

The effect of internal variability on anthropogenic climate projections

By ASGEIR SORTEBERG^{1*} and NILS GUNNAR KVAMSTØ², ¹*Bjerknes Centre for Climate Research, University of Bergen, Allègaten 70, N-5007 Bergen, Norway;* ²*Geophysical Institute, University of Bergen, Norway Bjerknes Centre for Climate Research, University of Bergen, Norway*

(Manuscript received 25 November 2005; in final form 27 June 2006)

ABSTRACT

A single model ensemble of five members, with a CO₂ increase of 1% per year, has been used to investigate the spread in climate change estimates. It is furthermore explored how this spread is related to internal variability of the climate system and the strength of the CO₂ forcing. As expected, the fraction of the globe where a statistically significant climate change could be detected increased with the strength of the CO₂ forcing. For temperature this fraction increased rapidly with integration time, while the increase of areas with statistically significant precipitation change was much slower. It is shown that increasing the averaging periods for estimating climate change from 20 to 40 yr reduced the spread due to internal variability by 40% in the Arctic regions. This is due to better sampling of decadal internal variability. In order to reduce sampling uncertainty below a fixed threshold, the number of ensemble members needed is much larger in the polar region than in the lower latitudes. This should be taken into account when designing modelling strategies (high-resolution modelling versus larger number of ensemble members) to reduce uncertainty in future polar climate change. At the time of doubling of CO₂, the variance of the climate change estimates, based on 20 yr averages from a single model ensemble, constitutes 13% and 42% of corresponding multimodel ensemble variances for temperature and precipitation, respectively. In the case of precipitation, this indicates that some caution should be used in attributing different climate change estimates to differences in model formulations.

1. Introduction

Simulation of future climate change encompasses a wide range of uncertainties. Internal variability of the real climate gives an upper limit to the accuracy of climate forecasts while additional uncertainty is associated with model deficiencies, future external forcings like solar variability, emission of greenhouse gases and particles. Given exact knowledge about future external forcing, the spread among model results may be due to real intermodel differences (parameterizations, level of sophistication and resolution), but may also be due to sampling of the natural internal variability of the model climates.

As computer capacity has increased, multiple realizations have been used to obtain uncertainty estimates of future climate projections. To span a representative range of possible realizations, such ensembles may be generated by a number of different models or by a single model with perturbed model physics and initial conditions.

In this study, we attempt to quantify the uncertainty contribution from internal variability using a single model with perturbed initial conditions. This contribution depends on several factors, including the length of the temporal averaging period, and the degree of spatial averaging (i.e. we expect zonal means to be less influenced by internal variability than grid point values). It seems likely that the increase of CO₂ with time will reduce the relative influence of the internal variability on the climate change estimates. Thus, the dependence of signal-to-noise ratio on an increasing forcing will be investigated by means of the generated five member ensemble. Any such dependency will be tested, for both 2 m temperature and precipitation. As mentioned, we also expect the climate signal to be better detected as the extent of geographical averaging increases. Obtaining a clear signal in this case will be at the expense of the detailed geographical and temporal information contained in the simulations. In this paper, we will however, not cover the effect of spatial averaging on the signal and noise, but rather focus on the effect of temporal averaging and the strength of the external forcing. An investigation of the behaviour of relative strengths of signal and noise will hopefully provide insight useful for designing more optimal simulations of global and regional climate change.

*Corresponding author.
e-mail: Asgeir.Sorteberg@bjerknes.uib.no
DOI: 10.1111/j.1600-0870.2006.00202.x

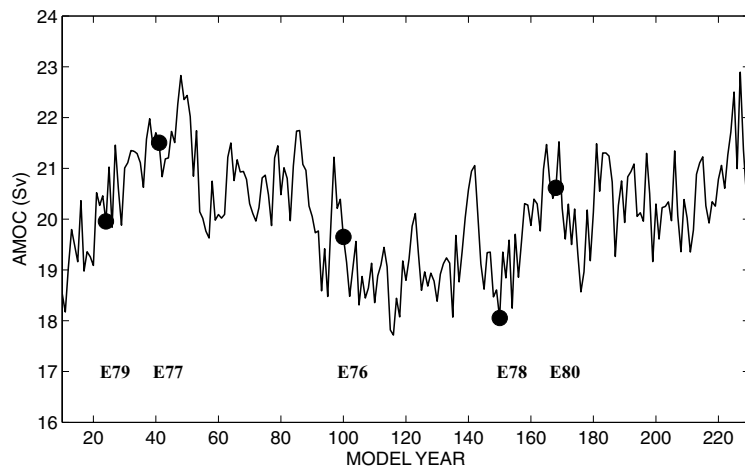


Fig. 1. The time series of simulated strength (Sv) of the AMOC in the control simulation with BCM. Arrows indicate the AMOC initial state for the five ensemble members named E76–80.

The methodology and experimental setup is given in Section 2 and the representation of the model variability is commented upon in Section 3. Section 4 discusses the dependence of the simulated signal and noise on temporal averaging and the external forcing strength, while Section 5 discusses the implication of this for the experimental design of climate change simulations. Section 6 compares the role of internal variability to model differences in climate change simulations, and finally, Section 7, discusses the role of internal variability in estimate spread from multimodel ensembles.

2. Methodology and experimental design

A five-member climate change ensemble simulation is carried out with different initial conditions taken from a 300-yr control simulation. All simulations are performed using the Bergen Climate Model (BCM; Furevik et al., 2003). Except for the 300-yr control run, which has constant external forcing, a 1% increase per year in atmospheric CO_2 has been applied to the five members of the climate change ensemble. All other atmospheric gases and aerosols are kept constant in the ensemble integrations. This is the same kind of experiment as was carried out in the second phase of the Coupled Model Intercomparison Project (CMIP2) (Meehl et al., 2000; Covey et al., 2003).

BCM consists of the atmospheric model ARPEGE/IFS (Déqué et al., 1994) coupled with a global version of the ocean model MICOM, including a dynamic–thermodynamic sea-ice model (Semtner, 1976; Hibler, 1979; Bleck et al., 1992). In this version of BCM, ARPEGE/IFS uses spectral truncation at wave number 63 with a horizontal grid mesh of about $2.8^\circ \times 2.8^\circ$ in latitude/longitude and 31 vertical levels from the surface to 10 hPa. MICOM is set up with a horizontal resolution of approximately $2.4^\circ \times 2.4^\circ$ with 24 layers in the vertical. Further technical descriptions together with descriptions of model climatology and variability in the control integration are given in Furevik et al. (2003).

The variability and stability of the Atlantic Meridional Overturning Circulation (AMOC) and other oceanic quantities in BCM and in MICOM-only runs have been investigated in a series of papers (Dutay et al., 2002; Gao et al., 2003; Nilsen et al., 2003; Bentsen et al., 2004). In general, the AMOC behaviour in the model system is found to be realistic. The AMOC sensitivity to CO_2 and fresh water forcing in BCM is at the lower end of the range when compared with other models (Otterå et al., 2003, 2004).

The initial conditions for the members of the BCM climate change ensemble have been taken from a 300-yr control integration. The strength of the AMOC is a good measure of the poleward oceanic heat transport and the members of the BCM climate change ensemble were initialized at different phases of the AMOC in the control run in order to span the internal variability of the AMOC (see Fig. 1). The maximum difference in AMOC initial state between the different members is around 3 Sverdrups (Sv) ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$). The members of the BCM climate change ensemble are integrated for 80 yr until doubling of CO_2 is reached. Results indicate that the BCM's AMOC has a 'memory' of one to two decades (Collins et al., 2006), thus the initial state of the AMOC is assumed to directly influence the simulation during the first few decades. However, the initial state may have indirect effects on the simulations for a longer time since it might affect the strength of other feedback processes in the system.

2.1. Analysis methods

At a given time (period), any simulated meteorological variable x , at a certain grid point or region, in the control run (ctl) can be written as the sum of a time mean (the hypothetical average over an infinite time period, $\overline{x_{\text{ctl}}}$) plus a time dependent deviation representing the variability:

$$x_{\text{ctl}} = \overline{x_{\text{ctl}}} + x'_{\text{ctl}}. \quad (1)$$

The time dependent component (x'_{ctl}) may in principle represent everything from hourly values to century scale anomalies. In this paper a control run average over a 30 yr period at the start of the individual CMIP2 integration is used as x_{ctl} . Thus, x_{ctl} will not be the same for each ensemble member and we denote the individual control periods as $x_{\text{ctl}(i)}$. A corresponding value in any of the BCM CMIP2 climate change ensemble members ($i = 1, 5$) can be written as a sum of the mean unperturbed climate ($\overline{x_{\text{ctl}}}$), the climate change signal (x_f) due to increased greenhouse gas concentrations and the internal climate variability of the ensemble member that is not averaged out by the chosen time and area averaging ($x'_{\text{cmip2}(i)}$):

$$x_{\text{cmip2}(i)} = \overline{x_{\text{ctl}}} + x_f + x'_{\text{cmip2}(i)}, \quad (2)$$

where $\text{cmip2}(i)$ represents the individual climate change simulation. The simulated difference between the control and each of the climate change simulation over a certain time can then be written as:

$$\Delta x_{\text{cmip2}(i)} = (x_{\text{cmip2}(i)} - x_{\text{ctl}(i)}) = x_f + (x'_{\text{cmip2}(i)} - x'_{\text{ctl}(i)}). \quad (3)$$

For the ensemble of climate change simulations, x_f is the anthropogenic signal which we want to detect and give a certain confidence value, while $(x'_{\text{cmip2}(i)} - x'_{\text{ctl}(i)})$ represent the component of the change which is due to differences in the internal variability of the control and climate change simulation over the chosen time period.

The simulated ensemble mean change may then be written as:

$$\langle \Delta x \rangle = x_f + \langle x'_{\text{cmip2}(i)} - x'_{\text{ctl}(i)} \rangle, \quad (4)$$

where $\langle \rangle$ indicates averaging over the ensemble members (in our case five simulations). We assume that the $\langle \Delta x \rangle$ estimate from n ensemble members is an unbiased estimate of the model's 'true' $\langle \Delta x \rangle$, that the ensemble members are independent and that there is no co-variability between the forced climate change component x_f and the variability term in eq. (2). Thus, an estimate of the ensemble's climate change variance ($\sigma_{\Delta x}^2$) is:

$$\begin{aligned} \sigma_{\Delta x}^2 &= \frac{\sum_{i=1}^n (\Delta x_{\text{cmip2}(i)} - \langle \Delta x \rangle)^2}{n-1} \\ &= \frac{\sum_{i=1}^n [(x'_{\text{cmip2}(i)} - x'_{\text{ctl}(i)}) - \langle x'_{\text{cmip2}(i)} - x'_{\text{ctl}(i)} \rangle]^2}{n-1}, \end{aligned} \quad (5)$$

where n is the number of simulations. This variance is thus a measure of the internal climatic *noise* that is left after the averaging (in this study $x'_{\text{cmip2}(i)}$ and $x'_{\text{ctl}(i)}$ are grid point averages over a few decades). For averaging periods of more than 20–30 yr there is no clear evidence in our simulations that the remaining internal variability is a function of the external forcing. An exception to this may be the Arctic region (this is further discussed in Section 5).

From the above expression it is clear that the climatic noise due to internal variability is a function of the spatial and temporal averaging of both the control and climate change simulations. In this study, a simulated climate change, Δx , is calculated as the

grid point (or regional) difference between a time mean value in a certain CMIP2 ensemble member (a 20 yr averaging period in all sections except Section 4.1, where we investigate the influence of sampling time on the climate change variability) and the mean value for the control integration over a 30 yr period at the start of the each CMIP2 simulation.

Under the assumption that the internal variability is independent of the strength of the external forcing, the signal-to-noise ratio may be calculated as the absolute value of the ratio between the ensemble mean change and the standard deviation of the individual climate change estimates:

$$\frac{S}{N} = \left| \frac{\langle \Delta x \rangle}{\sigma_{\Delta x}} \right|. \quad (6)$$

Assuming that the climate change estimates of the individual ensemble members (eq. 3) are normally distributed, we may calculate the $100(1 - \alpha)\%$ confidence interval for the true ensemble mean change given an infinite number of ensemble members ($\langle \Delta x \rangle_{\text{TRUE}}$) using the Student's t -test:

$$\langle \Delta x \rangle - t_{\alpha/2} \frac{\sigma_{\Delta x}}{\sqrt{n}} < \langle \Delta x \rangle_{\text{TRUE}} < \langle \Delta x \rangle + t_{\alpha/2} \frac{\sigma_{\Delta x}}{\sqrt{n}}, \quad (7)$$

where $t_{\alpha/2}$ is the upper $\alpha/2$ point of the t distribution and n the number of ensemble members.

Rearranging the above equation gives an estimate of the number of ensemble members required to calculate the climate change within a certain threshold ($\pm E$):

$$n = \left(\frac{t_{\alpha/2} \sigma_{\Delta x}}{E} \right)^2 \quad (8)$$

3. Simulated internal variability with BCM

An important issue in investigating the relationship between signal and noise is whether the model is able to generate a realistic internal variability. Together with BCM climatology, several aspects of the variability (e.g. AO and ENSO) are thoroughly discussed and found to be realistic in Furevik et al. (2003). In a previous study of the AMOC variability (Bentsen et al., 2004) and in an investigation of the NAO behaviour in different CMIP models (Kuzmina et al., 2005), BCM was found to be close to observations. The latter study also found that BCM responded with an increased NAO index to increased CO₂ levels. Further details on the high latitude response in the same experiment are reported in Sorteberg et al. (2005). As far as high frequency atmospheric variability is concerned, diagnostics shown in Doblas-Reyes et al. (1998) and Kvamstø et al. (2004) indicate that the atmospheric component of BCM exhibits realistic features in this regime as well. There is, however, a systematic error which is common to many GCMs, namely too zonally confined North Atlantic storm tracks (Doblas-Reyes et al., 1998; Lopez et al., 2000). It may also be noted that the zonally averaged annual temperature in the BCM control run (Fig. 2) exhibits strong decadal

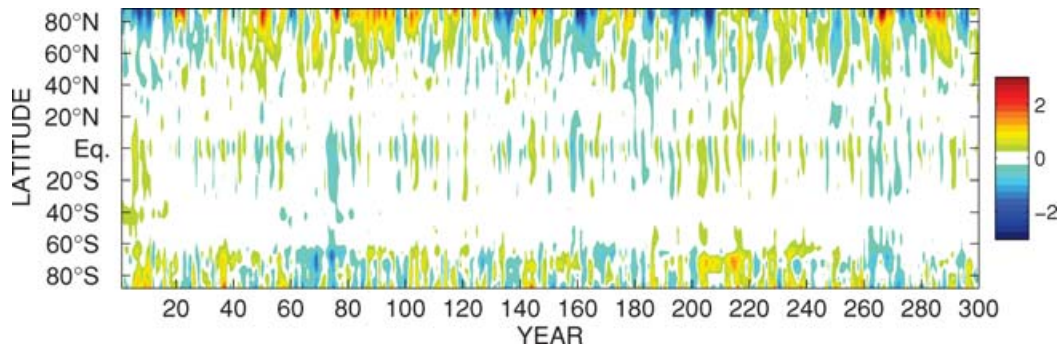


Fig. 2. Zonal means of annual 2 m temperature anomalies (°C) in the BCM control run.

fluctuations with amplitudes over 1°C in the Arctic region and shorter-term fluctuations in tropical temperatures due to ENSO. The simulated Arctic variability is comparable in magnitude to the observed variability (e.g. Johannessen et al., 2004, their fig. 1a).

As the representation of atmospheric variability in BCM can be taken to be state-of-the-art, we find the model to be a relevant tool for investigating the relationship between externally forced climate change and internal variability.

4. Mean response and signal-to-noise ratio

It is likely that the importance of internal variability relative to the ensemble mean climate change estimate ($\langle \Delta x \rangle$) will decline with time (and the growing strength of the external forcing). The relative importance of internal variability is also dependent on which variable one considers. Thus, in our 1% CO₂ increase per year simulations, we expect the signal-to-noise ratio (eq. 6) to increase with integration time and to be higher for weather parameters which are more strongly influenced by changes in external forcing. Figure 3 shows the BCM ensemble mean annual temperature and precipitation change (eq. 4), the standard deviation (eq. 5) and the signal-to-noise ratio (eq. 6) based on averages over years 60–79 when the external forcing is relatively strong (a doubling of CO₂). As in many other coupled models, the BCM has a global warming with a particularly pronounced amplification over the Arctic Ocean due to reduction of the sea-ice and changes in the boundary layer stratification (Fig. 3a). This is also the area containing the largest spread among the ensemble members (Fig. 3c). Hence, even though the Arctic may have a large sensitivity to anthropogenic greenhouse gas changes, its large internal variability makes it hard to attribute any climate change to changes in external forcings (natural or anthropogenic) when the forcing is weak. This is in agreement with other studies (e.g. Johannessen et al., 2004). However, as the forcing becomes stronger, the BCM simulations identify parts of the tropics and the Arctic as having the largest signal-to-noise ratio for temperature (Fig. 3e) and therefore (given a case with a globally equal distribution of observations) the regions where

one most easily could attribute a climate change to changed CO₂ levels. The picture of climate change is different for precipitation where the most pronounced changes are in the tropical areas (Fig. 3b) associated with an intensified wintertime Hadley cell (not shown). In the tropics the spread among the different members (Fig. 3d) appears to be correlated with the amplitude of the mean change, which inhibits high signal-to-noise ratios (Fig. 3f). The Arctic region has the largest signal-to-noise ratio for precipitation (Fig. 3e), but the signal-to-noise ratio is generally smaller than that for temperature worldwide.

Figure 4 shows how the area fraction of the earth, where 20-yr ensemble mean temperature and precipitation changes are significantly (95% level) different from zero (eq. 7), depends on integration time. Also included is the corresponding area where single realizations have significant changes (using the variance of the ensemble members, eq. 5). For a single realization it is possible to detect a significant annual mean temperature change over 60% of the earth's area around year 20 (mean over years 10–30). The corresponding area fraction of precipitation changes is about 15% and increases rather slowly with the strength of the external forcing to about 40% at year 70 (CO₂ doubling). As expected, the area fraction of the earth experiencing significant seasonal change is lower than that for annual change. During weak CO₂ forcing, the temperature change signal appears most robust during autumn and summer while the precipitation signal is strongest in winter and weakest in spring and autumn. The far lower fraction with significant precipitation change (compared to temperature change) reflects both the strong internal variability of precipitation and the weaker (relative to the internal variability) influence of increased CO₂ levels.

By comparing the area of significant change in the ensemble mean (lines) to corresponding areas for single realizations (shaded curves indicates the maximum and minimum significant area in the five realizations), the advantage of averaging is clearly evident (Figs. 4a and b). This is particularly pronounced for precipitation. The actual geographic areas where the results show statistically significant change around doubling of CO₂ (year 70 in Fig. 4) is given by the areas having a signal-to-noise ratio above 2 in Figs. 3e and f.

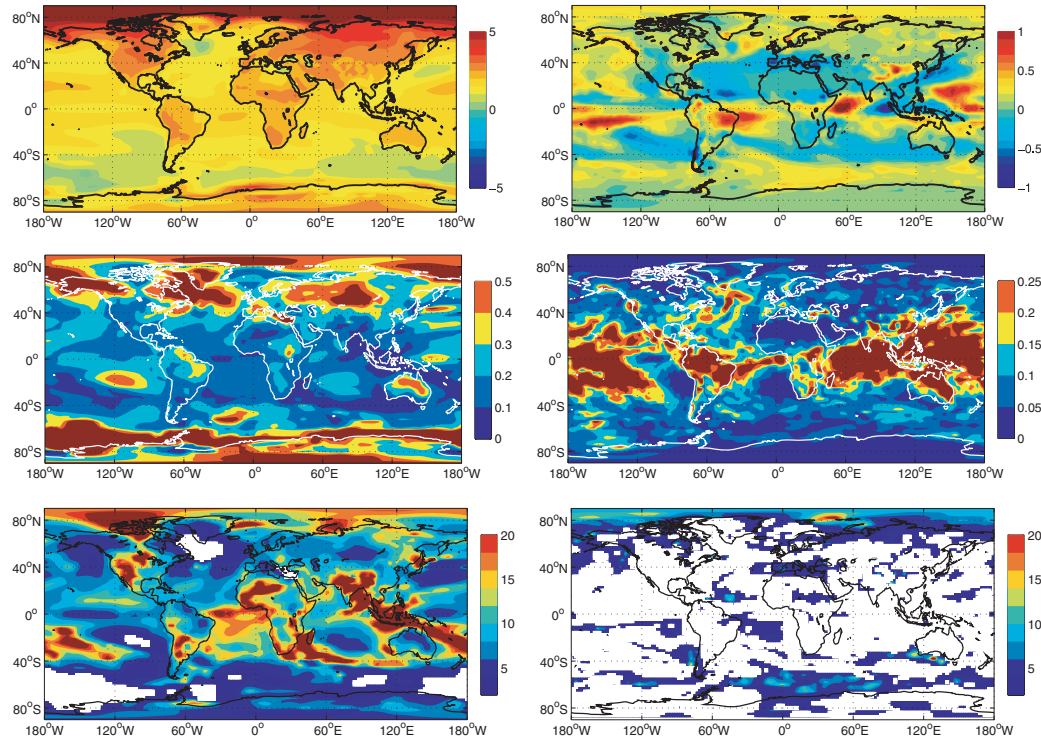


Fig. 3. Ensemble mean change of (a) 2 m temperature (T2m, °C) and (b) precipitation (mm d⁻¹). The standard deviation (square root of eq. 5) of estimated changes from the five ensemble members are shown for T2m (°C) in (c) and for precipitation (mm d⁻¹) in (d). (e) Signal-to-noise ratio (eq. 6) for T2m (°C⁻¹) and (f) signal-to-noise ratio for precipitation (mm d⁻¹). Areas in (e) and (f) where the signal-to-noise ratio is less or equal to 2 are masked with white colour. The mean changes and standard deviations are taken over year 60–79 (at doubling of CO₂) in the ensemble members. Changes are calculated relative to 30 yr means of the control run taken at the start of each CMIP2 simulation.

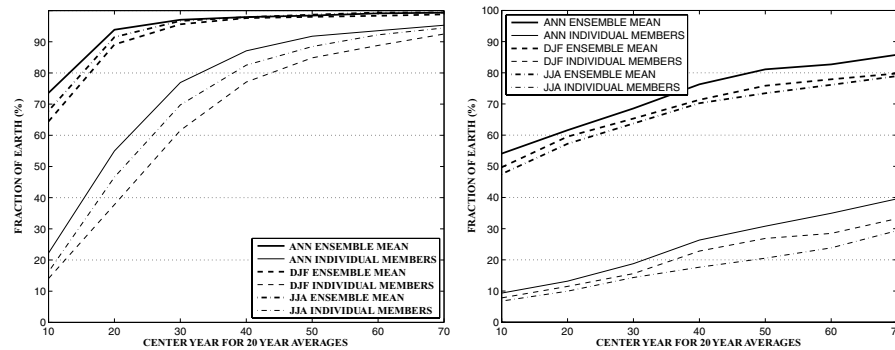


Fig. 4. Fraction (%) of global area which has a 95% significant change in (a) T2m and (b) precipitation in different seasons and the whole year. The heavy lines show the significant fractions for the ensemble mean (five members), while the light lines show the average fraction for the individual ensemble members. The fractions are shown as a function of integration time by calculating changes over running 20 yr periods (relative to the constant 30-yr control run means) in the ensemble members.

4.1. Impact of averaging time and external forcing

It is still the case that limited computer resources require a choice between high complexity (and resolution) and ensemble integrations with lower resolution in the design of numerical experiments for identifying climate change. Most regional downscaling simulations for instance, are performed over time slices of 15–30 yr (Giorgi et al., 2001). A reduction of noise, due to in-

ternal variability, can be obtained with a larger integration (and averaging) length. Under the assumption of purely stochastic internal variability with no serial correlation, the dependence of the climate change variance on averaging period (Δt) is given as:

$$\sigma_{\Delta x, \Delta t}^2 = \sigma_{\Delta x, \tau}^2 \left[\frac{(1/\Delta t) + (1/\Delta t_{\text{cl}})}{(1/\tau_{\text{cmip2}}) + (1/\tau_{\text{cl}})} \right]. \quad (9)$$

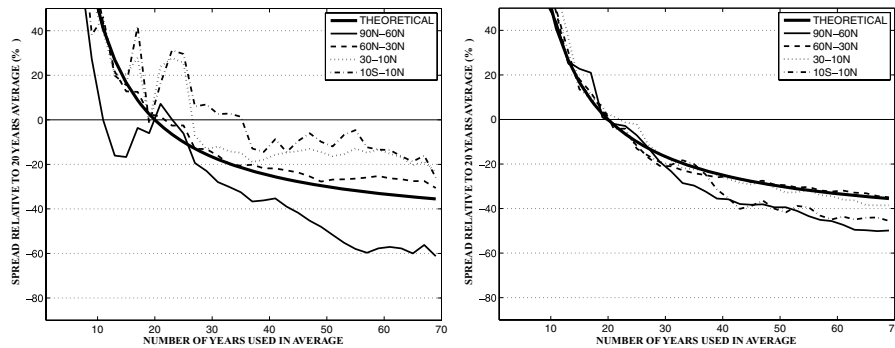


Fig. 5. (a) The mean T2m spread (given as the mean grid point change in the variance, eq. 5) for the different ensemble members as function of averaging period. (b) Same as (a), but for precipitation. The thick solid line shows the spread given by the theoretical function in eq. (9). The other lines are calculated by using eq. (5). The variance (in four latitude bands) is given in % relative to a 20 yr mean. Calculation of the variances are centred at year 40 (e.g. year 30 on the x -axis is the climate change variance calculated by using the mean changes in each simulation over year 25–55 relative to the mean 20 yr average over year 30–50). Changes are calculated for different periods centred at year 40 in the BCM CMIP2 simulations, with the x -axis giving the length of the averaging period.

In this theoretical expression τ is the reference averaging period for the control (τ_{ctl}) and CMIP2 (τ_{cmip2}) simulations which in our case have been set to 30 and 20 yr, respectively. Employing the reference values in eq. (5) gives the reference variance ($\sigma_{\Delta x, \tau}^2$). The theoretical variance estimate for $\Delta t \in [10, 70]$ yr is shown in Fig. 5 (note that $\Delta t_{\text{ctl}} = \tau_{\text{ctl}}$ for all values of Δt). For comparison, the variance (grid point spread) using eq. (5) for averaging lengths between 10 and 70 yr with the same τ_{ctl} is also included in Fig. 5. Estimates, using both eq. (5) and (9), were obtained for different latitudes.

In general, the spread decreases with averaging length in the same way as the theoretical estimate. The decrease is steep for the shortest intervals, and levels off asymptotically for values larger than about 30–50 yr. For annual 2 m temperatures, the variance in the climate change estimates for a 40 yr average is reduced by around 40% in the Arctic, 25% at mid-latitudes and 10–15% in the tropical and subtropical regions relative to the 20 yr average (Fig. 5a). The increased variance reduction for larger averaging times in the Arctic compared to the other regions may be explained by the model's larger decadal temperature variations in this region (Fig. 2). It is not clear why the tropical and subtropical regions have a smaller reduction than the theoretical approximation (eq. 9). One can only speculate that most of the variance is contained in timescales shorter than 10 yr and/or can be associated with sensitivity to the chosen reference interval (20 yr), but that has not been tested here.

The reduction of precipitation spread when increasing the averaging interval is comparable to that of temperature (Fig. 5b). The reduction of subtropical and mid-latitude variance with larger averaging interval follows the theoretical curve closely, while the Arctic and tropical variance reduction is slightly larger. For the latter regions, an increase in averaging interval from 20 to 40 yr gives a variance reduction of 40% in the simulated climate change estimates (Fig. 5b).

5. Dependence on ensemble size

As seen in Fig. 4, the areas where we can detect significant changes are small when the ensemble size is small and the external forcing is weak. Since the strength of the internal variability has a pronounced geographical distribution, the number of ensemble members needed to average out the internal variability will also have a geographical dependence. An approximate number of ensemble members needed to estimate the ensemble mean change at a certain confidence level within an estimate range $\pm E$ is given by eq. (8). This number will depend on the extent of temporal and spatial averaging used prior to the variance estimation. It is also likely that this number will vary with different models as well. Figure 6 shows the results from using eq. (8) to calculate the median number of ensemble members needed to estimate the annual mean temperature and precipitation change in individual grid squares within ± 0.2 K and ± 35 mm yr $^{-1}$, respectively for four latitude bands. The results are based on 20 yr averages during different stages in the CMIP2 integrations. For temperature (Fig. 6a) the two most prominent features are the larger number of ensemble members needed as we go northward and the fact that the ensemble size needed is quite insensitive to the strength of the CO $_2$ forcing, possibly with the exception of the Arctic region. The first point stresses the increase of decadal-scale internal variability with latitude and it is furthermore seen that the ensemble size must be increased by a factor of 6–8 to obtain the same level of certainty in the Arctic as in the tropics and subtropics. For precipitation the results are reversed. The tropical areas have a much stronger internal variability and therefore require a larger number of ensemble members there.

Given the fact that we have only five ensemble members, the uncertainty of the variance estimates and therefore the ensemble size needed, are large and should only be taken as indicative values.

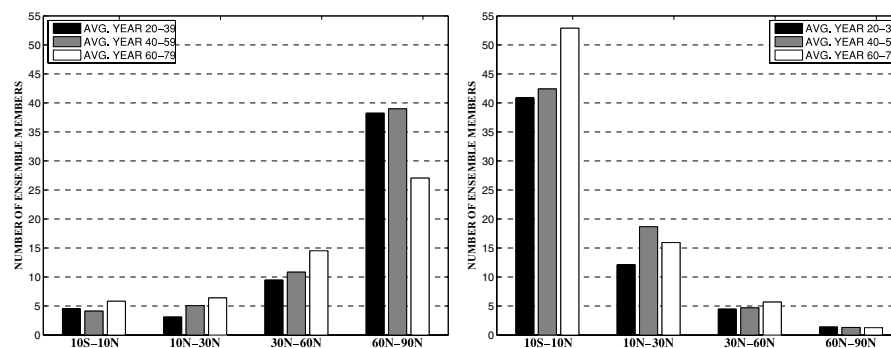


Fig. 6. (a) Median number of ensemble members needed (eq. 8) to estimate grid point temperature changes within ± 0.2 K with a 95% confidence for different latitude bands and CO₂ levels. (b) Same as (a) but for precipitation and ± 35 mm yr⁻¹ as uncertainty interval. The median number of ensemble members is based on 20-yr annual mean changes

The results indicate that the number of members needed is quite insensitive to the strength of the CO₂ forcing, possibly except for the Arctic. This is consistent with the assumption that the decadal-scale internal variability is not severely affected by the external forcing (eq. 5). This may not be true for the Arctic, as a reduced sea-ice and snow cover extent may lead to a reduced sea-ice albedo feedback and associated reduced decadal variability in the oceanic or atmospheric heat transport into the Arctic at the end of the simulation. Such indications are in agreement with the result of Räisänen (2002), who showed a predominant (but weak) decrease in interannual temperature variability in high latitudes around the doubling of CO₂. To test this one of our ensemble members was continued until 10 times CO₂ was reached. This integration showed that the internal variability in the Arctic was gradually reduced. Thus, in the case of the Arctic, the argument that the CO₂ forcing does not influence the internal variability will probably hold only for a moderate increase of CO₂ levels.

Note the strong dependence of the ensemble size on estimate range (E) in eq. (8). Thus, if the threshold is reduced with a factor of 2, the number of ensemble members needed would increase by a factor of 4.

6. The role of internal variability in multimodel ensemble spread

Figure 7 provides precipitation and temperature change from our five-member CMIP2 ensemble and for 15 other CMIP models (Covey et al., 2003) with similar experimental design. The change estimates are based on averages over years 20–39 and 60–79. Except in the Arctic, the BCM simulations give generally lower estimates than the multimodel ensemble mean precipitation change. BCM gives values which are fairly close to the multimodel mean temperature change. The spread among the different members of the BCM climate change ensemble, particularly when the CO₂ forcing is weak (Fig. 6a, c and e), constitutes a substantial fraction of the multimodel spread even when the data is averaged over such wide areas. As we approach grid point level, the ratio between the BCM spread and multi-

model ensemble spread increases. The average grid point ratio between the mean grid point variance in the climate change simulations (eq. 5) in the single model and multimodel variance using 20 yr averages is shown in Fig. 8. The spread among the different BCM ensemble members during weak forcing (years 20–39) is above 30% and 70% of the multimodel spread in all areas for temperature and precipitation, respectively. Räisänen, (2001) also found internal variability to be the dominating uncertainty source in early stages of the CMIP simulations. For temperature, the ratio is initially largest in the Arctic. At this stage the grid point BCM spread is almost 80% of the multimodel spread. It drops, however, to 5% around doubling of CO₂. The fact that the variance of the climate change estimates of the BCM ensemble spans a large range of the multimodel variance in early stages reflects the strong internal variability of BCM in this area. For precipitation, the BCM spread is higher (108%) than the multimodel variance in the tropics and the mid-latitudes when the CO₂ forcing is weak. This indicates that the internal precipitation variability simulated with BCM in these regions is high in comparison to many other models. As the CO₂ forcing increases (year 60–79), the ratio of the single to multimodel spread drops to around 20–45% (Fig. 8b). The spread calculations here are comparable to the IPCC values reported in Houghton et al. (2001), building on the indirect estimates of Räisänen (2001). Räisänen (2001) estimated that for annual grid point values, approximately 10% of the temperature and 30% of the precipitation spread (variance) in the multimodel climate change estimates may be attributed to internal variability at the time of doubling of CO₂. Our results for the whole globe indicate 7% and 38%, respectively. The slightly higher value for precipitation in this study is in line with the notion that the BCM has slightly stronger decadal internal variability in precipitation than many other coupled models participating in the CMIP project.

7. Conclusions

Using a single model ensemble with five members, the spread in the climate change estimates may be interpreted as the effect

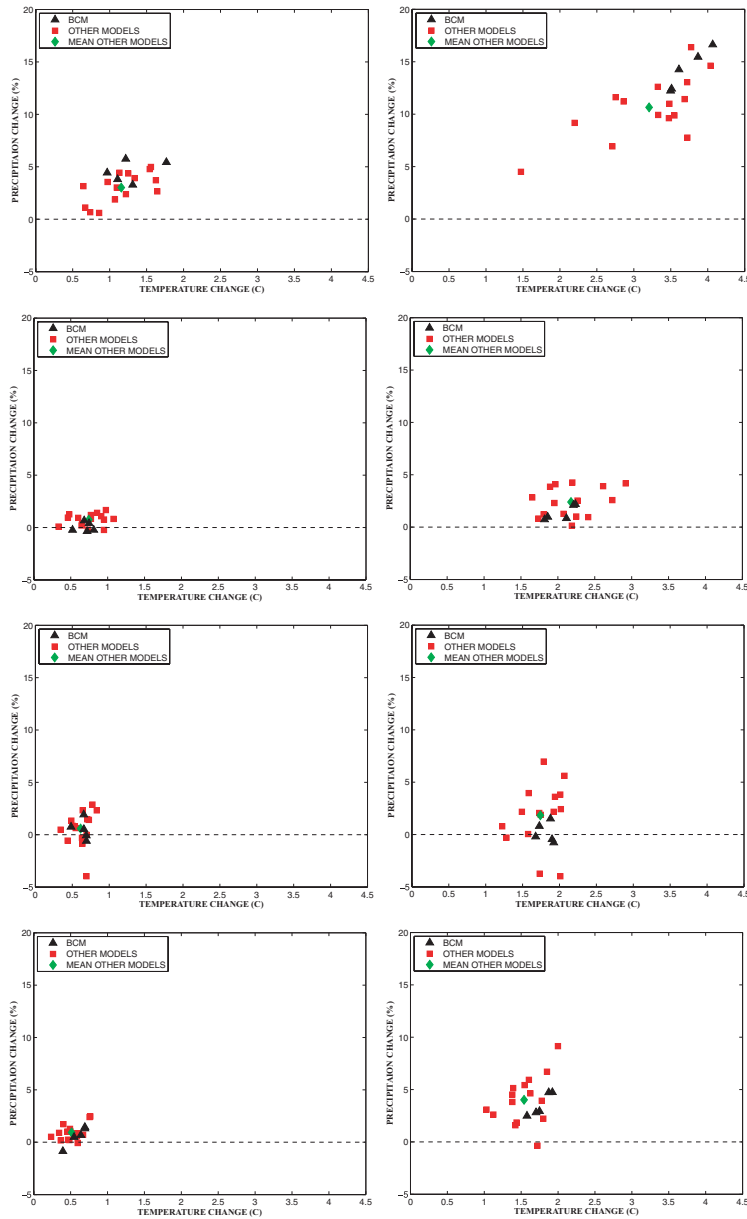


Fig. 7. Mean area averaged temperature and precipitation changes in the different BCM members (triangle) and from 15 other coupled models (squares) during weak (annual values averaged over year 20–39) and strong (annual values averaged over year 60–79) CO_2 forcing. Changes under weak (strong) forcing are shown in the left (right) column. Different panel rows show data in different latitude bands.

of internal variability. The ensemble simulation was carried out with a 1% CO_2 increase per year and the different members were initialized at different phases of the AMOC to ensure that the initial state of the ocean heat transport spanned the internal variability of the model. Due to the relatively small ensemble size, we have mainly focused on grid point mean statistics in different latitude bands.

By utilizing the framework of normal distributions, we have investigated the relation between simulated signal-to-noise ratio and geographical areas of statistically significant climate change. The largest warming was found in the Arctic where the spread among the different individual realizations was also large. For temperature, the signal-to-noise ratio at doubling of CO_2 was

largest in the tropical areas and the high Arctic and smallest at northern mid-latitudes and in the Southern Ocean. For precipitation the largest changes (in absolute values) are seen in the tropics due to an intensification of the Hadley cell. In relative terms (percentage change) the changes are largest in the Arctic and this is also the area where the signal-to-noise ratio is high. As expected, the area where a significant climate change could be detected increased with the strength of the CO_2 forcing and the number of ensemble members. The rate of increase is, however, dependent on which meteorological variable we consider. The ensemble mean temperature change (20 yr average) was significant in over 90% of the global area after 20 yr of integration. The same value for a single integration was 60%. In contrast,

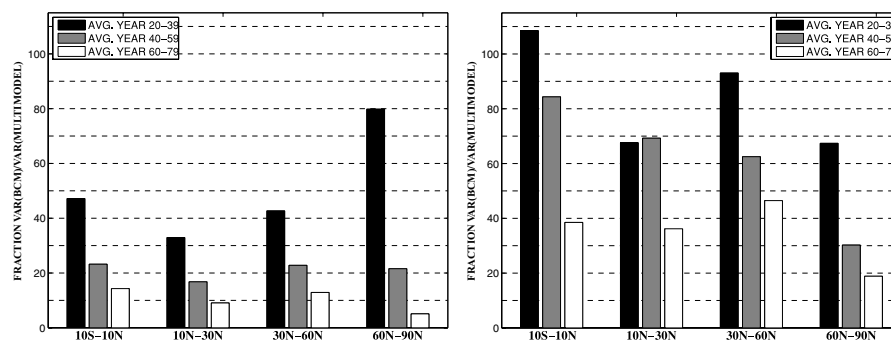


Fig. 8. (a) Average ratio (%) between the mean grid point variance of the 2 m temperature change estimates in the BCM ensemble and a multimodel ensemble. (b) Same as (a), but for precipitation. The spread is calculated using eq. (5). The variance of the climate change estimates are based on 20 yr averages if the CMIP2 simulations and 30 yr for the control simulation and shown as function of CO₂ forcing (weak: black, medium: grey and strong: white).

the fraction of the globe having a significant global precipitation response was 60% and 15% for the ensemble mean and individual realizations, respectively. At the time of CO₂ doubling, only 40% of the globe had a significant precipitation change in the individual realizations.

In general, the spread in the estimated climate change was reduced by increasing the length of the averaging period. The gross reduction was in accordance with the theoretical function derived under the assumption of internal variability being purely stochastic with no serial correlation. There is less agreement between the computed and theoretical ensemble spread for the Arctic and mid-latitude precipitation with a larger decline in the computed curve. This is in line with the large decadal internal variability in the region. However, due to the low number of ensemble members and the arbitrary choice of a 20 yr averaging (reference) period, one cannot rule out the possibility that the deviations between the curves are a result of sampling effects. The variance of temperature change estimates was typically reduced by 25% by increasing the averaging period from 20 to 40 yr. The impact on the spread of the precipitation change estimates was 40% for the mid- and high-latitudes and 25% for the tropics and subtropics.

It was found that the uncertainty of estimated climate change due to insufficient sampling of internal variability had a pronounced latitudinal dependence. To obtain the same level of absolute accuracy for an Arctic temperature change as for a tropical or subtropical temperature change, the number of ensemble members should be increased by a factor of 6–8. For precipitation this latitudinal dependence is reversed and hence the largest ensemble size is required in the tropics. It should be noted that if the accuracy of relative precipitation change was used instead of the absolute change, the regions influenced by the ITCZ, monsoonal or stormtrack variability would need the largest ensemble size.

For a moderate CO₂ increase (i.e. a CO₂ increase of 60% from the reference level around year 40–59) and averaging length of 20 yr, the grid point spread in climate change estimates for a sin-

gle model ensemble constitutes 20–30% of a multimodel ensemble spread for annual temperature and 50–90% for precipitation. Caution should therefore be taken in attributing different climate change estimates only due to differences in model formulations when the forcing is moderate. For precipitation this also holds for stronger CO₂ forcing.

8. Acknowledgments

This work has been done with support from the RegClim project (<http://regclim.met.no/>) funded by the Research Council of Norway. It has also received support from the Norwegian Super computing Committee (TRU) through a grant of computing time on an IBM Regatta computer at Parallab, University of Bergen. The CMIP modelling groups and the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at the Lawrence Livermore National Laboratory are acknowledged for conducting and making available the CMIP multimodel ensemble. The authors would also like to thank two anonymous reviewers for helpful and constructive comments. Finally, we acknowledge Dr. Alastair Jenkins for helpful comments on the text.

References

- Bentsen, M., Drange, H., Furevik, T. and Zhou, T. 2004. Simulated variability of the Atlantic meridional overturning circulation. *Clim. Dyn.*, **22**, 701–720.
- Bleck, R., Rooth, C., Hu, D. M. and Smith, L. T. 1992. salinity-driven thermocline transients in a wind-forced and thermohaline-forced isopycnal coordinate model of the north-Atlantic. *J. Phys. Oceanogr.* **22**, 1486–1505.
- Collins, M., Botzet, M., Carril, A., Drange, H., Jouzeau, A. and co-authors. 2006. Interannual to decadal climate predictability: a multi-perfect-model-ensemble study. *J. Climate* **19** (7), 1195–1203.
- Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P., Lambert, S. J. and co-authors. 2003. An overview of results from the Coupled Model Intercomparison Project. *Global Planet. Change* **37**, 103–133.

- Déqué, M., Dreveton, C., Braun, A., Cariolle, D., 1994. The ARPEGE/IFS Atmosphere Model: a contribution to the French Climate Modelling Community. *Clim. Dyn.* **10**, 249–266.
- Doblas-Reyes, F. J., Déqué, M., Valero, F. and Stephenson, D. B. 1998. North Atlantic wintertime intraseasonal variability and its sensitivity to GCM horizontal resolution. *Tellus* **50A**, 573–595.
- Dutay, J. C., Bullister, J. L., Doney, S. C., Orr, J. C., Najjar, R. and co-authors. 2002. Evaluation of ocean model ventilation with CFC-11: comparison of 13 global ocean models. *Ocean Model.* **4**, 89.
- Furevik, T., Bentsen, M., Drange, H., Kindem, I. K. T., Kvamstø, N. G. and co-authors. 2003. Description and evaluation of the Bergen climate model: ARPEGE coupled with MICOM. *Clim. Dyn.* **21**, 27–51.
- Giorgi, F., Hewitson, B., Christensen, J., Hulme, M. von Storch, H. and co-authors. 2001. Regional Climate Information: Evaluation and Projections (Chapter 10). In *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the IPCC* (eds J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell and C. A. Johnson) Cambridge University Press, Cambridge, pp. 739–768.
- Gao, Y. Q., Drange, H. and Bentsen, M. 2003. Effects of diapycnal and isopycnal mixing on the ventilation of CFCs in the North Atlantic in an isopycnal coordinate OGCM. *Tellus* **55B**, 837–854.
- Hibler, W. D. 1979. Dynamic Thermodynamic Sea Ice Model. *J. Phys. Oceanogr.* **9**, 815–846.
- Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., Dai, X., Maskell, K. and Johnson, C. A. 2001. *Climate Change 2001: The Scientific Basis*. Cambridge University Press, p. 881.
- Johannessen, O. M., Bengtsson, L., Miles, M. W., Kuzmina, S. I., Semenov, V. A. and co-authors. 2004. Arctic climate change: observed and modelled temperature and sea-ice variability. *Tellus* **56A**, 328–341.
- Kuzmina, S. I., Bengtsson, L., Johannessen, O. M., Drange, H., Bobylev, L. P. and Miles, M. W. 2005. The North Atlantic Oscillation and greenhouse-gas forcing. *Geophys. Res. Lett.* **32** (4), L04703, doi: 10.1029/2004GL021064.
- Kvamstø, N. G., Skeie, P. and Stephenson, D. B. 2004. Impact of Labrador sea-ice extent on the North Atlantic oscillation. *Int. J. Climatol.* **24**, 603–612.
- Lopez, P., Schmith, T. and Kaas, E. 2000. Sensitivity of the Northern Hemisphere circulation to North Atlantic SSTs in the ARPEGE climate AGCM. *Clim. Dyn.* **16**, 535–547.
- Meehl, G. A., Boer, G. J., Covey, C., Latif, M. and Stouffer, R. J. 2000. The Coupled Model Intercomparison Project (CMIP). *Bull. Am. Meteor. Soc.* **81**, 313–318.
- Nilsen, J. E. O., Gao, Y., Drange, H., Furevik, T. and Bentsen, M. 2003. Simulated North Atlantic-Nordic Seas water mass exchanges in an isopycnal coordinate OGCM. *Geophys. Res. Lett.* **30** (10), 1536, doi: 10.1029/2002GL016597.
- Otterå, O. H., Drange, H., Bentsen, M., Kvamstø, N. G. and Jiang, D. 2003. The sensitivity of the present-day Atlantic meridional overturning circulation to freshwater forcing. *Geophys. Res. Lett.* **30** (17), 1898, doi: 10.1029/2003GL017578.
- Otterå, O. H., Drange, H., Bentsen, M., Kvamstø, N. G. and Jiang, D. B. 2004. Transient response of the Atlantic meridional overturning circulation to enhanced freshwater input to the Nordic Seas-Arctic Ocean in the Bergen climate model. *Tellus* **56A**, 342–361.
- Räisänen, J. 2001. CO₂-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability. *J. Climate* **14**, 2088–2104.
- Räisänen 2002. CO₂-induced changes in interannual temperature and precipitation variability in 19 CMIP2 experiments. *J. Climate* 2395–2411.
- Semtner, A. J. 1976. Model For Thermodynamic Growth Of Sea Ice In Numerical Investigations Of Climate. *J. Phys. Oceanogr.* **6**, 379–389.
- Sorteberg, A., Furevik, T., Drange, H. and Kvamstø, N. G., 2005. Effects of simulated natural variability on Arctic temperature projections. *Geophys. Res. Lett.* **32** (18), L18708, doi: 10.1029/2005GL023404.