Detection of an anthropogenic influence on the observed changes in near-surface temperature and precipitation in Northern Europe

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Department
Geowissenschaften
der Universität Hamburg

vorgelegt von

Jonas Bhend

aus Bern, Schweiz

Hamburg, 2009

Als Dissertation angenommen vom
Department Geowissenschaften der Universität Hamburg
auf Grund der Gutachten von Professor Dr. Hans von Storch
und Dr. Burkhardt Rockel

Hamburg, den
Professor Dr. Jürgen Oßenbrügge
(Leiter des Department Geowissenschaften)

Abstract

Extensive evidence of a large-scale anthropogenic climate change has been collected during the last decades. In contrast, regional-scale climate change is less well understood. Thus, this study aims at contributing to the discussion of an observable human influence on changes in near-surface temperature and precipitation in the Baltic Sea catchment.

The spatial resolution of present-day global general circulation models (GCMs) is inappropriate for the simulation of climate over heterogeneous terrain and in regions with complex land-sea distribution. Therefore, we compare the spatial pattern of the change with regional climate model simulations. In contrast, the change in area-average temperature and precipitation is compared with GCM simulations, as dynamical downscaling with regional climate models is shown to have a minor influence on area-average quantities.

We compare the most recent trends in observed precipitation and temperature in the Baltic Sea catchment with estimates of the response to anthropogenic forcing from regional climate models. The observed change in temperature is consistent with regional climate change projections. In contrast, the model projections generally underestimate the magnitude of the recent change in precipitation. The recent change in the North Atlantic Oscillation does not fully account for the mismatch in simulated and observed precipitation changes.

In a formal detection and attribution analysis we assess the relative importance of anthropogenic and natural forcing and the role of internal variability in explaining the observed change in area-average temperature and precipitation using GCM simulations. The observed change in area-average temperature in the Baltic Sea catchment is very likely not caused by internal variability alone. Although anthropogenic forcing is the dominant forcing, we are not able to separate the anthropogenic and natural influence on the observed change in a two-signal attribution analysis. The results for precipitation are strongly dependent on the details of the analysis. We detect an external influence, but neither anthropogenic nor natural forcing alone provide plausible explanations for the observed change. Model simulated changes in precipitation are generally

much weaker than the observed changes, which corroborates the findings based on regional climate models.   The detection results for precipitation should be treated with caution, as present-day GCMs have severe limitations in simulating regional-scale precipitation.

Contents

1 Introduction

During the last decades, evidence of a global-scale human influence on the observed climate change has been detected in various thermal properties of the climate system, including near-surface temperature (Wigley and Raper, 1990, Stouffer et al., 1994, Santer et al., 1995, Hegerl et al., 1996, 1997, 2000, Barnett et al., 1999, Tett et al., 1999, Stott et al., 2000, Gillett et al., 2002b, among others), ocean heat content (Barnett et al., 2001), and sea surface temperatures (Gillett et al., 2008b). In contrast, the human influence in non-thermal properties has been detected only recently (e.g. sea level pressure, Gillett et al., 2003, precipitation, Zhang et al., 2007, and tropopause height, Santer et al., 2004). Even less is known about the anthropogenic climate change at the regional scale and on impact relevant parameters such as precipitation, wind, and extremes in climatic parameters. In recent years, regional-scale climate change and impact relevant parameters have thus come into focus (e.g. Kiktev et al., 2003, Christidis et al., 2005, Hegerl et al., 2006, Gillett et al., 2008a, Min et al., 2008). An anthropogenic effect on near-surface temperatures in Europe has been detected by Stott (2003), by Zhang et al. (2006) for the Mediterranean, and by Spagnoli et al. (2002) for France. Recently, Zorita et al. (2008) concluded that the number of record-warm years in northern Europe is significantly different from what is expected in a stationary climate. Apart from that, little is known about a potential anthropogenic influence in Northern Europe.

Therefore, this study aims at investigating on the detectability of an anthropogenic influence on observed temperature and precipitation changes in the Baltic Sea area.

In order to be able to discuss the detectability of an anthropogenic climate change in the Baltic Sea region, we need to introduce a few concepts and terminologies. The following paragraphs thus include a short introduction to the climate system, climate change, and the concepts of how we assess the detectability of anthropogenic change in observations of climatic parameters. Readers familiar with these concepts may well continue the lecture in chapter 2.

1.1 Climate change and the climate system

The climate system is a complex and interactive system. It consists of subsystems comprising the atmosphere, the hydrosphere, the land surface, the cryosphere, and the biosphere (Le Treut et al., 2007). The climate system is powered by the Sun; the temporally and spatially varying energy input due to the Earth's movement around the Sun causes motion in the atmosphere and the ocean and also various changes in the other components. In each of the subsystems and between the individual components of the climate system, a wealth of processes and interactions take place. Thus, the components of the climate system intrinsically generate variability. This variability is referred to as internal variability of the climate system.

Internal variability is present at all time scales. The atmosphere is the most unstable component of the climate system and thus changes most rapidly. Processes in the atmosphere, such as small-scale turbulence, synoptic systems, and stratosphere-troposphere exchange, cause variability on time-scales from virtually instantaneous to years. The oceans integrate the high-frequency variability from the atmosphere (Hasselmann, 1976) and damp temperature variations due to the thermal inertia of large water bodies. Another important source of variability is the coupling of different components of the climate system. The most well-known process in this respect is the El Niño-Southern Oscillation (ENSO). The ENSO phenomenon describes the coupling of a sea-level pressure oscillation in the Pacific, the Southern Oscillation, and changes in the ocean circulation and stratification leading to sea-surface temperature anomalies that are most noticeable in the tropical eastern Pacific, the El Niño/La Niña anomalies (see Neelin et al., 1998 for more discussion). El Niño/La Niña conditions occur every 3-7 years and have strong implications on the occurrence of floods and droughts in the tropics and to a lesser extent affect climate globally (Trenberth et al., 2007, and references therein). Further low-frequency variability in the climate system stems from changes in the ice-sheets and changing land cover and vegetation.

In addition to the variability generated in the climate system, climate also varies under the influence of external forcing mechanisms. External forcings can be either of natural origin, such as changes in the solar irradiance and changing atmospheric composition due to volcanic eruptions, or they can be anthropogenic, such as the increase in atmospheric greenhouse gas concentrations due to fossil fuel combustion. The response to external forcings is modified by often non-

linear processes and amplified or damped by feedback processes in the climate system. The first-order effect of external forcing is a change in Earth's energy balance. While changes in the solar irradiance directly affect Earth's energy balance, other forcings such as increasing stratospheric aerosol concentrations due to volcanic eruptions or increasing greenhouse gas concentrations alter Earth's energy balance through changing the radiative properties of the atmosphere. Forcing mechanisms such as changes in the tropospheric aerosol concentration and land cover changes also have a direct influence on the hydrological cycle.

The relative importance of the individual external forcing mechanisms at the global scale is assessed using the concept of radiative forcing (Ramaswamy et al., 2001). Radiative forcing is defined as the radiation imbalance at the tropopause after allowing stratospheric temperatures to readjust to radiative equilibrium but with the troposphere and land-surface temperatures remaining unadjusted. Thus, radiative forcing can only be computed in model experiments. Radiative forcing can be converted to changes in global mean temperature; the conversion factor is called climate sensitivity parameter. The quantification of climate sensitivity - though a central quantity for the understanding of climate change - is still a matter of ongoing debate (for a review see Knutti and Hegerl, 2008), and thus radiative forcing estimates are not easily transferred to temperature changes. In addition, radiative forcing is a useful concept to estimate and compare the global-scale equilibrium response to perturbations, but it does not necessarily convey information on the transient response of the climate system nor on the relative importance of the individual forcings at the regional scale.

The estimated contribution of individual radiative forcings to the total forcing from 1750 to 2005 as assessed by the Intergovernmental Panel on Climate Change (IPCC, Forster et al., 2007) is shown in figure 1.1. Accordingly, the radiative forcing due to long-lived greenhouse gases is +2.63 (+/-0.26) Wm^{-2}; the total radiative forcing due to human influences is +1.6 Wm^{-2} with a large uncertainty ranging from +0.6 to +2.4 Wm^{-2}. The large uncertainty of the anthropogenic radiative forcing estimate originates mainly from the uncertainties in the effect of aerosols on climate. We further note that forcings affecting local to continental scale climate such as the aerosol forcing or surface albedo changes are less well understood than the global-scale greenhouse gas forcing.

Persistent changes in climatic parameters both due to external forcing and internal variability are referred to as climate change. Climate change detection then deals with the decomposition of the observed climate change into an ex-
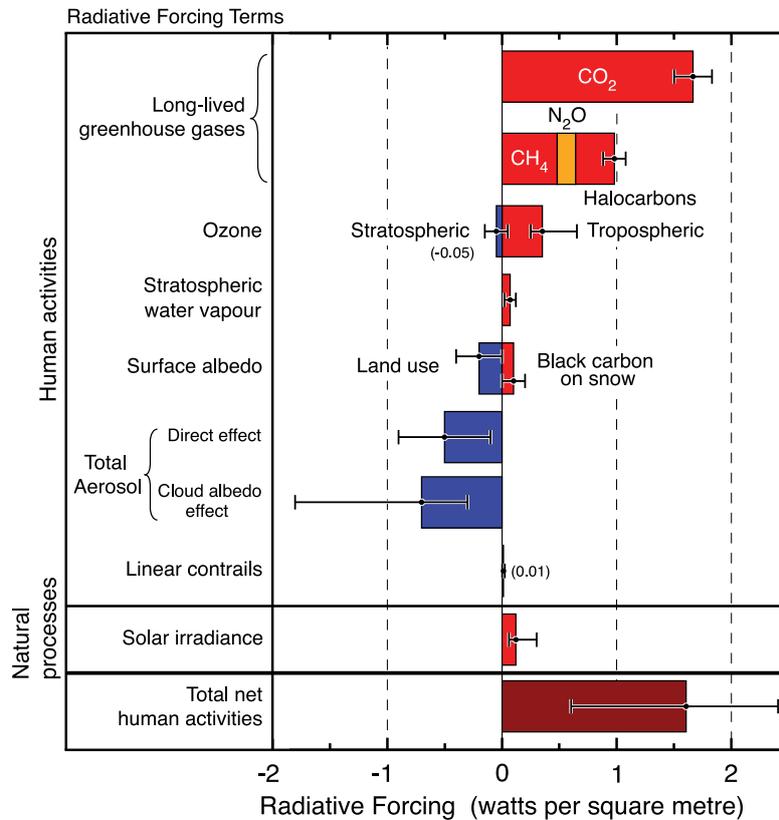
Figure 1.1: *Radiative forcing from 1750 to 2005 as assessed by the IPCC (adopted from Forster et al., 2007)*

ternally forced response and changes due to internal variability. Climate change detection is thus a signal in noise problem.

1.2 The concepts of detection and attribution

Following the definition from the third assessment report of the IPCC (Mitchell et al., 2001), detection is the process of demonstrating that the observed change is significantly different from internal variability in a statistical sense. This definition implies the following things: First, detection analysis does not provide explanations (possible forcing mechanisms) for the observed change. We only exclude with a certain probability of error the possibility of the observed change

being due to internal variability alone. Second, the crucial element in a detection analysis is the estimate of internal variability. If we underestimate internal variability, our detection statement will be overconfident and vice versa. Even though we do not have to specify an alternative hypothesis to be able to technically detect an externally forced change, adding knowledge on the expected change in the analysis will generally increase the power of the statistical test considerably.

Attribution, on the other hand, deals with the identification of the most likely explanation(s) for the observed change. Unequivocal attribution would require controlled experimentation with the climate system. This is not possible as we have no spare Earth to experiment with. Hence, attribution has to be based on physics-based models of the climate system (but not necessarily complex GCMs, see Schneider and Held, 2001, for an alternative approach). The attribution problem is usually formulated as a statistical fit with the observed change regressed on the model-derived responses to external forcings. In contrast to the detection problem, the null hypothesis of the attribution analysis is the desired outcome. That is, we try to *verify* the null hypothesis; a successful attribution statement is consequently less powerful than a detection statement. Instead, we rather think of attribution as a plausibility assessment. We attribute the observed change to a given forcing mechanism if the following three requirements are fulfilled with a specified probability of error:

- the observed change is not due to internal variability alone (detection),

- the observed change is consistent with the response to the hypothesized forcing mechanism(s),

- and the observed change is not consistent with the response to all other physically plausible forcing mechanisms.

The last requirement indicates that attribution is a never-ending task, since we have to reanalyze the data as our understanding of the forcing-response mechanisms in the climate system evolves.

1.3 Regional-scale detection and attribution

Compared to global-scale climate change detection, two main additional difficulties complicate detection and attribution at the regional scale. First, as a

consequence of the decreasing aggregation of data, the variability increases (see figure 1.2). This most often leads to decreasing signal-to-noise ratios of externally forced changes, and these changes are generally harder to detect in small domains (Zwiers and Zhang, 2003).
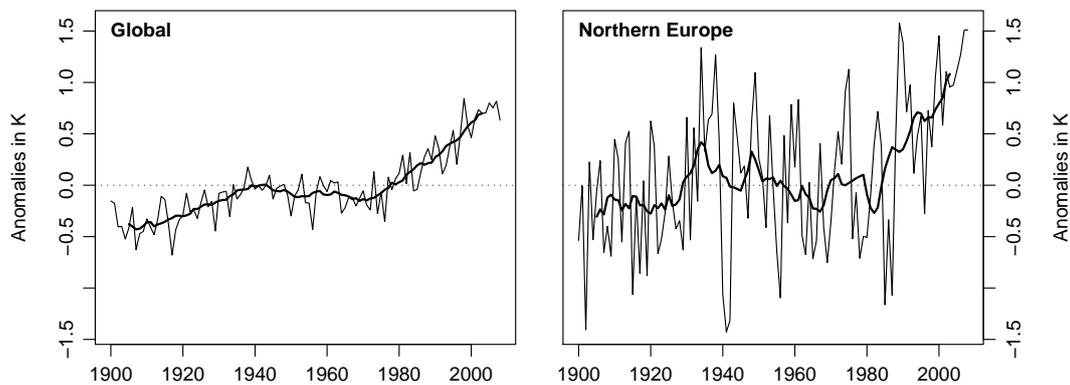

Figure 1.2: *Annual area average land-surface temperatures from 1900 to 2008 for global land except Antarctica (left) and Europe north of 50˚ N and west of 40˚ E (right) according to the CRUTEM3v data set (Brohan et al., 2006). The thick line denotes the 11-year moving average of the annual temperatures.*

Second, the potential limitations of present-day global general circulation models (GCMs) in reproducing regional-scale climate constrain our understanding of regional climate change. These potential limitations fall in two broad classes: resolution and complexity.

The coarse resolution of present-day GCMs with horizontal grid spacing of the order of 100-300 km can lead to considerable misrepresentation of regional climate in areas with strong relief and complex land-sea distribution (see figure 1.3). Processes such as snow-melt or changes in sea surface temperature and in sea-ice distribution are not well represented. Additionally it is known that the large-scale circulation response can be altered considerably by local topographic forcing (Christensen et al., 2007).

We can overcome resolution dependent limitations using empirical or dynamical downscaling, each of which has its own drawbacks. Empirical downscaling comes at the cost of loss in variability and is based on the assumption of stationarity of the empirical relationship used to downscale the desired quantities (Wilby et al.,
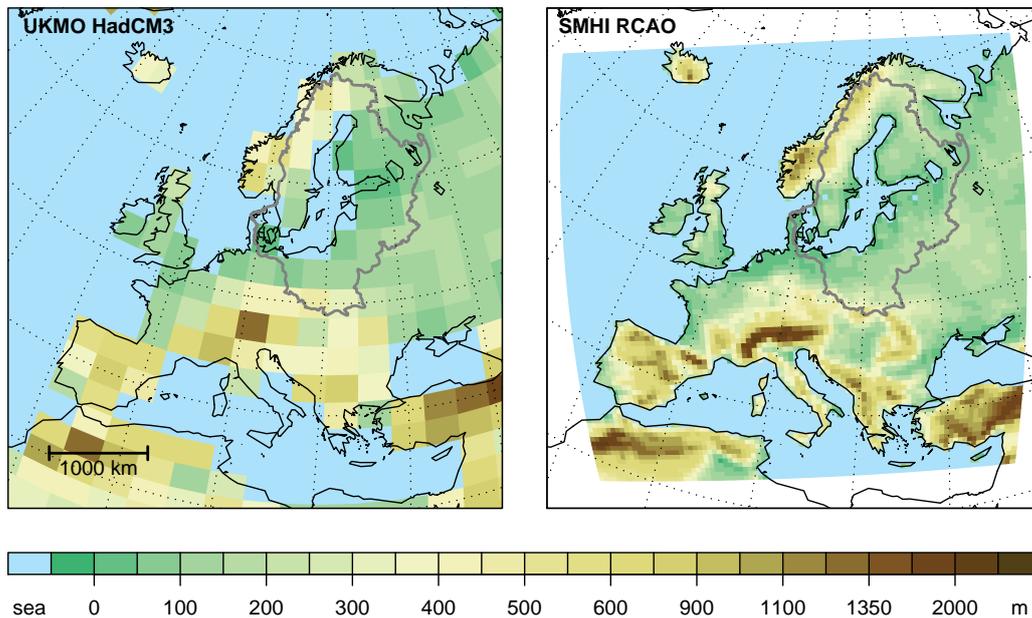
Figure 1.3: *Typical horizontal resolution and orography of present-day global climate models (left) and limited area models as run for ensemble simulations (right). The grey lines denote the Baltic Sea catchment.*

1998, von Storch and Zwiers, 1999). Dynamical downscaling using limited area models, on the other hand, is expensive in terms of compute time. Even though there have been significant efforts taken to downscale GCM simulations for Europe during the last two decades (e.g. the EU-Projects PRUDENCE, Christensen and Christensen, 2007, and ENSEMBLES, Hewitt and Griggs, 2004), simulations required for detection and attribution studies such as unforced control runs and single-forcing runs have not yet been performed with regional climate models (RCMs).

An alternative way to deal with the different spatio-temporal resolution of model and observation data is to upscale or aggregate the observed quantities. By doing so, we reduce small-scale and short-time fluctuations in the observed record that are not simulated by the models. However, upscaling is a promising approach only if the local processes do not influence climate and climate change at larger scales.

The second source of uncertainty in regional predictions arises from the varying complexity of climate models. Present-day coupled general circulation models include at least atmosphere, ocean, and sea-ice modules, the number of additional components included and the number of processes and forcings explicitly simulated varies significantly across models. All of these differences are important sources of uncertainty in regional projections; for climate change attribution, however, missing external forcing mechanisms are the most important uncertainty, since the local response may be dominated by forcings other than the increasing greenhouse gas (GHG) concentrations (e.g. land-use changes, industrial aerosols). This represents a major drawback for regional-scale attribution, as we cannot be sure that the most important explanations (the regional response to external forcings) have been included in the analysis.

One possible way to deal with the problem of potentially missing or misrepresented forcing mechanisms and processes in the models involves perfect-model studies. By carrying out the detection and attribution analysis for individual model simulations which are treated as pseudo-observations, we can identify the regionally important forcings in the model world. Getting the same positive detection and attribution both in the model world and in the real world (using of course model-derived climate change signals in both cases) then substantiates the finding that the hypothesized mechanisms are responsible for the observed change. On the other hand, if the detection results in the real world differ from the results using simulations as pseudo-observations, we cannot know whether the differences are due to missing forcing mechanisms or due to model biases.

1.4 Study domain

The Baltic Sea catchment is an ideal testbed for our understanding of regional climate change. The study domain covers an area of 1.7 million km^2 - slightly less than 20% of the European continent. The Baltic Sea catchment extends from 8.2 to 38.1° E and from 49 to 69.4° N and thus ranges from the temperate mid-latitudes to sub-polar regions (gray lines in figure 1.3). A wealth of different processes influence the climate in this region and potentially modify the large-scale response to external forcings. Features of the study domain and related processes include:

- a pronounced relief with mountains up to 2000 m at the northwestern divide and thus strong temperature and humidity gradients,

- variable land surfaces with numerous rivers and lakes and complex land-sea distribution,

- an inland sea with brackish waters and very distinct dynamics,

- intermittent snow-cover with high spatial and temporal variability due to the distinct orography,

- intermittent sea-ice in the Gulf of Bothnia and Gulf of Finland,

- wintertime cold-air outbreaks from the Arctic leading to deep convection with heavy snowfall, to name just a few.

Thus, the climate in the Baltic Sea area is expected to be difficult to predict, and consequently, climate change detection and attribution can help to identify gaps in our understanding of regional climate change.

1.5 Outline of the thesis

In the following chapter, we assess the influence of dynamical downscaling on the simulation of parameters relevant for regional scale detection. In chapters 3 and 4, we assess the spatial pattern of the change using information from regional climate models. In the absence of regional climate model simulations required for formal detection and attribution studies, we propose an alternative approach. In chapter 5 we analyze the temporal evolution of area-average temperature and precipitation in a formal detection and attribution analysis. The response patterns and estimates of the internal variability are derived from simulations with global climate models. The last chapter contains a summary of the most important results of the previous analyses along with a short outlook on potential future studies.

2 Regional climate in global models and limited area models

In this chapter, we discuss how potential model deficiencies due to inadequate resolution influence detection and attribution in northern Europe. Therefore, we compare the results of time-slice experiments with a coupled regional atmosphere-ocean model and the results of a set of simulations with present-day coupled atmosphere-ocean general circulation models (AOGCMs) with observation data. In contrast to AOGCMs, in which the Baltic Sea often features either as a part of the North Atlantic with an unrealistically wide opening towards the North Sea or as a lake with no exchange with the North Atlantic, the high resolution of the coupled regional climate models allows for a more realistic representation of the Baltic Sea. The misrepresentation of the Baltic Sea in present-day AOGCMs has supposedly large consequences for the simulation of the local climate (Räisänen et al., 2003, Kjellstrom et al., 2005).

2.1 Model data and observations used

The RCM data stem from the Rossby Centre regional atmosphere-ocean model (RCAO) of the Swedish meteorological and hydrological institute (SMHI, see Räisänen et al., 2004 and references therein). A total of six 30-year time slice experiments have been performed, three of which driven at the boundary with the global model HadAM3h (Jones et al., 2001b) and three driven with data from the ECHAM4/OPYC model (Roeckner et al., 1999). For each of the global models, one time slice simulation has been run with observed GHG and sulfate concentrations representing present-day conditions for the period 1961-1990 and two time slice experiments have been run representing possible futures for the period 2071-2100 with GHG and sulfate concentrations according to the IPCC SRES A2 and B2 scenarios.

The HadAM3h is an atmosphere only model and has been run with observed sea-surface temperatures in the present-day time slice to get the best possible

representation of the observed climate. For the time-slice from 2071 to 2100, the observed sea-surface temperatures have been combined with the climate change signal derived from a simulation with the coupled model HadCM3 (Gordon et al., 2000). The above mentioned limitations of present-day GCMs in representing the Baltic Sea are overcome when observed sea-surface temperatures are used as in the simulations driven with HadAM3h. In this setup, the comparison of the regional climate model simulation with its driving GCM simulation does not shed light on the question whether dynamical downscaling with a coupled regional climate model can help to overcome the shortcomings of global models at the regional scale. Thus, we do not analyze the RCM simulations driven with HadAM3h, but use the ECHAM4/OPYC driven simulations instead.

In addition to the GCM data used to drive the RCM at the boundaries, we include simulations from the World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project (CMIP3) multi-model dataset in the analysis (Meehl et al., 2007b). We use all of the available CMIP3 data except for the FGOALS1.0_g model (Yongqiang et al., 2002, 2004), which has a highly unrealistic sea ice distribution and should not be used for mid- to high-latitude studies (see documentation on this model on `http://www-pcmdi.llnl.gov/ipcc/model_documentation/ipcc_model_documentation.php`). Instead of future projections following the SRES A2 or B2 scenario as with the RCM and driving GCM data, we use the same CMIP3 projections following the SRES A1B scenario that are used in the formal detection and attribution analysis in chapter 5.

The observed temperature and precipitation in the Baltic Sea catchment is derived from gridded temperature and precipitation data sets. For temperature, we use the CRUTEM3v data from the Climatic Research Unit (CRU, Brohan et al., 2006, Jones and Moberg, 2003). CRUTEM3v is based on 2-m temperatures that have been interpolated to a 5° by 5° grid and are available as anomalies with respect to 1961 to 1990. The CRUTEM3v data are adjusted for changes in the variance due to a temporally changing number of stations contributing to the grid box averages (Jones et al., 2001a). The absolute temperatures needed to assess the climatological bias between simulations and observations are obtained by adding the gridded climatology of New et al. (1999) to the anomalies. Area-average precipitation is derived from the fourth version of the "Global Data Reanalysis Product" of the Global Precipitation Climatology Centre (GPCC, Schneider et al., 2008). These gridded precipitation fields have been interpolated from quality controlled rain-gauge data.

All model simulations and gridded observation data are interpolated to a common 5° by 5° longitude latitude grid using conservative remapping (Jones, 1999). Only grid boxes with a land-area fraction of at least 50 percent are considered in the interpolation. We compute area-average temperatures and precipitation from the gridded data and weight the individual grid boxes according to their contribution in area to the total area of the Baltic Sea catchment.

2.2 Analysis and Discussion

We compare three quantities relevant to the detection and attribution problem. First, we analyze the systematic bias in the model simulations by comparing the annual cycle of simulated area-average temperature and precipitation to the observations. Even though the systematic bias is not directly relevant to the detection and attribution question, it points to missing or misrepresented processes and is thus important in judging whether the model data are adequate. Second, we compare the variability of simulated area-average time series to the variability in the observations. Third, we investigate whether dynamical downscaling of GCM data with a coupled RCM alters the climate change response. Therefore we compare the change between the future (2071-2100) and present-time (1961-90) time slices available from the RCM with the respective change in GCM data.

2.2.1 Systematic biases in the present-day climate

Most of the GCMs feature a more pronounced annual cycle, with winters being too cold and and summers being slightly too cold on average (left panel in figure 2.1). The strong cold bias in the winter half-year in northern Europe is a well-known fact and has been attributed to weaker than observed westerlies in the CMIP3 ensemble (Christensen et al., 2007). Earlier studies based on different GCMs find a cold bias in the multi-model mean as well; this cold bias is most pronounced in spring (see chapter 3.3.2 in The BACC author team, 2008). The ECHAM4/OPYC simulation used to drive the coupled regional climate model RCAO shows a pronounced cold bias of about 2K in early spring, whereas the simulated temperature is close to the observations during summer and autumn. This bias is strongly reduced in the RCAO simulation. The higher spatial resolution and resulting more realistic snow cover and snow melt are likely candidates to explain the improved annual cycle of temperature in the RCM.
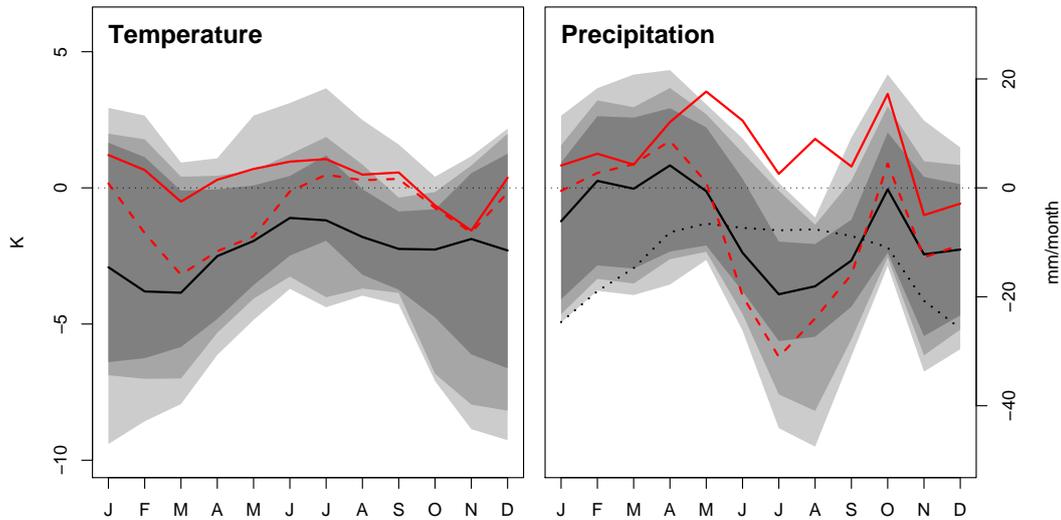
Figure 2.1: *Deviation of simulated from observed area-average temperature (left) and precipitation (right) for the period from 1961-90. The solid red line denotes the bias of the coupled regional climate model simulation, the dashed red line the bias of the GCM used to drive the regional model. The range and the middle 90% and 75% of all CMIP3 simulations are indicated in shades of grey along with the median of the CMIP3 simulations in black. The dotted black line denotes the uncorrected precipitation values (see text).*

Precipitation measurements suffer to varying degrees from systematic errors such as wind-induced undercatch and evaporation losses (e.g. Adam and Lettenmaier, 2003). In order to account for some of these systematic errors, we use a climatological correction for gridded precipitation data provided by Legates and Willmott (1990). Compared with the uncorrected observations, all of the GCMs simulate too much precipitation from autumn to spring and most of the GCMS simulate too little precipitation in summer (right panel in figure 2.1). After correcting the observations for wind-induced undercatch, most of the wet bias from autumn to spring is removed, whereas the dry bias in summer is now consistent across all models in the CMIP3 ensemble. The ECHAM4/OPYC simulation shows a bias very similar to the median of the bias of the CMIP3 ensemble, with the dry bias in summer being even stronger than the median of the bias in the CMIP3 ensemble. In contrast, we find no dry bias in summer in the regional climate model simulation. In addition, ECHAM4/OPYC shows much less variability in summer-

time precipitation than RCAO (see figure 2.2). Compared to the observed variability, both RCM and GCM underestimate the variability. The summertime dry bias (see figure 2.1) and the underestimation of precipitation variability are consistent features across the GCM simulations analyzed (not shown). These results hint to the fact that the precipitation parameterization in the global model is not adequate for the - mostly convective - precipitation in the Baltic Sea catchment area in summer.

Whether the higher resolution of the regional model leads to a better representation of convective precipitation or whether the parameterization in the regional model has been – deliberately or unintentionally – tuned to reproduce precipitation in northern Europe, we do not know. However, none of the 14 regional climate models analyzed in the PRUDENCE project shows a dry bias in summer over northern Europe (Jacob et al., 2007). This and the consistent dry bias in all GCM simulations analyzed indicates that tuning alone is not a plausible explanation of the improved representation of summertime precipitation in RCMs.
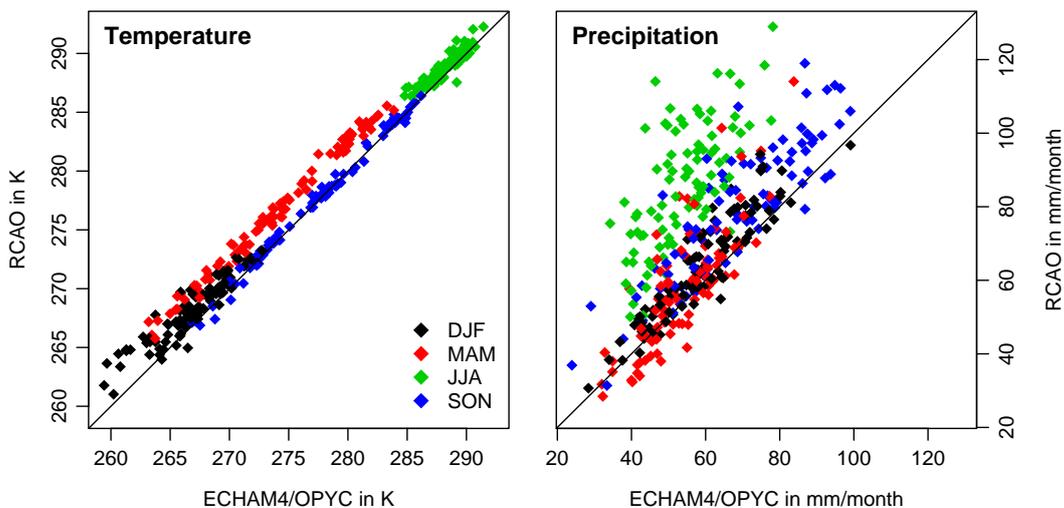

Figure 2.2: *Monthly area-average temperature (left) and precipitation (right) in the Baltic Sea catchment for the period from 1961 to 1990 as simulated in the regional coupled model RCAO and the global model ECHAM4/OPYC providing the boundary conditions for the regional climate model simulation. The colors indicate the seasons in which the monthly data fall.*

2.2.2 Variability in the present-day climate

We compare the standard deviation of simulated and observed seasonal temperature anomalies in the present-day time slice (1961-90). The variability of seasonal temperature anomalies as simlated by most of the GCMs is not significantly different from the observed variability according to a variance-ratio test (5% significance level, see figure 2.3). In addition, the variability in both the subset of the CMIP3 ensemble including anthropogenic forcing only (ANT) and the variability in the subset including anthropogenic and natural forcing (ALL) clusters around the observed variability. The variability of seasonal temperature anomalies as simulated by models with anthropogenic forcing only is not significantly different from the variability as simulated by models including both anthropogenic and natural forcings. Thus, natural forcing in GCM simulations has little influence on the variability of seasonal temperature in the Baltic Sea catchment. In contrast we find that simulated variability differs considerably across models.

Downscaling the GCM results with the coupled regional model RCAO decreases the variability in seasonal temperature anomalies considerably. Compared to the driving GCM, the RCM simulated temperature varies slightly less in all seasons. In winter, however, the difference in simulated variability is most pronounced. As can be seen from figure 2.2 this is mainly due to the fact that cold winters tend to be colder in ECHAM4/OPYC than in the regional model RCAO. Cold winters are also milder in the other RCM time slices compared to the respective GCM simulations used to drive the RCM. As for the cold bias in spring, this difference between RCM and GCM simulations could be related to excessive snow cover in the GCMs due to the coarse resolution. However, further analysis is needed to confirm this hypothesis.

The representation of variability in simulated and observed area-average precipitation is analyzed using relative seasonal precipitation anomalies from 1961 to 1990. The simulated variability in precipitation is lower than the variability observed for almost all the models analyzed (figure 2.4). This finding is in line with earlier studies concluding that variability in relative precipitation anomalies is underestimated in global models (The BACC author team, 2008, Räisänen, 2001). Downscaling with a coupled RCM leads to an increase in the variability, which is mainly due to better representation of convective precipitation in the regional model and thus stronger variability in summer and autumn (figure 2.2). However, as we have only time slices of 30 years available from the RCMs, we
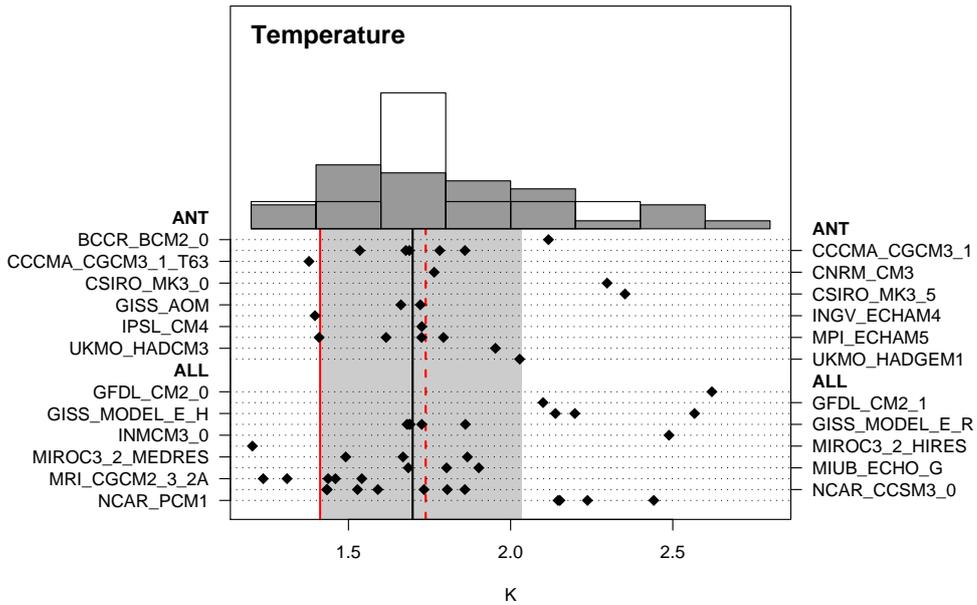
Figure 2.3: *Standard deviation of seasonal area-average temperature anomalies from 1961 to 1990. The vertical black line denotes the observed standard deviation according to the CRUTEM3v data along with its 95% confidence interval according to a variance ratio test. The diamonds and the histograms indicate the standard deviations as simulated in the CMIP3 ensemble. The solid red line is the standard deviation of seasonal temperature anomalies as simulated by the coupled regional climate model, the dashed red line corresponds to the standard deviation in the global model ECHAM4/OPYC used to drive the regional model at the boundaries.*

cannot investigate whether a more realistic representation of the Baltic Sea in RCM simulations has an influence on low-frequency variability. This limits our ability to judge, whether using RCM data could enhance our estimate of internal variability for regional-scale detection and attribution studies.
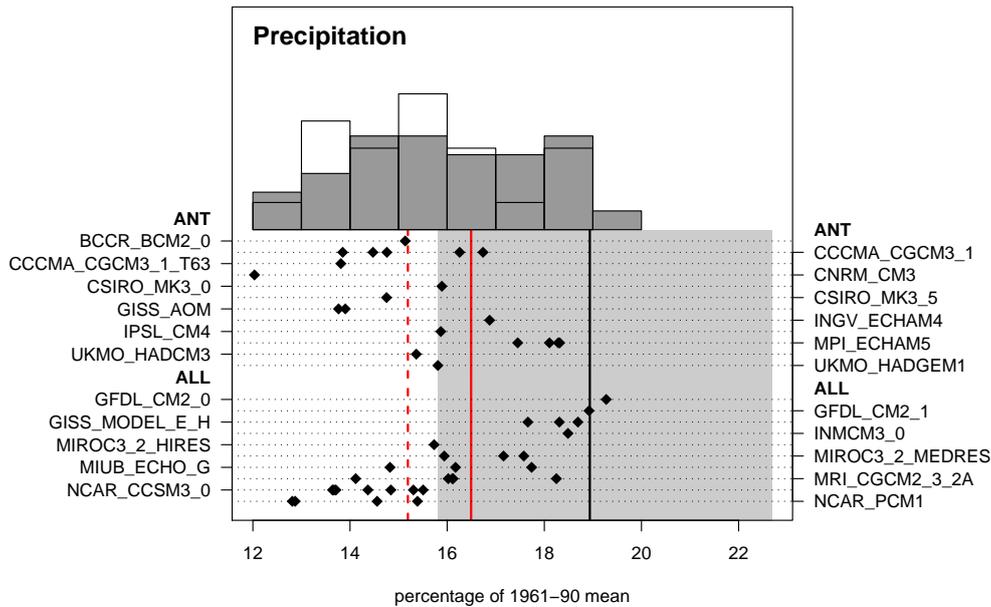

Figure 2.4: *As in figure 2.3 but for area-average relative precipitation anomalies from 1961-90 according to the GPCC v4 observation data.*

2.2.3 Climate change signal

The climate change signal is defined as the difference between the future (2071-2100) and present-day (1961-90) time slices. The area-average response to increasing GHGs is the well known warming, which is more pronounced in winter than in summer (Christensen et al., 2007, and references therein). Downscaling the GCM response with a coupled RCM results in the response being weaker by up to 1K per 110 years (see figure 2.5). A similar reduction of the climate change signal is also found when comparing the climate change signals derived from the coupled regional climate model driven with the HadAM3h model with the climate change signal derived from the HadCM3 simulations providing the

sea-surface temperature response for the HadAM3h simulation (not shown). The damped response in near-surface temperatures is likely due to the explicitly modeled Baltic Sea, as uncoupled RCMs do not show such a consistently reduced response (Christensen and Christensen, 2007).
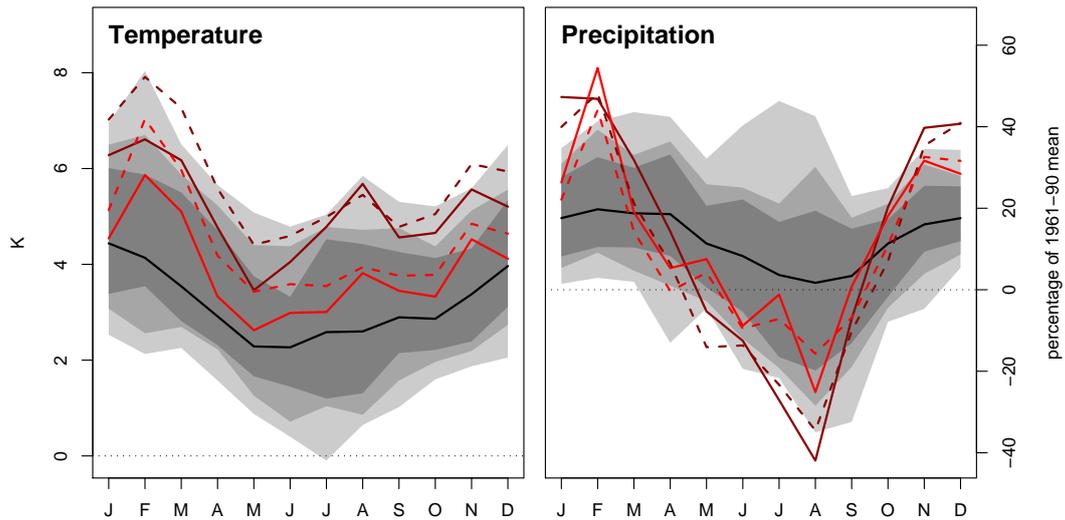

Figure 2.5: *Anthropogenic climate change signal defined as the difference in 30-year mean area-average temperature (left) and precipitation (right) for the period 2071-2100 minus 1961-90. The solid red lines denote the climate change signal of the coupled regional climate model simulation, the dashed red lines the climate change signal of the GCM for emissions according to the SRES A2 (dark red) and B2 scenario (red). The range and the middle 90 and 75 percent of the climate change signals derived from the CMIP3 simulations are indicated in shades of grey along with the median signal in black.*

The climate change signal for precipitation is not significantly altered after downscaling with the RCAO (figure 2.5). Even in summer, when precipitation formation is much less dependent on the boundary conditions by the driving GCM (Christensen and Christensen, 2007), the precipitation response is very similar in the RCM and driving GCM. This indicates that, even though we get a more realistic representation of precipitation in summer with RCMs (see section 2.2.1), the response to GHG forcing is not strongly dependent on the model resolution.

2.3 Conclusions

Area-average mean climate over the Baltic Sea is more realistically simulated in RCMs than in GCMs. The improvements are most obvious for temperature in spring, when temperature strongly depends on the simulation of snow cover and snow melt, and for precipitation in summer and autumn, when convective and thus small-scale precipitation formation dominates. Furthermore, the variability of area-average precipitation is slightly increased in the regional model and thus closer to the observed variability than the variability simulated in GCMs, whereas temperature variability is reduced in the regional model. Whether the reduction of temperature variability is due to the more realistic representation of the Baltic Sea or due to other processes we do not know. Also we do not know whether the coupled Baltic Sea influences low-frequency variability, as there are no long-term simulations with the coupled regional climate model available. The response of area-average precipitation to increasing GHG concentrations is similar in the GCM and RCM, even though precipitation in the RCM in summer and autumn is largely independent of the driving GCM. For temperature, in contrast, we find a consistent reduction of the GHG-induced warming in the RCM compared to its driving GCM.

We conclude, that GCM information can be used for detection and attribution studies with area-average temperature and precipitation in the Baltic Sea catchment. If we are interested in the spatial pattern of change, however, GCM simulations do not provide the necessary information, as important small-scale processes are not well resolved. Therefore, the analyses presented in the following two chapters compare the spatial pattern of change in observed precipitation and temperature to climate change projections derived from RCM simulations. In the last chapter we carry out a formal detection and attribution analysis with area-average quantities using GCM data.

3 Consistency of observed winter precipitation trends in northern Europe with regional climate change projections

This chapter is based on an atricle published in *Climate Dynamics* (Bhend and von Storch, 2008). The original manuscript has been slightly changed for editorial purposes.

3.1 Introduction

When asking for proof of anthropogenic effects on our changing climate, usually "detection and attribution studies" are sought after (Hasselmann, 1993, 1997, Allen and Tett, 1999). For formal detection we need significance of the results against natural internal variability. Consequently, the estimate of natural internal variability is the most crucial part of a detection study.

In the present paper, we pursue a different line of analysis – motivated by the missing or highly imperfect knowledge of natural variability in wintertime precipitation. We ask if the most recent trends are consistent with what contemporary regional climate models envisage as the response to increasing greenhouse gases (GHG) and changing aerosol concentrations. In this way, we offer the possibility to falsify the hypothesis of a presently observable anthropogenic signal. In this setup the recent change is (apart from uncertainties in the initial data and the preprocessing of the data) given, the response to anthropogenic forcing, however, has to be estimated. A testable hypothesis of the above research question is "observed change is drawn from the set of simulated responses to anthropogenic forcing". However, as we have reasons to believe that the estimated responses available do not represent unbiased versions of the "true" response, we refrain from formally testing the hypothesis and collect plausibility arguments instead.

Detection, i.e. rejection of the null hypothesis of "no anthropogenic signal" would be preferable, and the possible outcome of our analysis, namely "no falsification", is less interesting but nevertheless useful. However, it is important to be aware of the limitations of our approach. Our method cannot discriminate the plausibility of different forcing-effects but merely assess the consistency of recent changes with an a-priori assumed mechanism, in particular increasing levels of GHGs in the atmosphere. Furthermore, we cannot deduce a detection statement – "it is unlikely that the observed change is due to natural variability" – from a positive outcome ("anthropogenic forcing is a good explanation for the observed trend") of our analysis. Obviously, a regular "detection and attribution" analysis is more informative, but our method is applicable also in cases of considerably less data and without reference to sometimes hardly available estimates of natural variability.

If the recent trend fails to be consistent with the expected trend, then - given all assumptions are correct - in principle three reasons may be thought of: The model is insufficient (the expected signal is false), the natural variability overwhelms the signal, or more than the expected mechanism is at work, for instance decreasing concentrations of industrial aerosols in parallel to an increase of GHGs. However, due to the lack of meaningful estimates of the natural climate variability and the response to competing forcing mechanisms, we are not able to discriminate between these three reasons using the analysis as presented in this publication.

We focus on wintertime (DJF) precipitation in the Baltic catchment (denoted by solid grey contours in figures 3.2, 3.3, and 3.4) and northern European land areas (e.g. figure 3.2) as we expect the largest changes due to anthropogenic forcing to occur in this season. Furthermore, we know that wintertime precipitation is mainly large scale and thus more reliably simulated by climate models.

3.2 Materials and Methods

3.2.1 Observations

Trends in observation data are computed using the well-known gridded data set of the Climatic Research Unit (CRU TS 2.1, Mitchell and Jones 2005). These fields consist of monthly precipitation totals on a 0.5x0.5 degree latitude-longitude

grid available for the period from 1901 to 2002. The gridded observations have further been interpolated to the rotated latitude-longitude grid described in the next section in order to keep the effective grid box area comparable within the research domain. Trends in the observation data have been calculated using ordinary least squares linear regression.

It has been claimed that the CRU TS 2.1 data set is not suited for detection and attribution studies as the station series have not previously been homogenized and possible effects of urban development and land use changes are still present in the data. However, we expect the data to sufficiently reflect precipitation development in northern Europe during the last decades of the twentieth century due to the large number of stations entering the analysis and assuming that most of the inhomogeneities are not systematic.

3.2.2 Anthropogenic climate change signal estimates

The anthropogenic climate change signal is derived from time slice experiments with a regional climate model. Using well separated (in this case 110 years) time slices to estimate the anthropogenic climate change signal has the advantage of increasing the signal-to-noise ratio. In contrast, deriving the anthropogenic fingerprint from transient climate change simulations for the period under investigation (here 1973-2002) requires a large ensemble in order to get a noise free fingerprint. Such large ensembles of transient regional climate change experiments are not available at the moment. We try to capture the range of probable responses by looking at the different climate change projections, well aware of the fact that we might underestimate this range considerably.

The set of regional climate model simulations used in this paper consists of experiments run with the Rossby Centre regional Atmosphere-Ocean model (RCAO) of the Swedish Meteorological and Hydrological Institute (SMHI). These experiments have been carried out as part of the EU project PRUDENCE and are described in detail in Räisänen et al. (2004), Kjellström (2004), and references therein. The atmospheric part of the RCAO has been run on a rotated latitude-longitude grid with a grid spacing of 0.44° (approx. 49 km).

For each of the two different driving global models, the ECHAM4/OPYC (Roeckner et al., 1999) and HadAM3H (Gordon et al., 2000), a control run representing present day conditions (1961-1990) and two scenario runs (2071-2100) based on IPCC SRES A2 and B2 scenarios have been run. The regional anthropogenic

change signals for the different driving GCMs and scenarios have been defined as the difference between scenario and respective control run means scaled to change per decade. As we have only two "points" along the time axis, namely the change of 30 year mean precipitation from 1961-90 to 2071-2100, we scale the projections by assuming a linear development between 1961-2100; the validity of this assumption is discussed at the end of this section. The signal is further scaled to change per decade. Hence, we have a set of four regional climate change projections available, which are referred to as HadAM A2 (B2) and ECHAM A2 (B2) in the following according to the driving GCM and emission scenario used.

Underlying to our analysis are several assumptions, which are listed in the following:

First, we assume that our contemporary models are good enough for projecting anthropogenic climate change. We believe that they are, but we have to acknowledge that a conclusive proof of that assumption is not possible at this time.

Second, we presume that regional climate models - especially when coupled with an ocean model and thus resulting in much more realistic sea surface temperatures in the Baltic Sea - provide more realistic estimates of the present and future climates in this region than GCMs do.

Third, the response to anthropogenic forcing is assumed to be linear. This is supported by the analysis of climate change projections with different GHG forcings (SRES A2 and B2) which vary mainly in magnitude (Räisänen et al., 2004). Additionally, the global mean response to anthropogenic forcing is as a first approximation linear as well (Cubasch et al., 2001).

The significance of the climate change estimates has been assessed using the lookup table test as described in (Zwiers and von Storch, 1995). This modified t-test takes into account that the data in the two samples (1961-1990 and 2071-2100) are autocorrelated. In order to get a conservative estimate of the autocorrelation coefficient, we have set negative autocorrelation estimates to zero. Although this test accounts for temporal interdependence of the observations, the spatial correlation is not dealt with in particular. The abovementioned test indicates whether the estimated response could be due to internal model variability. The influence of high-frequency variability is expected to be small (this is supported by the results of the adjusted t-test). Still, as we do not use an

initial condition ensemble, we cannot quantify the influence of low-frequency variability on the climate change scenarios.

3.2.3 NAO representations and NAO signals

We use a station based NAO index, which is defined as the difference in normalized monthly sea level pressure (SLP) between Reykjavik and Gibraltar according to Jones et al. (1997). The NAO index in the set of regional climate model simulations has been derived accordingly from the respective SLP fields. The reference period for the normalization of SLP time series is 1961-1990. The variability in the dimensionless NAO index based on observations is higher than the variability based on the two different 1961-1990 representations in the regional model simulations with a standard deviation of 1.46 in the observations and 1.14 (1.09) in the HadAM (ECHAM) simulation.

The signal or fingerprint of the NAO is defined as the fraction of the variability in wintertime precipitation, which covaries with the NAO index. Thus, we regress the detrended precipitation time series on the detrended NAO index for each grid box separately using ordinary least squares estimation of the parameters of the linear regression. The slope of the regression is the NAO signal or fingerprint. This signal is removed from the observations by subtracting the product of the trend in the NAO index times the NAO signal from the trend in the observations. From the climate change projections, we remove the NAO fingerprint by simply subtracting the respective NAO fingerprint times the difference in the average NAO index for the periods 1961-90 and 2071-2100.

3.2.4 Comparing the patterns of change

The comparison of recent and expected trends R and E may be broken down by considering the dimensionless patterns $R^* = \frac{R}{\|R\|}$ and $E^* = \frac{E}{\|E\|}$, and the norms $\|R\|$ and $\|E\|$. The latter are the intensities of the pattern and are defined as follows:

$$\|R\| = \sqrt{\frac{1}{n} \sum_{i=1}^{n} R_i^2} \qquad \|E\| = \sqrt{\frac{1}{n} \sum_{i=1}^{n} E_i^2} \qquad (3.1)$$

The index $_i$ counts the spatial points $i = 1 \ldots n$.

The patterns are compared with the *pattern correlation coefficient* (PCC, equation 3.2), which is different from the centered pattern correlation and uncentered cross-moment introduced in Santer et al. (1993).

$$\rho_{R,E} = \frac{1}{n} \sum_{i=1}^{n} R_i^* E_i^* = \frac{\sum_{i=1}^{n} R_i E_i}{\sqrt{\sum_{i=1}^{n} R_i^2 \sum_{i=1}^{n} E_i^2}} \tag{3.2}$$

ρ is bound by 1, i.e., $|\rho| \leq 1$. We use uncentered pattern correlation because the information about a human induced change is both in the spatial mean and the spatial variability of the pattern.

Furthermore, we use a bootstrap test to investigate the range of PCCs of randomly generated trend fields with a spatial covariance structure similar to that of precipitation trends. Therefore, we repeatedly (200 times) randomly select thirty years of winter precipitation from the 102 available years and compute the trends. We then calculate PCCs for every two different trend fields once, giving us an ensemble of 19900 randomly generated PCCs. For this set of PCCs, we compute the percentiles. Additionally, as we know that the data might be auto-correlated, we repeat this experiment by randomly selecting groups of five and fifteen consecutive years, this technique is known as the moving blocks bootstrap (Wilks, 1997). It is shown, that for both the Baltic catchment and all of northern Europe, the distribution broadens and thus the percentiles increase (see table 3.1).

Table 3.1: *Percentiles of PCCs of trends from randomly selected precipitation fields for the Baltic catchment and northern Europe (in brackets).*

Percentiles	95th	98th	99th
1 year	0.592 (0.525)	0.690 (0.618)	0.742 (0.672)
5 years	0.623 (0.568)	0.718 (0.666)	0.769 (0.721)
15 years	0.688 (0.638)	0.792 (0.753)	0.852 (0.825)
NAO removed			
5 years	0.669 (0.602)	0.758 (0.697)	0.805 (0.749)

We use the results derived from the moving block bootstrap experiment with a block length of five years, however the selection of the block length is subjective. According to Wilks (1997), the appropriate block length is dependent on the

autocorrelation structure of the data under investigation and should thus be derived individually at each grid point. In contrast, one should use the same block length for all variables in multivariate problems. Thus, the ideal block length cannot be computed as at least one of the two criteria is violated. However, additional analyses with Wilks (1997) rule for choice of the block length have shown that - assuming first order autocorrelation - no block lengths larger than four are found within the domain.

The respective histograms of PCCs for the Baltic catchment and northern Europe are shown in figure 3.1. The percentiles of the random PCCs are listed in table 3.1, the uncertainties (2 standard deviations estimated from 20 repeats of the above experiment) for these percentile estimates range from 0.02 to 0.03 for the Baltic catchment (0.03 to 0.04 for northern Europe). The bootstrap estimates of random PCCs for the Baltic catchment and northern Europe after removing the NAO signal lie slightly higher as shown in the last row of table 3.1.
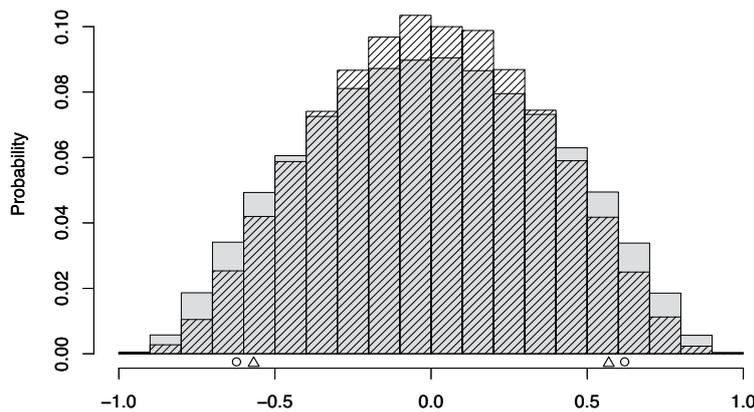

Figure 3.1: *Histogram of PCCs of trends of randomly selected precipitation fields. The shaded bars refer to the PCCs for the Baltic catchment, the hatched bars refer to northern European PCCs. The 5 and 95 percentiles for the Baltic catchment (northern Europe) are indicated by a triangle (circle).*

3.3 Observed and simulated changes in winter precipitation totals

3.3.1 Trends in observation data

We use 30-year trends in order to assess the most recent changes. On the one hand, the period under investigation should be sufficiently short, as we know from global and continental scale results, that the anthropogenic signal in temperature emerges in the last few decades from natural variability (Hegerl et al., 2007). The influence of natural variability on the observed trend reduces with increasing trend length on the other hand.

When different trend lengths are used, the pattern remains mostly unaffected. The magnitude of the trends, however, decreases with increasing trend length due to either a reduction of the fraction of trends induced by natural variability and/or due to a weaker anthropogenic signal in the mid-twentieth century (not shown).

Figure 3.2 shows the 30-year changes in seasonal winter (DJF) precipitation according to the gridded CRU data. The pattern is a general increase over most of northern Europe with regions of slight decrease in central Finland and southern Poland. The maximum positive (negative) trend within the Baltic catchment amounts to 31.01 (-14.28) mm per decade (seasonal totals). On average, the Baltic catchment has become wetter by 8.24 mm per decade. Highest relative rates of change are found over the Baltic states with rates of change up to 20.62% of 1961-1990 mean per decade, which corresponds roughly to a doubling of the seasonal precipitation during the period under investigation. An overview of the spatial statistics of the relative rates of change for the Baltic catchment based on the changes in CRU is given in table 3.2.

It is well known that a part of the precipitation trend, in particular along the Atlantic coastline, is strongly affected by the North Atlantic Oscillation (NAO, Lamb and Peppler, 1987, Wanner et al., 2001). The NAO signal in wintertime precipitation is a general increase (decrease) in precipitation with increasing (decreasing) NAO index over most parts of the domain under investigation, which is strongest in southern Norway and along the Norwegian coast (not shown). In view of our hypothesis, that the signal-to-noise ratio of anthropogenic climate change in the NAO would be low (Rauthe et al., 2004), we subtracted the NAO signal from the precipitation trends (Figure 3.3 left panel). Furthermore, a number of studies
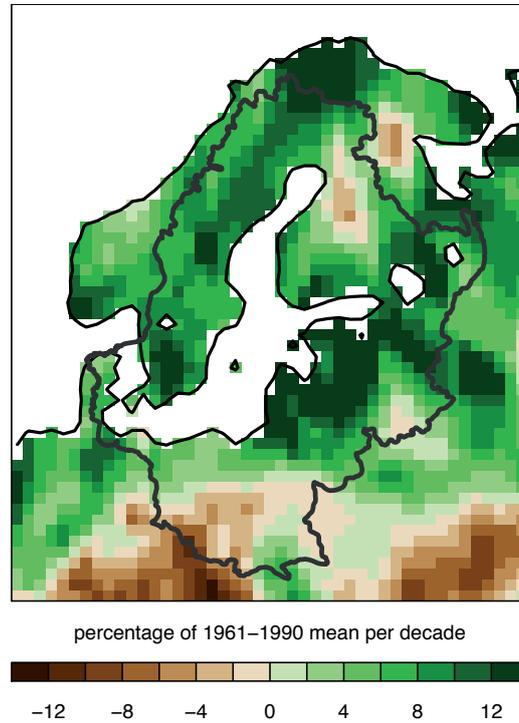
percentage of 1961–1990 mean per decade

−12   −8   −4   0   4   8   12

Figure 3.2: *Trends in winter (DJF) precipitation totals 1973-2002 in units of relative change compared to the 1961-1990 mean precipitation (according to the CRU TS 2.1 data), negative trends are marked by thin solid contours.*

Table 3.2: *Spatial statistics of the observed and expected changes (in units of percentage change of the respective 1961-90 mean per decade) for the Baltic catchment. Values in brackets refer to the spatial statistics for the Baltic catchment after removing the NAO signal.*

	CRU	range of RCAO scenarios
spatial mean	6.97 (4.97)	1.89 - 4.38 (1.79 - 3.70)
spatial std	5.27 (4.90)	0.42 - 0.97 (0.63 - 1.55)
intensity	8.73 (6.98)	1.94 - 4.49 (1.90 - 4.01)

conclude that the observed increase in the NAO is at least partially externally forced (Osborn et al., 1999, Gillett et al., 2002a, Gillett, 2005, among others), but the simulated trend in the NAO is generally smaller than observed (Gillett et al., 2002a, Osborn, 2004). Excluding the NAO signal thus also excludes a part of the variability from the observations, which is not reproduced by present day climate models. As a consequence, we expect the similarity of the patterns of the observed and simulated changes to increase and the intensities of the changes to converge.
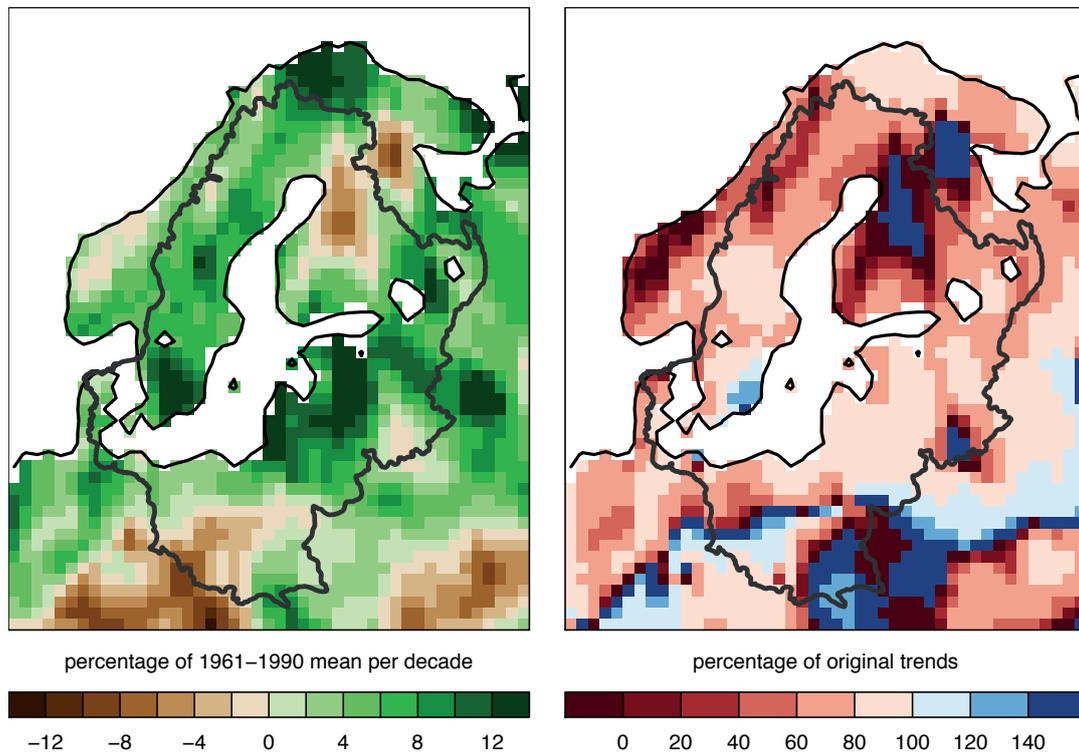

percentage of 1961–1990 mean per decade

−12   −8   −4   0   4   8   12

percentage of original trends

0   20   40   60   80   100   120   140

Figure 3.3: *Left panel: Trends in winter (DJF) mean precipitation 1973-2002, according to the CRU TS 2.1 data after the removal of the NAO signal. Right panel: percentage of trend after the removal of the NAO signal compared to the full trend shown in Figure 3.2. Negative trends and ratios are marked by thin solid black contours.*

The removal of the NAO signal leads to a considerable reduction in precipitation trends along the west coast of Europe as shown in the right panel of figure 3.3. The spatially averaged 30-year trend over the Baltic catchment decreases by 2.31

mm/decade when removing the NAO; the largest and smallest trends over the Baltic catchment are also reduced (31.01 mm/decade vs. 26.47 mm/decade; -14.28 mm/decade vs. -13.74 mm/decade). Nevertheless, the distribution is smoother (spatial standard deviation = 5.88 mm/decade) without the NAO than with the NAO (6.55 mm/decade). However, in central Finland and in southern parts of the Elbe catchment, subtracting the NAO signal increases the trends. These findings qualitatively apply to wider northern Europe as well.

3.3.2 Expected changes derived from climate change scenarios

Figure 3.4 shows the anthropogenic climate change projections for winter precipitation as derived from a pair of 30-year simulations, namely 2071-2100 and 1961-1990. Apart from the scenarios as introduced in section 3.2.2, the respective mean change of the scenarios forced with SRES A2 and B2, HadAM and ECHAM as well as the overall mean is shown in figure 3.4. The nine maps are to first order similar with increasing precipitation all over the Baltic catchment and in most of northern Europe. The major difference among the projections is located in an area just outside the Baltic catchment. Along the coastline in northwestern Norway, the HadAM simulations project a decrease in precipitation whereas the experiments driven with the ECHAM model show an increase.

Spatially averaged future changes in the Baltic catchment are larger in the ECHAM driven simulations (7.27 mm per decade) than in HadAM driven simulations (4.12 mm per decade) as shown in figure 3.4. In accordance with the stronger forcing, the mean response to the A2 emission scenario is higher than the response to B2 (6.59 and 4.80 mm per decade). Additionally, it is found that the spatial standard deviation of HadAM trends is lower (1.56 and 0.76 mm per decade for A2 and B2) than the standard deviation of the ECHAM simulations (2.38 and 1.39 respectively).

We have further assessed the significance of the projected anthropogenic change according to the regional model simulations. It is shown that for most of the region under consideration, the anthropogenic signal is hardly describable as a result of internal model variability. For the Baltic catchment, all of the changes as projected in the ECHAM driven runs have found to be significant at the 5% level (one-sided t-test taking into account autocorrelation, see section 3.2.4), whereas the fraction of significant changes is 97.7% (96.6%) for HadAM A2 (B2).
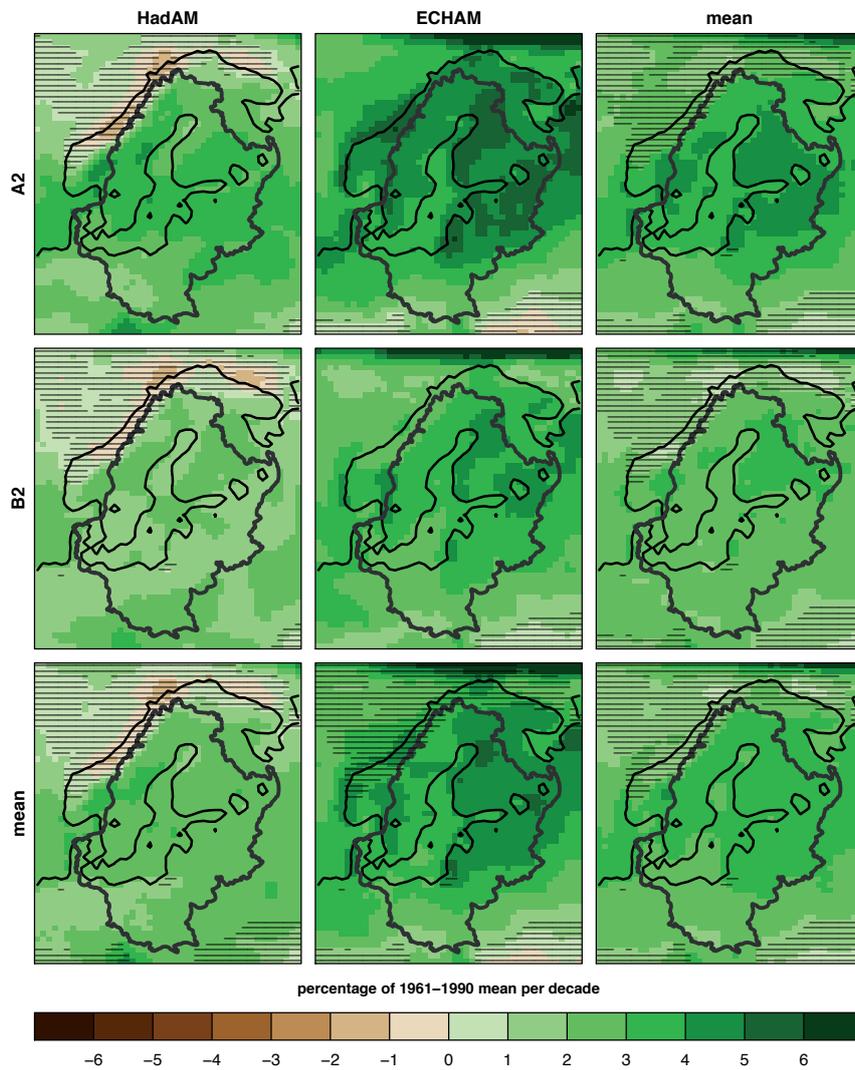
percentage of 1961–1990 mean per decade

Figure 3.4: *Anthropogenic climate change signal in DJF precipitation according to RCAO simulations with the four individual climate change projections in the upper left four panels and the respective mean of simulations driven with the same GCM (same emission scenario) in the bottom line (rightmost column). Negative projected changes are marked by thin solid contours. The hatched areas denote regions where the climate change projections are not significant at the 5% level (see section 3.2.2 for details); in the bottom line and rightmost column, areas where the projection of at least one of the four projections is not significant are hatched.*

Extending the research domain to all of northern Europe decreases the amount of grid boxes where the projections are significantly different from internal climate variability. Significant changes are found for 93.4 (90.9) percent of all land grid points in HadAM A2 (B2) and 90.7 (94.7) percent in ECHAM A2 (B2). As shown in figure 3.4, the regions where the projections are not significantly different from internal variability are located on the northwestern coast of Norway and Finland in the HadAM driven simulations, whereas insignificant changes are found in the southeastern part of northern Europe in the ECHAM driven simulations.

The similarity of the different climate change projections has been assessed using pattern correlation as introduced in section 3.2.4. As the observation data are available over land only, the PCCs between individual projections have been calculated using land grid boxes only as well. All of the climate change projections share very high PCCs of 0.941 (HadAM B2 with ECHAM A2) to 0.996 (ECHAM A2 with B2) for the Baltic catchment. When the area of interest is extended to wider northern Europe, the PCCs of projections driven with different GCMs are considerably reduced (0.831 to 0.928), whereas the patterns of simulations with different emission forcings with the same GCM are very similar with PCCs larger than 0.98.

Furthermore, the NAO signal has been removed from the simulations and climate change signals have been computed from the residuals. A consistent increase in the NAO index is found for all simulations. This change in the NAO index is stronger in the ECHAM simulations, with an increase in the difference between the normalized SLP series of the grid box Gibraltar and Reykjavik of 0.39 (0.61) per 110 years in the A2 (B2) simulation, than in the HadAM driven ones with an increase of 0.22 (0.20). The NAO signal in the simulations (not shown) is very similar to the NAO signal in the observations (e.g. figure 3.3), with increasing (decreasing) precipitation in northwestern Europe with increasing (decreasing) NAO index (not shown).

3.3.3 Correspondence of observed trends with anthropogenic climate change scenarios

The similarity of observed with possible future patterns of change is very much dependent on the scenario used (e.g. table 3.3). In general, the pattern of observed trends shows more similarity with the ECHAM driven simulations than

with HadAM ones. Furthermore, the stronger greenhouse gas forcing (SRES A2 scenario) leads to higher pattern correlation scores. Highest pattern similarity (0.85 for the Baltic catchment, 0.80 for northern Europe) is thus found when comparing the observed trends with the ECHAM A2 simulation.

Table 3.3: *Pattern correlation (see section 3.2.4) of precipitation in DJF (observation vs. simulation) for the Baltic catchment and northern Europe (in brackets). Significant (bootstrap test, 5 percent) results are labelled with an asterisk.*

$\rho_{R,E}$	CRU	NAO removed
HadAM A2	0.83* (0.63*)	0.76* (0.57)
HadAM B2	0.75* (0.54)	0.64 (0.45)
ECHAM A2	0.85* (0.80*)	0.75* (0.71*)
ECHAM B2	0.84* (0.77*)	0.74* (0.68*)

The observed PCCs based on the Baltic catchment all lie above the 95th according to the bootstrap introduced in section 3.2.4. This is not the case when extending the research area to all of northern Europe. The PCCs of the observed trends and ECHAM driven regional model simulations are significantly different from random PCCs whereas PCCs based on HadAM A2 lie close to and PCCs based on HadAM B2 lie below the 95th percentile of random PCCs.

The removal of the NAO signal leads to slightly different trend patterns in the observations as illustrated in figure 3.3 and in the climate change projections (not shown). In general, this causes a reduction of the PCCs which is most pronounced for the HadAM B2 simulation (reduction of 15 percent, table 3.3). In combination with the slightly higher significance levels when removing the NAO, PCCs are less often significantly different from random pattern correlation. For the Baltic catchment the pattern similarity is significant at the five percent level for all projections except the HadAM B2 scenario, whereas the PCCs for northern Europe fail to be significantly different from randomly generated PCCs for both HadAM simulations.

As a consequence of the normalization of the PCC by the intensities, no conclusions can be drawn about the similarity of the magnitude of the changes with pattern correlation. Therefore, we further compare the spatial mean change and the intensity of the change (table 3.2). In order to account for systematic biases

between model and observation data, we compare relative changes only. It is found that the changes in HadAM simulations are weaker than in ECHAM simulations and in accordance with the weaker GHG forcing in the SRES B2 driven projections weaker than in simulations run with SRES A2. Thus leading to an area mean change of 1.89 (HadAM B2) to 4.38 percent per decade (with respect to the 1961-90 mean) for the Baltic catchment. The spatial mean change in the observation data is considerably higher with 6.97 percent per decade for CRU for the period 1973-2002. The discrepancy between the climate change scenarios and observed changes is even larger when looking at the intensity of the change, and thus taking the spatial variability of the trend fields into account as well. For wider northern Europe, the spatial statistics show qualitatively similar features (not shown).

When looking at the spatial statistics after removing the NAO signal from the data, we find the following. As discussed in section 3.3.1, the removal of the NAO leads to a considerable reduction in the spatial mean trends for the period from 1973-2002. As the removal affects mainly the spatial mean trend, the intensities are still much higher in the observations after removing the NAO signal than in the climate change scenarios.

Both PCCs as well as the intensities are computed limited to regions of significant change in the climate change projections as well. The effect on the results for the Baltic catchment is negligible as only very few grid points are excluded from the analysis. For northern Europe, however, we find a considerable increase in pattern similarity statistics which is more pronounced for the HadAM projections than for projections driven with ECHAM. Weakest pattern similarity is still found when comparing the observations with HadAM B2. Focussing on regions with significant changes leads to a slight increase in the intensities of the climate change projections and thus decreases the ratio of intensities.

3.4 Discussion

3.4.1 Methodical considerations

It is shown that PCCs are sensitive to both the magnitude of the mean change and the pattern of the change as an extension of the analysis from the Baltic catchment to all of northern Europe and the comparison of results with A2 and

B2 emission scenarios show. Thus we conclude, that the method is illustrative even in situations where the climate change scenarios deviate to some extent.

3.4.2 Regional climate change scenarios

We use a set of climate change projections in our analysis indicating the range of the response to anthropogenic forcing. A priori, we consider all of the individual projections as possible and equally likely. The projected signal for the Baltic catchment is fairly consistent in both magnitude and pattern taking into account the differences in GHG forcing for the A2 and B2 scenarios (e.g. CO_2 induced radiative forcing of 4.42 (2.73) Wm^{-2} in 2100 with respect to 2000 for the SRES A2 (B2) scenario, Ramaswamy et al. 2001). However, there are still large discrepancies in the way different GCMs model the response of circulation (and as a consequence precipitation as well) to anthropogenic forcing as illustrated by the situation along the Norwegian coast.

3.4.3 Comparison of observed and simulated changes

Pattern correlation

The pattern of observed trends in winter precipitation in the Baltic catchment has been found to be consistent with all of the regional climate change scenarios used in this analysis (table 3.3). Furthermore, it is very improbable (with a probability of error of less than 5%) that the correlation found between patterns of observed and expected changes is random. This holds true for northern Europe as well, however, the observed patterns are considerably less consistent with the HadAM driven scenarios, due to a completely different response in some parts of northern Europe compared with the ECHAM scenarios. Even though the scenarios differ much more when extending the research area from the Baltic catchment to northern European land areas, we conclude that the recently observed pattern of change is consistent with the climate change scenarios at least for the projections driven with ECHAM.

Above findings are strengthened when limiting the analysis to regions where the changes in the scenarios have been found to be significant according to an adjusted t-test. The exclusion of regions with insignificant climate change

projections increases the signal-to-noise ratio of anthropogenic change (assuming that the climate models are able to model both the natural variability and the response to anthropogenic forcing correctly) and thus it is not surprising that we find better consistency between the modelled response and observed changes. Furthermore, the differences in consistency between HadAM and ECHAM driven simulations for northern Europe vanishes, as the regions with insignificant changes in the projections are in this case identical to the regions where the climate change scenarios differ most.

Further insight is gained when removing the NAO signal from both the observations and the climate change simulations. The correspondence of the changes in the residuals is considerably lower than in the full set, and we conclude consistency between observed and projected changes only for ECHAM driven scenarios (and HadAM A2 as well for the Baltic catchment).

The hypothesis that the signal-to-noise ratio of anthropogenic change in the NAO would be low is discussed in the following. The projected increase in the NAO index from 1961-1990 to 2071-2100 ranges from 0.2 in the HadAM simulations to 0.38 (0.61) in ECHAM A2 (B2). Thus, a consistent increase is shown among all four climate change projections, which is in line with the findings of Stephenson et al. (2006). However, the increase is significant on the five percent level for the ECHAM B2 scenario only according to a t-test as introduced in section 3.2.4. Thus most of the changes in the NAO in the set of model simulations could be due to internal variability alone, which in turn supports the basic assumption that the signal-to-noise ratio of anthropogenic change in the NAO is low. In opposition to these findings, Gillett et al. (2002a) find that the response of the NAO to GHG forcing in both the ECHAM4 and HadCM3 (a coupled version of the model used to drive the RCM simulations used in this study) is not explicable by internal variability alone. This discrepancy could result from differences in the definition of the NAO index; Osborn (2004) shows that the response to GHG forcing is much more model dependent when using a station based NAO index compared to pattern based indices.

Additionally, the considerable differences between ECHAM and HadAM driven simulations point again towards a different response of circulation to anthropogenic forcing in these models. Furthermore, it is interesting to see that the stronger greenhouse gas forcing leads to a weaker increase in the NAO in the ECHAM driven scenarios. This could either be a consequence of a nonlinear re-

37

sponse to the imposed forcing or due to the dominance of natural variability in the NAO estimates.

The observed trend in the NAO index of Jones et al. (1997) from 1973 to 2003 is 0.292 per decade. Thus, the projected increase in the NAO index amounts to 6-19% of the observed trend in the NAO only. It is a well-known fact that present day climate models underestimate the variability of the NAO and presumably also the response to increasing GHG (Osborn, 2004, Stephenson et al., 2006). Thus we conclude that either the projections of the NAO increase are correct and hence the signal-to-noise ratio of anthropogenic change in the NAO is low, or the response of the NAO to increasing GHG is stronger than simulated. The latter would have severe consequences for all conclusions drawn from these regional climate change projections as a stronger response of the NAO to GHG forcing would very likely lead to a stronger response of precipitation as well.

Magnitude of the rate of change

When comparing the spatial mean change, we find that the models underestimate the most recent rate of change by a factor of 1.4 (ECHAM A2) to 3.3 (HadAM B2) for the Baltic catchment. The same applies for all of northern Europe as well. In contrast, when removing the NAO signal, we find considerably better agreement of the observed area mean change with the climate change projections. Nevertheless, it is shown that the climate change simulations generally underestimate the observed change. Whether this mismatch in magnitude of the area mean changes is in any way significant is hard to infer from the data at hand. With respect to uncertainties in the observation data and interpolation, further experiments have shown, that for the Baltic catchment, the most recent area mean trends are significantly (with a probability of alpha-type error of 5 percent) different from the estimated area mean response when adding white noise with a standard deviation of 5.6 percent of the respective 1961-90 mean. However, as systematic biases have not been removed from the data (New et al., 1999, Mitchell and Jones, 2005) the error mainly due to wind-induced undercatch could be larger (Adam and Lettenmaier, 2003, Yang et al., 2005). Furthermore, the robustness of the conclusions to a shift of the period analyzed has been investigated. For the Baltic catchment area mean changes we find that 8 (11) of the last 10 (20) 30-year trends available are higher than the topmost anthropogenic change estimate (ECHAM A2). Whether these results indicate an emerging anthropogenic signal or fluctuations due to natural variability cannot be inferred.

Finally, it should be noted that even though we analyze relative precipitation changes, and thus systematic biases between the simulated and observed precipitation have little influence on the result, modelling deficiencies could severely influence the conclusions drawn.

As mentioned before, given all assumptions are correct, there are three possible reasons for this mismatch in the spatial mean change.

First of all, it could be due to the fact that the regional models are not able at all to simulate the response to anthropogenic forcing.

Second, the models could severely underestimate the response because more than the imposed forcing is at work or because the sensitivity to anthropogenic forcing is far to low. Gillett et al. (2004) and Zhang et al. (2007) conclude that the GCMs used in their global-scale detection and attribution studies considerably underestimate the response of circulation and precipitation changes to external forcing. However, Lambert et al. (2004) also find a strong influence of volcanic forcing. In contrast to the studies mentioned above, our simulations include anthropogenic changes only, thus we should keep in mind, that the response to natural forcing could be dominant in the observations. Apart from changes in aerosols due to volcanic eruptions, the main candidate for an additional forcing mechanism which could have a large effect on regional circulation and precipitation are industrial aerosols. According to (Räisänen et al., 2004), the ECHAM and the HadAM model both include changes in global scale aerosol concentrations, and treat the contribution to the radiative forcing explicitly. Furthermore, they argue, that the RCAO is not very sensitive to local changes in radiative forcing, since most of the climate change signal comes via the boundary conditions from the GCMs. However, mainly the indirect aerosol effect could cause strong and small-scale response in precipitation (Ramanathan et al., 2001) not captured in the models yet. Additionally, the scientific understanding of both the direct and indirect aerosol effects is still considered medium to low (Forster et al., 2007), and thus there is ample room for speculation.

Third, the signal-to-noise ratio of anthropogenic precipitation change is very small. Assuming that the model projections are right in both intensity and pattern, we conclude the following: a large fraction of the recent 30-year trends in wintertime precipitation are due to natural variability. The removal of the NAO signal leads to a considerable decrease in the ratio of the intensities as shown in table 3.2. Thus, by excluding the NAO signal we increase the signal-to-noise

ratio, which in turn supports the basic assumption, that the signal-to-noise ratio in the NAO is low.

Looking at the intensity of the change, and thus taking the spatial variability of the changes into account as well, the difference between observations and climate change simulations increases. Whether this is due to the fact that the climate change signal is large scale, and thus exhibits only little spatial variability over a small domain, cannot be determined. Alternatively, the different spatial scales represented in the gridded observations compared to the scales modelled in the RCM setup, could account for the difference in intensity as well.

3.5 Conclusions

Our analyses have shown that pattern correlation along with a comparison of the intensity of the changes as presented in this paper is suitable to assess the consistency of observed trends with climate change projections. The method as presented here is also applicable when natural variability estimates are missing and thus it is very useful when investigating climatic parameters for which long-term observations are not available and which are statistically less well behaved than surface temperature.

According to pattern correlation studies, anthropogenic forcing is a plausible explanation of the observed changes in wintertime precipitation over the Baltic catchment. Bootstrap experiments also show that it is unlikely that these pattern correlations are random. The situation is a little different when extending the area of interest to all of northern Europe. In this larger area, the climate model simulations project less consistent changes and consequently PCCs are only significantly different from random PCCs for some of the simulations. Thus the selection of the region under consideration has a great effect on the result. However, it is encouraging that we find consistency of the observed trends with regional climate change scenarios in regions where the different simulations project a consistent and significant change and less so in regions where the climate change scenarios differ.

The magnitude of the observed area mean change, however, is higher than the magnitudes as projected by the regional climate model. Hence, we cannot explain the observed trends in winter precipitation with increasing greenhouse gases alone. Both additional forcing mechanisms (such as the indirect aerosol

effect) not included in this model setup, and a general underestimate of the response to anthropogenic forcing are possible explanations for the mismatch in the rate of change. Additionally, another important factor possibly contributing to the trends in the observation data is natural variability, the importance of which cannot be inferred using the approach introduced here.

Acknowledgements

4 Is greenhouse gas forcing a plausible explanation for the observed warming in the Baltic Sea catchment area?

This chapter is based on an atricle published in *Boreal Environment Research* (Bhend and von Storch, 2009). The original manuscript has been slightly changed for editorial purposes.

4.1 Introduction

We compare observed changes in screen temperatures with the response to anthropogenic forcing, in our case changing greenhouse gas (GHG) and aerosol concentrations in the atmosphere, estimated from simulations with a regional atmosphere-ocean climate model. Since a large fraction of the variability of temperature in Northern Europe is linked to the North Atlantic Oscillation (NAO; van Loon and Rogers 1978, Lamb and Peppler 1987, for a review see Wanner et al. 2001), we also compare the changes after removing the fingerprint of the NAO from both the observations and the anthropogenic change estimates. Hypothesizing that the anthropogenic signal in the NAO is low, we expect an increase in correspondence of the observed and expected change after removing the NAO signal.

By comparing observed changes to the estimated response to a given forcing mechanism, we offer the possibility to falsify the hypothesis that the observed change is mostly related to this mechanism. Thus, we are looking for the consistency of the observed change with the change related to an *a-priori* known forcing (for a more in-depth discussion on the methodology see section 3.1). This consistency check could be thought of as a sort of "attribution without detection". In this set-up, the recent change is given; the expected signal, however, is unknown and has to be estimated from model simulations. As there are only a few simulations from coupled regional climate models available, we cannot

derive a reasonable estimate of the variability of the response to anthropogenic forcing (based on sufficiently many independent realizations). Therefore, we do not frame our examination as a formal statistical test with the null hypothesis "the observed change is drawn from the set of scenarios". We collect plausibility arguments from the quantitative comparison of the few scenarios with the observed change instead. Even if inferior to a full statistical test, our method generates instructive results in particular in cases where only little data exist.

If we conclude that the observed change is not consistent with the assumed forcing mechanism, then three possible reasons may be thought of: The simulated response to anthropogenic forcing is wrong, the assumed mechanism is overridden by another mechanism not accounted for in the simulations, or the signal-to-noise ratio of an anthropogenic change is too low and thus a large fraction of the observed trend is due to natural variability. On the other hand, if we find consistency of the observed change with the assumed forcing mechanism, we conclude that the assumed forcing is a plausible explanation of the observed change. Note, however, that consistency is not equivalent to the rejection of the "detection"-hypothesis, "the observed change is drawn from a set of changes due to natural variability."

4.2 Materials and Methods

4.2.1 Observed temperature trends

The observed changes are computed based on the monthly gridded temperature data set CRU TS 2.1 of the Climatic Research Unit (Mitchell and Jones, 2005). These data are available over land on a regular latitude longitude grid with 0.5̨ grid resolution for the period from 1901 to 2002. We define recent change as the trend over the last 30 years available, i.e. 1973–2002, and we estimate the slope of the regression line using least squares. The domain of interest is the Baltic Sea catchment area (Fig. 4.1).

4.2.2 Climate change scenarios

We use a set of simulations with the Rossby Centre regional Atmosphere-Ocean model (RCAO) of the Swedish Meteorological and Hydrological Institute (Kjell-
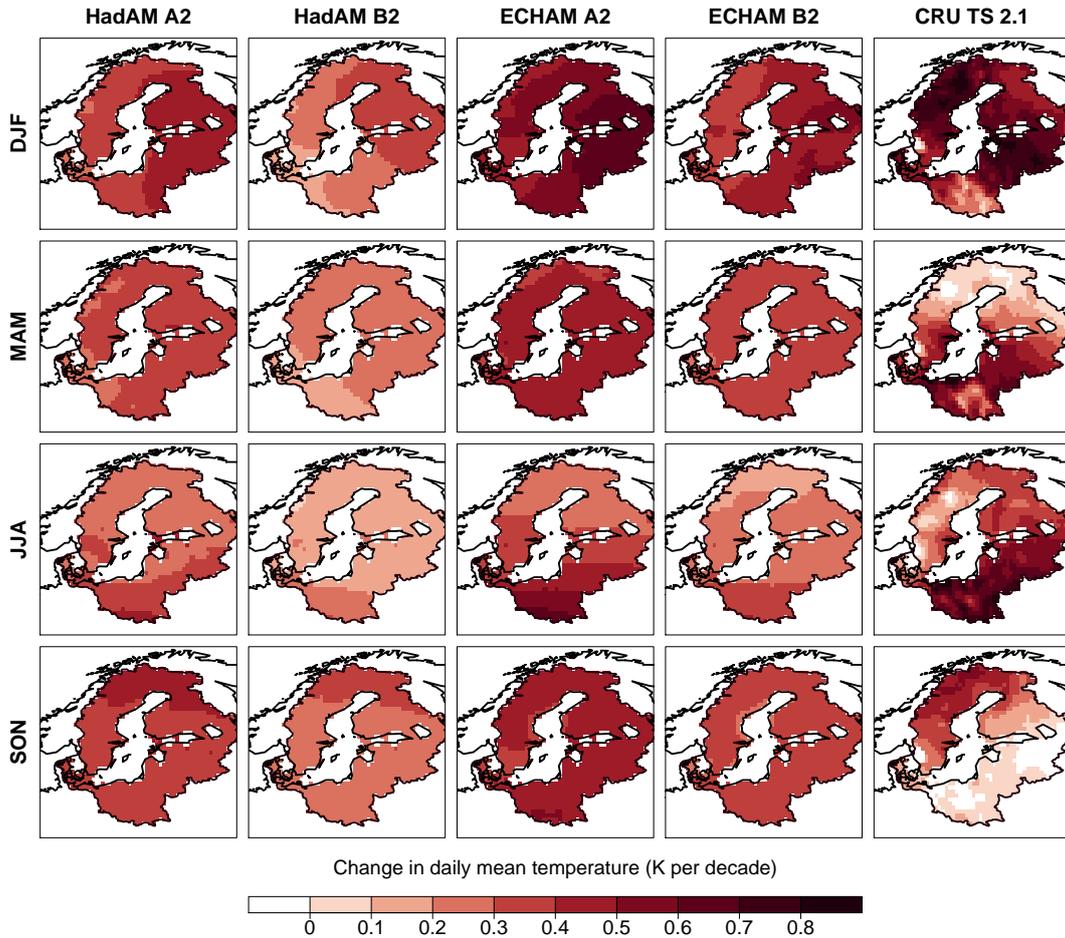
Change in daily mean temperature (K per decade)

Figure 4.1: *Seasonal anthropogenic climate change signal for daily mean temperature in the Baltic catchment according to the different climate change simulations with the RCAO model (first four columns) and the observed trend for the period 1973 to 2002 according to the CRU TS 2.1 data set (rightmost column).*

ström, 2004, Räisänen et al., 2004). The set of simulations includes six time slice experiments of 30 years driven with two different global models, the HadAM3h (Gordon et al., 2000) and the ECHAM4/OPYC (Roeckner et al., 1999). For each of the two different GCMs, one simulation with GHG and aerosol concentrations for the period 1961–1990 and two simulations for the period 2071–2100 using GHG and aerosol concentrations from the SRES A2 and B2 scenarios are available.

The response to anthropogenic forcing for a certain period would ideally be estimated from transient climate model simulations with and without anthropogenic forcing. However, as the low signal-to-noise ratio of the anthropogenic change on the regional scale would require a very large ensemble of transient simulations, which is presently not available, we estimate the anthropogenic signal from time slice experiments. In order to be able to do so, we have to make assumptions on the evolution of the anthropogenic signal. First, we assume that the spatial pattern of the response does not depend on the temporal evolution of the forcing. This hypothesis is supported by different analyses of global scale simulations (e.g. Meehl et al. 2007a: fig. 10.8). Second, we assume a linear development from 1961 to 2100, and thus we presumably overestimate the anthropogenic signal in the last decades of the 20th and the first decades of the 21st century. However, compared to the differences between the single realizations (due to different forcing and different driving GCMs) this possible overestimate is negligible.

We define the anthropogenic signal as the difference between the future (2071–2100) and control (1961–90) period mean climate. The resulting signal is further linearly scaled to change per decade. In accordance with the GCM and emission scenario used, we refer to the respective anthropogenic change estimates as HadAM A2 and B2 and ECHAM A2 and B2.

4.2.3 NAO representations and NAO signals

We use a station based NAO index, which is defined as the difference in normalized monthly sea level pressure (SLP) between Reykjavik and Gibraltar according to (Jones et al., 1997). The NAO index in the set of regional climate model simulations has been derived accordingly from the respective SLP fields. The reference period for the normalization of SLP time series is 1961–1990. The variability in the dimensionless NAO index based on observations is higher than the variability based on the two different 1961–1990 representations in the regional model

simulations with a standard deviation of 1.46 in the observations and 1.14 (1.09) in the HadAM (ECHAM) simulation.

The fingerprint of the NAO is defined as the fraction of the variability in monthly and seasonal mean temperature that covaries with the respective NAO index. We regress the detrended temperature time series on the detrended NAO index for each grid box separately using ordinary least squares estimation of the parameters of the linear regression. The slope of the regression is the NAO fingerprint. We remove the NAO fingerprint from the observations by subtracting the fingerprint times the trend in the NAO index from the trend in the observations. From the climate change projections, we remove the NAO fingerprint by simply subtracting the respective NAO fingerprint times the difference in the average NAO index for the periods 1961–90 and 2071–2100.

4.2.4 Pattern correlation

The comparison of the observed (O) and expected (E) pattern of change is carried out using pattern correlation. As the human influence on climate leads to both a change in the mean climate as well as a change in the spatial and/or temporal patterns, we use uncentered pattern correlation as described in Eq. 1.

$$R(O, E) = \frac{\sum_i O_i E_i}{\sqrt{\sum_i O_i^2 \sum_i E_i^2}} \tag{4.1}$$

Where $R(O, E)$ is the pattern correlation coefficient; O refers to the observed, E to the expected pattern of change and i represents the dimensionality of the pattern. We investigate patterns with different spatio-temporal resolution. These include the mean annual temperature changes both as area averages (O_{ann}, E_{ann}) and as spatially 0.5ą-distributed fields ($O_{ann,x}$, $E_{ann,x}$), and the seasonally and monthly resolved annual cycles of the change, again as area averages (e.g. O_{seas}, O_{mon}) and as spatially distributed fields (e.g. $O_{seas,x}$, $O_{mon,x}$). For the respective pattern correlation coefficients, the shorthand $R_{seas,x}$ is used for $R(O_{seas,x}, E_{seas,x})$.

4.3 Results

The observed change in annual near surface temperature (O_{ann}) for the Baltic Sea catchment area is 0.37 K per decade. The respective anthropogenic climate change signal (E_{ann}) ranges from 0.24 to 0.45 K per decade (for HadAM B2 and ECHAM A2 respectively). The observed change lies inside the range of expected changes and thus we conclude consistency of O_{ann} with E_{ann}. However, due to the strong interannual variability, the trend component in the observed 30-year time series accounts for only 11.7 percent of the total variance. After the signal of the NAO has been removed from both the observations and climate change simulations, we find a reduction of O_{ann} to 0.32 K per decade and a slight reduction in E_{ann} to 0.24 to 0.43 K per decade.

The annual cycle of the average anthropogenic change for the Baltic Sea catchment area is characterized by a stronger warming in E_{seas} in winter (DJF) than in summer (JJA, see Fig. 4.2). In contrast to E_{seas}, the annual cycle of the observed warming shows smallest warming in autumn. The differences between the expected and observed patterns increase, if we increase the temporal resolution and compare O_{mon} with E_{mon}. The anthropogenic change estimates derived from simulations driven with ECHAM show a pronounced annual cycle in E_{mon} with strongest warming in February and weakest warming in June and July. The simulations driven with HadAM show a slightly different pattern with smaller amplitude than the annual cycle of E_{mon} in ECHAM A2 and B2.

We find pronounced spatial variability in the pattern of observed changes ($O_{seas,x}$) with enhanced warming in the northern part of the Baltic Sea catchment area in winter (DJF) and autumn (SON) and in the southern part in spring (MAM) and summer (JJA). In the anthropogenic change estimates, however, spatial variability is much smaller and the annual cycle of the area mean changes dominates the combined spatio-temporal pattern (Fig. 4.1).

The pattern correlation coefficients for the different indices of surface temperature change and the different anthropogenic change estimates used are highest for the annual change with full resolution ($R_{ann,x}$), with values around 0.97 (Fig. 4.3). When investigating the annual cycle, we find highest pattern correlation scores for $R_{seas,x}$ with values larger than 0.9. With increasing spatial and temporal detail, the pattern similarity decreases. If the pattern of change is analyzed after the fingerprint of the NAO has been removed from both the observations
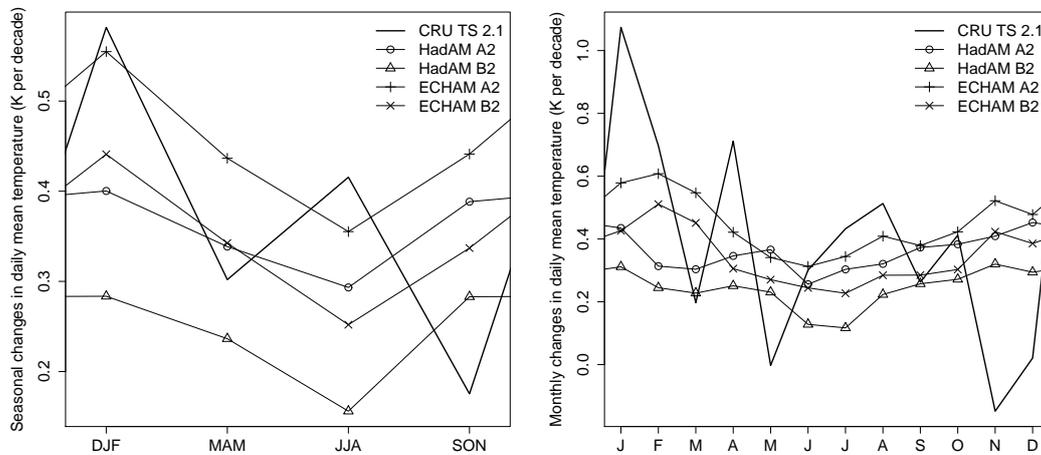
Figure 4.2: *Seasonal (left panel) and monthly (right panel) area mean changes in near surface temperature for the Baltic catchment. The observed trends for the period 1973 to 2002 are represented by the heavy black line, the different anthropogenic change estimates derived from a set of simulations with the RCAO model are represented by the thin black lines and different symbols.*

and climate change scenarios, we find slightly reduced pattern similarity compared to the pattern correlation with the full signal.

4.4 Discussion

The observed change in annual near surface temperature in the Baltic Sea catchment area is consistent with the anthropogenic change estimates derived from regional climate model simulations. For the annual area mean change, we assess consistency by comparing the magnitude of the simulated change with the observed trend. The magnitude of the observed trend, however, is strongly dependent on the trend length used. With increasing trend length, the magnitude of the most recent trend decreases almost linearly from 0.37 K per decade for 30-year trends to 0.06 to 0.07 K per decade for 70-year and longer trends (for a review of trends of different length in subregions and at individual stations see The BACC author team 2008). The observed decrease in magnitude of the most recent trends with increasing trend length could be due to the weaker
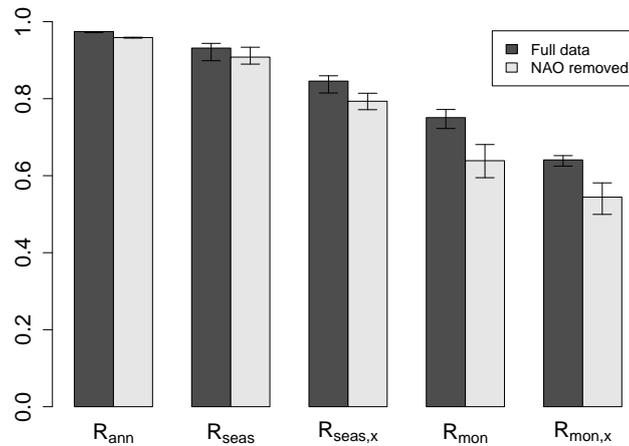
Figure 4.3: *Uncentered pattern correlation between different patterns of observed trends for the period 1973–2002 and the respective anthropogenic change patterns. The spatio-temporal resolution increases from left to right. The shaded bars indicate pattern correlation between the observations and the mean anthropogenic change signal of all 4 different members investigated, the whiskers denote the spread of pattern correlation of the observed change with the individual anthropogenic change signals.*

anthropogenic signal in the earlier parts of the 20th century or due to the decreasing influence of natural variability with increasing trend length. The most recent 30-year and 40-year trends in surface temperature are consistent with the anthropogenic change estimate; 50-year trends in annual area mean temperature of 0.23 K per decade lie slightly below the lowest anthropogenic change estimate (HadAM B2, 0.24 K per decade). We should note, however, that the anthropogenic change estimates are most likely biased high due to the assumption of a linear development of the anthropogenic signal from 1961–2002. From the above-presented evidence we conclude, that the anthropogenic forcing is a plausible explanation for the observed annual area mean trend and that this finding is robust to the choice of trend length. Without reference to an estimate of natural variability and additional forcings, however, we are not able to rule out alternative explanations. This is a major limitation of the method presented here.

The annual cycle in the climate change scenarios (E_{seas}) is different from the annual cycle in O_{seas}. The observations show larger amplitude between minimum

and maximum seasonal warming and a different structure with minimal warming occurring in SON instead of JJA as in the scenarios (left panel in Fig. 4.2). Increasing the temporal resolution results in a noisier pattern in the observations whereas the simulated annual cycle of changes remains smooth (right panel in Fig. 4.2). The lower amplitude in the simulated annual cycle of the temperature change could result from deficiencies in the representation of the large-scale circulation variability related to the NAO and atmospheric blocking in the driving GCM (D'Andrea et al., 1998, Osborn, 2004, Stephenson et al., 2006). Furthermore, the differences in O_{mon} from month to month point towards a low signal-to-noise ratio of an anthropogenic change in monthly area mean temperatures. Part of the observed variability in the annual cycle could be due to random variations and thus not explicable by the response to external forcing. We conclude that anthropogenic forcing is an insufficient explanation for the observed annual cycle in seasonal and monthly temperature changes in the Baltic Sea area.

The pattern similarity decreases with increasing spatial and temporal resolution (Fig. 4.3). Consequently, the signal-to-noise ratio of an anthropogenic change decreases with increasing spatio-temporal detail. This is in line with findings of various authors (e.g. Stott and Tett 1998, Zwiers and Zhang 2003). We conclude that either the anthropogenic climate change signal is large scale (compared to the size of the domain under investigation) or the RCM used is not able to simulate regional details with sufficient quality — given GHG concentrations as sole forcing.

Removing the NAO signal from the data previous to the analysis results in a slight reduction of the pattern correlation. We find that the signal-to-noise ratio of an anthropogenic change is reduced by removing the NAO signal. This indicates that our initial assumption of a negligible anthropogenic effect on the NAO may be incorrect. This conclusion is in line with Gillett et al. (2002b), who showed that the observed increase in the NAO index during the last decades of the 20th century is not attributable to natural variability alone and thus a significant part of the observed changes in the NAO is anthropogenic. However, anthropogenic forcing of the NAO remains a matter of debate, as present-day GCMs are not yet able to reproduce fully the observed NAO variability (Osborn, 2004, Stephenson et al., 2006).

4.5 Conclusions

Often it is assumed that the observed regional climate change is at least partly anthropogenic. As a plausibility check, we propose a simple consistency analysis. We therefore compare the most recent observed changes in near surface temperatures with an estimate of the anthropogenic signal as proposed by a set of regional climate model simulations. Comparing the annual area mean change in surface temperature in the Baltic catchment, we find good correspondence with the available climate change scenarios. The magnitude of the annual change is also similar after the NAO signal has been removed from both the observations and the simulations.

In order to further estimate the level of detail to which observed regional features could be explained by anthropogenic change we assess the pattern similarity of climate change patterns with different spatio-temporal resolution. In all cases, the similarity of observed and expected patterns of change decreases with increasing level of detail. Whether the loss in similarity with increasing spatial resolution is due to limitations in our ability in simulating small-scale structures of the climate change pattern, or a consequence of the climate change signal being in fact large scale we do not know. Nevertheless, it is shown that anthropogenic forcing is a plausible explanation for the observed area mean changes of annual temperature in the Baltic catchment. In contrast, anthropogenic forcing only partly explains the annual cycle of the observed change. Of course, we cannot exclude the possibility that both additional external forcings and internally generated variability may be similarly powerful in explaining the observed changes. Unfortunately, there are no comprehensive regional simulations on alternative forcing mechanisms available yet.

Acknowledgements

5 Detection and attribution of an anthropogenic influence on the observed change in temperature and precipitation

5.1 Introduction

In chapters 3 and 4 we have analyzed if the spatial pattern of the change in temperature and precipitation in the Baltic Sea catchment is consistent with regional climate change projections. In this chapter we focus on the temporal evolution of area-average temperature and precipitation.

The temperature evolution in northern Europe shows the same main features as global temperature evolution over the 20th century (figure 1.2). Two periods of distinct warming can be identified, the warming early in the 20th century until the 1930s, and the recent warming since the 1970s. These warming periods are interrupted by a period of temperature stagnation or even slight cooling around the 1950s. The most recent warming in the Baltic Sea catchment, however, is roughly twice as fast as the global warming, with regional amplification of the warming rates being largest in winter and spring (van Oldenborgh et al., 2009). The similarity of the global and northern European temperature evolution over the 20th century suggests that the same main causes lead to warming at the global and regional scale. The global warming is well reproduced by present-day climate models and its causes have been attributed to human influence (for a review see Hegerl et al., 2007). The regional amplification of the warming, on the other hand, indicates that the global-scale response is significantly modified at the regional scale.

Changes due to anthropogenic forcing are best expressed in linear trends or some more sophisticated metric that retains low-frequency variability, since the forcing is varying slowly as well. If we want to be able to distinguish between changes due to natural and anthropogenic forcings, however, trends are not a suitable quantity to look at for two reasons: First, the response to natural

forcing does not project well onto a linear trend, as solar forcing is rather a periodic oscillation and volcanic forcing is episodic with three major eruptions in the second half of the 20th century in 1963 (Mount Agung), 1982 (El Chichón), and 1991 (Mount Pinatubo). Second, trend patterns due to anthropogenic and natural forcing are highly correlated at the regional scale, as the spatial pattern and the annual cycle of the warming (or cooling) induced by either forcing are – as a first approximation – modified and amplified by the same processes and feedbacks. In contrast to other forcings, volcanic forcing induces cooling in summer and warming in winter over northern Europe (Shindell et al., 2004, Fischer et al., 2007). The response to volcanic forcing should thus be clearly distinguishable from the responses to anthropogenic or solar forcing, which rather lead to consistent warming or cooling in all seasons. The attribution of the observed change to volcanic forcing is complicated by the strong variability in the response to individual eruptions (Shindell et al., 2004) and by the limitations of present-day models in reproducing the circulation response to volcanic forcing and the resulting underestimation of winter warming in the models (Stenchikov et al., 2006). Nevertheless, it is advantageous to include the annual cycle of the change in the searched-for pattern. Therefore, we jointly analyze changes in all four seasons.

The global change in precipitation is - in contrast to the observed global warming - less well understood. Lambert et al. (2004) find that the evolution of global area-average land precipitation from 1940 to 1998 is significantly different from internal variability. The anthropogenic influence on global precipitation changes has recently been detected in the pattern of trends in annual zonally averaged precipitation by Zhang et al. (2007). The two studies illustrate that either the temporal or spatial pattern of precipitation change could be used to detect externally forced changes.

We use data from global AOGCMs in order to formally detect and attribute changes in near-surface temperature and precipitation. The analyses presented in chapter 2 reveal that dynamical downscaling has only a minor effect on the representation of the variability and response to anthropogenic forcing in area-average precipitation in the Baltic Sea catchment. The effect of downscaling on temperature, in contrast, is a consistent reduction in both variability and response to anthropogenic forcing. In addition, we note that there are significant systematic model biases in the representation of the mean climate in this region due to misrepresented small-scale processes such as snow-melt and convective precipitation. Furthermore, we stress that the variability in area-average precip-

itation is underestimated in all the models analyzed. As a first-order correction, we thus inflate the variability in the models to better match the observations. Nonetheless, we recommend to interprete detection and attribution results for precipitation with caution.

Precipitation and temperature observations and model simulations used in this chapter are presented in the following section. In section 5.3, we introduce the optimal fingerprint method used in the detection and attribution analysis. In section 5.4, we investigate if we can detect externally forced changes - that is, if the observed change in the Baltic Sea catchment is significantly different from internal variability. We further analyze the influence of the model used to derive the searched-for signals, the influence of the length of the time period analyzed, and the influence of the spatiotemporal resolution of the data on the detection result. In section 5.5, we investigate whether the relative importance of anthropogenic or natural drivers can be further quantified in a two-signal attribution analysis. The detection and attribution analyses are then further substantiated with perfect-model experiments in section 5.6. That is, we use model simulations as pseudo-observations to investigate potential detectability and attribution in the model world.

5.2 Data

5.2.1 Observed change

The observed change in near-surface temperature over land is estimated from the gridded monthly land surface temperature data of the Climatic Research Unit (CRU, Brohan et al., 2006, Jones and Moberg, 2003). This data set - referred to as CRUTEM3v - consists of homogenized and quality controlled station data from 1850 to present. The station data has been interpolated to a 5° by 5° latitude-longitude grid by averaging temperature anomalies of all the respective stations within one grid cell. Additionally, the variance of the grid-box time series has been adjusted in order to take into account changes in the number of stations that contribute to a grid box average (Jones et al., 2001a).

For precipitation, we use version 4 of the "Full Data Reanalysis Product" of the Global Precipitation Climatology Centre (GPCC, Schneider et al., 2008). These data include quality controlled rain gauge data from 1901 to 2007 that have

been interpolated to a 0.5° by 0.5° grid using the SPHEREMAP method (Shepard, 1968, Willmott et al., 1985). We use the GPCC's aggregated 2.5° by 2.5° gridded precipitation data in this study. We further aggregate the precipitation data to the 5° by 5° latitude-longitude grid of the CRUTEM3v data. The precipitation data are referred to as GPCC v4.

The number of rain-gauges included in the GPCC v4 gridded product varies considerably both in space and in time (figure 5.1). The average number of stations per 5°x5° grid box in the Baltic Sea catchment is considerably higher in the densely populated southern part of the Baltic Sea catchment (dashed line in figure 5.1) than in the rural northern part (dotted line). The number of rain-gauge measurements included in the GPCC database increases slowly during the first half of the 20th century with a remarkable drop in 1945 in the southern part of the Baltic catchment. Data coverage is considerably higher after 1950 especially in the southern part with best coverage from 1985 to 2000 and a rapid decrease of the number of rain-gauges included in the data set thereafter due to the delayed availability of non real-time measurements.
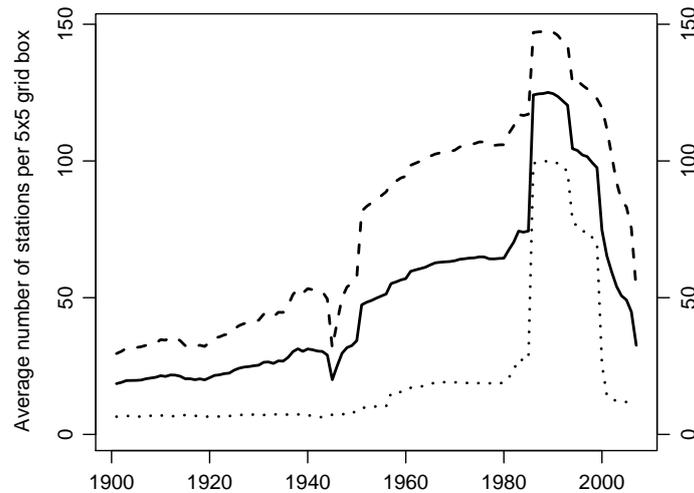
Figure 5.1: *Number of rain-gauge stations contributing on average to every 5°x5° grid box in the entire Baltic catchment (solid black line), the Baltic catchment south of 60° N (dashed line) and north of 60° N (dotted line).*

Additional observation and reanalysis data sets have been used in the detection and attribution analysis in order to confirm the results. As the results do not

depend on the data used to estimate the observed change (except for the rean-
alyzed precipitation), we limit the discussion to the above mentioned two data
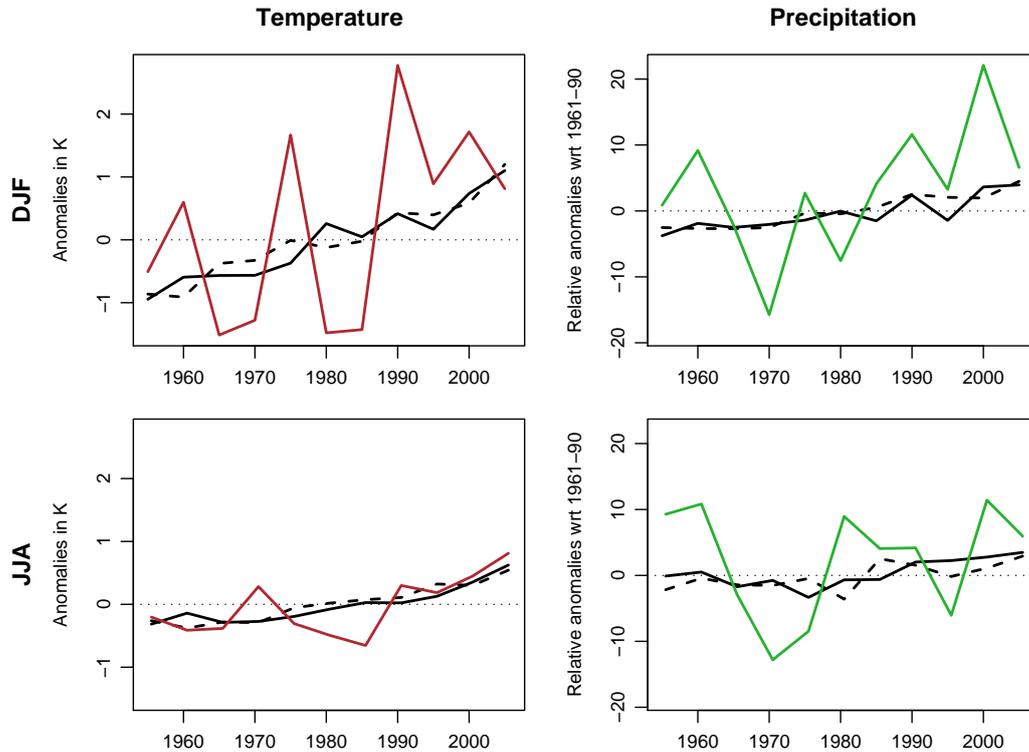sets for precipitation and temperature.

Figure 5.2: *Time series of observed 5-yearly averages of winter (DJF) and summer
(JJA) area-average temperature anomalies (left column, according to CRUTEM3v
in red) and relative precipitation anomalies (right column, according to GPCC v4
in green) along with the climate change signals as derived from the simulations
in the CMIP3 multi-model ensemble with anthropogenic forcing only (dashed
black lines) and anthropogenic and natural forcing (solid black lines).*

5.2.2 Climate change signals

The climate change signals are derived from model simulations with global
atmosphere-ocean general circulation models (AOGCMs) from the WCRP CMIP3

database (Meehl et al., 2007b). We refer to the estimated response to external forcing derived from GCM data as guess patterns in order to express the uncertainty in the hypothesized response. In order to minimize the influence of internal variability on the guess patterns, we average across a large number of simulations and models.

We divide the multi-model ensemble in two parts: one containing models with temporally varying anthropogenic forcing only, the other containing models with temporally varying anthropogenic and natural forcings over the 20th century (see table 5.1). In order to be able to derive guess patterns for the period from 1901 to 2007, for which both temperature and precipitation observations are available, we combine the 20th century runs that end in 2000 with the projections for the 21st century driven by anthropogenic forcings according to the SRES A1B emission scenario. This is a valid approach for three reasons: First, the scenario simulations have been continued from the 20th century simulations; combining them does therefore not introduce inhomogeneities. Second, the different emissions scenarios differ only marginally in the first few years of the 21st century (Nakićenović et al., 2000). Third, the response of the climate system to slowly evolving forcings is delayed due to the thermal inertia of the oceans (Hansen et al., 1985, Wigley and Schlesinger, 1985), thus differences between the projected and observed emissions have little impact on the climate system in the first few years. For the model simulations including natural forcings, this extension introduces a small bias, as the natural forcing is kept constant throughout the 21st century.

An overview of the number of ensemble members of the models used is given in table 5.1. For each model, these ensemble members differ only in the initial conditions and thus represent different realizations of internal variability. The individual models, however, differ in many respects including the number of external forcing mechanisms, the parameterization of sub-grid scale processes, and the number of processes, interactions, and components of the climate system explicitly modelled to name just a few. Thus, the CMIP3 multi-model ensemble is a very heterogeneous ensemble.

Two simple strategies how to compute a multi-model ensemble average from a heterogeneous ensemble can be thought of. First, each individual simulation is given equal weight, and thus the multi-model mean response is dominated by model responses of models with many simulations. Second, each model in the multi-model mean is given equal weight. Thus, the individual simulations

ALL		ANT	
GFDL CM2.0	1	BCCR BCM2.0	1
GFDL CM2.1	1	CCCMA CGCM3.1	5
GISS MODEL-E-H	3	CCCMA CGCM3.1 T63	1
GISS MODEL-E-R	5 (0)	CNRM CM3	1
INMCM3.0	1	CSIRO MK3.0	1
MIROC3.2 HIRES	1	CSIRO MK3.5	1
MIROC3.2 MEDRES	3	GISS AOM	2
MIUB ECHO-G	3	INGV ECHAM4	1
MRI CGCM2.3-2A	5	IPSL CM4	1
NCAR CCSM3.0	7	MPI ECHAM5	4
NCAR PCM1	4	UKMO HADCM3	1
		UKMO HADGEM1	1
TOTAL	34 (29)		20

Table 5.1: *Number of ensemble members of the individual models used in the detection and attribution analysis for temperature and precipitation (in brackets).*

of models with many ensemble simulations are given less weight. In the second case, the signal-to-noise ratio of the change in the multi-model mean is lower due to the down-weighting of individual simulations. Giving each model equal weight would be the method of choice if we were to believe that the models are independent and their response is centered around the true response. However, the models share common bias structures (Jun et al., 2008) and are thus not independent and there is intensive ongoing debate on how to deal with dependencies of model biases in the CMIP3 ensemble and multi-model ensembles in general (for a review on methods currently used see Tebaldi and Knutti, 2007). The results presented in this study are based on guess patterns derived from multi-model averages giving each simulation equal weight. Qualitatively, the conclusions are unaffected by this choice in most cases. However, due to the smaller estimation uncertainty of the hypothesized response, confidence intervals are narrower and the results are more robust to changes in the details of the analysis.

In addition to the guess patterns derived from the CMIP3 ensemble, we also use guess patterns of the anthropogenic and natural response derived from individual models in the analysis. These models are the HadCM3 (Gordon et al., 2000)

and PCM1.1 (Washington et al., 2000) models also included in the CMIP3 ensemble and the MIROC3.2 medium resolution model with high climate sensitivity (Hasumi and Emori, 2004, Yokohata et al., 2005, Ogura et al., 2008), which is a modification of the MIROC3.2 model with different cloud and precipitation parameterization. Of each of these models we use the ensemble mean of the four simulations of the climate of the 20th century driven with anthropogenic and the ensemble mean of the four simulations driven with natural forcings only as guess patterns.

Previous to the analysis, the model data and observations are interpolated to a common latitude-longitude grid with a resolution of 5° using conservative remapping (Jones, 1999). We use precipitation and temperature data over land only, since sea surface temperatures in the models are likely flawed due to the unrealistic representation of the Baltic Sea in the GCMs and since precipitation observations over water are not available for the period under investigation. Model grid boxes with a fractional land area of at least 50 percent are considered as land. We mask grid cells in the monthly data according to the missing value mask in the observed record. Annual averages are then computed for years with at least 10 months of data, seasonal averages are computed only for seasons with no missing data in the monthly records.

5.3 Optimal fingerprinting

When trying to answer the question whether an anthropogenic effect on climate has been observed, one often refers to detection and attribution studies (see section 1.2). A wealth of statistical methods has been used in the detection and attribution context, including correlation-based methods (Santer et al., 1993), optimal fingerprints or multivariate regression (Hasselmann, 1979, 1997, Allen and Tett, 1999), and Bayesian methods (Hasselmann, 1998, Schnur and Hasselmann, 2005). We use a variant of optimal fingerprinting to analyze the detectability of an anthropogenic signal in this study.

5.3.1 Fingerprinting as linear regression

The underlying assumption in detection and attribution studies is that the climate ψ can be linearly decomposed into deterministic responses to external forc-

ings ψ_i (the signal) and internal climate variability ν (see equation 5.1). That is, climate change detection and attribution is a signal in noise problem.

$$\psi = \sum_i \psi_i + \nu \qquad (5.1)$$

Furthermore, we assume, that the signal patterns are known, but their scaling is unknown. Thus, we account for uncertainty in the magnitude of the forcing and uncertainty in the magnitude of the response to the forcing - the climate sensitivity. Given these assumptions, we can reformulate the detection and attribution problem as a linear regression problem (see equation 5.2) with the observed change y being a linear combination of the scaled signals x_i (with scaling a_i) plus random noise from internal variability u. We use ordinary least squares (OLS) to estimate the parameters a_i.

$$y = \sum_i a_i x_i + u \qquad (5.2)$$

In this framework, we express the detection and attribution problem as follows: We detect an externally forced change, if any of the scaling factors a_i is significantly different from zero, that is, if we are not able to explain the observed change y with internal variability u alone. We attribute the change to the respective forcings, if the scaling factors are significantly different from zero but not significantly different from one.

In contrast to classical regression analysis, the significance of the scaling factors is not computed based only on the distribution of the residuals. Instead, the distribution of the scaling factors a_i under the null-hypothesis $a_i = 0$ is assessed from fits of the regression model to independent segments of long control simulations (i.e. climate model simulations with external forcings kept constant). The longer record thus used for significance testing ensures that changes due to low-frequency components of the internal variability are not erroneously identified as response to forcing.

5.3.2 Accounting for noise-contamination in the signal estimates

The signals x_i are *a priori* unknown and have to be estimated from climate model simulations. Therefore, these estimated signals are contaminated by internal

variability as well, especially if small ensembles of model simulations are used to estimate the transient response to a given forcing. Following Allen and Stott (2003) and Stott (2003), we thus extend the regression model by an additional error term u_i that accounts for internal variability in the estimated signals (equation 5.3). This is the total least squares (TLS) solution to linear regression. We further assume that both the internal variability u and the noise contamination in the signals u_i have the same covariance structure, the magnitude of the noise contamination in the signals compared to the noise in the observations, however, is reduced by the square-root of the number of ensemble members used to estimate the signal. By factoring estimation uncertainty of the climate change signals into the analysis, we avoid a systematic negative bias of the scaling factors. It is important to notice, however, that model biases and the corresponding uncertainty in the response patterns are not explicitly included in the analysis.

$$y = \sum_i a_i(x_i - u_i) + u \qquad (5.3)$$

Furthermore, we stress that for climate change detection - the statement that the observed change is significantly different from internal variability - estimation uncertainty in the guess pattern is of second importance. Although the power of detection assessment may be reduced by including additional uncertainties, TLS is unlikely to lead to erroneous positive detection of an externally forced change. If we want to quantify the similarity of the simulated and observed change, avoiding the negative bias of the OLS solution is important. Using OLS, we would on average conclude that the model simulated changes are weaker than the observed changes even though they are in fact similar. Adding estimation uncertainty of the guess patterns in the analysis removes this negative bias, and from the scaling factors derived using TLS we would also conclude that the model simulated and the observed change are - on average - similar if the underlying responses are in fact similar.

A full description of the estimation procedure for TLS regression is presented in (Allen and Stott, 2003) and we do not attempt to discuss the estimation in detail here. Nevertheless, important properties of scaling factors estimated using TLS have to be pointed out, as these differ markedly from the respective properties when using OLS. First, the TLS solution minimizes the squared distances perpendicular to the best fit line. Thus, if there is little correlation between the

observed change and the signal, no particular direction for the line relating simulated and observed changes will be preferred, and consequently, the scaling on the signal will not necessarily be close to zero as with OLS. Second, if the noise-contamination in the signal u_i is strong and thus the unknown noise-free response could approach zero, the scaling a_i could be very high or even infinite. Thus, strong noise-contamination of the estimated signals leads to wide confidence intervals that may even be open-ended, i.e. include infinite scaling (see figure 5.11).

5.3.3 Signal-to-noise optimization

The a_i are unbiased and hence optimal estimators only if the noise components in u are independent and identically distributed (*iid*). Therefore, the regression model is not fitted using the original observation and model data but using a transformed version thereof. A suitable transformation resulting in *iid* residuals is given by using (a truncated set of) the empirical orthogonal functions (EOFs) of the control simulation weighted by the square root of their eigenvalues, since the control represents our estimate of internal variability. This transformation corresponds to the signal-to-noise optimization in the optimal fingerprint method as described in (Hasselmann, 1993).

When using spatial information in the detection and attribution analysis, it is obvious that we would want to rely on the first few EOFs only, as we expect the models to reliably represent large-scale features but not the small-scale details. Allen and Tett (1999) propose a consistency test that can be used to determine the truncation in the phase space for which model simulated variability is consistent with internal variability derived from the observations. The null-hypothesis is: the control simulation of climate variability is an adequate representation of the variability in the real world in the respective truncated phase-space. As we require the transformation to result in *iid* residuals, we can then test wether the residuals behave like mutually independent random noise (for an in-depth discussion please refer to Allen and Tett, 1999). If we reject the null-hypothesis, the transformation that is based on simulated variability does not result in *iid* residuals. Testing with an increasing number of EOFs retained in the analysis can then provide an upper bound on the number of EOFs of the simulated variability that are consistent with the observed record. We note, however, that the power of this test is generally low.

When using temporal patterns in the detection analysis, however, there is no obvious reason why we should truncate the EOFs at all, as the models are not *a priori* expected to underestimate high-frequency variability (corresponding to EOFs with small eigenvalues) - at least not on time-scales relevant for the detection problem. Thus, we do not truncate the series of EOFs when using area-average quantities, however, we allow for a maximum of twenty EOFs in the case of space-time patterns and choose the final truncation according to the above mentioned consistency test.

Using a truncated set of scaled EOFs of the control for the transformation, however, can lead to erroneous attribution results if the signal does not project well on the truncated set of EOFs used. For the detection, any variable transformation resulting in iid residuals could be applied, for attribution, in contrast, the transformation should not influence the relative importance of the different signals. Therefore, we use EOFs derived from the all-forcings runs (i.e. the simulations from the CMIP3 database described in section 5.2.2) for the detection and attribution analysis, in order to assure that all of the searched-for signals project well on the first few EOFs.

The problem of large (small) eigenvalues being positively (negatively) biased when estimated from a small sample (see von Storch and Hannoschöck, 1985 for further discussion) is taken care of by using an independent segment of the control for significance testing. The estimate of internal variability used for signal-to-noise optimization is derived from the intra-ensemble spread of model simulations from models with more than two ensemble members. The estimate of internal variability used to compute confidence intervals is derived from overlapping segments of the pre-industrial control runs of all CMIP3 models. Segments starting every ten years are used. For the computation of the confidence intervals we assume the corresponding degree of freedom of non-overlapping segments, which represents a conservative estimate of the spread of scaling factors under the null hypothesis. We inflate the control run variability for models with significantly lower variance in the forced simulations compared to the observed variance (variance ratio test at the 5% level, figure 2.4).

5.4 Single-signal detection

We focus on changes in the second half of the 20th century in the detection and attribution analysis in order to minimize the effects of sparse and temporally changing observational coverage (e.g. figure 5.1). The observed trend from 1953 to 2007 in area-average temperatures in the Baltic Sea catchment is well reproduced by both the anthropogenic and the all-forcings guess pattern as derived from the CMIP3 ensemble (see figure 5.2). In addition, the anthropogenic and all-forcings guess patterns differ only marginally, thus indicating that anthropogenic forcing is the dominant forcing for temperature during the last 50 years. The observed trend in precipitation in winter, in contrast, is stronger than predicted by the models. Furthermore, the difference between the anthropogenic and all-forcings guess patterns (dashed and solid black lines in the right-hand panels of figure 5.2) in precipitation are of comparable magnitude to the hypothesized response to anthropogenic forcing and thus it is less clear whether anthropogenic or natural forcing is the dominant forcing. The variability in 5-year averages of temperature in winter and precipitation summer and winter is strong compared to the hypothesized responses to anthropogenic or anthropogenic and natural forcings. This illustrates, that the signal-to-noise ratio of externally forced changes compared to internal variability is low.

We assess whether the observed change can be explained by natural internal variability alone. Therefore, we fit the observations to the all-forcings (ALL) and anthropogenic (ANT) signals as derived from the CMIP3 ensemble individually. Additionally, the observations are also regressed on the response to natural forcings only, which is the difference between the all-forcings and anthropogenic guess patterns. The thus constructed guess pattern is more strongly contaminated by internal variability than either of the initial guess patterns. Its noise contamination corresponds to the noise-contamination of an ensemble with 12.6 members.

5.4.1 Detection with time series of area-average anomalies

We investigate the detectability of an externally forced signal using time series of seasonal anomalies of area-average temperature and precipitation from 1953 to 2007. To further suppress variability, we use time series of non-overlapping 5-yearly averages of the seasonal anomalies (see figure 5.2). Thus, the vector

of the observed change used in the detection analysis includes the anomalies of average winter (DJF) temperatures from 1953 to 1957 as its first entry, anomalies of average spring (MAM) temperatures from 1953 to 1957 as its second entry, and so forth with the last element being the anomalies of average autumn (SON) temperatures from 2003 to 2007.

We find a detectable external signal in seasonal temperature anomalies from 1953 to 2007(figure 5.3). Both the anthropogenic and the all-forcings guess patterns are detected with 10% risk of error. The best-fit scalings for the anthropogenic and all-forcings guess patterns are very similar and close to one, indicating that either of the proposed responses is a plausible explanation for the observed change. From the small difference in scaling on the anthropogenic and all-forcings response follows, that the response to natural forcings is rather weak and unimportant in explaining the observed change. This is further confirmed by using the natural guess pattern in the single-signal detection analysis (rightmost column in figure 5.3). We have to amplify the natural signal as derived from the CMIP3 ensemble considerably to best fit the observations.

When we neglect noise contamination of the guess patterns (the ordinary-least-squares solution to the regression problem, see section 5.3), the best-fit scalings on the natural guess pattern is close to zero. In contrast, the best-fit scalings on both the all-forcings and the anthropogenic guess patterns are close to one but only the anthropogenic guess pattern is detected (not shown). Thus, we conclude that natural forcing alone is no plausible explanation of the observed change, whereas anthropogenic forcing alone is a plausible explanation of the observed temperature change from 1953 to 2007.

The influence of the time period analyzed is investigated using windows of decreasing length from 1903, 1953, 1963, and 1973 to 2007. We detect an anthropogenic signal in the observed temperature changes for all time periods, the all-forcings signal, however, can only be detected during the last 55 to 45 years. The size of the confidence intervals decreases with increasing length of the time window, thus indicating that the increase in significance due to increasing sample size outperforms the increasing signal-to-noise ratio due to the accelerated warming towards the end of the 20th century. The all-forcings signal explains the observed warming since the 1950s very well, with best-fit scalings being close to one. The evolution of temperatures in the first half of the 20th century, however, is not well reproduced in the all-forcings simulations. Furthermore, we note
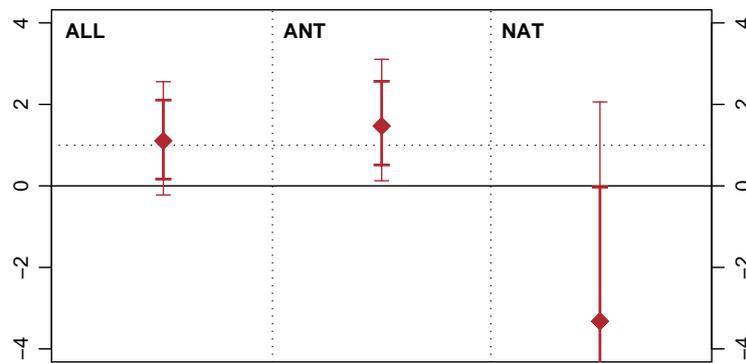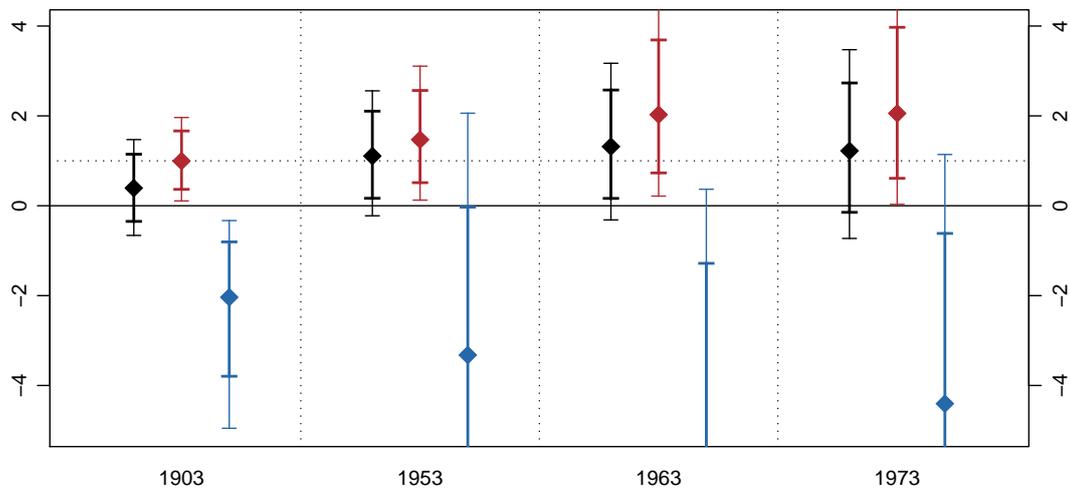
Figure 5.3: *Scaling factors and corresponding confidence intervals of a single-signal detection analysis with time series of 5-yearly averages of seasonal area-average temperature anomalies from 1953 to 2007 according to the CRUTEM3v data. The guess patterns are derived from simulations of the CMIP3 ensemble with all forcings (left), anthropogenic forcings (middle), and the difference between all and anthropogenic forcing guess patterns (right). The diamonds indicate the best-fit scaling on the guess patterns, thick vertical lines denote the 90% confidence interval about the best-fit scaling, the thin vertical lines denote the 90% confidence interval when control variability is doubled.*

that the recent warming is stronger than the predicted warming due to anthropogenic forcing, corroborating the findings of van Oldenborgh et al. (2009).

Successful detection of an external influence depends both on the estimate of the internal variability and on the signal that is searched for. However, only the estimate of variability is critical, as looking for a signal that is far away from the true response simply reduces our detection skill. In contrast, errors in the estimate of variability will lead to false detection results. If we underestimate variability, we will be overconfident about a potentially detectable signal and vice versa. Thus, we double the control variability used to estimate the confidence intervals about the best-fit scaling (e.g. thin lines in figures 5.3 and 5.4). With the control variability doubled, only the anthropogenic signal in area-average temperature anomalies can be detected with a probability of error of 10%.

The detectable anthropogenic signal is slowly varying (see figure 5.2), and thus a misrepresentation of the low-frequency variability in the control runs used can lead to false detection results. We compare the power spectra of observed and

Figure 5.4: *Scaling factors and corresponding confidence intervals of a single-signal detection analysis with time series of 5-yearly averages of seasonal area-average temperature for different time periods. From left to right the time periods used in the analysis are the 105 years from 1903 to 2007, 55 years from 1953 to 2007, 45 years from 1963 to 2007, and the 35 years from 1973 to 2007. The guess patterns used in the single-signal analysis stem from the CMIP3 all-forcings (black) and anthropogenic (red) simulations. The natural forcing only guess pattern (blue) is the difference between the all-forcings and anthropogenic guess patterns.*

simulated seasonal area-average temperature in the Baltic Sea region. We find no clear evidence of a consistent underestimation of the low-frequency variability in the model simulations (see figure 5.5). However, the spread across the models used is wide and we cannot directly assess internal variability as we cannot extract the internal variability from the observations. Instead, we compare the forced simulations with the observations.

In addition to the analysis of the low-frequency variability, we check if the variability of the residuals from the fit is consistent with internal variability in the models as proposed by Allen and Tett (1999). We test whether the residuals from the fit behaves like mutually independent, normally-distributed random noise in the coordinate system defined by the weighted EOFs of model simulated variability. It can be shown that for the TLS approach, the sum of squares of the
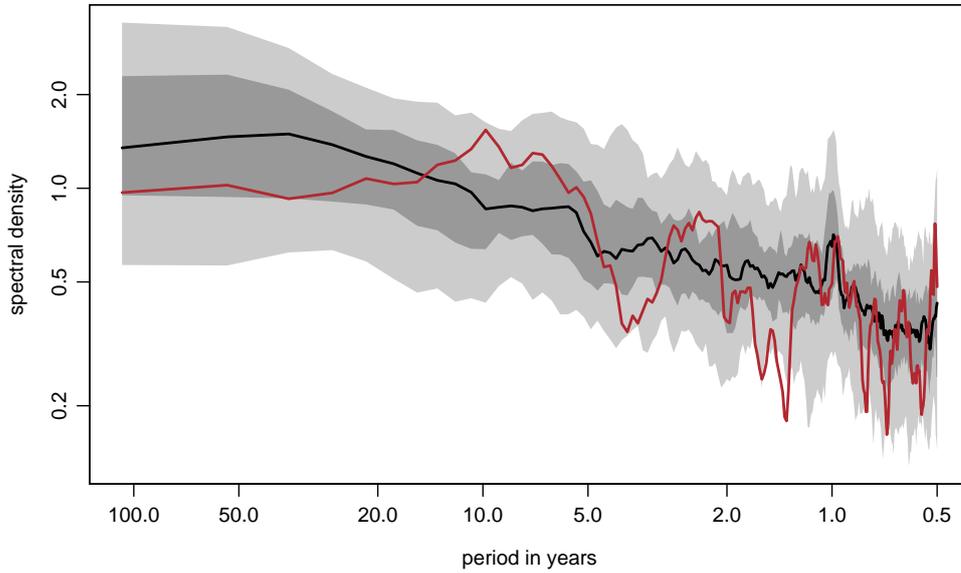
Figure 5.5: *Smoothed periodogram of seasonal area-average temperature anomalies for the period from 1901 to 2007. The red line denotes the observed spectrum according to the CRUTEM3v data, the median of all forced CMIP3 simulations is plotted in black along with the interquartile range in dark gray and the range from the 5th to the 95th percentile in light gray. The periodogram is smoothed with a modified Daniell smoother of length 9.*

residuals r^2 is expected to follow an F-distribution (Allen and Stott, 2003), with:

$$r^2 \sim (k - m)F_{(k-m),n} \tag{5.4}$$

where k is the number of EOFs retained in the analysis, m the number of signals, and n the number of statistically independent segments from the control simulation used. Instead of the F distribution, we test against the χ^2 distribution, which is the appropriate test for the residuals from an OLS regression (Allen and Tett, 1999), and which provides a conservative estimate of model consistency in the presence of noise-contamination of the guess patterns. In figure 5.6, the evolution in dependence of the truncation of $(k - m)/r^2$, which corresponds to the cumulative variance ratio of simulated vs. observed variance, is shown along with the respective 10% and 90% confidence levels. According to the consistency test, the model simulated variability is consistent with the residual variability for all truncations. The models slightly overestimate low-frequency variability (corre-

sponding to the first few EOFs). This is consistent with the slightly lower spectral density found in the comparison of the power spectra for multi-decadal variability (figure 5.5).

We further analyze wether the detection analysis is robust to changes in the number of EOFs retained in the analysis. In general, we find that in cases in which the signal-to-noise ratio allows us to successfully detect an external influence, the results are not sensitive to the truncation used (right panel in figure 5.6). On the other hand, if the signal-to-noise ratio is low, the best-fit scaling gets - as expected - heavily dependent on the details of the analysis (not shown).

Number of EOFs retained in the analysis
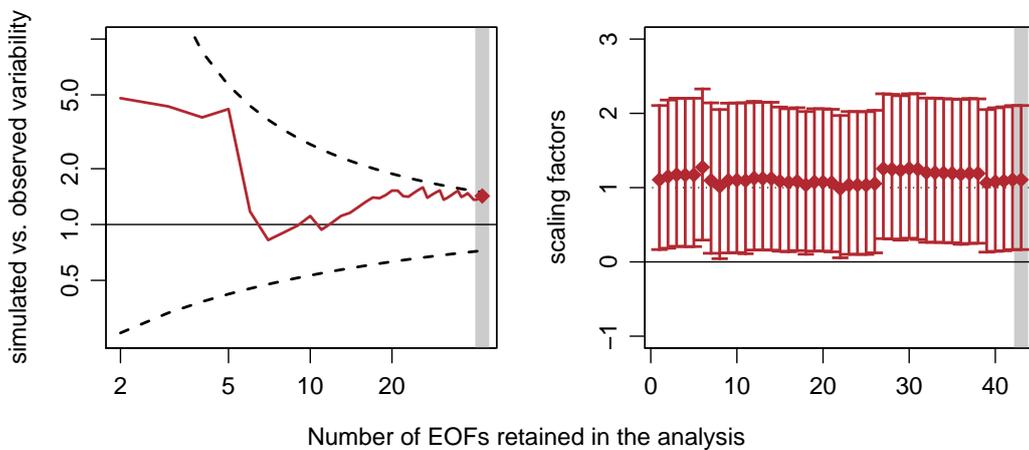
Figure 5.6: *Ratio of simulated and observed residual variance (left, see text) and scaling factors and confidence intervals (right) in dependence of the truncation for single-signal detection using time series of 5-yearly averages of seasonal area-average temperature from the CRUTEM3v data and the multi-model all-forcings guess pattern from 1953 to 2007. The dashed line denotes the 10% and 90% confidence levels within which simulated residual variance is consistent with the observed residual variance according to a χ^2 test. The grey bar indicates the truncation used.*

The results presented so far rely on the guess patterns from the CMIP3 multi-model ensemble. If we use HadCM3, PCM1.1, and MIROC3.2 single-forcing runs to estimate the signals, we have to restrict the analysis to the years before 2000, since most of the simulations have not been carried on into the 21st century. In correspondence with the length of the period chosen in the analysis of the most recent data, we analyze the 55 years of 5-yearly averages from 1943 to 1997. For

all three models, the all-forcings guess pattern is the sum of the anthropogenic and natural forcings only simulations. Furthermore, we also use the average of the three models providing single-forcing simulations for detection (apricot scaling and uncertainty in figure 5.7).
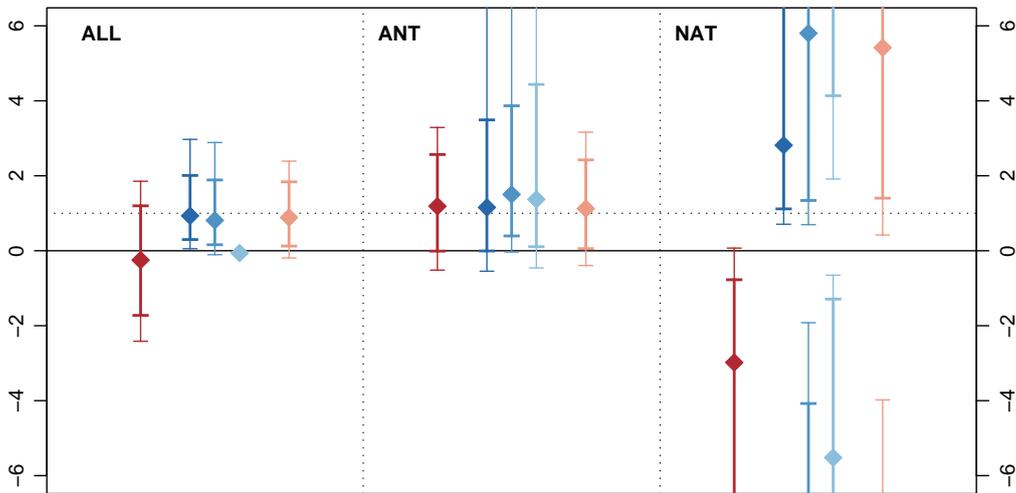

Figure 5.7: *Scaling factors and confidence intervals in dependence of the models used to derive the climate change signals for single-signal detection with seasonal temperature from 1943 to 1997 according to the CRUTEM3v data. From left to right the simulations used to derive the guess patterns stem from the CMIP3 multi-model ensemble (darkred), the HadCM3 (darkblue), the MIROC3.2 (blue), and the PCM1.1 (lightblue) single-forcing runs, and the average of the three models with single-forcing runs (apricot).*

We find a detectable external signal using most of the guess patterns from the individual models (figure 5.7). The all-forcings guess patterns from HadCM3 and MIROC3.2 explain the observed change well, with best-fit scalings on these signals being close to unit scaling, whereas the all-forcings signals derived from PCM1.1 and the multi-model mean provide poor explanations. The confidence intervals about the best-fit scalings on the all-forcings signal from HadCM3 and MIROC3.2 are even narrower than the confidence interval on the multi-model mean signal, illustrating the improvement in detection skill when the searched for signal is close to the true response. Although the 10 most recent years used in the detection analysis with the multi-model mean (see figure 5.3) help to in-

crease the signal-to-noise ratio, using a better estimate of the signal - as with HadCM3 and MIROC3.2 guess patterns - is similarly helpful.

Best-fit scalings on the anthropogenic guess patterns agree well across models and are close to unit scaling. The 90% confidence intervals about the best-fit scalings on the anthropogenic guess patterns include unit scaling for all models analyzed, whereas they do not include unit scaling for the natural guess patterns. Thus, we conclude that anthropogenic forcing alone is a plausible explanation for the observed change from 1943 to 1997, whereas natural forcing alone is no plausible explanation independent of the model(s) used to derive the signal patterns.

The scalings on the anthropogenic guess patterns do not reflect the fact that equilibrium climate sensitivity - being the globally averaged warming resulting from an increase in atmospheric GHG concentrations - differs considerably across the models analyzed. The equilibrium sensitivity in the CMIP3 ensemble ranges from 2.1 to 4.4° C for a doubling of atmospheric CO_2 concentrations (see table 8.2 in Randall et al., 2007). The PCM1.1 model marks the lower end of the simulated sensitivities with 2.1° C, the HadCM3 model lies in the middle of the CMIP3 range with equilibrium sensitivity of 3.3° C, and the high-sensitivity version of the MIROC3.2 has an equilibrium sensitivity of 6.2° which is far beyond the sensitivities simulated in the CMIP3 ensemble (Ogura et al., 2008). The reasons why models with different equilibrium climate sensitivities reproduce the observed record similarly well are manyfold and include compensating ocean heat uptake rates and compensating net aerosol effects (Knutti, 2008).

In contrast to temperature, detection results for seasonal precipitation in the Baltic Sea catchment are less consistent. Even though we detect external influences in the observed changes from 1943 to 1997 for all three different guess patterns, we have to amplify the signals considerably to best fit the observations (see figure 5.8). As for temperature, the best-fit scaling on the anthropogenic and all-forcings guess pattern is similar, thus indicating that the natural response as derived from the CMIP3 ensemble is unimportant in explaining the observed change in precipitation.

In contrast to temperature, the observed change in area-average precipitation is considerably stronger than simulated in GCMs. If we use guess patterns from individual models, we have to amplify the anthropogenic and all-forcings guess patterns by factors of three to ten (figure 5.9), the natural guess patterns have to be amplified even more. This misrepresentation of observed precipitation
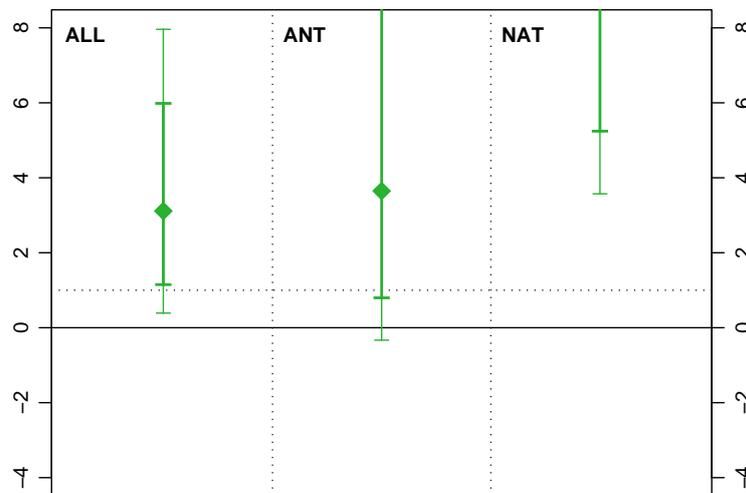
Figure 5.8: *As in figure 5.3 but for time series of 5-yearly averages of seasonal precipitation anomalies according to the GPCC v4 data.*

changes in present-day climate models is a well-known fact (Zhang et al., 2006, and chapter 3) and the limited skill in simulating changes in sea-level pressure and sea-surface temperature over Europe and the North Atlantic have been identified as possible causes (G. J. van Oldenborgh, pers. communication). We conclude that according to changes in area-average precipitation from 1943 to 1997, neither the combined anthropogenic and natural forcing, nor anthropogenic or natural forcing alone are plausible explanations for the observed change (see figure 5.9).

In contrast to temperature, the signal-to-noise ratio of externally forced changes in precipitation compared to internal variability is lowest for the 45 years from 1963 to 2007 and not for the longest time period investigated (not shown). This indicates that the anthropogenic and natural signals in the first half of the 20th century is not well-known.

5.4.2 The influence of space-time resolution on the detection result

We analyze the influence of spatiotemporal resolution on the result of the detection analysis presented in the previous section. Therefore, we repeat the single-
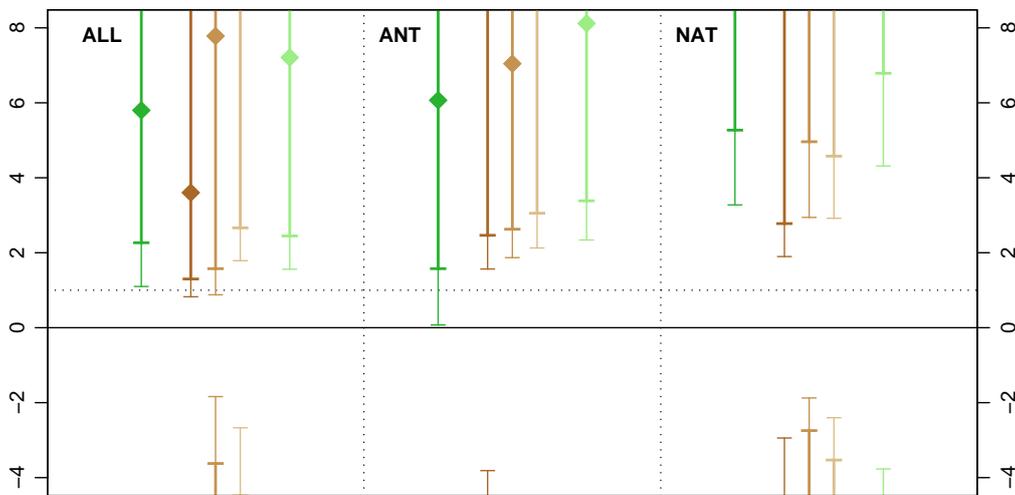
Figure 5.9: *As in figure 5.7 but for relative seasonal precipitation anomalies according to the GPCC v4 data.*

signal analysis with different levels of aggregation of the data. The detection vector in the four cases analyzed contains seasonal area-average trends from 1953 to 2007 in the most aggregated case (a vector of length 4), time series of 5-yearly averages of seasonal area-average anomalies (as in the previous section), seasonal trends for all of the fifteen 5°x5° grid boxes contributing to the Baltic Sea catchment (a vector of length 60), and finally 5-yearly averages of seasonal anomalies for all grid boxes (a vector of length 660).

We detect the anthropogenic and all-forcings guess patterns using either area-average anomalies, area-average trends, or spatiotemporal anomalies; no externally forced change can be detected using the spatial pattern of trends. Highest signal-to-noise ratio is achieved with time series of 5-yearly averages of seasonal area-average temperatures in the Baltic Sea catchment (dark red scalings in figure 5.10). For different space-time resolutions of the natural guess patterns we get different signs on the best-fit scaling. Such an erratic behavior of the best-fit scalings is expected for weak signals and low signal-to-noise ratios when using the TLS solution to regression (see section 5.3). We further find that the best-fit scalings and uncertainty ranges for the natural signal are strongly dependent on the truncation used, whereas they are stable for the anthropogenic and all-forcings signals for truncations larger than two. These first EOFs describe the

low-frequency (in the case of time series) and/or large-scale (in the case of patterns with spatial information) change in the Baltic Sea catchment.
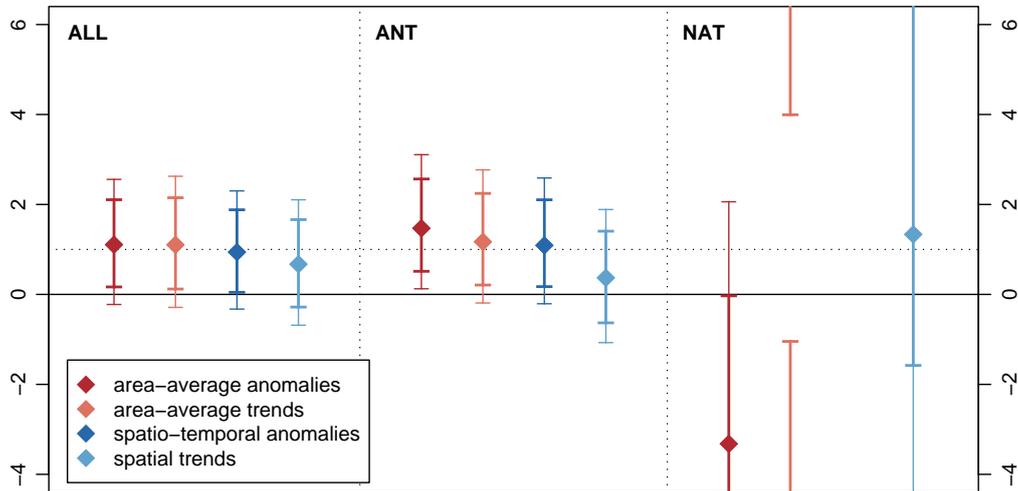

Figure 5.10: *Single-signal detection analysis with seasonal near-surface temperature from 1953 to 2007 in dependence of the space-time resolution of the patterns used (corresponding to figure 5.3). For each column from left to right these are: time series of 5-yearly averages of seasonal area-average temperature (darkred), seasonal trends in area-average temperatures (apricot), time series of 5-yearly averages of seasonal temperature at the individual grid boxes (darkblue), and seasonal trends at the individual grid boxes (lightblue).*

In contrast to temperature, including the spatial pattern of precipitation changes into the analysis increases the signal-to-noise ratio. When including the spatial pattern of the change in the analysis, the scaling on the anthropogenic and all-forcings guess patterns are consistent with unit scaling and the best-fit scalings are much closer to one, whereas natural forcing alone is still no plausible explanation of the observed change (figure 5.11)

It is important to notice, however, that the varying spatial coverage with rain-gauge stations is expected to introduce a significant bias when spatial variability is included in the analysis. The effect of changing coverage on area-average precipitation, on the other hand, is expected to be small.
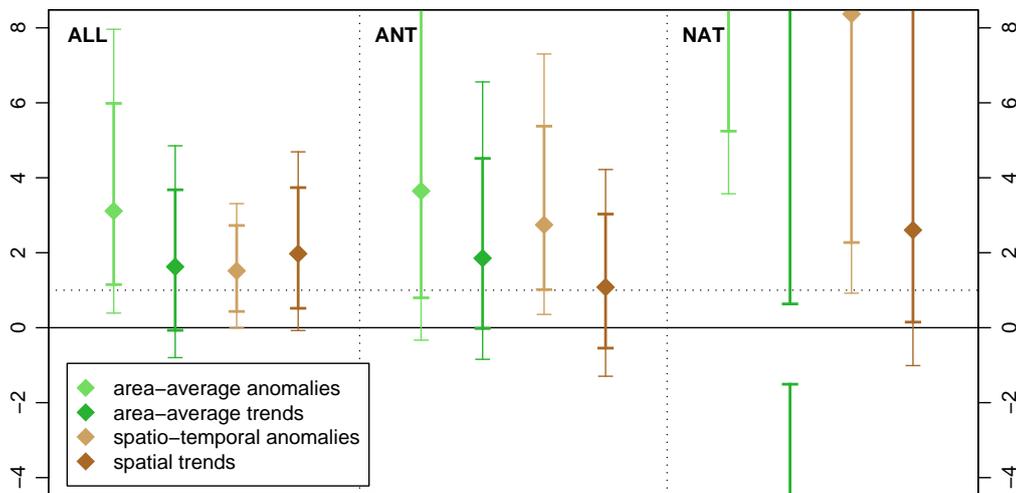
Figure 5.11: *As in figure 5.10 but for relative seasonal precipitation anomalies with respect to the 1961-90 mean derived from GPCC v4.*

5.5 Two-signal attribution

In order to distinguish between different forcing mechanisms, we jointly fit the all-forcings and anthropogenic guess patterns (anthropogenic and natural guess patterns for HadCM3, PCM1.1, and MIROC3.2) to the observations. In contrast to the single-signal analysis with the all-forcings guess pattern, this allows us to take into account scaling and estimation uncertainty in both patterns individually. For the CMIP3-derived guess patterns, the scaling on the anthropogenic and natural signals can then be computed from the two-signal analysis with the all-forcings and anthropogenic guess patterns after the fit. The scaling on the natural signal corresponds to the scaling on the fitted all-forcings guess pattern, the scaling on the anthropogenic signal corresponds to the sum of the scalings on the fitted anthropogenic and all-forcings guess patterns, as both contain the response to anthropogenic forcing. In comparison to fitting the anthropogenic and natural guess patterns (being the difference between all-forcings and anthropogenic guess patterns) directly, fitting the original guess patterns and computing the corresponding scalings after the fit is advantageous, as the original guess patterns are less noise-contaminated than the (indirectly derived) natural guess pattern.

We find a detectable external signal when fitting both guess patterns simultaneously; zero scaling on both signals is not within the 90% confidence ellipsoid around the best-fit scaling (see figure 5.12). Furthermore, the best-fit scaling on the anthropogenic signal is close to one and negative on the natural signal as in the single-signal analysis. The confidence ellipsoid extends further in the direction of the response to natural forcing, thus the signal-to-noise ratio on the natural signal is lower than for the anthropogenic signal. Even though the observed warming is technically not attributable to anthropogenic forcing, as zero scaling on the anthropogenic signal is within the confidence ellipsoid about the best-fit scaling, anthropogenic warming is clearly the more likely explanation for the observed change.
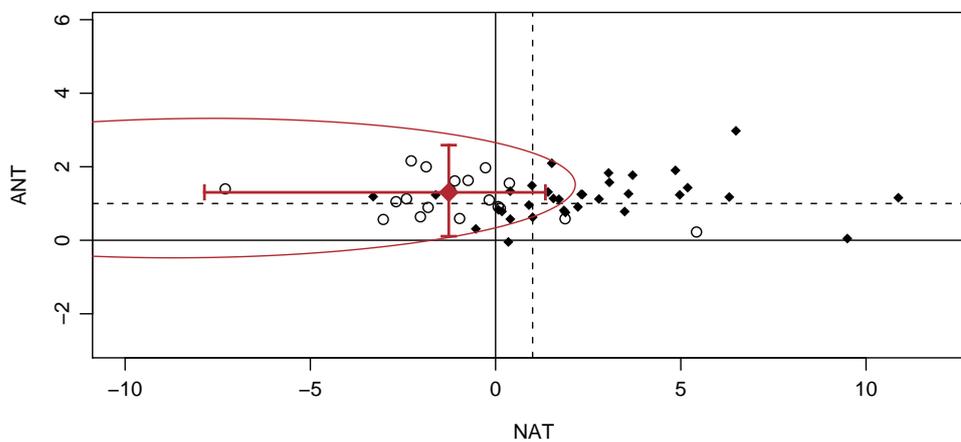
Figure 5.12: *Two-signal attribution analysis with time series of 5-yearly averages of area-average seasonal temperature in the Baltic Sea catchment from 1953 to 2007 using the CRUTEM3v data. The red diamond indicates the best-fit scaling on the anthropogenic and natural signals, the red line denotes the joint 90% confidence interval and the vertical and horizontal lines indicate the one-dimensional 90% confidence intervals about the best-fit scaling. Additionally, the best-fit scalings for the two-signal analysis using the individual CMIP3 simulations as pseudo-observations are shown.*

The low signal-to-noise ratio of naturally forced changes is either due to the natural response being weak compared to internal variability or due to the fact that the response derived from the multi-model ensemble is far from the true response. In order to shed more light on this issue, we rerun the two-signal analysis

with the individual CMIP3 simulations as pseudo-observations. We find that the best-fit scaling on the natural signal is positive for most of the simulations including anthropogenic and natural forcing, but negative for most of the simulations with anthropogenic forcing only (see figure 5.12). However, the best-fit scalings of the natural guess patterns for the all-forcings simulations (black diamonds in figure 5.12) vary considerably across models. The natural signal can be detected in 14 of the 25 all-forcings simulations used as pseudo-observations for which an external influence can be detected, which is remarkable given the wide spread of the best-fit scalings. We conclude that even though the response to natural forcing of most model simulations is closer to the multi-model response than the response in the observations, the models do not agree well on the natural response.

The guess patterns used in the two-signal analysis stem from a heterogeneous multi-model ensemble. We implicitly assume that the difference between the anthropogenic and the all-forcings guess pattern from the CMIP3 ensemble is the natural signal. Parts of this difference, however, can be attributed to systematic model biases, as different models are used to estimate the anthropogenic and all-forcings guess patterns. Thus we compare the results based on the multi-model mean patterns with attribution results based on guess patterns derived from individual models (see figure 5.13).

Corresponding to the single-signal analysis, the best-fit scaling on the natural response differs considerably across models in the two-signal analysis as well (figure 5.13). In contrast to the single-signal analysis, however, also the best-fit scalings on the anthropogenic guess patterns differ considerably across models in the two-signal analysis. Nevertheless, we find a detectable external influence in the period from 1943 to 1997 for all of the models except the PCM1.1 (light-blue line in figure 5.13) and the combined single-forcing ensemble (apricot line in figure 5.13).

The attribution results are not robust to changes in the time period considered. In general, the most recent change is dominated by anthropogenic forcing, whereas the attribution results depend strongly on the models when changes in the early 20th century are also taken into account (not shown).

The two-signal analysis with either area-average or spatiotemporal precipitation anomalies reveals no clear picture, except that the observed change is very likely not due to internal variability alone. The best-fit scalings vary considerably depending on the model used to derive the signal patterns, the time period ana-
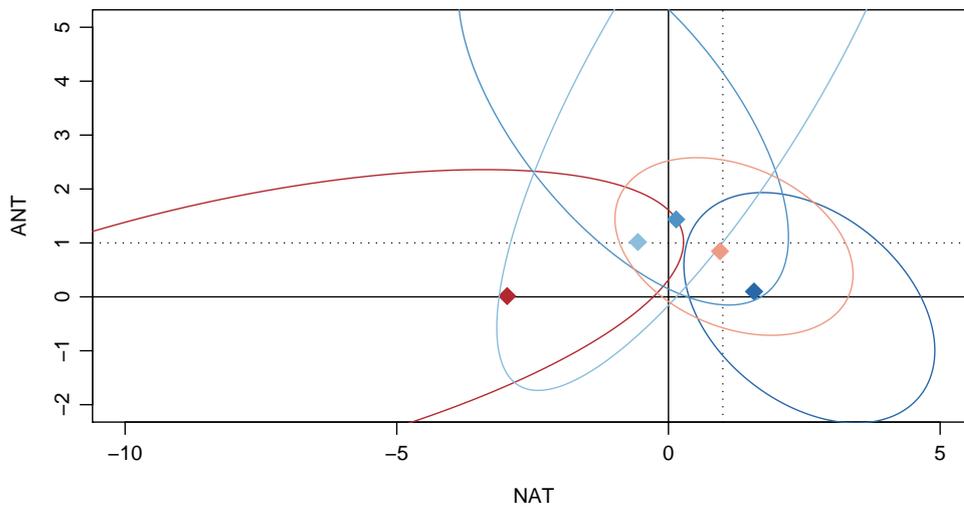
Figure 5.13: *Two-signal attribution analysis with 5-yearly averages of area-average seasonal temperature in the Baltic Sea catchment from 1943 to 1997 according to the CRUTEM3v data. The signals used in the analysis are estimated from simulations of the CMIP3 multi-model ensemble (darkred), the HadCM3 (darkblue), the MIROC3.2 (blue), and the PCM1.1 (lightblue) single-forcing runs, and the average of the three models with single-forcing runs (apricot).*

lyzed, and the spatiotemporal resolution of the pattern (not shown). Thus, we conclude, that even though it is unlikely that the observed change in precipitation is due to internal variability alone, we are not able to reproduce nor explain the observed change.

The anthropogenic and natural guess patterns are correlated for the different models analyzed and thus the major axes of the confidence ellipsoids are tilted; this correlation explains part of the variability in best-fit scalings across models. Correlation of the signals further results in wide confidence intervals about the best-fit scaling, as a strong response to natural forcing can be compensated with a strong response to anthropogenic forcing. Consequently, we cannot yet distinguish individual forcings in the Baltic Sea catchment.

Various strategies can be chosen to increase the skill in attributing the observed change to potential forcing mechanisms. These strategies include decreasing the noise contamination in the model signals by increasing the ensemble used to

estimate the signals, or increasing the signal-to-noise ratio of externally forced changes for example by looking at longer records. The second approach leads to smaller confidence intervals when temperature changes from 1903 to 1997 are analyzed (not shown). However, the best-fit scalings on the signals vary even more strongly across models when the first half of the 20th century is included in the analysis.

5.6 Detection and attribution in a surrogate climate

We identify four main sources of uncertainty that are not explicitly included in the detection analysis presented above. These are:

- Systematic model biases of the GCMs used to derive the hypothesized responses to different forcings and to estimate the internal variability (see chapter 2).

- Additional forcing mechanisms not accounted for in the GCM simulations such as land-use change.

- Scale mismatch between observed quantities measured at discrete locations and simulated quantities representing spatial averages.

- Sources of variability in the observations not included in the models such as inhomogeneities due to station relocations or due to changing spatial coverage.

We repeat the detection and attribution analysis using model simulations as pseudo-observations in order to analyze the potential detectability and attribution in cases unaffected by above uncertainties. The remaining sources of uncertainties are explicitly dealt with in the detection analysis. These are the uncertainties related to the estimation of both the hypothesized signals and internal variability from finite samples. The first of which is dealt with by including noise-contamination of the guess patterns into the analysis, which results in the total-least-squares formulation of the detection problem (see section 5.3). Estimation uncertainty of the internal variability is taken into account by using independent segments of the control simulations for the signal-to-noise optimization and the hypothesis testing. Successful detection and attribution of a regional climate change in the model world thus depends exclusively on the signal-to-noise ratio of the externally forced changes compared to internal variability and the experimental setup.

5.6.1 Detection and attribution with a perfect model

We repeat the detection and attribution analysis with time series of 5-yearly averages of seasonal, area-average temperature and precipitation from 1943 to 1997 using model simulations as pseudo-observations. For each of the three models providing simulations driven by natural and anthropogenic forcing only, we carry out the detection and attribution analysis with the respective all-forcings simulations as pseudo-observations. The all-forcings guess pattern in the single-signal analysis is then the sum of the natural and anthropogenic guess patterns in order to use simulations independent from the pseudo-observations for the construction of the guess patterns.

It is important to notice that even though the model used to derive the hypothesized signals of natural and anthropogenic change is perfect in that its systematic model bias is zero and all of the forcings and processes influencing the pseudo-observations are known, the response to natural and anthropogenic forcing is still uncertain, as we have to estimate it from a small set of simulations. The uncertainty of the guess patterns contributes to the uncertainty on the best-fit scalings (due to the unknown contribution of internal variability) and thus broadens the confidence intervals which in turn limits our skill in detecting and attributing externally forced changes.

Given a perfect model, we detect an external influence on time series of seasonal area-average temperature from 1943 to 1997 in the Baltic Sea catchment in 15 of the 36 cases analyzed (figure 5.14). The best-fit scalings vary considerably between the different simulations used, indicating, that the signal-to-noise ratio of the externally forced changes compared to internal variability is low. Using the all-forcings guess pattern to detect changes in the all-forcings simulations is beneficial according to the HadCM3 and MIROC3.2 results. This is somewhat surprising, as the all-forcings guess patterns used in the analysis are more strongly contaminated by internal variability than the anthropogenic and natural guess patterns due to the way they are constructed. Nevertheless, this nicely illustrates how looking for a pattern close to the true response increases our skill of detecting an externally forced change.

As in the single-signal analysis, the best fit scaling factors vary considerably across simulations in the two-signal analysis. This variability can again be explained in parts with the correlation of the two patterns. The natural and anthropogenic guess patterns as derived from the HadCM3 and MIROC3.2 (PCM1.1) models are
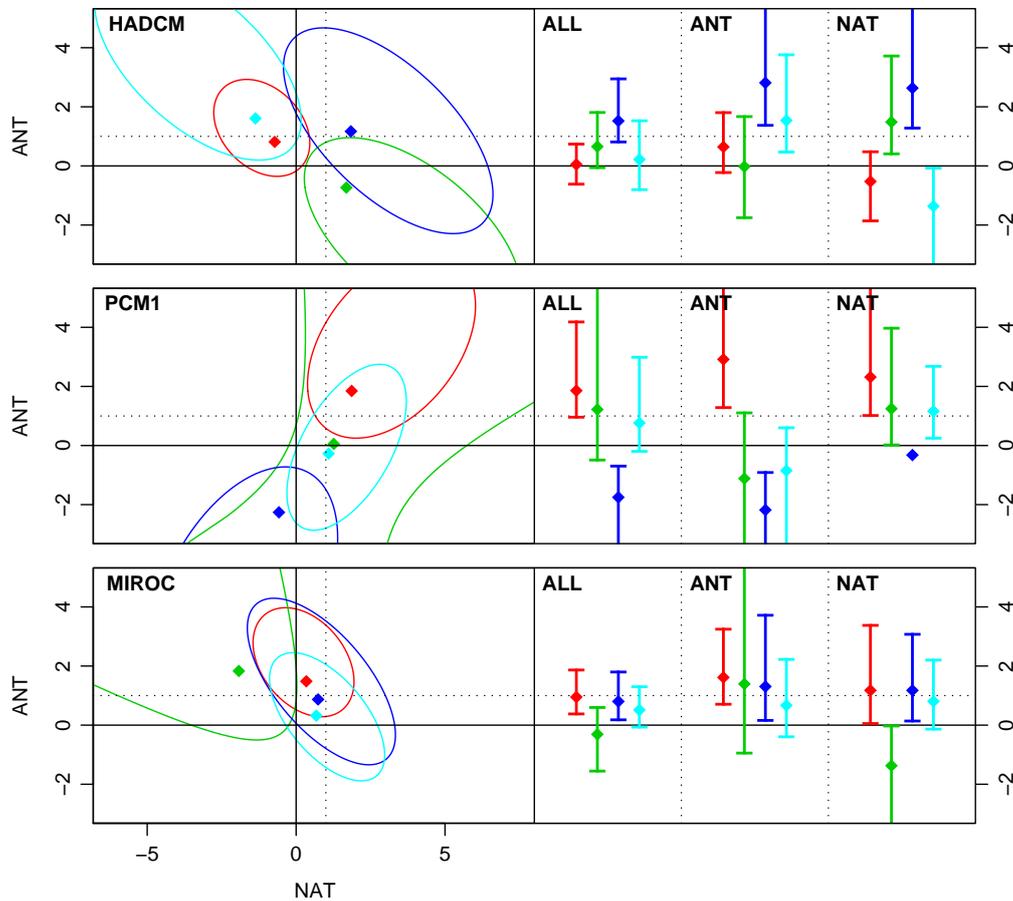
Figure 5.14: *Two-signal (left) and single-signal (right) analysis using time series of 5-yearly averages of seasonal area-average temperature from 1943 to 1997 from the HadCM3, PCM1.1 and MIROC3.2 all-forcings simulations as pseudo-observations. The searched-for signals have been derived from the anthropogenic and natural forcing simulations of the respective models. The all-forcings guess pattern in the single-signal analysis is the sum of the natural and anthropogenic guess pattern. Diamonds denote best-fit scalings, the ellipsoids and the vertical lines indicate the 90% confidence intervals about the best fit scaling. The colours denote the four different simulations of each model used as pseudo-observations. These simulations differ in initial conditions only.*

correlated (anti-correlated). Consequently, the major axis of the respective confidence ellipsoids is tilted by approximately 45° with respect to the horizontal line (see figures 5.13 and 5.14). That is, a positive scaling on the natural response can be compensated by negative scaling on the anthropogenic response and vice versa. Thus, the influence of individual forcing mechanisms can not be separated.

The two-signal analysis further reveals that neither of the forcings clearly dominates in the period from 1943 to 1997. Nevertheless, an external influence is detected in 7 out of 12 cases and the scaling on both the anthropogenic and natural guess patterns is consistent with unit scaling in 3 of the 7 cases for which an external influence is successfully detected in the two-signal analysis. Furthermore, we are able to separately detect both guess patterns individually in the two-signal analysis for one simulation with the PCM1.1 model (red diamond and line in the middle panel of figure 5.14) and thus we can formally attribute the simulated change to combined anthropogenic and natural forcing.

Perfect-model detection and attribution with time series of 5-yearly averages of area-average precipitation anomalies is carried out with the PCM1.1 and MIROC3.2 model data only, as HadCM3 all-forcings simulations were not readily available at the time. For both models, scaling factors on all of the three searched-for signals in the single-signal analysis vary considerably across the four different all-forcings simulations used as pseudo-observations (not shown). Consequently, we conclude that the signal-to-noise ratio of the externally forced response to either forcing compared to internal variability is low. Nevertheless, an external signal is detected in the single-signal analysis in 5 out of 12 cases using PCM1.1 and in 4 out of 12 cases using MIROC3.2 data. In 4 out of 8 cases, the anthropogenic guess pattern can be separately detected and is consistent with unit scaling in the two-signal analysis. In contrast, the natural guess pattern cannot be separately detected in the two-signal analysis.

We conclude that we would not expect to be able to detect an external influence in area-average temperature and precipitation from 1943 to 1997 even if we had a perfect model. Even with the all-forcings guess pattern, best-fit scalings are not always consistent with unit scaling. The low signal-to-noise ratio of externally forced changes in precipitation and the thereby resulting strong noise-contamination of the guess patterns derived from small ensembles inhibits successful attribution in the period from 1943 to 1997 even if we had a perfect model.

5.6.2 Time of emergence

Above perfect-model study illustrates, that external changes in seasonal temperatures during the 20th century cannot be reliably detected with climate signals derived from small ensembles even if we had a perfect model. The question thus arises, whether either the increasing anthropogenic signal in the near future or the use of imperfect but less noise-contaminated guess patterns significantly increase our skill in detecting externally forced changes.

We perform a single-signal analysis with guess patterns derived from all the CMIP3 models providing at least 3 ensemble simulations for the 20th and 21st century. Thus we have a total of 11 (10 for precipitation) signal patterns at hand, 9 (8) from individual models (see table 5.1) plus the all-forcings and anthropogenic forcing multi-model patterns as described in section 5.2.2. Depending on the model, the guess patterns include either anthropogenic (GHG and sulfate) forcing only or anthropogenic and natural (solar and volcanic) forcing. Furthermore, we use the 54 (49 for precipitation) individual simulations as pseudo-observations. For all combinations of guess patterns and pseudo-observations, we analyze the detectability of the signal in the simulations for different time periods, the shortest including 20 years from 1953 to 1972, the longest ranging from 1953 to 2097 (upper panel in figures 5.15 and 5.16). Due to the increasing length of the time window used for the detection analysis alone, we expect an increase of the signal-to-noise ratio. Additionally, a decreasing number of independent segments from the control runs is available for time windows of increasing length. This artificially increases the estimated signal-to-noise ratio of external changes for the longest periods analyzed. In order to partly avoid this effect, we reduce the number of control run segments used for the analysis of short time windows. In addition, we also analyze detection using moving windows of fixed length of 55 years, the first ranging from 1903 to 1952, the second from 1908 to 1957, and so forth with the last ranging from 2043 to 2097 (bottom row in figures 5.15 and 5.16).

Above experimental setup allows us to investigate the effect of the detection strategy and the signal patterns on potential detectability. We define potential detectability as the fraction of simulations for which an anthropogenic signal can be detected at the 90% level. Correspondingly, potential attribution is the fraction of simulations consistent with unit scaling of the respective signal. We further distinguish three cases:
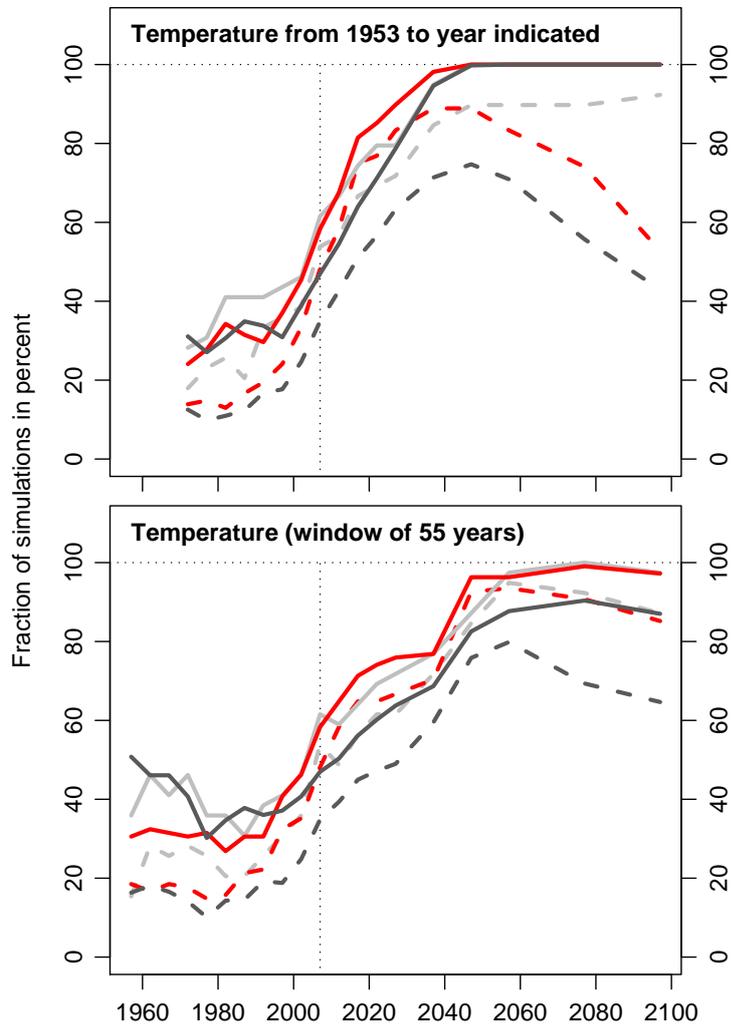
Figure 5.15: *Emergence of a detectable external signal in 5-yearly averages of area-average seasonal temperature from 1953 to the year indicated (top panel) and for windows with fixed length of 55 years (bottom panel). All the simulations from the CMIP3 ensemble are used as pseudo-observations. Signal patterns are derived from models with at least three ensemble members (see text). In the perfect-model analysis, the signal patterns and pseudo-observations stem from the same model (light grey), whereas in the imperfect-model case, signals and observations stem from different models (dark grey). The detection results with the multi-model mean signals are displayed in red. The solid lines denote the fraction of simulations showing detection at the 90% level, the dashed lines indicate the fraction of simulations additionally consistent with unit scaling (attribution). The vertical line indicates the year 2007.*

- The perfect-model case: both the simulations used as pseudo-obervations and the simulations used to derive the guess patterns stem from the same model. The guess pattern is constructed from all the available simulations except the one used as pseudo-observations. Thus, the guess patterns are estimated from 2 to 6 simulations depending on the model.

- The imperfect-model case: the simulation used as pseudo-observations stems from a different model than the simulations used to estimate the guess pattern. Depending on the model, the guess pattern is estimated from 3 to 7 simulations.

- The multi-model mean case: the guess pattern is the multi-model mean as described in section 5.2.2.

The guess patterns in the perfect-model case are thus more strongly contaminated by internal variability than the guess patterns in the imperfect-model case. On the other hand, the perfect-model patterns are generally closer to the true response than the imperfect-model patterns.

Anthropogenic changes in both temperature and precipitation clearly emerge from noise during the 21st century (see figures 5.15 and 5.16). Potential detectability increases in all cases with increasing length of the time period analyzed (upper panels in figures 5.15 and 5.16). Potential detectability also increases with the increasing forcing towards the end of the 21st century (lower panels in figures 5.15 and 5.16). The signal-to-noise ratio of externally forced changes is lower for precipitation than for temperature and thus potential detectability is generally lower for precipitation. The results on potential detectability of precipitation changes support the findings of Giorgi and Bi (2009), who conclude that in northern Europe, GHG-induced precipitation changes with respect to the period from 1980 to 1999 are expected to significantly emerge around 2040.

Potential detectability is higher for guess patterns from a perfect model than for guess patterns from an imperfect model (light gray versus dark gray lines in figures 5.15 and 5.16). Perfect model guess patterns and multi-model mean guess patterns, however, perform equally well except for windows of fixed length with precipitation for which perfect model signals increase potential detectability (lower panel in figure 5.16). Hence, in general, the effect of decreasing noise contamination in the guess pattern with increasing sample size compensates for the fact that the multi-model mean guess pattern is generally further away from the true response than the perfect model guess pattern. This is in contradiction
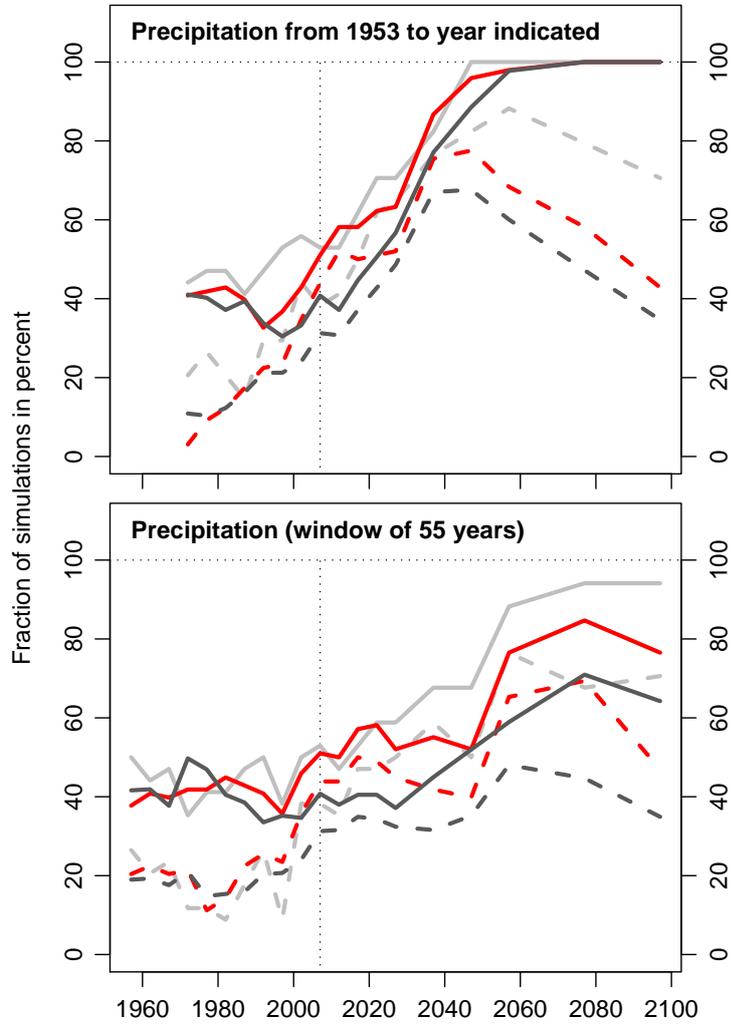
Figure 5.16: *As in figure 5.15 but for 5-yearly averages of area-average seasonal precipitation from 1953 to the year indicated (top panel) and for windows with fixed length of 55 years (bottom panel).*

to what we find when analyzing detection of observed temperature changes for which two out of three individual models perform better than the multi-model mean (all-forcings simulations in figure 5.7).

Potential attribution is much more dependent on the signal used than potential detection. When using time windows of increasing length (upper panels in figures 5.15 and 5.16), potential attribution with imperfect model and multi-model mean guess patterns decreases towards the end of the 21st century. This implies that different models do not agree very well on their regional response to anthropogenic forcing - which is consistent with the uncertainties in the quantification of climate sensitivity at the global scale (see chapter 1).

Potential attribution of temperature changes with perfect-model guess patterns approaches the theoretical upper bound of 90% for the longest time periods (upper panel in fig 5.15). The decrease of potential attribution of precipitation changes using perfect-model guess patterns for time periods larger than 110 years indicates, that we underestimate the internal variability for these time periods significantly. This might be either due to the decreasing number of independent segments of the control run available for longer time periods, due to the pooling of control runs from different models, or due to the preindustrial variability not being representative for internal variability in a climate with strong GHG forcing.

Potential attribution results indicate that we expect precipitation changes to be consistent with anthropogenic forcing for the period from 1953 to 2007. However, this is not the case for observed precipitation. Thus, we conclude that potential attribution further supports above findings that either the models lack important forcings and processes or the models underestimate anthropogenic precipitation change considerably at the regional scale.

5.7 Conclusions

We are able to detect an external influence on the observed temperature changes in the Baltic Sea catchment. The detection is robust across a variety of space-time resolutions and valid for different time periods analyzed. The spatial pattern of the change, however, contains little additional information for the detection of external influences, which corroborates previous findings (Bhend and von Storch, 2009). Furthermore, the detection result is largely independent of

the choice of model from which we derive the expected climate change signals. Using the multi-model mean pattern as done in this study has the advantage of little noise-contamination in the searched-for signal and the longest possible record can be analyzed. The comparison with detection results using guess patterns from individual models reveals that detection depends on the guess pattern used and that the multi-model mean pattern is probably not an optimal choice.

Both the anthropogenic and all-forcings signal fit the observed warming well, indicating that the response to natural forcing is rather weak and unimportant in explaining the observed change. Further insight into the relative importance of the different forcing mechanisms is provided by the two-signal attribution analysis. For the multi-model mean we find further evidence that anthropogenic forcing is the dominant forcing for the period from 1953 to 2007, whereas for the period from 1943 to 1997 the relative importance of the forcings varies considerably across models. We conclude, that the global models do not agree on the response to natural forcing in the second half of the 20th century but they agree on the response to anthropogenic forcing.

We note that regional-scale attribution is difficult as the different signals are often highly correlated and thus not distinguishable using regression-based detection methods. A potential approach to deal with the problem of correlated signals is to include additional variables in the analysis, such as jointly attributing precipitation and temperature changes (in analogy to the approach to regional projections by Tebaldi and Knutti, 2007).

We also detect an external influence on the observed precipitation changes in the study domain. Two major caveats, however, reduce the confidence in this finding. First, variability in simulated precipitation is generally lower than the observed variability. Second, the models are not able to reproduce the magnitude of the observed precipitation change, indicating that present-day GCMs either lack important additional forcing mechanisms, or that they significantly underestimate the regional-scale precipitation response to anthropogenic and/or natural forcing.

Acknowledgements

6 Conclusions and Outlook

This work contributes to the understanding of regional climate change in northern Europe. We have analyzed recent changes in both near-surface temperature and precipitation in the Baltic Sea catchment using different approaches. The spatial pattern of the change is investigated using high-resolution gridded observations and regional climate model data. As the necessary simulations for a formal detection and attribution analysis are not available, we propose a different approach. We compare the observed change with climate change projections and assess whether anthropogenic forcing is a plausible explanation of the observed change. In a second step, we investigate the detectability of an anthropogenic signal in area-average temperature and precipitation using data from global climate models.

The observed spatial pattern of change in winter precipitation from 1973 to 2002 is consistent with the pattern derived from regional climate change projections. The observed area-mean increase in precipitation, however, is considerably stronger than the simulated response to anthropogenic forcing. The increase in precipitation in northern Europe goes together with a change in the North Atlantic Oscillation (NAO) towards more westerly wind situations over the North Atlantic (Hurrell, 1995, Wanner et al., 2001, Thompson and Wallace, 2001). The observed change in the NAO, however, is not reproduced in present-day climate models (Osborn, 2004, Stephenson et al., 2006). Therefore, we remove the influence of the NAO on the observed and simulated changes in precipitation previous to the analysis. The observed change in the NAO does not fully explain the mismatch between the observed increase in precipitation and the projected changes. Thus, the following three causes for the mismatch in simulated and observed rates of change in precipitation may be thought of: changing atmospheric greenhouse gas and sulfate concentrations are not the dominant forcing and other forcing mechanisms such as changes in stratospheric aerosol concentrations due to volcanic eruptions have a major impact on precipitation in this period, the modeling system is significantly flawed in its ability to predict the precipitation response to anthropogenic forcing, or the observed trend is strongly influenced by internal variability.

In a formal detection analysis using data from global climate models, we find that the observed change in area-average precipitation from 1953 to 2007 is very likely not due to internal variability alone. However, neither the response to anthropogenic forcing (GHG and sulfate) nor the response to natural forcing (solar irradiance and volcanic aerosol changes), as derived from GCMs alone, provides a plausible explanation for the observed change. Furthermore, we note that most GCMs significantly underestimate precipitation variability. Although we inflate model derived variability accordingly, we recommend to interpret detection of an externally forced change in precipitation with caution.

Area-average temperature trends from 1973 to 2002 are consistent with estimates of the anthropogenic signal derived from regional climate change projections. Increasing the spatiotemporal resolution, however, decreases the similarity of the simulated and observed patterns of change. In addition, the anthropogenic signal derived from regional climate model simulations is fairly uniform, indicating that the projected warming is not strongly modified locally.

We assess the relative importance of different possible explanations for the observed change in area-average temperature in a detection and attribution analysis. The observed warming is very likely not due to internal variability alone. The hypothesized response to anthropogenic forcing as derived from different models is consistent with the observed warming, whereas the response to natural forcing alone is not consistent. Although this indicates that anthropogenic forcing is the dominant forcing in the second half of the 20th century, formal attribution – the quantitative separation of the influence of competing forcing mechanisms – is not possible. The relative importance of anthropogenic and natural forcing cannot be further quantified in a two-signal analysis (as done by Stott et al., 2006) due to three main reasons:

First, the low signal-to-noise ratio of externally forced changes compared to internal variability in the Baltic Sea catchment inhibits a formal attribution of anthropogenic or natural forcing in a two-signal analysis. Due to the increasing anthropogenic forcing, however, attribution of the observed warming to human influence will become possible in the near future. This is further supported by detection analyses using simulated data as pseudo-observations.

Second, the regional-scale response to natural forcing is not well known and varies considerably across models. The natural response as derived from multi-model ensembles thus tends to be weak and the natural response as derived from individual models is often far away from the true response – both cases

lead to large uncertainty ranges about the best-fit scaling of the natural guess pattern in regression based detection analyses.

Third, the response to different forcings can often not be clearly distinguished at the regional scale, as the guess patterns are correlated. Correlation of natural and anthropogenic guess patterns leads to very wide confidence ellipsoids in a two-signal analysis. A potential improvement could be achieved through the joint analysis of different variables such as precipitation and temperature. Due to the limited skill in simulating precipitation changes, however, only marginal improvements may be expected.

We explicitly include the influence of internal variability and uncertainty in the magnitude of the response to different forcings in the detection and attribution analysis. The uncertainty due to systematic model biases and the uncertain temporal evolution of the forcings, however, are not taken into account. A generalization of the regression-based detection and attribution analysis that uses inter-model variance of the guess patterns as an extra uncertainty has been developed and successfully applied at the global scale (Huntingford et al., 2006). Given the strong inter-model variability of model derived responses at the regional scale, explicit treatment of the uncertainty in the guess patterns is certainly an important issue for future regional-scale detection and attribution studies.

Adding additional sources of uncertainty in the analysis tends to further decrease the already low signal-to-noise ratio at the regional scale and successful detection and attribution is thus further complicated. On the other hand, there is room for improvement of the detectability of external changes at the regional scale. A few strategies with the potential to increase the signal-to-noise ratio at the regional scale are outlined below.

Inclusion of information from global-scale detection and attribution analyses in regional analyses could help to better constrain the regional results. For example, in a Bayesian formulation of the detection and attribution problem as proposed by Hasselmann (1998) and Schnur and Hasselmann (2005), global-scale information could be included as prior probabilities on the scaling factors for individual guess patterns. A caveat of this strategy is that for regions with very low signal-to-noise ratio, the outcome of a detection analysis is based almost exclusively on the global-scale information (S.Leroy, personal communication). Interpretation and communication of findings about regional quantities based on data from outside the region is expected to be difficult.

Further development of climate models used to infer about regional changes help to decrease the uncertainties in the guess patterns and thus increase the detectability of externally forced changes.

Simulations of multi-thousand ensembles of models, at present mainly used to assess the influence of model formulation on climate change projections (Piani et al., 2005, 2007), could be used either to reduce the uncertainties in the guess patterns or to explicitly deal with any of the above mentioned sources of uncertainty at the simulation level.

The use of quantities that describe the modification of the global-scale signals at the regional scale more generally could help to reduce inter-model differences in the response to different forcings. Good and Lowe (2006) analyze the difference between the regional area-average precipitation response to increasing GHG concentrations and minimum and maximum responses at the grid box level within a region and find that the ratio of small-scale trends to the area-average trends are very similar across models.

In addition to the suggestions listed above, the increasing anthropogenic forcing in the foreseeable future and the resulting strengthening of the signal will further help to reduce uncertainties in our understanding of regional climate change. Better understanding of regional climate change and the resulting confidence in regional projections in turn will help to prepare for the potentially harmful or beneficial effects of climate change.

Abbreviations

ALL	simulation with anthropogenic and natural forcing
ANT	simulation with anthropogenic forcing only
AOGCM	coupled Atmosphere-Ocean General Circulation Model
BALTEX	The Baltic Sea Experiment
CMIP3	Coupled Model Intercomparison Project
CRU	Climatic Research Unit
DWD	Deutscher Wetterdienst
ENSEMBLES	Ensembles-Based Predictions of Climate Changes and Their Impacts
ENSO	El Niño-Southern Oscillation
EOF	Empirical Orthogonal Function
GCM	General Circulation Model
GEWEX	Global Energy and Water Cycle Experiment
GHG	greenhouse gases
GPCC	Global Precipitation Climatology Centre
IPCC	Intergovernmental Panel on Climate Change
IPCC SRES	IPCC Special Report on Emissions Scenarios
NAO	North Atlantic Oscillation
NAT	simulation with natural forcing only
OLS	Ordinary Least Squares
PCC	Pattern Correlation Coefficient
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PRUDENCE	Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects
RCAO	Rossby Centre Regional Atmosphere-Ocean Model
RCM	Regional Climate Model
SLP	Sea Level Pressure
SMHI	Swedish Meteorological and Hydrological Institute
TLS	Total Least Squares
WCRP	World Climate Research Programme
WGCM	Working Group on Coupled Modelling

List of Figures

List of Tables

Bibliography

Adam, J. C. and Lettenmaier, D. P. (2003). Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research-Atmospheres*, 108(D9):4257.

Allen, M. R. and Stott, P. A. (2003). Estimating signal amplitudes in optimal fingerprinting, part i: theory. *Climate Dynamics*, 21(5-6):477–491.

Allen, M. R. and Tett, S. F. B. (1999). Checking for model consistency in optimal fingerprinting. *Climate Dynamics*, 15(6):419–434.

Barnett, T. P., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., Rasmusson, E., Roeckner, E., Ropelewski, C., Santer, B., and Tett, S. (1999). Detection and attribution of recent climate change: A status report. *Bulletin of the American Meteorological Society*, 80(12):2631–2659.

Barnett, T. P., Pierce, D. W., and Schnur, R. (2001). Detection of anthropogenic climate change in the world's oceans. *Science*, 292(5515):270–274.

Bhend, J. and von Storch, H. (2008). Consistency of observed winter precipitation trends in northern europe with regional climate change projections. *Climate Dynamics*, 31(1):17–28.

Bhend, J. and von Storch, H. (2009). Is greenhouse gas forcing a plausible explanation for the observed warming in the Baltic Sea catchment area? *Boreal Environment Research*, 14:81–88.

Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F. B., and Jones, P. D. (2006). Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research-Atmospheres*, 111(D12).

Christensen, J. and Christensen, O. (2007). A summary of the prudence model projections of changes in european climate by the end of this century. *Climatic Change*, 81(0):7–30.

Christensen, J., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P. (2007). *Climate Change 2007: The Physical Science Basis*, chapter Regional Climate Projections, pages 847–940. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Christidis, N., Stott, P. A., Brown, S., Hegerl, G. C., and Caesar, J. (2005). Detection of changes in temperature extremes during the second half of the 20th century. *Geophysical Research Letters*, 32(20):L20716.

Cubasch, U., Meehl, G., Boer, G., Stouffer, R., Dix, M., Noda, A., Senior, C., Raper, S., and Yap, K. (2001). Projections of future climate change. In Houghton, J., Ding, Y., Griggs, D., Noguer, M., van der Linden, P., Dai, X., Maskell, K., and Johnson, C., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, pages 99–181. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

D'Andrea, F., Tibaldi, S., Blackburn, M., Boer, G., Déqué, M., Dix, M., Dugas, B., Ferranti, L., Iwasaki, T., Kitoh, A., et al. (1998). Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979-1988. *Climate Dynamics*, 14(6):385–407.

Fischer, E. M., Luterbacher, J., Zorita, E., Tett, S. F. B., Casty, C., and Wanner, H. (2007). European climate response to tropical volcanic eruptions over the last half millennium. *Geophys. Res. Lett.*, 34:–.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D., Haywood, J., Lean, J., Lowe, D., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schulz, M., and Dorland, R. V. (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*, chapter Changes in Atmospheric Constituents and in Radiative Forcing, pages 129–234. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Gillett, N. P. (2005). Northern hemisphere circulation. *Nature*, 437(7058):496–496.

Gillett, N. P., Allen, M. R., McDonald, R. E., Senior, C. A., Shindell, D. T., and Schmidt, G. A. (2002a). How linear is the Arctic Oscillation response to greenhouse gases? *Journal of Geophysical Research-Atmospheres*, 107(D3):4022.

Gillett, N. P., Stone, D. A., Stott, P. A., Nozawa, T., Karpechko, A. Y., Hegerl, G. C., Wehner, M. F., and Jones, P. D. (2008a). Attribution of polar warming to human influence. *Nature Geoscience*, 1(11):750–754.

Gillett, N. P., Stott, P. A., and Santer, B. D. (2008b). Attribution of cyclogenesis region sea surface temperature change to anthropogenic influence. *Geophysical Research Letters*, 35(9).

Gillett, N. P., Weaver, A. J., Zwiers, F. W., and Wehner, M. F. (2004). Detection of volcanic influence on global precipitation. *Geophysical Research Letters*, 31(12):L12217.

Gillett, N. P., Zwiers, F. W., Weaver, A. J., Hegerl, G. C., Allen, M. R., and Stott, P. A. (2002b). Detecting anthropogenic influence with a multi-model ensemble. *Geophysical Research Letters*, 29(20):1970.

Gillett, N. P., Zwiers, F. W., Weaver, A. J., and Stott, P. A. (2003). Detection of human influence on sea-level pressure. *Nature*, 422(6929):292–294.

Giorgi, F. and Bi, X. (2009). Time of emergence (toe) of ghg-forced precipitation change hot-spots. *Geophys. Res. Lett.*, 36:–.

Good, P. and Lowe, J. (2006). Emergent behavior and uncertainty in multimodel climate projections of precipitation trends at small spatial scales. *Journal of Climate*, 19(21):5554–5569.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A. (2000). The simulation of sst, sea ice extents and ocean heat transports in a version of the hadley centre coupled model without flux adjustments. *Climate Dynamics*, 16(2-3):147–168.

Hansen, J., Russell, G., Lacis, A., Fung, I., Rind, D., and Stone, P. (1985). Climate response times: Dependence on climate sensitivity and ocean mixing. *Science*, 229(4716):857–859.

Hasselmann, K. (1976). Stochastic climate models .1. theory. *Tellus*, 28(6):473–485.

Hasselmann, K. (1979). *Meteorology of Tropical Oceans*, chapter On the signal-to-noise problem in atmospheric response studies, pages 251–259. Royal Meteorological Society.

Hasselmann, K. (1993). Optimal fingerprints for the detection of time-dependent climate-change. *Journal of Climate*, 6(10):1957–1971.

Hasselmann, K. (1997). Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dynamics*, 13(9):601–611.

Hasselmann, K. (1998). Conventional and bayesian approach to climate-change detection and attribution. *Quarterly Journal of the Royal Meteorological Society*, 124(552):2541–2565.

Hasumi, H. and Emori, S. (2004). K-1 coupled GCM (MIROC) description. K-1 technical report no. 1, Center for Climate System Research (CCSR), University of Tokyo; National Institute for Environmental Studies (NIES); Frontier Research Center for Global Change (FRCGC).

Hegerl, G., Zwiers, F. W., Braconnot, P., Gillett, N., Luo, Y., Orsini, J. M., N. Nicholls, J. P., and Stott, P. (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*, chapter Understanding and Attributing Climate Change, pages

663–745. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hegerl, G. C., Hasselmann, K., Cubasch, U., Mitchell, J. F. B., Roeckner, E., Voss, R., and Waszkewitz, J. (1997). Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Climate Dynamics*, 13(9):613–634.

Hegerl, G. C., Karl, T. R., Allen, M., Bindoff, N. L., Gillett, N., Karoly, D., Zhang, X. B., and Zwiers, F. (2006). Climate change detection and attribution: Beyond mean temperature signals. *Journal of Climate*, 19(20):5058–5077.

Hegerl, G. C., Stott, P. A., Allen, M. R., Mitchell, J. F. B., Tett, S. F. B., and Cubasch, U. (2000). Optimal detection and attribution of climate change: sensitivity of results to climate model differences. *Climate Dynamics*, 16(10-11):737–754.

Hegerl, G. C., von Storch, H., Hasselmann, K., Santer, B. D., Cubasch, U., and Jones, P. D. (1996). Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *Journal of Climate*, 9(10):2281–2306.

Hewitt, C. and Griggs, D. (2004). Ensembles-based predictions of climate changes and their impacts. *EOS*, 85(52):566–567.

Huntingford, C., Stott, P. A., Allen, M. R., and Lambert, F. H. (2006). Incorporating model uncertainty into attribution of observed temperature change. *Geophysical Research Letters*, 33(5):L05710.

Hurrell, J. W. (1995). Decadal trends in the north-atlantic oscillation - regional temperatures and precipitation. *Science*, 269(5224):676–679.

Jacob, D., Bärring, L., Christensen, O., Christensen, J., de Castro, M., Déqué, M., Giorgi, F., Hagemann, S., Hirschi, M., Jones, R., Kjellström, E., Lenderink, G., Rockel, B., Sánchez, E., Schär, C., Seneviratne, S., Somot, S., van Ulden, A., and van den Hurk, B. (2007). An inter-comparison of regional climate models for europe: model performance in present-day climate. *Climatic Change*, 81(0):31–52.

Jones, P. D., Jonsson, T., and Wheeler, D. (1997). Extension to the North Atlantic Oscillation using early instrumental pressure observations from gibraltar and south-west iceland. *International Journal of Climatology*, 17(13):1433–1450.

Jones, P. D. and Moberg, A. (2003). Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001. *Journal of Climate*, 16(2):206–223.

Jones, P. D., Osborn, T. J., Briffa, K. R., Folland, C. K., Horton, E. B., Alexander, L. V., Parker, D. E., and Rayner, N. A. (2001a). Adjusting for sampling density in grid box land and ocean surface temperature time series. *Journal of Geophysical Research-Atmospheres*, 106(D4):3371–3380.

Jones, P. W. (1999). First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, 127(9):2204–2210.

Jones, R., Murphy, J., Hassell, D., and Taylor, R. (2001b). Ensemble mean changes in a simulation of the european climate of 2071-2100 using the new hadley centre regional modelling system hadam3h/hadrm3h. Technical report, Hadley Centre, Met Office, Bracknell, UK.

Jun, M., Knutti, R., and Nychka, D. (2008). Spatial analysis to quantify numerical model bias and dependence. *Journal of the American Statistical Association*, 103(483):934–947.

Kiktev, D., Sexton, D., Alexander, L., and Folland, C. (2003). Comparison of modeled and observed trends in indices of daily climate extremes (vol 16, pg 3560, 2003). *Journal of Climate*, 16:3560–3571.

Kjellström, E. (2004). Recent and future signatures of climate change in europe. *Ambio*, 33(4-5):193–198.

Kjellstrom, E., Doscher, R., and Meier, H. E. M. (2005). Atmospheric response to different sea surface temperatures in the Baltic Sea: coupled versus uncoupled regional climate model experiments. *Nordic Hydrology*, 36(4-5):397–409.

Knutti, R. (2008). Why are climate models reproducing the observed global surface warming so well? *Geophysical Research Letters*, 35(18).

Knutti, R. and Hegerl, G. (2008). The equilibrium sensitivity of the earth's temperature to radiation changes. *Nature Geoscience*, 1:735–743.

Lamb, P. J. and Peppler, R. A. (1987). North-atlantic oscillation - concept and an application. *Bulletin of the American Meteorological Society*, 68(10):1218–1225.

Lambert, F. H., Stott, P. A., Allen, M. R., and Palmer, M. A. (2004). Detection and attribution of changes in 20th century land precipitation. *Geophysical Research Letters*, 31(10):L10203.

Le Treut, H., Somerville, R., Cubasch, U., Ding, Y., Mauritzen, C., Mokssit, A., Peterson, T., and Prather, M. (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*, chapter Historical Overview of Climate Change Science, pages 93–127. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Legates, D. R. and Willmott, C. J. (1990). Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology*, 10(2):111–127.

Meehl, G., Stocker, T., Collins, W., Friedlingstein, P., Gaye, A., Gregory, J., Kitoh, A., Knutti, R., Murphy, J., Noda, A., Raper, S., Watterson, I., Weaver, A., and Zhao, Z.-C.

(2007a). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, chapter Global Climate Projections, pages 747–845. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

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

Min, S. K., Zhang, X. B., and Zwiers, F. (2008). Human-induced arctic moistening. *Science*, 320(5875):518–520.

Mitchell, J., Karoly, D., Hegerl, G., Zwiers, F., Allen, M., and Marengo, J. (2001). *Detection of Climate Change and Attribution of Causes*, pages 695–738. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Mitchell, T. D. and Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, 25(6):693–712.

Nakićenović, N. and Intergovernmental Panel on Climate Change Response Strategies Working Group (2000). *Special report on emissions scenarios*. Cambridge University Press : Published for the Intergovernmental Panel on Climate Change, Cambridge [England]; New York.

Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F. F., Wakata, Y., Yamagata, T., and Zebiak, S. E. (1998). Enso theory. *Journal of Geophysical Research-Oceans*, 103(C7):14261–14290.

New, M., Hulme, M., and Jones, P. (1999). Representing twentieth-century space-time climate variability. part i: Development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate*, 12(3):829–856.

Ogura, T., Emori, S., Webb, M. J., Tsushima, Y., Yokohata, T., Abe-Ouchi, A., and Kimoto, M. (2008). Towards understanding cloud response in atmospheric GCMs: The use of tendency diagnostics. *Journal of the Meteorological Society Of Japan*, 86(1):69–79.

Osborn, T. J. (2004). Simulating the winter North Atlantic Oscillation: the roles of internal variability and greenhouse gas forcing. *Climate Dynamics*, 22(6-7):605–623.

Osborn, T. J., Briffa, K. R., Tett, S. F. B., Jones, P. D., and Trigo, R. M. (1999). Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Climate Dynamics*, 15(9):685–702.

Piani, C., Frame, D. J., Stainforth, D. A., and Allen, M. R. (2005). Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophysical Research Letters*, 32(23).

Piani, C., Sanderson, B., Giorgi, F., Frame, D. J., Christensen, C., and Allen, M. R. (2007). Regional probabilistic climate forecasts from a multithousand, multimodel ensemble of simulations. *Journal of Geophysical Research-Atmospheres*, 112(D24).

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,.

Räisänen, J. (2001). Hiilidioksidin lisääntymisen vaikutus pohjois-euroopan ilmastoon globaaleissa ilmastomalleissa (the impact of increasing carbon dioxide on the climate of northern europe in global climate models). *Terra*, 113:139–151.

Räisänen, J., Hansson, U., Ullerstig, A., Döscher, R., Graham, L. P., Jones, C., Meier, H. E. M., Samuelsson, P., and Willén, U. (2003). GCM driven simulations of recent and future climate with the rossby centre coupled atmosphere - Baltic Sea regional climate model RCAO. Technical report, Rossby Centre, SMHI, Norrköping, Sweden.

Räisänen, J., Hansson, U., Ullerstig, A., Doscher, R., Graham, L. P., Jones, C., Meier, H. E. M., Samuelsson, P., and Willen, U. (2004). European climate in the late twenty-first century: regional simulations with two driving global models and two forcing scenarios. *Climate Dynamics*, 22:13–31.

Ramanathan, V., Crutzen, P. J., Kiehl, J. T., and Rosenfeld, D. (2001). Atmosphere - aerosols, climate, and the hydrological cycle. *Science*, 294(5549):2119–2124.

Ramaswamy, V., Boucher, O., Haigh, J., Hauglustaine, D., Haywood, J., Myhre, G., Nakajima, T., Shi, G., and Solomon, S. (2001). Radiative forcing of climate change. In Houghton, J., Ding, Y., Griggs, D., Noguer, M., van der Linden, P., Dai, X., Maskell, K., and Johnson, C., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, pages 99–181. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Randall, D., Wood, R., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R., Sumi, A., and Taylor, K. (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*, chapter Climate Models and Their Evaluation, pages 589–662. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Rauthe, M., Hense, A., and Paeth, H. (2004). A model intercomparison study of climate change-signals in extratropical circulation. *International Journal of Climatology*, 24(5):643–662.

Roeckner, E., Bengtsson, L., Feichter, J., Lelieveld, J., and Rodhe, H. (1999). Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle. *Journal of Climate*, 12(10):3004–3032.

Santer, B. D., Taylor, K. E., Wigley, T. M. L., Penner, J. E., Jones, P. D., and Cubasch, U. (1995). Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dynamics*, 12(2):77–100.

Santer, B. D., Wigley, T. M. L., and Jones, P. D. (1993). Correlation methods in fingerprint detection studies. *Climate Dynamics*, 8(6):265–276.

Santer, B. D., Wigley, T. M. L., Simmons, A. J., Kallberg, P. W., Kelly, G. A., Uppala, S. M., Ammann, C., Boyle, J. S., Bruggemann, W., Doutriaux, C., Fiorino, M., Mears, C., Meehl, G. A., Sausen, R., Taylor, K. E., Washington, W. M., Wehner, M. F., and Wentz, F. J. (2004). Identification of anthropogenic climate change using a second-generation reanalysis. *Journal of Geophysical Research-Atmospheres*, 109(D21).

Schneider, T. and Held, I. M. (2001). Discriminants of twentieth-century changes in earth surface temperatures. *Journal of Climate*, 14(3):249–254.

Schneider, U., Fuchs, T., Meyer-Christoffer, A., and Rudolf, B. (2008). Global precipitation analysis products of the gpcc. Technical report, Global Precipitation Climatology Centre (GPCC), Deutscher Wetterdienst.

Schnur, R. and Hasselmann, K. I. (2005). Optimal filtering for bayesian detection and attribution of climate change. *Climate Dynamics*, 24(1):45–55.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly spaced data. In *Proc. 23rd ACM Nat. Conf.*, pages 517–524. Brandon/Systems Press, Princeton, NJ.

Shindell, D. T., Schmidt, G. A., Mann, M. E., and Faluvegi, G. (2004). Dynamic winter climate response to large tropical volcanic eruptions since 1600. *J. Geophys. Res.*, 109:–.

Spagnoli, B., Planton, S., Deque, M., Mestre, O., and Moisselin, J. M. (2002). Detecting climate change at a regional scale: The case of france. *Geophysical Research Letters*, 29(10):1450.

Stenchikov, G., Hamilton, K., Stouffer, R. J., Robock, A., Ramaswamy, V., Santer, B., and Graf, H.-F. (2006). Arctic oscillation response to volcanic eruptions in the ipcc ar4 climate models. *J. Geophys. Res.*, 111:–.

Stephenson, D. B., Pavan, V., Collins, M., Junge, M. M., and Quadrelli, R. (2006). North Atlantic Oscillation response to transient greenhouse gas forcing and the impact on european winter climate: a cmip2 multi-model assessment. *Climate Dynamics*, 27(4):401–420.

Stott, P. A. (2003). Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophysical Research Letters*, 30(14):1728.

Stott, P. A., Kettleborough, J. A., and Allen, M. R. (2006). Uncertainty in continental-scale temperature predictions. *Geophysical Research Letters*, 33(2):L02708.

Stott, P. A. and Tett, S. F. B. (1998). Scale-dependent detection of climate change. *Journal of Climate*, 11(12):3282–3294.

Stott, P. A., Tett, S. F. B., Jones, G. S., Allen, M. R., Mitchell, J. F. B., and Jenkins, G. J. (2000). External control of 20th century temperature by natural and anthropogenic forcings. *Science*, 290(5499):2133–2137.

Stouffer, R. J., Manabe, S., and Vinnikov, K. Y. (1994). Model assessment of the role of natural variability in recent global warming. *Nature*, 367(6464):634–636.

Tebaldi, C. and Knutti, R. (2007). The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences And Engineering Sciences*, 365(1857):2053–2075.

Tett, S. F. B., Stott, P. A., Allen, M. R., Ingram, W. J., and Mitchell, J. F. B. (1999). Causes of twentieth-century temperature change near the earth's surface. *Nature*, 399(6736):569–572.

The BACC author team (2008). *Assessment of Climate Change in the Baltic Sea Basin*. Springer Berlin, Heidelberg, New York. accepted.

Thompson, D. W. J. and Wallace, J. M. (2001). Regional climate impacts of the northern hemisphere annular mode. *Science*, 293(5527):85–89.

Trenberth, K., Jones, P., Ambenje, P., Bojariu, R., Easterling, D., Tank, A. K., Parker, D., Rahimzadeh, F., Renwick, J., Rusticucci, M., Soden, B., and Zhai, P. (2007). *Climate Change 2007: The Physical Science Basis*, chapter Observations: Surface and Atmospheric Climate Change, pages 235–336. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

van Loon, H. and Rogers, J. (1978). The Seesaw in Winter Temperatures between Greenland and Northern Europe. Part I: General Description. *Monthly Weather Review*, 106(3):296–310.

van Oldenborgh, G. J., Drijfhout, S., van Ulden, A., Haarsma, R., Sterl, A., Severijns, C., Hazeleger, W., and Dijkstra, H. (2009). Western Europe is warming much faster than expected. *Climate of the Past*, 5:1–12.

von Storch, H. and Hannoschöck, G. (1985). Statistical aspects of estimated principal vectors (eofs) based on small sample sizes. *Journal of Climate and Applied Meteorology*, 24:716–724.

von Storch, H. and Zwiers, F. (1999). *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, United Kingdom.

Wanner, H., Bronnimann, S., Casty, C., Gyalistras, D., Luterbacher, J., Schmutz, C., Stephenson, D. B., and Xoplaki, E. (2001). North Atlantic Oscillation - concepts and studies. *Surveys in Geophysics*, 22(4):321–382.

Washington, W. M., Weatherly, J. W., Meehl, G. A., Semtner, A. J., Bettge, T. W., Craig, A. P., Strand, W. G., Arblaster, J., Wayland, V. B., James, R., and Zhang, Y. (2000). Parallel climate model (pcm) control and transient simulations. *Climate Dynamics*, 16(10-11):755–774.

Wigley, T. M. L. and Raper, S. C. B. (1990). Natural variability of the climate system and detection of the greenhouse-effect. *Nature*, 344(6264):324–327.

Wigley, T. M. L. and Schlesinger, M. E. (1985). Analytical solution for the effect of increasing co2 on global mean temperature. *Nature*, 315(6021):649–652.

Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J., and Wilks, D. S. (1998). Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, 34(11):2995–3008.

Wilks, D. S. (1997). Resampling hypothesis tests for autocorrelated fields. *Journal of Climate*, 10(1):65–82.

Willmott, C., Rowe, C., and Philpot, W. (1985). Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *The American Cartographer*, 12:5–16.

Yang, D. Q., Kane, D., Zhang, Z. P., Legates, D., and Goodison, B. (2005). Bias corrections of long-term (1973-2004) daily precipitation data over the northern regions. *Geophysical Research Letters*, 32(19):L19501.

Yokohata, T., Emori, S., Nozawa, T., Tsushima, Y., Ogura, T., and Kimoto, M. (2005). Climate response to volcanic forcing: Validation of climate sensitivity of a coupled atmosphere-ocean general circulation model. *Geophysical Research Letters*, 32(21).

Yongqiang, Y., Rucong, Y., Xuehong, Z., and Hailong, L. (2002). A flexible coupled ocean-atmosphere general circulation model. *Advances in Atmospheric Sciences*, 19(1):169–190.

Yongqiang, Y., Xuehong, Z., and Yufu, G. (2004). Global coupled ocean-atmosphere general circulation models in lasg/iap. *Advances in Atmospheric Sciences*, 21(3):444–455.

Zhang, X., Zwiers, F., Hegerl, G. C., Lambert, F. H., Gillett, N. P., Solomon, S., Stott, P. A., and T., N. (2007). Detection of human influence on twentieth-century precipitation trends. *Nature*, 448:461–465.

Zhang, X. B., Zwiers, F. W., and Stott, P. A. (2006). Multimodel multisignal climate change detection at regional scale. *Journal of Climate*, 19(17):4294–4307.

Zorita, E., Stocker, T. F., and von Storch, H. (2008). How unusual is the recent series of warm years? *Geophysical Research Letters*, 35:–.

Zwiers, F. W. and von Storch, H. (1995). Taking serial-correlation into account in tests of the mean. *Journal of Climate*, 8(2):336–351.

Zwiers, F. W. and Zhang, X. B. (2003). Toward regional-scale climate change detection. *Journal of Climate*, 16(5):793–797.

Wetzel, P. (2005): **Interannual and Decadal Variability in the Air-Sea Exchange of CO₂.**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 7/2004, pp. 77

Stier, P. (2005): **Towards the Assessment of the Aerosol Radiative Effects - A Global Modelling Approach.** Reports on Earth System Science, Max Planck Institute for Meteorology, No. 9/2004, pp. 111

Zuo, X. (2005): **Annual Hard Frosts and Economic Growth.**
Department of Economics, University of Hamburg, Hamburg, pp. 112

Jung, M. (2005): **Carbon sequestration options in the international climate regime.**
Department of Economics, University of Hamburg, Hamburg, pp. 119

Zhou, Y. (2005): **Economic Analysis of Selected Environmental Issues in China**
Department of Economics, University of Hamburg, Hamburg, pp. 101

Devasthale, A. (2005): **Aerosol Indirect Effect in the Thermal Spectral Range as Seen from Satellites**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 16/2005, pp. 70

Zandersen, M. (2005): **Aerosol Valuing Forest Recreation in Europe: Time and Spatial Considerations**
Department of Economics, University of Hamburg, Hamburg, pp. 125

Xuefeng Cui (2005*):* **Interactions between Climate and Land Cover Changes on the Tibetan Plateau**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 17/2005, pp. 125

Stehfest, Elke (2005*):* **Modelling of global crop production and resulting N2O emissions**
Zentrum für Umweltsystemforschung Universität Kassel pp. 125

Kloster, Silvia (2006*):* **DMS cycle in the ocean-atmosphere system and its response to anthropogenic perturbations.** Reports on Earth System Science, Max Planck Institute for Meteorology, No. 19/2006, pp. 82

Criscuolo, Luca (2006*):* **Assessing the Agricultural System and the Carbon Cycle under Climate Change in Europe using a Dynamic Global Vegetation Model**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 21/2006, pp. 140

Tiwari, Yogesh Kumar (2006*):* **Constraints of Satellite Derived CO2 on Carbon Sources and Sinks**
Technical Reports, Max-Planck-Institut für Biogeochemie, No.7/2006, pp.125

Schurgers, Guillaume (2006*):* **Constraints Long-term interactions between vegetation and climate - Model simulations for past and future -** Reports on Earth System Science, Max Planck Institute for Meteorology, No. 27/2006, pp. 135

Ronneberger, Kerstin Ellen (2006*):* **The global agricultural land-use model KLUM - A coupling tool for integrated assessment -** Reports on Earth System Science, Max Planck Institute for Meteorology, No. 26/2006, pp. 123

Woth, Katja (2006*):* **Regionalization of global climate change scenarios: An ensemble study of possible changes in the North Sea storm surge statistics**
Department for Earth Sciences, University of Hamburg, Hamburg, pp. 97

Hoelzemann, Judith Johanna (2006*):* **Global Wildland Fire Emission Modeling for Atmospheric Chemistry Studies**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 28/2006, pp. 206

Gaslikova, Lidia (2006*):* **High-resolution wave climate analysis in the Helgoland area**
Department for Earth Sciences, University of Hamburg, Hamburg, pp. 90

Grossmann, Iris (2006*):* **Future perspectives for the Lower Elbe Region 2005–2030: Climate Trends and Globalisation**
GKSS-Forschungszentrum, Geesthacht, pp. 175

Narayan, Caroline (2006*):* **CO2 fluxes and concentration patterns over Eurosiberia: A study using terrestrial biosphere models and the regional atmosphere model REMO**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 29/2006, pp. 242

Vizcaino, Miren (2006): **Long-term interactions between ice sheets and climate under anthropogenic greenhouse forcing Simulations with two complex Earth System Models**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 30/2006, pp. 187

Schwoon, Malte (2006): **Managing the Transition to Hydrogen and Fuel Cell Vehicles – Insights from Agent-based and Evolutionary Models –**
 Reports on Earth System Science, Max Planck Institute for Meteorology, No. 32/2006, pp. 132

Link, Peter Michael (2006): **Modeling the economic impacts of changes in thermohaline circulation with an emphasis on the Barents Sea fisheries**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 33/2006, pp. 185

Li, Qian (2006): **Climatological analysis of planetary wave propagation in Northern Hemisphere winter**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 35/2006, pp. 153

Weis, Philipp (2006): **Ocean Tides and the Earth's Rotation - Results of a High-Resolving Ocean Model forced by the Lunisolar Tidal Potential**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 36/2006, pp. 115

Heistermann, Maik (2006): **Modelling the Global Dynamics of Rain-fed and Irrigated Croplands**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 37/2006, pp. 152

Kristina Trusilova (2006): **Urbanization impacts on the climate in Europe**
Technical Reports, Max Planck Institute for Biogeochemie, No. 9/2006, pp. 82

Xiuhua Zhu (2007): **Low frequency variability of the Meridional Overturning Circulation**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 39/2007, pp. 158

Christoph Müller (2007): **Climate Change and Global Land-Use Patterns — Quantifying the Human Impact on the Terrestrial Biosphere**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 41/2007, pp. 126

Sven Kotlarski (2007): **A Subgrid Glacier Parameterisation for Use in Regional Climate Modelling**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 42/2007, pp. 179

Daniela Matei (2007): **Decadal Variability: Internal Variability and Sensitivity to Subtropics**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 44/2007, pp. 107

Adetutu Aghedo (2007): **The impact of african air pollution: A global chemistry climate model study**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 45/2007, pp. 142

Melissa Anne Pfeffer (2007): **The Relative Influences of Volcanic and Anthropogenic Emissions on Air Pollution in Indonesia as Studied With a Regional Atmospheric Chemistry and Climate Model**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 46/2007, pp. 119

Felix Landerer (2007): **Sea Level and Hydrological Mass Redistribution in the Earth System: Variability and Anthropogenic Change**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 47/2007, pp. 115

Angelika Heil (2007): **Indonesian Forest and Peat Fires: Emissions, Air Quality, and Human Health**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 50/2007, pp. 142

Manu Anna Thomas (2008): **Simulation of the climate impact of Mt. Pinatubo eruption using ECHAM5**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 52/2008, pp. 161

Martin Jung (2008): **Uncertainties of terrestrial carbon cycle modelling: Studies on gross carbon uptake of Europe**
Technical Reports - Max Planck Institute for Biogeochemistry, No. 11/2008, pp. 142

Jörg Winterfeldt (2008): **Comparison of measured and simulated wind speed data in the North Atlantic**
GKSS 2008/2 pp. 142

Katsumasa Tanaka (2008): **Inverse Estimation for the Simple Earth System Model ACC2 and its Applications,** pp. 296

Krzysztof Stanislaw Katrynski (2008): **Quantifying the terrestrial carbon budget of middle latitude ecosystems: the role of Volatile Organic Compounds**
Technical Reports - Max Planck Institute for Biogeochemistry, No. 12/2008, pp. 170

Dorothea F. Banse (2008): **The influence of aerosols on North Atlantic cyclones**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 53/2008, pp. 94

Birgit Hünicke (2008): **Atmospheric forcing of decadal Baltic Sea level variability in the last 200 years: A statistical analysis**
GKSS 2008/4 pp. 100

Francesca Guglielmo (2008): **Global cycling of semivolatile organic compounds in the marine and total environment -A study using a comprehensive model**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 54/2008, pp. 119

Heinz Jürgen Punge (2008): **The quasi-biennial oscillation: Representation in numerical models and effects on stratospheric transport and trace gases**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 55/2008, pp. 91

Rita Seiffert (2008): **Impact of small-scale fluctuations on climate sensitivity and its stochastic analysis**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 58/2009, pp. 95

Gabriela Sousa Santos (2008): **The effect of halogens on global tropospheric ozone**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 59/2009, pp. 143

Diana Rechid (2008): **On biogeophysical interactions between vegetation phenology and climate simulated over Europe**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 61/2009, pp. 130

Christine Schleupner (2008): **GIS as integrating tool in Sustainability and Global Change**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 62/2009, pp. 184

Jana Sillmann (2009): **Extreme Climate Events and Euro-Atlantic Atmospheric Blocking in Present and Future Climate Model Simulations**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 63/2009, pp. 89

Hui Wan (2009): **Developing and testing a hydrostatic atmospheric dynamical core on triangular grids**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 65/2009, pp. 153

Julia Pongratz (2009) **A model estimate on the effect of anthropogenic land cover change on the climate of the last millennium**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 68/2009, pp. 136

Malte Heinemann (2009) **Warm and sensitive Paleocene-Eocene climate**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 70/2009, pp. 113

Kerstin Verzano (2009) **Climate Change Impacts on Flood Related Hydrological Processes: Further Development and Application of a Global Scale Hydrological Model**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 71/2009, pp. 146

Kai Lessmann (2009) **Endogenous Technological Change in Strategies for Mitigating Climate Change**
pp. 174

Dritan Osmani (2009) **Stability and Burden Sharing Emissions in International Environmental Agreements, A Game -Theoretic Approach**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 73/2009, pp. 176

David Anthoff (2009) **Equity and Climate Change: Applications of FUND**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 74/2009, pp. 169

Claas Teichmann (2010) **Climate and Air Pollution Modelling in South America with Focus on Megacities**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 76/2009, pp. 167

Declan O'Donnell (2010) **Towards the Assessment of the Climate Effects of Secondary Organic Aerosols**
Reports on Earth System Science, Max Planck Institute for Meteorology, No. 77/2009, pp. 113