual and seasonal basis. The plots on the left side are showing percentages of the total uncertainty, whereas those on the right hand side are presenting the absolute values of variance. In the right figures also the standard deviation (SD) is written above the bars.

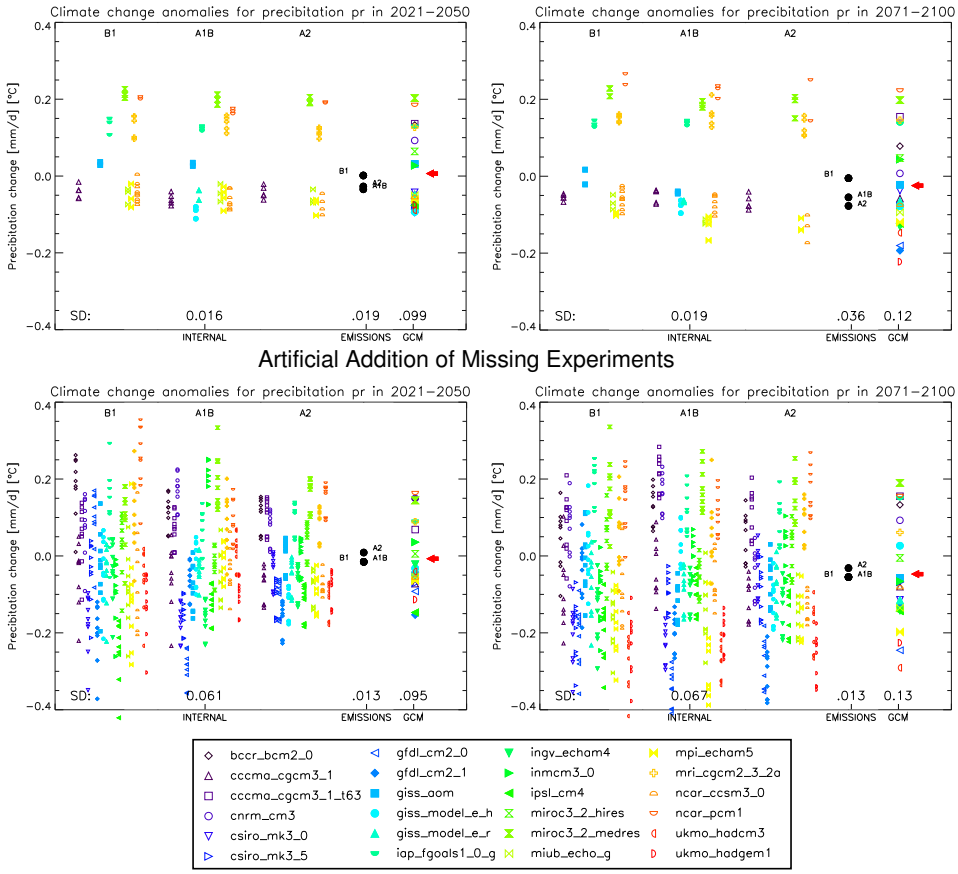


Figure 3.41: Same as figure 3.39 but for pr.

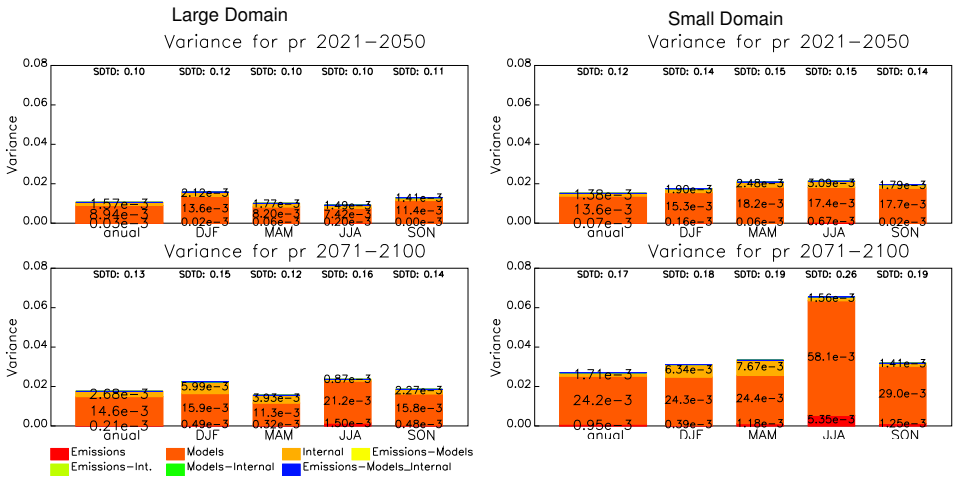


Figure 3.42: Same as the right plots in figure 3.40 but for pr.

uncertainty. Since the corresponding variance is small compared to the total variance, this behavior is not critical. The emission scenario uncertainty is approximately as big as the internal uncertainty. The bulk of variance is covered by the GCM formulation and is approximately one magnitude larger than those from the other uncertainty sources. Compared with the *tas* analyses the uncertainty range of the *pr* climate change signal is relatively constant during the 21st century.

The results from the ANOVA are displayed in figure 3.42. The plots with the uncertainty ranges in percentages can be found in the digital appendix. Again the results are similar in both uncertainty estimation methods. The GCM formulation is the biggest source of uncertainty in the *pr* climate change signal. The season with the highest uncertainty is summer, especially in the small domain.

3.4.3 Pressure at Sea Level (*psl*)

The uncertainty estimation in figure 3.43 shows that also in the case of sea level pressure the GCM formulation contributes the largest part to the uncertainties. The standard deviation of the mean GCM formulation climate change signals is doubling from the first half of the 21st century to the end. The standard deviation of internal variability and emission scenarios have approximately the same magnitude and both show only a slight temporal increase. The results from the analysis of the artificial adding of missing data are similar and depicted in the digital appendix.

The findings of the ANOVA method (figure 3.44) show similar results as figure 3.43 on the annually basis. However, the seasonal evaluation shows that internal variability in the climate system can cause a notable amount

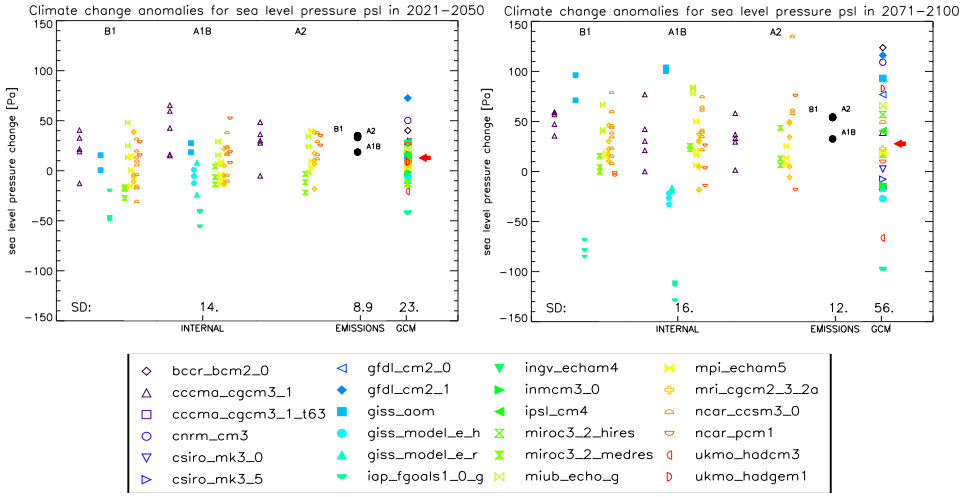


Figure 3.43: Same as upper plots in figure 3.39 but for *psl*.

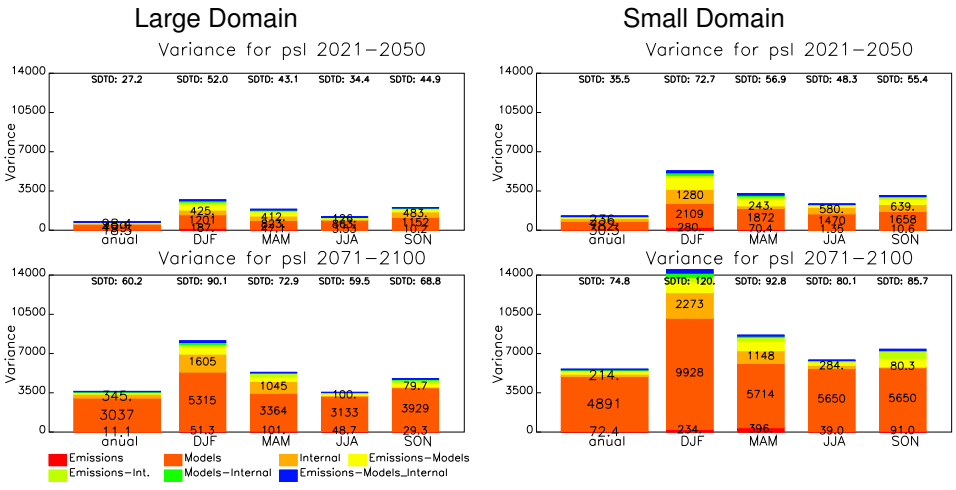


Figure 3.44: Same as right plots in figure 3.40 but for *psl*.

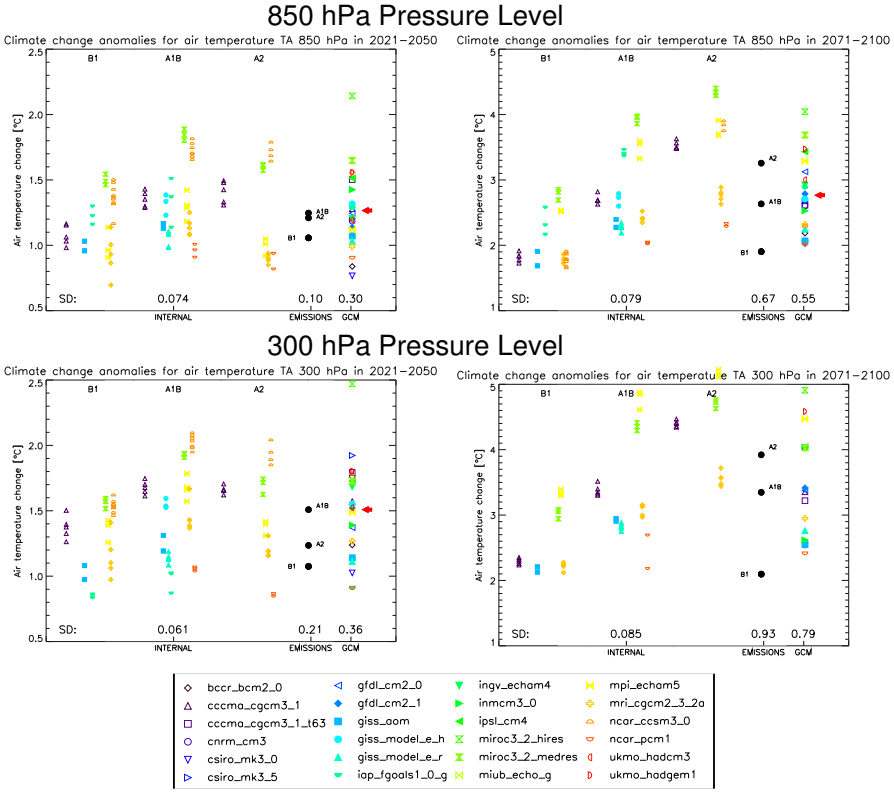


Figure 3.45: Same as upper plots in figure 3.39 but for *ta* at the 850 hPa and the 300 hPa pressure level.

of uncertainty especially in the winter and spring season. Furthermore there is a visible amount of uncertainty in the combination of emission scenarios and GCM formulation.

Again, as in the *tas* and *pr* analyses, the uncertainties are increasing on the small domain. For *psl* the winter season has the highest total uncertainty.

3.4.4 Air Temperature (*ta*)

The results from the estimation of uncertainties in the climate change signal of *ta*, at the 850 hPa pressure level (see figure 3.45), are similar to those of the *tas* evaluation in figure 3.39. At the 300 hPa pressure level the characteristic of the distribution stays the same, but the amount of total uncertainty is increasing. The results at 700 hPa and 500 hPa are similar to those at 850 hPa and 300 hPa and the amount of total uncertainty is in between. The corresponding plots can be found in the digital appendix.

The ANOVA (figure 3.46) shows the dominance of GCM formulation uncertainty, especially in the first half of the 21st century, and the increasing variance in the emission scenarios until the end of the 21st century. As for

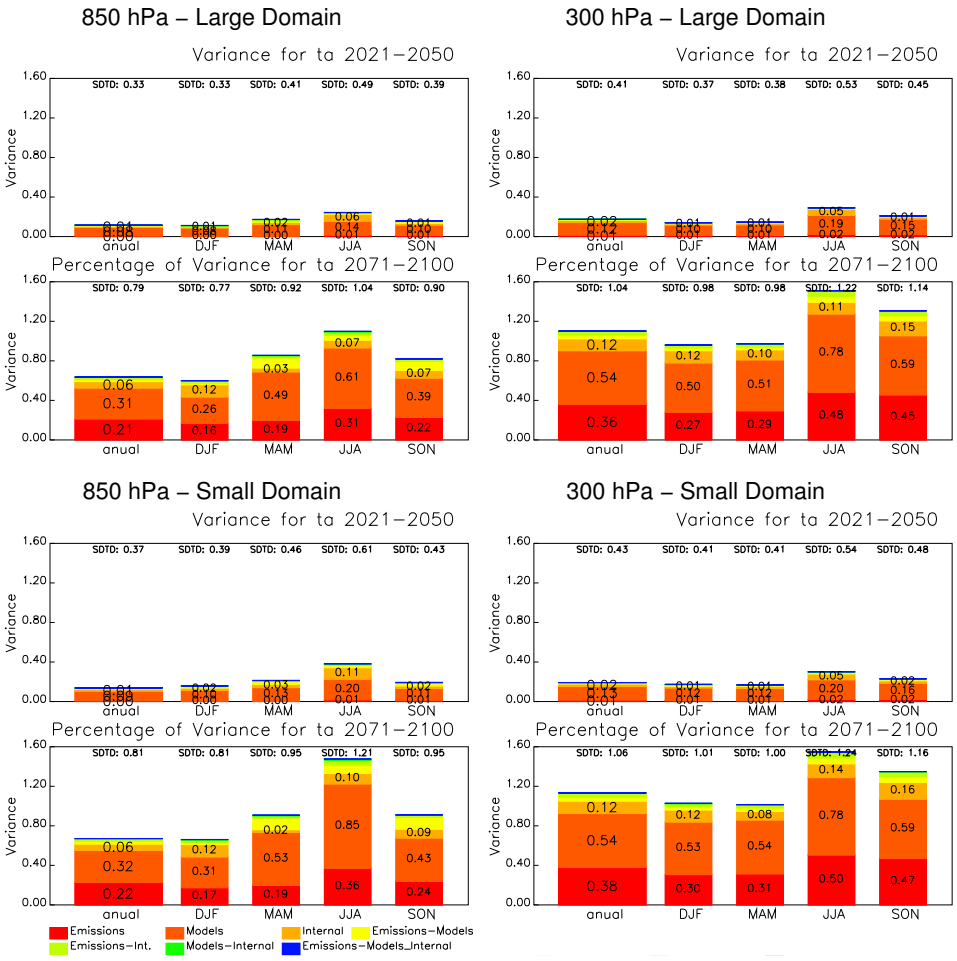


Figure 3.46: Same as the right plots in figure 3.40 but for *ta* at the 850 hPa and the 300 hPa pressure level.

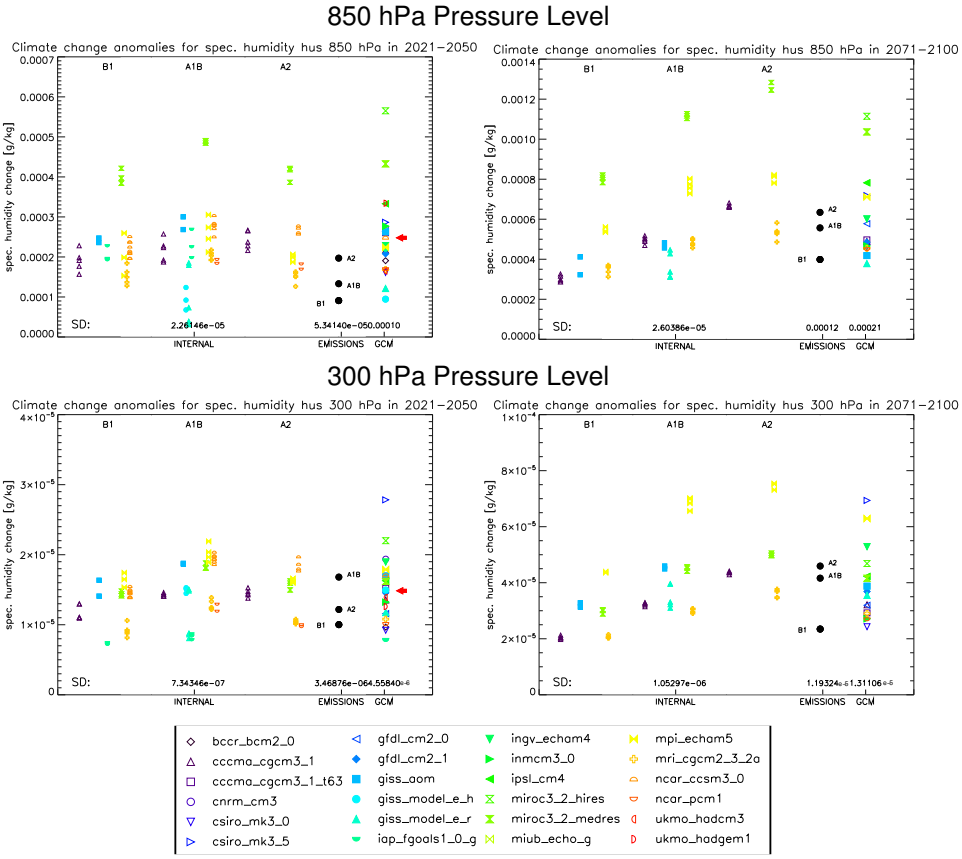


Figure 3.47: Same as the right plots in figure 3.40 but for specific humidity (*hus*) on the 850 hPa and the 300 hPa pressure level.

tas, the total amount of variance is more than doubling from the first half of the century to the end at all pressure levels. The uncertainty composition is nearly the same for both domains, but the amount is bigger in the small one. The most uncertain period is, as for *tas* and *pr*, the summer season and the season with the lowest amount of total variance is winter.

3.4.5 Specific Humidity (*hus*)

Specific humidity is inhomogeneous distributed in the atmosphere and is rapidly decreasing with height. Because of that, the plots in figure 3.47 have different scales on the ordinate to guarantee the visibility of the results.

At both displayed pressure levels, the major source of uncertainty is the GCM formulation, followed by emission scenario and internal uncertainties. Especially the second is increasing during the 21st century because of the connection between *hus* and *ta*.

The ANOVA evaluations (figure 3.48) show similar results. Also here

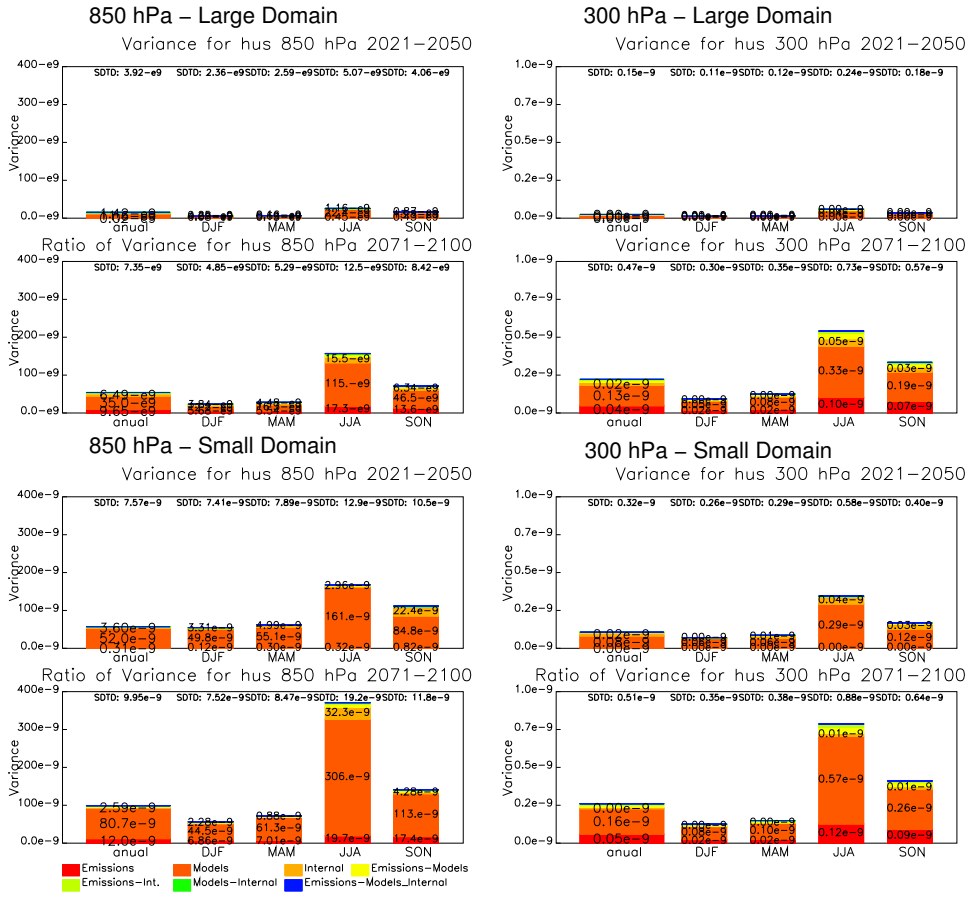


Figure 3.48: Same as the right plots in figure 3.40 but for *hus* on the 850 hPa and the 300 hPa pressure level.

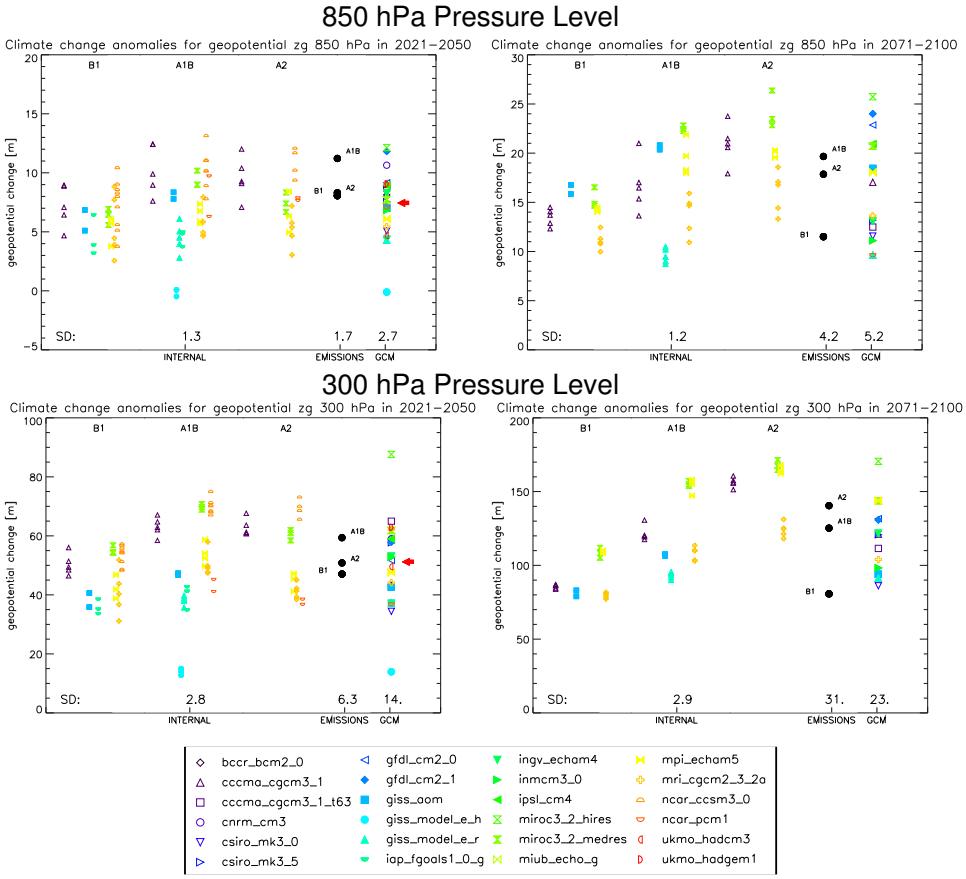


Figure 3.49: Same as upper plots in figure 3.39 but for geopotential height (zg) at the 850 hPa and the 300 hPa pressure level.

the major uncertainty source is the GCM formulation. The ratio of the emission scenario is increasing during the end of the 21st century. This effect enhances with increasing altitude. Internal effects are negligible. There is a strong increase of uncertainty between the 2021 to 2050 and the 2071 to 2100 climatologies, especially in the large domain. Furthermore, the variances are generally larger on the small domain than on the large one. The season with the highest amount of variance is summer.

3.4.6 Geopotential Height (zg)

The upper plots in figure 3.49 show the uncertainty estimation for the geopotential height (zg) at the 850 hPa pressure level for the large domain. In the first half of the 21st century the internal- and emission scenario uncertainties are approximately half as big as the GCM uncertainty. While the internal uncertainties are temporal constant until the end of the 21st century, the GCM and especially the emission scenario uncertainties are increasing. A

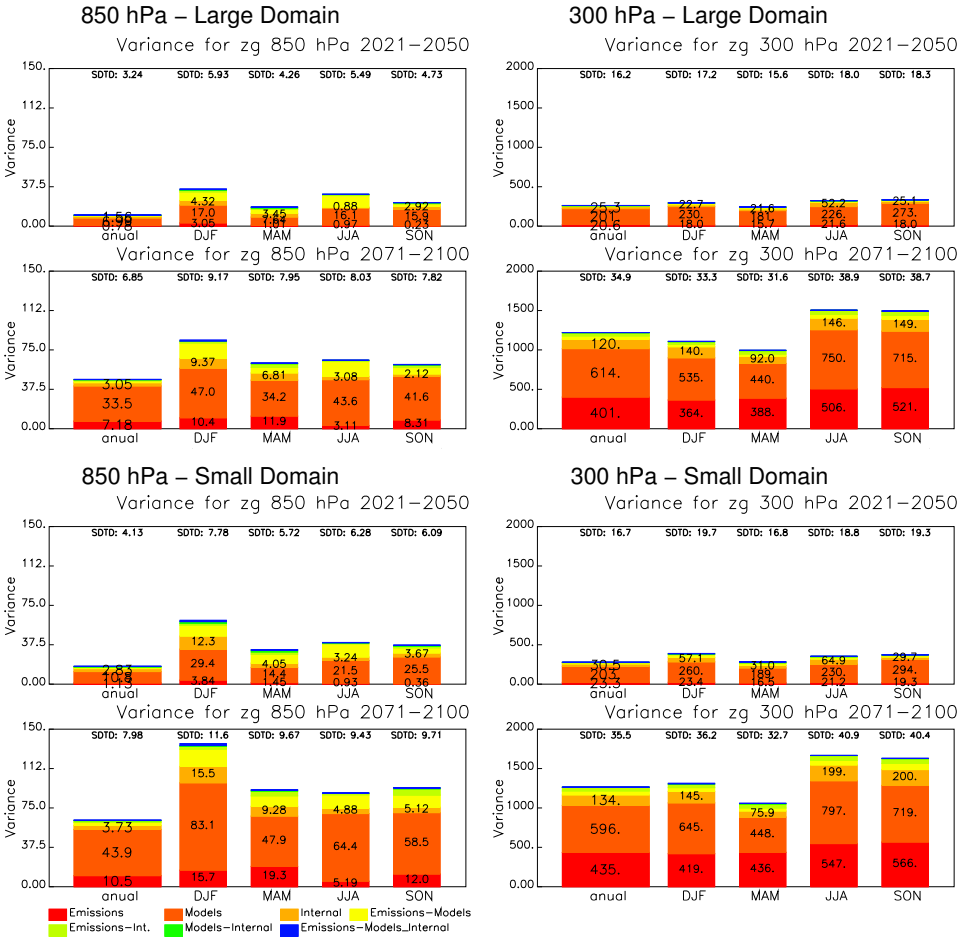


Figure 3.50: Same as the right plots in figure 3.40 but for geopotential height (zg) at the 850 hPa and the 300 hPa pressure level.

similar result is obtained at the 300 hPa level (lower plots in figure 3.49). As for the other variables the uncertainty range is increasing with height. A small amount of variance is covered by internal processes while GCM formulation and especially emission scenarios, in the 2071-2100 climatology, are contributing the largest parts of uncertainty.

The results from the ANOVA for zg are visible in figure 3.50. At the 850 hPa pressure level (left side) the major uncertainty source is the GCM formulation. The contribution of emission scenarios to the variance is increasing with time. The internal uncertainty plays a minor role. The major difference between the two domains is the stronger increase of total variance in the small domain. In both domains the season with the largest uncertainty at 850 hPa is winter.

This is changing with increasing altitude where the summer and autumn

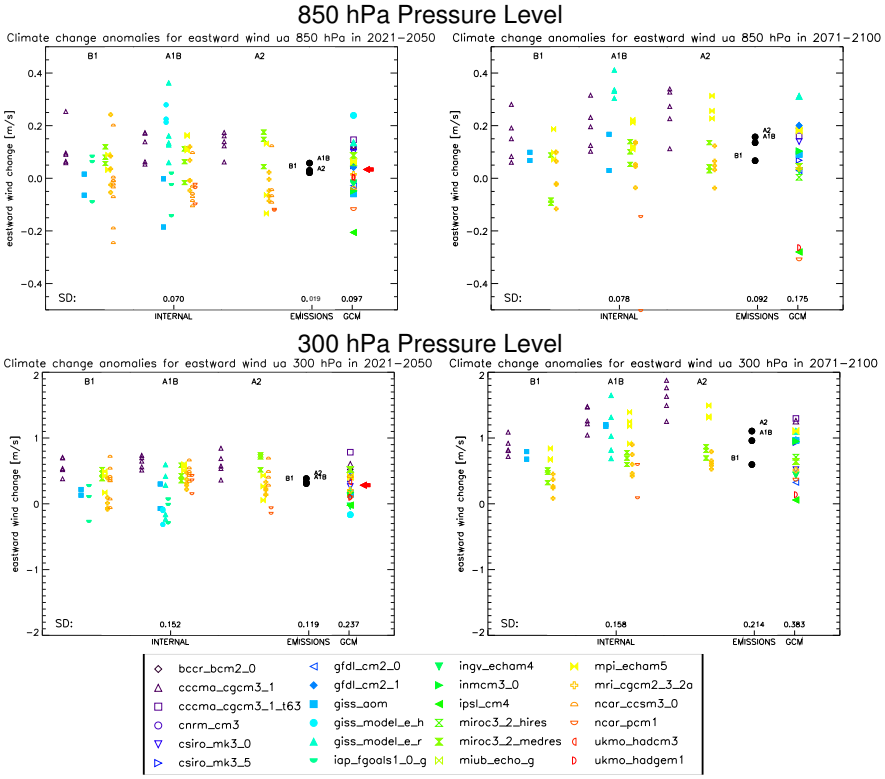


Figure 3.51: Same as upper plots in figure 3.39 but for eastward wind (*ua*) at the 850 hPa and the 300 hPa pressure level.

climate change signal are getting more uncertain. Furthermore, compared with higher pressure levels, a very strong increase of the emission scenario uncertainty is visible during the 21st century. In the 2071 to 2100 climatology one third of the total variance is added from this uncertainty source. Comparing the results from the two domains at 300 hPa (upper and lower plot on the right hand side of figure 3.50) there is nearly no difference in amount and distribution of variance visible.

3.4.7 Eastward Wind (*ua*)

The uncertainty estimation for eastward wind (*ua*) is showing that the GCM formulation is the biggest uncertainty source. In the first half of the century, the emission scenarios have nearly no contribution to uncertainty, but internal effects are important. While the latter are staying approximately constant until the end of the century, the GCM uncertainty is doubling and the emission scenario uncertainties are quadrupling at the 850 hPa pressure level. The effects are staying similar in higher altitudes, but their temporal growing rates are smaller.

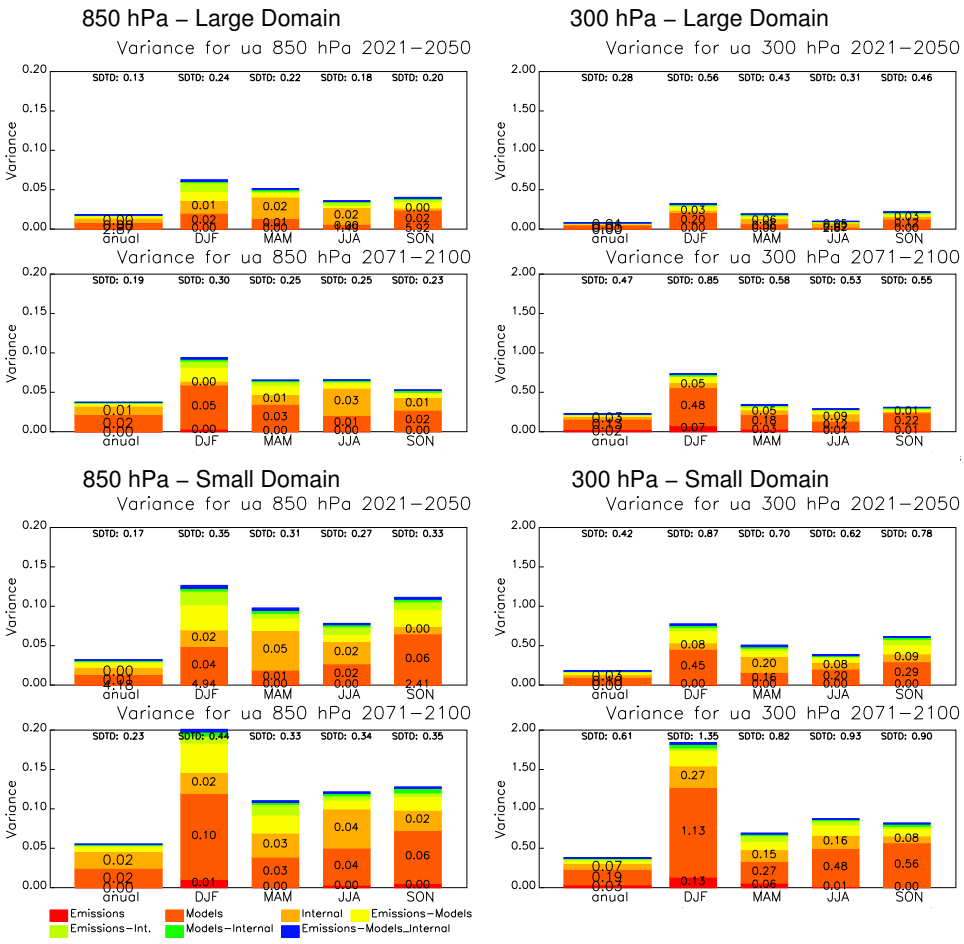


Figure 3.52: Same as the right plots in figure 3.40 but for eastward wind at the 850 hPa and the 300 hPa pressure level.

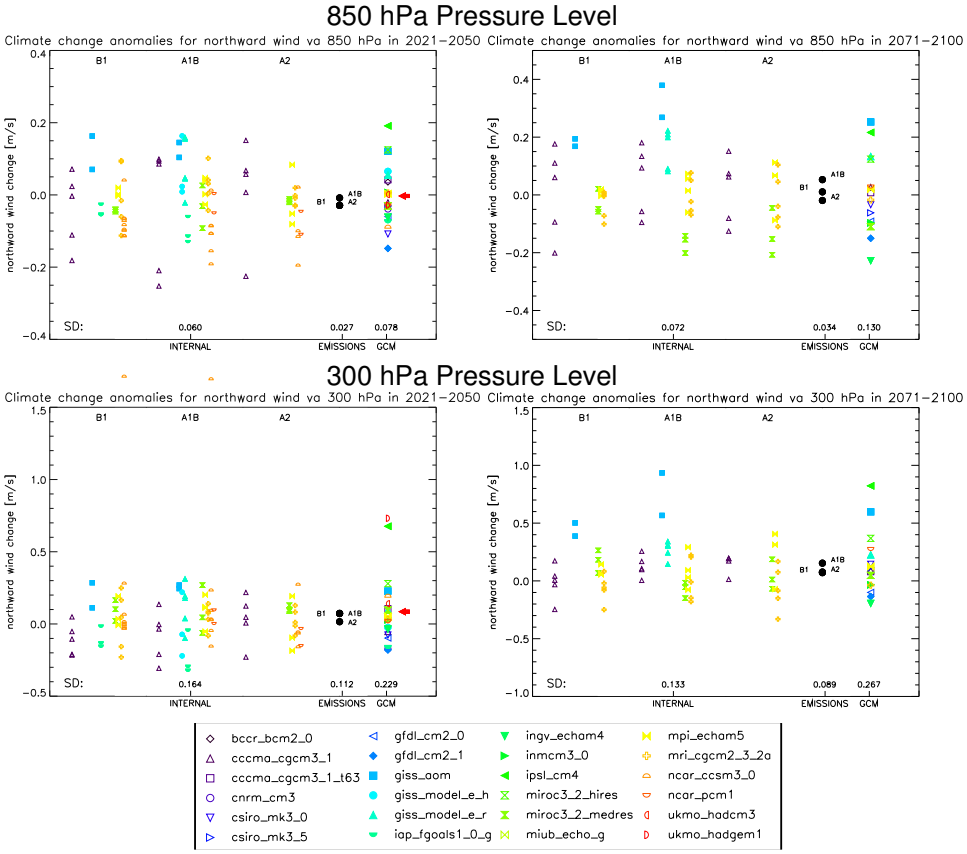


Figure 3.53: Same as upper plots in figure 3.39 for northward wind (*va*) at the 850 hPa and the 300 hPa pressure level.

The results of the ANOVA in figure 3.52 show similar characteristics. However, there are some differences like the small contribution of emission scenarios to the total variance at the end of the 21st century. Only in winter there is a notable effect visible. But the combinations of the GCM formulation and emission scenario uncertainty (yellow bars) are more important (especially on the small domain) than it was for other parameters. The two uncertainty analyses are in good agreement with the relatively large uncertainty of internal effects. Nevertheless, the GCM formulation is the major source of uncertainty. The season with the highest uncertainty is winter.

3.4.8 Northward Wind (*va*)

In case of northward wind the largest variance is carried by the GCM formulation followed by internal effects and emission scenarios (figure 3.53). The absolute amount of uncertainty is increasing with altitude and time.

Figure 3.54 shows the results of the ANOVA for northward wind veloc-

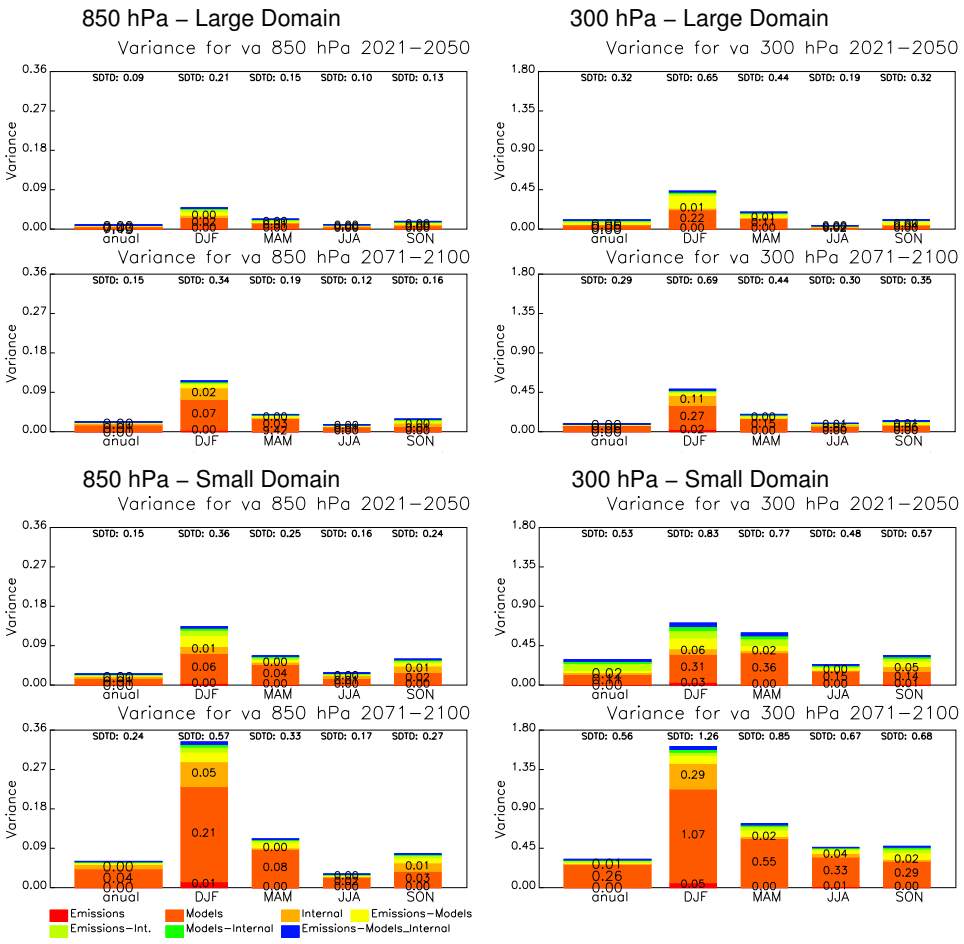


Figure 3.54: Same as the right plots in figure 3.40 but for northward wind at the 850 hPa and the 300 hPa pressure level.

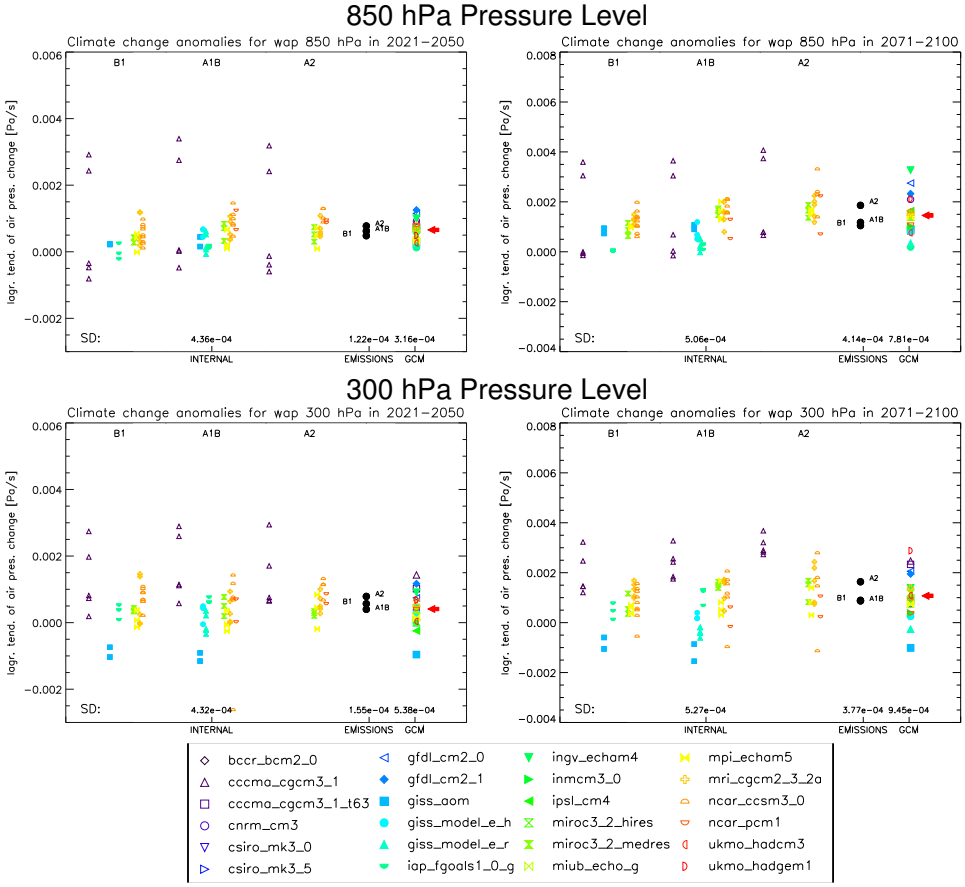


Figure 3.55: Same as upper plots in figure 3.39 but for lagrangian tendency of air pressure (*wap*) at the 850 hPa and the 300 hPa pressure level.

ity. The most striking feature is the high total variance in winter. Especially in the 2071-2100 climatology the uncertainty amount is three to four times higher than for the annual average. Generally, the uncertainty is approximately twice as large in the small domain as it is in the large one. As in the previous analyses, the GCM formulation is the dominant uncertainty source.

3.4.9 Lagrangian Tendency of Air Pressure (*wap*)

The uncertainty estimation of the vertical air movement (lagrangian tendency of air pressure *wap*) is shown in figure 3.56. There is no uncertainty source which is especially dominant or especially small compared to the others. In the first half of the 21st century the variance of GCM formulation and internal effects are nearly equal. Until the end of the century the emission scenario uncertainty is increasing to the amount of the internal uncertain-

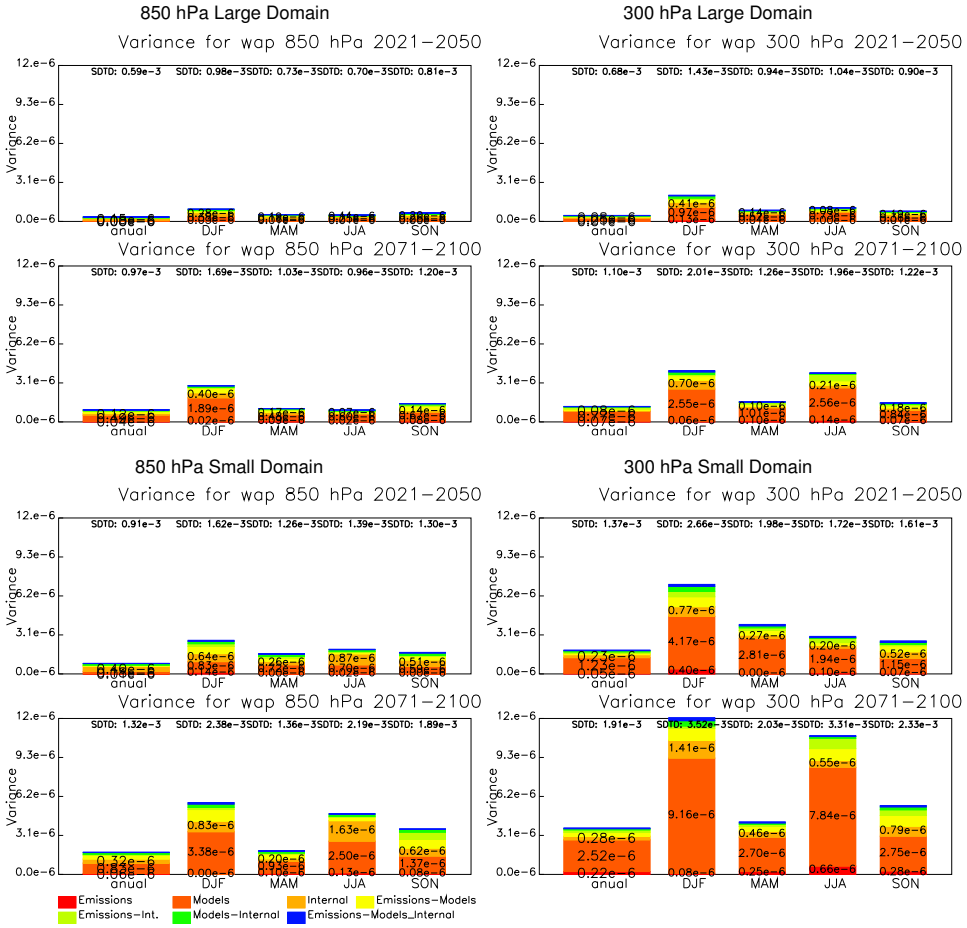


Figure 3.56: Same as the right plots in figure 3.40 for lagrangian tendency of air pressure at the 850 hPa and the 300 hPa pressure level.

ties. The variance due to GCM formulation is doubling from the first half of the 21st century to the end and contributes the biggest part of the total uncertainty.

Due to the ANOVA results in figure 3.56 the major source of uncertainty is the GCM formulation. Emission scenario uncertainties are small for both climatologies. However, the emission scenario-model interaction uncertainty is important and has approximately the same value as the internal effects. The uncertainties are increasing with time, altitude and in the small domain. The season with the largest total variance is winter as well as summer in the second half of the century.

Chapter 4

Discussion

The results of the preceding chapter supply the foundations to answer the initially asked questions:

- which GCMs are suitable to provide initial and boundary conditions for RCMs in the Alpine region?
- how do the climate projections of the GCMs look like?
- what are the sources of uncertainty in GCM climate change signals?

In the following, the findings from chapter 3 are interpreted and compared with known results from literature. The used methods are evaluated and the validity as the limitations of the results are discussed.

Model Performance

The GCM evaluations show that there is no single GCM which universally outperforms all other models from the CMIP3 ensemble. Furthermore, the MPI depends on the considered parameters and season (see figure 3.11 and 3.10), but the difference between the two domains is relatively small (see figure 3.10). However, there are some GCMs which show an above average performance in most of the evaluated fields. According to the dark blue boxes in the outer right column in figure 3.12 these top performing models are listed in tabular 4.1.

Remarkable is the especially good performance of the multi model mean (row y in figure 3.12). This is also visible in every other evaluation method (e.g., Taylor plots figure 3.5, bias maps figure 3.3 or seasonal circle analyses figure 3.1) and is in good agreement with scientific literature (e.g. [Reichler and Kim, 2008a], [Weigel, 2008] or [Gleckler et al., 2007]).

So far there is no published analysis of the performance of the CMIP3 GCMs in the Alpine region, but Reichler and Kim [2008a] calculated the

MPI globally. Furthermore, they presented unpublished evaluations in a talk at the annual meeting of the EGU (European Geophysical Union) in 2008 where they pointed out that: "Good models do well over all regions and all quantities" [Kim and Reichler, 2008]. The top five models from their global evaluation are listed and compared with the obtained results in this work in tabular 4.1.

Table 4.1: Comparison of the five GCMs with the best MPI with the top five models (global MPI) from a study done by Reichler and Kim [2008a].

Rank	MPI (Alpine region)	MPI Reichler (global)
1	hadcm3	echam5
2	hadgem1	hadgem1
3	echam5	gfdl_cm2_1
4	miroc3_2_hires	hadcm3
5	ingv_echam4	ingv_echam4

Even if the domain, the evaluation period (Reichler and Kim used 1979 to 1999) and the parameter composition is different, the results are very similar. This shows that the MPI is a temporal and spatial robust index if enough parameters are considered. Also other analyses lead to a similar GCM selection. Gleckler et al. [2007] used different evaluation methods with focus on the global mean annual cycle with the result: "In the extra-tropics, for example, we find the UKMO-HadCM3, UKMOHadGEM1, GFDL-CM2.1, MICROC3.2 (hires), and MPIECHAM5 errors are smaller than those found in the "typical" model by more than 10 %." [Gleckler et al., 2007].

It seems that the MPI is a very useful tool to get an overview of the general model performance, but it also has its limitations. Weaknesses of a model can be easily overlooked when they are hidden within one single quantity. For example, the well performing echam5 GCM has a poor performance of spring air temperature on each considered pressure level. Furthermore, quantities like the daily cycle, the simulation of extreme events or climatological trends, which are also important for RCMs, are disregarded in the MPI calculation. A good approach to choose a GCM for regional climate modeling would be to look first at the MPI for the different domains and different seasons and then look into more detail like parameter specific indices, maps or Taylor plots.

To find the reasons why certain GCMs are performing better or worse in the investigated domains is beyond the scope of this study. However, the MPI seems not to depend on flux adjustment or grid resolution. There are only four flux adjusted models in the ensemble (cccma_cgcm3_1, cccma_cgcm3_1_t63, inmcm3_0 and mri_cgcm2_3_2a), which have an average performance and do not show any systematic biases. Furthermore, the grid resolution of the best performing models is very different and ranges from a

T106, L56 resolution, in case of the miroc3_2_hires, to a T42, L19 resolution for the ukmo_hadgem1 and the ingv_echam4 GCMs.

GCM Climate Change Signals in the 21st Century

In general, GCMs do not show a special high or low climate change signal for all parameters (see figure 3.36). Nevertheless, there are some GCMs, which have high or low climate change signals for a broad range of parameters. For example, the miroc3_2_hires GCM has a very strong climate change signal in nearly each considered field with exception of the lagrangian tendency of air pressure and precipitation. Averaged over all parameters it has a 40 % higher mean climate change signal than the ensemble mean. On the other side of the spectra there is the ncar_pcm1 GCM, which has an overall weak climate change signal (with exception of northward wind and geopotential height in the small domain). Its average climate change signal is more than 30 % lower than the GCM ensemble mean signal.

The differences between specific GCMs and the multi model mean are relatively constant in all emission scenarios. This means that models with a strong climate change signal in the A2 scenario generally have a strong climate change signal in the B1 and A1B scenarios too. The same is valid for the domain size. A weak signal in the large domain indicates a weak signal in the small one.

Not many studies are comparable with the above mentioned results. Many puplications are focusing on the temperature increase caused by a doubling of CO_2 in the atmosphere (this is generally called climate sensitivity) or on the transient climate response[1]. By considering that the climate change signals in the 21st century in the Alpine region are comparable to the global climate sensitivity (equilibrium temperature for a doubling of CO_2) the results from the IPCC fourth assessment report can be compared with the results from section 3.2.1 for surface temperature (see table 4.2).

Three of the top five models are the same compared with the transient climate response and there are even four matches in the comparison with the climate sensitivity experiment. Considering the differences between the compared quantities this result increases the confidence in the method, which is used in this study.

The following listing of parameters gives an overview of the climatological changes in the annual and seasonal mean climate. It focuses on the climate change signal 2071 to 2100 minus 1971 to 2000 in the large domain (the results are very similar in the small domain).

[1]The IPCC defines the the transient climate response as: "... the change in the global surface temperature, averaged over a 20-year period, centred at the time of atmospheric carbon dioxide doubling, that is, at year 70 in a 1 % per year compound carbon dioxide increase experiment with a global coupled climate model. It is a measure of the strength and rapidity of the surface temperature response to greenhouse gas forcing." [IPCC, 2007].

Table 4.2: Comparison of the five GCMs with the strongest temperature at surface climate change signal (reference: 1971 to 2000, forecast: 2071 to 2100 in the large domain) with the five GCMs with the highest transient climate response and the highest equilibrium global climate response (climate sensitivity) [IPCC, 2007].

Rank	climate change signal	transient climate response	climate sensitivity
1	miroc3_2_hires	miroc3_2_hires	ipsl_cm4
2	miroc3_2_medres	mpi_echam5	ukmo_hadgem1
3	ipsl_cm4	mri_cgcm2_3_2	miroc3_2_hires
4	ukmo_hadcm3	miroc3_2_medres	miroc3_2_medres
5	ukmo_hadgem1	ipsl_cm4	mpi_echam5

- **Temperature at Surface:** significantly increasing in all GCM simulations and in all scenarios (see section 3.2.1). Half of the ensemble shows an increasing inter-annual seasonal variability in winter. The season with the highest climate change signal is summer.
- **Precipitation Flux:** majority of models are showing a decrease of precipitation in spring, summer and autumn . The winter precipitation is likely to increase. The seasonal variability is projected to rise in all seasons. The most sensitive season is summer.
- **Pressure at Sea Level:** no clear result, but slightly more than half of the models are showing a significant rise in sea level pressure and seasonal variability. The season with the strongest climate change signal is winter.
- **Air Temperature:** the increasing is significant at all considered pressure levels and gets stronger in higher levels of the troposphere. The seasonal variability is significantly increasing in winter and also in spring at the 300 hPa level. The most sensitive season is the summer.
- **Specific Humidity:** is significantly increasing in all seasons, scenarios and pressure levels. At high pressure levels the seasonal variability is increasing especially in winter and also in the others seasons at the 500 hPa and 300 hPa level. The most sensitive season is the summer.
- **Geopotential Height:** a significant rise is visible in all pressure levels and seasons. Most of the models show an increase in seasonal variability in winter. Summer is the most sensitive season.
- **Eastward Wind:** is increasing in winter and spring especially in high altitudes. The most sensitive seasons are also winter and spring. Approximately half of the GCMs show an increase of seasonal variability.
- **Northward Wind:** there is an increase in winter and a decrease during the other seasons. On the annual time scale there are nearly

no changes visible. The changes in monthly variability are unclear. Winter is the most sensitive season.

- **Lagrangian Tendency of Air Pressure:** the majority of GCMs project an increase especially during the transition periods (see section 3.2.8). More than half of the ensemble is showing a significant increase of seasonal variability. There is no season which has an especially strong climate change signal.

In general the significance of the climate change signals are changing more between the seasons and the pressure levels than between the domains and the emission scenarios.

It is possible to show that the top six overall high performing GCMs (concerning the model performance index annually) have partly a stronger climate change signal as the worst performing six GCMs. In table 3.1 the 36 scenario- parameter combinations are listed which differ in the mean climate change signal of the top and the worst six model ensembles by more than 90 % confidence level. 17 of them are also significant at a 95 % and 9 at the 97.5 % confidence level. The significant variables are mainly air temperature (*ta* and *tas*) and parameters which are related to temperature like specific humidity and geopotential height. Furthermore, also eastward wind speed and the lagrangian tendency of air pressure are showing significantly higher climate change signals. The variances of the climate change signals of the high and low performing GCM ensembles are showing no significant trends.

This means that a model with a good model performance index (which is based on a low spatial root mean square averaged over all parameters) shows for some parameters a higher climate change signal (e.g., a stronger increasing temperature in the 21st century) than a GCM with a bad MPI. For example, the differences between the high and low performing ensemble are typically between 0.4 K and 1.2 K for temperature.

However, linear correlation analyses between the MPI of the GCMs and their climate change signals are showing no significant results (see figure 3.37 and 3.37). This means that the general statement "well performing GCMs have different climate signals as worse performing GCMs" is not valid.

Uncertainty Estimation

The dominant uncertainty source in the climate change signal for all parameters on annually and seasonal basis and for both domains is the GCM formulation. In the 2021 to 2050 climatology more than two thirds and in the 2071 to 2100 climatology 60 % of the total variance (average over all parameters, annually in the large domain) are added by this uncertainty source (see table 4.3). By the end of the century the second highest value is carried by the emission scenario uncertainty with 15.6 % and followed by

the internal (initial condition) uncertainty with 12.3 %. The rest, which is missing to 100 %, is commonly small and is contributed by the combination terms of formula 2.31.

In table 4.3 it is visible that the results in both domains are only varying by a few percent. It is important that the emission scenario uncertainties are initiated by long term effects. In the first half of the 21st century they only contribute about three to four percent of the total variance. At the end of the century these values raise to over 15 % and get more important than the initial condition uncertainties. However, the contribution of uncertainty to the different groups is strongly depending on the regarded parameter. For example for air temperature the emission scenario uncertainty carries over 35 % of the total variance at the end of the century whereas this ratio is only 1 % for northward wind. A large contribution of the emission scenario uncertainty can also be found for geopotential height and specific humidity. The results from the ANOVA for all parameters, pressure levels and seasons are given in Table 4.5

Table 4.3: Average uncertainty of the climate change signal in percent for the groups GCM formulation, emission scenario and internal uncertainties for both climatologies and both domains.

Climatology	**Large Domain**			**Small Domain**		
	Emission	GCM	Internal	Emission	GCM	Internal
2021 to 2050	3.8	66.2	13.6	3.5	62.0	16.8
2071 to 2100	15.7	60.1	12.3	15.31	61.2	11.73

The major difference between the two domains is that the total amount of variance is always bigger on the small one. A possible reason for this is that a larger fraction is covered by topographically complex areas in the small domain as in the large one.

Comparing the two applied methods, the main statement, that the GCM formulation is the major source of uncertainty in the climate change signal of GCMs, stays the same. However, there are differences in the amount and temporal growing rate of internal uncertainties and emission scenario uncertainties. Rowell [2006], who used the uncertainty estimation technique in a paper in 2006 stated that this method: "... aims to provide an initial estimate of the relative importance of the uncertainty arising from RCM formulation, and compare it with the other sources of uncertainty ..." [Rowell, 2006]. Therefor, this method is helpful to get a quantitatively overview of the uncertainty sources, but the ANOVA technique can provide more reliable results.

The obtained results from the ANOVA are in good agreement with published findings. For example, Rowell [2006] studied the surface temperature and precipitation uncertainties of the PRUDENCE RCMs (see section 1.2.2)

for the United Kingdom and found that: "The largest source of uncertainty, for both variables and in all seasons, is the formulation of the global coupled model." [Rowell, 2006]. He also found out that the emission scenario uncertainties are almost negligible for the total precipitation uncertainty by the end of the 21st century in the UK.

In a study of Déqué et al. [2008] the PRUDENCE RCMs where used to investigate the uncertainty sources of the climate change signals for surface temperature with the ANOVA method over whole Europe. Their results were that: "The uncertainty introduced by the choice of the driving GCM is generally larger than the other three sources." [Déqué et al., 2007]. For the Alps they obtained the percentage of total variance which are diosplayed in tabular 4.4(reference period 1961 to 2000 and climate change 2071 to 2100).

Table 4.4: Comparison of the contribution to the total variance (in percent) by GCM formulation, the emission scenarios and the internal variability in the Alpine domain (Déqué) and the small domain. The 21 % of uncertainty from RCM formulation, which occurs in Déqué et al. [2007], was equally distributed to the other uncertainty sources.

Uncertainty	**Déqué et al. [2007]**	**Own results**
Emission Scen.	27.03	36.52
GCM	61.25	52.73
Internal	11.72	10.12
Total StD	1.0	0.82

Even if Déqué et al. [2007] were using RCM data and did not use the same setup of GCMs or emission scenarios the results are similar. There are no studies published to compare the obtained results for other parameters and pressure levels. Nevertheless, the good agreement for surface temperature and precipitation gives confidence to the obtained results.

In this study a further result is that the total amount of uncertainty is strongly increasing with time. In many cases the standard deviations of the climate change signals are more than doubling from the first half of the 21st century to the end. The gray shaded background in plots like 3.13 shows that the models are able to simulate the trend of the climate change in the 20th century before they start to diverge. This is a known phenomena, which is discussed in a paper of Khiel [2007]. Since there are no defined forcings for the 20th century and GCMs are evaluated by their performance in this period, the GCMs are trimmed to fit the observed data. So, GCMs are able to simulate the temperature increase of the last hundred years and differ only by 25 % (the range is approximately plus 0.5 K to 0.7 K). Considering the climate sensitivity forecasts (equilibrium temperature for doubling CO_2) the GCMs differ by a factor of two (the range lies between 1.5 K to 4 K). Kiehl found: "... that the total anthropogenic forcing for a wide range

of climate models differs by a factor of two and that the total forcing is inversely correlated to climate sensitivity." [Kiehl, 2007]. This means that the increase of uncertainty at the end of the 21st century is mainly caused by tuning the GCMs to the postindustrial climate change.

Table 4.5: Overview of the ANOVA results for all seasons and all parameters in the large domain. Shown are the percentage of the total variance of the GCM formulation, the emission scenario and the initial condition uncertainty. The missing percentages are carried by mixed terms, which in general are small and not listed in this table. In the column StdDev. the total standard deviation of the corresponding variables are shown. The standard deviations of *hus* and *wap* are too small, so the decimal power is written under the annual values (for *hus* e-9 and for *wap* e-3). Shown are the annual mean ANOVA values (big numbers) and the values of the different seasons (smaller numbers) as it is described on the bottom of the table.

	Uncertainty in climate signal 2021 to 2050												Uncertainty in climate signal 2071 to 2100											
	Emissions			GCM			Inter			StdDev.			Emissions			GCM			Inter			StdDev.		
tas	5.56	4.57 6.25	2.61 6	85	84.6 69.5	83.6 82.8	9.18	10.4 23.8	13.3 10.7	0.33	0.31 0.41	0.35 0.36	35.9	32.8 33	34.8 32.2	52.2	45.6 58.3	55.2 56	11.5	21.3 7.89	9.56 11.4	0.73	0.72 0.87	0.67 0.81
pr	0.32	0.17 2.22	0.62 0.28	84.7	86.3 81.3	81.6 89	14.9	13.4 16.3	17.7 10.9	0.1	0.12 0.1	0.1 0.11	1.21	2.21 6.38	2.05 2.61	83.4	70.9 89.8	72.6 85.1	15.2	26.7 3.7	25.2 12.2	0.13	0.15 0.16	0.12 0.14
psl	0.3	6.9 0.29	2.5 0.5	83	44 72	44 57	9.5	15 10	22 23	27.2	52 34.4	43.1 44.9	2.4	0.63 1.3	1.9 0.61	67	65 88	63 82	13	19 2.8	19 1.6	60.2	90.1 59.5	72.9 68.8
ta850	6.1	2.94 6.28	2.21 7.56	74.4	74.6 57.5	67.9 66.5	10.8	44.8 55.8	57.4 48.4	0.33	0.33 0.49	0.41 0.39	33.4	27.6 28.9	23 27.8	49.1	44.8 55.8	57.4 48.4	10.1	20.5 7.02	4.61 9.45	0.79	0.77 1.04	0.92 0.9
ta700	7.57	5.09 7.98	4.79 8.03	74.4	75.5 61.8	70.9 72.5	10.1	7.92 22.4	10.4 10.6	0.33	0.32 0.42	0.33 0.37	37	33.5 33.3	39.9 36.4	42.8	39.9 49	41.2 44.7	12.5	19.3 9.41	10.3 11	0.82	0.79 0.99	0.74 0.89
ta500	8.65	7.16 8.88	6.61 8.84	76.2	75.3 67.3	75.2 76.8	8.21	8.17 16.4	7.03 6.99	0.38	0.37 0.44	0.36 0.42	36.1	33.5 35.3	36.9 36.7	41.8	41.3 44.7	41.8 42.3	14.7	18.5 12.2	13.9 13.1	0.95	0.93 1.05	0.88 1.04
ta300	8.95	7.87 7.79	7.91 9.57	72.9	75.5 65.6	74.3 75.1	12.1	8.31 20.4	10.2 8.6	0.41	0.37 0.53	0.38 0.45	32.7	29 32	30.3 34.8	49.2	52.1 52.3	52.9 45.7	10.9	12.8 7.87	10.4 11.5	1.04	0.98 1.22	0.98 1.14
zg850	7.39	8.63 3.22	5.56 1.05	66.2	48.1 53.2	41.8 71	14.7	12.2 2.9	18.9 13	3.24	5.93 5.49	4.26 4.73	15.2	12.3 4.77	19.1 13.5	71.4	55.8 66.8	54.8 68	6.51	11.1 4.73	10.9 3.47	6.85	9.14 8.03	7.95 7.82
zg700	7.28	5.68 4.86	5.36 2.82	73.1	68.3 67.8	62.2 78.5	12.2	10.3 15.5	13.9 10	5.41	7.21 5.85	5.92 6.5	25.3	20.1 24.5	28.3 27	61.7	64.3 61.2	55.3 57.8	6.11	8.78 7.72	6.97 6.1	10.9	11.8 11.8	11.1 12
zg500	7.36	5.68 5.98	5.62 4.14	76.6	75 69.8	71.6 81.4	10.1	8.04 15.7	10 8.26	9.52	10.9 10.3	9.7 10.9	31.1	29 30.9	36.4 32.6	53.9	54.1 53.7	48.3 51.8	8.09	10.7 8.86	7 8.17	19.5	19.4 21.7	18.5 21.6
zg300	7.88	6.07 6.68	6.45 5.39	76.9	77.6 69.9	74.5 81.6	9.66	7.66 16.1	8.88 7.52	16.2	17.2 18	15.6 18.3	32.9	32.9 33.6	38.9 34.9	50.3	48.2 49.9	44.2 47.9	9.87	12.6 9.71	9.24 10	34.9	33.3 38.9	31.6 38.7
hus850	0.14	1.58 1.75	1.94 2.65	75.7	83.1 87.3	84.2 84	9.27	4.58 4.51	2.49 5.28	3.9 e-9	2.36 5.07	2.59 4.08	17.8	24 11	19.7 19.2	64.6	51.9 73.5	58.2 65.6	11.9	16.2 9.91	15.9 8.93	7.35 e-9	4.85 12.5	5.29 8.42
hus700	4.2	3.61 3.59	2.4 4.67	85.8	83.4 85.1	83.3 84	2.36	2.94 4.46	1.77 1.96	1.84 e-9	1.3 3.27	1.48 2.01	21.8	25.3 16.3	22.6 22.1	58.9	47.5 67.6	57.8 58.6	13.6	19.9 10	13.6 13.2	4.89 e-9	3.17 8.49	3.63 5.39
hus500	7.29	5.27 5.13	6.21 8.96	79.3	82 79.4	76.6 75.7	4.23	3.3 7.06	4.22 3.12	0.7 e-9	0.61 1.12	0.57 0.77	23.8	24.1 19.9	24.1 23.4	54.6	54.9 59.6	55.7 53.9	13.8	13.2 12	12.5 14.4	2.19 e-9	1.69 3.25	1.79 2.51
hus300	5.46	3.7 4.35	4.3 6.67	79.6	81.8 79.4	79.2 79.3	7.38	7.7 7.87	8.35 5.47	0.15 e-9	0.11 0.24	0.12 0.18	21.7	21.8 19	19.4 22.6	59.9	62.4 62.4	64.7 58.3	9.27	6.65 9.27	7.04 10.2	0.47 e-9	0.3 0.73	0.35 0.57
ua850	0.16	0.58 0.04	0.45 0.15	48.2	32.1 17.5	25.9 60.5	28.9	25.9 59.3	53.1 6.54	0.13	0.24 0.18	0.22 0.2	0.28	3.77 1.4	0.66 1.55	58.2	59.7 30.1	52.9 50.3	29	5.13 52.9	18.4 30	0.19	0.3 0.25	0.25 0.23
ua700	0.2	0.57 0.24	0.65 0.32	58.1	44.2 19.1	22.9 68.6	23.3	24.6 66.5	52.9 5.14	0.15	0.28 0.2	0.23 0.23	2.73	5.35 0.79	4 0.57	60.5	64 32.2	53 63.3	23.9	9.04 50.2	19.2 19.1	0.24	0.39 0.32	0.29 0.28
ua500	0.54	0.98 0.09	0.36 0.65	64.6	55.3 19.7	27 64.7	15.6	15.2 64.4	43.9 8.6	0.19	0.38 0.25	0.29 0.31	7.54	7.27 2.79	7.92 1.85	61.1	68.4 33.3	55.5 71.3	18.1	8.2 46	15.6 10.5	0.33	0.58 0.41	0.39 0.39
ua300	0.81	1.09 0.03	0.06 0.81	67.1	64.4 27.9	35.9 59	14.3	10 52.8	35.9 15.6	0.28	0.56 0.31	0.43 0.46	13	10.5 4.68	10.4 4.14	58.4	66.3 42.8	54.3 73.9	14.8	7.97 31.7	17.2 5.94	0.47	0.85 0.53	0.58 0.55
va850	0.79	2.42 2.1	4.89 2.33	59	52.3 50.6	56.9 43	15.9	11.5 22.1	8.35 19.4	0.09	0.21 0.1	0.15 0.13	0.81	3.25 1.63	0.24 1.05	69.5	60.1 68.2	78 40.3	16.8	23 7.16	2.87 28.8	0.15	0.34 0.12	0.19 0.16
va700	1.05	1.88 1.71	3.87 2.72	54.9	53.1 51.6	60.6 51.6	7.81	9.93 10.2	6.66 12.6	0.12	0.27 0.13	0.19 0.17	1.33	4.35 1.8	0.16 1.55	73.1	59.3 64.8	78.2 46.4	11.6	23.5 11.3	2.32 19	0.17	0.41 0.16	0.24 0.2
va500	0.87	1.05 1.58	3.13 3.14	51.3	51.5 59.5	60.6 46.9	5.13	7.52 7.92	6.25 11.2	0.19	0.41 0.15	0.28 0.22	1.91	5.2 2.19	0.14 2.57	73.5	55.3 64.7	77.4 52.4	9.81	26.2 14	2.03 12.2	0.22	0.54 0.22	0.32 0.26
va300	0.62	0.5 0.15	2.11 2.95	48.1	50.7 61.9	60 41.2	4.33	4.35 6.49	5.44 9.25	0.32	0.85 0.19	0.44 0.32	2.35	5.49 2.1	0.06 4.15	76.9	56.5 63.8	78.2 55.5	5.55	23.9 18.2	1.57 8.49	0.29	0.69 0.3	0.44 0.35
wap850	2.74	3.1 2.2	2.9 0.26	25.2	39 51	29 38	43.4	28 23	36 31	0.59 e-3	0.98 0.7	0.73 0.81	5.2	0.87 2.93	8.92 5.69	51	65.4 64.1	45.5 46.3	13.1	13.9 7.76	16.2 9.68	0.97 e-3	1.69 0.96	1.03 1.2
wap700	3.12	4.3 0.6	3.1 0.34	40	47 64	43 42	27.5	21 3.8	22 25	0.61 e-3	1.27 0.86	0.87 0.88	8.3	0.89 4.72	8.77 8.91	54.9	68.1 52.8	58 43.6	8.13	13.2 12.8	9.2 10.5	0.96 e-3	1.96 1.38	1.14 1.19
wap500	3.05	5.6 0.71	2.6 0.84	53.8	49 71	52 44	18	18 3.1	16 23	0.72 e-3	1.58 1.06	1.03 1.03	7.17	1.08 4.14	7.34 7.55	62	66.1 63.4	63.8 48.4	6.7	14.9 8.02	5.62 11.9	1.11 e-3	22.2 1.83	1.36 1.31
wap300	2.89	6.6 0.57	2 1.5	53.8	56 67	58 44	17.9	20 7.7	16 24	0.68 e-3	1.43 1.04	0.94 0.9	5.75	1.48 3.78	6.43 5.2	63.4	62.7 65.9	63.4 55.6	7.03	17.1 5.51	6.24 12.5	1.10 e-3	2.01 1.96	1.26 1.22
Means:	**3.75**	**3.84** **3.16**	**3.38** **3.45**	**66.22**	**63.3** **60.3**	**59.4** **65.1**	**13.58**	**13.4** **21**	**18.9** **13.5**				**15.7**	**14.7** **13.4**	**16** **15.2**	**60.12**	**57.4** **58.5**	**58.6** **56.1**	**12.28**	**15.7** **14.4**	**10.8** **12**			

Annual	**DJF**	**MAM**
	JJA	**SON**

Chapter 5

Conclusion

The results of this study indicate that the choice of the forcing GCM is essential for regional climate modeling in the Alpine region. The variance in the climate change signal (averaged over a broad spectra of parameters) is caused by over (60 ± 10) % by the GCM formulation. The initial condition uncertainty has a typical value of (12 ± 5) % and the emission scenario uncertainty (4 ± 3) % in the middle and (16 ± 13) % at the end of the 21st century. Therefor, the emission scenario selection is only important for long term simulations and for specific parameters (for *tas*, *ta*, *zg* and *hus* it has a mean of 28 % and for *pr*, *psl*, *ua*, *va* and *wap* the mean is 4 %).

The GCM evaluation for the 21th century showed that there is a large spread in model quality and that there are some models, which perform better than others, even if there is no perfect model. A special case is the multi model mean, which outperforms even the best GCMs. However, for the reliability of future climate projections it is important, but not sufficient that GCMs have a good performance of the contemporary climate [McAvaney et al., 2001]. It is also important to look at the climate sensitivity of the models [Raper et al., 2001]. Even if it is possible to show that for some parameters (temperature, specific humidity, geopotential height, eastward wind and vertical air movement) an ensemble of well performing GCMs are showing a higher climate change signal at the end of the 21st century than a ensemble of worse performing GCMs, linear correlation analyses showed that there is no general connection between the performance of a GCM and the strength of its climate change signal.

Because even well performing GCMs are showing a broad spread of future climate change signals, the most accurate way to make regional climate projections is to use an ensemble of good performing GCMs, which differ in their climate sensitivity and future projections.

Bibliography

[Barker et al., 2007] Barker, T., L. Bernstein, P. Bosch, Osvaldo C., Z. Chen, R. Christ, O. Davidson, W. Hare, S. Huq, D. Karoly, V. Kattsov, Z. Kundzewicz, J. Liu, U. Lohmann, M. Manning, T. Matsuno, B. Menne, B. Metz, M. Mirza, N. Nicholls, L. Nurse, R. Pachauri, J. Palutikof, M. Parry, D. Qin, N. Ravindranath, A. Reisinger, J. Ren, K. Riahi, C. Rosenzweig, M. Rusticucci, S. Schneider, Y. Sokona, S. Solomon, P. Stott, R. Stouffer, T. Sugiyama, R. Swart, D. Tirpak, C. Vogel, and G. Yohe (2007), Climate Change 2007: Synthesis Report, Summary for Policymakers, pp. 52, *http://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf*

[Charney et al., 1979] Charney G. J., A. Arakawa, D. J. Baker, B. Bolin, R. E. Dickinson, R. M. Goody, C. E. Leith, H. M. Stommel, and C. I. Wunsch (1979), Carbon Dioxide and Climate: A Scientific Assessment, report, pp. 22, Climate Research Board, Washington D.C.

[Christensen and Christensen, 2007] Christensen H. J., and B. O. Christensen (2007), A summary of the PRUDENCE model projections of changes in European climate by the end of this century, *Climatic Change*, Vol. *81*, pp. 7-30.

[Cebon et al., 1998] Cebon P., U. Dahinden, H. Davis, D. M. Imboden, and C. C. Jaeger, 1998, *Views from the Alps, Regional Perspective on Climate Change*, pp. 515, MIT Press, Camebridge, Massachusetts, London, England.

[Collins et al., 2001] Collins M., S. F. B. Tett, and C. Cooper (2007), The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustment, *Climate Dynamics*, Vol. *17*, pp. 61-81.

[Davies, 1976] Davies H. C. (1976), A lateral boundary formulation for multi-level prediction models, *Quarterly Journal of the Royal Meteorological Society*, Vol. *102*, pp. 405-418.

[Déqué et al., 2007] Déqué M., D. P. Rowell, D. Lüthi, F. Giorgi, J. H. Christensen, B. Rockel, D. Jacob, E. Kjellström, M. de Castro, and

B. van den Hurk (2007), An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, *Climatic Change*, Vol. *81*, pp. 53-70.

[Dickenson et al., 1989] Dickenson, R. E., R. M. Errico, F. Giorgi, and G. T. Bates (1989), A regional climate model for western United States, *Climate Change*, Vol. *15*, pp. 383-422.

[Easterling and Stern, 1999] Easterling, E. W. , and C. W. Stern (1999), *Making Climate Forecasts Matter*, 192 pp., National Academy Press, Washington DC.

[Eckey et al., 2002] Eckey H. F., R. Kosfeld, and M. Rengers (2002), *Multivariate Statistik, Grundlagen - Methoden - Beispiele*, 442 pp., Gabler, Wiesbaden.

[ECMWF, 2009] ECMWF European Centre for Medium-Range Weather Forecasts, http://www.ecmwf.int, (September 22, 2009)

[Efron and Tibshirani, 1993] Efron B., and J. R. Tibshirani (1993), *An Introduction to the Bootstrap*, pp. 436, Chapman & Hall, New York.

[Etling, 2008] Etling D. (2008), *Theoretische Meteorologie: Eine Einführung*, Vol. 3, pp. 376, Springer, Berlin.

[Geer, 1996] Gerr W. I. (1996), *Glossary of Weather and Climate, with Related Oceanic and Hydrologic Terms*, pp. 272, American Meteorological Society, Boston.

[Giorgi and Bates, 1989] Giorgi F., and G. T. Bates (1989), The climatological skill of a regional model over complex terrain, *Monthly Weather Review*, Vol. *117*, pp. 2325-2347.

[Giorgi and Mearns, 2002] Giorgi, F., and L. O. Mearns (2002), Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, *Journal of Climate*, Vol. *15*, pp. 1141-1158.

[Giorgi and Mearns, 1991] Giorgi F., and L. O. Mearns (1991), Approaches to regional climate change simulation: A review, *Reviews of Geophysics*, Vol. *29*, pp. 191-216.

[Gleckler et al., 2007] Gleckler P. J. , K. E. Taylor, and C. Doutriaux (2007), Performance metrics for climate models, *Journal for Geophysical Research*, Vol. *113*, D06104, doi:10.1029/2007JD008972.

[Goosse et al., 2008] Goosse H., P. Y. Barriat, W. Lefebvre, M.F. Loutre, and V. Zunz (2008), *Introduction to climate dynamics and climate modeling*, Online textbook available at http://www.climate.be/textbook.

[IPCC, 2001] Houghton, J.T.,Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.) (2001), Climate change 2001: the scientific basis, contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, pp. 881, Cambridge University Press, Cambridge, United Kingdom and New York, NY

[IPCC, 2007] Solomon S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (eds.) (2007), IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 996, Cambridge University Press, Cambridge, United Kingdom and New York, NY.

[Jacob et al., 2007] Jacob, D., L. Barring, J. H. Christensen, M. de Castro, M. Deque, F. Giorgi, S. Hagemann, G. Lenderink, B. Rockel, E. Sanchez, C. Schaer, S. I. Seneviratne, S. Somot, A. van Ulden, B. van den Hurk (2007), An inter-comparison of regional climate models for Europe: model performance in present-day climate, *Climatic Change*, Vol. *81*, pp. 31-52.

[Jones et al., 1976] Jones R. G., J. M. Murphy, and M. Noguer (1976), Simulation of climate change over Europe using a nested regional-climate model. I: Assessment of control climate, including sensitivity to location of lateral boundaries, *Quarterly Journal of the Royal Meteorological Society*, Vol. *121*, pp. 1413-1449.

[Ki Min and Hensen, 2007] Ki Min S., and A. Hense (2007), Hierarchical evaluation of IPCC AR4 coupled climate models with systematic consideration of model uncertainties, *Climate Dynamics*, Vol. *29*, doi:10.1007/s00382-007-0269-2.

[Kiehl, 2007] Kiehl, T. J. (2007), Twentieth century climate model response and climate sensitivity, *Geophysical Research Letter*, Vol. *34*, L22710, doi:10.1029/2007GL031383

[Kim and Reichler, 2008] Kim J., and T. Reichler (2008), Regional Performance of the IPCC-AR4 Models in Simulating Present-Day Mean Climate, EGU General Assembly, Vienna, Austria, http://www.inscc.utah.edu/j̃kim/, April 13-18.

[Klonk and Klein Tank, 2008] Klok E. J., and A. M. G. Klein Tank (2008), Updated and extended European dataset of daily climate observations, *International Journal of Climatology*, vol. *29*, issue 8, pp. 1182-1191. (2004

[Kraus, 2004] Kraus H. (2004), Die Atmosphäre der Erde, Eine Einführung in die Meteorologie, pp. 422, Springer, Berlin, Heidelberg.

[Leonhart, 2004] Leonhart R. (2004), *Lehrbuch Statistik, Einstieg und Vertiefung*, 496 pp., Hans Huber, Bern.

[Liljequist and Cehak, 2006] Liljequist G. H., and K. Cehak (2006), Allgemeine Meteorologie, Vol. 3, pp. 412, Springer, Berlin.

[Lorenz and Jacob, 2005] Lorenz P., and D. Jacob (2005), Two-way coupling of global and regional climate models, Max Planck Institute for Meteorology, Hamburg, Germany, RCM Workshop, Boulder.

[Maxino et al., 2008] Maxino C. C., B. J. McAvaney, A. J. Pitman, and S. E. Perkins (2008), Ranking the AR4 climate models over the Murray-Darling Basin using simulated maximum temperature, minimum temperature and precipitation, *International Journal of Climatology*, Vol. *28*, pp. 1097-1112.

[McAvaney et al., 2001] McAvaney, B.J, C.Covey, S. Joussaume, V. Kattsov, A. Kitoh, W. Ogana, A.J. Pittman, A.J. Weaver, R.A. Wood, and Z.C. Zhao (2001), *Model Evaluation. Climate Change 2001: The Scientific Basis*, Chapter 8, pp. 471-524, *Cambridge University Press*, Cambridge, United Kingdom and New York, NY.

[McGregor, 1996] McGregor J. L. (1996), Regional Climate Modeling, *Meteorology and Atmospheric Physics*, Vol. *63*, pp. 105-117.

[Meehl et al., 2007] Meehl G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bulletin of the American Meteorological Society*, Vol. *88*, pp. 1383-1394.

[Murphy et al., 2004] Murphy J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, Vol. *430*, pp. 768-772.

[PCMDI, 2009] PCMDI Program for Climate Model Diagnosis and Intercomparison (2009), www-pcmdi.llnl.gov, (September 22, 2009).

[Perkins and Pitman, 2008] Percins S. E., and A. J. Pitman (2008), Do weak AR4 models bias projections of future climate changes over Australia?, *Climate Change*, Vol. *93*, pp. 527-558.

[Pierc et al., 2009] Pierc D. W. , T. P. Barnett, B. D. Santer, and P. J. Gleckler (2009), Selecting global climate models for regional climate change studies, *Proceedings of the National Academy of Sciences*, Vol. *106*, pp. 8441-8446.

[Räisänen, 2007] Räisänen J. (2007), How reliable are climate models?, *Tellus*, Vol. *59A*, pp. 2-29

[Raper et al., 2001] Raper S. C. B., J. M. Gregory, and T. J. Osborn (2001), Use of an upwelling-diffusion energy balance climate model to simulate and diagnose A/OGCM results, *Climate Dynamics*, Vol. *17*, pp. 601-613.

[Reichler and Kim, 2008a] Reichler T., and J. Kim (2008a), How Well do Coupled Models Simulate Today's Climate?, *Bulletin of the American Meteorological Society*, Vol. *89*, pp. 303-311.

[Reichler and Kim, 2008b] Reichler T., and J. Kim (2008b), Uncertainties in the climate mean state of global observations, reanalyses, and the GFDL climate model, *Journal of Geophysical Research*, Vol. *113*, 2008

[Roger, 2006] Roger S. (2006), http://www.meteo.physik.uni-muenchen.de/r̃oger/Einfuehrung_Teil_II/Teil_II_L09_SS2006.pdf, (September 22, 2009)

[Rowell, 2006] Rowell R. D. (2006), A demonstration of the uncertainty in projections of UK climate change resulting from regional model formulation, *Climatic Change*, Vol. *79*, pp. 243-257.

[Sachs and Hedderich, 2006] Sachs L., and Hedderich J. (2006), *Angewandte Statistik, Methodensammlung mit R*, vol. 12, 702 pp., Springer, Berlin Heidelberg.

[Nakicenovic et al., 2007] Nakicenovic N., J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grübler, T. Y. Jung, T. Kram, E. L. La Rovere, L. Michaelis, S. Mori, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Riahi, A. Roehrl, H.-H. Rogner, A. Sankovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor, and Z. Dadi (2007) *Special Report on Emissions Scenarios*, Cambridge University Press, Cambridge, United Kingdom and New York, NY

[Sterl, 2004] Sterl A. (2004), Notes and correspondence On the (In)Homogeneity of Reanalysis Products, *Journal of Climate*, Vol. *17*, pp. 3866-3873.

[Storch et al., 2000] Storch H., H. Langenberg, and F. Feser (2000), A Spectral Nudging Technique for Dynamical Downscaling Purposes, *Monthly weather review*, Vol. *128*, pp. 3664-3673.

[Storch et al., 1999] Storch H., S. Güss, and M. Heimann (1999), *Das Klimasystem und seine Modellierung, Eine Einführung*, pp. 255, Springer, Berlin Heidelberg

[Storch and Zwiers, 1999] Storch H., and W. Zwiers (1999), *Statistical Analysis in Climate Research*, pp. 484, Cambridge University Press, Cambridge

[Taylor, 2000] Taylor E. K. (2000), *Summarizing Multiple Aspects of Model Performance in a Single Diagram*, PCMDI Report, pp. 29, Vol. **55**, Program for climate model diagnosis and intercomparison University of California, Lawrence Livermore National Laboratory, Livermore California.

[Noda et al., 2008] Noda A., S. Kusunoki, and M. Nakamura (2008), Development of Super High Resolution Global and Regional Climate Models, Annual Report of the Earth Simulator Center April 2007 - March 2008, pp. 271-279, Japan Meteorological Agency, University of Tokyo Download: http://www.jamstec.go.jp/esc/publication/annual/annual2007/index.html, (September 22, 2009).

[Trenberth, 1992] Trenberth E. K. (1992), Climate System Modeling, pp. 818, Cambridge University Press, Cambridge.

[Ulden and Oldenborgh, 2006] Ulden A. P., and G. J. Oldenborgh (2006), Large-scale atmospheric circulation biases and changes in global climate simulations and their importance for climate change in Central Europe, *Atmospheric Chemistry and Physics*, Vol. *6*, pp. 863-881.

[Uppala et al., 2005] Uppala, S. M., P. W. Kallberg , A. J. Simmons, U. Andrae, V. D. C. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. J. Hoskins, L. Isaksen, P. A. E. M. Janssen, Jenne R., A. P. McNally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P.

Viterbo, and J. Woollen (2005), The ERA-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society*, Vol. *131*, pp. 2961-3012.

[Waldron et al., 1996] Waldron K. M., J. Peagle, and J. D. Horel (1996), Sensitivity of a spectrally filtered and nudged limited area model to outer model options, *Monthly weather review*, Vol. *124*, pp. 529-547.

[Wang et al., 2004] Wang Y., L. R. Leung, J. L. McGregor, D. Lee, W. Wang, Y. Ding, and F. Kimura (2004), Regional Climate Modeling: Progress, Challenged, and Prospects, *Journal of the Meteorological Society of Japan*, Vol. *82*, No. 6, pp. 1599-1628.

[Weigel, 2008] Weigel A. P.,M. A. Liniger, and C. Appenzeller (2008), Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?, *Quarterly Journal of the Royal Meteorological Society*, Vol. *134*, pp. 241-260.

[WGNE, 2009] WGNE at PCMDI, *WGNE Standard Diagnostics of Mean Climate*, http://www-pcmdi.llnl.gov/projects/amip/OUTPUT/WGNEDIAGS/, (September 22, 2009)

List of Figures

1.1 Overview of the general development of a climate Model [Goosse et al., 2008]. 4
1.2 Short history of GCM development. 6
1.3 Uncertainty propagating from emission scenarios over internal, GCM and RCM uncertainties to climate impact model uncertainties. The red frame marks the part of the "uncertainty chain", which is the topic of this thesis. 7
1.4 Model Performance Index I^2 for different model generations (rows) and different models (circles). The better the model the smaller the index. The diameter of the circles indicate the 95% confidence interval. Numbers and letters mark different models. Models marked with a red circle use flux correction. Grey circles show the average of all model performance indices (MPIs) in one model generation whereas black circles mark the MPI of the multi-model-mean of one model generation. The green circle REA belongs to the NCEP/NCAR reanalysis. In the last row the preindustrial control experiment (PICTRL) of the CMIP-3 project is displayed [Reichler and Kim, 2008a]. 9
1.5 The development of components in GCMs since the mid 70th until the fourth IPCC Assessment reports [IPCC, 2007]. (FAR (1990): First Assessment Report -, SAR (1995): Second Assessment Report -, TAR (2001): Third Assessment Report -, AR4 (2007): Fourth Assessment report of the IPCC). 10
1.6 Europe for different model resolutions: T21 (ca. 500 km), T42 (ca. 250 km), T63 (ca. 180 km), T106 (ca. 100 km) [IPCC, 2007]. . . . 11
1.7 Summary of the data availability at the PCMDI. Colored boxes indicate that some variables (not necessarily all) are stored in the multi model dataset while shaded colors mean that there is an ensemble of runs for the corresponding model and experiment. The meaning of the colors is described in the legend [PCMDI, 2009]. . . 15

2.1 Monthly mean data available at the PCMDI data portal [PCMDI, 2009]. 22
2.2 Station network of the E-OBS Dataset (August 2007) [Klok, 2007]. 27

2.3 Normalized root mean square (RMS) errors of reanalysis/GCM data validated with the mean of multiple observations and averaged over all "physics" and "dynamics" quantities for different seasons (DJF - winter, MAM - spring, JJA - summer, SON - autumn, AVG - annual) and regions (NH - northern hemisphere, TR - tropes, SH - southern hemisphere, GL - global). (NCEP/NCAR (NNR), NCEP/DOE (NDR), ERA-40 (ERA), JRA-25 (JRA) and GFDL CM2.1 climate model (GFD)). Since the ERA 40 data were used as reference for the dynamic parameter analyze there are no results of the evaluation of these data (gray shaded boxes). [Reichler and Kim, 2008b]. 28

2.4 The two investigated domains (red: S=32°, N=60°, E=345°, W=35°), (green: S=38°, N=52°, E=0°, W=22°). 30

2.5 The azimuthal position of the simulated field (test) represents the correlation coefficient R. The distance between the origin and the test field is the variance of σ_s and the distance from the origin to the reference field is σ_o. The RMS E' is represented by the distance from the reference field to the simulated field [Taylor, 2000]. 34

2.6 Theoretical structure of a box-whiscer plot. 35

3.1 Spatial mean of the large domain annual cycle for *tas*, *pr*, *psl*, *ta* and *hus* in the 1961 to 2000 climatology. Thin collored lines are different GCMs while the thick red line is corresponding with the reference data. 49

3.2 Same as figure 3.1 for *zg*, *ua*, *va*, and *wap*. The legend is the same as in figure 3.1. 50

3.3 Spatial bias of the multi model mean (left), the iap_fgoals1_0_g (upper right) and the ukmo_hadgem1 (lower right) GCM minus the E-OBS reference dataset annually for *tas* in the 1961 to 2000 climatology. 51

3.4 Seasonal bias for *pr* in the multi model mean field. 52

3.5 Taylor diagram for *tas* (red symbols), *pr* (blue symbols) and *psl* (yellow symbols) for the annually climatology in the large and small domain. Displayed are the spatial RMS error, the spatial correlation coefficient and the centered spatial variability. 53

3.6 Taylor diagrams for *ta* (red), *hus* (yellow), *zg* (blue), *ua* (light blue), *va* (violet) and *wap* (green) in the pressure levels 850 hPa, 700 hPa, 500 hPa and 300 hPa for the annual climatology in the large domain. 54

3.7 The MPI for the large domain, annually and for all seasons. Each circle is corresponding to a GCM from the legend. The colors and sizes of the circles are showing the uncertainty range (95 % confidence interval) of the single MPIs. The x-axis is scaled logarithmically so that differences in small MPIs are better visible. The circle with the red edging shows the multi model mean performance (model y). 55
3.8 Same as figure 3.7, but for the small domain. 56
3.9 The performance indices for *ta*, *hus*, *zg*, *ua*, *va* and *wap* in the large domain, annually on 850 hPa. For reasons of visibility the radii of the circles are constant. The uncertainty of the values is displayed as filling color. 57
3.10 Overview of the MPIs for different GCMs, domains and seasons. . 58
3.11 Performance indices for all GCMs and parameters on the large domain annually. 59
3.12 Overview of the model performance indices for all variables, domains and seasons. Different GCMs are plotted in the rows whereas different parameters are plotted in the columns. The outer right column shows the mean performance over all parameters. Blue colors are corresponding to good MPIs, whereas red colors represent bad performing models. The partition of the squares is explained in the lower right corner. 60
3.13 Temporal development of two meter surface temperature *tas* from 1960 to 2100. 63
3.14 Box plots of surface temperature *tas* for all domains, climatologies, seasons and experiments. 64
3.15 Significance analyses of the climate change signal (left side) and changes in inter-seasonal variability (right side) for *tas*. The upper side of the plot shows the climate change signals for 2021 to 2050 and the lower part those for 2071 to 2100. Different bars are showing the annual mean climatologies and different seasonal climatologies. Each bar is separated in boxes, where the three columns are presenting the three emission scenarios B1, A1B and A2 and the 24 rows show different GCMs. Each box is divided into two triangles where the upper one shows the significance of the climate change signal of the large domain and the lower one the significance of the small domain corresponding to the description on the right bottom. Red colors are indicating a significant higher mean or monthly variance in *tas* and blue colors are showing a significant decrease in mean or monthly variance of *tas*, respectively. 65
3.16 Temporal development of the spatial mean monthly precipitation in scenario B1, A1B and B2 on the large and small domain. 66
3.17 Same as figure 3.14 but for precipitation flux pr. 67
3.18 Same as figure 3.15 but for pr. 68

3.19 Temporal development of the spatial mean yearly sea level pressure in scenario B1, A1B and B2 in the large and small domain. 69

3.20 Same as figure 3.14 but for sea level pressure *psl*. 70

3.21 Same as figure 3.15 but for pressure at sea level. 71

3.22 Mean air temperature (*ta*) rise at 850 and 300 hPa for the large domain. 71

3.23 Significance in the change of mean air temperature at the four different pressure levels, the three different emission scenarios and the two different climatologies in the small and large domain. 72

3.24 Same as figure 3.23 but for changes in monthly variance of *ta*. . . . 73

3.25 Changing of specific humidity in the spatial average of the small domain at the 850 hPa and 300 hPa level. 74

3.26 Same as figure 3.24 but for *hus*. 75

3.27 Temporal development of the mean geopotential height on the large domain. 76

3.28 Same as figure 3.23 but for *zg*. 77

3.29 Temporal changing in the spatial mean state of eastward *ua* and northward *va* wind in the large domain at 850 hPa and 300 hPa. . 78

3.30 Box whisker plots for the eastward (left side) and northward wind velocity (right side) for all domains, seasons and scenarios at the 300 hPa pressure level. 79

3.31 Test for significant changes in the mean horizontal wind fields *ua* (upper half) and *va* (lower half) in the end of the 21ct for all pressure levels, domains and scenarios. 80

3.32 Temporal development of the lagrangian tendency of air pressure *wap* in the large domain. 81

3.33 Box Whisker Plot of the climate change signal of *wap*. 82

3.34 Same as figure 3.23 but for *wap*. 83

3.35 Same as figure 3.24 but for *wap*. 84

3.36 Relative differences in percent between ensemble-mean climate change signal and GCM climate change signals for all GCMs, parameters, emission scenarios and both domains and climatologies. Red colors indicate stronger climate change signals, blue colors are corresponding with weak climate change signals. Rows in the plot are different GCMs, columns display different parameters and pressure levels. The subdivision of single rectangles is explained in the lower right corner. The mean climate sensitivity is displayed in the last column (equally weighted average over all parameters). 86

3.37 Analyze of the rank correlation between the performance indices of GCMs and their climate sensitivity for the parameters *tas*, pr, *psl*, *hus* and *zg*. The last three are analyzed at four pressure levels (850 hPa, 700 hPa, 500 hPa and 300 hPa). Different colors and symbols are showing different emission scenario experiments (black (cross) is B1, green (triangle) is A1B and red (box) is A2). The colored lines are representing the linear fit of the data. The coefficient of determination (r^2) is written on the right top of each subplot. . . . 87
3.38 Same as figure 3.37 but for *ua*, *va* and *wap*. 88
3.39 Differences in the uncertainty estimation for *tas* in both climatologies (2021 to 2050 and 2071 to 2100) for the original data matrix (upper two graphs) and the artificially filled data matrix (lower two graphs) in the large domain. The two graphs on the left hand side are showing the climate change signal for the 2021 to 2050 climatology whereas the plots on the right are showing the 2071 to 2100 climatology. The colored symbols are corresponding to the GCMs (see legend). The standard deviation (SD) of the three uncertainty groups are written on the bottom of each plot. 92
3.40 Temperature at surface ANOVA for both climatologies, domains and on annual and seasonal basis. The plots on the left side are showing percentages of the total uncertainty, whereas those on the right hand side are presenting the absolute values of variance. In the right figures also the standard deviation (SD) is written above the bars. 93
3.41 Same as figure 3.39 but for pr. 94
3.42 Same as the right plots in figure 3.40 but for pr. 95
3.43 Same as upper plots in figure 3.39 but for *psl*. 96
3.44 Same as right plots in figure 3.40 but for *psl*. 96
3.45 Same as upper plots in figure 3.39 but for *ta* at the 850 hPa and the 300 hPa pressure level. 97
3.46 Same as the right plots in figure 3.40 but for *ta* at the 850 hPa and the 300 hPa pressure level. 98
3.47 Same as the right plots in figure 3.40 but for specific humidity (*hus*) on the 850 hPa and the 300 hPa pressure level. 99
3.48 Same as the right plots in figure 3.40 but for *hus* on the 850 hPa and the 300 hPa pressure level. 100
3.49 Same as upper plots in figure 3.39 but for geopotential height (*zg*) at the 850 hPa and the 300 hPa pressure level. 101
3.50 Same as the right plots in figure 3.40 but for geopotential height (*zg*) at the 850 hPa and the 300 hPa pressure level. 102
3.51 Same as upper plots in figure 3.39 but for eastward wind (*ua*) at the 850 hPa and the 300 hPa pressure level. 103
3.52 Same as the right plots in figure 3.40 but for eastward wind at the 850 hPa and the 300 hPa pressure level. 104

3.53 Same as upper plots in figure 3.39 for northward wind (va) at the 850 hPa and the 300 hPa pressure level. 105
3.54 Same as the right plots in figure 3.40 but for northward wind at the 850 hPa and the 300 hPa pressure level. 106
3.55 Same as upper plots in figure 3.39 but for lagrangian tendency of air pressure (wap) at the 850 hPa and the 300 hPa pressure level. . 107
3.56 Same as the right plots in figure 3.40 for lagrangian tendency of air pressure at the 850 hPa and the 300 hPa pressure level. 108

List of Tables

1.1 Summary of the sources of uncertainty [Rowell, 2006]. 17

2.1 Model overview of the CMIP3 GCMs part I [PCMDI, 2009]. 20

2.2 Model overview of the CMIP3 GCMs part II [PCMDI, 2009]. . . . 21

2.3 Climate variables, pressure level (for *tas*, *pr* and *psl* at ground level, others at 850, 700, 500 and 300 hPa) and corresponding reference data, which are used for the analyses. 23

2.4 Average daily counts of various types of observation supplied to the ERA40 data assimilation, for five selected periods [Uppala,2005]. . 29

3.1 Parameters with significant differences in the climate change signal (reference: 1971 to 2000 and forecast: 2071 to 2100) between the six GCMs with the best MPIs and the six GCMs with the worst MPIs. The mean climate change signals of the ensembles are given as the difference from the multi model mean climate change signal. 90

4.1 Comparison of the five GCMs with the best MPI with the top five models (global MPI) from a study done by Reichler and Kim [2008a].110

4.2 Comparison of the five GCMs with the strongest temperature at surface climate change signal (reference: 1971 to 2000, forecast: 2071 to 2100 in the large domain) with the five GCMs with the highest transient climate response and the highest equilibrium global climate response (climate sensitivity) [IPCC, 2007]. 112

4.3 Average uncertainty of the climate change signal in percent for the groups GCM formulation, emission scenario and internal uncertainties for both climatologies and both domains. 114

4.4 Comparison of the contribution to the total variance (in percent) by GCM formulation, the emission scenarios and the internal variability in the Alpine domain (Déqué) and the small domain. The 21 % of uncertainty from RCM formulation, which occurs in Déqué et al. [2007], was equally distributed to the other uncertainty sources.115

4.5 Overview of the ANOVA results for all seasons and all parameters in the large domain. Shown are the percentage of the total variance of the GCM formulation, the emission scenario and the initial condition uncertainty. The missing percentages are carried by mixed terms, which in general are small and not listed in this table. In the column StdDev. the total standard deviation of the corresponding variables are shown. The standard deviations of *hus* and *wap* are too small, so the decimal power is written under the annual values (for *hus* e-9 and for *wap* e-3). Shown are the annual mean ANOVA values (big numbers) and the values of the different seasons (smaller numbers) as it is described on the bottom of the table. 117

Abstract:
In this study the performance and climate change signals over Europe of the 24 General Circulation Models (GCMs) used in the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC) are analyzed, to provide a comprehensive basis for the selection of lateral boundary conditions for Regional Climate Models (RCMs) operated in the Alpine region. Further, various components of uncertainty in the GCM climate change signals (GCM formulation, emission scenario and internal variability) are quantified using analysis of variance (ANOVA) techniques.
For evaluation, nine parameters (three on the ground level and six in four different pressure levels) are compared against observed and reanalysis data in the climate period 1961 to 2000. Concerning the future scenarios, the temporal development of the GCM ensemble results are analysed until 2100. The climate change signals for the periods 2021 to 2050 and 2071 to 2100 minus the reference period 1971 to 2000 are evaluated in detail using the ANOVA technique. Furthermore, the correlation between climate change signal and the model performance is analyzed.
The results of this study can be used as guideline for the selection of driving data for RCM simulations in the Alpine region. They give an overview of the ability of state of the art GCMs to reproduce the contemporary climate in an orographically complex area and show the sources of uncertainty in the climate change signals of multiple climate variables.

Zum Inhalt:
In dieser Studie wird die Güte und die Klimawandelsignale jener 24 globalen Klimamodelle (GCMs) über Europa analysiert, die im vierten Sachstandsbericht (AR4) des Intergovernmental Panel on Climate Change (IPCC) zur Anwendung kamen. Ziel ist es, eine umfassende Grundlage für die Auswahl von Randbedingungen für regionale Klimamodelle (RCMs) im Alpenraum zu liefern. Weiters werden Komponenten der Unsicherheit in den Klimaszenarien (GCM Formulierung, Emissionsszenario, interne Variabilität) mit Hilfe der Varianzanalyse (ANOVA) quantifiziert.
Für die Evaluierung der Modellgüte werden neun Parameter (drei in Bodenhöhe und sechs in vier Druckniveaus) in der Klimaperiode 1961 bis 2000 mit Mess- und Reanalysedaten verglichen. Bezüglich der Klimaänderung wird die zeitliche Entwicklung des GCM Ensembles bis 2100 untersucht, wobei die Änderungssignale für die Perioden 2021 bis 2050 und 2071 bis 2100 (verglichen mit der Referenzperiode 1971 bis 2000) im Detail betrachtet, und für die ANOVA Analyse herangezogen werden. Zusätzlich wird der Zusammenhang zwischen Modellgüte und Klimaänderungssignal untersucht.
Die Resultate dieser Studie können als Richtlinie für die Auswahl von Antriebsdaten für RCMs im Alpenraum dienen. Sie geben einen Überblick darüber, wie gut moderne GCMs den vorherrschenden Klimawandel in einem orographisch komplexen Gebiet nachbilden können und zeigen die Unsicherheitsquellen in den Klimasignalen verschiedenster Klimaparameter.

Wegener Center for Climate and Global Change
University of Graz
Leechgasse 25
A-8010 Graz, Austria
www.wegcenter.at
ISBN 978-3-9502615-7-8

Wegener Zentrum für Klima und Globalen Wandel
Karl-Franzens-Universität Graz
Leechgasse 25
8010 Graz, Austria
www.wegcenter.at
ISBN 3-200-00407-X