

Multi-model decadal potential predictability of precipitation and temperature

G. J. Boer¹ and S. J. Lambert¹

Received 10 January 2008; accepted 31 January 2008; published 5 March 2008.

[1] A first multi-model estimate of the long timescale potential predictability of precipitation is obtained based on over 8000 years of data from the control simulations of 21 state-of-the-art coupled climate models. The analysis also updates earlier estimates of the potential predictability of temperature to provide a consistent estimate for these basic climate parameters. Long timescale potential predictability is found mainly over the oceans at middle to high latitudes and predominantly where the surface is connected to the deeper ocean. Precipitation's modest potential predictability resembles an attenuated version of that for temperature on these timescales. Over land, predictability is largely absent for precipitation and comparatively weak for temperature where it is found over the northern and western parts of northern hemisphere land masses bordering the oceans. Regions exhibiting potential predictability indicate where we may hope to find predictive skill at long timescales and also points indirectly to the processes involved. **Citation:** Boer, G. J., and S. J. Lambert (2008), Multi-model decadal potential predictability of precipitation and temperature, *Geophys. Res. Lett.*, 35, L05706, doi:10.1029/2008GL033234.

1. Introduction

[2] A first analysis of the long timescale potential predictability of precipitation is undertaken together with an update and extension of previous results for temperature. We adopt a multi-model ensemble approach following the methodology of Boer [2004] and apply it to the collection of coupled model data, referred to as the CMIP3 dataset [Meehl et al., 2007], generated for the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report [Intergovernmental Panel on Climate Change (IPCC), 2007]. The result is an internally consistent estimate of the long timescale potential predictability for these two basic climatological parameters.

[3] Potential predictability is not predictability in the classical sense but an indirect approach which attempts to quantify the fraction of long-term variability that may be distinguished from the internally generated unpredictable natural variability "noise" at the timescale of interest. The long term variability "signal" that rises above this noise, if it exists and is of appreciable magnitude, is presumed to be the result of processes in the physical system which are predictable with sufficient knowledge and information. Potential predictability has been studied in the observations of both real and modelled systems [e.g., Madden, 1976;

Zwiers, 1996; Rowell, 1998; Rowell and Zwiers, 1999; Boer, 2004] and there is evidence of its correspondence with standard predictability [e.g., Griffies and Bryan, 1997; Boer, 2000; Collins, 2002; Collins and Sinha, 2003]. Model-based investigations of long timescale potential predictability are undertaken because the observational record is too short to permit reliable estimates. The results apply to the real system to the extent that the models appropriately reproduce the variability of the coupled climate system.

[4] The assumption is that the unforced variability of the system may be represented in the form $X = \nu + \varepsilon$ where ν is an internally generated long timescale potentially predictable component and ε an unpredictable shorter timescale "noise" component with variances $\sigma^2 = \sigma_\nu^2 + \sigma_\varepsilon^2$. The potentially predictable variance fraction *ppvf* is $p = \sigma_\nu^2 / \sigma^2$ and this is what we seek to estimate. We test the null hypothesis that $p = \sigma_\nu^2 = 0$ and estimate confidence limits $p_l < p < p_u$. The basic results are plots of statistically stable values of the *ppvf* where it is of sufficient magnitude to be of interest.

2. Coupled Models and Simulations

[5] The data used in the study consist of the preindustrial control simulations of the models listed in Table 1 from

Table 1. Control Simulation Data Available in the WCRP CMIP3 Dataset

Model Identifier	Control (years)	Initial Data Resolution
bccr	1850–2099	128 × 64
cccma1	1850–2850	96 × 48
cccma2	1850–2199	128 × 64
cnrm	1930–2429	128 × 64
csiro	1871–2250	192 × 96
gfdl1	1–500	144 × 90
gfdl2	1–500	144 × 90
giss_am	1850–2100, 1850–2199	90 × 60
giss_ah	1880–2279	72 × 46
giss_er	1901–2400	72 × 46
iap	1850–2199, 1850–2199, 1850–2199	128 × 60
ingv	1761–1860	320 × 160
inmcm	1871–2200	73 × 45
ipsl	1860–2359	96 × 72
miroc_hr	1–100	320 × 160
miroc_mr	2300–2799	128 × 64
miub	1860–2200	96 × 48
mpi	2150–2655	192 × 96
mri	1851–2200	128 × 64
ncar_c	280–509, 300–799	256 × 128
ncar_p	451–1039	128 × 64
ukmo_h	1859–2199	96 × 73
ukmo_g	1860–2099	192 × 145

¹Canadian Centre for Climate Modelling and Analysis, Environment Canada, University of Victoria, Victoria, British Columbia, Canada.

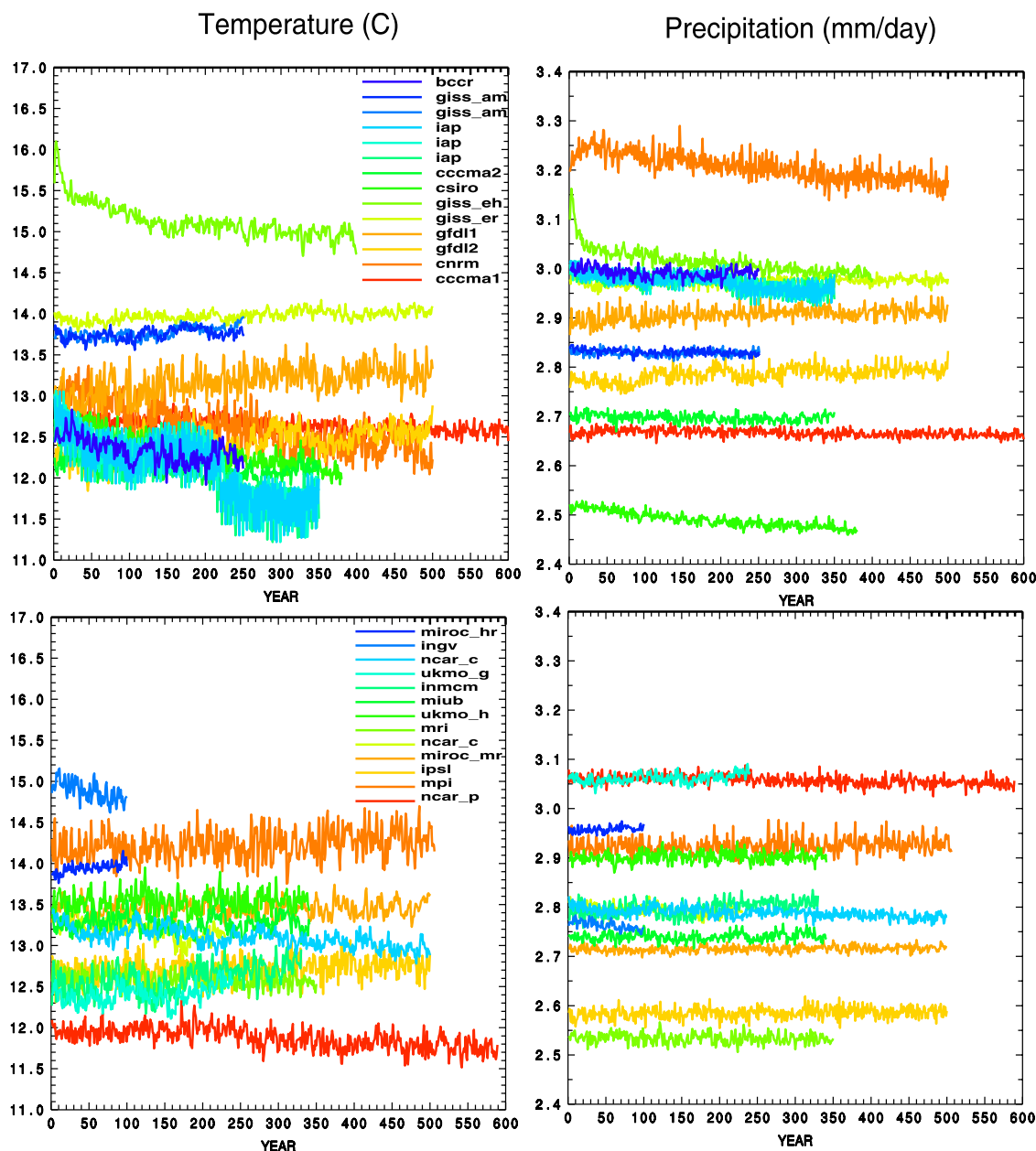


Figure 1. Timeseries of global annual mean temperature ($^{\circ}\text{C}$) and precipitation (mm/day) from the preindustrial control simulations of current climate models in the CMIP3 data set. The results are plotted over two panels to avoid excessive overlap of the curves.

the CMIP3 dataset. In brief, the control simulations are intended to be simulations of the equilibrium climate for “preindustrial” conditions which differ only modestly from current conditions. A number of modelling groups provide multiple control simulations although most provide a single long simulation. The data are re-expressed on a common 128×64 grid for the analysis described here.

[6] The Multi-Model Ensemble (MME) approach pools the statistics of the collection of individual models to better estimate climate population parameters. Pooled multi-model statistics are generally in better agreement with observation-based estimates than are the results from individual models. This has been demonstrated for climatological mean quan-

ties [Lambert and Boer, 2001] and also for second order climatological statistics (variances and covariances) that arise in the Lorenz energy cycle in the work by Boer and Lambert [2008]. The MME approach is also adopted in seasonal forecasting applications (Special Issue on DEMETER, *Tellus, Series A*, 57, 2005) and in the climate applications of IPCC [2007, chap. 10].

[7] The control simulations listed in Table 1 are intended to represent an equilibrium climate state and to exhibit the natural variability of the climate system unforced by external factors such as GHG changes, volcanoes, solar variation etc. Figure 1 plots the timeseries of global annual average values of temperature and precipitation for each of the

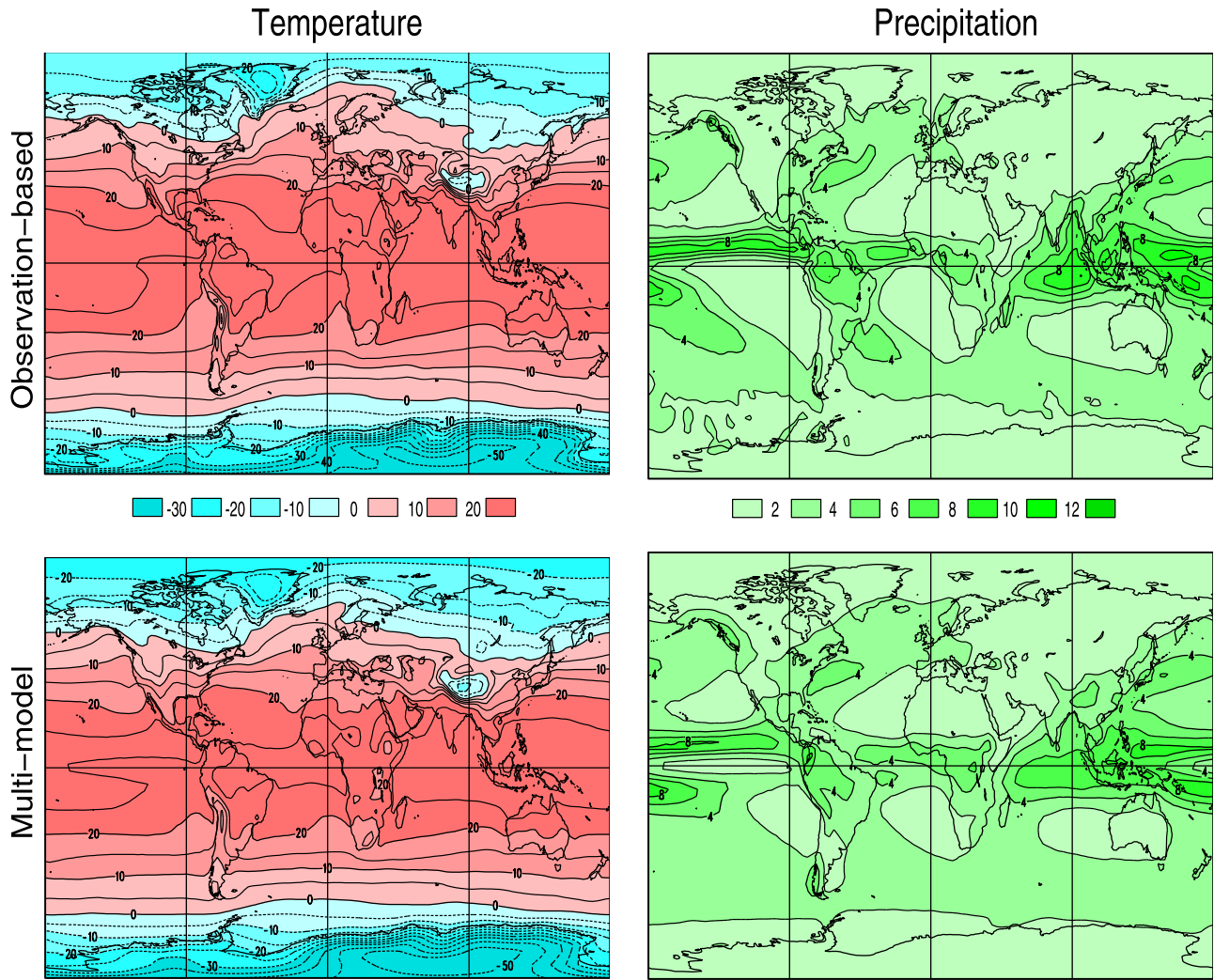


Figure 2. (top) Observation-base climatological annual mean temperature ($^{\circ}\text{C}$) and precipitation (mm/day) from ERA40 reanalysis [Uppala *et al.*, 2005] for temperature and Xie and Arkin [1996] for precipitation. (bottom) The corresponding multi-model ensemble mean climatological annual mean temperature and precipitation distributions.

individual control simulations listed in Table 1. The values are plotted on two panels to avoid too many overlapping curves and are ordered by the length of the simulation with the longer timeseries “underneath” and the shorter curves above and so are not ordered alphabetically. It is apparent from Figure 1 that a number of the models display “climate drift” in their global mean temperatures and, to a lesser extent, visually at least, in their global mean precipitation rates. Drift in global means is a reflection of larger local values and it is important to eliminate drift before analyzing potential predictability since drift will be interpreted as potentially predictable variance. This is discussed below.

[8] Figure 2 plots the geographical distribution of the time and ensemble (indicated by braces) means $\{\bar{X}\}$ and Figure 3 the multi-model pooled standard deviation of annual means $\hat{\sigma}_X = \sqrt{\{(X - \bar{X})^2\}}$ (after drift removal).

Observation-based climatological estimates of these quantities are also plotted, based on the ERA40 reanalysis [Uppala *et al.*, 2005] for temperature and the 17-year Xie-Arkin data set [Xie and Arkin, 1996] for precipitation.

The intent of the comparisons of Figures 2 and 3 is to show that the mean and variance structures of the variables are reasonably well simulated as a necessary condition for further analysis. The investigation of the details of the differences is not the intent of the paper and has its own difficulties since the observationally-based estimates of precipitation and its variability, for instance, will have appreciable uncertainty of their own.

[9] The pooled multi-model quantities are generally in good agreement with the observation-based estimates although there are some deficiencies such as an indication of a the “cold tongue” in the tropical eastern Pacific temperature and the associated “split ITCZ” in model results in Figure 2. The averaged ensemble mean temperature is somewhat colder than the observation-based estimates in Figure 2 but this is partially due to the cooler climate of 1850 to which the preindustrial control refers. The ensemble standard deviations of temperature and precipitation in Figure 3 are in very reasonable accord with the observation-based values. The main area of disagreement is over the very high southern latitudes which cover a comparatively small

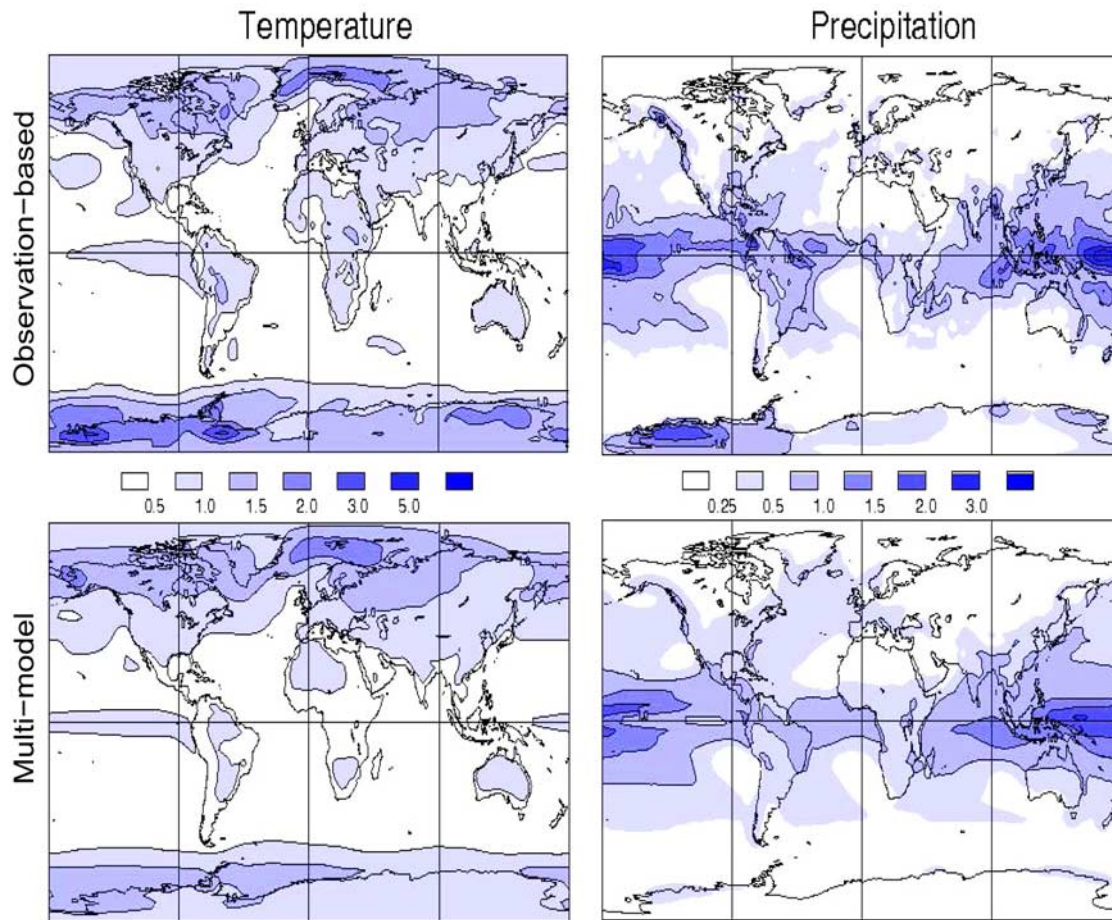


Figure 3. As for Figure 2 but for the standard deviations of annual mean temperature and precipitation.

area and which are not well observed in any case. Overall, the reasonable correspondence of the multi-model statistics with observationally-based values supports further analysis.

3. Climate Drift

[10] As in the work by *Boer* [2004], drift is identified with long timescale trends which are removed by fitting orthogonal polynomials of increasing order (linear, quadratic, etc.) to the data at each point. Linear trends account for an appreciable percentage of the overall drift variance for temperature, and quadratic and higher order curves account for comparatively little. Drift in precipitation is considerably weaker than that of temperature as a proportion of the overall variance. The stepwise fitting of orthogonal polynomials in time and the removal of the associated “drift” rapidly converges in the sense that higher order polynomials account for little additional variance. This motivates the removal of third order polynomial trend in temperature and second order polynomial trend in precipitation at all points even if individual model results do not evidence significant drift. Removal of higher order polynomials has little effect in that case and, even if some small amount of true variability rather than “drift” were to be removed, the result will be conservative in the sense of under- rather than over-estimating potential predictability. Results from two models with data difficulties of one kind or another are omitted from the potential predictability calculations which are

based on the results of 23 control simulations of varying lengths from 21 models comprising 8850 years of data in all.

4. Multi-model Potential Predictability of Precipitation and Temperature

[11] As described by *Boer* [2004] the statistical model used is much like that of *Rowell* [1998] and *Zwiers* [1996] termed the RZ approach. The multi-model approach assumes that model results may be considered as independent samples from the population of models produced with current knowledge and that statistics may be pooled across the model ensemble. The resulting statistics are based on a large amount of data so confidence bands are very narrow and statistical significance is not an issue.

[12] The basic results of the analysis are given in Figure 4 which displays the estimated potential predictability variance fractions $p = \sigma_v^2 / \sigma^2$ for precipitation and temperature. Values are positive and plotted on a more or less logarithmic scale in order to represent also the weak values of potential predictability which penetrate into the land areas. The *ppv* values for 5-, 10- and 25-year means are plotted. Values less than 2% are not coloured.

[13] The results for temperature represent an extension and update of the results given by *Boer* [2004] but based on current models and considerably more data. In the current estimate, land areas display comparatively little potential

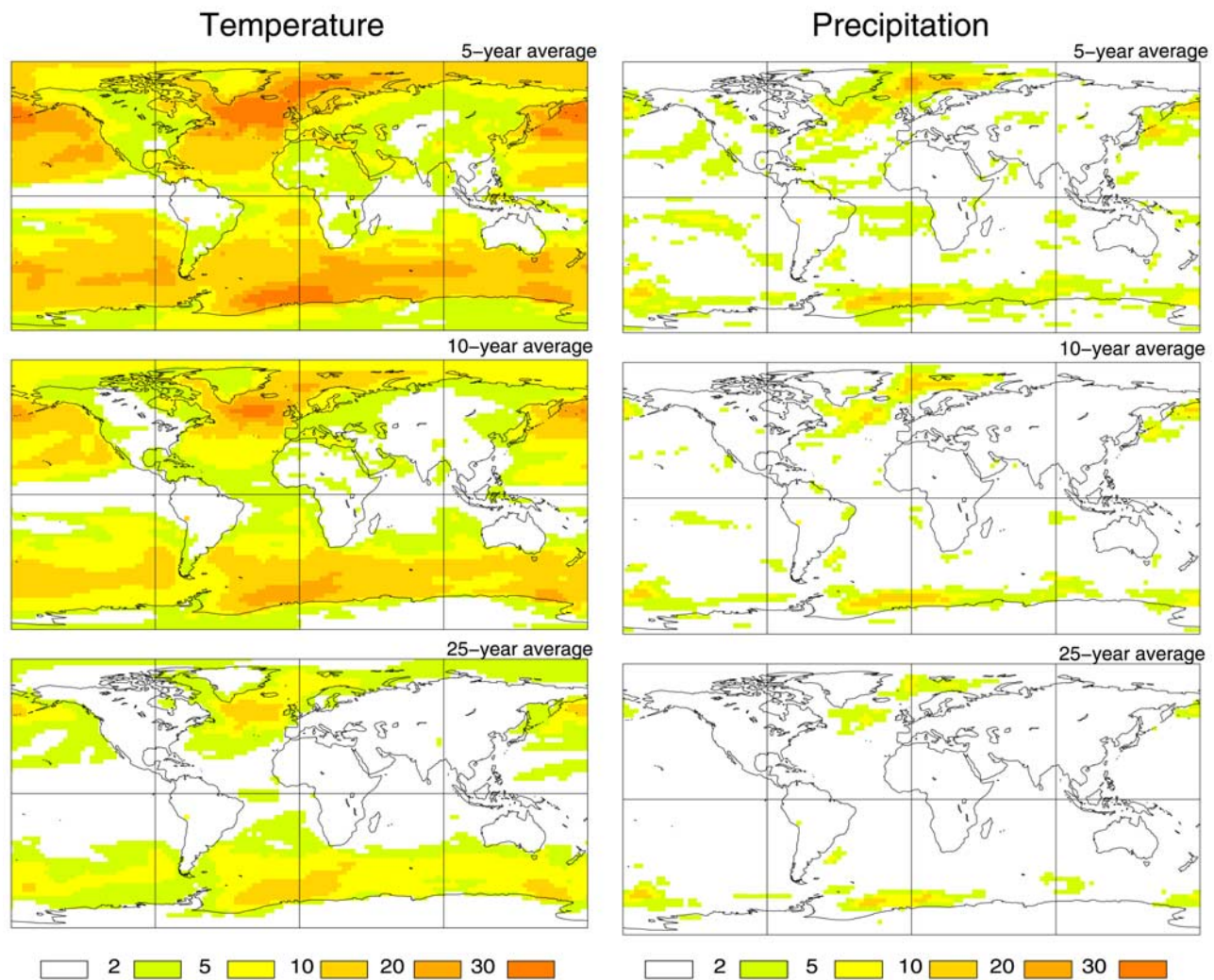


Figure 4. The estimated potential predictability variance fraction $p = \sigma_v^2 / \sigma^2$ expressed as a percentage, for 5-, 10- and 25 year averages for temperature and precipitation.

predictability compared to oceans and equatorial regions display comparatively little potential predictability compared to higher latitudes. Temperature results are broadly similar to the earlier results although there are some differences in detail. Land areas display very modest temperature *ppvf* values although, for shorter timescales, some values bordering the oceans exhibit values of between 5 and 10%. There are intriguing hints of long timescale predictability in the Himalayan region which suggests a potential cryospheric mechanism (*Barnett et al.* [1989] and subsequent studies). Otherwise the comparative strength and short timescales of temperature variability over land mitigate against long timescale potential predictability.

[14] The results for precipitation in Figure 4 (right) represent the first estimate of the long timescale potential predictability of this important climatological parameter. The *ppvf* is, however, considerably smaller than that for temperature and is modest at best. The areas where the *ppvf* is large enough to be coloured reside mainly over the oceans and the precipitation *ppvf* pattern appears as an attenuated version of the temperature pattern. This is in accord with the strong contemporary correlation of annual means of temperature and precipitation over the oceans (not shown) and

suggests that on these timescales the temperature over the oceans drives the associated precipitation.

[15] The smallness of the potential predictability in the tropical region is perhaps counterintuitive in that ENSO and other tropically-based mechanisms are usually thought of as providing predictive skill on seasonal and interannual timescales. We are dealing with even longer timescales than this, however, and tropical processes have comparatively short timescales by this measure. Similar results are found by *Fraedrich and Blender* [2003], for instance, via a different analysis. According to the multi-model estimates, middle to high latitude ocean areas in both hemispheres are the main sites of potential predictability. The largest values of long timescale potential predictability are associated with regions where the surface layer makes contact with the deeper ocean beneath.

5. Summary

[16] A first long-timescale multi-model estimate of the potential predictability of precipitation is obtained by pooling 8850 years of simulated data from the control runs of 21 coupled climate models. The associated analysis of the

potential predictability of temperature updates earlier results to provide a consistent multi-model estimate of these two basic climate parameters. The potential predictability variance fraction (*ppvf*) $p = \sigma_p^2 / \sigma^2$ estimates the fraction of the long timescale variance that is not a consequence of averaging short timescale unpredictable behaviour and hence may be ascribed to long timescale physical processes which are presumed to be predictable with enough information. The *ppvf* in this study deals with the unforced climate behaviour and hence the potential predictability of internally generated long timescale variability. To the extent that the multi-model estimates reflect the behaviour of the coupled climate system, the spatial distribution of the *ppvf* indicates where we may hope to find predictive skill at long timescales and also provides indirect information on the processes involved.

[17] For temperature, the largest values of the *ppvf* are found over the oceans at middle to high latitudes and are mainly associated with regions in which the surface is connected to the deeper ocean. This is presumably the case in the North Atlantic region where the connection to the meridional overturning circulation has been the topic of considerable study as summarized, for instance, by Latif *et al.* [2006, and references therein]. In the North Pacific the situation is rather less direct as discussed also by Latif [2006] where long timescale variability and predictability processes are reviewed. The connection between the surface and the deep Southern Ocean is a feature of climate change simulations where warming is thereby retarded [IPCC, 2007, chap. 10] and also of other predictability studies [e.g., Collins, 2002]. Potential predictability over land is comparatively weak although there are some regions of modest values over northern and western regions of northern hemisphere land masses. The weakness of the land *ppvf* arises not so much because the long timescale potentially predictable variability is small but rather because the short timescale variability is large over land resulting in a small *ppvf*.

[18] These first estimates of the *ppvf* for precipitation indicate that it is a comparatively pale reflection of the *ppvf* for temperature and correlation analysis (not shown) indicates that on long timescales temperature variations drive precipitation variations. Potential predictability for precipitation is found largely in the North Atlantic region and the Southern Ocean with little if any suggestion of potential predictability over land.

[19] Potential predictability given by the *ppvf* is a particular measure of long timescale variability. As noted above, small values do not necessarily mean that predictability is absent but rather that the long timescale potentially predictable variance is, on average, small compared to the total variability. While it is reasonable to measure potential predictability in this way, it might be possible to identify

and exploit specific situations of enhanced predictability for which the *ppvf* is larger than the average value. If so, skill might be obtained in regions with small average *ppvf* if prediction is restricted to these situations. It remains to investigate this possibility further, over land for instance, and also to more clearly identify the physical mechanisms which give rise to the average *ppvf* shown here.

[20] **Acknowledgments.** We are grateful to Oleg Saenko and Slava Kharin for their helpful comments.

References

- Barnett, T. P., L. Dümenil, U. Schlese, E. Roeckner, and M. Latif (1989), The effect of Eurasian snow cover on regional and global climate variations, *J. Atmos. Sci.*, **46**, 201–661.
- Boer, G. J. (2000), A study of atmosphere-ocean predictability on long time scales, *Clim. Dyn.*, **16**, 469–477.
- Boer, G. J. (2004), Long time-scale potential predictability in an ensemble of coupled climate models, *Clim. Dyn.*, **23**, 29–44.
- Boer, G. J., and S. J. Lambert (2008), The energy cycle in atmospheric models, *Clim. Dyn.*, **30**, 371–390.
- Collins, M. (2002), Climate predictability on interannual to decadal time scales: The initial value problem, *Clim. Dyn.*, **19**, 671–692.
- Collins, M., and B. Sinha (2003), Predictability of decadal variations in the thermohaline circulation and climate, *Geophys. Res. Lett.*, **30**(6), 1306, doi:10.1029/2002GL016504.
- Fraedrich, K., and R. Blender (2003), Scaling of atmosphere and ocean temperature correlations in observations and climate models, *Phys. Rev. Lett.*, **90**, 108501.
- Griffies, S. M., and K. Bryan (1997), A predictability study of simulated North Atlantic multidecadal variability, *Clim. Dyn.*, **13**, 459–487.
- Intergovernmental Panel on Climate Change (2007), *Climate change 2007: The Physical Science Basis—Contribution of the Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, 996 pp., Cambridge Univ. Press, Cambridge, U. K.
- Lambert, S. J., and G. J. Boer (2001), CMIP1 evaluation and intercomparison of coupled climate models, *Clim. Dyn.*, **17**, 83–106.
- Latif, M. (2006), On North Pacific multidecadal climate variability, *J. Clim.*, **19**(12), 2906–2915.
- Latif, M., M. Collins, H. Pohlmann, and N. Keenlyside (2006), A review of predictability studies of Atlantic sector climate on decadal time scales, *J. Clim.*, **19**, 5971–5987.
- Madden, R. A. (1976), Estimates of natural variability of time-averaged sea level pressure, *Mon. Weather Rev.*, **104**, 942–952.
- Meehl, G., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. Mitchell, R. Stouffer, and K. Taylor (2007), The WCRP CMIP3 multimodel dataset: A new era in climate change research, *Bull. Am. Meteorol. Soc.*, **88**, 1383–1394.
- Rowell, D. P. (1998), Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations, *J. Clim.*, **11**, 109–120.
- Rowell, D. P., and F. Zwiers (1999), The global distribution of sources of atmospheric decadal variability and mechanisms over the tropical Pacific and southern North America, *Clim. Dyn.*, **15**, 751–772.
- Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, **131**, 2961–3012.
- Xie, P., and A. Arkin (1996), Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions, *J. Clim.*, **9**, 840–858.
- Zwiers, F. (1996), Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2, *Clim. Dyn.*, **12**, 825–847.

G. J. Boer and S. J. Lambert, Canadian Centre for Climate Modelling and Analysis, Environment Canada, University of Victoria, P.O. Box 1700, Victoria, BC, Canada V8W 2Y2. (george.boer@ec.gc.ca)