Evaluation of monthly mean data fields from the ECHAM4/OPYC3 control integration.

R.E. Benestad, I. Hanssen-Bauer, E.J. Førland, K. Iden,
& O.E. Tveito

4th. draft: DNMI, May 14, 1999

Contents

1 Introduction

Global general circulation models (GCMs) are used to make future scenarios for climate change. The climate models are still crude, partly because of low spatial resolution and partly because all climate processes are not yet understood. Because of these shortcomings it is important to know their strengths and weaknesses in order to interpret the model results.

It is crucial to appreciate the uncertainties involved in making the future climate scenarios. There are two main sources of uncertainties: *a)* in the prescribed greenhouse gas forcing and *b)* in the model (mis)representation of the climate processes. The uncertainty in the greenhouse gas emission stems from the fact that the emission of CO_2 is a function of economic activity and political decisions, and we cannot know what the economy will be like in the future nor what the politicians will dictate. One way to deduce the uncertainties in the climate predictions due to anthropogenic emission of greenhouse gases is therefore to estimate the sensitivity of the mean climate state with respect to different CO_2 concentrations.

The uncertainties associated with model discrepancies can be assessed through model-observation comparison. The focus of this report will primarily be on the large scale differences between observed and model statistics of present day climate integrations. This model evaluation will provide a basis for future regional downscaling employing dynamical and statistical models.

The global coupled climate model from *Max-Planck-Institute für Meteorologie* (MPI) which is called ECHAM4/OPYC3 will be used in the *RegClim*[1] project as a basis for statistical as well as dynamical downscaling of future climate scenarios for Norway. It is therefore of great interest to know how well this particular model describes the present climate in the vicinity of the Nordic countries, as it is assumed that any misrepresentation of present climatic features will be a result of model weaknesses which also may affect the simulation of future climate. Therefore, a model evaluation is carried out here by comparing the model control climate (CTL) and historical climate observations. In this respect, it is important to keep in mind that the CTL results are not entirely comparable with the observations of the 20th century, as the statistics of the historical observations do not necessarily correspond to those of the 1990s due to long term climatic variations. However, it will be assumed in this report that the primary climatic features in the early part of this century were similar to those of the present day climate despite a temperature trend, and hence that a good correspondence between the CTL results and the historical data is necessary for the model results to be

[1] Regional Climate Developement Under Global Warming

realistic.

The outline of this report is as follows. The first section will give an overview of the main features of the ECHAM4/OPYC3. This section is mainly technical and may be skipped by those not interested in the model technicalities. The observational data sets are described next, followed by the model evaluation where both mean fields and standard deviation maps from ECHAM4/OPYC3 control integration are compared with corresponding observations. The evaluated model diagnostics include surface temperature (T(2m)) sea level pressure (SLP), 500hPa geopotential height (Φ_{500}), 500hPa temperature (T_{500}) and sea surface temperature (SST). In section 6, the leading **e**mpirical **o**rthogonal **f**unctions (EOFs) from the control integration are compared with EOFs from the observational data sets ("observed EOFs"). In this section, an evaluation is carried out with respect to the spatial EOF patterns, the eigenvalues, and the spectral characteristics of the **p**rincipal **c**omponents (PCs, which indicate the temporal variations of the spatial patterns described by the EOFs). We will refer to the most prominent EOF (the EOF associated with highest variance) as the "leading EOF". Finally, time series from a selection of locations in the CTL results were evaluated against the best available observations.

2 Description of the MPI ECHAM4/OPYC3 coupled model integrations

2.1 Model details

The ECHAM (*Roeckner & Coauthors*, 1996) model is based on the European Centre for Medium Weather Forecast (ECMWF) IFS model, but has undergone several modifications. The acronym "ECHAM" stands for "EC", which refers to the origin of the model, i.e. ECMWF, and "HAM", for Hamburg (the Max-Planck institute) where the model was modified for climate studies. In particular, the parameterisation schemes have been optimised for climate simulations. The reference resolution of the atmospheric model is T42 (which approximately corresponds to $2.8° \times 2.8°$) with cycle 31 and 19 vertical layers.

"OPYC3" (*Oberhüber*, 1993) stands for Ocean and isopycnal coordinates, and the model is designed for studies of large scale ocean circulation. The ocean model includes a sea-ice model which decouples the ocean from extreme atmospheric winter conditions in the Arctic and facilitates a realistic treatment of the salinity budget by melting and freezing sea-ice.

The model features can be summarised as following:

- Model Diagnostics

ECHAM4 Prognostic variables: Vorticity, Divergence, Temperature, Log surface pressure, water vapour, and cloud water.

ECHAM4 SST and Ice from OPYC3 through coupling and from COLA/CAC AMIP SST and Sea-Ice data set.

OPYC3 The OPYC3 model is a hydrostatic isopycnal primitive equation model fully coupled to a surface mixed layer. A snow and sea-ice model with prognostic equations for ice depth and concentration is included. The model is designed for large scale processes, and the basic quantities conserved are: momentum, energy, mass, potential vorticity, salt and heat.

- Physical Parameterisation schemes

ECHAM4 Roughness length uses the Charnock formula over sea and constant values over sea-ice. Over land, the roughness length is a function of vegetation and orography.

ECHAM4 Vegetation is taken as the fraction of grid area covered by vegetation.

ECHAM4 Albedo at sea is a function of zenith angle. Over bare land, the albedo is based on satellite data, and for sea-ice the albedo is a function of temperature. In snow covered region, the albedo is a function of temperature and fractional forest.

ECHAM4 Physical radiation uses a two-stream approximation with 6 spectral intervals for the terrestrial part and 4 intervals for the solar part. The gaseous absorbers are H_2O, CO_2 and O_3, where CO_2 and O_3 as well as aerosols are prescribed. The optical depth is affected by the cloud cover and the emissivity is a function of the cloud water path and the radiation scheme includes a continuum absorption scheme. The cloud overlap is random apart for contiguous clouds where the overlap is maximum. The diurnal cycle is included and the radiation time step is 2 hours.

ECHAM4 The cloud scheme includes a cloud water transport equation, subgrid scale condensation and cloud formation with different thresholds for convective and stratiform clouds, temperature dependent partitioning of liquid/ice phase, rain formation by conversion of cloud drops, sedimentation of ice crystals, evaporation of cloud water, and evaporation of precipitation.

ECHAM4 The convection scheme is based on a mass flux scheme for deep, shallow and mid-level convection, where clouds are represented by a bulk model which includes updraughts and downdraughts. The convective mass transport and stratocumulus convection are parameterised.

ECHAM4 The planetary boundary layer scheme estimates the surface fluxes of momentum, heat, moisture and cloud water from the roughness length and Richardson number.

ECHAM4 The land-surface processes include heat transfer with a 5-layer soil model and a water budget for soil moisture, interception moisture and snow. The effects of vegetation includes stomatal control on evapotranspiration and interception of rain and snow. The run-off scheme is based on catchment and considerations including sub-grid scale variations of field capacity. The sea-ice temperatures are calculated from the surface energy budget.

ECHAM4 The horizontal diffusion includes gravity wave drag, surface stresses due to gravity waves, orographic forcing and a Richardson number dependent vertical momentum flux scheme.

OPYC3 The surface salinity is relaxed towards observed values employing an annual relaxation scheme. The fresh water input due to river run-off is included in the surface salinity relaxation, and is not prescribed or calculated directly.

OPYC3 The mixed-layer model treats entrainment and detrainment differently and the surface layer is fully coupled to the isopycnal model.

OPYC3 The convection scheme sets the grid values to the corresponding vertical average values of the mixed layer and the underlying unstable layers.

- Model numerics

ECHAM4 Horizontal representation: Spectral transform, Triangular truncation (T21/**T42**/T63/T106).

ECHAM4 Vertical representation: Hybrid coordinate system, Second order finite differences, 19 vertical levels.

ECHAM4 Time integration: Semi-implicit and Leap Frog with time filters. Time step for T42 resolution is 24 minutes (T21 - 40 min./T63 - 15 min./ T106 - 12 min.).

ECHAM4 Orography and land-sea mask from US Navy data set.

OPYC3 The discretisation in space is implemented on a Arkawa B-grid, and the model can define its own grid or use T21, T42 or T106 Gaussian grids.

OPYC3 The advection scheme for momentum and scalar quantities is based on either the Upstream (first order) or the Crowley scheme (optional). Both numerical schemes may suffer from artificial numerical diffusion, but the latter is less diffusive than the Upstream method and less susceptible to overshoot. A predictor-corrector is adopted to allow larger time steps.

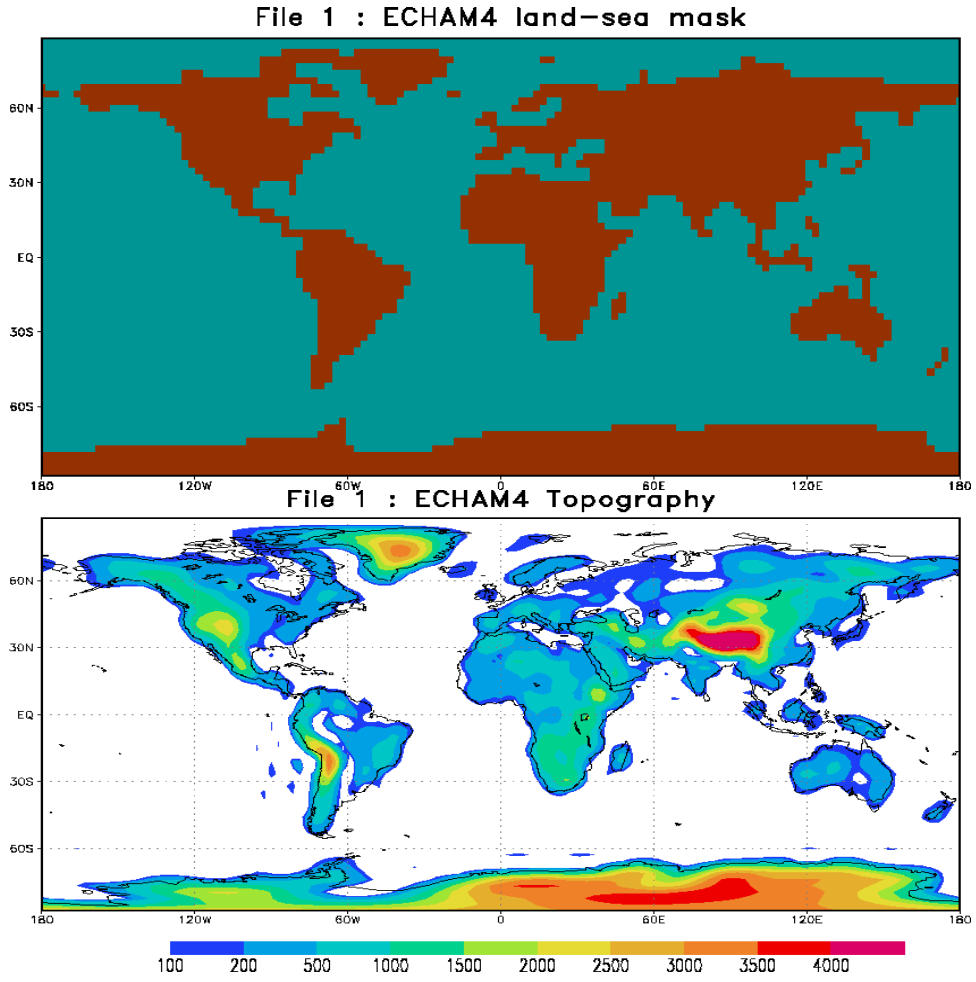

Figure 1: The global land-sea mask used in the ECHAM4 model (upper) and the ECHAM4 model topography (lower) [The figures were obtained from the IPCC Web site].

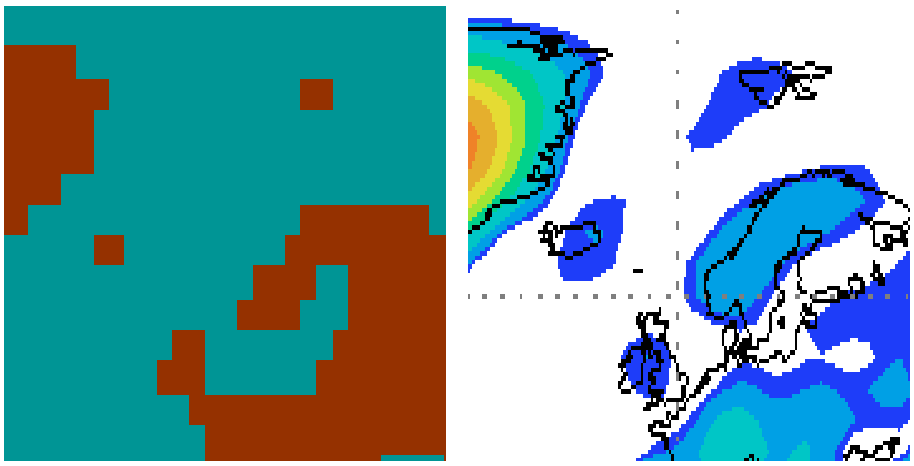

Figure 2: The land-sea mask for the Nordic countries used in the ECHAM4 model (left) and the ECHAM4 model topography (right) [The figures were obtained from the IPCC Web site].

The ECHAM4/OPYC3 model land-sea mask and topography are shown in figure 1. It is evident that the spatial resolution and the orographic details are much too coarse to give a good description of the regional climate in Scandinavia (figure 2). The representation of the Norwegian mountains in the model is not realistic, and the model will therefore not give a realistic description of the topographical influence on the Norwegain climate. Downscaling of the large scale features to local climate parameters is therefore needed to give future climate scenarios for Norway. Furthermore, Italy and Denmark have been dug away and the British Channel is not represented in the model. On the other hand, the model topography was also modified so that the straits between "Denmark" and Sweden, between Greece and Turkey, and between Spain and Morocco were much wider than in the real world[2]. All these modifications to the model landscape may have important implications for the global model circulation (The removal of some Indonesian islands may have even greater significance for the global circulation as this most probably affects the Indonesian throughflow).

2.2 The model integrations

As the results from the model control integration (CTL) are discussed here an outline of the model CTL integration is given. The coupled model was spun up for 100 model years before starting the CTL. Here, the results from

[2]No political motives for this.

the CTL model years 101-400 will be presented.

If a coupled model is not completely spun-up before being used in experiments, then there is a risk of model drift due to the fact that the oceanic and atmospheric components have not reached equilibrium (*Marotzke*, 1996). A constant flux adjustment was added to the transfer of heat between the atmospheric and oceanic components in order to avoid artificial climate drift. The climate drift in the CTL is primarily caused by a mismatch in the fluxes between the atmospheric and oceanic models, possibly due to misrepresentation of physical processes (*Bengtsson*, 1996) or incomplete spin-up. Artificial climate drift represents a serious problem for making projections of possible future climates due to global warming as the difference between real and "fictitious" warming often is indistinguishable. The application of a flux correction compensates for the climate drift, but does not really solve the underlying problem. The flux correction, however, is partly justified by the fact that present knowledge of the real fluxes between the oceans and the atmosphere is poor and that the model representation of these therefore is far from perfect.

The CTL was forced with prescribed concentrations of CO_2, CH_4, and N_2O which corresponded to the 1990 values (*IPCC* (1990), table 2.5). The concentration of industrial gases, such as CFCs, were taken to be zero. There was no SO_4 cycle, and both O_3 and aerosols were prescribed according to their climatological distribution. The CTL was integrated over 300 model years.

3 Description of observational data sets

The observations discussed in connection with model validation were taken from the NMC (now NCEP) ds195.5 data set (monthly mean sea level pressure, 500hPa geopotential heights and temperatures), *Jones* et al. (1998) data set from the University of East Anglia (UEA) (surface temperatures), and the GISST2.2 data sets (sea surface temperatures, see table 1). Although the NCAR ds010.0 and UEA sea level pressure (SLP) data sets contained longer observational record, the NMC data was preferred to the others as it had better coverage over the Arctic and similar spatial resolution to the ECHAM4 data. A comparison between the NCAR, UEA, and NMC SLP suggested that the main features in the different data sets were similar (*Benestad*, 1998b). The NMC sea level pressure (SLP) and the 500hPa geopotential heights records spanned the 1946-1994 period whereas the 500hPa temperatures only started in 1962. The GISST2.2 data set contained sea surface temperature (SST) observations for the 1903-1994 period. Further

Table 1: Data set sizes. The superscript* indicates that the data set contained many diagnostics, including surface (T(0m) and T(2m)) and sea surface temperatures (SST), sea level pressure (SLP), 500hPa geopotential heights (Φ_{500}), and 500hPa temperatures (T_{500}).

Data set	Time interval	Number of months used.	Number of years used.
ECHAM4/OPYC3 CTL*	100-200	1188	99
ECHAM4/OPYC3 CTL*	200-300	1200/1199	100/99
ECHAM4/OPYC3 CTL*	300-400	1200	100
NMC ds195.5 SLP	1946-1994	587	49
NMC ds195.5 Φ_{500}	1946-1994	587	49
NMC ds195.5 T_{500}	1962-1994	395	33
GISST2.2 January SST	1903-1994	92 (Jan/Jul)	92
Jones surface temperatures	1854-1994	1080	90 (1905-1994)
UKMO SLP	1880-1983	552	46 (1900-1945)

description of the data sets is given in *Benestad* (1998b).

The time series comparisons in section 7 were based on both climate data from the Norwegian Meteorological Institute's data base and pressure maps from a merged data set, based on the NMC ds195.5 between 1946-1994 and SLPs from the UK Meteorological office SLP (UKMO SLP) (*Jones*, 1987) before 1946. The latter data set had a coarser grid spacing (5°N-S $\times$10°E-W) than the NMC, but comparisons based on common gridpoints during the period 1946-1982 indicated no major differences.

All the CTL data sequences were not all exactly 100 year long, as can be deduced from table 1, and the reason for this was the organisation of the GRIB data. The GRIB file ctl_mmA_20306 (CTL year 203, June) did not contain geopotential height data, and therefore some of the data sequences were shorter than 100 years.

4 Methods and computational details

The analysis and most of the figures were made in Matlab, and the same code was used for the same analysis applied to the various data fields. The same number of observations were used in the computation of the differences between the mean fields and standard deviations. The Matlab codes estimating these differences were called *pldmean.m* and *pldstdv.m* respectively. The mean field comparisons included all the seasons (12 calendar months) as well

as the mean values of individual calendar months. The standard deviations described the annual cycle in addition to other variability.

The mathematical details of EOF analysis are described in a note by *Benestad* (1999b), and the EOF patterns described here correspond to the type often called 'S-mode' in the literature. The EOF analysis described in section 6 was applied to data on a common grid with similar number of data points in time and space. The data were de-trended before the EOF analysis, by estimating the linear trend at each grid point and subtracting this trend from the respective time series. The EOFs were only calculated for one calendar month at the time (the January or July months, representing the winter and summer time variability), and therefore were not affected by the annual cycle. The EOFs were computed by the Matlab code *pca.m* (which calls *eof.m*).

The spectral analysis was based on the maximum entropy method (MEM), together with a window function in order to reduce the leakage between the different frequencies in the periodogram estimates (*Press* et al., 1989, p.465). The spectral analysis used a window width (or "the number of poles") shorter than the time series length in order to reduce the standard deviation (from 100% for ordinary FT) of the periodogram estimates. The spectral analysis was implemented by the code *xpcacf.m* and the window functions were generated by *wfngen.m*.

Time series for various locations were generated by linear interpolation of the gridded data. The extraction of these time series was done by the Ferret script *picklocs.jnl*, and the subsequent analysis was made in SAS/Excel.

5 Mean values and standard deviations of the gridded data sets.

5.1 Surface temperatures

5.1.1 Annual mean values

The annual mean surface temperatures over a 60-year interval (model years 141-200) of CTL were compared with the 60 most recent years of available observations in the *Jones* et al. (1998) data set. The left panel in figure 3 shows the mean observed values, and the grey areas denote locations with missing data. Figure 3 indicates that there were large areas with missing data especially over the oceans, however, the coverage for Scandinavia was complete for the entire period. The right panel shows the corresponding model annual mean surface temperatures, and the left panel in figure 4 presents the

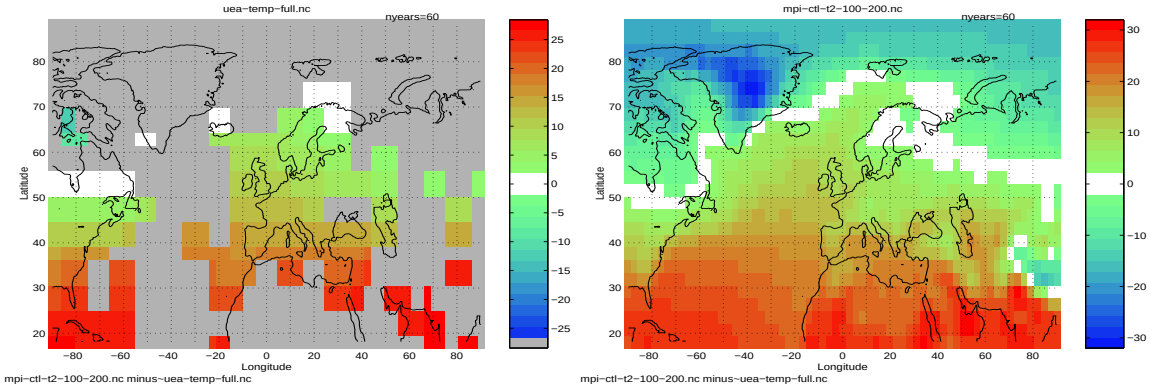

Figure 3: The mean gridded T2 field from *Jones* et al. (1998) is shown in the left panel and the mean T2 fields from ECHAM4/OPYC3 on the right. The results shown here are for model years 141-200 and 1935-1994.

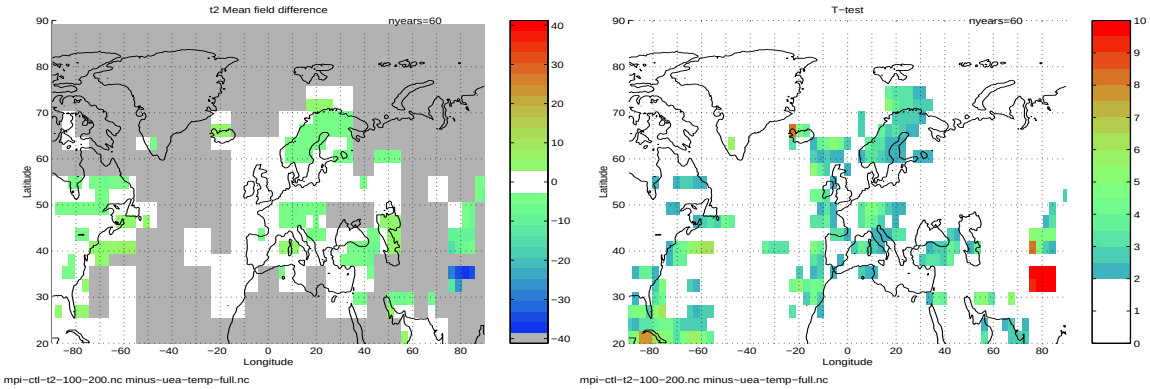

Figure 4: The difference between the mean T2 fields from ECHAM4/OPYC3 and *Jones* et al. (1998) (left) and the results from a two-sample Student's T-test with a null hypothesis of the mean values of the two data sets being different (right). Large values (red) indicate areas where the mean values are significantly different.

difference between the mean model values and the mean observed temperatures.

The model indicated much too cold annual mean conditions over the Himalayas ($\Delta T \approx 30°$C), which is due to the fact that the UEA temperatures (the 1961-1990 climatology) had been reduced to sea level (with a lapse rate of 6°C/km) whereas the model temperatures were for the actual altitude of the given locations. Therefore, the cold bias does not imply a model defect, but indicates different references. It is therefore expected that the model surface temperatures will be colder than those of the observations over Europe and Scandinavia, with a cold bias ranging from 0°C near the sea level to 3°C at 500m height (model topography).

The ECHAM4/OPYC3 CTL temperatures were similar to the observations over most of Europe, however, the model appeared to have a slight cold bias over most of Scandinavia. The right panel in figure 4 shows the scores from a Student's T-test, designed to test the null hypothesis stating that the difference between the CTL and observations was due to chance. The T-test scores of 2.0 and higher, shown in the panel as coloured regions, indicated that there was less than 5% probability that these differences were due to chance. Therefore, the cold model bias over Scandinavia was statistically significant. Most of these differences over Scandinavia may, however, be explained in terms of different reference height used in the model and observations.

A simple test was carried out to investigate whether the model-observation differences may have been due to sampling fluctuations associated with slow variations in the temperatures. This test cannot conclusively prove that the model results were stationary in case the mean differences between the two periods were insignificant, however, data that *did give* substantially different means for different periods may be regarded as non-stationary for time scales longer than 60 years. Interdecadal and centennial oscillations, if present, may for instance introduce different trends and means in the consecutive 60-year sequences, however, over longer time periods the net effect of such oscillations may diminish.

Three arbitrary 60-year long periods of the CTL data were compared to give a rough idea of how much the 60-year annual mean model temperatures vary (figure 5). The mean temperatures did not vary much from period to period, and the greatest differences between the CTL years 141-200 and 241-200 of around 1.5°C (left) were found over Svalbard. A Student's T-test indicated statistically significant differences above the 95% confidence level between Iceland and Greenland and over Svalbard (not shown). The greatest differences between 241-300 and 341-400 were of the order 0.3°C over eastern Greenland (right), and a Student's T-test statistics (not shown)

could not reject the null hypothesis that the temperatures of these different periods were the same. It was therefore assumed that most of the model-observation differences were most probably due to model misrepresentation and not due to sampling errors. There were greater differences between the CTL temperatures from the years 141-200 and the later periods than between the 241-300 and 341-400 intervals.

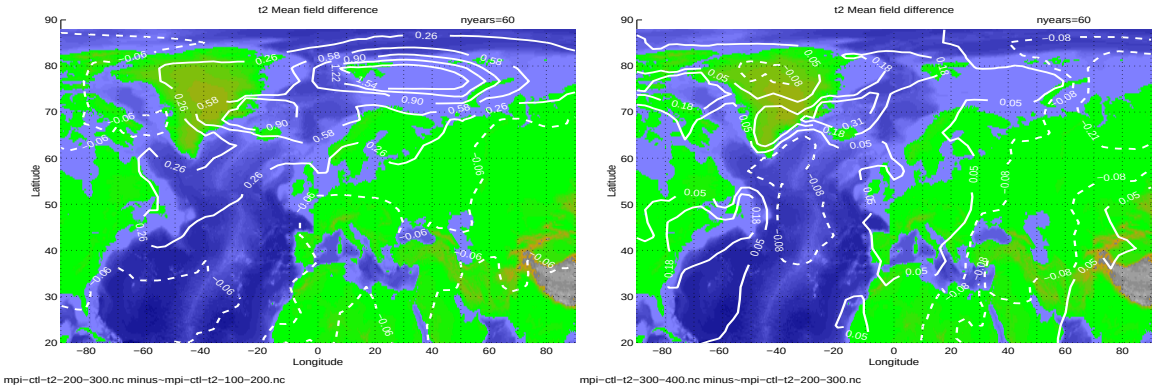

Figure 5: The difference between the CTL T2m from the CTL years 141-200 and 241-200 (left) and from the CTL years 241-300 and 341-400 (right).

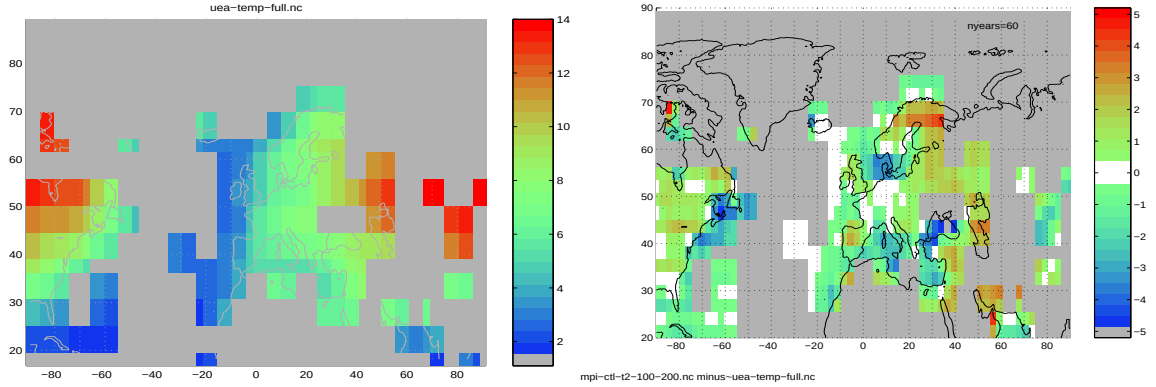

Figure 6: The standard deviation of the T2m field from *Jones* et al. (1998) is shown in the left panel and the difference between the stdv SLP fields from ECHAM4/OPYC3 and NMC ds195.5 on the right.

5.1.2 Standard deviation

Figure 6 shows the observed T(2m) standard deviation distribution (left panel) and the comparison between the model and observed T(2m) standard deviations. It is evident that the surface air temperature varies more

over land than over the oceans, and the greatest temperature amplitudes were found in areas with little maritime influence.

The difference between model and observed variability in the right panel indicates that there were some serious model misrepresentations. The model indicated too strong temperature fluctuations especially over northern Scandinavia, but also too weak variability over the coastal regions near the North Sea, the Newfoundland and around the Mediterranian. Less than 0.5°C of these differences could be explained as sampling fluctuations (not shown).

5.1.3 Evaluation of the different seasons

A comparison between the January mean surface temperatures over a 60-year period indicated that the January mean CTL temperatures were too warm in the maritime regions but too cold over land (not shown). These results do not necessarily imply that the winter time model land-sea contrasts was too sharp, the observations and the model surface temperatures used different altitude references.

The model CTL and UEA July temperatures had statistically significant differences over most of Scandinavia and the model had a cold bias in Southern Scandinavia (not shown, see chapter 7). Again, the observations had been reduced to the sea level.

5.2 Sea level pressure

5.2.1 Annual mean values

Figure 7 shows the mean SLP field of NMC (left), ECHAM4 (right), and the difference between the two mean SLP fields is given in the left panel of figure 8. The same number of data points were used in both the NMC and ECHAM4 SLPs when the differences were calculated. Similar results were obtained for the NCAR ds010.0 and UEA SLPs (not shown).

Figure 8 indicates that the greatest misrepresentations of the mean SLPs were over the Himalayas, in the Arctic and over Greenland, while the mean errors were close to zero in southern Scandinavia. A strong gradient in the mean SLP errors may have more serious implications regionally than the absolute errors, as the average atmospheric circulation is affected by such gradients. The model evaluation indicated a substantial meridional gradient in the mean SLP errors north of Norway, which implies a too weak mean westerly geostrophic flow over the Barents Sea and Fenno-Scandinavia. There were also some large SLP errors over parts of Egypt and Libya.

Because of sparse observational network in the Arctic (*Jones*, 1987), it is

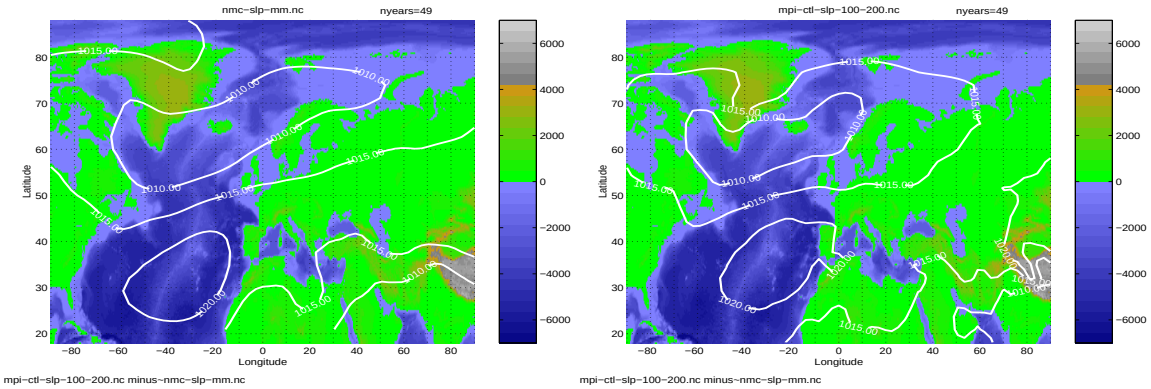

Figure 7: The mean SLP field from NMC ds195.5 is shown in the left panel and the mean SLP fields from ECHAM4/OPYC3 on the right.

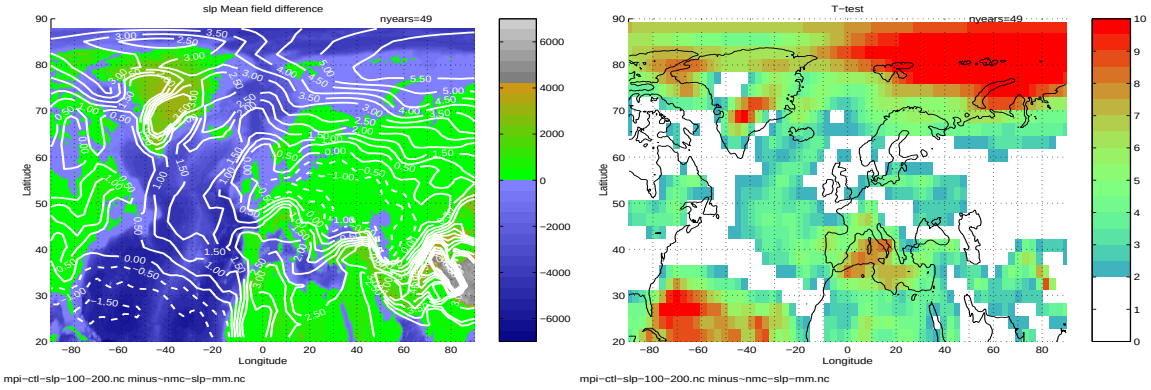

Figure 8: The difference between the mean SLP fields from ECHAM4/OPYC3 and NMC ds195.5 (left) and the results from a two-sample Student's T-test with a null hypothesis of the mean values of the two data sets being different (right). Large values (red) indicate areas where the mean values are significantly different. The comparison was made between the model years 151-200 and 1946-1994.

difficult to say whether the model-observation differences there were entirely due to model misrepresentation or whether errors in the observations also contributed to the differences.

A Student's T-test (*Wilks*, 1995) was used to assess the statistical significance of the difference between the mean SLP fields (figure 8, right). The T-test scores are functions of the mean value differences divided by the root of the respective sample variances (accounting for autocorrelation), and scores greater than 2.0 represent significant $\Delta\overline{SLP}$ at 95% confidence level or higher. The test revealed statistically significant mean errors in the Arctic north of Russia, over the Caribbean, and over the Mediterranean. The high scores over the Caribbean were a consequence of small SLP variance (figure 10). The mean errors over the British isles, southern Scandinavia, Finland and central and eastern Europe were not statistical significant. In these regions, the SLP mean differences were relative small compared with the SLP variability.

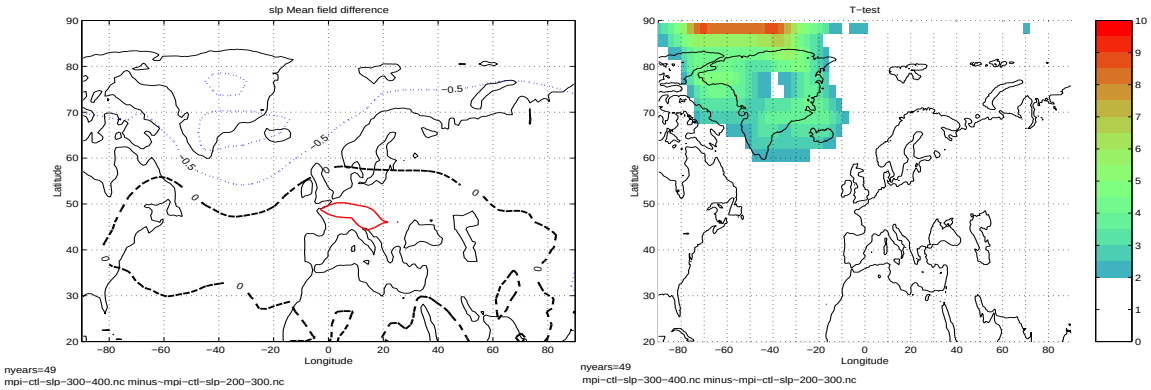

Figure 9: The difference between the CTL SLPs from the years 151-200 and years 251-300 (left) and corresponding Student's T-statistics.

A simple stationarity test was applied to the CTL SLP fields by comparing the mean values over several arbitrary 49 year long sequences. The differences between the mean SLP fields from the intervals corresponding to CTL years 251-300 and 351-400 are shown in the left panel of figure 9, and the Student's T-test score in the right panel. The largest long-term SLP variability was seen over Greenland and the Arctic regions (left panel), but the statistical test suggested significantly different mean SLPs only over Greenland and the adjacent maritime regions. The differences in the mean fields from the two different periods may therefore indicate the presence of interdecadal and centennial oscillations. Such oscillations may explain some of the differences between simulated and observed SLPs shown in figure 8.

Machenhauer et al. (1998) reported a common model SLPs bias in a belt running east-west across Europe. These low SLPs were generally located between two centres of too high SLP values over northern and southern Europe. Figure 9, however, gives an indication of sampling fluctuations associated with the model SLPs, suggesting that some of the errors in the SLP bias associated with the storm track activity may be due to decadal and interdecadal variability.

5.2.2 Standard deviation

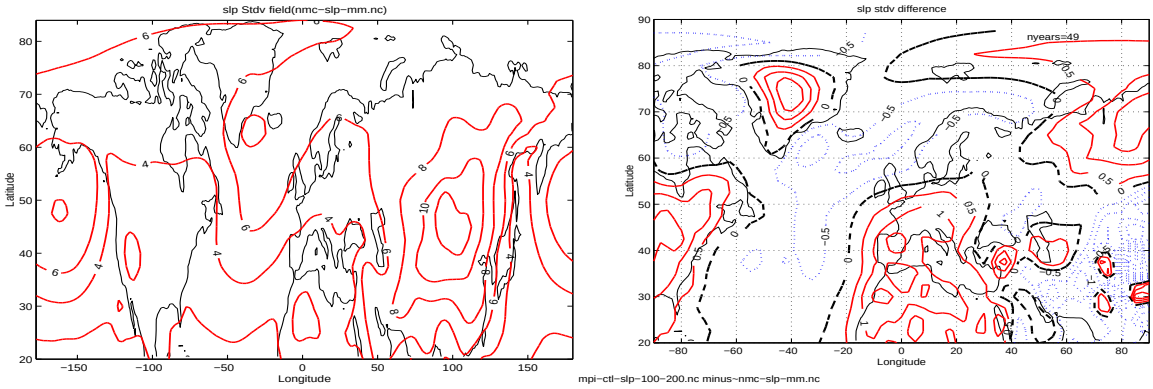

Figure 10: The standard deviation (stdv) of the SLP field from NMC ds195.5 is shown in the left panel and the difference between the SLP stdv fields from ECHAM4/OPYC3 and NMC ds195.5 on the right. The comparison was made between the model years 151-200 and 1946-1994.

The magnitude of the SLP variability (all time scales shorter than 49 years) are shown in figure 10. The observations (left panel) suggested high variability over the Himalayas, in the North Pacific, and southwest of Iceland. With the exception of the Himalayas, the SLP variance maxima were found over the ocean interior, and the smallest standard deviations were found in the sub-tropics.

The differences between the observed and model SLP standard deviation indicated too much model variability over Greenland, in the Arctic north of Russia, and over the continents. The model predicted too little SLP variability over the Atlantic ocean.

The difference between the SLP standard deviation for the CTL years 151-200, 151-300 and 351-400 was investigated (not shown) to see if there were substantial slow changes in the model variance over time that would introduce non-stationarities in the analysis. The magnitude of the differences in the standard deviation fields was less than 0.5hPa except for over

southern Greenland and the Greenland-Iceland Sea, where the difference in the variance was slightly above 0.5hPa between the intervals corresponding to the CTL years 151-200 and 251-300. The comparison between the model years 251-300 and 351-400 also indicated some differences in the SLP standard deviation for the two latter periods. The largest differences of about 0.5hPa were found over a region stretching from the North Sea to southern Greenland. In other words, sampling fluctuations due to long term model variability may account for a good part of the differences seen between the model and observed standard deviation.

5.2.3 Evaluation of the different seasons

A comparison between the January mean SLPs from the 151-200 and 251-300 model year intervals revealed a similar pattern as figure 8 (left panel), but with larger errors (not shown). A corresponding Students T-test nevertheless gave no significant errors over the North Atlantic ocean, due to greater winter time SLP variability in the maritime regions. In July (not shown), on the other hand, the error magnitudes were smaller than in January, but due to weaker summer time variability, the errors over large regions of the North Atlantic ocean now scored higher than 2.0 in the Students T-test. The regions over Greenland and the Himalayas had similar errors in the winter and the summer, suggesting that some of the SLP misrepresentation may be due to a poor description of orographic features.

5.3 500hPa geopotential heights

5.3.1 Annual mean values

Figure 11 shows the mean NMC Φ_{500} (left) and ECHAM4 Φ_{500} (right) fields for the 1946-1994 period and the model years 151-200 respectively. The differences between the model and the observations are shown in the left panel of figure 12. The smallest mean errors, $\Delta\overline{\Phi_{500}}$, were seen over southern Scandinavia and the Baltics and the greatest $\Delta\overline{\Phi_{500}}$ values were found over the Arctic. A Student's T-test indicated insignificant differences in $\overline{\Phi_{500}}$ over Scandinavia, the Baltics, North America and Eurasia, but significant differences elsewhere (figure 12, right). In general, the ECHAM4 500hPa isobar was at higher altitudes than indicated by the observations.

Interdecadal and centennial variations in the ECHAM4/OPYC3 CTL Φ_{500} fields, if present, may be responsible for the statistically significant differences in the mean fields of different 49-year periods (not shown). In parts of the Arctic, the values of Φ_{500} were non-stationary for time scales smaller

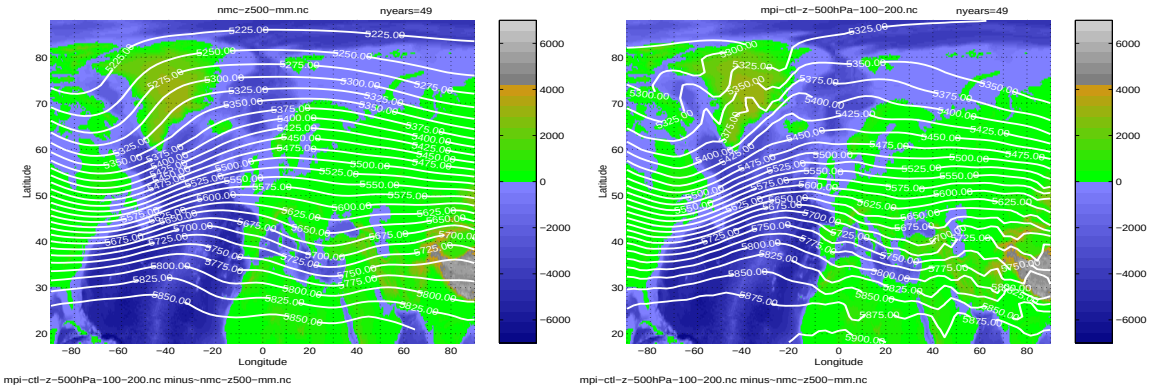

Figure 11: The mean 500hPa geopotential heights field from NMC is shown in the left panel and the mean ECHAM4/OPYC3 500hPa geopotential heights on the right. The results are for the 1946-1994 period and the model years 151-200 respectively.

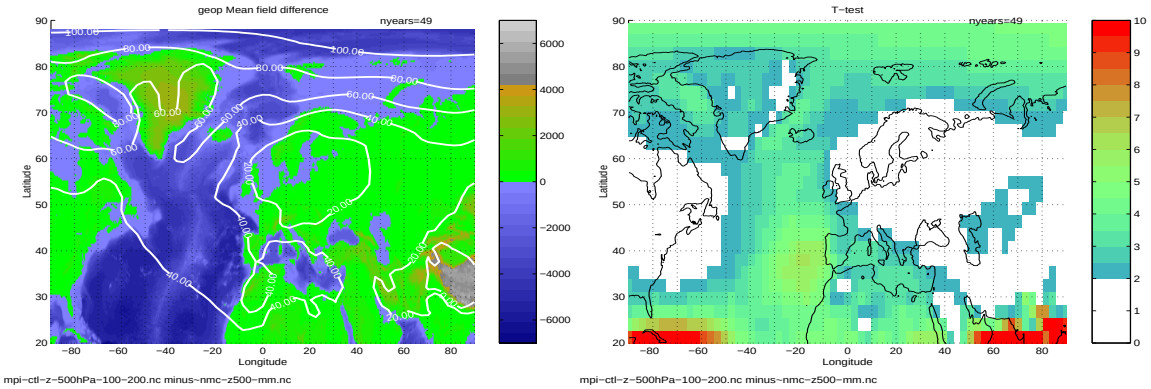

Figure 12: The difference between the ECHAM4/OPYC3 and NMC 500hPa geopotential heights (left) and the test of differences in NMC and ECHAM/OPYC3 mean 500hPa geopotential heights using a two-sample Student's T-test (right).

than 50 years according to a comparison between the model years 251-300 and 351-400, whereas the CTL intervals, 151-200 and 251-300 did not give any evidence of non-stationarity. Although statistically insignificant, the estimated values of $\Delta\overline{\Phi_{500}}$ for model years 151-200 and 251-300 had a spatial structure that bore a weak resemblance to the NAO, suggesting that the presence of weak interdecadal and centennial oscillations cannot be ruled out. This structure was absent in the difference field constructed from the intervals based on CTL years 251-300 and 351-400. Hence, there was no clear evidence that the model 500hPa geopotential heights were non-stationary on the 50-year time scale.

5.3.2 Standard deviation

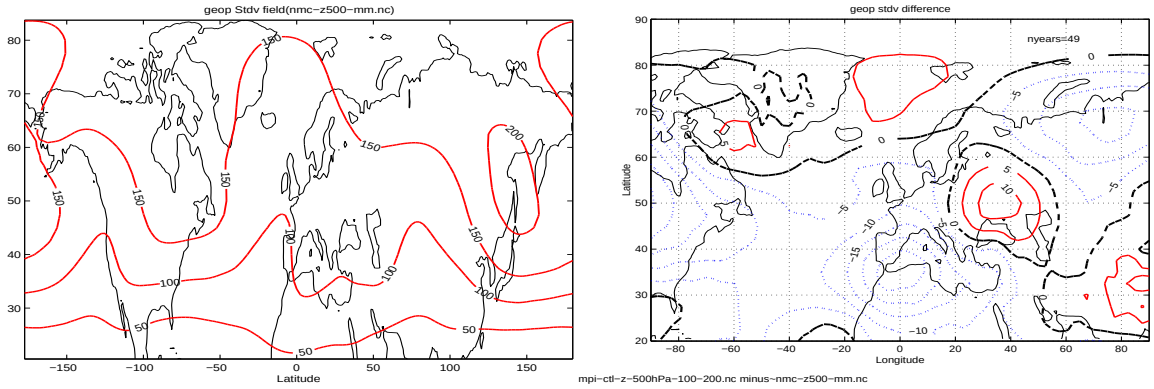

Figure 13: Same as in figure 10, but showing standard deviations of Φ_{500} instead of SLP. The results are for the 1946-1994 period and the model years 151-200 respectively

The observed (NMC) Φ_{500} standard deviations (1946-1994) and the differences between the model and NMC standard deviations are shown in figure 13. The differences in the model and observed Φ_{500} standard deviations were small over the North Atlantic, Fennoscandia and the Baltics, and the largest differences were found over Ukraine, North-western Africa, North America, and over the northern Ural mountain range. The difference plot in the right panel shows indications of standing wave features, extending northeastwardly from northwestern Africa to the Kara Sea. A close inspection of the left panel in figure 13 may indicate that this wave train structure in the difference field was not a spurious model feature, but rather due to an actual phenomenon which was not captured by the model.

The standard deviation of the CTL Φ_{500} field between the model years 151-200 and 251-300 (not shown) varied as much as 10m southeast of Green-

land, and between model years 251-300 and 351-400 large standard deviation differences were found over the Labrador Sea (5m), the Bay of Biscaya (4-5m) and over Greece (3m). Although the slow variations in the Φ_{500} amplitudes in general were less than 25% of the greatest differences between observations and CTL years 151-200, these results indicate that a 49-year long period probably is too short for model evaluation.

5.3.3 Evaluation of the different seasons

The differences between the January mean values of CTL (years 151-200) and NMC (1946-1994) Φ_{500} over the Mediterranean, Greenland and the Arctic were of the order of 60m (not shown), while over Scandinavia, the model Φ_{500} were approximately 40m too high. The scores from a Student's T-test indicated that most of the model-observation differences were statistically significant.

The differences between the July 500hPa geopotential heights (not shown) were small over western Mediterranean (around zero), but large over the Arctic (greater than 120m). The mean July errors over Scandinavia were between 20 and 40m. In July the errors were insignificant only over small areas, including the Iberian peninsula and northwest Africa, northeastern Russia, and the northeastern USA.

5.4 500hPa temperatures

5.4.1 Annual mean values

The comparison between the NMC (1962-1994) T_{500} and ECHAM4 (model years 167-200) T_{500} fields in figures 14 and 15 suggests that the 500hPa isobar was in general too cold in the model, and especially over Scandinavia where the magnitude of the cold bias was greater than 1.0°C. The largest model CTL T_{500hPa} discrepancies were found over the Himalayas, but a confidence test only indicated significant model-observation differences over the Carribian, northeastern Africa and the middle East.

The corresponding 33-year annual mean model 500hPa isobar (not shown) was slightly too high over Europe (20-60m) in addition to being too cold, indicating slightly too cold conditions in the higher altitudes. The fact that the SLP errors over Scandinavia were small furthermore suggests that the model produced too warm conditions near the surface. This cooling aloft and warming near the ground implies that the hydrostatic stability over Scandinavia may have been too weak in the CTL results. It is possible that the crude representation of the vertical density profile (19 vertical levels)

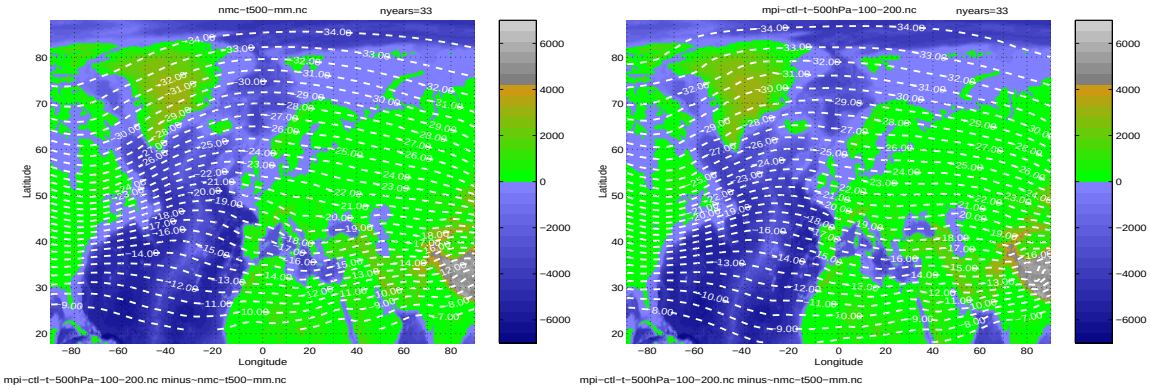

Figure 14: The mean 500hPa temperature field from NMC is shown in the left panel and the mean 500hPa temperature field from ECHAM4/OPYC3 (right). The results are for the 1962-1994 period and the model years 167-200 respectively.

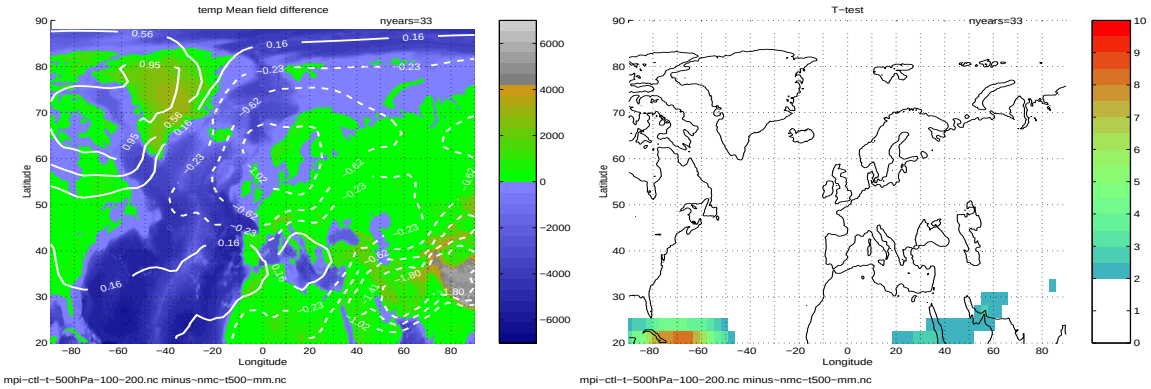

Figure 15: The difference between the ECHAM4/OPYC3 and NMC 500hPa temperatures (left) and the test of differences in NMC and ECHAM/OPYC3 mean 500hPa geopotential temperatures using a two-sample Student's t-test (right).

and/or the rough topographical effects in the model may have introduced systematic model errors.

A comparison between the various 33-year T_{500} mean fields revealed only wave-like differences which were statistically insignificant in limited regions (not shown). As was the case for both SLPs and 500hPa geopotential height, the comparison between the consecutive 33-year mean values could not rule out the presence of interdecadal and centennial variations, and the test indicated that the T_{500} fields were non-stationary over time scales longer than 30 years.

Upper air observations tend to be less accurate than surface observations, and some of the discrepancies between the observations and the model at the 500hPa isobar surface may have been due to observational errors.

5.4.2 Standard deviation

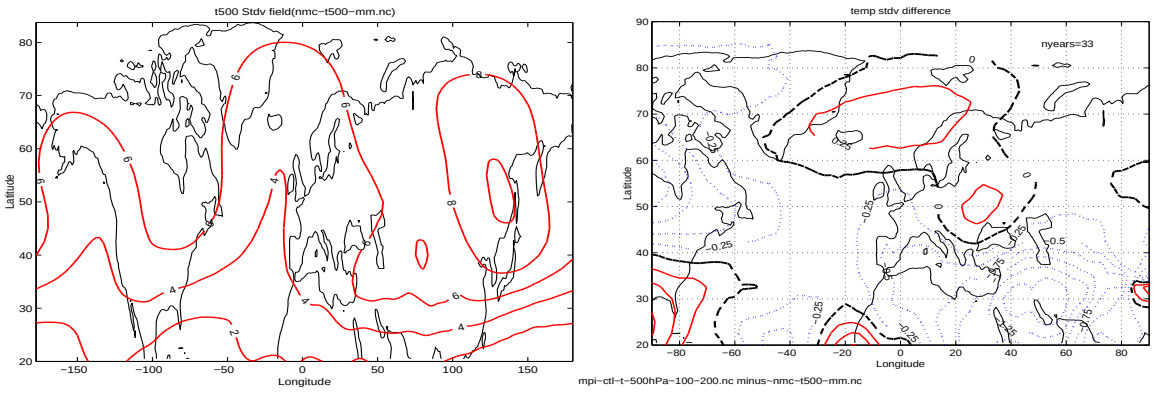

Figure 16: Same as in figure 10, but showing standard deviations of T_{500} instead of SLP.

The standard deviation fields, shown in figure 16 suggest large variability over the continents (left) and the differences between the observations and model values (right) suggested that the model described too much variability over the Norwegian Sea and over Belorussia. The model described too little variability in the T_{500} fields over the Middle East, south-western Europe, and eastern Canada.

The greatest differences between T_{500} standard deviation fields from the CTL year 167-200 and 267-300 sequences were 0.25°C over Iceland, Finland and Belorussia (not shown). The standard deviation differences between the intervals 267-300 and 367-400 were of similar order, but with maximum magnitude of 0.3°C over the Labrador Sea.

5.4.3 Evaluation of the different seasons

The differences between the mean model and observed January T_{500} fields indicated errors of +2°C over Hudson Bay, -1°C over Scandinavia, and +1°C over part of the Mediterranean sea (not shown). Most of the differences between the model and observations were statistically insignificant, with the exceptions over the North Sea, Hudson Bay, and Tadzhikistan (not shown).

A similar comparison between the model and observed July 500hPa temperatures (not shown) indicated small ($|\Delta T_{500}| < 1$°C) and negative mean errors over Scandinavia. The worst T_{500} misrepresentation was seen over the Middle East (magnitude > 5°C), the Caribbian (magnitude > 2°C), Greenland (magnitude > 1°C) and over Romania (magnitude > 1°C). The errors in T_{500} were only statistically significant near these maxima as well as over the British isles and northern Russia.

Since both the winter and summer errors were smaller in magnitude than the annual mean errors, the worst model misrepresentation of the T_{500} fields took place during spring and/or autumn.

5.5 SST

5.5.1 Annual mean values

The mean mid latitude SST field described relatively warm sea surface in the eastern part of the Atlantic ocean and colder surface waters in the west (figure 17, left). This asymmetry has been attributed to the heat transport by the ocean currents (*Grötzner* et al., 1998; *Latif*, 1998; *Sutton & Allen*, 1997), and the Gulf stream transports warm water eastward and poleward whereas the relatively cold western ocean basin is a consequence of the cold western boundary current. The OPYC3 ocean model has a low spatial resolution and may not resolve the Gulf current well enough to give a realistic description of the SST fields. The difference between the observed SSTs from GISST2.2 and the model surface temperatures in the right panel of figure 18 suggests substantial mean SST errors away from the ocean interior. The largest differences between observed and model SSTs were seen along the coast of Newfoundland. The comparison shown here only included the January months (because of memory limitations as the GISST2.2 data set had a spatial resolution of $1° \times 1°$).

Observations have suggested possible North Atlantic SST oscillations with time scales of around 10 years (*Deser & Blackmon*, 1993), 35-60 years (*Timmermann* et al., 1998; *Lau & Weng*, 1995; *Delworth* et al., 1993), 76 years (*Schlesinger & Ramankutty*, 1994), and 80-90 years (*Appenzeller* et al., 1998). The implications of such low frequencies being present are that 100-year and

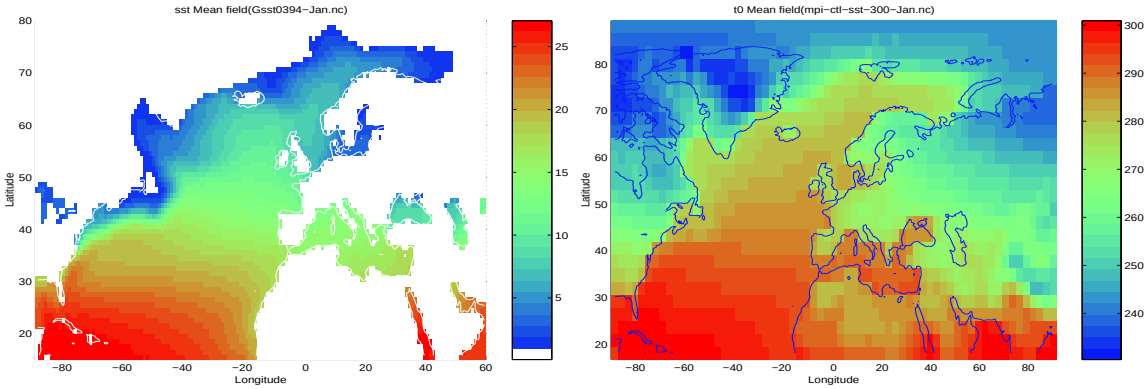

Figure 17: The mean January SST field from GISST2.2 is shown in the left panel and the difference between the mean surface temperature fields from ECHAM4/OPYC3 and GISST2.2 SSTs on the right.

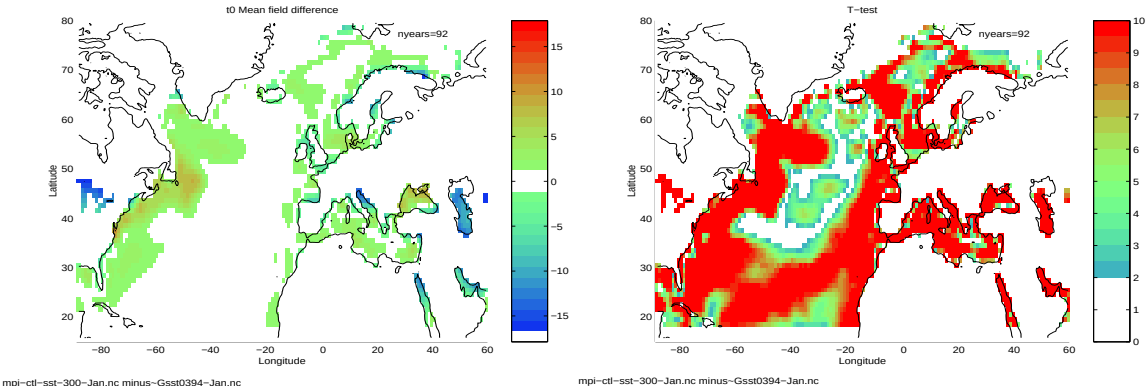

Figure 18: The difference between the January mean surface temperature fields from ECHAM4/OPYC3 and GISST2.2 SSTs (left panel). The test of differences in mean SSTs from GISST2.2 and ECHAM/OPYC3 (t_0) using a two-sample Student's t-test (right).

shorter SST records may not be sufficiently long to achieve a stationary sample. Bearing this reservation in mind, we find significant differences in the north-western Atlantic (figure 18, right panel), although these discrepancies do not conclusively demonstrate any ocean model shortcomings. A Student's T-test identified statistically significant differences in the mean SSTs and the test scores suggested that the probability that the SST differences were due to chance was less than 5% almost everywhere (figure 18, right panel).

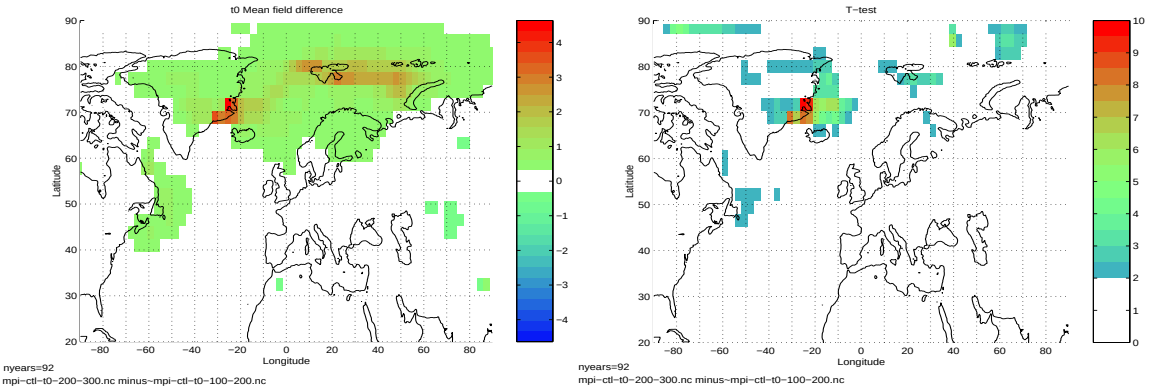

Figure 19: The difference between the CTL SSTs from the years 109-200 and the years 209-300 (left) and corresponding Student's T-statistics (all seasons).

The surface temperatures were tested for stationarity by comparing the mean values for the model years 108-200 with 208-300 and 208-300 with 308-400. The differences (figure 19) between the 108-200 and the 208-300 model year intervals over most of the North Atlantic were small and a two-sample Student's T-test could not prove that the SSTs here were non-stationary. In the Greenland-Iceland Sea and over the Arctic, on the other hand, the differences in the temperatures scored higher than 2.0 in the Student's T-test. The differences between the mean SSTs between the model years 108-200 and 208-300 were substantially greater than between 208-300 and 308-400. These results suggested only strong interdecadal and centennial variability at the high latitudes, and oscillations in the Arctic with time scale longer than 90 years were prominent in the early ECHAM4/OPYC3 model CTL results. The fact that the long-term SST variability were reduced by a factor of around 5 after the first 100 years may indicate that spin-up effects still were present in the beginning of the CTL integration.

The locations where the model differed most from observations, such as along the US east coast, exhibited relatively weak long-term changes. Most of the differences between the model results and the observations may therefore

be attributed to model misrepresentation of the SSTs.

5.5.2 Standard deviation

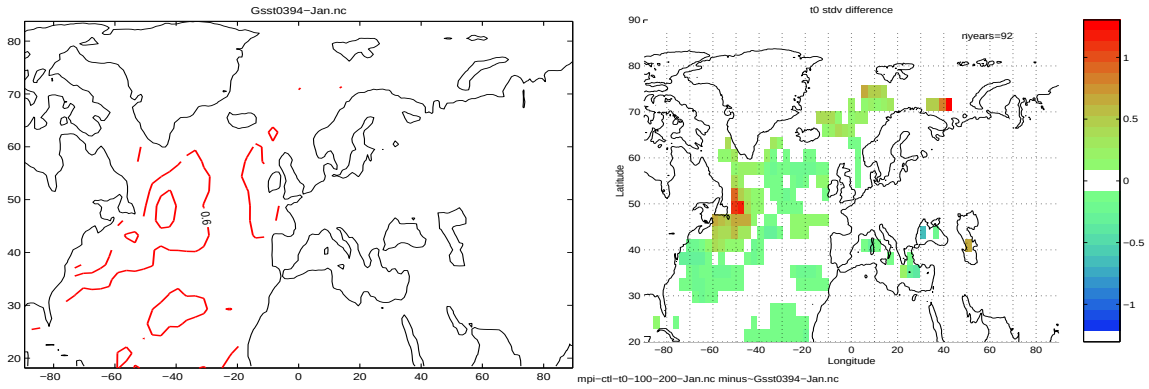

Figure 20: Same as in figure 17, but showing standard deviations instead of mean values.

The SST standard deviations are shown in figure 20. Strongest SST variability was found off the coast of Newfoundland and along the Gulf stream extension (left panel) where large differences between the model and observed mean SST also were found.

The greatest changes in the sea surface temperature variability between the different CTL periods were seen over the Arctic regions, and isolated grid boxes differences with greater than 2°C were seen near the east coast of Greenland and Iceland (not shown). The SST amplitudes in the Labrador Sea did also show large variations (0.5-1°C) over the centuries, and it is possible that surface temperatures and ice cover in the polar region reflect low frequency oscillation such as variations in the thermohaline circulation or the "Arctic Oscillation" (*Thompson & Wallace*, 1998).

5.5.3 July SSTs

A comparison between the model and observed July SSTs (not shown) indicated that the mean summer time model SSTs between Iceland and Svalbard and the maritime regions adjacent to Greenland were too warm by around 4°C, whereas the North Sea and the Norwegian coast were around 2°C colder than the observations. The mean SST errors were generally smaller than 1°C in the North Atlantic ocean, although a small region off the coast of Newfoundland was about 2°C too warm. The scores from a Student's T-test (not shown) indicated a very high confidence that most of the differences between

the model and observed SSTs in the Nordic Seas were not due to sampling fluctuations. The SST differences in most of the North Atlantic and Labrador Sea were regarded as statistically significant above the 95% confidence level.

Christensen et al. (1998) found seasonal biases in the ECHAM4/OPYC3 SSTs, with too warm sea surface in the winter and too cold in the summer. They attributed these systematic seasonal SST biases to a seasonally independent flux correction.

6 EOF Analysis

The statistical downscaling models used in the *RegClim* project make predictions based on the relationship between historical patterns of circulations and past records of local climate variables. It is therefore important that the model anomalies have similar spatial structures to the observed anomalies if the model results are to be taken as predictors for the downscaling models. The data fields describing the circulation patterns can be decomposed into principal component analysis (PCA) products (*Wilks*, 1995), also known as Empirical Orthogonal Functions (EOFs), in order to reduce the degrees of freedom, remove noise, and minimize the computational demands. The EOFs may be thought of as (geographically weighted) eigenvectors in data space (*North* et al., 1982; *Peixoto & Oort*, 1992), with each vector describing a unique normal mode[3] spatial structure. Here the spatial structure, variance and spectral characteristics of the the model EOFs are compared with the observations.

In order to compare the EOFs from the model results and observations, the data were interpolated onto the same grid (similar to the ECHAM4 grid) and the same number of data points were used in the EOF analysis. The data were de-trended by subtracting the best fit linear trend at each grid point prior to PCA. A similar analysis on gridded data that were not detrended and had different resolution and time lengths gave other results, suggesting that this analysis was sensitive to the trend, the spatial resolution and/or the length of the data record. Therefore, the EOFs shown here should not be interpreted as physically meaningful beyond the use in downscaling models, and are only shown here as a comparison between the model and observations.

The sense of the EOFs (i.e. the direction of the eigenvectors) shown here is arbitrary, as the spatial EOF patterns indicate which regions have variability that is correlated and which locations that are anti-correlated. It

[3] "normal mode" refers to an orthogonal vector describing a time dependent or oscillating structure.

is the PCs that determine which anomalies that are positive and which are negative. Therefore, in this EOF comparison, it does not matter whether the EOFs are "inverted" (all anomalies of opposite polarity) or not.

The principal components from both the CTL results and the observations were subject to spectral analysis in order to compare the dominant time scales in the model results and the observations. The results discussed here were obtained using the maximum entropy method (*Ghil & Yiou*, 1996; *Press* et al., 1989). It is important to note that the data records in this spectral analysis were short (42 years for SLP and Φ_{500}, 32 years for T_{500}, and 92 years for SST), and therefore the lowest frequency power density peaks (with time scales longer than 20 (50) years) were not captured by the spectral analysis described here. The lowest harmonics obtained from the spectral analysis corresponded to periods similar to the record length, and therefore were associated with high sampling uncertainties.

Different trials with various filter types (square, Bartlett, Tukey-Hanning, and Parzen, see *Press* et al. (1989) p. 465 for more details) and window widths (14, 20, 30, 42 year) indicated that the results of the spectral analysis were sensitive to both window size as well as filter type. The ordinary square windows gave the least credible results, and the smaller window width gave more stable spectral estimates at the expense of the spectral resolution. Here a window width of 20 years was used for comparisons with the NMC and UEA data and 50 years for the GISST2.2-model SST comparison, representing a compromise between stability and capturing decadal variability. The spectral results were based on spectral analysis with a Tukey-Hanning window type. The spectral power densities at the high frequency end may be subject to aliasing problems (despite low-pass filtering), and should therefore be regarded with care (*Wilks*, 1995, p. 347).

6.1 January surface temperatures

6.1.1 Variance and Spatial Structures

Figure 21 shows the spatial structures of the 2 leading model (upper panel and third from the top) and observed (second from the top and bottom panel) EOFs for January mean temperatures[4]. The observations contained a large number of missing data wholes, shown in grey in the figure.

[4]The reason why there were differences in the regions of valid data between figure 3 and figure 21 was that the mean difference test was carried out on a spatial grid on which the grid points were computed using a nearest neighbour interpolation whereas the EOFs were estimated on a grid employing linear interpolation. There were also holes in different regions during the different seasons.

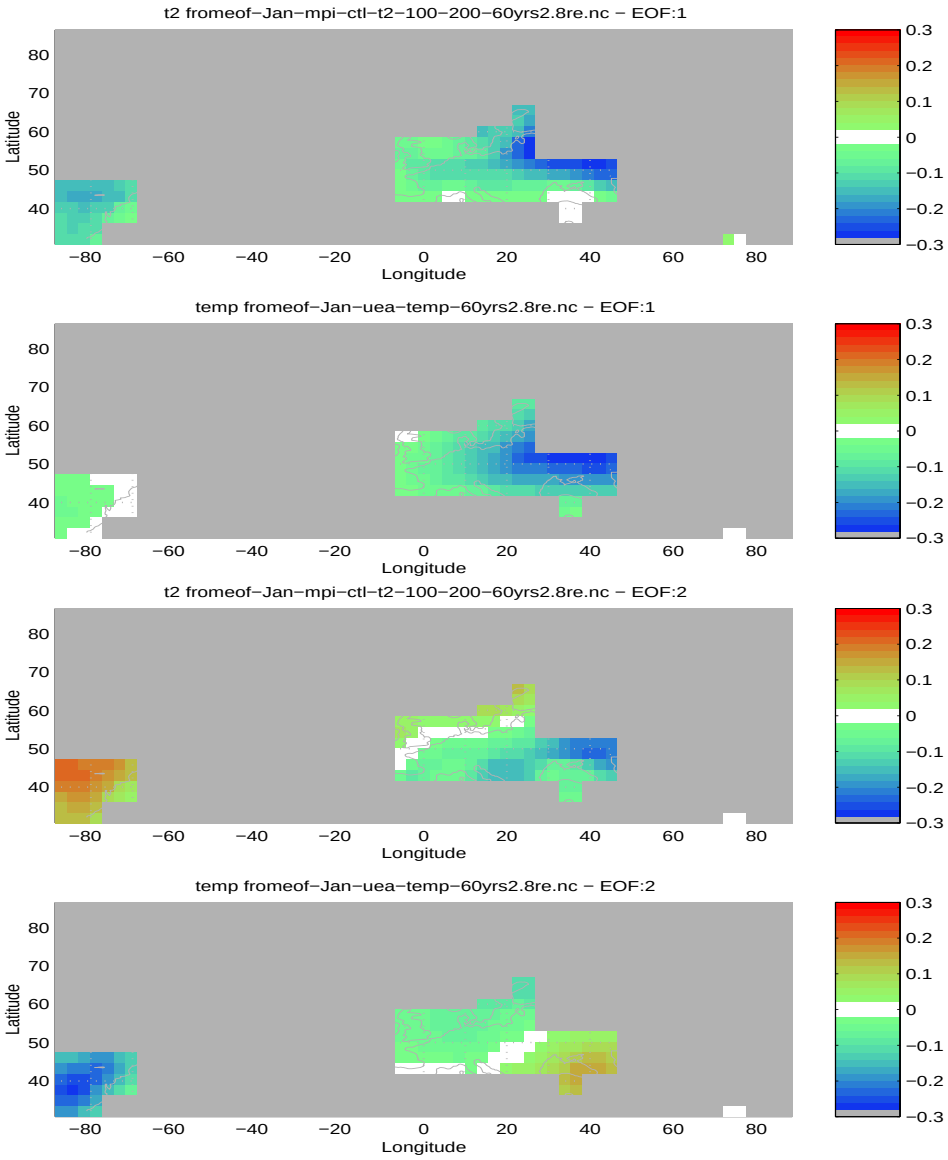

Figure 21: The 2 leading January EOFs of ECHAM4 CTL and *Jones* et al. (1998) surface temperatures. The first and third panels are the two leading model EOFs and the second and fourth panels the corresponding observed EOFs.

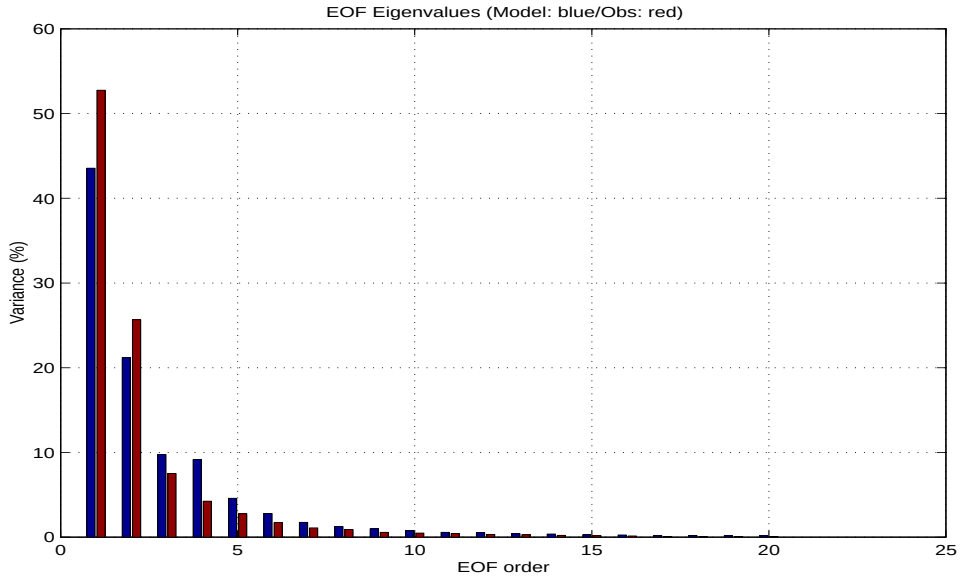

Figure 22: The eigenvalues of the 20 leading modes for MPI ECHAM4 CTL surface temperatures (first columns) and *Jones* et al. (1998) surface temperatures (second columns).

Both model and observed leading EOFs (two upper panels) indicated strong variability over eastern Europe and western Russia. These temperature anomalies were coherent with a weak signal over the eastern USA.

The second model EOF (third from the top) indicated strong variability over the south-eastern USA anti-correlated with prominent variability over eastern Europe. A tendency of variability out of phase with surface temperatures over eastern Europe can also be seen near the British isles and over the North Sea. The second observed EOF (bottom panel) also indicated strong variability in the USA correlated with fluctuations over the North Sea and northern Europe, but with opposite polarity to those anomalies over eastern Europe.

A comparison between the third order EOFs (not shown) indicated similar east-west temperature dipole structures, but with substantially stronger weights over the Baltic Sea in the observations. The eastern centre of action was also found further north in the model than in the observations. The observed EOF furthermore indicated a correlation between the surface temperatures over Chechenya and the eastern USA which was weak in the model.

The main features of the fourth order EOFs were also roughly similar, although the dipole pattern described by the model was alinged along a

north-south axis whereas the observations indicated a northeast-southwest oriented dipole.

The leading model EOF accounted for around 43% of the total variance while the leading observed EOF described nearly 52% (figure 22). The second observed EOF also described more of the variance than the corresponding model EOF.

6.1.2 Spectral Characteristics

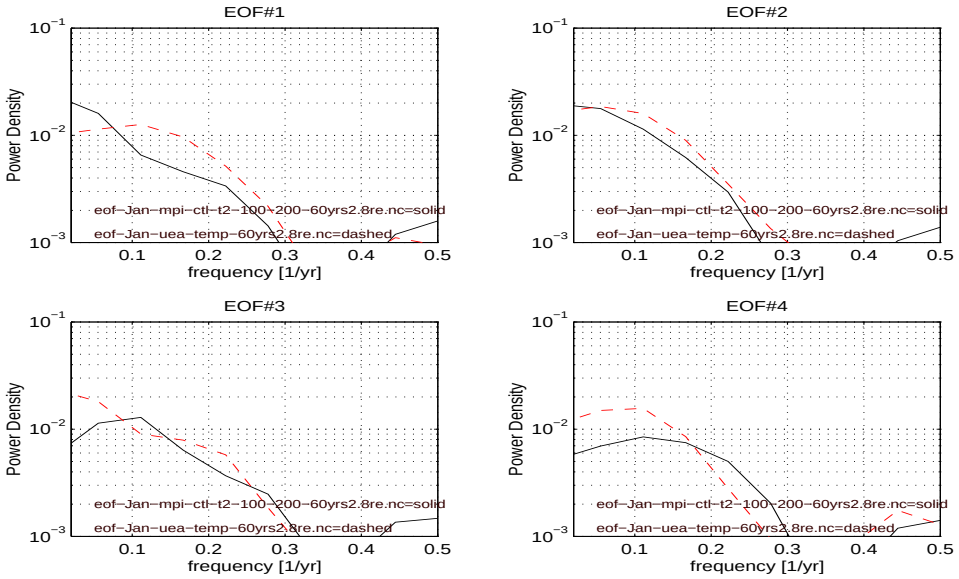

Figure 23: The power density spectra associated with the 4 leading January EOFs where the solid lines represent MPI ECHAM4 surface temperatures and the dashed lines *Jones* et al. (1998) surface temperatures. The PCs had been smoothed with a 3-point Gaussian low-pass filter to remove the high frequency signal.

The power spectrum of the leading model EOF was characterised by a red-noise spectrum on which hints of weak spectral peaks were superimposed (figure 23). The observations also indicated red noise charactersitics, but with a more prominent oscillations of time scales of around 10 years. The power spectra for the second EOFs were similar, but the power spectra of the third and fourth order indicated different spectral properties. The difference between power spectra for the higher order EOFs may be a result of the differences in the spatial structures of these modes, but may also indicate shortcomings in the model description of the temporal evolution of these features.

6.1.3 Sampling fluctuations

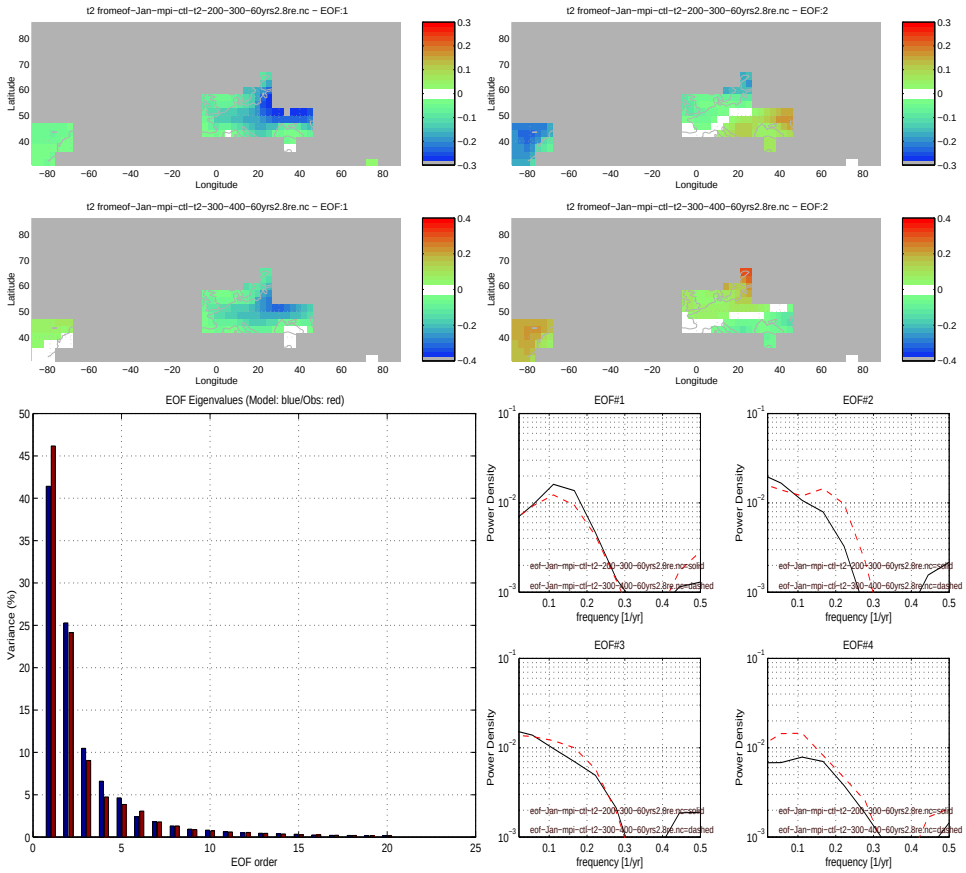

Figure 24: The two leading EOF pairs from CTL years 241-300 (top) and 341-400 (second from the top) and corresponding eigenvalues (lower left) and power spectra (lower right).

EOF analysis applied to different 60-year long intervals suggested that the main features of the 2 leading EOFs were robust (figure 24), and the eigenvalues of the leading EOFs varied by as much as 5% suggesting that the relative strength of the leading EOF pattern was not constant. Some of the grid point time series contained a clear linear trend (not shown), which was removed prior to the analysis. However, the fact that such trends were clearly evident also suggests non-stationarity on these time scales in the unprocessed data.

The power spectra of the leading EOFs from CTL years 241-300 and 341-400 (figure 24) were similar for the leading EOFs but different for the second EOF. The differences in the spectra of the second EOF, however, is a result of sampling fluctuations, and gives an indication of the statistical

uncertainty of the EOF and spectral analysis.

The similarity between the patterns and the spectral properties of the various intervals suggest that most of the model-observational T(2m) PCA differences were real and not due to sampling fluctuations.

6.2 July surface temperatures

Figure 25 shows the spatial structures of the two leading model and observed July T(2m) EOFs. It is evident from the comparison of the leading EOFs (two upper panels) that there was a systematic error in the model representation of the surface temperature anomalies, simulating too much variance in western Europe and too little in eastern Europe. The second EOFs were roughly similar, although there were some differences in the details. The model placed the eastern centre of action too far west, and the line of zero variability ran northwest-southeast in the model, whereas the observations suggested a more south-north alignment.

6.3 January SLP

6.3.1 Variance and Spatial Structures

The leading model and observed January SLP EOFs were similar (upper two panels in figure 26), describing a North Atlantic Oscillation (NAO) type structure. The model produced too strong anomalies over central Iceland compared to the observations, and the southern centre of action was further northwest than in the observations. The eigenvalues in figure 27 indicate that the leading model EOF described more variability than the corresponding observed EOF, accounting for almost 40% of the variance whereas the NMC SLP EOF described almost 35%. The model consequently slightly underestimated the importance of the higher order EOFs.

The second EOFs in the ECHAM4 SLP and NMC SLP data sets both described a west-east dipole structure, and their corresponding eigenvalues were of similar magnitude. Slight differences in the second EOF spatial structures suggested that the model centre of action over the North Atlantic was too strong and that the model SLP anomalies in central Asia were too weak and too far southeast compared to the observed EOF pattern.

The third model EOF (not shown) gave hints of the North Atlantic storm track, although dominated by SLP anomalies over eastern Russia. The third observed NMC EOF, on the other hand, showed a west-east dipole pattern with centres of action over the UK and the Urals, and was similar to the second NMC EOF with the whole pattern shifted eastwards by approximately

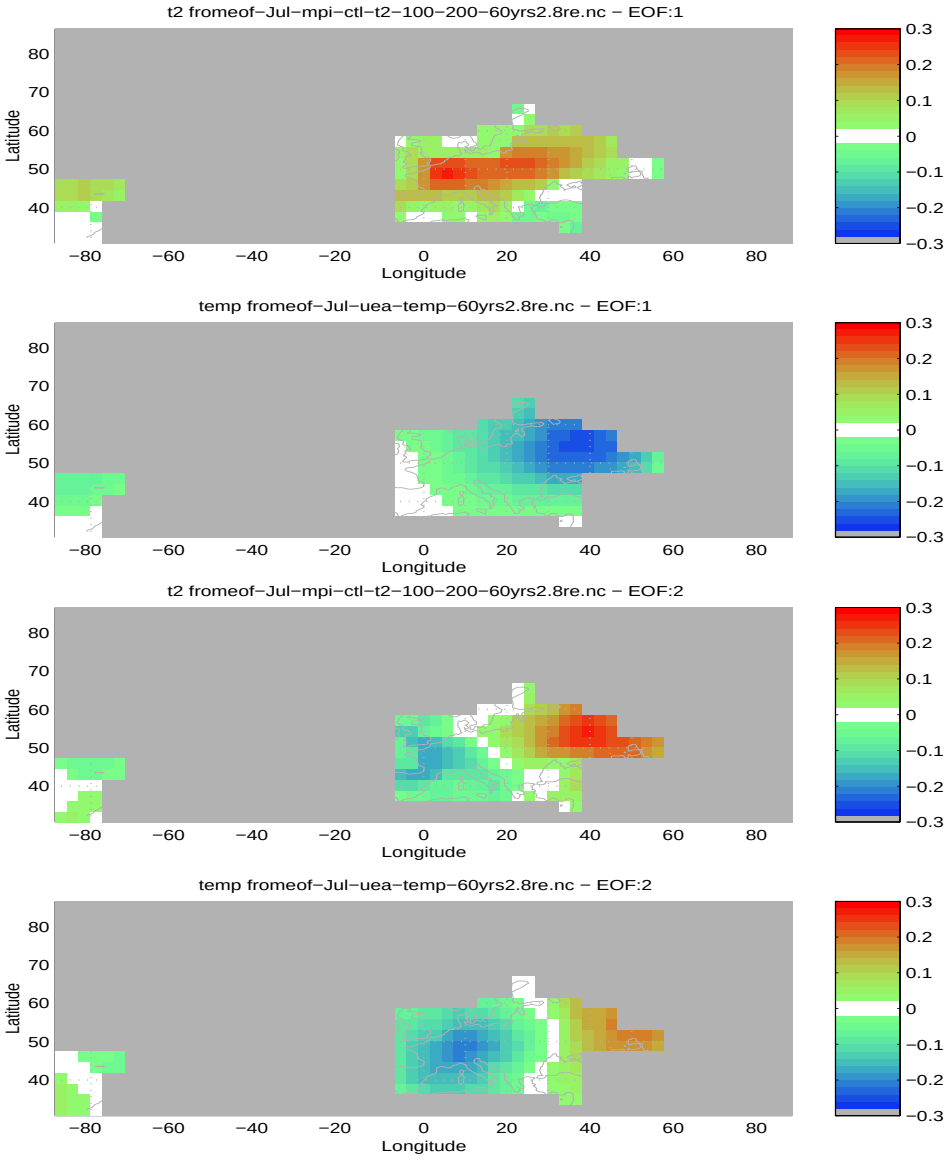

Figure 25: The 2 leading July EOFs of ECHAM4 CTL and UEA T(2m). The first and third panels are the two leading model EOFs and the second and fourth panels the corresponding observed EOFs.

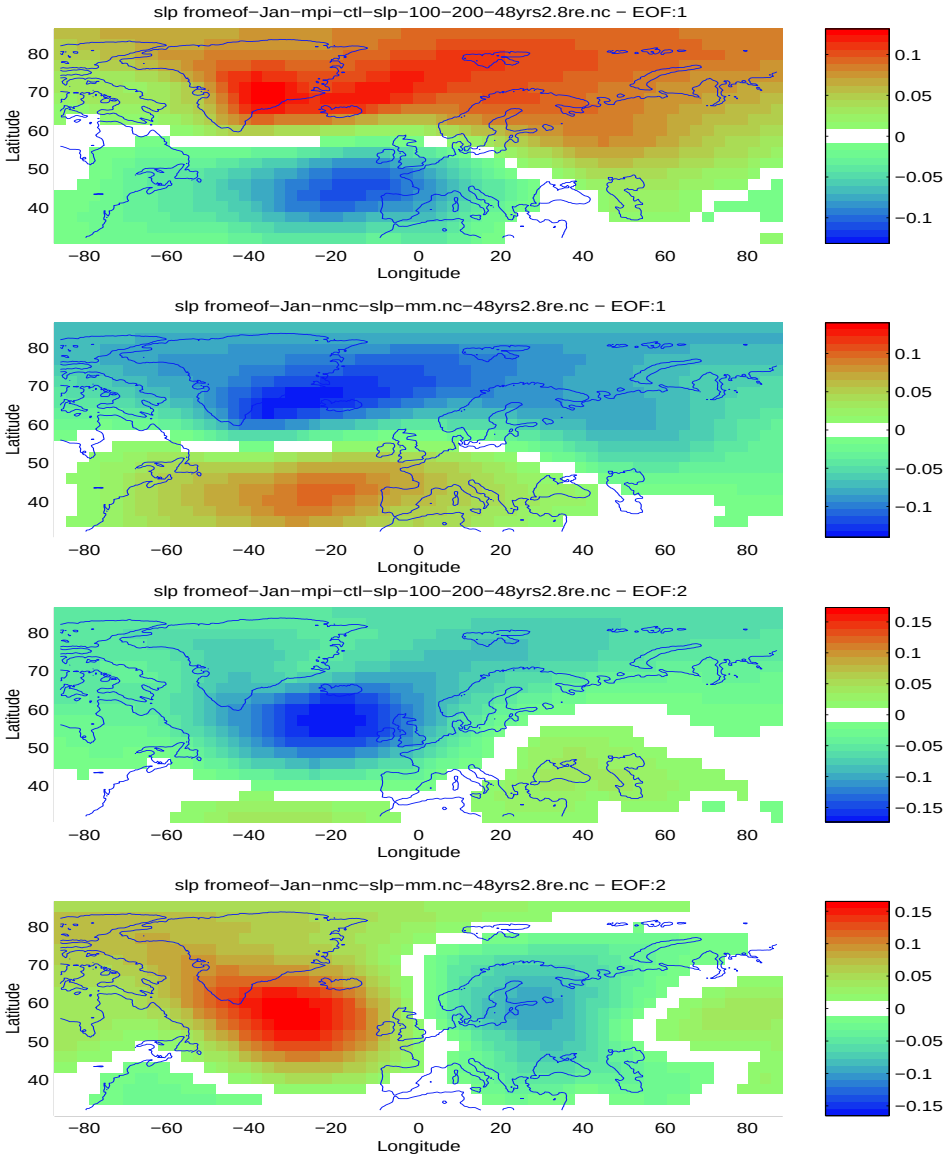

Figure 26: The 2 leading January EOFs of ECHAM4 CTL and NMC SLPs. The first and third panels are the two leading model EOFs and the second and fourth panels the corresponding observed EOFs.

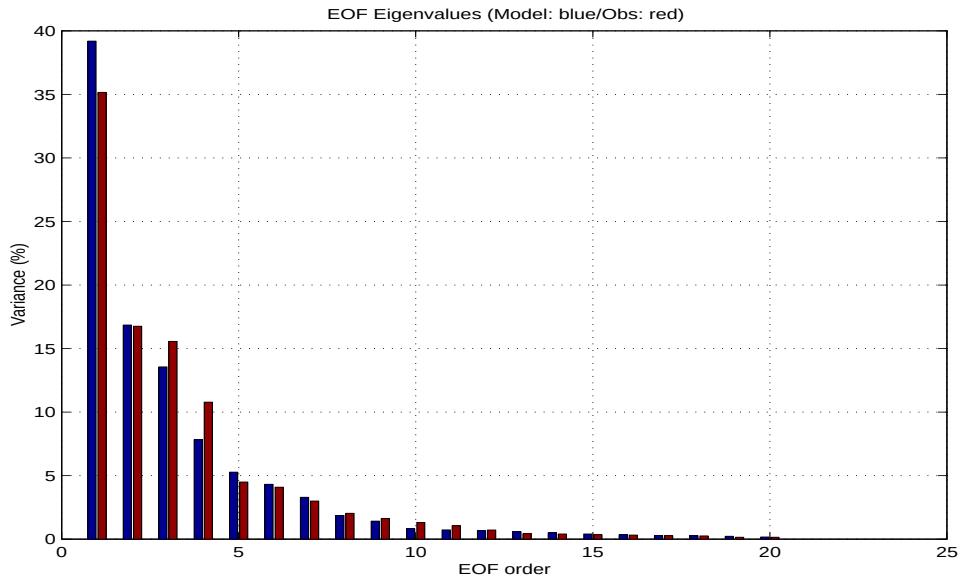

Figure 27: The eigenvalues of the 20 leading modes for MPI ECHAM4 CTL SLPs (first columns) and NMC SLP (second columns).

30°. The second and third NMC EOFs had similar eigenvalues, suggesting that they probably were degenerate (*North* et al., 1982) and that they together described a propagating wave feature. The North Atlantic storm track was seen in NMC EOF 4, whereas corresponding ECHAM4 EOF described a west-east dipole structure (not shown). The ordering of observed EOFs 2,3 and 4 shown here is arbitrary as these appeared to be degenerate.

The differences in the EOF structures can be related to the comparison between the standard deviations in section 5.2.2, and the NAO like structures in the leading EOFs are consistent with strongest variance over Iceland. The northwestward shift in the northern NAO dipole SLP maximum of the model results was in line with too large model standard deviation values seen over central Greenland in section 5.2.2.

The higher order eigenvalues were similar for the model results and the observations. The higher order EOF patterns (not shown) were increasingly more and more different, which partly may be due to orthogonality restrictions (All EOFs in the same set are orthogonal) and partly a result of the model's inability to give a good description of small scale details or the possibility that the higher order EOFs only described noise.

Model misrepresentations of EOF patterns may in general have consequences for the downscaling models, however, in this case the optimal January SLP NMC canonical correlation (CCA) and multivariate regression

(MVR) models used the EOF predictor combination 1, 5, 6, 11, 13, 15, 16, and 20 (*Benestad*, 1998a, 1999a), where the leading EOF represented a significant contribution (0.72 of 0.92) to the January temperatures at Lista and Oksøy fyr. Hence, the differences between the higher order model and observed EOFs were not critical for the January NMC SLP models (at Lista and Oksøy fyr).

6.3.2 Spectral Characteristics

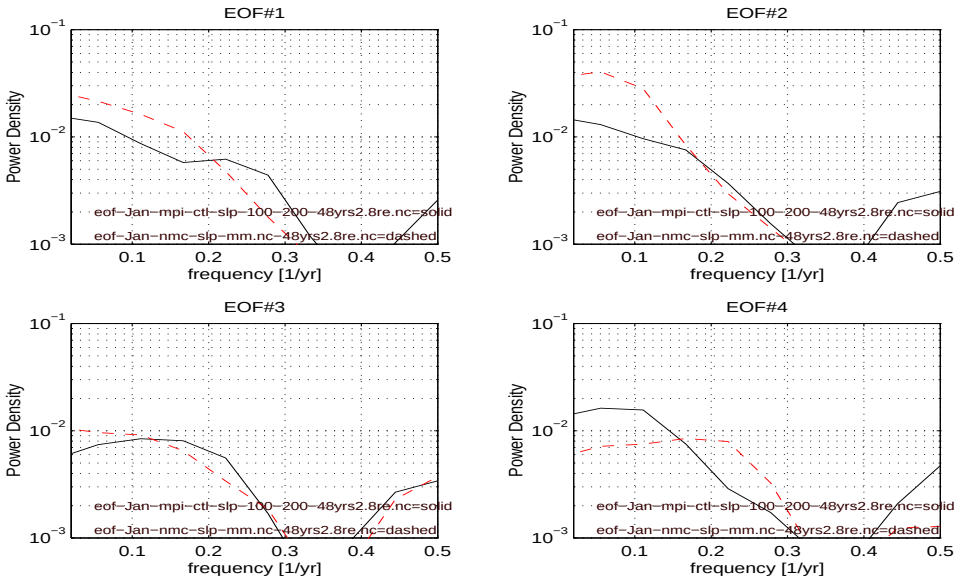

Figure 28: The power density spectra associated with the 4 leading January EOFs where the solid lines represent MPI ECHAM4 SLP and the dashed lines NMC SLP. The PCs had been smoothed with a 3-point Gaussian low-pass filter to remove the high frequency signal.

All the power spectra describing the time evolution of each EOF had a red noise character, but with various spectral peaks superimposed (figure 28). The leading observed SLP EOF (dashed line), for instance, had most power associated with low frequencies as expected for a red noise process, but the model EOF (solid line) also had a spectral peak in the frequency range of 3-5 years in addition to the red noise spectrum.

The second order EOF spectra did not exhibit characteristic time scales, and their red noise properties were far from similar.

The fact that the higher order model and observed EOFs were different may explain why the power spectra of the higher order EOFs did not match

each other. In summary, the comparison between the EOFs of ECHAM4 and the NMC SLPs revealed that the model captured SLP anomalies associated with the NAO realistically, although the NAO was too prominent in the model and was associated with too high frequencies.

6.3.3 Sampling fluctuations

A comparison was made between the SLP EOFs of different 48-year periods of the ECHAM4 CTL model integration (not shown), by examining three 48-year-long sequences from the CTL model years 152-200, 252-300, and 352-400.

The eigenvalues were not stationary on time scales shorter than 48 years. EOF analysis on different CTL 48-year long periods indicated that the variance of the leading EOF could vary by as much as 7%. The spatial EOF patterns for the third and higher order EOFs were substantially different, suggesting that the differences may have been due to sampling uncertainties rather than model misrepresentation.

The leading EOF from the 152-200 CTL year period described a red noise process which also displayed a 4-5 year oscillation whereas no such oscillation was evident in neither the 252-300 nor 352-400 year period. The power spectra for the CTL 252-300 and 352-400 year periods, however, were similar.

In summary, the comparison between different 48-year CTL periods indicated non-stationary behaviour in the model results, especially in the beginning of the CTL data record. Hence the discrepancies between the observed and model EOFs were not conclusive evidence of model misrepresentation.

6.4 July SLP

The leading model and observed July SLP EOFs described roughly similar features, but the leading model EOF showed strong anomalies west of Ireland whereas the corresponding observations described the North Atlantic storm track as a stretch of high variability running from the southwest to the northeast as far as the British isles. Coherent with the storm track in the observations was an extensive region with SLP anomalies of anti-correlated anomalies (figure 29).

The second model EOF (third from the top) contained traces of the North Atlantic storm track, but had also a zonally wide region of anomalies of opposite sense extending from Iceland over northern Scandinavia to northeastern Russia. The observations (bottom panel), on the other hand, indicated a NAO like structure.

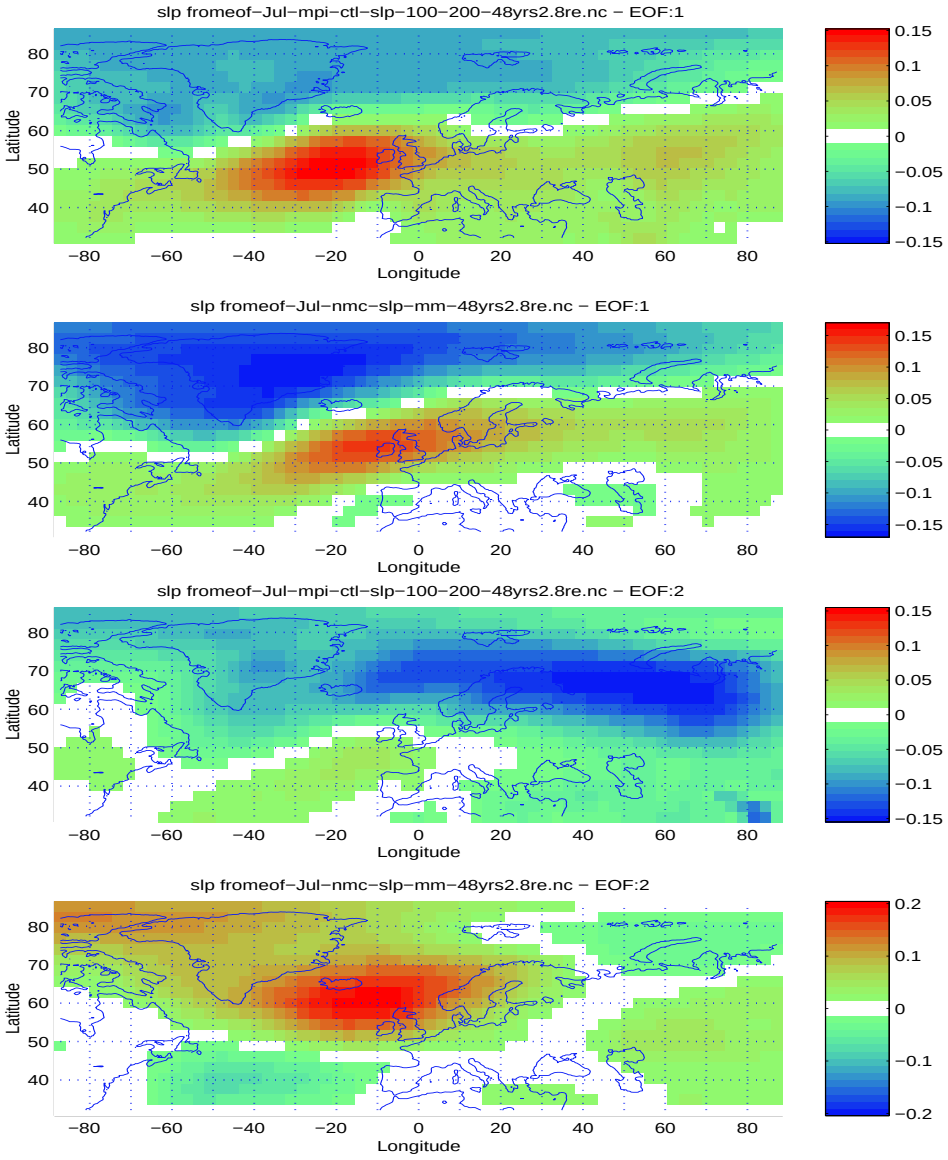

Figure 29: The 2 leading July EOFs of ECHAM4 CTL and NMC SLPs. The first and third panels are the two leading model EOFs and the second and fourth panels the corresponding observed EOFs.

Hints of a standing wave train pattern was seen in the third model and observed EOFs (not shown), with maxima of alternating polarity over Bermuda, the interior North Atlantic (50°N and 30°E), Scandinavia, and the Urals. The anomalies over the Ural range were weak in the model results.

The fourth EOFs (not shown) of model results and observations were roughly similar, although model described too much variability in the Arctic north of Russia. The observations indicated a tripole pattern with maxima over Labrador and central Europe and with opposite sign over Iceland.

The two leading July EOFs of both model results and observations accounted for about 23% and 16% respectively of the total variance. The model appeared to underestimate the importance of the third EOF as the model eigenvalue was about 10% compared to around 14% for the observations. The two leading EOFs may have been degenerate according to *North* et al. (1982), and it is therefore possible that some of the differences between the two leading EOFs may be due to this degeneracy.

The leading model EOF had a more prominent spectral peak associated with slightly higher frequency than the observations suggested. (time scale of about 6 years). The second model EOF resembled a red noise spectrum while observations had a prominent spectral peak at around 5 years and had no red noise properties. The third and fourth EOFs had vaguely similar spectral characteristics, although the third model EOF had a sharper spectral peak at around 6 years which was absent in the observations.

The fact that ECHAM4/OPYC3 described approximately similar SLP structures for both January and July suggests that the MPI model gave a good description of the annual cycle. The winter model anomalies were more similar to the corresponding observations than the summer anomalies, as the second order July EOFs were more different than the corresponding January EOFs.

6.5 January 500hPa geopotential heights

6.5.1 Variance and Spatial Structures

Figure 30 shows the results from a similar analysis as for figure 26, but applied to the January ECHAM4 and NMC Φ_{500} fields. Although the leading January Φ_{500} EOFs had similar overall structures, there were some important differences. The leading model EOF (upper panel) weights described a dipole pattern with centres of action over Greenland and over the North Atlantic, but the Greenland maximum in the observations (panel 2 from the top) was located to the south of the model maximum and the model anomalies over the North Atlantic were too strong. The Atlantic storm track, visible in the

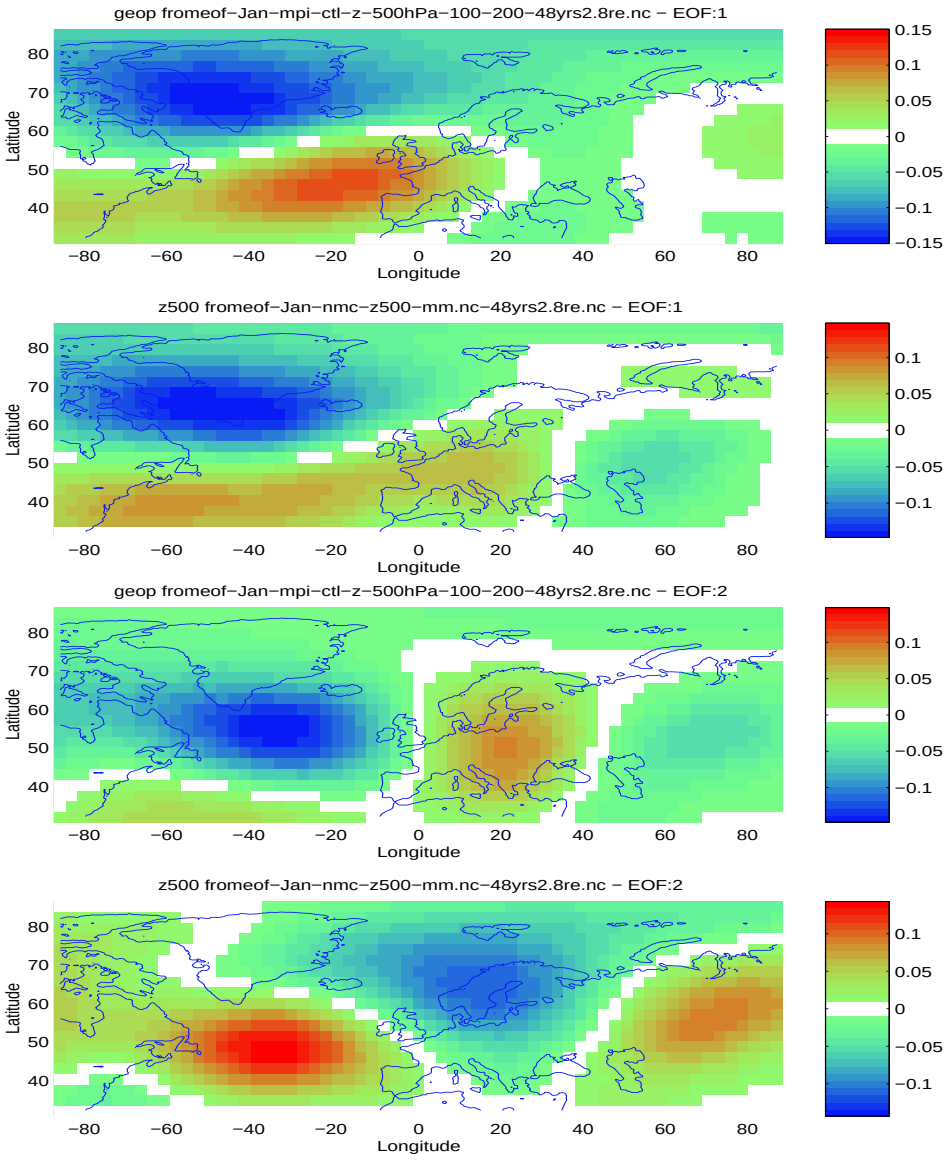

Figure 30: The 2 leading January Φ_{500} EOFs of ECHAM4 CTL and ECHAM4 CTL NMC 500hPa geopotential heights.

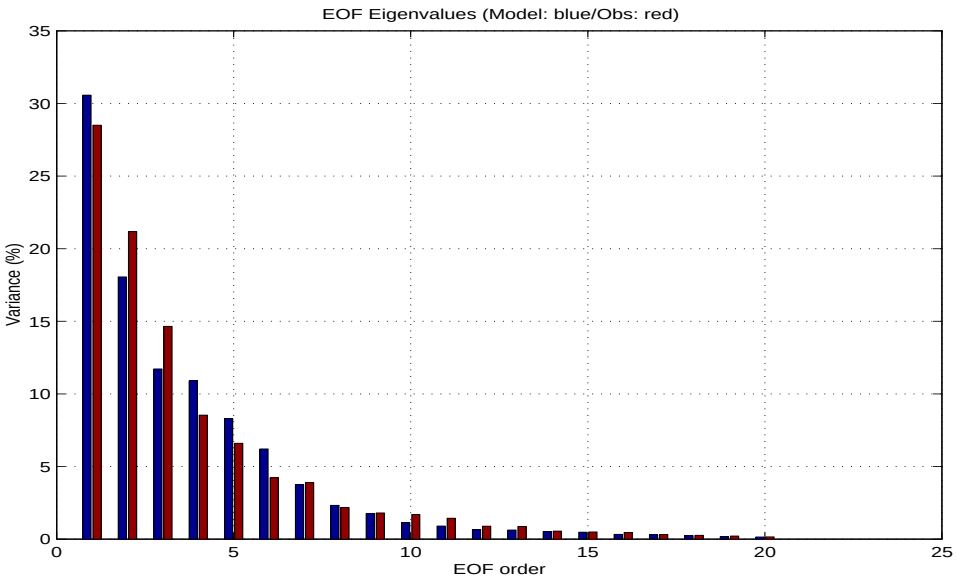

Figure 31: The eigenvalues of the 20 leading modes for MPI ECHAM4 CTL Φ_{500} (first columns) and NMC Φ_{500} (second columns). Window length=21 years, Parzen window

leading Φ_{500} EOFs as a long stretch of high variability (red colour in the two upper panels of figure 31), was seen too far north in the model results. Models in general tend to simulate too strong storm track activity (*Doblas-Reyes* et al., 1998), which is also the case here (over the North Atlantic). The model furthermore described too strong anomalies in the Arctic. Other important differences between the model and observed leading Φ_{500} EOFs were seen over the Kola peninsula, the Barents Sea and over Kazakhstan.

The two lower panels in figure 30 show the second model and observed January EOFs, and it is evident that their spatial structures bore some resemblance although there were some important differences. The model described stronger anomalies over the North Atlantic whereas the observations indicated strongest anomalies over Scandinavia.

The third EOFs (not shown) differed as there were strong anomalies over the Urals in the model and a maximum over the British Isles and the Irish Sea in the observations. The model fourth Φ_{500} EOF (not shown) described strong anomalies over the North Sea, and was substantially different to the corresponding NMC EOF characterised by strongest loadings over the Arctic region east of Greenland.

The regions with large standard deviation differences in section 5.3.2 corresponded with the locations where the 2 leading model and observed EOFs

indicated different weights. The results of the EOF evaluation were therefore consistent with the earlier analysis on the standard deviation.

Figure 31 shows how much of the total variance that the different January EOFs described. The leading model EOF was more prominent than the correpsonding observed EOF, whereas model may have underestimated the importance of the second and third order EOFs. The important EOFs for the statistical downscaling model based on Φ_{500} were number 1, 4, 9, 11, 13, 14, 15, and 19. The leading EOF contributed with 0.75 to the total correlation score of 0.91 at Hellisøy fyr (*Benestad*, 1998a, table 21, p. 64), suggesting that the leading EOF was more important for the local temperatures than the higher order EOFs.

6.5.2 Spectral Characteristics

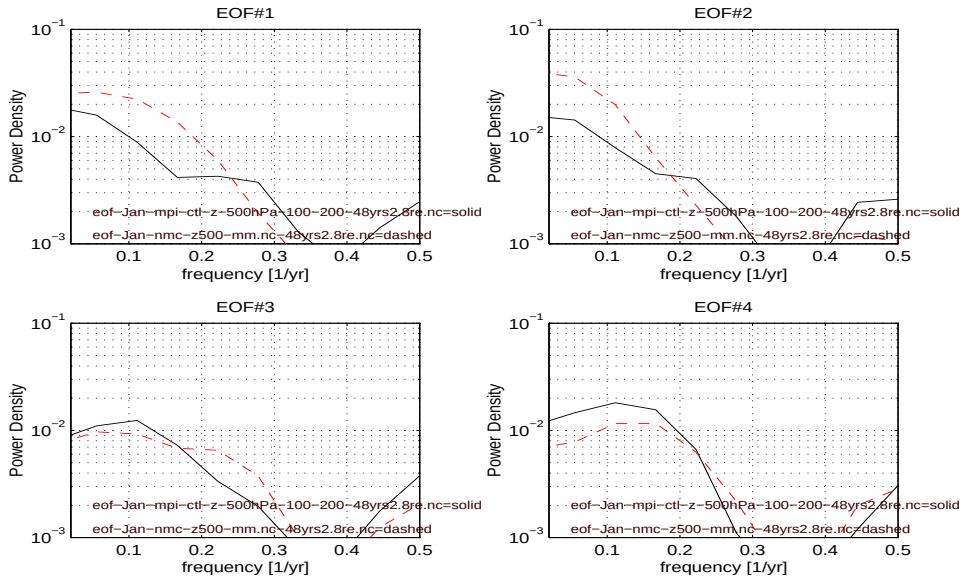

Figure 32: The power density spectra associated with the 4 leading January EOFs where the solid lines represent MPI ECHAM4 Φ_{500} and the dashed lines NMC Φ_{500}. The PCs had been smoothed with a 3-point Gaussian low-pass filter to remove the high frequency signal.

The spectral analysis of the January Φ_{500} PCs, shown in figure 32, indicates that the leading observed EOFs (dashed line) could be described as red noise processes whereas the model results also described oscillations at a time scale of around 3-4 years for the leading and 4-5 years for the second EOFs. The third and fourth power spectra were roughly similar, and in both cases the

modal structures could be described red noise processes. The third model EOF, however, appeared to have a preferred time scale of around 10 years on top of the red noise spectrum.

6.5.3 Sampling fluctuations

As for the comparison between different 48-year periods for the SLPs, the 4-5 year frequency peak of the Φ_{500} PCs was only evident in the 152-200 CTL year sequence. The two leading EOFs were roughly similar for the three periods examined, whereas the higher order ones were substantially different. The eigenvalues, however, were roughly similar, suggesting an uncertainty in the variance of the leading EOF by approximately 2%.

6.6 July 500hPa geopotential heights

Both the leading July EOFs (not shown) in model results and the observations described prominent ϕ_{500} anomalies west of Ireland, although the model did not capture a secondary maxima of similar sign over the Urals. The anomalies in the observations were aligned from the southwest to the northeast and may be connected with the North Atlantic storm track whereas no such structure was evident in the model.

The second model EOF (not shown) showed stretches of high variance running from the southwest to the northeast while the corresponding observations described strong anomalies over southern Scandinavia embedded inside a large scale anomaly of opposite polarity. This structure did resemble the wave train structure in the leading observed EOF, and it is therefore possible that the second model EOF may account for some of the differences between the leading model and observed EOFs.

In the third EOF (not shown), the model results exhibited strong anomalies over the Baltic region whereas observations indicated a prominent west-east dipole with centres of action south of Iceland and the Kola peninsula.

The fourth model and observed EOFs (not shown) were different, as the model described anomalies over southwestern Russia and weak opposite anomalies over the North Sea while the observations revealed a prominent west-east dipole with maxima over the North Atlantic and central Europe.

The 20 leading model and observed July EOFs had similar eigenvalues and a spectral analysis suggested that the leading model summer EOF had features that oscillated at a higher frequency (time scale of about 6 years) than was evident in the observations (time scale $\approx$ 9 years). The second model EOF exhibited a spectral peak at about 4-5 years while the observations had a characteristic time scale of about 10 years. The power spectra for

higher order EOFs were different, which may be a result of the corresponding EOFs being different.

6.7 January 500hPa temperatures

6.7.1 Variance and Spatial Structures

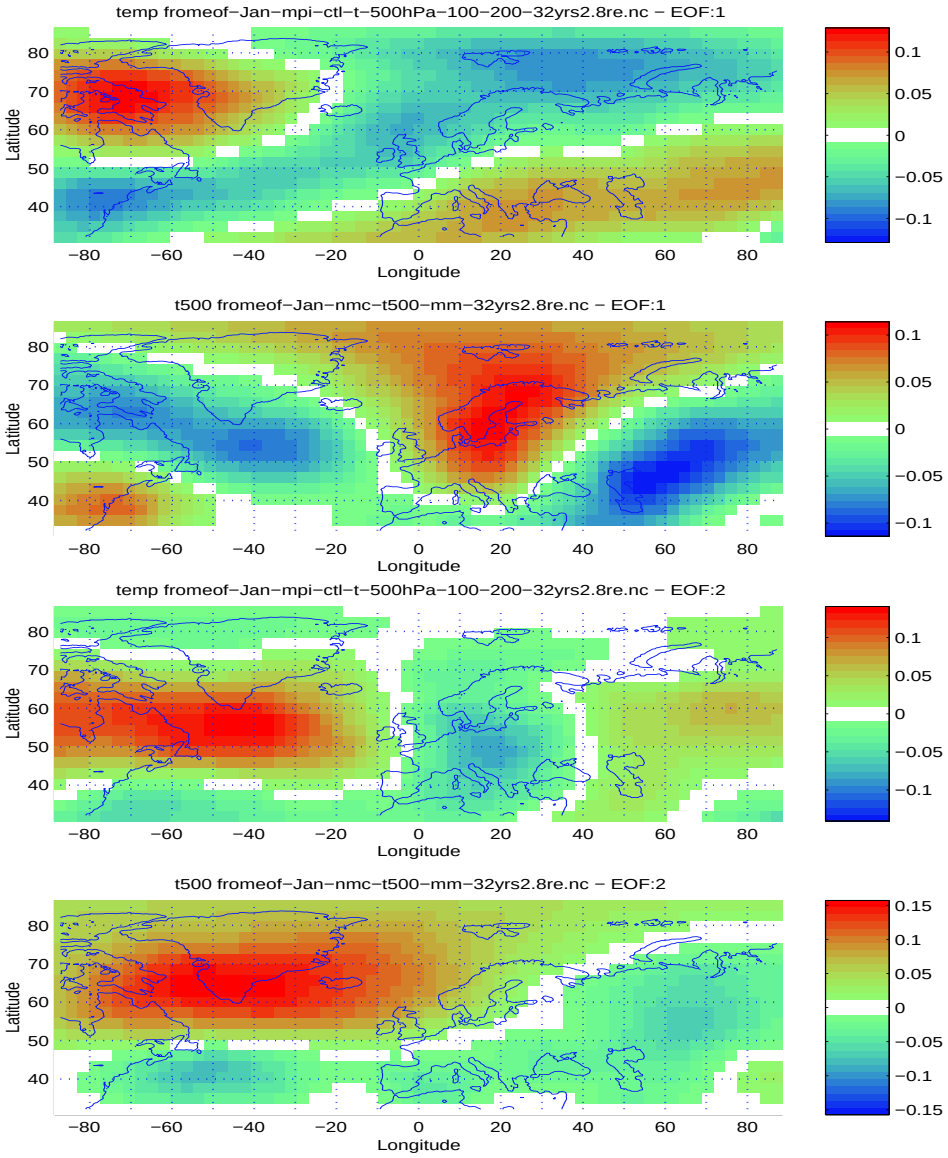

Figure 33: The 2 leading January T_{500} EOFs of ECHAM4 CTL and ECHAM4 CTL NMC.

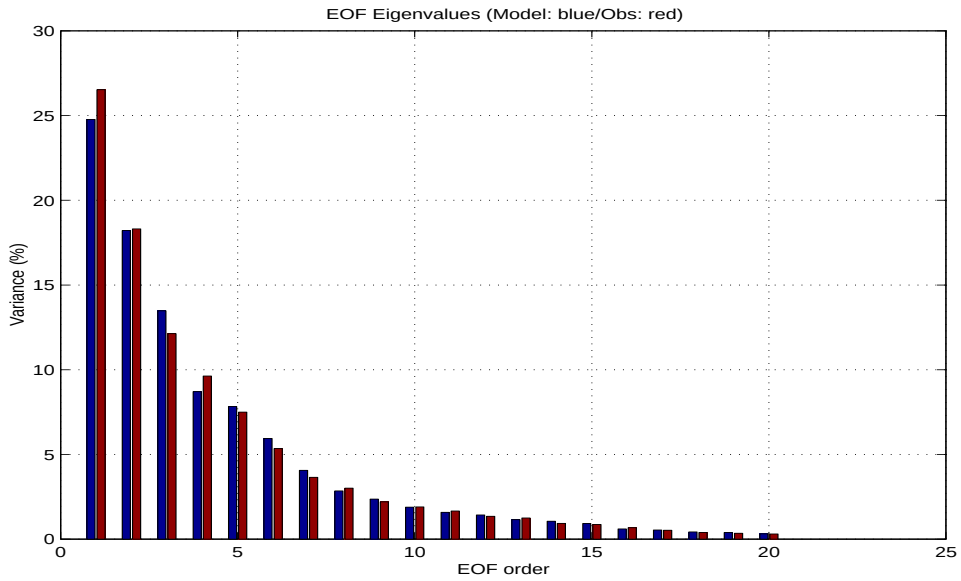

Figure 34: The eigenvalues of the 20 leading EOFs for MPI ECHAM4 CTL T_{500} (first columns) and NMC T_{500} (second columns) fields.

The leading model and observed January T_{500} EOFs in figure 33 (top panel) described distinct spatial structures, although both indicated strong anomalies over the Arctic region. The model EOF (top panel) described a extensive spatial structure stretching from the southwest to the northeast, sandwiched between two regions with anomalies of opposite sign. The leading NMC EOF (second from the top), on the other hand, described a stationary wave pattern with maximum anomalies south of Greenland, the Baltic and Kazakhstan.

The second January EOFs of the model results (third from the top) and observations (bottom panel) both were characterised by a north-south dipole pattern over the western North Atlantic basin. The model EOF, however, also gave indications of a west-east dipole structure with centres of action over the North Atlantic and eastern Europe.

The third model and observed EOFs had a stretch of prominent anomalies running from the southwest to the northeast across the North Atlantic (not shown). Both EOFs also displayed strong loadings over Russia of similar sign to the anomalies across the North Atlantic and over the Labrador Sea and eastern Europe of opposite polarity.

The fourth model and observed EOFs both described prominent anomalies of similar polarity over northern Europe and the Hudson Bay, although the model maxima were seen too far west (not shown). The observations furthermore described prominent anomalies of opposite sign over eastern Green-

land which were weak or absent in the model results.

The differences between the model and observed T_{500} variability was studied by plotting the differences in their respective standard deviations in section 5.4.2. There were large discrepancies over the Norwegian-Iceland Sea, northern Russia, the Persia Gulf, Ukraine, which bore little resemblance to the the difference between the 2 leading EOF structures. This observation, however, does not imply that the EOF analysis is inconsistent with the standard deviation fields, but the T_{500} may be subject to prominent small scale variability.

The leading model EOF described a slightly smaller proportion of the total variance than the corresponding observational analysis (figure 34). In general, however, the model and observed EOFs accounted for similar amounts of variance. The optimal T_{500} downscaling model for the Norweagian January temperatures was based on the EOF combination 1, 2, 3, 5, 6, 7, 9, 11, 12, 14, and 15.

6.7.2 Spectral Characteristics

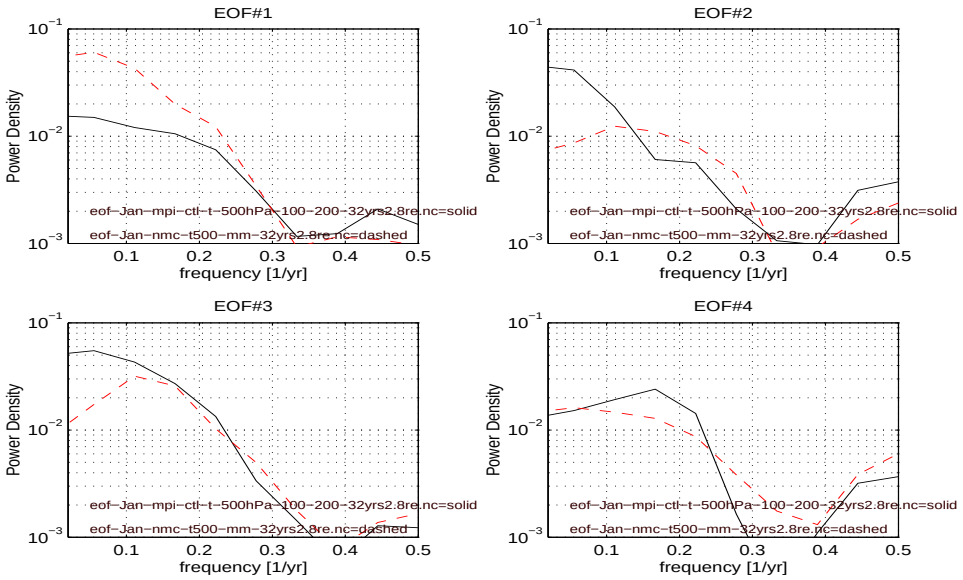

Figure 35: The power density spectra associated with the 4 leading January EOFs where the solid lines represent MPI ECHAM4 T_{500} and the dashed lines NMC T_{500}. The PCs had been smoothed with a 3-point Gaussian low-pass filter to remove the high frequency signal.

Figure 35 shows the power spectra of the 4 leading January PCs. Both leading EOF structures described a weak oscillation with time scales of about

4-5 years on top of the red noise although the observed spectrum had a more prominent low frequency variability. The second model EOF was associated with too much low frequency variability and a strong red noise character. The corresponding observed EOF, on the other hand, described oscillations with preferred time scales of 5-10 years.

The third EOFs were associated with different spectral properties, where the model indicated a red noise type process whereas the observations indicated a distinct spectral peak at 6-10 years. The fourth model EOF, however, had a periodicity of around 5 years, while the fourth observational EOF indicated a stronger red noise character.

The differences in the power spectra may reflect the differences in the spatial structures of these EOFs. It is possible that the short time intervals examined and sampling fluctuations may have influenced the EOF order, as the spectral power spectrum of the leading observed EOF was more similar to the second model EOF's power spectrum. Furthermore, swapping the order of model EOFs 3 and 4 may also give a better agreement between the power spectra.

6.7.3 Sampling fluctuations

The red noise character of the leading model EOFs for different intervals were substantially different, suggesting that 32 years is too short for model evaluation (figure 36). Furthermore, the leading EOF from the CTL years 368-400 showed a spatial structure that vaguely resembled the wave pattern in the observations. The leading spatial EOF patterns showed completely different features between CTL years 168-200 and 368-400, and the variance accounted for by the leading EOF varied by approximately 5% between the different intervals.

6.8 July 500hPa temperatures

Both leading July EOFs of the model results and observations (not shown) described two southwest-northeast running stretches of opposite polarity across the North Atlantic, but the anomalies in the model were located further north.

The second model and observed EOFs (not shown) were characterised by a southwest-northeast aligned structure of strong variability, most likely connected with the North Atlantic storm track. The model simulated too strong storm track anomalies over the North Atlantic.

The comparison between the third EOFs (not shown) suggested that the model described too strong anomalies over Ireland and too weak variability to

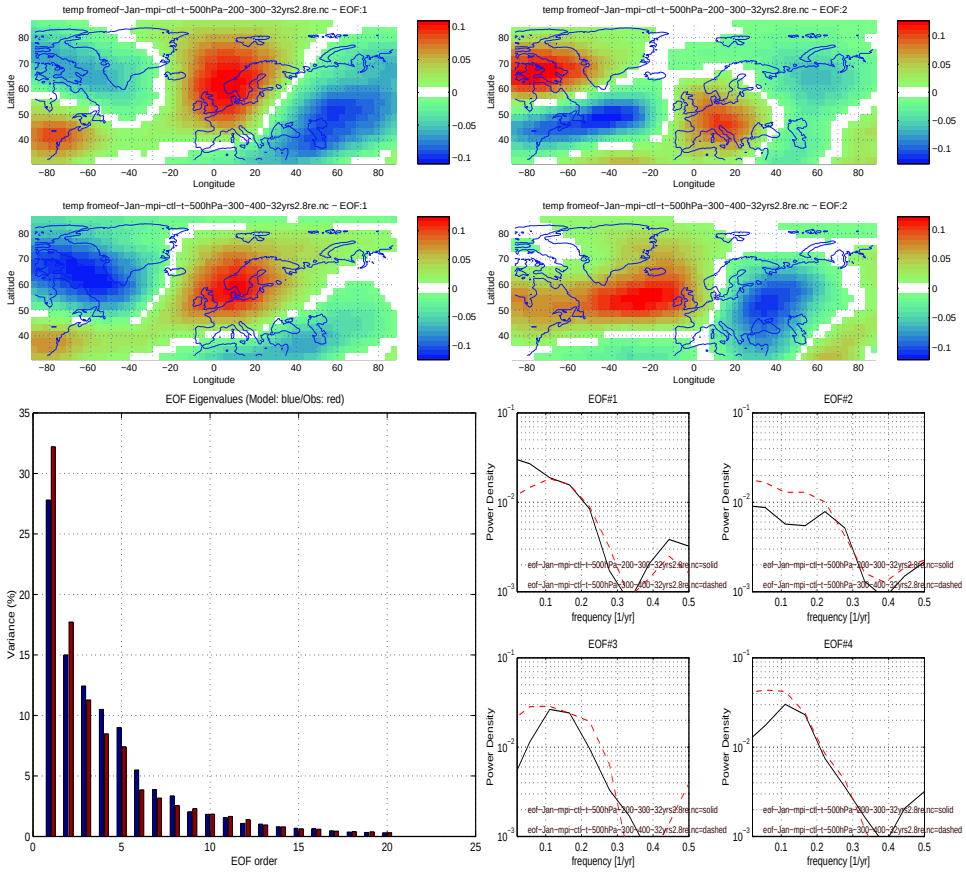

Figure 36: The two leading January EOF pairs from CTL years 268-300 (top) and 368-400 (second from the top) and corresponding eigenvalues (lower left) and power spectra (lower right).

the east of the British isles as the corresponding anomalies in the observations were located over the North Sea.

The fourth model EOF (not shown) had many small scale structures whereas the observations indicated strong anomalies over the Barents Sea and weak anomalies with opposite sign to the south.

The leading model EOF contributed to more of the total variance than corresponding observed EOF. The first and second leading observed EOFs, however, were degenerate according to *North* et al. (1982)'s criterion. The model overestimated the importance of third and fourth EOFs with respect to variance.

The power spectra of the principal components suggested that the leading observed July EOF was associated with a weak spectral peak of around 5 years that was absent in the model results. The second and third order EOFs had largely similar spectral properties.

6.9 January SST

6.9.1 Variance and Spatial Structures

It is important that the model gives a realistic description of the SSTs since the SSTs are used to determine the heat fluxes between the atmosphere and oceans in the coupled model. Large discrepancies in the SST statistics may indicate that there was a problem related to the coupling. Furthermore, if the SSTs are used as boundary values for regional modelling (for instance time slice integrations or statistical downscaling), then misrepresentations in the SSTs can give misleading results. The two leading January SST EOFs from the model results and observations are shown in figure 37, and it is apparent that the OPYC3 model description of the January SST anomalies for the control run was associated with serious errors in the Nordic Seas, despite some similarities in the North Atlantic large scale structures.

The leading model EOF (upper panel) indicated much too strong SST variability close to the coast of Newfoundland, and the SST anomalies were associated with too small spatial scales. *Kushnir & Held* (1996) argued that the spatial scale of the SST anomalies is important for the degree of atmospheric response to SSTs. The leading model EOF also produced too high variance in the Barents Sea, North Sea, and Kattegat. In contrast to the observations, the leading model EOF accounted for too much of the total variance compared to the second EOF.

The amount of variance described by the January EOFs is shown in figure 38, and the two leading GISST2.2 EOFs accounted for similar amounts of variance, and were degenerate according to *North* et al. (1982). The second

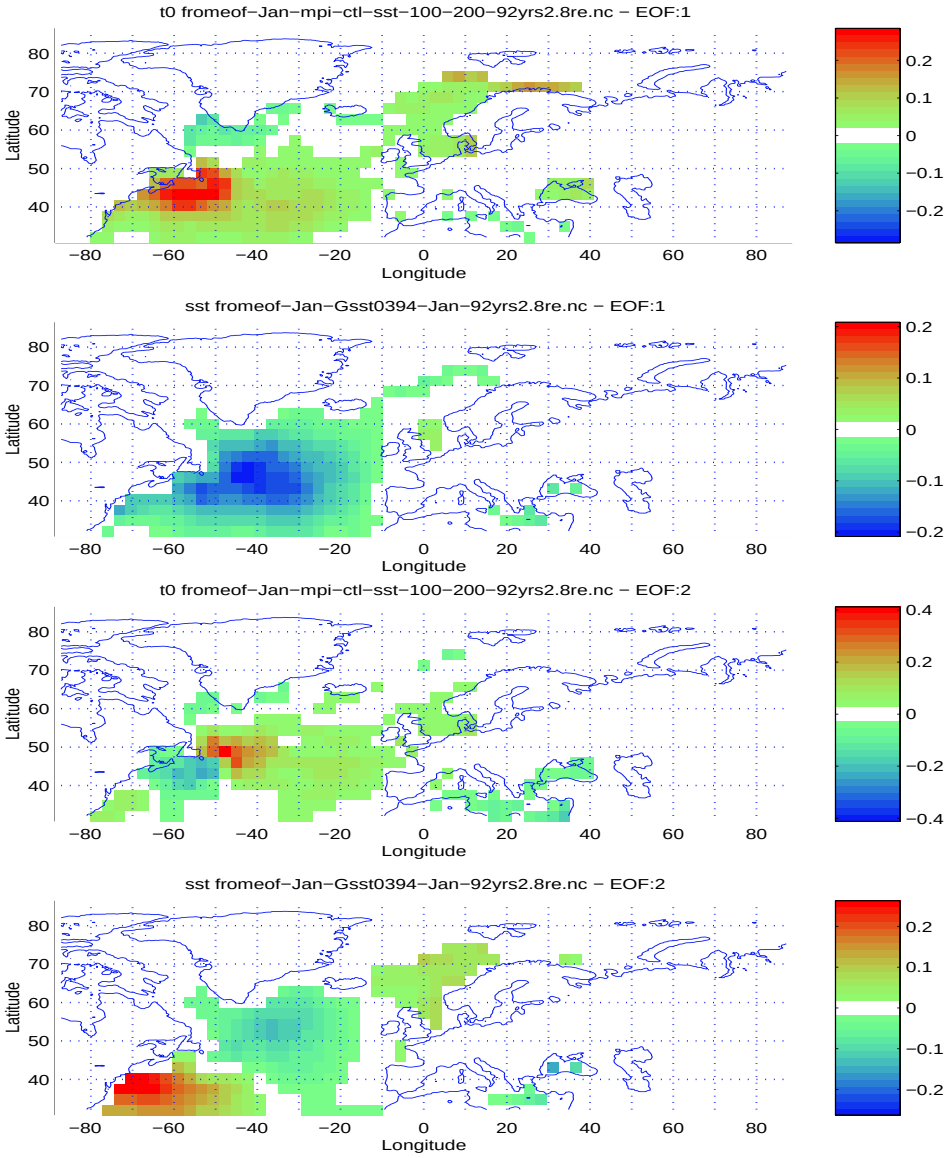

Figure 37: The 2 leading January EOFs of GISST2.2 SSTs and ECHAM4 CTL.

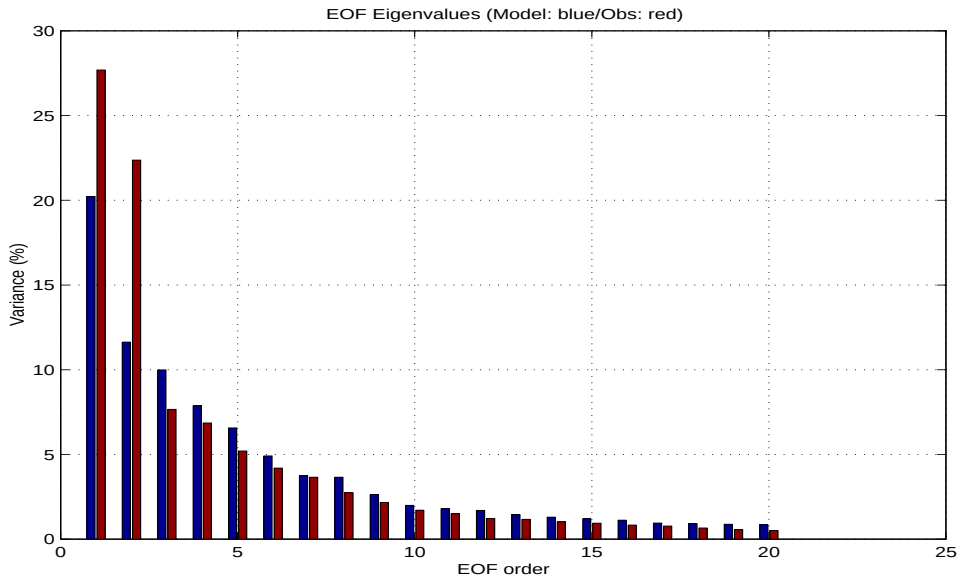

Figure 38: The eigenvalues of the 20 leading modes for MPI ECHAM4 CTL SSTs (first columns) and GISST2.2 SSTs (second columns).

model EOF represented anomalies with too small spatial scales, and had a similar eigenvalue to the third model EOF suggesting that they may be degenerate.

The third model SST EOF (not shown) described strong anomalies east of Newfoundland and south of Greenland, while the observations in contrast, exhibited a wave like structure running along the western boundary with maxima of alternating polarities off New England, Newfoundland and Greenland-Iceland Sea. The fourth model and observed EOFs contrasted by the fact that the observations were dominated by a northwest-southeast dipole structure whereas the model EOF described apparently noisy small scale structures (not shown).

The large differences in the standard deviation fields along the Newfoundland coast in section 5.5.2 were consistent with the results from the EOF comparison.

The predictors of the optimal January SST statistical downscaling model included EOFs 1, 2, 8, 9, 10, 11, 14, and 17. The leading January EOF of the North Atlantic SSTs accounted for 80% (0.60) of the total correlation score (0.75) at station Flisa, and hence the higher order EOFs were not of critical importance for the statistical downscaling model (for Flisa).

6.9.2 Spectral Characteristics

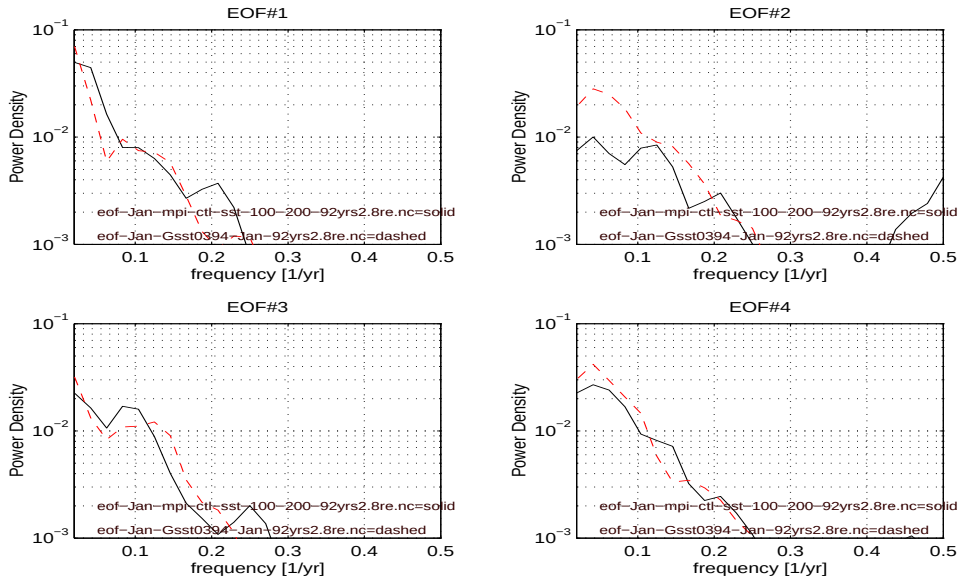

Figure 39: The power density spectra associated with the 4 leading January EOFs where the solid lines represent MPI ECHAM4 SLP and the dashed lines GISST2.2 SSTs. The PCs had been smoothed with a 3-point Gaussian low-pass filter to remove the high frequency signal.

The power spectra associated with the SST structures described by the 4 leading January EOFs are shown in figure 39. Both leading model and observed EOFs described spatial structures which were associated with oscillations with a time scale of approximately 10 years, although the model SSTs also oscillated with a time scale of approximately 5 years. The second order model EOF was associated with too weak variability at low frequencies, and had preferred time scales of around 8 years and 5 years.

6.9.3 Sampling fluctuations

The leading EOF of the different periods were similar (not shown), describing strong variability just east off Newfoundland.

The second EOF of CTL years 108-200 indicated little variability in the Kattegat compared to the second EOFs for the 208-300 and 308-400 year periods. The strong CTL 208-300 year small scale anomaly (one grid box) in Kattegat appeared to be coherent with a large scale northeast-southwest dipole structure stretching along the western boundary, with similar polarity in the southwest. In 308-400 CTL results, on the other hand, the Kattegat

anomaly was coupled with small scale anomalies east of Newfoundland whose maximum weights with similar polarity as in Kattegat were seen north-east. The anomalies during the CTL years 308-400 in the western basin of the North Atlantic ocean were similar to those in 108-200 year period. The inconsistency in the relationship between SSTAs in Kattegat and the North Atlantic from period to period is an indication that there is no robust relationship between the local SSTAs in the Nordic seas and the large scale anomalies in the North Atlantic.

The third and fourth EOF structures in the different periods were not robust as the different periods gave different EOF patterns (not shown). For the third EOF, the 108-200 model year period had strong northeast-southwest dipole structure along North America's east coast and in the Labrador sea, whereas the 208-300 period produced strong weights near Kattegat and southeastern Mediterranean, and the 308-400 model years only in the southeastern Mediterranean.

The fourth EOFs were generally characterised by strong small structure anomalies. During the model years 108-200, strong small scale anomalies were seen near Newfoundland and Southeast Mediterranean. The 208-300 CTL year interval gave large values in two grid boxes east of Newfoundland, whereas the 308-400 period gave strong weights in Kattegat and southeast Mediterranean. The eigenvalues were of the same orders of magnitude although the leading eigenvalue varied by as much as 2.5%. Similar power spectra were found for the two leading EOFs regardless of period chosen, although the leading EOF for the 108-200 and 208-300 CTL year periods only gave red noise spectra while the power spectrum for the leading EOF for 302-400 CTL years suggested a weak 5-year oscillation as well. The third and higher order EOFs were different.

6.10 July SST

Model results indicated strong anomalies in the leading July EOF in shape of a tongue confined between 40°N and 50°N and extending from the west to around 10°E (not shown). Significant anomalies were also found near the ice edge in the Polar Sea. The observations described prominent SST anomalies in the North Atlantic interior although only between 60°W and 20°W, but gave no indications of strong variance in the Arctic.

The second model EOF for the July month (not shown) had strong anomalies in the vicinity of the Barents Sea and White Sea while the observations described maximum variability in the central North Atlantic.

The third model summer time EOF described too strong variability in the Barents Sea and White Sea region whereas the observations indicated

prominent large scale SSTAs in the North Atlantic in the shape of a west-east dipole structure (not shown).

The model described large anomalies near the ice edge in the fourth EOF. The observations, on the other hand, had a north-south dipole in the North Atlantic south of Greenland.

The model underestimated the importance of 5 leading EOFs, and the leading model EOF accounted for only 14% whereas the leading observed SSTs EOF described 29% (not shown). The three leading EOFs had similar power spectra, however, and all had spectral peaks with $\tau \approx 7$ years on top of red noise characteristics.

7 Comparison of gridpoint values

7.1 Temperature and sea level pressure at selected grid points

7.2 Surface temperatures

7.3 SLP

8 Conclusions

The main conclusions of the ECHAM4/OPYC3 model evaluation can be summarised as follows:

- Climate records shorter than 50 years are really too short for model evaluation or comparison due to possible non-stationarity as a result of very low frequency variability.
- Important differences in the mean SLP fields over the Arctic may have implications for the geostrophic wind over the Svalbard and the Barents Sea.
- The model gave a reasonable description of the mean 500hPa geopotential heights, but the 500hPa temperature anomaly structures may potentially suffer from serious model misrepresentation.
- ECHAM4/OPYC3 has serious problems with the description of SSTs, as the model SSTs were associated with too small spatial scales in the North Atlantic and too much variability in the Nordic Seas.
- The differences between the EOFs of the model results and observations present a problem for the downscaling of the model data to local climate variability. Some of these problems can partly be overcome by regressing the spatial model EOF patterns onto those of the observations.
- The MPI model gave a good description of the seasonal variability.
- Possible spin-up biases still present in the 100-200 model years. Don't use this period for analysis or prediction. Same probably applies to GHG.
- Oscillating features simulated by the model often had different preferential time scales to those in the observations.
- Indication of misrepresentation of upper air diagnostics: both mean values and EOFs of 500hPa temperatures appeared to be inconsistent with observations/expectations. One problem, however, is that observations of high altitude may not be very good.
- Some of the errors may be reduced by subtracting the CTL from the GHG results if these errors apply equally to both integrations. Question if these errors are affected by changes in the mean climate due to CO_2 forcing.

- Predictors extending over too small regions in downscaling of the GCMs may give misleading results as the EOF analysis indicated that the climatic features often were displaced.

- Downscaling should be based on SSTs from the entire North Atlantic ocean and not just from the Nordic seas where there are large systematic errors.

References

Appenzeller, C., Stocker, T.F., & Anklin, M. 1998. North Atlantic Oscillation Dynamics Recorded in Greenland Ice Cores. *Science*, **282**, 446–449.

Benestad, R.E. 1998a. *CCA applied to Statistical Downscaling for Prediction of Monthly Mean Land Surface Temperatures: Model Documentation.* Klima 28/98. DNMI.

Benestad, R.E. 1998b. *Description and Evaluation of the Predictor Data sets used for Statistical Downscaling in the RegClim.* Klima 24/98. DNMI.

Benestad, R.E. 1999a. *MVR applied to Statistical Downscaling for Prediction of Monthly Mean Land Surface Temperatures: Model Documentation.* Klima 2/99. DNMI.

Benestad, R.E. 1999b. *S-mode and T-mode EOFs from a GCM modeller's perspective: Notes on the linear algebra.* Tech. rept. DNMI.

Bengtsson, L. 1996. The Climate response to the Changing Greenhouse Gas Concentration in the Atmosphere. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability.* NATO ASI series, vol. 44. Springer.

Christensen, O.B., J.H., Christensen, Machenhauer, B., & Botzet, M. 1998. Very High-Resolution Climate Simulations over Scandinavia - Present Climate. *Journal of Climate*, **11**, 3204–3229.

Delworth, T., Manabe, S., & Stouffer, R.J. 1993. Interdecadal Variations of the Thermohaline Circulation in a Coupled Ocean-Atmosphere Model. *Journal of Climate*, **6**, 1993–2011.

Deser, C., & Blackmon, M.L. 1993. Surface climate variations over the North Atlantic ocean during winter: 1900-1989. *Journal of Climate*, **6**, 1743–1753.

Doblas-Reyes, F.J., Déqué, M., Valero, F., & Stephenson, D.B. 1998. North Atlantic wintertime intraseasonal variability and its senaitivity to GCM horizontal resolution. *Tellus*, **50A**, 573–595.

Ghil, M., & Yiou, P. 1996. Spectral Methods: What they Can and Cannot Do for Climate Times Series. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability.* NATO ASI series, vol. 44. Springer.

Grötzner, A., Latif, M., & Barnett, T.P. 1998. A Decadal Climate Cycle in the North Atlantic Ocean as Simulated by the ECHO Coupled GCM. *Journal of Climate*, **11**, 831–847.

IPCC. 1990. Climate Change: The Scientific Assessment. IPCC. Camebridge University Press.

Jones, P. D., Raper, S. C. B., Bradley, R. S., Diaz, H. F., Kelly, P. M., & Wigley, T. M. L. 1998. Northern Hemisphere surface air temperature variations, 1851–1984. *J. Clim. Appl. Met.*, **25**, 161–179.

Jones, P.D. 1987. The early twentieth century Arctic high - fact or fiction? *Climate Dynamics*, **1**, 63–75.

Kushnir, Y., & Held, I.M. 1996. Equilibrium Atmospheric Response to North Atlantic SST Anomalies. *Journal of Climate*, **9**, 1208–1220.

Latif, M. 1998. Dynamics of Interdecadal Variability in Coupled Ocean-Atmosphere Models. *Journal of Climate*, **11**, 602–624.

Lau, K.-M., & Weng, H. 1995. Climate signal detection using wavelet transform: How to make a time series sing. *Bull. Amer. Meteor. Soc.*, **76**, 2391–2402.

Machenhauer, B., Windelband, M., Botzet, M., Christensen, J.H., Déqué, M., Jones, R.G., Ruti, P.M., & Visconti, G. 1998. *Validation and Analysis of Regional Present-day Climate and Climate Change Simulations over Europe*. Tech. rept. 275. Max Planck-Institute für Meteorlogie.

Marotzke, J. 1996. Analysis of Thermohaline Feedbacks. *In:* Anderson, D.L.T., & Willebrand, J. (eds), *Decadal Variability*. NATO ASI series, vol. 44. Springer.

North, G.R., Bell, T.L., & Cahalan, R.F. 1982. Sampling Errors in the Estimation of Emperical Orthogonal Functions. *Monthly Weather Review*, **110**, 699–706.

Oberhüber, J.M. 1993. Simulation of the Atlantic circulation with a coupled sea ice-mixed layer isopycnal general circulation model. Part 1: Model description. *Journal of Physical Oceanography*, **22**, 808–829.

Peixoto, J.P., & Oort, A.H. 1992. *Physics of Climate*. AIP.

Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T. 1989. *Numerical Recepies in Pascal*. Cambridge University Press.

Roeckner, E., & Coauthors. 1996. *The atmospheric general circulation model ECHAM4: Impact of model physics and resolution.* Tech. rept. 93. Max Planck-Institute für Meteorlogie.

Schlesinger, M.E., & Ramankutty, N. 1994. An Oscillation in the global climate system of period 65-70 years. *Nature*, **367**, 723–726.

Sutton, R.T., & Allen, M.R. 1997. Decadal predictability of North Atlantic sea surface temperature and climate. *Nature*, **388**, 563–567.

Thompson, W.J., & Wallace, J.W. 1998. ? *Geophys. Res. Lett.*, **25**, 1297–1300.

Timmermann, A., Latif, M., Voss, R., & Grötzner, A. 1998. Northern Hemispheric Interdecadal Variability: A Coupled Air-Sea Mode. *Journal of Climate*, **11**, 1906–1931.

Wilks, D.S. 1995. *Statistical Methods in the Atmospheric Sciences.* Orlando, Florida, USA: Academic Press.

9 Appendix