

Sensitivity of climate change signals deduced from multi-model Monte Carlo experiments

Heiko Paeth*, Andreas Hense

Meteorologisches Institut, Universität Bonn, Auf dem Hügel 20, 53121 Bonn, Germany

ABSTRACT: Near-surface air temperature and precipitation are frequently presumed to be indicators of anthropogenic climate change; in this study, they are compared with each other in terms of their simulated signal-to-noise ratio. Special emphasis is given to the characteristic time scales of temperature and rainfall signals. By means of analysis of variance, based on an ensemble of 4 greenhouse-gas (GHG) induced, long-term, coupled climate model experiments, the contributions of internal noise and external GHG forcing to total variability are quantified. The part of the variance accounted for externally is a valuable measure for the predictability of a meteorological variable. Further, the sensitivity to different climate models is evaluated, based on multi-model ensembles. With regard to temperature, the GHG forcing accounts for more than 90 % of the total variance in the tropics and 50 to 70 % in the extra-tropics, implicating a warming trend all over the globe. On the other hand, precipitation is largely induced by internal variability. The GHG scenario accounts for at best 30 % of the variance in rainfall. The corresponding trend pattern is more differentiated, predicting dryer conditions in the subtropical regions and rising precipitation elsewhere. The temperature signal, arising from 1970 onward, is evident on time scales of 30 yr and longer. For precipitation, there is a weak but statistically significant low-frequency signal on a regional scale, which is apparent for 60 yr time slices but not at shorter time scales. At present, precipitation, in contrast to temperature, represents an unfavourable detection variable for GHG-induced climate change. The super-ensemble approach reveals that intermodel variations contribute the major part to total variability in high latitudes, whereas the GHG impact on temperature is still evident in the tropics. Large model uncertainties occur in the regions of sea ice margins, mountains and over Antarctica.

KEY WORDS: Climate change · Detection variable · Signal-to-noise ratio · Multi-model ensemble · ANOVA

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

There is a consensus among the scientific community (Houghton et al. 2001) that the climate of the Earth is affected by anthropogenic activities such as greenhouse-gas (GHG) and aerosol emissions. Thus, it is a principal challenge for climatologists to detect indicators of man-made climate change in observational data. While the physical mechanism of, for example, an enhanced greenhouse effect due to anthropogenic GHG emissions is plausible (Houghton et al. 1990, 1992, Flohn et al. 1992), in practice statistical evidence

for a causal relationship between increasing GHG concentrations and recent climate change is hard to find, as other factors such as the sun, the oceans or volcanic eruptions may induce variability on the same decadal to interdecadal time scales as the GHGs (Houghton et al. 1990, 1992). One possible approach to distinguish the GHG effect from the others is the use of climate model simulations with clearly defined forcing scenarios. The model predicts a spatio-temporal pattern ('fingerprint') of GHG-induced climate change, which in turn has to be found in the spatial and temporal structure of observational data sets in order to judge whether or not and to what extent the real climate corresponds to that prescribed model scenario (Houghton

*E-mail: hpaeth@uni-bonn.de

et al. 1990). In recent years substantial scientific work has been devoted to such signal analyses (e.g. Santer et al. 1996, North & Stevens 1998, Barnett et al. 2001, Schneider & Held 2001).

One of the factors in competition with the possible anthropogenic climate signal is the natural variability (Hunt 1998, Hulme et al. 1999) which arises from the nonlinear, chaotic interactions of large-scale turbulent flows in the atmosphere and ocean. This natural variability can only be described by statistical methods. Therefore, the fingerprinting has to be embedded into a statistical approach as well. Special detection methods have been described, such as optimal fingerprinting (Hasselmann 1979, 1993), in order to deal with the natural variability in observational time series (Hegerl et al. 1996, 1997, 2001, Hegerl & North 1997, Paeth & Hense 2001).

But there is another aspect which has to be taken into consideration: the climate model itself produces internal variability due to the complexity of model equations and independent of the external forcing conditions (Keen & Murphy 1997). The predicted fingerprint which enters the above optimal methods as a deterministic component does not necessarily represent the climate response to the prescribed forcing scenario. Instead it may itself be contaminated by the model's inherent internal variability being unpredictable with regard to the changing external conditions, such as GHG concentrations. Efficient signal detection will rely on a detection variable which is dominated by the external forcing rather than by the internal variability of the model. One instrument which allows us to quantify the relationship between external and internal model variability is the realization of so-called Monte Carlo experiments (Brankovic et al. 1990, Cubasch et al. 1994, Paeth et al. 2002). They consist of several simulations with the same climate model based on identical model physics and forcing scenarios but varied initial conditions which are assumed to be the major source for internal variability besides other sources, such as randomized parametrizations (von Storch 2001). From the resulting ensemble members it can be estimated whether the external conditions dominates time-space evolution of a specific variable or whether the latter is mainly determined by perturbed initial conditions. The classical statistical tool to quantify this signal-to-noise ratio of detection variables is 1-way analysis of variance (1-way ANOVA), which divides total variability into internal and external contributions (von Storch & Zwiers 1999). In this context, we present a spatially, temporally and spectrally differentiated approach.

Climate change detection studies are confronted with the general uncertainty of climate models which usually differ in terms of resolution, parametrization

and forcing (Barnett et al. 1998, 2000, Hegerl et al. 2000). In principle, one does not know which model projection of future climate is most realistic, since the ability to reproduce observed present-day climate as only 1 realization of the climate system is not necessarily sufficient evidence for the reliability of model predictions. Therefore, it is of basic interest to quantify intermodel variations with regard to simulated climate change and internal variability. This question can be addressed by applying again a classical statistical tool, 2-way analysis of variance (2-way ANOVA). To obtain comparable intermodel variability, it is necessary to have a more or less identical external GHG forcing. Otherwise an uninterpretable mixing between intermodel variability and the forcing would occur. Therefore, our results are based on a super-ensemble (Kharin & Zwiers 2002) of two 4-member ensembles only: the German ECHAM3 ensemble (Roeckner et al. 1992) and the Hadley Centre coupled model GHG experiments (HADCM2) (Johns et al. 1997, Giorgi & Francisco 2000), in which the forcing is reasonably similar.

In this modelling study, we completely neglect the natural external factors such as solar variability and volcanic activity in order to isolate and quantify the impact of an enhanced GHG effect.

Among the large number of possible detection variables we have chosen annual means of near-surface air temperature at 2 m above ground (T2M) and annual sums of precipitation (PRE). This choice is obvious for several reasons: According to the physical understanding of the atmospheric radiation balance, T2M should be directly affected by an enhanced GHG effect, maybe with some modification by possible feedback processes which would induce additional uncertainty. This is actually the case for PRE, whose response to rising GHGs is more complex (Mahlman 1997), influenced, for example, by advection processes due to changing circulation patterns (Paeth et al. 1999, Ulbrich & Christoph 1999, Wibig 1999) or feedbacks via vegetation and land-cover changes (Chase et al. 2000, Levis et al. 2000). However, both variables, T2M and PRE, are among the major variables for climate impact studies determining the regional effects of a global climate change. The present work presents the relative contributions to the total variability of intermodel variability, internal chaotic variability and externally forced variations. This gives the spatial structure as well as the characteristic time scales of the GHG-induced signal and a measure of its predictability.

This study is an expansion of the method presented in the Intergovernmental Panel on Climate Change (IPCC) 2001 report, in which the multi-model ensemble mean climate change is split up into the common

forced component, intermodel variations and internal variability (Houghton et al. 2001, p. 537). In addition, our method provides insight into the statistical significance of the contributions and takes into account the interaction between different forcings and the climate response, as the model experiments compared with each other in the IPCC report are subject to different forcing scenarios, e.g. GHG only versus GHG + sulphate aerosols. This aspect is also quantified in our study. Moreover, we offer a more detailed and quantitative view into the spatial, temporal and spectral characteristics of climate-change fingerprints.

The following section gives some information about the climate model data and the statistical method. Thereafter, the results of the spatial, temporal and spectral 1-way ANOVA applied to the ECHAM3 simulations as well as the model intercomparison using 2-way ANOVA are described. The last chapter summarizes the main results and offers some conclusions for detection studies and recommendations for future climate modelling strategies.

2. DATA AND STATISTICS

2.1. Monte Carlo experiments

In this study, we refer to an ensemble of 4 simulations with GHG forcing according to the transient IPCC scenario SA90 (Houghton et al. 1990) realized with the coupled general circulation model (GCM) ECHAM3/LSG (Roeckner et al. 1992, Voss et al. 1998) for the period 1880–2049 and according to a 4-member ensemble produced by the HADCM2 covering the period 1860–2099 (Johns et al. 1997, Giorgi & Francisco 2000). The latter's GHG forcing is slightly weaker than that in ECHAM3, according to IS92A (Houghton et al. 1992), which amounts to somewhat less than 1 % yr⁻¹.

Both models are fully coupled atmosphere-ocean climate models. ECHAM3 is based on a modified weather forecast GCM of the European Centre for Medium-range Weather Forecasts and is run at the Max Planck Institute for Meteorology in Hamburg (Roeckner et al. 1992, Voss et al. 1998). The GHG scenario consists of observed concentrations until 1985 followed by an increase of about 1 % yr⁻¹. In this paper, these runs are referred to as NEIN (new early industrial), characterizing the clearly defined GHG forcing (Hegerl et al. 1996). The spectral resolution with wave number 21 (T21) has been interpolated to a regular 5° to 10° grid.

In addition to the original run, we have realized 3 further NEIN simulations, according to the Monte Carlo setup. This means that all realizations are driven

with the same forcing scenario and model physics, but with different initial states of atmospheric and oceanic conditions in January 1880. In this context, we have chosen 3 states of the corresponding control run (Voss et al. 1998), each 200 yr apart from each other. The resulting ensemble of 4 model runs then offers the possibility to extract internally and externally induced contributions to the total variability. Note that these realizations represent a relatively small number with respect to other Monte Carlo experiments, e.g. in numerical weather forecasting.

Validating the data of the ECHAM3/LSG coupled model has led to good agreement with the main features of observed geopotential height and temperature. Compared with previous model versions, precipitation is reproduced more realistically, but reveals some misjudgement predominantly in the subtropical regions (Roeckner et al. 1992). However, GHG-forced climate model experiments are supposed to underestimate natural variability as additional external factors are not considered. This is especially the case for the ECHAM3/LSG model (Paeth & Hense 2001). On the other hand, the prescribed GHG Scenario A of the IPCC 1990 report actually seems to be overestimated with regard to recent progress in climate protection and emission reductions. Thus, GHG concentrations are supposed to rise by less than 1 % yr⁻¹ (Houghton et al. 1992). As a consequence, the signal-to-noise ratio within the NEIN ensemble may suggest an overestimated GHG signal. This has to be taken into account when interpreting the results in a quantitative way.

HADCM2 provides an ensemble of 4 GHG integrations covering the period 1860–2099 with a GHG forcing slightly weaker than in ECHAM3, amounting to somewhat less than 1 % yr⁻¹ according to IS92A (Houghton et al. 1992). HADCM2 is a grid point model with a resolution of 3.75° zonally and 2.5° meridionally. This is also interpolated to the joint regular analysis grid with a resolution of 5° to 10°. The Hadley Centre model also reproduces the main features of the observed climate system in a reliable way (Johns et al. 1997). The fact that this model does not conserve mass (Osborn et al. 1999) is irrelevant to the variables considered in this study.

2.2. One-way and two-way analysis of variance

1-way ANOVA is a classical univariate statistical method to distinguish between internal and external components of total variability, given an ensemble of model simulations with identical external forcing (Scheffé 1963, von Storch & Zwiers 1999). Given the data (e.g. T2M at a single grid point) X_{jk} with $j = 1 \dots J$ time steps and $k = 1 \dots M$ ensemble members, a linear

model can be defined, consisting of the overall mean μ , a signal common to all ensemble members β_j —the so-called treatment effect—and the residual ε_{jk} , which is unpredictable and due to perturbed initial conditions:

$$X_{jk} = \mu + \beta_j + \varepsilon_{jk} \quad (1)$$

2-way ANOVA can be applied to a set of several model ensemble experiments with similar or identical external forcing. In our case this super-ensemble contains only 2 different models (ECHAM3/LSG and HADCM2). The fact that the climate-change scenarios slightly differ from each other, SA90 (Houghton et al. 1990) versus IS92A (Houghton et al. 1992), will be discussed in Section 3.4. For the 2-way ANOVA, again a linear model for the data X_{ijk} of run k in ensemble i at time j is set up (von Storch & Zwiers 1999):

$$X_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{jk} \varepsilon_{jk} \quad (2)$$

In addition to the overall mean μ , the common forcing signal β_j and the residual ε_{ijk} as mentioned above, the so-called block effect α_i and an interaction coefficient γ_{ij} are introduced. The term α_i represents the systematic difference between the climate models relative to the overall mean, while γ_{ij} describes the variations of the temporal means of each individual run around the overall mean after removing α_i and β_{ij} . This component can either be interpreted as how differently the climate models react to an identical external forcing or as the impact of slightly different forcings.

The statistical significance of each of the components α , β , γ is assessed by hypothesis testing. For the 1-way ANOVA one assumes the random effects version of Eq. (1) with statistically independent variables β_j and ε_{jk} . Then the null hypothesis to be tested is

$$H_{01W} : \text{var}(\beta_j) = 0 \quad (3)$$

In the 2-way ANOVA case we will consider a combination of fixed (the model differences) and random effects (the external forcing and the interaction with the model, independent of the errors ε) in which case the null hypotheses to be tested are

$$\begin{aligned} H_{02W} : \alpha_1 &= \alpha_2 = 0 \\ &: \text{var}(\beta_j) = 0 \\ &: \text{var}(\gamma_j) = 0 \end{aligned} \quad (4)$$

The tests are performed with the help of the partition of the total sum-of-squares SS_{tot} into the sum-of-squares connected to the components of the 2 linear models (Eqs. 1 and 2).

$$\begin{aligned} \text{1-way ANOVA} : SS_{\text{tot}} &= SS_{\beta} + SS_{\varepsilon} \\ \text{2-way ANOVA} : SS_{\text{tot}} &= SS_{\alpha} + SS_{\beta} + SS_{\gamma} + SS_{\varepsilon} \end{aligned} \quad (5)$$

The calculation of the different sums-of-squares is described in von Storch & Zwiers (1999). The tests are performed using the fact that the sums-of-squares

under the appropriate null hypotheses are χ^2 distributed random variables and that ratios of sum-of-squares are Fisher F -distributed, both with appropriately chosen degrees of freedom. By comparing the ratios with the critical values from a Fisher F table either the null hypothesis in Eq. (3) or the 3 different null hypotheses in Eq. (4) can be accepted or rejected at a given level of significance.

The ratios of the individual sums-of-squares to the total are named in the following as the variance accounted for, for example, for the external forcing effect β_j :

$$\text{ex var}_{\beta} = \frac{SS_{\beta}}{SS_{\text{tot}}} \cdot 100 \quad (6)$$

In the present study, we refer to annual means of T2M and annual sums of PRE, leaving aside the seasonal response to the GHG forcing, because special emphasis is given to the longer time scales from decadal to centennial. The 1-way ANOVA will be carried out in 3 different contexts, applying it to:

- the full time series from 1880 to 2049 at every grid point, giving insight into the spatial distribution of the variance accounted for by the various components
- selected area averages in a moving time window with a width of either 30 or 60 yr to characterize the transient temporal variability
- selected area averages for low-pass filtered data with variable filter length to characterize the relevant time scales of the variability components ('spectral approach')

Note that in every case care has been taken to estimate the number of independent time samples with the help of the autocorrelation function (von Storch & Zwiers 1999).

3. RESULTS

3.1. Spatial approach

In this section the 1-way ANOVA is based on the whole integration period between 1880 and 2049 ($J = 170$ without the autocorrelation adjustment), applied to every gridpoint of the regular grid. Fig. 1a shows the GHG-induced variance for T2M in percent and the level of statistical significance. Between 1880 and 2049 the GHG forcing partly accounts for about 90% of the ECHAM3/LSG near-surface temperature variance in the low latitudes. In the extra-tropics, variance accounted for is generally in the range of 50 to 70%. However, there are some regions where internal variability clearly dominates with values larger than 80%. These places are in the North Atlantic region

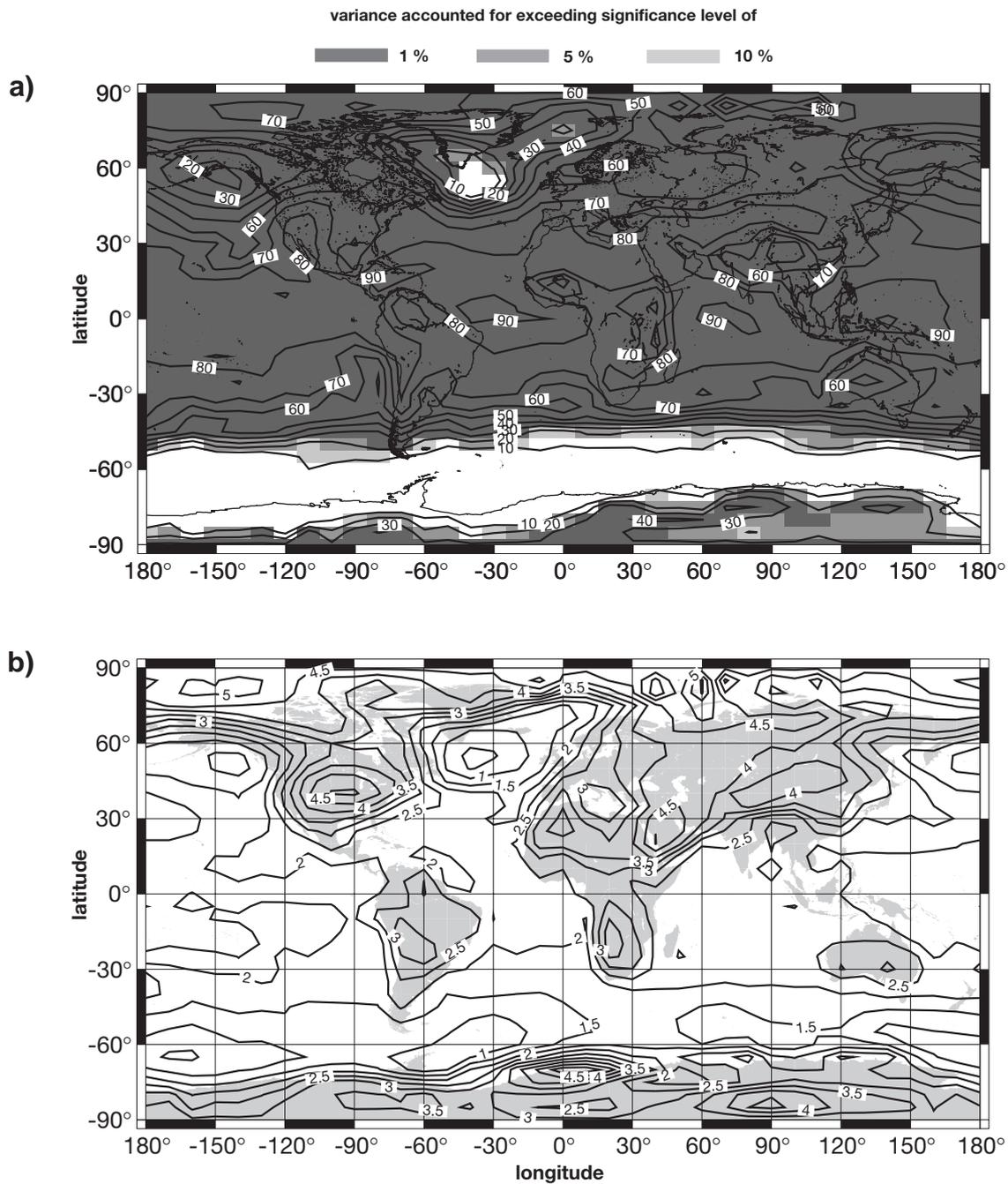


Fig. 1. (a) 1-way ANOVA of annual means of 2 m temperature at every grid point over the whole model period (1880–2049) for 4 ensemble members with CO₂-only forcing (NEIN ensemble): percent variance accounted for by external GHG forcing (contour lines) and significance of GHG influence (grey shading). (b) Trend pattern of annual means of 2 m temperature over the whole model period: total temperature change (contour lines)

south of Iceland, the Aleutian region, and the southern hemisphere mid-latitudes. Those regions are characterized by intense cyclone activity, enhanced natural variability and a strong atmosphere-ocean interaction, all masking the external GHG influence. As a consequence, such regions are largely unfavourable for

regionally based detection studies. Over large areas of the globe north of 40°S the null hypothesis of a zero external signal has to be rejected at a 1% significance level.

Fig. 1b indicates the temperature change of the ensemble mean over the whole model period based on

the linear trend. Since the ensemble mean estimates the external effect β_j , this pattern is a good measure for the GHG signal. It shows the well-known structure (Houghton et al. 2001) of a 3 to 4 K temperature change over the continents and a 2 K increase over wide oceanic areas. The weakest temperature increase, less than 1 K, occurs over the North Atlantic and over the Southern Ocean. These changes can not be distinguished from the internal variability.

Nevertheless, with respect to simulated internal variability T2M is an appropriate detection variable for GHG-induced climate change. This is particularly true for signal analyses based on global fingerprints, since simulated T2M variability in the ECHAM3/LSG model is dominated by GHG radiative forcing, except for some specific regions with intense natural variability. However, in real climate this GHG-induced fingerprint might compete with fingerprints from natural external factors.

In Fig. 2 the same methodical approach is shown for annual sums of PRE for the full simulation period, 1880–2049. It is quite obvious that for the major part of the globe the GHG influence is not statistically significant (Fig. 2a). There are some isolated regions where external variance reaches 10 to 20 % of total PRE variance, particularly over northern Europe, the Mediterranean, India, the Southern Ocean (up to 30 %), central South America, and parts of the Arctic. This means that between 1880 and 2049 PRE is almost everywhere exclusively influenced by mechanisms of natural variability. Even in the regions with statistically significant GHG impact, internal noise accounts for nearly 80 % of total PRE variance. As a consequence, it will be hard to detect this weak GHG signal (Fig. 2b) in observations. This also holds for regional climate change studies or the comparison of observations with 1 individual scenario run.

Fig. 2b again presents the trend pattern of long-term PRE change. Compared with T2M, this pattern is more differentiated, with positive and negative tendencies. The strongest PRE increase occurs in the inner tropics, especially in the monsoon regions of India, Indonesia, and Western Africa, where annual sums are 100 to 300 mm higher than at the beginning of industrialisation. Over the mid-latitudes and polar caps PRE is rising by 50 to 100 mm. In contrast, drier conditions are predicted for Central North America and the subtropical latitudes, particularly at the oceanic locations of subtropical highs and over the Mediterranean. The annual sum PRE decrease amounts to 50 to 150 mm in these areas, which are characterized by naturally arid conditions anyway. Relative to the total amount of annual PRE, the decrease in the subtropics as well as rising PRE in the mid- and high latitudes are more important than the strong increase in the tropics, where rainfall amounts to annually 2000 mm and more.

Of course, such PRE changes undoubtedly would have an essential impact on human activity and culture with regard to agriculture and natural resources. Particularly, this holds for the subtropical dry regions, e.g. the Sahel, as well as the Central Asian wheat belt. But on the other hand, it is hardly possible to say that the GHG forcing causes the projected PRE changes, even despite the fact that the ECHAM3/LSG model tends to overestimate the signal-to-noise ratio. Since observational data contains even more competing influences, natural and man-made, the detection problem for PRE is still aggravated.

3.2. Temporal approach

In order to gain an insight into the onset and the sensitivity of the T2M and PRE climate change signals to the time series' length, a second approach using 1-way ANOVA was applied. It is based on multi-decadal periods, continuously running through the whole integration period, 1880–2049. As the PRE signal is only evident over some selected regions, the 1-way ANOVA now refers to such regional area means. We have chosen northern Europe as an extra-tropical region, India as a monsoon region, Central South America as the inner tropics, and the Mediterranean as representative of the subtropical latitudes. Each area includes 28 grid points.

In Fig. 3 the 1-way ANOVA is based on the data within running 60 yr periods of T2M. This period corresponds to the expected characteristic multi-decadal time scale of the climate's response to GHG forcing (Paeth et al. 1999). The results of the 1-way ANOVA are attributed to the central year of the period in question. For northern Europe, the solid black line indicates the GHG-induced contribution, SS_{β} , to the total variability, SS_{tot} . It becomes statistically significant for the model period 1940–2000 (centred on 1970) onward at a significance level of 5 % and for the 60 yr period 1960–2020 (centred on 1990) at 1 %. Until the end of the integration period in 2049, the GHG scenario is the predominant influence on the future T2M fluctuation, amounting to 65 % of the total variance within the 60 yr windows. As to be expected, the corresponding annual mean T2M time series of the ensemble mean, representing the common GHG signal, shows a positive trend from 1980 onward. The variations around this ensemble mean are small within the ensemble (grey shading). Thus, we can define a statistically significant GHG-induced climate change signal for simulated T2M over northern Europe using 60 yr time windows starting with 1940–2000. Similar result are found for the other regions: India, Central South America, and the Mediterranean. These regions even reveal a clearer signal, with 80 to 90 % of the GHG-induced

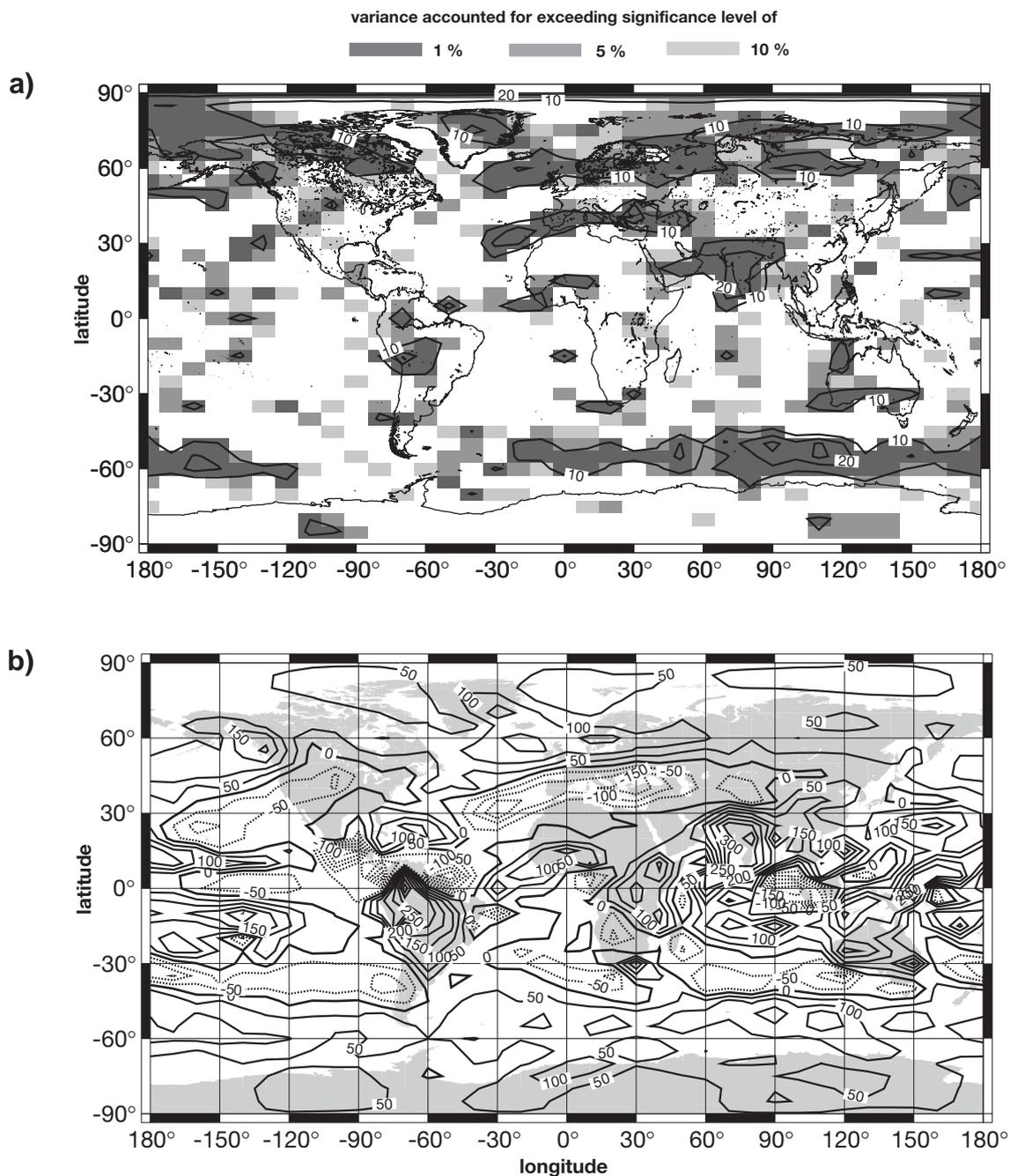


Fig. 2. (a) 1-way ANOVA of annual sums of precipitation at every grid point over the whole model period (1880–2049): percent variance accounted for by external GHG forcing (contour lines) and significance of GHG influence (grey shading). (b) Trend pattern of annual sums of precipitation over the whole model period: total precipitation change (contour lines)

variance. In addition, the T2M signal occurs 20 yr earlier than in the mid-latitudes.

Fig. 4 shows the results for T2M when we apply the 1-way ANOVA in running 30 yr windows. In northern Europe, the previously clear GHG impact is considerably reduced to 40% and less on this shorter time scale. This is probably due to the fact that in the mid-

latitudes natural variability has a strong interdecadal component which overlies the GHG signal on the 30 yr scale. For example, Paeth et al. (1999) could show that the 30 yr fluctuations of the North Atlantic Oscillation are mainly due to natural variability, whereas longer periods reveal a significant GHG impact. In the tropical and subtropical regions, the T2M signal is still per-

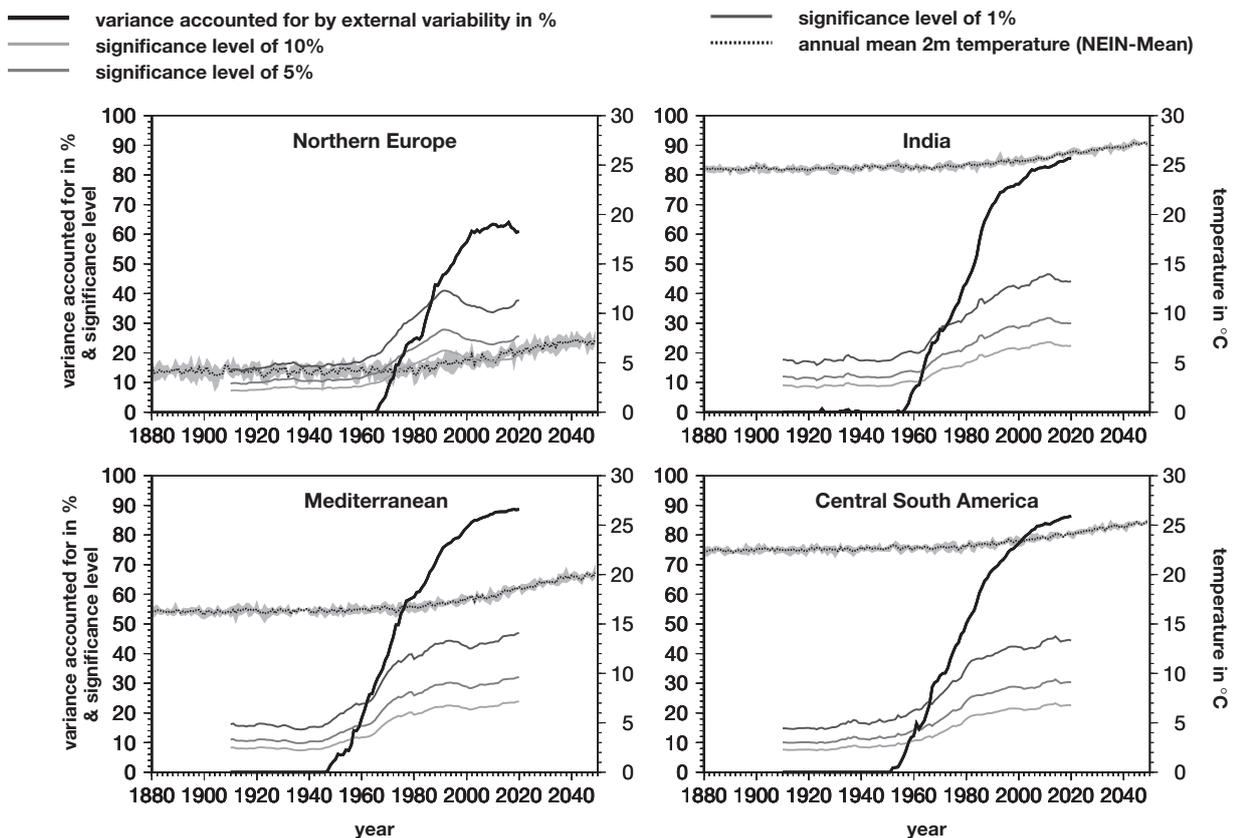


Fig. 3. 1-way ANOVA of annual means of 2 m temperature in selected regions over continuous 60 yr periods: externally accounted for variance (solid black line), corresponding levels of significance (3 solid grey-scale lines), ensemble mean of 2 m temperature (dotted line), and minimum/maximum range of 2 m temperature within the ensemble (grey shading). The x-axis legend refers to the middle year of each period within the integration period 1880–2049

sistent and statistically significant at a 1% level, with the ratio of external variance to total variance decreasing to 70%. The signal onset is shifted into the period 1955–1985. In summary, we can conclude that a stable and statistically meaningful climate change signal relative to the model's internal variability can be extracted from the near-surface temperature of ECHAM3/LSG scenario runs provided that signals are derived from time periods of 30 to 60 yr centered on 1970 and later.

In Fig. 5 the same analysis is shown for PRE. Note that the right-hand ordinate for the annual sums of precipitation changes between the 4 sub-plots, because annual sums of PRE are quite different from region to region. Considering 60 yr windows over northern Europe, the GHG influence is quite weak, reaching about 15% (rejection of the null hypothesis at the 5% level) for the period 1960–2020 and 25% for the end of the simulation (rejection of the null hypothesis at the 1% level). This means that even for the simulation period 1990–2049 about $\frac{3}{4}$ of the PRE variability is still naturally induced. This is also apparent from the

strong variations around the ensemble mean (grey shading). As a consequence, an increase in PRE after 1980 in a single realisation (and therefore probably also in nature) is not necessarily due to the imposed GHG forcing. The same conclusion can be drawn for the Mediterranean, where decreasing rainfall is only marginally (20%) attributed to rising GHG without hardly reaching the significance level of 1%. In the Indian monsoon area, where PRE enhancement is strongest, with 150 to 200 mm over the 60 yr period 1990–2049, there appears to be a clearer PRE signal, reaching nearly 40% of total variability from the period 1960–2020 onward. A weaker signal exists in central South America, although the positive PRE trend is also remarkable. In Fig. 6, the variability of PRE in 30 yr periods is addressed. Compared to Fig. 5, the weak PRE signals in the 60 yr period almost disappear in all 4 regions. Even in India, external variance is reduced to non-significant values. Externally forced model changes in PRE over 30 yr contribute only a minor part to the total variability. Thus a statistically meaningful climate change signal for PRE in the ECHAM3/LSG

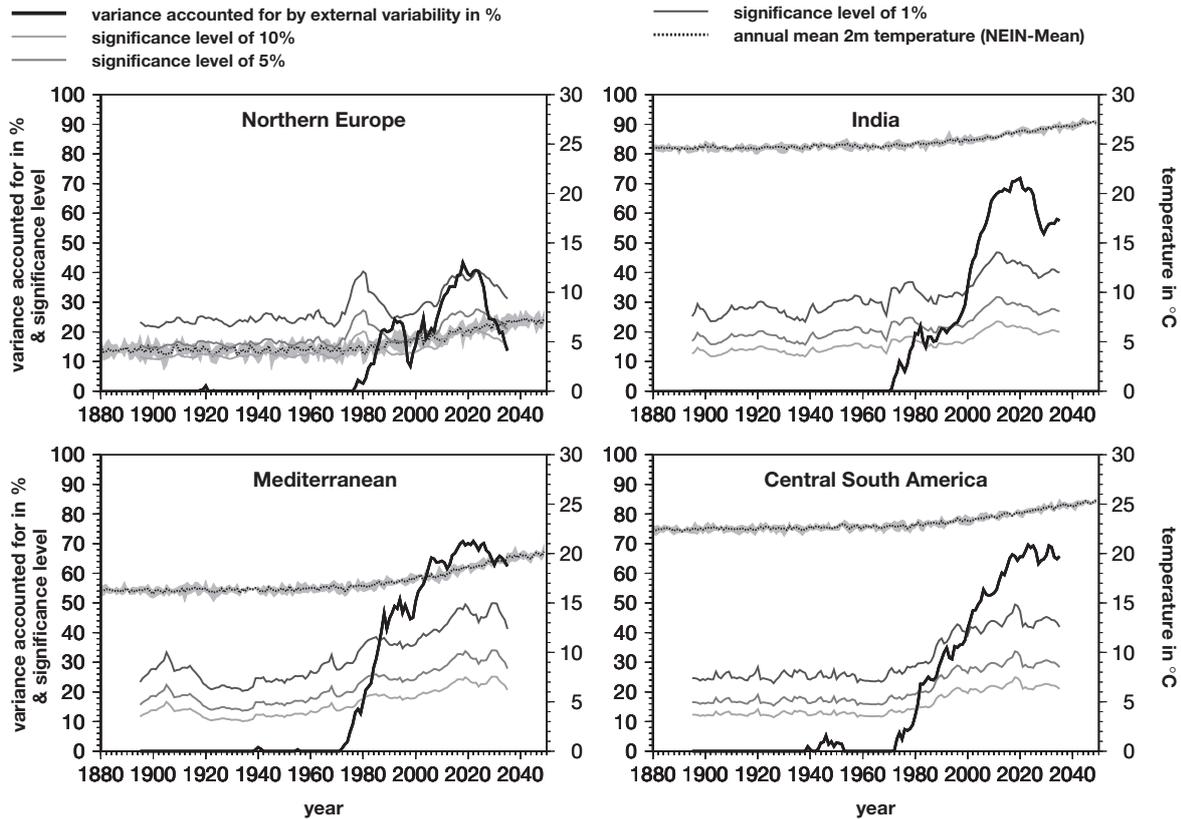


Fig. 4. 1-way ANOVA of annual means of 2 m temperature in selected regions over continuous 30 yr periods: externally accounted for variance (solid black line), corresponding levels of significance (3 solid grey-scale lines), ensemble mean of 2 m temperature (dotted line), and minimum/maximum range of 2 m temperature within the ensemble (grey shading)

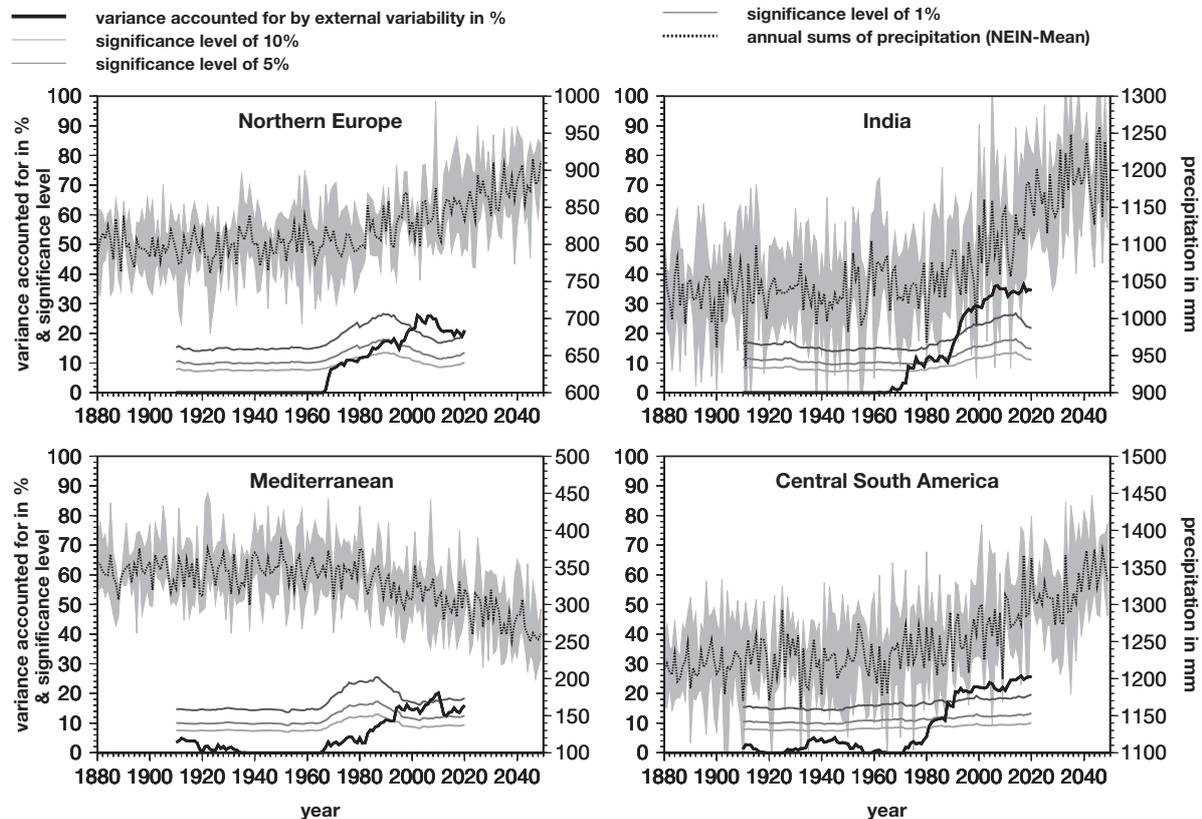


Fig. 5. 1-way ANOVA of annual sums of precipitation in selected regions over continuous 60 yr periods: externally accounted for variance (solid black line), corresponding levels of significance (3 solid grey-scale lines), ensemble mean of precipitation (dotted line), and minimum/maximum range of precipitation within the ensemble (grey shading)

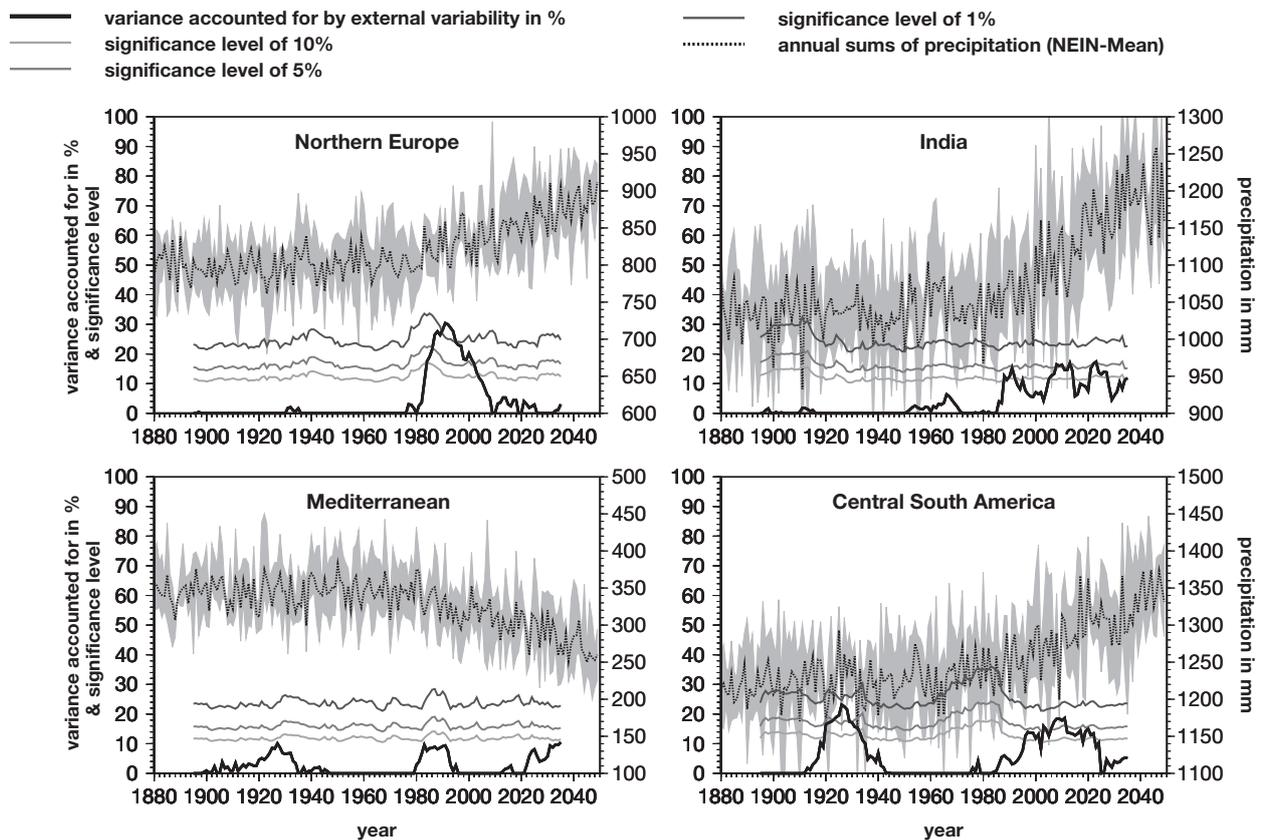


Fig. 6. 1-way ANOVA of annual sums of precipitation in selected regions over continuous 30 yr periods: externally accounted for variance (solid black line), corresponding levels of significance (3 solid grey-scale lines), ensemble mean of precipitation (dotted line), and minimum/maximum range of precipitation within the ensemble (grey shading)

cannot be derived on the basis of 30 yr windows, while it is possible on a 60 yr basis, although the explained variance is weak. Simulated PRE reveals a statistically significant GHG-induced signal on long time scales, but even in the regions with strongest PRE response, variability is predominantly due to noise.

3.3. Spectral approach

Another interesting question, related to climate change research, addresses the part of the spectrum of climate variability which is most affected by the GHG forcing. One might consider that the above conclusions, especially for PRE, result from a strong interannual variability while internal unpredictable variability on longer time scales might be much weaker. Therefore, a third variant of the 1-way ANOVA was applied to the T2M and PRE time series. Prior to the 1-way ANOVA all time series were low-pass filtered, using running means between 1 and 40 yr. The 1 yr 'filter' does not change the annual values at all, while a

k -year filter ($k \in [2, 40]$) dampens the variability on time scales shorter than k yr before the 1-way ANOVA is carried out. This method is different from the temporal approach, where annual means over different periods were examined without removing any fluctuations. Here, emphasis is not on the time evolution of the signal and the dependence on the time series' length, but on the time scales of T2M and PRE variability which are best predictable from the GHG scenario.

Fig. 7 shows externally accounted for T2M variance and corresponding levels of significance as a function of the filter length, k , for the whole integration period, 1880–2049. At first sight, it becomes clear that interannual to multi-decadal T2M fluctuations are almost exclusively determined by the external GHG forcing. In northern Europe the strongest signal occurs on the multi-decadal time scale, accounting for 95% of total variability. With regard to the year-to-year changes, where extra-tropical natural variability has a strong influence, the curve turns slightly downward, to 70%. This feature is common to all selected regions. However, in the tropics, the signal peaks are rather shifted

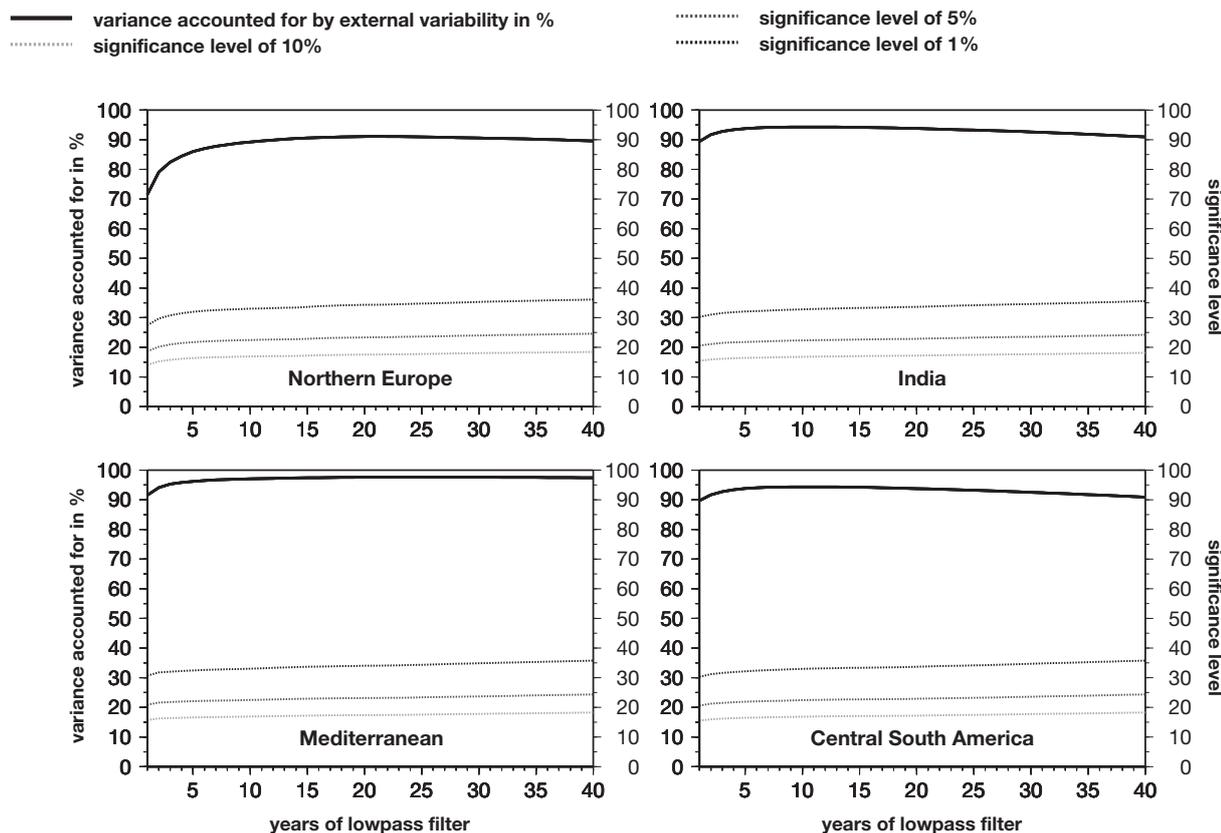


Fig. 7. Spectral 1-way ANOVA of annual means of 2 m temperature in selected regions for the period 1880–2049, based on running-mean low-pass filters between 1 and 40 yr: externally accounted for variance (solid black line) and corresponding levels of significance (3 grey lines)

to the decadal time scale, although all values, 1 to 40 yr, are generally found on a high level of around 95%. Thus, model T2M is a reliable detection variable for temperature fluctuations on interannual and longer time scales.

In Fig. 8 the same is illustrated for PRE. Compared with T2M, the signal is again less striking. It is the interannual variability with time scales less than 5 to 10 yr which is responsible for the strong drop in GHG forcing related variability relative to the total variability. Decadal and interdecadal variations of PRE can be separated reasonably well into GHG-induced variability and internal variability. Northern Europe and the Mediterranean reveal a dominating external influence on the decadal to multi-decadal time scale, amounting to 80%. Over India and South America the impact even rises to 85%.

Nonetheless, it is a promising result for future detection studies that decadal PRE fluctuations may have some predictability potential with respect to rising GHG concentrations, although it has to be called to mind that the ECHAM3/LSG generally overestimates the signal-to-noise ratio. It has to be assumed that similar constraints hold for observations. Present time

series of rainfall observations are mostly too short to handle the decadal time scale in a statistically proper way. Therefore, the use of PRE in climate change analysis might come in some decades, when observational time series will have been continued.

The analysis of the HADCM2 data with the 1-way ANOVA brought up very similar results with respect to the strength of the GHG-induced signal as the ECHAM3/LSG data. Therefore, those 1-way ANOVA results are not repeated here.

3.4. Super-ensemble approach

As described in Section 2.2 the uncertainty factor of imperfect climate model projections can be addressed by using 2-way ANOVA based on the ECHAM3/LSG and HADCM2 ensembles which are driven by increasing GHG concentrations into the 21st century. As outlined above, this method compares the intermodel variability with respect to the common GHG signal and the internal variance.

Fig. 9 shows the super-ensemble components of total T2M variability in a spatial approach for the overlap-

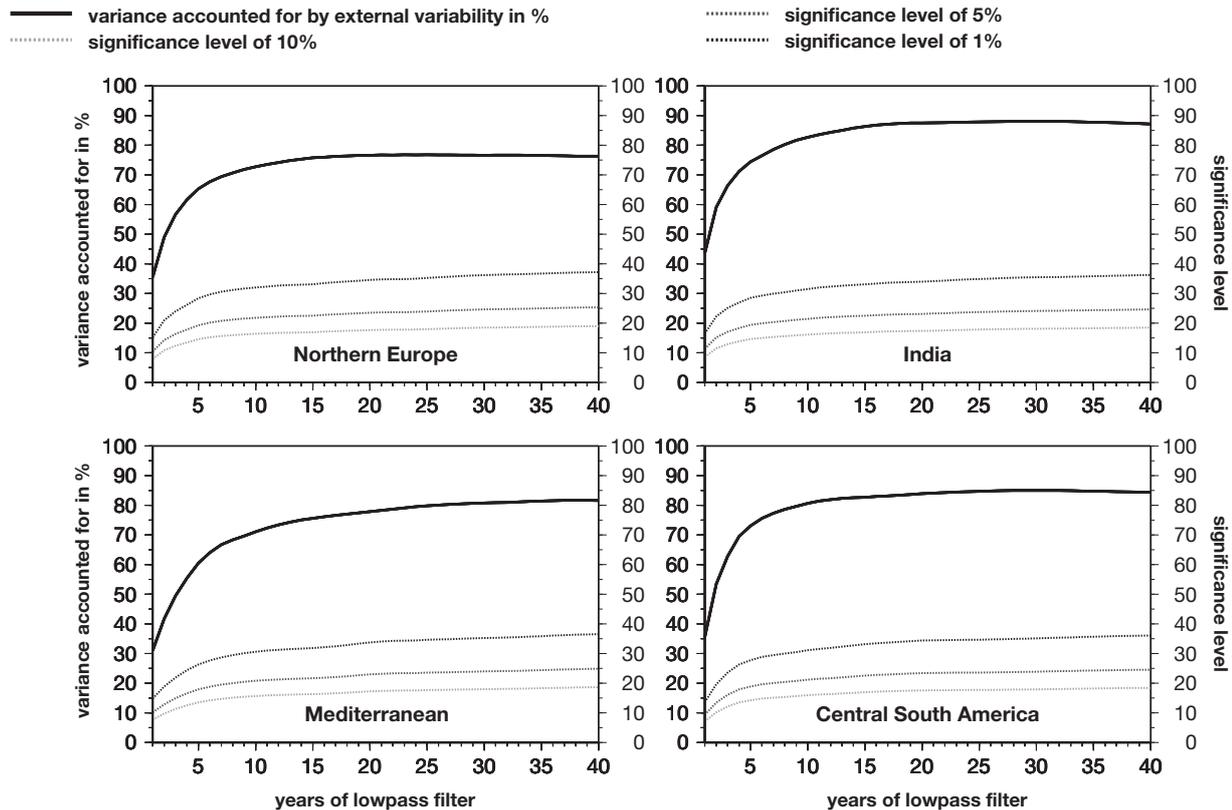


Fig. 8. Spectral 1-way ANOVA of annual sums of precipitation in selected regions for the period 1880–2049, based on running-mean lowpass filters between 1 and 40 yr: externally accounted for variance (solid black line) and corresponding levels of significance (3 grey lines)

ping period of both models, which is 1880–2049. Each panel represents the contribution by 1 of the 4 terms in Eq. (2) in percent – the block effect, the overall treatment effect, the interaction coefficient and the residual. Dark shading indicates a strong influence.

Although relative internal variability (SS_e/SS_{tot} , bottom right) still is minor, except over regions with intense cyclone frequency, the impact of increasing GHG on temperature (top right) is much less pronounced than in Fig. 2, where it is based on the 1-model ensemble. Over wide areas of the tropics and subtropics and the Northern Hemisphere mid-latitude continents, the GHG signal is consistent between the 2 climate models. Particularly, over the tropical Atlantic and Indonesia the joint GHG impact amounts to more than 90%. An interesting fact is that the global structure of the joint GHG signal is very similar to the data availability of near-surface temperature (e.g. Jones et al. 1999). In view of the large internal and intermodel variability, for example, the Southern Hemisphere simulations cannot be reliably interpreted. Therefore missing observations on the Southern Hemisphere do not harm a detection study.

The strong signal deterioration occurs over mountainous regions, the sea-ice margins, the Southern

Ocean and Antarctica. Especially in those regions, the ECHAM3/LSG and HADCM2 lead to systematically different results (top left), predicting fundamentally different responses of T2M with regard to the radiative forcing. Actually, sea ice and mountain effects are subject to large uncertainties in climate modelling. Therefore, the upper left panel is a quantitative reflection of the general 'Achilles' heel' of state-of-the-art climate models (von Storch & Bray 1999) with respect to the sensitivity of the climate system to, e.g. resolution-based, differences in orography and sea ice margins. Note that this figure only refers to 2 climate models. Uncertainty could be much stronger when the study is extended to a greater number of climate models. Lastly, temperature projection into the 21st century is barely sensitive to the different GHG forcings, SA90 versus IS92A, accounting for less than 10% of total variance (bottom left).

With regard to PRE (Fig. 10), the weak GHG signals over some regions in Fig. 2 completely disappear due to pronounced intermodel variations. Total rainfall variability is split up into model uncertainty and unpredictable climate noise when unfiltered interannual variability is included (cf. Giorgi & Francisco 2000). With the knowledge of the previous results, particu-

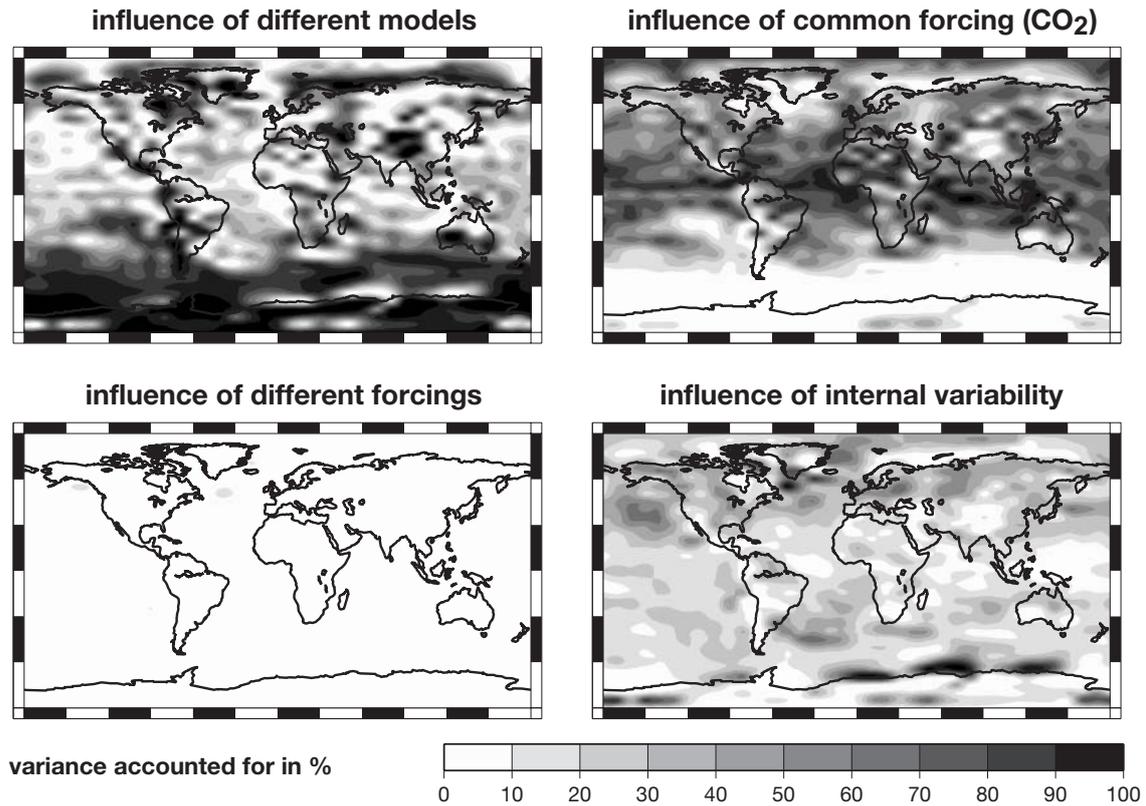


Fig. 9. 2-way ANOVA of annual means of 2 m temperature over the globe, based on the overlapping period 1880–2049 of ECHAM3/LSG and HADCM2: contribution in % to total variance by different model parametrizations, different GHG forcings, common GHG forcing, and internal variability

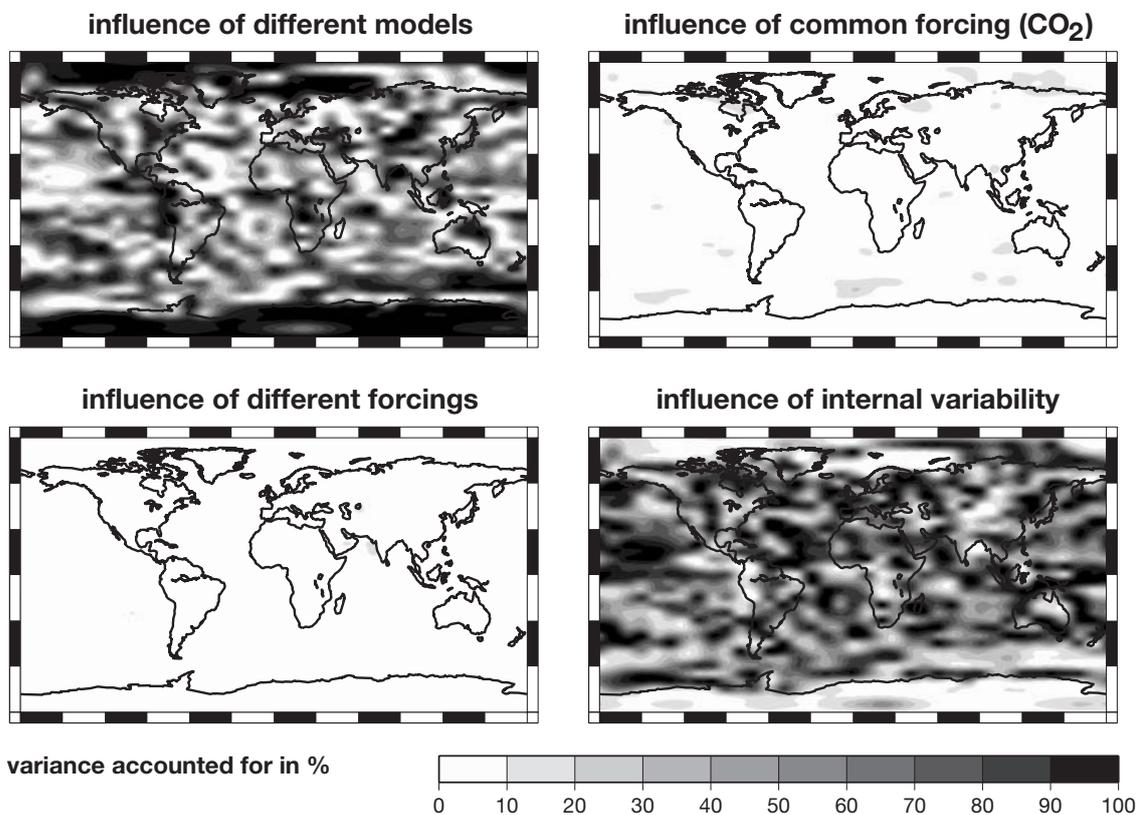


Fig. 10. 2-way ANOVA of annual sums of precipitation over the globe, based on the overlapping period 1880–2049 of ECHAM3/LSG and HADCM2: contribution in % to total variability by different model parametrizations, different GHG forcings, common GHG forcing, and internal variability

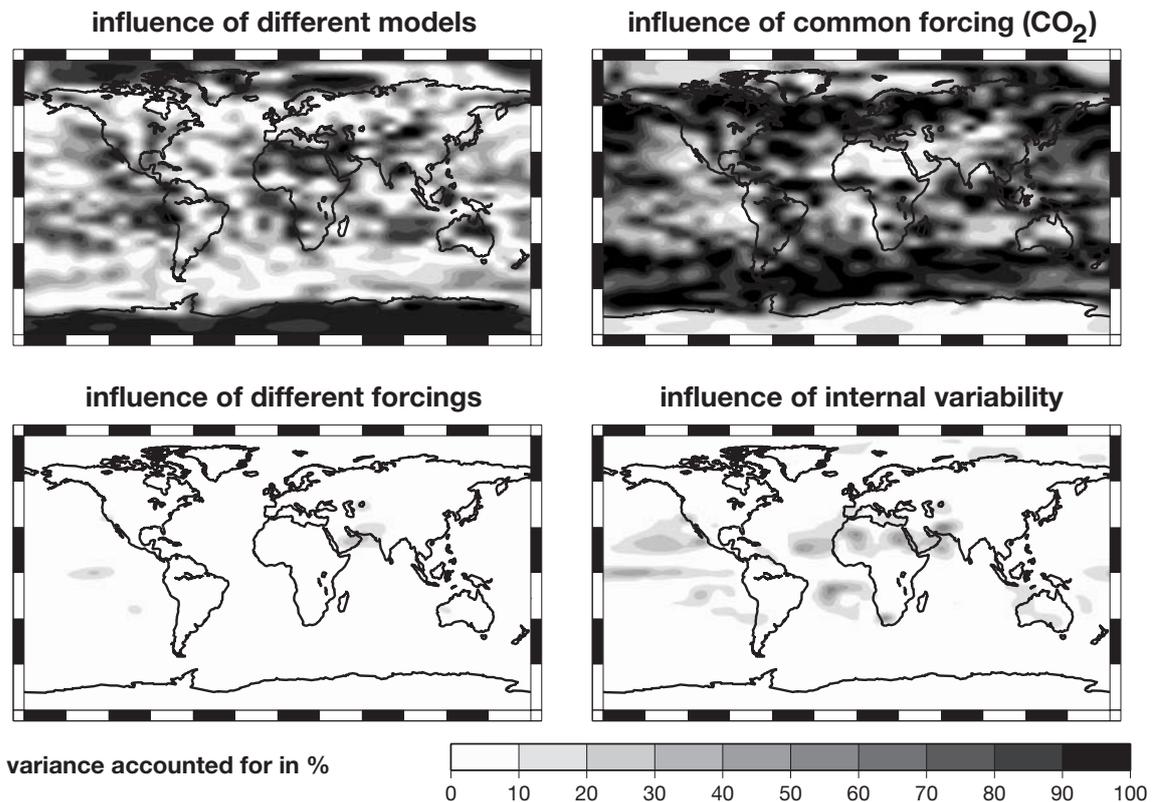


Fig. 11. 2-way ANOVA of 10 yr filtered annual sums of precipitation over the globe, based on the overlapping period 1880–2049 of ECHAM3/LSG and HADCM2: contribution in % to total variance by different model parametrizations, different GHG forcings, common GHG forcing, and internal variability

larly the spectral 1-way ANOVA (Fig. 8), the PRE signal should emerge from behind internal variability when decadal-mean rainfall anomalies are considered. This is demonstrated in Fig. 11, where 2-way ANOVA is based on 10 yr running mean low-pass-filtered PRE time series. Indeed, internal variability (bottom right), one of the main contributors to total rainfall variance in Fig. 10, is almost completely replaced by the common GHG forcing (top right), which accounts for the major part of rainfall fluctuations in the inner tropics and mid-latitudes. However, the model uncertainty still persists in the high latitudes, mountainous areas and, additionally, in the subtropics. Although this is a promising result in terms of climate change detection, one has to keep in mind that nowadays observational time series of PRE are still too short to address the decadal time scale properly. Note that due to the filtering the degrees of freedom are largely reduced by a factor of 10 at least.

However, the strong and almost global temperature signal in Fig. 1 is also partially blurred out due to systematic differences between the climate models. Optimal conditions for detecting GHG-induced warming are revealed over the inner-tropical oceans, where the

GHG impact seems to be present with all considered climate change simulations.

4. SUMMARY AND CONCLUSIONS

Based on the ECHAM3 ensemble, the increasing GHGs have a great impact on simulated temperature variability between 1880 and 2049, implying (the well-known) global warming pattern with highest amplitude over the Northern Hemisphere continents. When this impact is compared statistically, using 1-way ANOVA, with the unpredictable internal variability, it is clearly significant, reaching 90% of total T2M variability in the inner tropics. No effect is found in regions with strong natural variability, for example, due to frequent cyclone activity in the Northern and Southern Hemisphere mid-latitudes. PRE variability within the same period is weakly influenced by the GHG forcing. Actually, there are a few regions with a significant impact compared to the internal model variability, but this amounts to 30% at best. The PRE trend pattern reveals increasing annual sums of rainfall in the inner tropics, the monsoon areas, and the mid- and high latitudes, whereas in the sub-

tropical arid regions drier conditions are predicted. This pattern seems to represent a climate change fingerprint, but attribution to anthropogenic GHG emissions fails due to the intense natural variability of PRE.

Applying the 1-way ANOVA within 60 yr time windows moving through the integration period, the signal onset becomes visible, here presented for selected regions with the strongest PRE response. For T2M and PRE the signal first emerges within the period 1940–2000. Although slightly weaker in northern Europe, the T2M signal is striking everywhere, accounted for 90% of total variance. Precipitation is again barely induced by the GHG forcing, with at most 40% of total variance in India.

Based on shorter periods of 30 yr which correspond to the length of many observational time series, the T2M signal is largely persistent, while for precipitation it completely disappears, indicating that natural variability clearly dominates PRE variability within 30 yr time slices.

Applying the 1-way ANOVA to low-pass-filtered values of PRE and T2M reveals high predictability from GHG on all time scales of T2M variability, from the interannual to the multidecadal scale. Increasing GHG concentrations have some clear influence on decadal to multidecadal PRE fluctuations. The strong drop in the relative amount of GHG-related variability with respect to total variability can be attributed to the interannual PRE variations on time scales less than 5 yr.

The model intercomparison reveals some remarkable systematic differences between the ECHAM3/LSG and HADCM2 coupled climate models, even with respect to near-surface temperature, which generally is supposed to be only mildly noisy. This is particularly true over higher mountain regions, sea ice margins and the Southern Hemisphere high latitudes. However, in the inner tropics the GHG warming signal is still persistent, indicating the importance of radiative forcing in the low latitudes. Rainfall variability is entirely attributed to intermodel variations and internal noise, whereas the GHG impact is completely blurred out.

Thus, we can conclude that near-surface air temperature is a very favourable detection variable for defining climate change signals on all time scales from annual to decadal and for rather short time series from 30 yr upward. This is in total agreement with the theoretical idea of a GHG-induced modification of the radiation balance. However, this finding is mostly confined to the low and mid-latitudes. Polar, mountainous and cyclo-genetic areas should be masked out from global fingerprints. On the other hand, precipitation represents an unsuitable variable for defining anthropogenic climate change, since external GHG forcing is blurred out by strong natural variability. This is especially true for time series of 30 yr and shorter. However,

some good predictability can be found when all PRE variations on time scales less than 10 yr are removed. Unfortunately, observational data mostly do not cover sufficiently long periods to handle the decadal time scale of PRE changes. Therefore, it is likely that in the future PRE will play a more important role in climate change detection than nowadays.

Acknowledgements. This work was supported by the German Ministry of Research and Education under grant 07/VKV01. We thank Drs Reinhard Voss and Ulrich Cubasch from the DKRZ for providing the data sets of the ECHAM3/LSG climate model in Hamburg. Particularly, we would like to acknowledge the extensive support by Dr. Reinhard Voss during the realization of the NEIN ensemble.

LITERATURE CITED

- Barnett TP, Hegerl GC, Santer B, Taylor K (1998) The potential effect of GCM uncertainties and internal atmospheric variability on anthropogenic signal detection. *J Clim* 11: 659–675
- Barnett TP, Hegerl GC, Knutson T, Tett SFB (2000) Uncertainty levels in predicted patterns of anthropogenic climate change. *J Geophys Res* 105(D12):15525–15542
- Barnett TP, Pierce DW, Schnur R (2001) Detection of anthropogenic climate change in the world's oceans. *Science* 292:270–274
- Brankovic C, Palmer TN, Molteni F, Tibaldi S, Cubasch U (1990) Extended range predictions with the ECMWF models: time lagged ensemble forecasting. *Q J R Meteorol Soc* 116:835–866
- Chase TN, Pielke RA, Kittel TGF, Nemani RR, Running SW (2000) Simulated impacts of historical land cover changes on global climate in northern winter. *Clim Dyn* 16:93–105
- Cubasch U, Santer BD, Hellbach A, Hegerl G, Höck H, Maier-Reimer E, Mikolajewicz U, Stössel A, Voss R (1994) Monte Carlo climate change forecasts with a global coupled ocean-atmosphere model. *Clim Dyn* 10:1–19
- Flohn H, Kapala A, Knoche HR, Mächel H (1992) Water vapour as an amplifier of the greenhouse effect: new aspects. *Meteorol Z NF* 1:122–138
- Giorgi F, Francisco R (2000) Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Clim Dyn* 16: 169–182
- Hasselmann K (1979) On the signal-to-noise problem in atmospheric response studies. Shaw DB (coord) *Meteorology over the tropical oceans*. Royal Meteorological Society, Reading, p 251–259
- Hasselmann K (1993) Optimal fingerprints for the detection of time-dependent climate change. *J Clim* 6:1957–1971
- Hegerl GC, North GR (1997) Comparison of statistically optimal approaches to detecting anthropogenic climate change. *J Clim* 10:1125–1133
- Hegerl GC, von Storch H, Hasselmann K, Santer BD, Cubasch U, Jones PD (1996) Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *J Clim* 9: 2281–2306
- Hegerl GC, Hasselmann K, Cubasch U, Mitchell JFB, Roeckner E, Voss R, Waszkewitz J (1997) Multi-fingerprint detection and attribution analysis of greenhouse gas,

- greenhouse gas-plus-aerosol and solar forced climate change. *Clim Dyn* 13:613–634
- Hegerl GC, Stott PA, Allen MR, Mitchell JFB, Tett SFB, Cubasch U (2000) Optimal detection and attribution of climate change: sensitivity of results to climate model differences. *Clim Dyn* 16:737–754
- Hegerl GC, Jones PD, Barnett TP (2001) Effect of observational sampling error on the detection of anthropogenic climate change. *J Clim* 14 198–207
- Houghton JT, Jenkins GJ, Ephraums JJ (eds) (1990) Climate change. The IPCC Scientific Assessment. Cambridge University Press, Cambridge
- Houghton JT, Callandar BA, Varney SK (eds) (1992) Climate change 1992. The supplementary report to the IPCC Scientific Assessment. Cambridge University Press, Cambridge
- Houghton JT, Ding Y, Griggs DJ, Noguer M, Van der Linden PJ, Dai X, Maskell K, Johnson CA (eds) (2001) Climate change 2001. The scientific basis. Cambridge University Press, Cambridge
- Hulme M, Barrow EM, Arnell NW, Harrison PA, Johns TC, Downing TE (1999) Relative impacts of human-induced climate change and natural climate variability. *Nature* 397:688–691
- Hunt BG (1998) Natural climatic variability as an explanation for historical climatic fluctuations. *Clim Change* 38: 133–157
- Johns TC, Carnell RE, Crossley JF, Gregory JM, Mitchell JFB, Senior CA, Tett SFB, Wood RA (1997) The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup, and validation. *Clim Dyn* 13:103–134
- Jones PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its changes over the past 150 years. *Rev Geophys* 37:173–199
- Keen AB, Murphy JM (1997) Influence of natural variability and the cold start problem on the simulated transient response to increasing CO₂. *Clim Dyn* 13:847–864
- Kharin VV, Zwiers FW (2002) Climate predictions with multi-model ensembles. *J Clim* 15:793–799
- Levis S, Foley JA, Pollard D (2000) Large-scale vegetation feedbacks on a doubled CO₂ climate. *J Clim* 13:1313–1325
- Mahlman JD (1997) Uncertainties in projections of human-caused climate warming. *Science* 278:1416–1417
- North GR, Stevens MJ (1998) Detecting climate signals in the surface temperature record. *J Clim* 11:563–577
- Osborn TJ, Briffa KR, Tett SFB, Jones PD, Trigo RM (1999) Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Clim Dyn* 15:685–702
- Paeth H, Hense A (2001) Signal analysis of the atmospheric mean 500/1000 hPa temperature north of 55° N between 1949 and 1994. *Clim Dyn* 18:345–358
- Paeth H, Hense A, Glowienka-Hense R, Voss R, Cubasch U (1999) The North Atlantic Oscillation as an indicator for greenhouse-gas induced climate change. *Clim Dyn* 15: 953–960
- Paeth H, Hense A, Hagenbrock R (2002) Comments on 'Twentieth-century trends of Arctic precipitation from observational data and a climate model simulation'. *J Clim* 15:800–803
- Roeckner E, Arpe K, Bengtsson L, Brinkop S, Dümenil Land 9 others (1992) Simulation of the present-day climate with the ECHAM model: impact of model physics and resolution. Rep No. 93, Max-Planck-Institut für Meteorologie, Hamburg
- Santer BD, Taylor KE, Wigley TML, Johns TC and 9 others (1996) A search for human influence on the thermal structure of the atmosphere. *Nature* 382:39–46
- Scheffé H (1963) The analysis of variance. Wiley, New York
- Schneider T, Held IM (2001) Discriminants of twentieth-century changes in Earth surface temperature. *J Clim* 14: 249–254
- Ulbrich U, Christoph M (1999) A shift of the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing. *Clim Dyn* 15:551–559
- von Storch H (2001) Statistics: an indispensable tool in dynamical modelling. In: von Storch H, Flöser G (eds) Models in environmental modelling. GKSS School of Environmental Research, Geesthacht, and Springer-Verlag, Berlin, p 203–217
- von Storch H, Bray D (1999) Perspectives of climate scientists on global change. Conference Proceedings, Climate change policy in Germany and the United States, Berlin, June 15–18, 1997. Publ German Am Acad Council 7:33–48
- von Storch H, Zwiers FW (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge
- Voss R, Sausen R, Cubasch U (1998) Periodically synchronously coupled integrations with the atmosphere-ocean general circulation model ECHAM3/LSG. *Clim Dyn* 14:249–266
- Wibig J (1999) Precipitation in Europe in relation to circulation patterns at 500 hPa level. *Int J Climatol* 19:253–269

Editorial responsibility: Hans von Storch, Geesthacht, Germany

*Submitted: August 19, 2001; Accepted: May 21, 2002
Proofs received from author(s): September 4, 2002*