Geophysica (1994), 30, 1-2, 3-30

A Comparison of the Results of Seven GCM Experiments in Northern Europe

Jouni Räisänen

Department of Meteorology, University of Helsinki
P.O. Box 4 (Hallituskatu 11), FIN-00014 University of Helsinki, Finland

(Received: April 1994; Accepted: September 1994)

Abstract

Results of seven GCM experiments in northern Europe are described. Two of the five transient experiments, named herein after the atmospheric model components as ECHAM1 and ECHAM2, are from MPI (the Max-Planck-Institut für Meteorologie, Germany), the remaining three from GFDL (the Geophysical Fluid Dynamics Laboratory, USA), NCAR (the National Center for Atmospheric Research, USA), and UKMO (the United Kingdom Meteorological Office). The two equilibrium experiments are those of GFDL and UKMO. The ability of the models to simulate present climate and their response to enhanced greenhouse forcing are both studied. The quantities included in the analysis are sea-level pressure, surface air temperature, and precipitation. Precipitation data were unfortunately available for only three experiments.

Taking both winter and summer into account, the climatological sea-level pressure distribution in northern Europe is best reproduced in the ECHAM2 control simulation. The simulated winter temperatures are generally too cold; in some models the systematic error for the continental northern Europe is of the order of 10 °C. The most realistic winter temperatures are given by the GFDL simulations. In summer, the systematic errors are much smaller, but the north-south gradient is too large in several models. The ECHAM1, ECHAM2, and NCAR models all show some skill in simulating the geographical distribution of annual precipitation, but the simulated seasonal cycle differs from the observed. In the ECHAM1 and NCAR simulations, in particular, summer precipitation is severely underestimated in most parts of northern Europe.

The available experiments generally indicate a larger CO_2-induced warming in northern Europe in winter than in summer. For those four transient experiments in which the doubling of CO_2 was achieved, the average warming in Finland at that time is 3 to 4 °C in winter and slightly over 2 °C in summer. ECHAM1, ECHAM2, and NCAR all predict a slight increase in the area average precipitation for continental northern Europe, but the seasonal cycle and geographical details of this increase vary from model to model. The model results indicate a small but widespread pressure decrease in northern Europe in summer, and there are also some hints of a slight northward shift of the Icelandic low in winter.

1. *Introduction*

In summer 1993, the Intergovernmental Panel on Climate Change (IPCC) initiated a project entitled "Evaluation of Regional Climate Simulations" (*IPCC WG I*, 1993). Among the aims of this project are to assess the performance of current general circulation models (GCMs) in simulating contemporary climate and to intercompare the response of different GCMs to increasing CO_2 on regional scales. For this purpose, a public data base was

collected at the Deutsches Klimarechenzentrum (DKRZ[1]), which contains data from four modelling centres: the Max-Planck-Institut für Meteorologie (MPI) in Hamburg, the Geophysical Fluid Dynamics Laboratory (GFDL) in Princeton, the United Kingdom Meteorological Office (UKMO) in Bracknell, and the National Center for Atmospheric Research (NCAR) in Boulder. Altogether, there are results from five transient greenhouse simulations[2], from two 2 x CO_2 equilibrium simulations, and from the seven corresponding control runs intended to simulate the present climate.

Some features of the transient experiments are given in Table 1. In the two equilibrium experiments (GFDL and UKMO), shallow mixed-layer oceans were used instead of the oceanic GCMs. In the UKMO equilibrium experiment, in addition, an earlier version of the UKMO atmospheric model with coarser horizontal resolution (5° x 7.5°) was used than in the corresponding transient experiment. Regardless of the original resolution of the models, all the results in the data set were interpolated at DKRZ to the T21 gaussian grid. In this grid, the width of the gridboxes is slightly over 600 km in the meridional direction and in southern Finland about 300 km in the zonal direction (see Fig. 1). The results of the greenhouse runs were averaged over a 10-year (ECHAM1, ECHAM2, NCAR, and UKMO) or 20-year (GFDL) time slice, which in four of the five transient experiments is centred at the doubling of CO_2. In the NCAR experiment, however, the increase of CO_2 in the middle of this time slice (the last 10 years of the 70-year simulation) was only about 65 %. The results of the control simulations are from a 10-year time slice with the exception of the GFDL experiments, in which a 100-year averaging period was used.

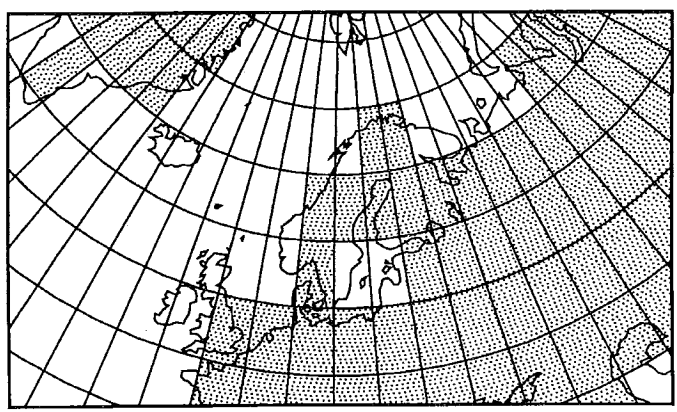

Figure 1. Gridboxes in the T21 grid in northern Europe. "Land gridboxes" are shaded.

1 DKRZ works in close co-operation with the Max-Planck-Institut für Meteorologie in Hamburg.

2 In this paper, "transient simulation" has the same meaning as "time-dependent simulation" in Section 6.4 of *Houghton et al.* 1990. For an explanation of "equilibrium simulation", see *Houghton et al.* 1990, Section 6.2.

This paper describes some of the simulation results in northern Europe. The two parts of the IPCC regional model intercomparison project are both covered: control simulations are compared with observational data in section 2, and the model-simulated CO_2-induced climate change is described in section 3. The analysis comprises those three quantities - sea-level pressure, surface air temperature, and precipitation - that are the first three on the list of priority defined by *IPCC WG I* (1993). In addition, a comparison between the climatological wintertime temperature sounding for a station in northern Finland and the simulated temperature profiles in two models is included at the end of section 2.2. Many aspects of the discussion are restricted to the situation in winter (December-February = DJF) and in summer (June-August = JJA), because, for example, data for spring (March-May = MAM) and autumn (September-November = SON) were not available for the GFDL and UKMO experiments. In addition, although climate in northern Europe is characterized by large daily and interannual variations, the present analysis is confined to the mean state only. This choice was largely dictated by the rather limited amount of variability information in the IPCC data set.

Table 1. Summary of the five transient experiments. The resolution of the models is given using the convention in Table B1 of *Houghton et al.* (1992): for example, the GFDL atmospheric model is a spectral model with rhomboidal truncation to maximum wave number 15 (roughly equivalent to a gridbox size of 4.5° in latitude and 7.5° in longitude) and 9 vertical levels. CLW = Cloud Liquid Water. (Updated from *Houghton et al.* (1992) using unofficial documentation provided by DKRZ.)

	ECHAM1 [a] (Cubasch et al. 1992)	ECHAM2 [a]	GFDL (Stouffer et al. 1989, Manabe et al. 1991, 1992)	NCAR (Washington and Meehl, 1989 [b])	UKMO (Murphy 1992)
Resolution (atmosphere)	T21L19 (5.6° x 5.6°)	T21L19 (5.6° x 5.6°)	R15L9 (4.5° x 7.5°)	R15L9 (4.5° x 7.5°)	2.5° x 3.75° L11
Resolution (ocean)	4° x 4° L11	T42L9 (2.8° x 2.8°) [c]	4.5° x 3.75° L12	5° x 5° L4	2.5° x 3.75° L17
Features	prognostic CLW, quasi-geostrophic ocean	prognostic CLW, isopycnic ocean, viscous-plastic sea-ice	no diurnal cycle, isopycnal ocean diffusion	no diurnal cycle, no run-off	prognostic CLW, isopycnal ocean diffusion
Increase of CO_2 (greenhouse run)	IPCC scenario A for equival. CO_2	IPCC scenario A for equival. CO_2	1% / year (compound)	1% / year (linear)	1% / year (compound)

[a] The "ECHAM1" and "ECHAM2" experiments have been conducted at MPI. These acronyms refer to the atmospheric models; the oceanic model components are denoted as LSG and OPYC, respectively.

[b] Washington and Meehl only describe the first 30 years of the NCAR experiment.

[c] Increased meridional resolution (1°) between 30°S and 30°N.

2. *Comparison of control simulations with observations*

There are at least two reasons to validate climate models against observational data, both of which arise from the fact that present models are far from perfect. First, this kind of

comparison can be useful for modellers, as information about model weaknesses may help in the development of more realistic models. Another, more practical point concerns the model-produced estimates of the climatic effects of increasing CO_2 and other greenhouse gases. As these estimates vary a lot from model to model, some way is needed to distinguish the more likely results from the less probable. As a first guess one can suppose that a model's ability to simulate climate changes correlates with its ability to simulate the present climate. Thus, a model with a good control climate may be more likely to provide reasonable estimates of climate changes than one which only simulates the present climate with great difficulty. It must be emphasized, however, that this relationship is only an approximate one: due to the complexity of the climate system, it is not necessarily the case that exactly the same factors are important for the simulation of present climate as for climate changes.

To estimate the quality of the control simulations is unfortunately not a straightforward matter. One of the important questions concerns the area used in the evaluation. Even at the regional level, climate changes are likely to be determined primarily by very large-scale processes. This suggests that a wide comparison domain should be used even when the area of principal interest is fairly small. On the other hand, it is evident that the local performance nevertheless deserves some extra attention. For example, when climate changes in northern Europe are considered, a deficiency in the simulated sea-level pressure pattern over Scandinavia is certainly more worrying than one on the opposite side of the globe. In view of these arguments, two comparison domains were used in the case of sea-level pressure: one covering northern Europe and the northeastern parts of the Atlantic Ocean, the other the whole northern extratropics. For the climatological temperature and precipitation distributions, however, no full-cover global-scale data sets were available for the present study. The analysis for these quantities is therefore restricted to northern Europe only, which undeniably is a shortcoming in estimating the reliability of the greenhouse simulations. Be that as it may, however, the investigation of model performance in a limited area is motivated by its potential to reveal regional features that easily become overlooked in global-scale analysis. A nice example of an earlier study of this kind is that conducted by *Whetton et al.* (1994) for the Australian region.

2.1 Sea-level pressure

Sea-level pressure is a very fundamental meteorological variable, since its distribution largely determines the atmospheric circulation near the surface. The winter (DJF) and summer (JJA) averages of this quantity in the various control simulations, together with the climatological means, are shown for northern Europe in Figs. 2 and 3. The climatological values are based on the ECMWF (European Centre for Medium Range Weather Forecasts) operational analyses for the 10-year period from 1979 to 1989 (*Hoskins et al.* 1989).

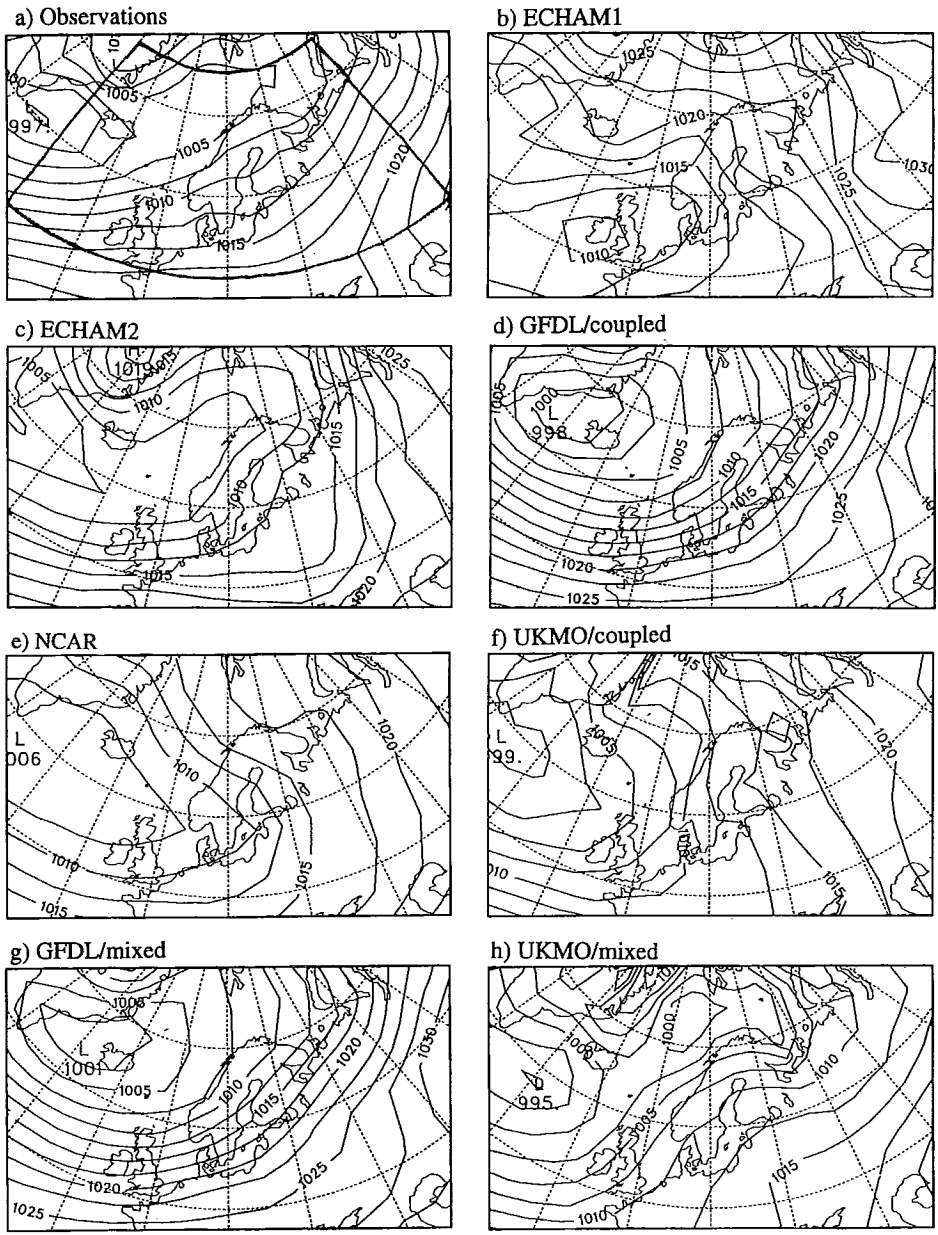

Figure 2. Distribution of mean sea-level pressure in northern Europe in December-February as given by (a) ECMWF analyses 1979-1989, and (b)-(h) control simulations. Coupled models: (b) ECHAM1, (c) ECHAM2, (d) GFDL, (e) NCAR, and (f) UKMO. Models with mixed-layer ocean: (g) GFDL, and (h) UKMO. Contour interval is 2.5 hPa.

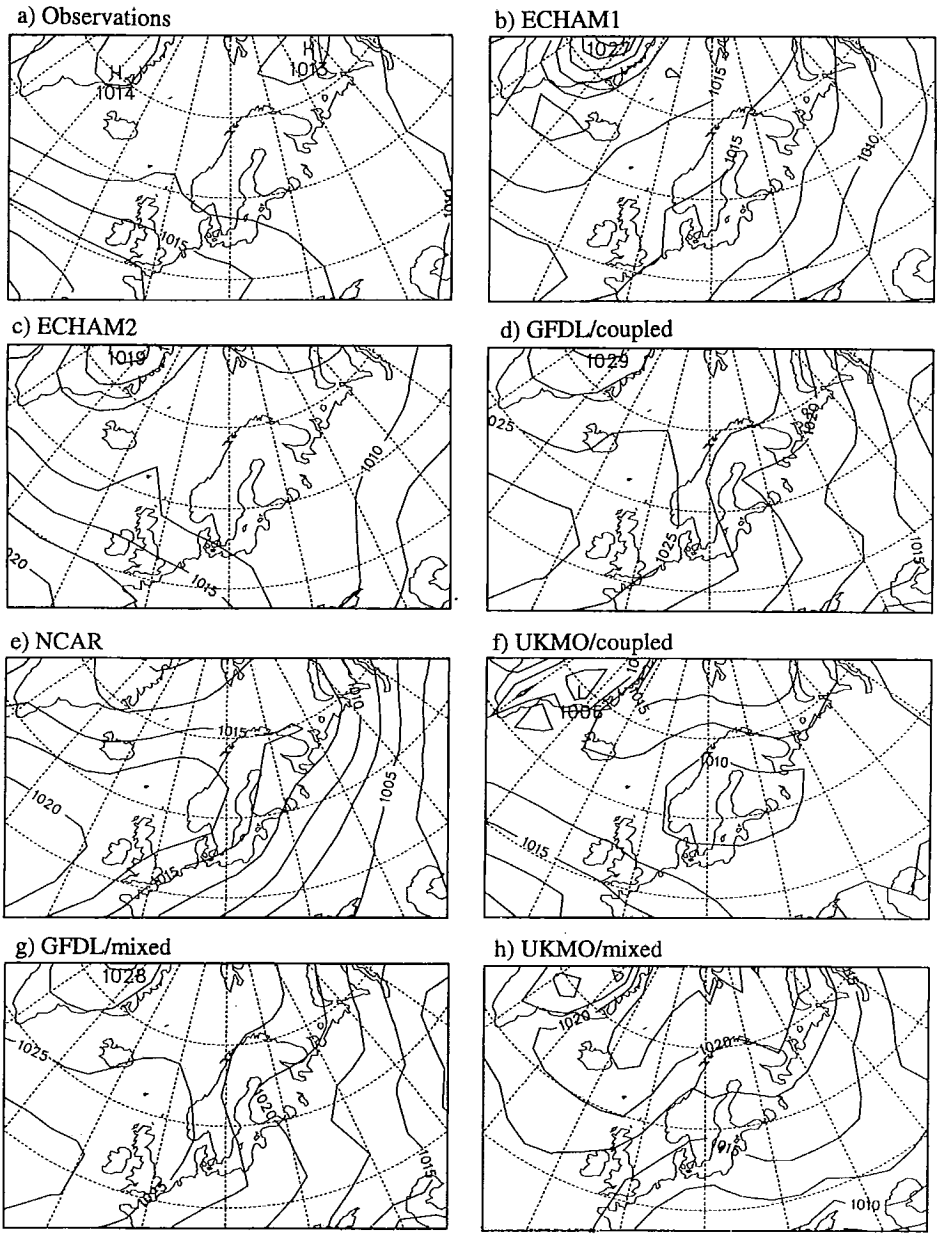

Figure 3. As Figure 2 but for June-August.

In winter (Fig. 2), most models qualitatively simulate the North Atlantic low pressure area centred west of Iceland. In the ECHAM1 simulation, however, this low is located much too far to the south and there is actually a moderately strong high pressure area north of 60°N. At a more detailed level, there are errors in the strength and location of the pressure patterns in the other simulations as well. For example, the observed northeastward extension of the Icelandic low toward the Arctic Ocean is totally lacking in the NCAR and the UKMO/coupled runs. In these simulations, as well as in the ECHAM1 run, the average geostrophic wind direction in Finland is therefore from the south or southeast, not from the southwest as the ECMWF analyses show. It is worth noting that the two GFDL simulations give very similar pressure fields but there are large differences between the UKMO/coupled and UKMO/mixed simulations. This is probably because, unlike in the GFDL experiments, even the atmospheric models were somewhat different in the two UKMO experiments.

In addition to the subjective comparison between simulated and observed fields, we find it useful to quantify the results by means of simple statistical calculations. In Table 2, two commonly used quantities are employed: correlation coefficient and root-mean-square (RMS) error. Although these two quantities give partly overlapping information, neither of them is needless: The RMS error gives a good overall measure of the simulation errors, but its interpretation is complicated because this quantity reacts to all types of differences (those in pattern structure, the amplitude of the patterns, and the mean value) between the simulated and observed fields. In contrast, the correlation coefficient only measures the similarity in pattern structure, which is perhaps the most interesting aspect in comparison of observed and modelled pressure distributions. The calculations were performed for the area (50°-75°N, 30°W-60°E), which is bordered by the bold lines in Fig. 2a, and cosine-weighting was applied to compensate for the decreasing zonal width of the gridboxes with increasing latitude.

Table 2. Statistical comparison between the simulated sea-level pressure fields and the ECMWF analyses for the area (50°-75°N, 30°W-60°E). The unit for RMS error is hPa.

a) December - February

	Coupled models					Mixed-layer ocean	
Model	ECHAM1	ECHAM2	GFDL	NCAR	UKMO	GFDL	UKMO
Correlation	0.05	0.87	0.94	0.37	0.53	0.95	0.90
RMS error	11.6	3.0	4.8	6.3	5.7	5.0	4.3

b) June - August

	Coupled models					Mixed-layer ocean	
Model	ECHAM1	ECHAM2	GFDL	NCAR	UKMO	GFDL	UKMO
Correlation	0.53	0.87	0.63	0.48	0.40	0.64	-0.24
RMS error	2.4	1.1	10.8	4.4	2.0	10.4	6.5

Although this statistical analysis only reflects the performance of the models in a rather limited domain, it is interesting to note that the model-to-model differences are not small. In winter, two of the five coupled models (GFDL and ECHAM2) and the two mixed-layer models score a rather high (of the order of 0.9) correlation with the climatological pressure field, but the remaining three models succeed much worse. This holds especially for ECHAM1, in which the Icelandic low is seriously misplaced. The picture given by the RMS errors is in this season broadly similar to that given by the correlation coefficients.

In summer, the correlation coefficients are in most cases lower than in winter. With the exception of the northern edge of the subtropical high over the Atlantic, however, the climatological pressure variations in the calculation domain are very small. Correlation coefficients can thus be low even when absolute errors are moderate. The RMS values show that this is indeed the case for some models.

Regardless of the measure used, the ECHAM2 simulation performs best in summer. In the Atlantic sector, in particular, the simulated pressure distribution is almost identical to that obtained from the ECMWF analyses. After this model, however, the two measures give a somewhat different ordering. The poor performance of the two GFDL simulations in the RMS error comparison results from markedly too high simulated pressure values all over the domain. However, this result is partly an artefact: as also noted by *Boer et al.* (1992), the global mean sea-level pressure in the GFDL simulations is about 6.5 hPa too high. Since it is the pressure differences, not the absolute values that are meteorologically interesting, it might actually be more relevant to subtract this global bias before the calculations. When this was done, the resulting summertime RMS errors were near 5 hPa for both GFDL simulations. In winter, in contrast, this correction did not alter the RMS values substantially.

Qualitatively, the errors in summer also vary from model to model. A general feature, present at least in the ECHAM1, NCAR and GFDL runs, is a somewhat too strong pressure fall towards the interior of the Eurasian continent. The northern edge of the subtropical high over the Atlantic is well simulated in the ECHAM2 and UKMO/coupled runs, but it protrudes too far north in the GFDL and NCAR simulations. In the UKMO/mixed simulation, which succeeds by far the worst in the correlation comparison, the north-south pressure gradient over the Atlantic is reversed: instead of a northward extending subtropical high, the pattern is dominated by a southward extending Arctic high.

To get a better estimate of the potential reliability of the models, it is, as discussed earlier, most likely useful to investigate their performance also in a wider domain. The statistical computations were therefore repeated for the whole northern extratropics (30°-90°N). Some mountainous areas in the Himalayas, Rocky Mountains, and Greenland were however excluded, because sea-level pressure is a well-defined quantity only in relatively low areas. (The exact borderline between the included and excluded areas was

a surface height of 1000 m in the GFDL orography field.) The results are presented in Table 3.

The relative performance of the models over this larger area is in most respects similar to that in northern Europe. There are, however, some minor differences. The ECHAM2 simulation now unambiguously gives the best results in both the winter and the summer. In summer, the UKMO/coupled simulation definitely performs second best, and compared to the results for northern Europe, it also succeeds moderately well in winter. The two GFDL simulations, on the other hand, rank slightly worse in the whole extra-tropics, at least in winter. (The RMS errors for GFDL are again contaminated by the global bias; the "corrected" values are near 5 hPa in winter and about 3.5 hPa in summer.) In most cases, the correlation coefficients for this wider area are higher than they were for northern Europe alone, implying that large-scale pressure variations are simulated more reliably than regional features. It is perhaps worth noting that, in contrast to what was suggested by one of the reviewers, this improvement is not a trivial result associated with the simulation of strong equator-pole gradients in the zonal-mean distribution: in the case of sea-level pressure, the meridional gradients are actually not very strong in the Northern Hemisphere (see *Houghton et al.* 1990, p. 101). In fact, the subtraction of the zonal means from the observed and simulated fields increased the correlation coefficients in a larger number of cases than it decreased them (not shown).

Table 3. As Table 2 but for the area 30°-90°N.

a) December - February

	Coupled models					Mixed-layer ocean	
Model	ECHAM1	ECHAM2	GFDL	NCAR	UKMO	GFDL	UKMO
Correlation	0.59	0.93	0.87	0.76	0.84	0.86	0.86
RMS error	9.0	3.4	7.2	5.0	4.3	7.2	5.5

b) June - August

	Coupled models					Mixed-layer ocean	
Model	ECHAM1	ECHAM2	GFDL	NCAR	UKMO	GFDL	UKMO
Correlation	0.81	0.94	0.79	0.79	0.86	0.80	0.12
RMS error	3.0	1.6	8.1	5.5	2.4	7.8	6.3

2.2 *Surface air temperature*

For the great audience, surface air temperature at about two metres height is perhaps the most interesting single climatic variable. In the IPCC data set, however, this quantity was available only for the two MPI simulations. For GFDL and NCAR, the temperatures at the lowest model level, located in both cases at roughly 80 metres above the surface, are used in the following. For the two UKMO experiments, the only available temperature

was that of the actual land or sea surface. Although surface temperature was in fact available for all models, it was not used if not necessary, since observational data were only available for air temperature. In most cases, the differences between surface temperatures and the temperatures used in the comparison were actually unimportant. However, as briefly discussed later in this section, there were some exceptions.

The observation-based and model-simulated temperature fields are shown in Figs. 4 (winter) and 5 (summer). The observational temperature data were obtained (with the kind help of Eino Hellsten and Heikki Tuomenvirta at the Finnish Meteorological Institute) from a data bank maintained by the Carbon Dioxide Information Analysis Center (CDIAC) in Oak Ridge, Tennessee, USA. The mean temperatures were primarily taken from the period 1961-1990, but some synoptic stations with data available for only the last 20 years were also included. To make a comparison with the model results feasible, values for the T21 grid were calculated by taking an unweighted average of the available station values within each gridbox.

As in the case of sea-level pressure, the intercomparison results are summarized by using a few statistical quantities (Table 4). Again, the calculation domain is bounded from the south by the latitude 50°N and from the east by the longitude 60°E, but only those gridboxes are included which are classified as land in the models and for which observational data were available. Among these 28 gridboxes (the shaded area in Fig. 4a), the number of stations in each ranged from 1 to 16, with an average of 3.6. In addition to the two quantities used in the previous subsection, Table 4 shows the systematic model errors. These figures were obtained by subtracting the observation-based area averages (-7.5 °C in winter and 16.2 °C in summer) from the model-simulated values.

Table 4. Statistical comparison between simulated and observed temperatures in land areas of northern Europe (see text). The unit for RMS error and systematic error is °C.

a) December - February

	Coupled models					Mixed-layer ocean	
Model	ECHAM1	ECHAM2	GFDL	NCAR	UKMO	GFDL	UKMO
Correlation	0.89	0.89	0.88	0.89	0.88	0.86	0.92
RMS error	5.5	6.0	3.2	10.2	12.7	3.5	10.0
Systematic error	-4.3	-4.8	-1.6	-8.8	-11.8	-1.8	-9.5

b) June - August

	Coupled models					Mixed-layer ocean	
Model	ECHAM1	ECHAM2	GFDL	NCAR	UKMO	GFDL	UKMO
Correlation	0.90	0.89	0.93	0.90	0.98	0.93	0.84
RMS error	3.6	2.8	2.1	5.4	1.9	2.2	3.0
Systematic error	+1.8	-2.1	-0.2	+3.5	-1.7	+0.6	-2.6

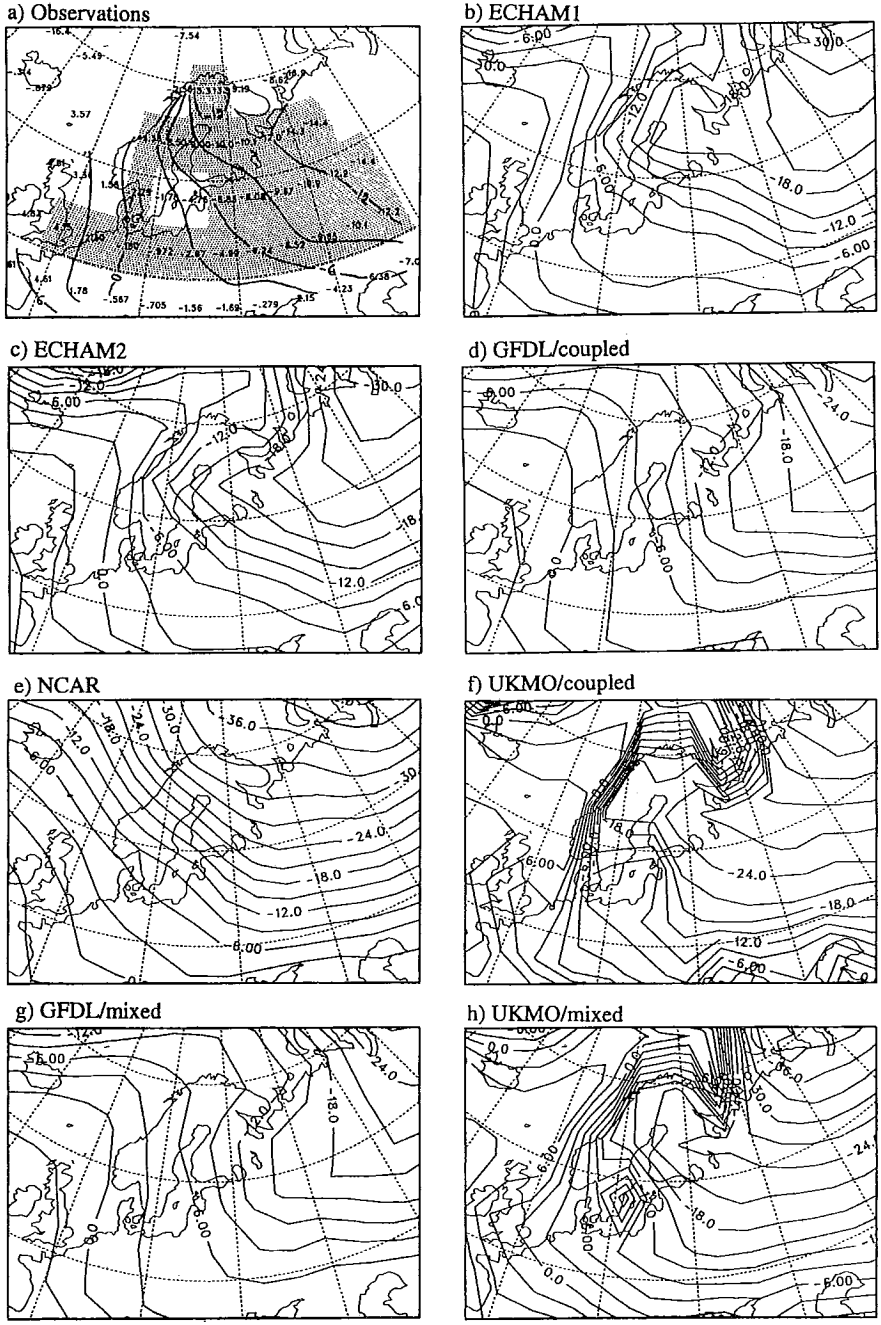

Figure 4. Distribution of surface or near-surface temperature (see text) in northern Europe in December-February as given by (a) observations, and (b)-(h) control simulations. Coupled models: (b) ECHAM1, (c) ECHAM2, (d) GFDL, (e) NCAR, and (f) UKMO. Models with mixed-layer-ocean: (g) GFDL, and (h) UKMO. Contour interval is 3 °C, and the shading in (a) marks the area used in the statistical comparison.

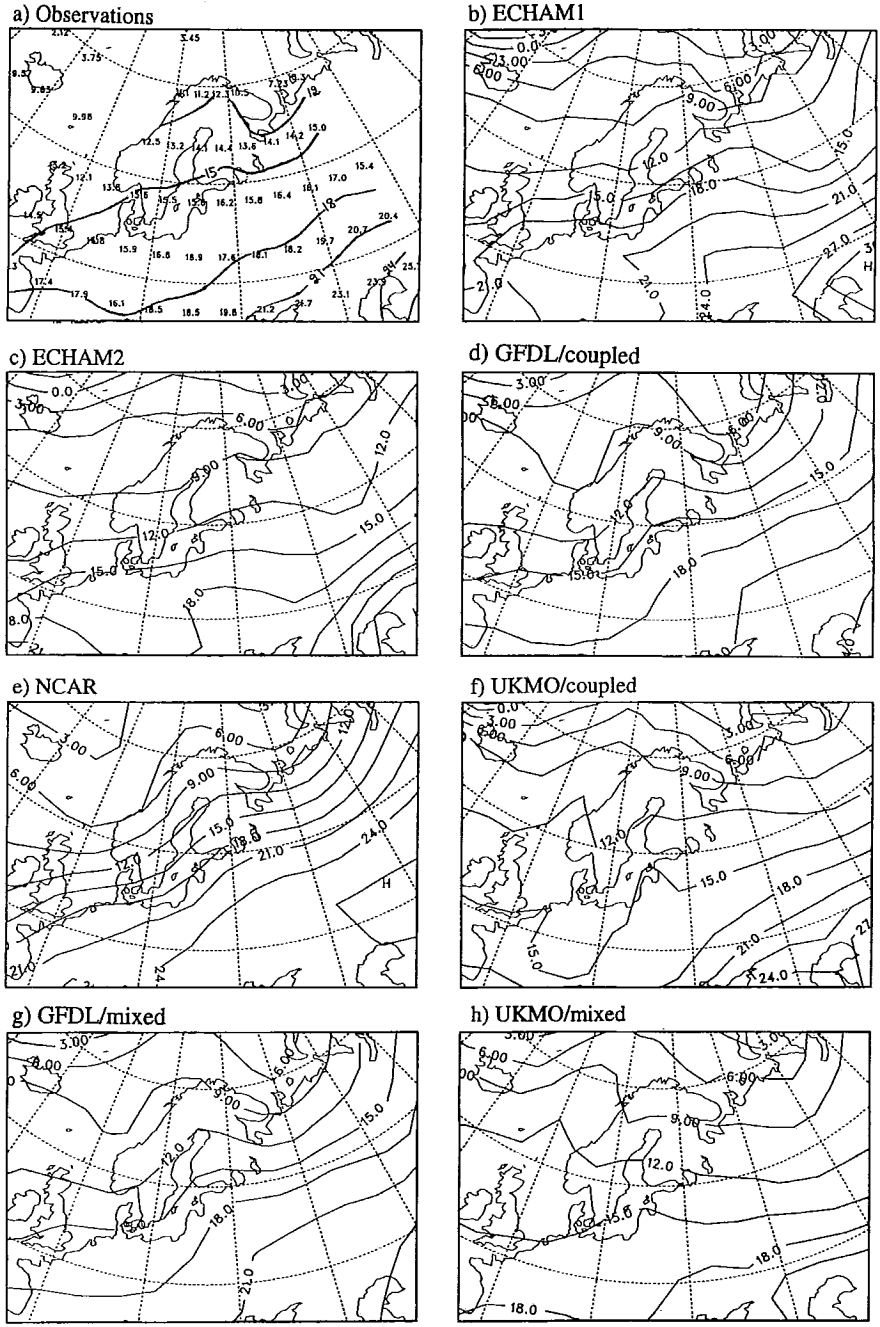

Figure 5. As Figure 4 but for June-August.

As indicated by the high correlation coefficients, all the models qualitatively simulate the general decrease in temperature from southwest to northeast in winter and from southeast to northwest in summer. In winter, however, the simulated temperatures are almost invariably too cold. The average cold bias is almost 12 degrees in UKMO/coupled, roughly 9 degrees in UKMO/mixed and NCAR, and between 4 and 5 degrees in the two MPI simulations. In many cases, most clearly in the NCAR simulation, this error grows towards northeast, thus making the horizontal temperature gradient too strong. In the two GFDL runs, however, the cold bias is on the average quite small and it is actually reversed e.g. in northern Finland. With the exception of the NCAR run, the other models also simulate a local amelioration in temperature in this area, in contrast to the observations. This feature is possibly connected with the coarse resolution of the models, which tends to spread the warming effect of the Arctic Ocean too far south.

Considering the calculation area as a whole, there is no such clear systematic tendency for the simulated temperatures to be too low in summer, as in winter. Only ECHAM2 and the two UKMO models exhibit an appreciable cold bias, and in the NCAR and ECHAM1 simulations the area average temperatures are actually too warm. However, the errors are not similar across the whole area. A feature that is most evident in the NCAR and ECHAM1 runs, but also identifiable in most other simulations, is that the north-south temperature gradient is too strong. In northern Finland, the simulated summer temperatures are at least a couple of degrees too cold in all models. This is also the case for most models in the southern and central parts of the country, but the GFDL and NCAR temperatures are near the observed values.

Apart from the UKMO simulations, the results described above have been calculated using air temperatures at 2 metres (ECHAM1 and ECHAM2) or about 80 metres height (GFDL and NCAR) instead of the actual surface temperatures. In winter, this choice was not important: although wintertime surface temperatures exceed air temperatures by several degrees in all models in some sea areas (which were omitted in the statistical comparison), the corresponding differences in land areas are small compared to the large differences between the models. In summer, however, the land-surface temperatures in the NCAR and the two GFDL simulations are much higher than the lowest level air temperatures. Averaged over the calculation domain, this difference is between 3.5 and 4 °C in all these three simulations. It seems evident that the 80 metres temperatures were the better choice: although the highest temperatures, at least in daytime, are measured near the ground in summer, the average difference between 2 and 80 metres probably does not exceed the dry-adiabatic value of 0.8 °C.

In addition to the surface and near-surface temperatures, temperatures at higher atmospheric levels were available for the GFDL and NCAR simulations. No detailed analysis of these upper-level temperatures was made, but it is interesting to look at one example. Figure 6 represents the average December-February temperature sounding at Sodankylä (67.4°N, 26.7°E) for 1965-1980 (*Finnish Meteorological Institute* 1984)

together with the simulated GFDL and NCAR soundings at the nearest gridpoint (69.2°N, 28.1°E). Clearly, the temperatures in the NCAR simulation are several degrees too cold throughout the troposphere, but by far the largest errors occur within the lowest two kilometres above the surface. The GFDL temperatures are also slightly too cold in the free atmosphere. In the surface layer, however, this simulation is a few degrees too warm because the model is essentially unable to reproduce the low-level inversion present in the climatological sounding.

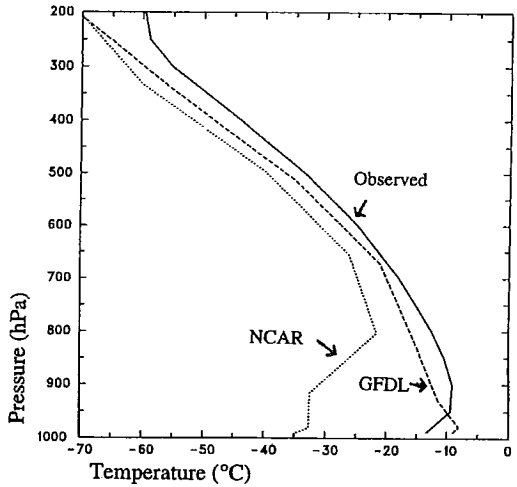

Figure 6. Average December-February temperature sounding at Sodankylä (67.4°N, 26.7°E) in 1965-1980 and the soundings for the nearest gridpoint in the GFDL/coupled and NCAR control simulations.

2.3 Precipitation

The IPCC data set contains precipitation data for only three models: ECHAM1, ECHAM2, and NCAR. Since results for all four seasons are available for these models, we begin this subsection by briefly discussing the annual averages. The annual simulated and climatological precipitation distributions are shown in Fig. 7. The source of climatological data was the same as that for temperature, though the sample of stations was a little larger. The same 28 land gridboxes that were used in the case of temperature (see Fig. 4a) were also used in the calculations for precipitation, and for these gridboxes, the average number of stations was 5.3, the range being from 1 to 19. The observed distribution shows a general, though certainly not monotonous, decrease of precipitation from the Atlantic coast towards the continental interior, and this decrease is also qualitatively simulated by all three models. Even at the annual level, however, there are large differences in the details. The correlation coefficients with the observed annual precipitation are 0.41 for ECHAM1, 0.59 for ECHAM2, and 0.57 for NCAR.

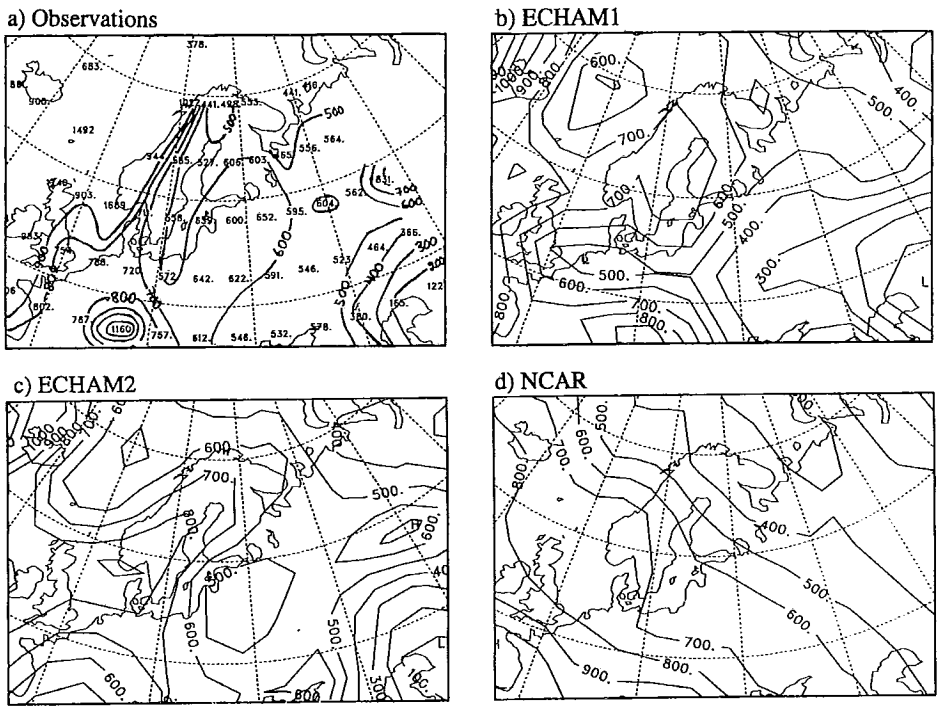

Figure 7. Annual average precipitation (mm) in northern Europe as given by (a) observations, and the control simulations of (b) ECHAM1, (c) ECHAM2, and (d) NCAR. Contour interval is 100 mm.

Taking an average over the whole domain, all three models simulate a smaller than observed annual precipitation, but this deficit exceeds 10 % only in the ECHAM1 simulation (see Table 5). However, the observed and simulated annual cycles are different. The models invariably simulate a minimum of precipitation in summer and a maximum in winter. According to the observational data, in contrast, summer is the wettest season and winter the second driest.

The excessive summertime dryness of the models is undoubtedly a real error. No such clear conclusions can be drawn, however, from the larger than observed simulated precipitation in winter. This is because precipitation measurements tend to give an underestimate of the real precipitation, and this systematic error is especially serious for snowfall. In southern and central Finland, for example, the actual winter precipitation for 1961-1990 is estimated to be almost 35 % larger than the measured precipitation (*Reijo Solantie*, personal communication). This error, which is mainly of aerodynamical origin, is however not expected to be constant across the whole of northern Europe. Appreciable variations in it may arise from variations in the portion of precipitation falling as snow and those in

windiness, and also because different rain gauges have been used in different countries. In the coldest parts of the area, in particular, the problems with snowfall measurement also affect the observational record in spring (March-May) and autumn (September-November). According to the figures given by *Solantie* (personal communication), the appropriate corrections for southern and central Finland are near 23 % in spring and about 15 % in autumn. For the totally liquid precipitation in summer, the correction is estimated to be 6 %.

The differences in the area averages of the observed and simulated seasonal precipitation fields are accompanied by differences in the geographical patterns (not shown). A striking example occurs in summer in the ECHAM1 and ECHAM2 simulations: while the observed spatial distribution is fairly uniform with three-month totals around 200 mm in most of northern Europe, these two models produce a pronounced precipitation maximum over Finland and Sweden. In these areas, the summer precipitation in the ECHAM1 and ECHAM2 simulations is fairly realistic, but, on the other hand, precipitation in areas south of 60°N is underestimated more severely than the area averages in Table 5 indicate. In the NCAR simulation, in contrast, the underestimate is more evenly distributed, and even Finland and Sweden receive less than a half of the observed summertime precipitation.

Table 5. Average seasonal precipitation totals (mm) for land areas of northern Europe as given by control simulations and observations.

	DJF	MAM	JJA	SON	Annual
ECHAM1	152	113	67	131	464
ECHAM2	177	142	120	151	591
NCAR	175	155	65	155	551
Observations	120	119	198	171	608

3. *Climate changes induced by increasing CO_2*

This section describes the model-simulated climate changes, defined as the differences between the greenhouse runs and the control runs. The same three quantities are included in the analysis as in the previous section, but in different order. Since, at least at the qualitative level, surface air temperature is the quantity for which the uncertainties in the expected changes are smallest, the simulation results for this quantity are discussed before those for precipitation and sea-level pressure.

3.1 *Surface air temperature*

The wintertime temperature changes in the transient experiments in northern Europe are shown in Fig. 8, and the corresponding fields for summer in Fig. 9. In addition to the results of the individual experiments, composite fields (Figs. 8f and 9f) for both seasons are included. These composites were calculated to illustrate the most common features in the transient experiments, and they were obtained as the arithmetic average of the ECHAM1,

ECHAM2, GFDL, and UKMO results. The NCAR results were omitted, since (as noted in the introduction) the doubling of CO_2 was not achieved in this experiment.

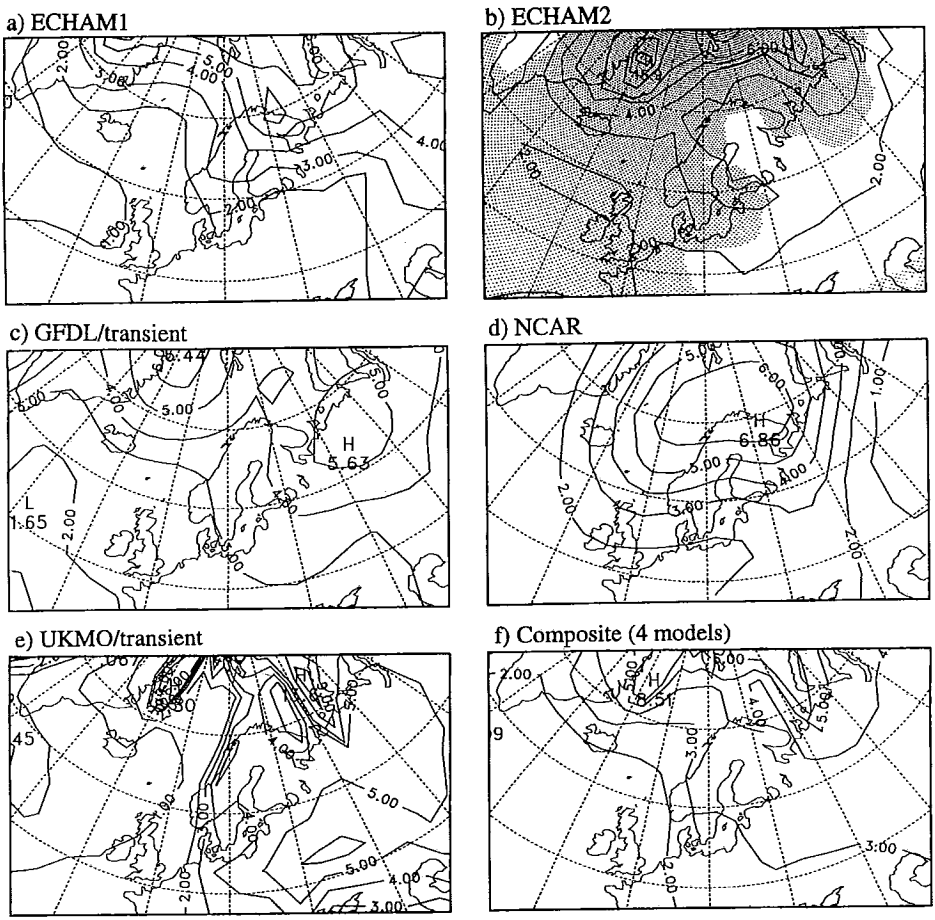

Figure 8. Change in surface or near-surface temperature in December-February in the transient experiments: (a) ECHAM1, (b) ECHAM2, (c) GFDL, (d) NCAR, (e) UKMO, and (f) average of (a), (b), (c), and (e). Contour interval is 1 °C up to 6 °C and thereafter 3 °C. The shading in (b) marks the areas where the increase is significant at the 95 % level.

With the exception of the UKMO experiment, in which marginal cooling occurs over the Atlantic Ocean south of the southern tip of Greenland, warming is simulated over the whole area portrayed in Figs. 8 and 9 in all models and in both seasons. Because CO_2-induced warming is a widely known result from earlier studies (see e.g. *Houghton et*

al. 1990), this qualitative similarity between the various experiments is not surprising. Nevertheless, since this point becomes more important when precipitation and sea-level pressure are considered, it is worthwhile to remember that the differences between the greenhouse and control simulations are not completely determined by genuine long-term changes in the model climate. On the contrary, interannual variability together with the limited sample sizes (for most models, an averaging period of only 10 years was used) may introduce large amounts of noise. To estimate the statistical significance of the simulated changes, two-tailed *t*-tests were performed for the ECHAM2 experiment, which was the only experiment with information about interannual variability available in the present data set. From the shading in Figs. 8b and 9b it can be seen that in both the winter and summer, the simulated warming in this experiment is significant at the 95 % confidence level over most of the domain. In winter, however, relatively wide continental areas exist where, according to this criterium, the simulated temperature change of about 2 °C is not significantly different from zero.

As is most clearly shown by the composite fields, the simulated warming in the land areas of northern Europe is generally stronger in winter than in summer. A qualitatively similar seasonal cycle in the warming is also indicated over the Atlantic, though the difference between winter and summer is small. In the individual experiments, there are local exceptions to this general rule even in the land areas, but they are restricted to south of 60°N. Over Finland, the warming in the composite field is between 3 and 4 °C in winter and slightly over 2 °C in summer. In the NCAR experiment, which was not included in the composite, the difference between winter and summer is even larger. On the other hand, this difference is slightly smaller in the GFDL experiment, which gave the most realistic temperatures in the control simulation, and in ECHAM2, which in most respects provided the best sea-level pressure distributions.

As for comparison, the five transient experiments give the following global annual averages (taken as the mean of December-February and June-August for those experiments for which data for the other seasons were unavailable) of warming: ECHAM1 1.4, ECHAM2 1.8, GFDL 2.1, NCAR 1.0 and UKMO 1.7 °C. In northern Europe, the simulated warming generally exceeds these figures in winter. The ratio of winter warming in Finland to the global average is smallest, near 1.5, in the ECHAM2 experiment. At the other extreme, NCAR simulates six times the global average warming in the northern parts of the country. The summertime warming values in northern Europe are much nearer the global averages; in Finland the ratio is almost invariably between 1 and 1.5.

In both the winter and summer, the composite fields indicate a minimum of warming over the central Atlantic and a gradual increase towards the continental interior. In summer, however, the west-east gradient is generally fairly small, and even in winter it is absent in the ECHAM2 and NCAR experiments. In summer, the warming also has a minimum in the Arctic Ocean in most experiments, but it is generally very strong in the same area in winter. The latter feature is possibly related to the northward retreat of sea-ice. The finer

details in the geographical pattern are highly variable from experiment to experiment, and any attempt to interpret them is probably rendered useless by the effects of interannual variability.

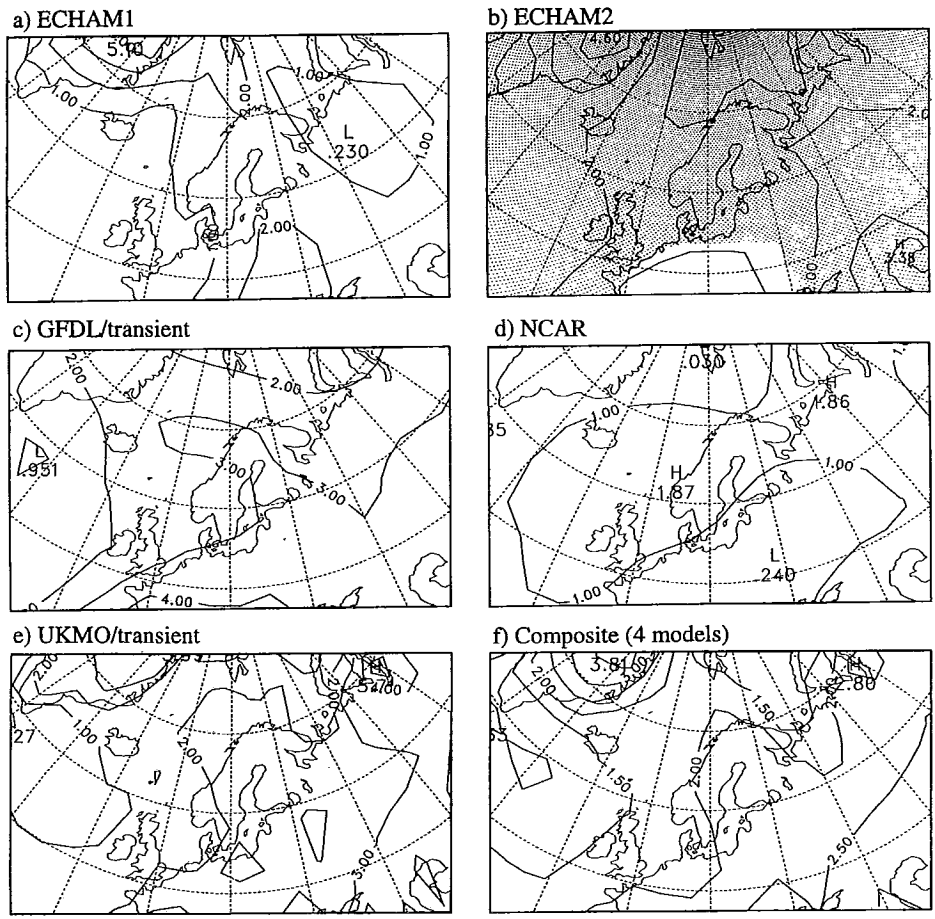

Figure 9. As Figure 8 but for June-August and with a contour interval of 0.5 °C in (f).

Results for spring and autumn were available only for ECHAM1, ECHAM2 and NCAR. For Finland, and also for the land areas of northern Europe as a whole, the ECHAM2 simulation indicates a slightly smaller warming in spring than in any other season. In the other cases, however, the simulated warming in spring and autumn is

generally between the values for summer and winter. With the exception of some areas near the coast of the Arctic Ocean, ECHAM1 and NCAR simulate a somewhat larger warming in spring than in autumn.

Figure 10 represents the simulated warming in the two 2 x CO_2 equilibrium experiments (GFDL and UKMO). As expected, the warming is stronger in these experiments than it is in the obviously more realistic transient experiments of the same modelling centres. The difference is most striking over the Atlantic Ocean and somewhat smaller in the continental Europe. In the case of UKMO, however, there is a marked minimum of warming over the Atlantic even in the equilibrium experiment. In Finland, the ratio of transient warming to equilibrium warming is almost the same for GFDL and UKMO, being near 0.7 in both summer and winter. The global average of this ratio is roughly 0.6 in both experiments: the average transient and equilibrium warmings, taken as the mean of winter and summer, are 2.1 °C and 3.6 °C for GFDL, and 1.7 °C and 2.7 °C for UKMO.

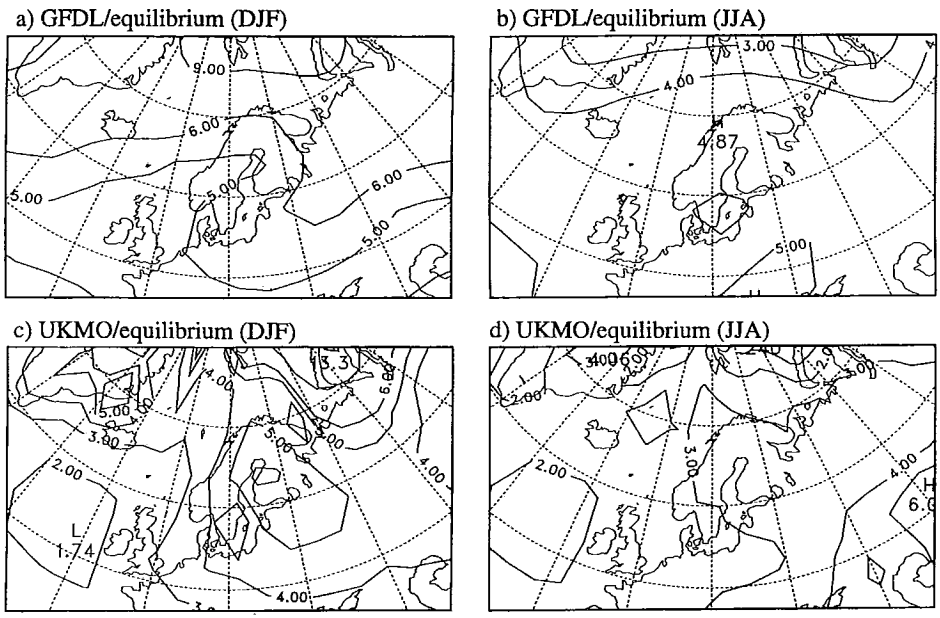

Figure 10: Warming in northern Europe in the equilibrium experiments: (a) GFDL, winter, (b) GFDL, summer, (c) UKMO, winter, and (d) UKMO, summer. Contour interval is 1 °C up to 6 °C and thereafter 3 °C.

3.2 *Precipitation*

The three experiments for which precipitation data were available all indicate a slight increase of annual precipitation in most parts of northern Europe (see Fig. 11). The increases in the area average precipitation, taken over the land gridboxes north of 50°N and west of 60°E, are 5 % for ECHAM1, 8 % for ECHAM2, and 14 % for NCAR. Although not dramatic, these increases clearly exceed the corresponding global averages, which are 2.4 % for ECHAM1 and ECHAM2, and 3.2 % for NCAR.

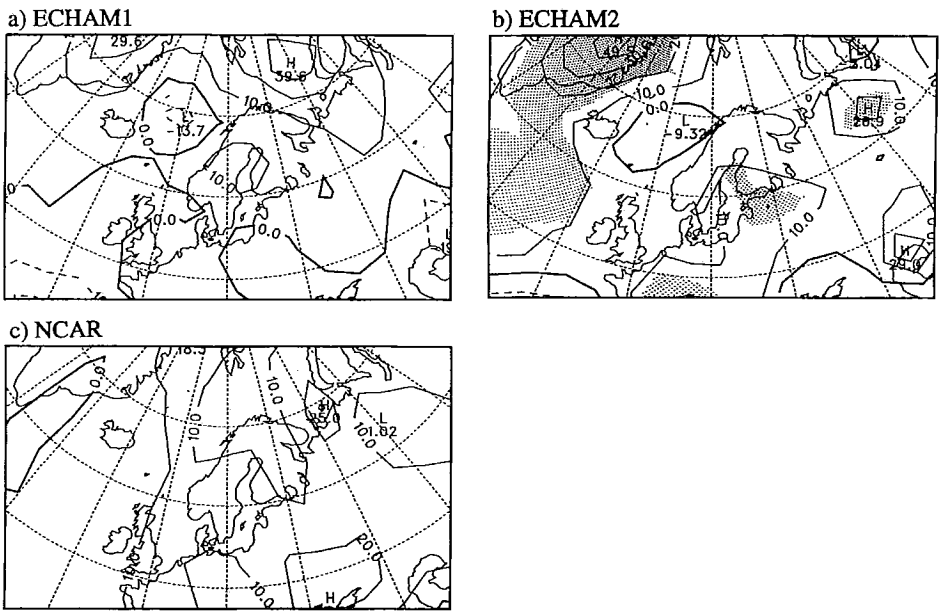

Figure 11. Simulated changes in annual precipitation, expressed in percent: (a) ECHAM1, (b) ECHAM2, and (c) NCAR. Contours are drawn at every 10 % with a bold zero line. The shading in (b) marks the areas where the change is significant at the 95 % level.

The geographical patterns of the simulated annual changes consist of fairly small-scale details, which vary a lot from model to model. While some of these variations undoubtedly originate in the differences between the models, others most likely reflect the influence of interannual variability. In the case of precipitation, this variability is indeed a larger problem than in the case of temperature: in the ECHAM2 experiment, the

simulated annual changes are statistically significant at the 95 % level only in a few gridpoints over the continental northern Europe (see Fig. 11b). There is, though, a somewhat wider area of significant increase extending from Greenland to the Atlantic Ocean. In the simulated changes for the three-month seasons (not shown), the details vary from model to model even more than in the case of annual averages, and at least in the ECHAM2 experiment, their statistical significance is mostly low. Even at the seasonal level, however, the area average changes for the whole of northern Europe are in all cases at least marginally positive (see Table 6a). In the two MPI simulations, the largest average increase occurs in summer. In the NCAR experiment, in contrast, summer is the season with the smallest increase.

The corresponding seasonal changes for Finland alone are represented in Table 6b. These values were calculated by using the three gridboxes that are centred within the borders of the country at (63.7°N, 22.5°E), (63.7°N, 28.1°E) and (69.2°N, 28.1°E). The simulated changes are generally positive in Finland as well, and it is difficult to find any systematic differences from the values for the whole of northern Europe. The NCAR experiment indicates a slight decrease of summer precipitation in Finland, but the unrealistically low precipitation in the control simulation suggests that not much weight can be attached to this result.

Table 6. Simulated seasonal changes in area average precipitation (%).

a) Land areas of northern Europe north of 50°N and west of 60°E

	DJF	MAM	JJA	SON	Annual
ECHAM1	+4	+4	+14	+4	+5
ECHAM2	+4	+1	+17	+13	+8
NCAR	+12	+20	+6	+13	+14

b) Finland

	DJF	MAM	JJA	SON	Annual
ECHAM1	+10	+4	+20	+11	+11
ECHAM2	+1	+12	+7	+9	+7
NCAR	+8	+10	-6	+18	+10

3.3 Sea-level pressure

The simulated changes in mean sea-level pressure are shown in Figs. 12 (winter) and 13 (summer). As in section 3.1, the composite fields presented in Figs. 12f and 13f give the average for those four transient experiments (ECHAM1, ECHAM2, GFDL, and UKMO) in which the doubling of CO_2 was achieved.

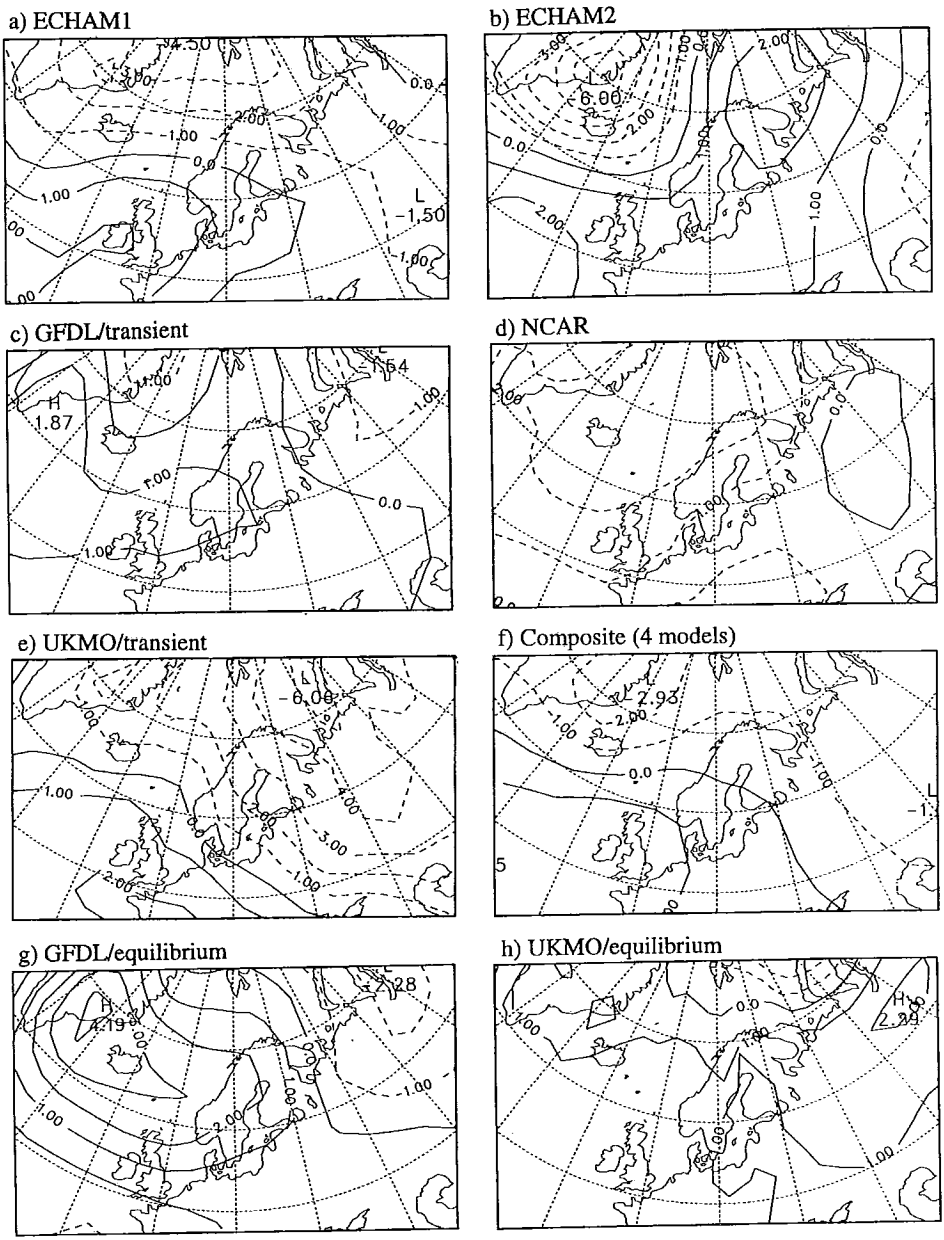

Figure 12. Simulated changes in mean sea-level pressure in December-February. Transient experiments: (a) ECHAM1, (b) ECHAM2, (c) GFDL, (d) NCAR, (e) UKMO, and (f) average of (a), (b), (c), and (e). Equilibrium experiments: (g) GFDL, and (h) UKMO. Contour interval is 1 hPa.

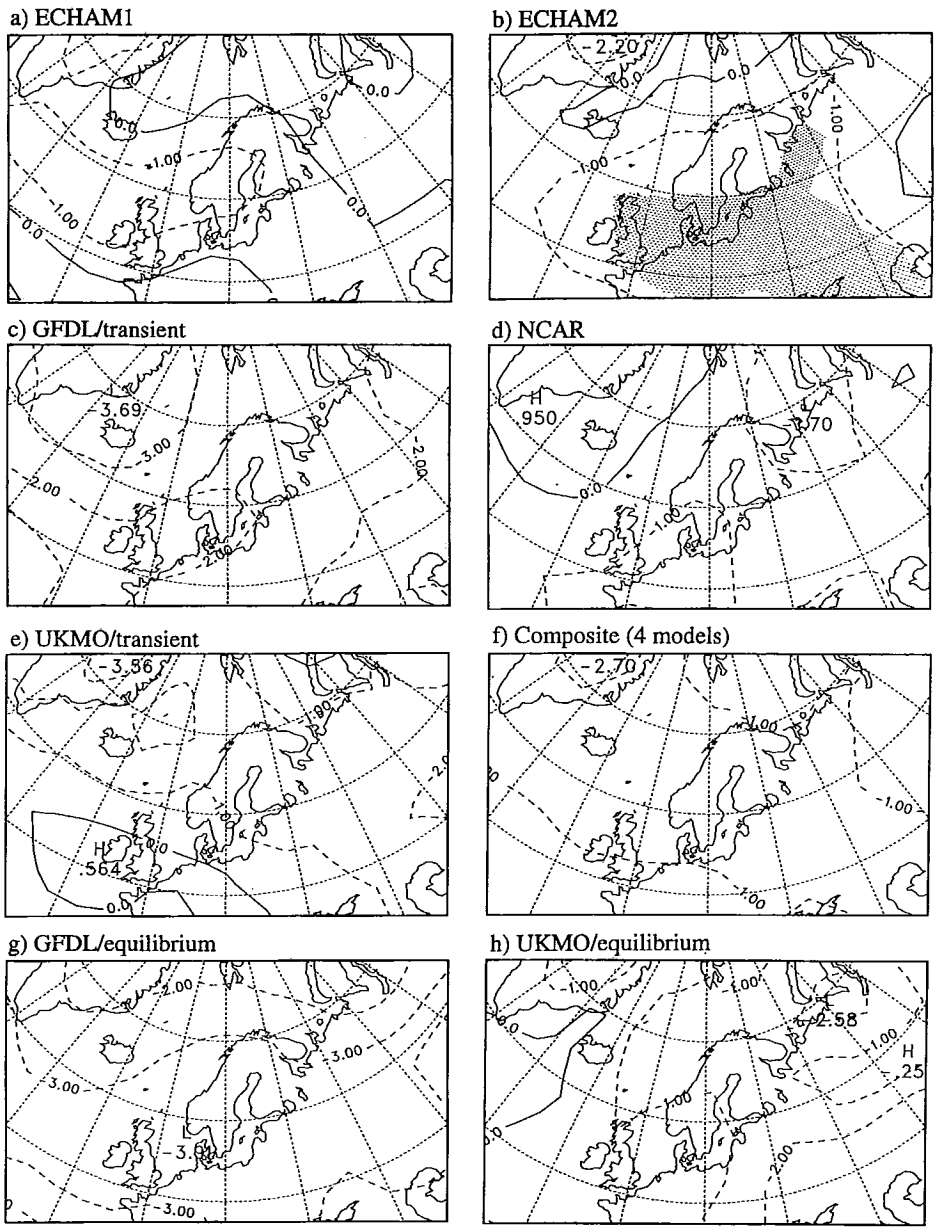

Figure 13. As figure 12 but for June-August. The shading in (b) marks the areas where the change is significant at the 95 % level.

In winter, relatively large pressure changes (locally up to 6 hPa) are simulated in many experiments. Over the continental northern Europe, there are large differences from model to model, but in the Atlantic sector the results of the *transient* simulations show a somewhat better agreement. In the composite field, there is an area of pressure decrease centred near the east coast of Greenland, but south of the 60th latitude, pressure increases slightly instead. This pattern, which indicates a slight northward shift of the Icelandic low and a strengthening of westerly winds over the northern North Atlantic, is very evident in three of the five individual transient experiments (ECHAM1, ECHAM2, and UKMO), and some hints of it can also be seen in the GFDL experiment. Nevertheless, the large year-to-year variations in the wintertime circulation make it impossible to rule out the possibility that this similarity between the individual experiments just occurs by chance: in the ECHAM2 experiment, the simulated changes are actually non-significant at the 95 % level everywhere within the area portrayed. Moreover, one of the two equilibrium experiments (GFDL) gives a pressure change pattern which is in marked contrast with the transient simulations. In this experiment, there is an extended area of simulated wintertime pressure increase centred between Iceland and Greenland, which implies a weakening of the Icelandic low but not a shift in its location.

In summer, the simulated pressure changes are in most models smaller than in winter. At least in most parts of northern Europe, however, all seven experiments indicate a slight summertime pressure decrease (typically of the order of 1-2 hPa). In spite of its smallness, this decrease may actually be a more reliable result than the possible wintertime changes discussed above: in the ECHAM2 experiment (Fig. 13b), the decrease is significant at the 95 % level in an area covering roughly a half of the continental northern Europe (note that the interannual variability is much weaker in summer than in winter). The simulated decrease is, however, generally very widespread and there are no large gradients, which implies that the corresponding changes in the average wind field are weak.

4. *Summary and discussion*

This paper briefly described results from seven GCM experiments, five of which have been conducted with coupled atmosphere-ocean models and two with atmospheric models with a mixed-layer ocean. The main area of interest was northern Europe with some special attention given to Finland. The ability of the models to simulate present climate and their response to increased CO_2 were both examined. The quantities included in the analysis were sea-level pressure, surface air temperature and precipitation.

Considerable differences were found in the ability of the models to reproduce the climatological sea-level pressure distribution in northern Europe. Of the five coupled models, the most realistic simulations were given by ECHAM2 and GFDL in winter, and by ECHAM2 in summer. After ECHAM2 and GFDL, the wintertime ranking in terms of performance of the coupled models was, according to the correlation coefficient and RMS error comparisons, UKMO, NCAR, and ECHAM1. In summer, however, no unequivocal

ranking after ECHAM2 was obtained. In the whole northern extratropics north of 30°N, the ECHAM2 simulation performed best in both seasons. In other respects, too, the comparison results for this wider area were mostly similar to those for northern Europe alone, although some minor differences were detected. As regards the two models with mixed-layer ocean, that of GFDL gave very similar results to the coupled model of the same modelling centre. There were, however, appreciable differences between the UKMO/coupled and UKMO/mixed pressure simulations, the latter of which was good in winter but poor in summer.

Most simulations substantially underestimate winter temperatures in northern Europe, and this error typically grows towards the northeast. The average cold biases for the land areas north of 50°N and west of 60°E are 12 °C for UKMO/coupled and about 9 °C for NCAR and UKMO/mixed. By far the most realistic winter temperatures are given by the two GFDL simulations, in which this bias is less than 2 °C. In summer, the temperature biases for the whole of northern Europe are much smaller than in winter, and their sign varies from model to model. Over Finland, however, the simulated temperatures tend to be some degrees too cold in summer as well.

The three models with precipitation data available (ECHAM1, ECHAM2, and NCAR) all qualitatively simulate the large-scale east-west gradient in annual precipitation, but the geographical details are not very reliable. The simulated seasonal cycle - a minimum in summer and a maximum in winter in all models - is also at variance with observations. The larger than observed winter precipitation is at least partly explained by the large systematic errors in measurement of snowfall, but the summertime rain deficit, which is most serious in the ECHAM1 and NCAR simulations, is an unambiguous model error.

In northern Europe, the greenhouse runs in the available experiments clearly indicate a larger increase of temperature in winter than in summer. In Finland, for example, the average warming for those four transient experiments in which the doubling of CO_2 was achieved is 3-4 °C in winter and slightly over 2 °C in summer. In the two equilibrium experiments, the warming is larger than it is in the transient experiments, but the difference is not as striking in continental northern Europe as over the Atlantic Ocean.

The ECHAM1, ECHAM2 and NCAR experiments all predict a slight increase of annual precipitation in Finland and also in most other parts of northern Europe. However, the differences between the results of the individual models allow few conclusions about the seasonal and geographical variations of this increase. Appreciable differences between models also exist in the simulated changes of wintertime sea-level pressure, but there are some indications of a northward shift of the Icelandic low. In summer, the models typically simulate a small but widespread pressure fall in northern Europe.

In constructing climate scenarios for impact studies, an important issue concerns the timing of the model results. With the exception of the NCAR experiment, the results of the transient greenhouse runs were from a time slice centred at the doubling of CO_2.

According to the IPCC "business-as-usual" scenario, which takes into account other greenhouse gases (e.g. CH_4, N_2O, and the CFCs) in addition to CO_2, an equivalent increase in radiative forcing is expected to take slightly over 60 years (*Houghton et al.* 1990, Policymakers Summary, Fig. 6). Nevertheless, the timing issue is a problematic one, at least for two reasons. First, many transient experiments suffer from the "cold start" phenomenon (*Hasselmann et al.* 1993): the warming in the first decades of the greenhouse run is much slower than it is at the later stages of the experiment. This is thought to be an artefact of the experimental design, which basically arises because transient greenhouse runs are generally started from an equilibrium stage rather than from an already warming stage. For the ECHAM1 experiment, for example, the resulting delay in the warming is estimated to be about 15 years (see Fig. 10 of *Hasselmann et al.* 1993). Second, what is perhaps even more important, climate has its own natural variability. To predict the long-term natural variations, accurate knowledge of the initial state of the oceanic circulation would obviously be needed. Since such knowledge is not available at present, it is actually more appropriate to regard the simulations as sensitivity tests rather than as genuine predictions.

Acknowledgements

The temperature and precipitation observations used in the comparison were extracted from the CDIAC data base by Heikki Tuomenvirta and partially processed by Eino Hellsten. The author is likewise grateful to Reijo Solantie for the information about precipitation measurement errors, and to Eero Holopainen and Timothy Carter, as well as the two anonymous reviewers, for their comments on the manuscript. The helpful attitude of Victor Ocana at DKRZ is also acknowledged.

References

Boer, G. J., K. Arpe, M. Blackburn, M. Déqué, W. L. Gates, T. L. Hart, H. le Treut, E. Roeckner, D. A. Sheinin, I. Simmonds, R. N. B. Smith, T. Tokioka, R. T. Wetherald, and D. Williamson, 1992: Some results from an intercomparison of the climates simulated by 14 atmospheric general circulation models. *J. Geohys. Res*, **97**, 12,771-12, 786.

Cubasch, U., K. Hasselmann, H. Höck, E. Maier-Reimer, U. Mikolajewicz, B. D. Santer, and R. Sausen, 1992: Time-dependent greenhouse warming computations with a coupled ocean-atmosphere model. *Climate Dynamics*, **8**, 55-69.

Finnish Meteorological Institute, 1984: Statistics of radiosonde observations 1961-1980. *Meteorological yearbook of Finland*, **61-80**, part 3. Government Printing Centre, Helsinki, Finland, 178 pp.

Hasselmann, K., R. Sausen, E. Maier-Reimer, and R. Voss, 1993: On the cold start problem in transient simulations with coupled atmosphere-ocean models. *Climate Dynamics*, **9**, 53-61.

Hoskins, B. J., H. H. Hsu, I. N. James, M. Masutani, P. D. Sardesmukh, and G. H. White, 1989: *Diagnostics of the Global Atmospheric Circulation Based on ECMWF analyses 1979-1989*. WCRP - 27, WMO/TD - NO. 326.

Houghton, J. T., G. J. Jenkins, and J. J. Ephraums (Eds.), 1990: Climate Change. The IPCC Scientific Assesment. Cambridge University Press, Great Britain, 365 pp.

Houghton, J. T., B. A. Callander, and S. K. Varney (Eds.), 1992: Climate Change 1992. The Supplementary Report to the IPCC Scientific Assesment. Cambridge University Press, Great Britain, 200 pp.

IPCC WG I, 1993: Evaluation of regional climate simulations: Terms of reference. An information letter distributed by the IPCC Technical Support Unit (Bracknell, UK) in August 1993.

Manabe, S., R. J. Stouffer, M. J. Spelman, and K. Bryan, 1991: Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. *J. Climate*, **4**, 785-818.

Manabe, S., M. J. Spelman, and R. J. Stouffer, 1992: Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. Part II: Seasonal response. *J. Climate*, **5**, 105-126.

Murphy, J. M., 1992: A prediction of the transient response of climate. *Climate Research Technical Note CRTN 32*. Hadley Centre, Bracknell, Great Britain, 27 pp.

Stouffer, R. J., S. Manabe, and K. Bryan, 1989: Interhemispheric asymmetry in climate response to gradual increase of atmospheric CO_2. *Nature*, **342,** 660-662.

Washington, W. M., and G. A. Meehl, 1989: Climate sensitivity due to increased CO_2: Experiments with a coupled atmosphere and ocean general circulation model. *Climate Dynamics*, **4**, 1-38.

Whetton, P. H., P. J. Rayner, A. B. Pittock, and M. R. Haylock, 1994: An assessment of possible climate change in the Australian region based on an intercomparison of general circulation modeling results. *J. Climate*, 7, 441-463.