Correlation between Inter-Model Similarities in Spatial Pattern for Present and Projected Future Mean Climate

Manabu Abe[1], Hideo Shiogama[1], Julia C. Hargreaves[2], James D. Annan[2], Toru Nozawa[1], and Seita Emori[1, 3]

[1]*National Institute for Environmental Studies*, *Tsukuba*, *Japan*

[2]*Research Institute for Global Change*, *JAMSTEC*, *Yokohama*, *Japan*

[3]*Center for Climate System Research*, *The University of Tokyo*, *Kashiwa*, *Japan*

Abstract

When an averaging method is used for model future projection with weights determined according to the model performance in the present climate, generally, the stationarity of the model performance between present and future is implicitly assumed. Here we investigate this assumption using multi-model data. We consider the correlation between inter-model similarities in the spatial pattern for the present-day climate and future climate change for surface air temperature, precipitation and sea level pressure on global and zonal domains in the seasonal time scale. We further extend previous work by devising a bootstrap method to estimate the statistical significance of all correlations, which have previously not been estimated. Most of the correlation coefficients for precipitation were significant, but moderate or low in the absolute value. Many of those for the other variables were not significant. Also, we discuss the magnitude of the inter-model similarity used in this work.

1. Introduction

There have been a number of climate change impact studies where multi-model climate change projections have been weighted according to how well the models simulate the present-day climate (e.g., Giorgi and Mearns 2002, 2003; Murphy et al. 2004; Tebaldi et al. 2004, 2005; Annan et al. 2005; Piani et al. 2005; Nohara et al. 2006). Using a weighted average could provide a better projection than a uniform average of multi-model projections (Min and Hense 2006). However, actually there is no widely accepted measures for assessing climate model performances as whole, and also it is unclear how performance against the observations translates into future simulations (Räisänen et al. 2005; Tebaldi and Knutti 2007). The geographical pattern of climate variables should be a key factor for quantifying performance of three dimensional climate models, in addition to the magnitude of global mean in any variable such as global mean in surface air temperature. A pattern similarity statistic is an objective assessment of the model's geographical performance against observations, and may be used to define a weight for a multi-model average (e.g., Nohara et al. 2006). The main assumption governing the use of weights is that of stationarity of the relationship between observed and simulated climate, estimated in the 20th century and applied to future simulations. However, a direct assessment of future model performance is, of course, not possible as there are no observations of the future climate, and there is no consensus as yet as to the best way to evaluate future climate change projections. There is a need for an objective method for demonstrating the relevance of a given test of current climate of a model to the reliability of the model's enhanced greenhouse response.

Recently, to demonstrate the relevance of particular tests of the present-day model performance to the reliability of the enhanced greenhouse gas response at the regional scale, Whetton et al. (2007) (hereafter W07) tested whether the similarity in the present-day model climate could account for the similarity in the model's enhanced greenhouse gas response in the multi-model space. W07 found correlations of moderate magnitude to be a common occurrence, suggesting that assessment of the present regional climate simulation of models may be valuable. Their approach is based on recognizing the correspondence between the two concepts; that if models which are more similar to one another for the present climate also have more similar responses to greenhouse gas forcing, then models which are more similar to the real climate can be expected to have more similar responses to the real world under greenhouse gas forcing. Because the first concept is directly testable using a multimodel-space, although the second is not, they investigated the correlation as the first step of the approach. Then, they expect tests with the most discriminating power in the multi-model space would be selected for testing models against observations.

In this paper, to better understand the connection between present-day model performance and future climate change projection in the multi-model space, we use a similar approach to W07 to investigate further the similarities in the spatial patterns of different model runs of both present and future climates, where the future climate change is represented by a response to increase greenhouse gases. In W07, future climate changes were normalized by global mean change in surface air temperature, so the correlations between the similarity of different models in both patterns of present regional climate and patterns of regional enhanced greenhouse response did not consider differences in magnitude of global average warming response among models. In this study, the future climate change is not normalized by global mean change in surface air temperature, in order to consider the spatial response pattern including the differences in magnitude of global average warming response. In addition, our target is larger-scale mean climate, rather than regional-scale. We also investigate the degree to which the similarities between the models agree for the two modelled climate states, by correlating the present similarities with the future similarities. After obtaining the correlations between the present and future similarities, we test their significance with a bootstrap method, extending the work of W07 which did not attempt this. Also, we show scatter plots of two example cases to compare the future similarities with the present similarities, discussing the implications for weighting methods where it was assumed that future climate change could be tightly constrained by the present climate model performance.

Corresponding author: Manabu Abe, Climate Risk Assessment Section, Center for Global Environmental Research, National Institute for Environmental Studies, 16-2, Onogawa, Tsukuba, Ibaraki 305-8506, Japan. E-mail: abe.manabu@nies.go.jp.

2. Data

In this study, 21 GCM model simulations forced with SRES A1B emission scenario (Meehl et al. 2007b) were chosen as the future climate projection, and the '20th Century Climate in Coupled Model' (20C3M) simulations were used as the present climate (Meehl et al. 2007a). The selected models appropriate to this study are BCCR-BCM2.0, CGCM3.1(T47), CGCM3.1(T63), CNRM-CM3, CSIRO-Mk3.5, GFDL-CM2.0, GFDL-CM2.1, GISS-AOM, GISS-EH, GISS-ER, UKMO-HadCM3, UKMO-HadGEM1, INM-CM3.0, IPSL-CM4.0, MIROC3.2(hires), MIROC3.2 (medres), ECHOG-G, ECHAM5/MPI-OM, MRI-CGCM 2.3.2, CCSM3, and PCM (web site: WCRP CMIP3 multi-model database; see http://www.pcmdi.llnl.gov/ipcc/about_ipcc.php). All the model data is linearly interpolated to a spatial resolution of 2.5° × 2.5°. The variables used in this study are precipitation (PRCP), surface air temperature (T2), and sea level pressure (SLP). We define the present mean climate for each variable as its average for 1981–2000, and the future climate change is defined to be the difference between the average for 2081–2100 and 1981–2000. For some models, the final year in these intervals was not available, in which case intervals of 19 years were used.

We used reanalysis T2 and SLP from ERA-40 (Uppala et al. 2005) and monthly mean precipitation data produced by GPCP (Adler et al. 2003) for the same period (1981–2000) to represent present-day observed climate with global coverage.

3. Method

A variety of distance functions, or "metrics", have been proposed for evaluating the performance of climate models in comparison to recent observational data (e.g., Watterson 1996; Taylor 2001; Gleckler et al. 2008). Here we use these same metrics to measure the similarity of pairs of climate models. The M statistic of Watterson (1996), which combines the pattern correlation coefficient and root mean square error, was used and abbreviated as M. We also calculated the centred pattern correlation coefficient, which is called R hereafter. Thirdly, with a similar way to calculation of the Climate Prediction Index (CPI) in Murphy et al. (2004), we normalized the root mean square error by mean of the spatial average of the modelled interannual standard deviation in a model pair to uniform units of the statistic and we call this statistic RMS. From 21 climate models it is possible to pick 210 different pairs. For each pair, we calculated the similarity between the two models for annual and seasonal (MAM, JJA, SON, and DJF) mean in present climate and future climate change, to investigate how the metrics vary between seasons. The metrics for global (GLB) and four zonal domains (northern high latitudes, NH: 75 – 45°N, northern mid-latitudes, NM: 45–15°N, equatorial region EQ: 15°N–15°S, and southern mid-latitudes, SM: 15–45 °S), were calculated to examine the variation of the metrics between different zonal bands.

After calculating the metrics, we calculated the correlation coefficients between the present and future metrics, calling this correlation coefficient the "present-future correlation coefficient" for convenience. Correlations for combinations between the same variables, domains and seasons were assessed in this study, because we wanted initially to evaluate correlations between the same variables, which we expected to be generally stronger than more distant relationships.

We applied a simple bootstrap method to test the statistical significance of the present-future correlation coefficient, since the effective number of degrees of freedom is difficult to determine because there will be dependence among the 210 model pairs (W07; Jun et al. 2008). There are 21 model samples for both the present climate and the future changes. In the set of 21 original (true) models, the future change is related to the present climate since both the future change and the present climate were simulated by the same model. From these samples, we generated the set of 21 pseudo-models where the future change is unrelated to the present climate by splicing the 21st century projections to randomly reordered 20th century simulations. Based on each set of 21 pseudo-models, we calculated the inter-model metrics for present climate and future projection in the same way as for the original models, and then obtained the present-future correlation coefficient for the pseudo-models. The bootstrap replication number is 10,000, for each of which we generated a different set of pseudo-models. From this set of replicates, the mean present-future correlation coefficient was almost zero. When the present-future correlation coefficient for the real models is taken to differ significantly from zero at the 95% level according to this null distribution, we describe the present-future correlation coefficient as significant. In addition to the statistical significance, we also consider the magnitude of the coefficients, which we describe as low, moderate, and high according to the ranges 0–0.3, 0.3–0.7 and 0.7–1 respectively.

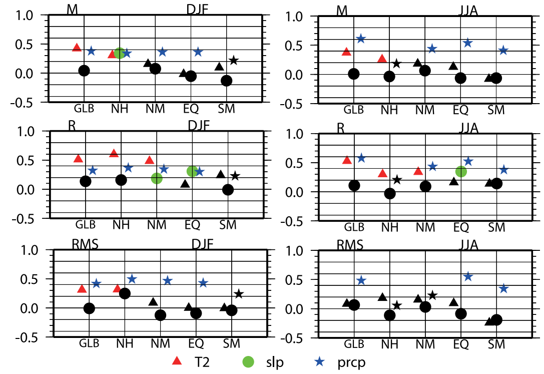

Fig. 1. The present-future correlation coefficients of the inter-model metrics (inter-model similarities) in future climate change projection with those in present climate on defined domains for December–February (DJF) and June–August (JJA). Top, middle and bottom figures show that in cases with a) M-statistic, b) centred pattern correlation (R) and c) normalized RMS errors, respectively. Triangle, circle, and star are for temperature (T2), mean sea level pressure (SLP), and precipitation (PRCP), respectively. The red, green, and blue marks indicate significant coefficients, but black marks not-significant coefficients.

4. Results and discussions

Figure 1 shows the present-future correlation coefficients for all the domains in DJF and JJA. A black mark means the coefficient is not statistically significant at the 95% level. As the present-future correlation coefficients vary according to domain, season, variables, and the metrics used (those in other seasons are shown in Supplement 1), it would be difficult to recommend a single combination of these as a metric for comprehensive assessment of the present-future relation. Thus, we have to understand what it is that the statistic evaluates in the model and apply the statistics appropriate for each particular purpose.

For PRCP, the present-future correlations in most domains are significant, but either moderate in magnitude, or low. Over all three metrics, the present-future correlation coefficients in GLB and EQ in DJF and JJA are moderate. The present-future correlation in JJA PRCP for the NH is quite different from that in DJF. For tropical precipitation this study obtained a moderately strong correlation in contrast to the low correlation of

W07. To check further an effect of using precipitation change in mm day^{-1} on this contrast, the present-future correlation coefficient with M was calculated with percent change of precipitation in future. The correlation coefficients on tropical precipitation in JJA and SON are significantly moderate. Therefore, using of precipitation change in mm day^{-1} does not strongly affect the contrast between results in this study and W07. Further, in this study the analysis is made over larger spatial scales than in W07, meaning that the metrics should be dominated by similarity of the large-scale features such as the Intertropical Convergence Zone (ITCZ). Our results therefore suggest that the zonal scale analysis of modeled present-day precipitation may be more useful for assessing the future performance of models than the analysis on the smaller regional scales.

For T2, present-future correlations for GLB and NH in DJF are significant and moderate. In JJA, for M and R, the GLB and NH also have significant moderate coefficients. In NH, the future change similarity in the spatial pattern of T2 is strongly related to the similarity in the current spatial pattern of T2. This result is consistent with the results in the northern high-latitude region in W07. So, as surface albedo feedback is particularly critical for climate change prediction in the northern hemisphere (e.g., Winton 2006), the predicted future climate in NH may be sensitive to the present-day spatial pattern of surface temperature depending on land process including snow amount and its seasonality and expression of sea ice. For SLP most of the present-future correlation coefficients are not statistically significant for the three metrics, although there are a few cases where there are significant correlations of low or moderate strength.

Since we did not normalize the model results by the global climate change, we expect the effect of the global climate change to enter into the RMS statistic, but not into R: The centred pattern correlation, R, is a measure of the similarity of the spatial pattern of variable fields rather than of the domain-mean, but the root mean square error, RMS, is very sensitive to differences in the domain-mean. To analyse this further, we calculated the correlation of the present and future metrics with the future domain-mean difference. As expected, for RMS there are strong correlations between the future change metrics and the future domain-mean differences, particularly for T2 and SLP, but much lower correlations for the R statistic. However, neither the present-future correlation coefficients nor the present-day metrics show much correlation with the future domain-mean difference (see Supplement 2). Therefore, the future climate sensitivity difference cannot be estimated from the present-day metrics for T2 and SLP. On the other hand, for precipitation, such a relation was not clear. Further research is needed.

We give here examples showing scatter plots of the present and future metrics for two cases where the present-future correlations were statistically significant for all metrics, according to our bootstrap method. Figures 2a–c and 2d–f show the metrics for JJA PRCP in EQ and DJF T2 in NH, respectively. The blue dots in the figures indicate present-day model performances of each model, which were defined as the similarity to the observations (GPCP in Figs. 2a–c and ERA40 in Figs. 2d–f). The present-future correlation coefficients in both R and RMS for JJA PRCP in EQ are similar at 0.52 and 0.50, respectively (Figs. 2b–c). The regression coefficients are, however, rather different at 0.69 and 0.18, respectively. In R for DJF T2 in NH, the present-future correlation coefficient is 0.60, and the regression coefficient is 3.73, much larger than in M or RMS, where the regression coefficients are below 1.0. The difference between the correlation coefficient and the regression coefficient depends on the relative variance of the metrics for the present and future. The cases with larger regression coefficients indicate that future model performance is

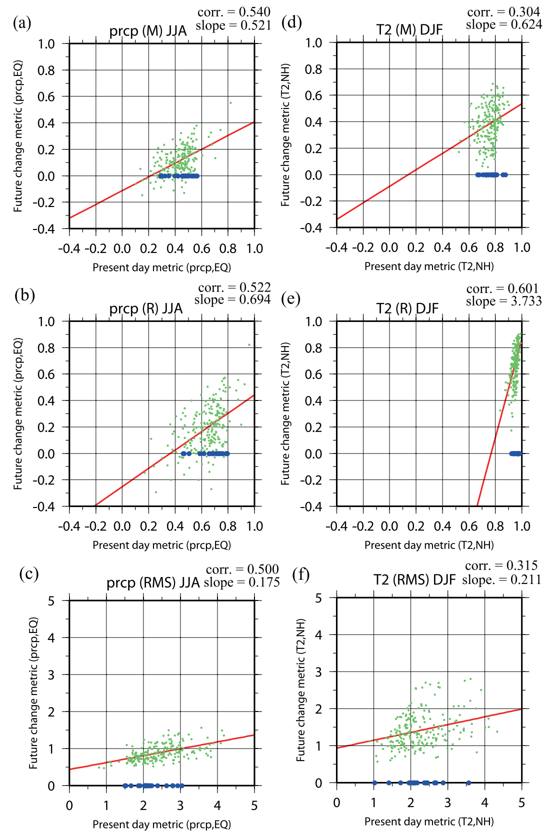

Fig. 2. Scatter plots for the present and future change metrics on (a–c) JJA PRCP in EQ and (d–f) DJF T2 in NH. Green dots indicate the metrics of every 210 model pairs in both present and future. Blue dots indicate present climate model performance with observation data (GPCP or ERA40-T2), for each model. Red line is the regression line.

strongly dependent on the present-day model. A small regression coefficient indicates that improvements in the current model performance alone are unlikely to result in substantial changes in the reliability of projection. Therefore, not only correlation but also regression of the future change metrics with the present metrics may be useful indicators of model performance.

The current model performances compared to the observations in every case are within the same range as the present inter-model metrics. Therefore, the present-future relations among models may also provide useful information for evaluating the actual reliability of future climate change prediction.

Models are routinely verified using the present day climate, so it is expected that, as we have found, the present-day metrics in R and M are all higher than the future metrics, probably due to tuning with the observations for better simulation of the present-day climate (Fig. 2 and Supplement 3). However, direct comparison in RMS metric between the present and the future is difficult due to the normalization method applied in this study. In addition, the magnitude of the future metrics can provide a guide as to the likely reliability of model predictions. For example, even though there is a clear positive correlation in Fig. 2b, the future change metrics in R are all rather low (around 0.2). If all the models substantially disagree with each other and we have no method for choosing the best one, we cannot possibly believe now that a specific one of them is good. Therefore uncertainty concerning this aspect of future climate change is high, and it is unlikely that any of these models can be considered reliable as a predictor of precipitation changes in this region. In contrast, Fig. 2e indicates a much higher reliability for temperature change patterns, with a future R-statistic of at least 0.6 even for the worst models, due to greater similarity among models in temperature response pattern than other variables. In this case predictions of temperature change patterns may already be considered reasonably robust, even though the present-future correlation suggests that further model improvement can be expected to lead to even better results.

5. Summary

To assess the assumption of the stationarity of the model performance between present and future and also to reveal or better understand the statistical relationships between climate change projection and climate model performance, we investigated correlations between inter-model similarities in future climate change projection and the present-day mean climate, based on a similar approach to that in W07. Since the effective number of degrees of freedom is not clear in this study, as also mentioned by W07, we tested the significance of the present-future correlation coefficient using a bootstrap method. Most of the present-future correlation coefficients for precipitation were significant, but moderate or low in absolute value. In the tropics, particularly, the present-day metrics for precipitation may be indicative of the reliability of future change prediction of precipitation. In contrast, W07 did not achieve a strong result for precipitation in the tropics. We suggest that the difference in the results may be due to this analysis considering larger spatial scales. The present-future correlations for surface temperature are mostly moderate and significant in northern high-latitudes, lower and less reliably significant in mid-northern latitudes and even lower and not significant in other parts of the globe, while the sea level pressure results were mostly low and not significant everywhere.

There are more cases with significant correlation coefficient in M and R rather than in RMS. On the other hand, in R, although a lot of cases with significant correlation were seen, there is a concern that R can measure only similarity in the spatial amplification pattern. Thus, a metric using the M statistic could be suitable for further investigation about a relation between present model performance and future model performance with the inter-model similarities. In addition to the present-future correlation coefficient, the regression coefficient also may be a useful indicator of the present-future relationship. However, for some cases, despite significant correlations of good strength, the future metrics were very low across all model pairs, meaning that little information about the future climate could be obtained from the metric. We are, therefore, doubtful whether future projections can be tightly constrained using only the simple model performance presented here for the present-day climate.

Acknowledgements

We acknowledge the international modeling group for contributing their data to the CMIP 3 database, and the PCMDI for collecting and archiving the model data. This work was supported by the Global Environment Research Fund (S-5) of the Ministry of the Environment, Japan.

Supplements

Supplement 1 shows the present-future correlation coefficients of the inter-model metrics in future climate change projection with those in present climate for annual mean (ANN) and two seasons (March – May (MAM) and September–November (SON)).

Results of supplementary analyses are described in Supplements 2, and 3.

References

Annan, J. D., J. C. Hargreaves, R. Ohgaito, A. Abe-Ouchi, and S. Emori, 2005: Efficiently constraining climate sensitivity with paleoclimate simulations, *SOLA*, **1**, 181–184.

Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P.-P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-present), *J. Hydrometeor.*, **4**, 1147–1167.

Giorgi, F., and L. O. Mearns, 2002: Calculation of average, uncertainty range, and reliability of regional climate chanegs from AOGCM sumulations vai the "relaibility ensemble averageing" (REA) method, *J. Climate*, **15**, 1141–1158.

Giorgi, F., and L. O. Mearns, 2003: Probability of regional climate change based on the Reliability Ensemble Averageing (REA) method, *Geophys, Res. Lett.*, **30**, 1629, doi:10.1029/2003GL017130.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models, *J. Geophys. Res.*, **113**, D06104, doi:10.1029/2007JD008972.

Jun, M., R. Knutti, and D. W. Nychka, 2008: Spatial analysis to quantify numerical model bias and dependence: How many climate model are there?, *J. Amer. Statis. Assoc.*, **103**, 934–947.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007a: The WCRP CMIP3 multimodel dataset-A new era in climate change research, *Bull. Amer. Meteor. Soc.*, **88**, 1838–1394.

Meehl, G. A., T. F. Stocker, W. D. Collins, P. Friedlingstein, A. T. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, S. C. B. Raper, I. G. Watterson, A. J. Weaver, and Z.-C. Zhao, 2007b: Global Climate Projections. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Min, S.-K., and A. Hense, 2006: A Bayesian approach to climate model evaluation and multi-model averaging with an application to global mean surface temperature from IPCC AR4 coupled climate models, *Geophys. Res. Lett.*, **33**, L08708, doi:101029/2006GL025779.

Murphy, J. M., D. M. H. Sexton, D. N. Barnet, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, **430**, 768–772.

Nohara, D., A. Kitoh, M. Hosaka, and T. Oki, 2006: Impact of climate change on river discharge projected by multi-model ensemble, *J. Hydrometeor.*, **7**, 1076–1089.

Räisänen, J., 2007: How reliable are climate models?, *Tellus Ser. A*, **59**, 2–29.

Tebaldi, C., and R. Knutti, 2007: The use of the multi-model ensemble in probabilistic climate projection, *Philos. Trans. R. Soc. Sea. A*, **365**, 2053–2075.

Tebaldi, C., L. O. Mearns, D. Nychka, and R. L. Smith, 2004: Regional probabilities of precipitation change: A Baysian anaysis of multimodel simulations, *Geophys. Res. Lett.*, **31**, L24213, doi:10.1029/2004GL021276.

Tebaldi, C., R. W. Smith, D. Nychka, and L. O. Mearns, 2005: Quantifying uncertainty in projections of regional climate change: A Baysian approach to the analysis of multi-model ensembles, *J. Climate*, **18**, 1524–1540.

Uppala, S. M., P. W. Kållberg, A. J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, and co-authors, 2005: The ERA-40 re-analysis, *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.

Watterson, I. G., 1996: Non-dimensional measures of climate model performance, *Int. J. Climatol.*, **16**, 379–391.

Whetton, P., I. Macadam, J. Bathols, and J. O'Grady, 2007: Assessment of the use of current climate patterns to evalate regional enhanced greenhouse response patterns of climate models, *Geophys. Res. Lett.*, **34**, L14701, doi: 10.1029/2007GL030025.

Winton, M., 2006: Surface albedo feedback estimates for the AR4 climate models, *J. Climate*, **19**, 359–365.

Manuscript received 11 July 2009, accepted 29 August 2009
SOLA: http://www.jstage.jst.go.jp/browse/sola/