Estimating regional climate change uncertainty in Japan at the end of the 21st century with mixture distribution

Shunya Wakamatsu1, Kenji Oshio2, Koji Ishihara2, Hirokazu Mura2, Takuya Nakashima2 and Tsuyoshi Inoue3

1Ministry of Education, Culture, Sports, Science and Technology, Japan
2Japan Meteorological Agency, Japan
3Mitsubishi Research Institute Inc., Japan

Abstract:

To facilitate accurate assessments of the regional impacts of global warming, and make informed decisions about appropriate measures to mitigate them, detailed global warming projections with uncertainties are needed. The Ministry of Environment of Japan and the Japan Meteorological Agency performed 21 different multi-scenario and multi-ensemble experiments in Japan using the regional climate model MRI-NHRCM with a horizontal resolution of 20 km. To estimate the total range of uncertainty due to natural fluctuations and the variety of experimental runs by a single climate model with multi-physics and multi-SST ensembles under each greenhouse gas emission scenario, a unique statistical method that combined a mixture distribution and bootstrap resampling was adopted. Based on three models that adopted the Yoshimura scheme as a cumulus convection parameterization, annual mean temperatures in Japan were projected to rise significantly by 1.1 ± 0.4°C, 2.0 ± 0.4°C, 2.6 ± 0.6°C, and 4.4 ± 0.6°C under the RCP2.6, RCP4.5, RCP6.0, and RCP8.5 scenarios, respectively, at the end of the 21st century relative to the end of the 20th century (ensemble means ± standard deviations). In contrast, changes in future annual precipitation over Japan were projected to be statistically insignificant.

KEYWORDS global warming; regional climate model; uncertainty; mixture distribution; bootstrap resampling

INTRODUCTION

To facilitate accurate assessment of the regional impacts of global warming, as well as informed decisions regarding appropriate measures to mitigate them, detailed global warming projections with uncertainties are needed. To make these impact assessments and to formulate a national plan for adaptation, the Ministry of Environment of Japan and the Japan Meteorological Agency performed 21 kinds of multi-scenario and multi-ensemble simulations of the climate over Japan with the Earth Simulator, which was one of the world’s fastest supercomputers in 2013 and 2014. However, how to treat the uncertainties of these results is unknown because we cannot calculate the standard deviation appropriately from a few perturbed runs.

Global warming projections using climate models generally contain uncertainties associated with (i) future greenhouse gas (GHG) emissions (scenario uncertainty), (ii) climate models (model uncertainty) and (iii) natural fluctuations (internal variability) (Hawkins and Sutton, 2009; Ishihara, 2010).

Scenario uncertainty reflects the variety of emission scenarios; these uncertainties are inevitable in the case of global warming projections of radiative forcing and depend strongly on both future political decisions and economic development. Recently, four Representative Concentration Pathways (RCPs; Moss et al., 2010), which are GHG concentration trajectories (RCP2.6, RCP4.5, RCP6.0, and RCP8.5), were adopted for the Coupled Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al., 2012), which provides multi-model datasets under specific emission scenarios. These results have also been used as basic materials for the fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC, 2013).

Model uncertainty comes from both the imperfections of our understanding about physical processes that are part of our climate system and the limitations of the computational resources available for model calculations (Tebaldi and Knutti, 2007). To estimate this uncertainty, projection results made by multiple climate models are needed under each RCP emission scenario. Because detailed projection results in Japan are provided mainly by the Regional Climate Model (RCM) developed by the Meteorological Research Institute (MRI) of the Japan Meteorological Agency (JMA), it is very difficult to estimate model uncertainty under present circumstances.

Internal variability requires consideration of both the natural variability inherent in the atmosphere and the ocean and the uncertainty in its estimation which comes from a lack of samples due to limitations in computational resources. Generally, climate models cannot reproduce inter-annual natural variability very well. Moreover, detailed projection results with RCMs are often conducted for only a short period of time (e.g., 20 years). Consequently, it is difficult to adequately reproduce a variety of inter-annual samples. To
reduce this uncertainty, increasing data samples using statistical resampling techniques such as the bootstrap resampling method (Efron, 1979) and Monte Carlo simulations have been used in several previous studies (e.g., Shiogama et al., 2007; Ishihara, 2010).

To estimate uncertainty based on multiple experimental results, each uncertainty mentioned above should be considered. Based on a suggestion proposed by Hawkins and Sutton (2009) for partitioning uncertainty, Yip et al. (2011) used analysis of variance (ANOVA), and this approach was later improved by, for example, Northrop and Chandler (2014). These approaches to estimating total uncertainty as a sum of partitioning variance are the key to summarize results that comes from several experimental designs.

In this study, the ANOVA approach with a mixture distribution and traditional bootstrap resampling was adopted to quantify uncertainties in future changes in annual mean temperature and annual precipitation in Japan from a climate model which has limited climate sensitivity. With larger climate model datasets, this approach enables us to estimate global warming projections over Japan with uncertainties and contributes greatly to the development of a national plan for adaptation.

MODELS AND EXPERIMENTS

All detailed simulations of future meteorological conditions in Japan were conducted with the MRI non-hydrostatic regional climate model (NHRCM) MRI-NHRCM20 (e.g., Sasaki et al., 2011, 2012), which was developed by the MRI. This climate model is based on an operational non-hydrostatic mesoscale model and covers all regions of Japan (117.43–160.93°E, 19.48–46.61°N) with a horizontal resolution of 20 km. The horizontal grids are 211 × 175, and the vertical component has 40 layers. This NHRCM uses the Kain-Fritsch (KF) scheme (Kain and Fritsch, 1993) to parameterize cumulus convection. This scheme is known to accurately depict convective activities around Japan. The MRI-NHRCM20 needs not only concentrations of greenhouse gases, ozone, and aerosols, but also atmospheric conditions calculated by a Global Climate Model, which in this case was the Atmospheric General Circulation Model (AGCM) named MRI-AGCM3.2H with a resolution of 60 km (e.g., Mizuta et al., 2012; Endo et al., 2012). The horizontal grids of the AGCM are 640 × 320, and the vertical component has 60 layers.

In the calculations with MRI-AGCM3.2H, the present climate and future climate corresponded to the periods from 1 September 1984 to 31 August 2004 and from 1 September 2080 to 31 August 2100, respectively. The future climate experiments were conducted under four GHG emission scenarios: RCP2.6, RCP4.5, RCP6.0, and RCP8.5. As for sea surface temperature (SST), an observational dataset (HadISST) (Rayner et al., 2003) was used for the present climate, and three kinds of SST patterns categorized by applying cluster analysis to the SST distributions of the CMIP5 models (Mizuta et al., 2014) were used for the future climate to conduct ensemble experiments under each RCP scenario. Due to the normalization used in this categorization process, the global average of the SST changes was almost the same for all three SST patterns under each RCP scenario. Therefore, experimental settings in this study overlap climate model uncertainty by focusing on the representation of SST change but not on climate sensitivity inherent in each CMIP5 model.

Because future climate changes under the RCP8.5 scenario tended to show a significant warming signal, which is important for assessing the risk of natural disasters, we conducted more experiments under the RCP8.5 scenario. The experiments involved changing not only SST boundary conditions but also the cumulus convection parameterization of the AGCMs. The cumulus convection parameterization is a mechanism used in numerical forecasting models to redistribute heat and moisture in the vertical direction, and it affects the behavior of precipitation (e.g., Japan Meteorological Agency, 2015). In these experiments, we adopted the Arakawa-Schubert (AS) scheme (Arakawa and Schubert, 1974; Randall and Pan, 1993), the Yoshimura (YS) scheme (Yoshimura et al., 2015) and the KF scheme. An AGCM with a different cumulus convection parameterization can be regarded as a different AGCM. Therefore, each present climate was reproduced using an NHRCM that was driven by the output from an AGCM for each cumulus convection parameterization.

As a result, in these experiments there were three kinds of present climate with different AGCM cumulus convection parameterizations, and three kinds of future climate with different SST boundary conditions and one AGCM cumulus convection parameterization under each of the RCP2.6, RCP4.5 and RCP6.0 scenarios. Under the RCP8.5, there were nine kinds of future climate with different SST boundary conditions and different AGCM cumulus convection parameterizations (see Table I).

Before estimating the uncertainty in future climate changes, the reproducibility of the three kinds of present experiments was checked, with only the bias of daily precipitation was corrected. For further details, see Text S1 and Figure S1.

METHODOLOGY

As mentioned above, in this study, three kinds of model results (members) \( (N_m = 3) \) were used in the RCP2.6, RCP4.5, and RCP6.0 experiments, and nine members \( (N_m = 9) \) were used in the RCP8.5 experiment. In all experiments, the present and future climate consisted of 20-year time series of data \( (N_t = 20) \). Because the total number of members \( N_m \) differed under each RCP scenario, overall uncertainty was quantified for each RCP scenario in this study.

Estimating uncertainty with a mixture distribution

Based on Hawkins and Sutton (2009), with an assumption of normal distribution, the model output data \( x(m,t) \) for each model member \( (m = 1, 2, \ldots, N_m) \) and year \( (t = 1, 2, \ldots, N_t) \) (e.g., annual average value) could be expressed as follows:

\[
    x(m,t) = \mu + \alpha(m) + \epsilon(m,t)
\]

(1)

Here,

\( \mu = x(\cdot,\cdot) \) is the ensemble mean of all model members and times, \( \alpha(m) = x(m,\cdot) - x(\cdot,\cdot) \) is the deviation of time-averaged model data mean from the overall ensemble mean, \( \epsilon(m,t) = x(m,t) - x(m,\cdot) \) is the error term, which is independent and
identically distributed for model member \( m = 1, 2, \ldots, N_m \) and time \( t = 1, 2, \ldots, N_t \).

The model uncertainty \( M \) is the sample variance of time-averaged model means around the entire ensemble mean:

\[
M = \frac{1}{N_m} \sum_{m=1}^{N_m} \left( \frac{1}{N_t} \sum_{t=1}^{N_t} \left( x(m, t) - \mu \right) \right)^2
\]

where \( \mu \) is the time-averaged mean of each model member.

In Equation (2), the denominator is not \( N_m - 1 \) but \( N_m \). The reason is that multi-model ensembles are “ensembles of opportunity” (Tebaldi and Knutti, 2007), which means that the ensemble is somewhat affected by non-scientific issues such as the interests of research groups or funding. Hence we do not think of available models as samples from a wider population of possible models. Therefore, we did not use an unbiased estimate of the variance to take account of random effects in Equation (2). From this perspective, the model uncertainty \( M \) is not expected to cover all of the model uncertainties. In this study, the uncertainty of climate sensitivity is also not included because of the experimental setting of the normalized SST conditions.

The internal variability \( V \) is the variance of each member around the model mean:

\[
V = \frac{1}{N_m(N_t-1)} \sum_{m=1}^{N_m} \sum_{t=1}^{N_t} \left( x(m, t) - x(\cdot, \cdot) \right)^2
\]

where \( \sigma_{\text{eq}}^2 \) is the variance of each model member output.

From Equations (2) and (3), the total uncertainty \( T \) is defined as:

\[
T = M + V = \frac{1}{N_m} \sum_{m=1}^{N_m} \left[ \left( \mu_m - \mu \right)^2 + \sigma_{\text{eq}}^2 \right] = \sum_{m=1}^{N_m} \sigma_{\text{eq}}^2
\]

where \( \sigma_{\text{eq}}^2 \) is the variance of a mixture distribution (Burnham and Anderson, 2002).

**Bootstrap method**

In addition to the above equation, uncertainty due to small sample size (\( N_t = 20 \)) was also taken into account by a traditional bootstrap method (Efron, 1979). The steps are as follows:

1. Generate \( N_t \) samples \( x(m, t) \) from the original samples
\(x(m, t)\) by allowing duplication.

2. Estimate the mean \(\mu\) and variance \(\sigma_{\text{eq}}^2\) from \(x(m, t)\).

3. Iterate the above steps \(N = 10,000\) times.

To estimate the total uncertainty from the above bootstrap samples, we used the Law of Total Variance (Weiss et al., 2005), which explains that the total variance in a random variable \(Y\) is equal to the sum of the expectation value of the variance and the variance of the expectation value when \(Y\) is conditioned on some random variable \(X\). That is:

\[
\text{Var}[Y] = \text{Var}[E[Y|X]] + E[\text{Var}[Y|X]]
\]

Based on Equation (5), the uncertainty due to bootstrap resampling \(x(m, t)\) is expressed as follows:

\[
\text{Var}[x(m, t)] = \text{Var}[E[x(m, t) | x(m, t)]]
\]

\[+E[\text{Var}[x(m, t) | x(m, t)]]\]

(6)

The terms on the right side can be rewritten in terms of each bootstrap sample’s mean \(\mu\) and variance \(\sigma^2\) as follows:

\[
E[x(m, t) | x(m, t)] = \mu
\]

(7)

\[
E[\text{Var}[x(m, t) | x(m, t)]] = E[(\sigma_{\text{eq}}^2)] = \sigma_{\text{eq}}^2
\]

(8)

Here, we used the approximation that the mean of the bootstrap distribution equals the mean of the original data (Berendsen, 2011).

Using Equations (7) and (8), we rewrote Equation (6) as follows:

\[
\text{Var}[x(m, t)] \approx \text{Var}[\mu] + \sigma_{\text{eq}}^2 = T'
\]

(9)

\[
T' = \frac{1}{N-1} \sum_{i=1}^{N} (\mu - \frac{1}{N} \sum_{i=1}^{N} \mu)^2 + \frac{1}{N} \sum_{i=1}^{N} [\mu_i - \mu]^2 + \sigma_{\text{eq}}^2
\]

(10)

Therefore, with the use of Equation (4), the total uncertainty \(T'\) can be expressed with the bootstrap term \(B\) as:

\[
T' = T + T = B + M + V
\]

(11)

\[
B = \frac{1}{N-1} \sum_{i=1}^{N} (\mu - \frac{1}{N} \sum_{i=1}^{N} \mu)^2
\]

(12)

RESULTS

Uncertainty of annual average temperature change

Figure 1 shows the future changes in annual temperature averaged over Japan (boxes) and their standard deviations (whiskers). For reference purposes, standard deviations of the present climate that correspond to interannual variations are also shown as whiskers in each panel of Figure 1. Figure 1(a) shows the future temperature changes and their standard deviations under each RCP scenario. Each result consists of three kinds of experiments that were projected based on the YS scheme and different SST conditions. The temperature changes were 1.1 ± 0.4°C (mean ± standard deviation, the same applies hereafter), 2.0 ± 0.4°C, 2.6 ± 0.6°C, and 4.4 ± 0.6°C under the RCP2.6, RCP4.5, RCP6.0, and RCP8.5 GHG emission scenarios, respectively. The simple implication of these results is that the more future GHG emissions increase, the more the temperature and its range of uncertainty increase. Moreover, the results also show that future temperature change is significant under all emission scenarios, even though inter-annual variations were taken into account.

A comparison of Figure 1(a) with Figures 1(b) and 1(c) also shows the differences in future changes due to different convection schemes under the RCP8.5 emission scenario. The fact that these mean values are almost the same indicates that there are no differences in temperature responses due to differences in cumulus convection parameterizations if the GHG emission scenario is the same.

Figure 1(d) shows the result of all RCP 8.5 scenario experiments. The future change of annual temperature under the RCP8.5 scenario is projected to be 4.4 ± 0.6°C, which is the same as the result in Figure 1(a) based on only the YS scheme. According to the IPCC (2013), the projected change in global mean surface temperature is 2.6–4.8°C (5–95% model range, ensemble mean: 3.7°C) by the end of the 21st century compared to the end of the 20th century under the RCP8.5 scenario. Our result is placed above the ensemble mean in IPCC (2013) because the projected future temperature change in Eastern Asia including Japan based on CMIP5 global models tends to be larger than the global average value (Christensen et al., 2013).

Estimated uncertainty of annual precipitation change

Figure 2 shows the future changes in annual precipitation averaged over Japan (boxes) and their standard deviations.
UNCERTAINTY OF REGIONAL CLIMATE CHANGE

The details of all figures are the same as Figure 1. According to Figure 2(a), it is clear that annual precipitation is projected to increase slightly under all RCP scenarios, but the changes are not significant because of their large standard deviations. In contrast, Figures 2(a)–(c) show different results, despite being based on the same GHG emission scenario. Among the three kinds of cumulus convection parameterizations, the KF scheme caused the largest change in precipitation, not only for annual precipitation in the present but also for the future change under RCP8.5.

Uncertainty analysis

In this study, total uncertainty was expressed as the sum of internal variability \( V \), model uncertainty \( M \), and bootstrap variance \( B \). Figure 3 shows the fraction of each variance to total uncertainty. It is clear that internal variance is the largest but insufficient to explain all the uncertainty. \( B \) is the term to compensate for the lack of estimation about \( V \) but it accounts for a small part of the whole uncertainty. This means that 20-year samples in each experiment are adequate to a certain extent to capture \( V \). \( M \) does not depend on the GHG emission amount. Therefore, it seems that model uncertainty should always be considered in each global warming projection since it can be so large even when the magnitude of future change is not large (see RCP6.0 case of temperature uncertainty in Figure 3). This result, however, should be taken with caution in terms of making comparisons because it depends on the numbers of samples from the provided ensemble (see Table I).

DISCUSSION AND CONCLUSIONS

In this study, ranges of uncertainty due to internal variability, multiple experiments, and resampling were estimated by adopting a mixture distribution and bootstrap resampling to experimental runs by a single climate model with multi-physics and multi-SST ensembles under each RCP scenario. Based on this analysis, all projected changes in temperature were statistically significant, unlike future changes in annual precipitation.

However, we must recognize that this approach is not perfect, because it depends on the following assumptions.

(i) All samples in all experiments are statistically independent of each other and have the same occurrence probability.

(ii) All data follow a normal distribution.

(iii) Appropriate population means and variances can be estimated by bootstrap resampling.

These assumptions are not always true, especially (i) and (ii).

We regarded the above assumptions as reasonable in this study for the following reasons. First, the results for different years and for different experiments in this study can be considered statistically independent and also to have the same likelihood of occurrence. This conclusion follows from the fact that in general the value for a particular year is independent of the value in the preceding year, and experimental values of SST and cumulus parameterizations were made without allowing duplication of the data source.

With respect to the assumption that the data follow a normal distribution, it is well known that annual values follow a normal distribution in the cases of average temperature and
precipitation (e.g., Japan Society of Hydrology and Water Resources, 1997). In addition to these assumptions related to our method, this study is based on the following concepts.

(i) We consider not only the scenario and model uncertainty but also internal variability as a part of uncertainty because internal variability can also be affected by climate change.

(ii) We suggest a unique statistical method which enables us to optimize the experimental runs which have limited climate sensitivities.

ACKNOWLEDGMENTS

This work used the dataset produced and provided in the FY2013 and FY2014 climate change projection project conducted by the Ministry of the Environment of Japan (MOEJ). This dataset was created by using climate models developed by the Meteorological Research Institute, in cooperation with the Japan Meteorological Agency, and the SOUSEI Program of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

The DIAS dataset is archived and provided under the framework of the Data Integration and Analysis System (DIAS), through the National Key Technology, Marine Earth Observation Exploration System. The official names of the data are “Global Climate Change Projection Data by MOEJ (in cooperation with JMA)” (http://search.diasjp.net/en/dataset/GCM60_ADAPT2013) and “Regional Climate Change Projection Data by MOEJ (in cooperation with JMA)” (http://search.diasjp.net/en/dataset/NHRCM20_ADAPT2013).

SUPPLEMENTS

Text S1. Reproducibility check and bias correction
Figure S1. Scatter plots between observations and the present climate reproduced with NHRCM

REFERENCES


Weiss NA, Holmes PT, Hardy M. 2005. A course in probability, Addison-Wesley, Boston, USA; 816.