Asia-Pac. J. Atmos. Sci., 52(2), 237-249, 2016
DOI:10.1007/s13143-016-0022-z
pISSN 1976-7633 / eISSN 1976-7951

Spatial Analysis of Future East Asian Seasonal Temperature Using Two Regional Climate Model Simulations

Yura Kim[1], Mikyoung Jun[2], Seung-Ki Min[3], Myoung-Seok Suh[4], and Hyun-Suk Kang[5]

[1]*Department of Statistics, University of Michigan, Ann Arbor, USA*
[2]*Department of Statistics, Texas A&M University, College Station, USA*
[3]*School of Environmental Science and Engineering, Pohang University of Science and Technology, Pohang, Korea*
[4]*Department of Atmospheric Sciences, Kongju National University, Kongju, Korea*
[5]*National Institute of Meteorological Sciences, Jeju, Korea*

(Manuscript received 30 October 2015; accepted 25 April 2016)

Abstract: CORDEX-East Asia, a branch of the coordinated regional climate downscaling experiment (CORDEX) initiative, provides high-resolution climate simulations for the domain covering East Asia. This study analyzes temperature data from regional climate models (RCMs) participating in the CORDEX - East Asia region, accounting for the spatial dependence structure of the data. In particular, we assess similarities and dissimilarities of the outputs from two RCMs, HadGEM3-RA and RegCM4, over the region and over time. A Bayesian functional analysis of variance (ANOVA) approach is used to simultaneously model the temperature patterns from the two RCMs for the current and future climate. We exploit nonstationary spatial models to handle the spatial dependence structure of the temperature variable, which depends heavily on latitude and altitude. For a seasonal comparison, we examine changes in the winter temperature in addition to the summer temperature data. We find that the temperature increase projected by RegCM4 tends to be smaller than the projection of HadGEM3-RA for summers, and that the future warming projected by HadGEM3-RA tends to be weaker for winters. Also, the results show that there will be a warming of 1-3°C over the region in 45 years. More specifically, the warming pattern clearly depends on the latitude, with greater temperature increases in higher latitude areas, which implies that warming may be more severe in the northern part of the domain.

Key words: Bayesian hierarchical model, CORDEX-East Asia, nonstationary covariance function, regional climate model, spatial regression

1. Introduction

COrdinated Regional climate Downscaling EXperiment (CORDEX) is a program sponsored by the World Climate Research Programme (WCRP). Its goal is to organize an international coordinate framework to produce improved generation of regional climate change projections worldwide. CORDEX consists of multiple ensembles (both dynamically and statistically downscaled versions) of climate variables over several regional domains around the globe. For more details, see Giorgi et al. (2009). CORDEX-East Asia is a branch of the CORDEX initiative for an East Asian region where weather and climate extremes, such as heat wave and heavy rainfall events, are expected to occur more frequently with stronger amplitude under global warming (Min et al., 2015). Currently, outputs are available for 72 years from 5 Regional Climate Models (RCMs), HadGEM3-RA, RegCM4, SNU-MM5, SNU-WRF, and YSU-RSM, constrained by a Global Climate Model (GCM), HadGEM2-AO; these include 27 year simulations for the current (1979-2005) climate and 45 year simulations for the future (2006-2050) climate. More information on the experimental set-up and other details can be found at http://cordex-ea.climate.go.kr.

In the research literature, there have been several studies on the CORDEX-East Asia data. Using the simulation results from 1989 to 2008, Suh and Oh (2012) examined the prediction skills of five ensemble methods for temperature and precipitation. Using the precipitation simulation results from 1989 to 2006, Park et al. (2013) studied the impact of boundary conditions for the RegCM4 model. Moreover, Park et al. (2016) analyzed the performance of five RCMs in simulating temperature and precipitation extremes in the summer over East Asia.

There also have been GCM-based studies for the East Asia region. Baek et al. (2013) presented GCM simulation results from HadGEM2-AO under four Representative Concentration Pathway (RCP) scenarios. Their work is focused on presenting simulation results for climate change in the 21st century, including global and regional changes in temperature and precipitation. The future East Asian climate was projected to be warmer and wetter than the global mean change. Min et al. (2004) also analyzed multiple GCM simulation results for climate change in the 21st century over East Asia (20-50°N, 100-145°E). Their multi-model analysis consistently suggested that East Asia will likely become warmer and wetter.

However, more rigorous statistical analysis of the data from CORDEX-East Asia is scarce, in contrast to the North American Regional Climate Change Project (NARCCAP), another international program for regional climate change

Corresponding Author: Mikyoung Jun, Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77840, USA.
E-mail: mjun@stat.tamu.edu

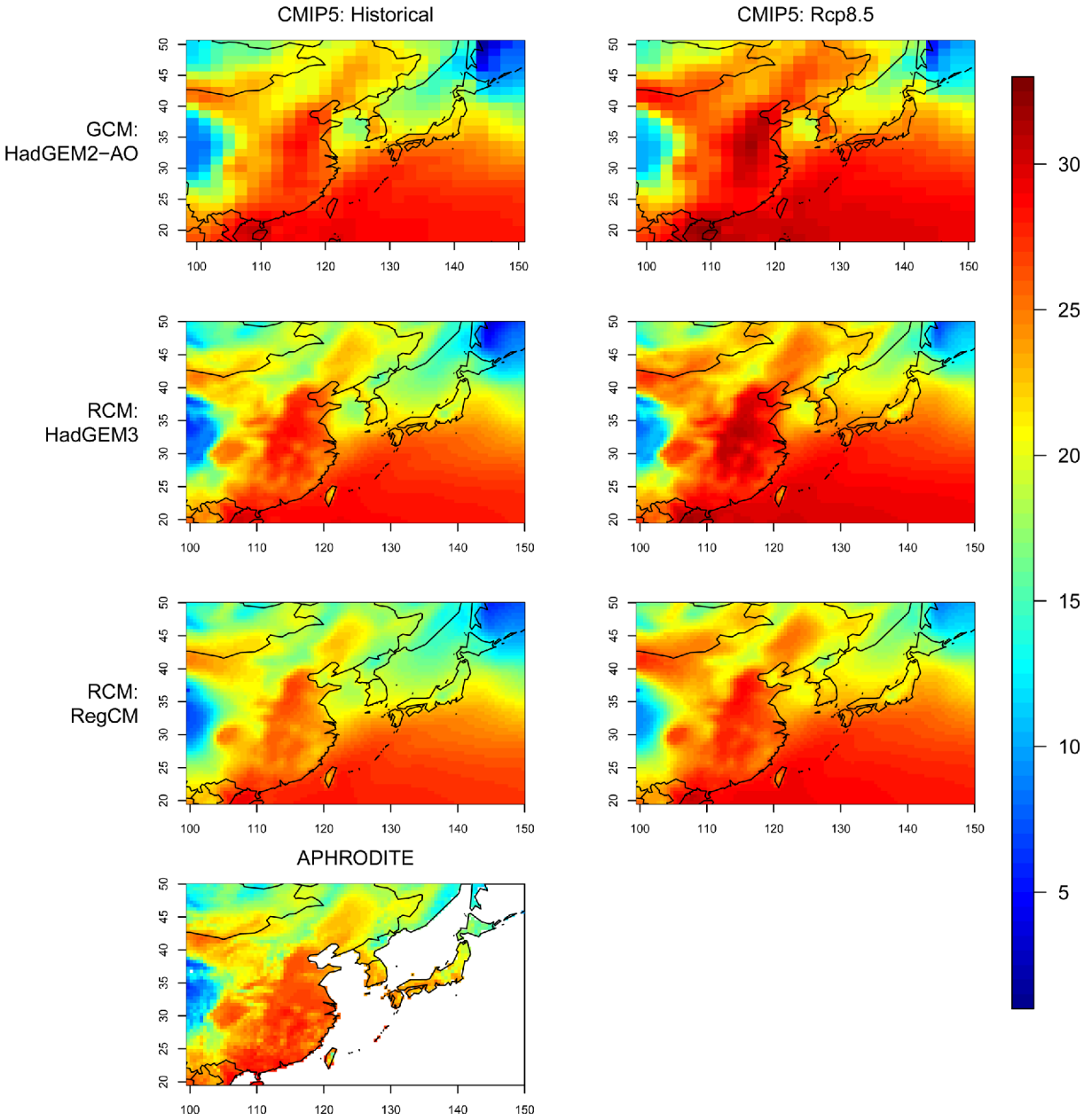

Fig. 1. Spatial map of surface air temperatures (unit: °C) by GCM (HadGEM2-AO) and two RCMs (HadGEM3-RA and RegCM4) for the historical (left) and RCP8.5 (right) scenario experiments. Each figure shows JJA averages, averaged over 27 years. APHRODITE data for the years 1979-2005 (unit: °C) is shown in the last row.

projection experiments (http://www.narccap.ucar.edu). For example, Sain et al. (2011) assessed differences between two downscaling methods and their projections of summer temperature and precipitation using NARCCAP outputs. Greasby and Sain (2011) used a hierarchical Bayesian spatial model approach to characterize the temperature profiles for seasons of the year, current and future, using NARCCAP outputs. Kang et al. (2011) used a Bayesian hierarchical spatial model to combine multiple regional climate model outputs from the NARCCAP experiment. Most of these studies find that spatial dependence in the NARCCAP outputs is significant. Thus, accounting for such spatial dependence structures gives more *statistically efficient* analysis results. In other words, ignoring spatial dependence may lead to inaccurate results with greater uncertainty (i.e., statistical inefficiency). We expect that CORDEX-East Asia outputs exhibit strong spatial dependence as well, similar to the NARCCAP outputs, and we need to take this into account in our statistical models.

Our goal in this paper is to provide a statistically rigorous analysis of outputs in the CORDEX-East Asia experiment that accounts for the spatial dependence structure of the data. In particular, we assess how the mean state of the temperature

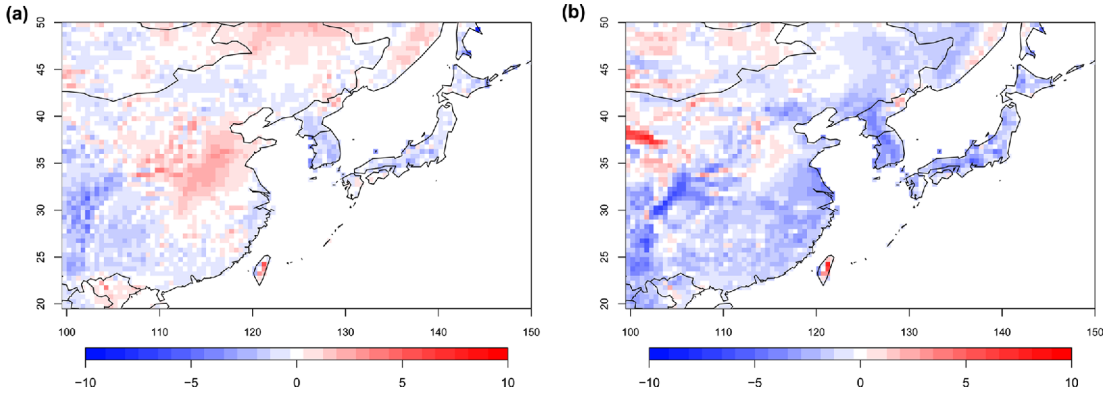

Fig. 2. Model bias (i.e. differences between model output and observation) of two RCMs, (a) HadGEM3-RA, (b) RegCM4, for JJA mean temperatures, averaged over 1979-2005 (unit: °C). APHRODITE data are used as observations.

varies across different RCM outputs. The two RCMs we consider are HadGEM3-RA, developed by the Met Office Hadley Centre (Davies et al., 2005), and RegCM4, developed by the International Centre for Theoretical Physics (Giorgi et al., 2012). For the current climate, we consider the historical simulations (1979-2005), and for the future climate, we consider the outputs with RCP8.5 runs (2024-2050), which assume a high CO_2 concentration scenario in the future describing a rising radiative forcing pathway leading to 8.5 W m^{-2} in 2100 (Moss et al., 2010). Within the entire domain of the CORDEX-East Asia experiment, we focus on the domain covering part of China, Korea, and Japan (latitude range 20°-50°N and longitude range 100°-150°E). See Fig. 1 for a map of the domain considered in this work.

The rest of this paper is organized as follows. Section 2 describes the statistical method and computational details. Section 3 presents the results of our analysis. Finally, we conclude in Section 4 with a possible extension of our work. An appendix provides details about our statistical models and their implementation.

2. Data and Statistical Method

a. Data description

We use historical surface temperature data from 1979 to 2005 as well as future projections with the RCP8.5 scenario from 2024 to 2050: a total of 27 years of data for both the current and the future climate. The RCM outputs from the CORDEX archive come as monthly averages, and we take seasonal averages over June-August (JJA) and December-February (DJF) for each given year. We consider two RCM outputs, from HadGEM3-RA and RegCM4, which have slightly different horizontal resolutions of 0.44° and 50 km, respectively. To deal with this grid discrepancy, we used linear interpolation with a triangulation scheme to obtain the data on common grids of 0.5° × 0.5°. The interpolated data has 61 (latitude) by 101 (longitude) grid points over the East Asian domain (Fig. 1).

Figure 1 shows the spatial distributions of historical and future JJA mean temperatures from the two different RCMs and observations. We can clearly see east-west as well as north-south patterns, having warm summers over eastern China. The RCM results we use are all driven from the same GCM, HadGEM2-AO. The effect of GCM boundary forcing on the outputs of RCMs is known to be significant (Christensen et al., 2013; Park et al., 2016). Figure 1 displays the effect of GCM on the RCM outputs. We note that, for the Boreal summer temperature, the HadGEM3-RA RCM produces slightly higher temperatures than the RegCM4 model.

For further evaluation of the two RCMs, we compare observations with the RCM outputs. We used the Asian Precipitation-Highly Resolved Observation Data Integration Towards Evaluation of Water Resources (APHRODITE; Yatagai et al., 2012) data for the years 1979-2005 as observations (shown in Fig. 1). APHRODITE is the only data over monsoon Asia (15°S-55°N, 60-155°E) that provides high-resolution observations, both spatially and temporally. We calculate seasonal means from daily mean temperatures on a grid with a 0.5° × 0.5° resolution, which was created based on quality-controlled station observations (Yasutomi et al., 2011). The number of stations used for this data set was 1.5-3 times larger than the number used for the Global Telecommunication System (GTS) reports. The monthly mean temperature climatology was shown to be comparable to other gridded temperature products (Yasutomi et al., 2011).

Figure 2 shows model biases, i.e., differences between model outputs and observations. We notice that even though HadGEM3-RA produced higher temperature data than RegCM4 for the Boreal summer over the region, it gives lower temperatures than the observed data over a large portion of the region, including the Korean Peninsula. This cold bias in JJA is known to be associated with a cold sea surface temperature (SST) bias in HadGEM2-AO GCM (Park et al., 2016). Overall, RegCM4 tends to produce lower temperatures than HadGEM3-RA. Although one could incorporate observations into the statistical model fitting procedure, which would be equivalent to giving weights to climate models based on model

biases (e.g., Hawkins and Sutton, 2009; Salazar et al., 2011), we do not use the observation data for the statistical analysis in the next section.

b. Functional ANOVA model

Our statistical model is based on the functional Analysis of Variance (ANOVA) approach. It is called a "functional" ANOVA because each factor in the ANOVA structure comes in functional form (i.e. as a function of spatial locations). The categorical factors include the choice of RCM and the choice of the period for GCM (historical run or future RCP8.5 scenario run). The functional ANOVA approach is a useful tool for comparing the data with respect to four different combinations of RCM/period. Several works have analyzed climate model outputs with the functional ANOVA approach, including Kaufman and Sain (2010) and Salazar et al. (2011). Salazar et al. (2011) considered discrepancies between observations and regional climate model outputs in the space-time context. Our approach is similar to the one in Kaufman and Sain (2010), and we only consider model outputs in spatial context in the functional ANOVA framework.

Let $Y_{ijt}(\boldsymbol{s})$ denote the output from the ith RCM model ($i = -1$ for HadGEM3-RA and $i = 1$ for RegCM4), the jth period ($j = -1$ for the current run and $j = 1$ for the future run) at time t ($t = 1,\ldots,27$) and location $\boldsymbol{s}$. Then we define

$$Y_{ijt}(\boldsymbol{s}) = \mu(\boldsymbol{s}) + i\alpha(\boldsymbol{s}) + j\beta(\boldsymbol{s}) + ij(\alpha\beta)(\boldsymbol{s}) + \gamma(t - 14) + \epsilon_{ijt}(\boldsymbol{s}). \quad (1)$$

Here, $\mu(\boldsymbol{s})$ represents the grand mean temperature at location $\boldsymbol{s}$. Two main effects, $\alpha(\boldsymbol{s})$ and $\beta(\boldsymbol{s})$, represent the deviation from the grand mean for the chosen RCM and time period, respectively. The interaction term, $(\alpha\beta)(\boldsymbol{s})$, enables us to differentiate the rate of temperature increase with two different RCMs. Finally, γ represents the temperature change in one year, which we assume fixed for spatial location $\boldsymbol{s}$. Similarly to Kaufman and Sain (2010), we assume that the terms, μ, α, β, and $(\alpha\beta)$ follow a Gaussian distribution.

The term μ in Eq. (1) may require the most complex spatial structure of all the terms, since all other terms describe deviations from μ and μ itself represents the grand mean over the domain. This term is modeled through a nonstationary covariance model. In particular, we allow the variance to be different over the land and the ocean. That is, for σ_μ, $\eta > 0$, we let

$$\mu(\boldsymbol{s}) = \begin{cases} \mu_0 + \sigma_\mu e^{\eta} W(\boldsymbol{s}), & \text{if } \boldsymbol{s} \in \text{land}, \\ \mu_0 + \sigma_\mu W(\boldsymbol{s}), & \text{if } \boldsymbol{s} \in \text{sea}, \end{cases} \quad (2)$$

where W is a mean zero Gaussian process. The covariance structure of W is given by a Matérn correlation function (Stein, 1999) that only depends on distance:

$$C(d;\phi,\nu) = \frac{1}{2^{\nu-1}\Gamma(\nu)}\left(\frac{d}{\phi}\right)^{\nu} \mathcal{K}_\nu\left(\frac{d}{\phi}\right), \quad (3)$$

with ϕ, $\nu > 0$ and $\mathcal{K}_\nu$ a modified Bessel function of order ν. That is, we have $\text{Cov}\{W(\boldsymbol{s}), W(\boldsymbol{u})\} = C(|\boldsymbol{s} - \boldsymbol{u}|; \phi_\mu, \nu_\mu)$. Here, $|\boldsymbol{s} - \boldsymbol{u}|$ denotes the distance between the two locations $\boldsymbol{s}$ and $\boldsymbol{u}$. Note that the covariance model in Kaufman and Sain (2010) is isotropic (which implies constant variance over the entire domain), and this isotropic assumption may be limited for the μ term in Eq. (1). We fix $\nu_\mu = 2$, similarly to Kaufman and Sain (2010).

We assume that the mean structure of the grand mean temperature process μ at location $\boldsymbol{s}$, $\mu_0(\boldsymbol{s})$, can be well modeled by the following three factors: the latitude of location $\boldsymbol{s}$, the elevation at location $\boldsymbol{s}$, and whether $\boldsymbol{s}$ is located on the land or the sea. The term μ_0 is modeled through a regression equation that depends on the latitude (L) and elevation (A, unit: km) in the following way:

$$\mu_0(\boldsymbol{s}) = \begin{cases} \mu_{00} + \mu_{01}\sin\{L(\boldsymbol{s})\} + \mu_{02} + \mu_{03}A(\boldsymbol{s}), & \text{if } \boldsymbol{s} \in \text{land} \\ \mu_{00} + \mu_{01}\sin\{L(\boldsymbol{s})\}, & \text{if } \boldsymbol{s} \in \text{sea}. \end{cases} \quad (4)$$

The structure in Eq. (4) is useful for explaining the large-scale variation of the μ term. The data set on elevation and land/sea mask is obtained from the Joint Institute for the Study of the Atmosphere and Ocean (http://research.jisao.washington.edu/data_sets). To deal with the discrepancy of the grid resolution for different RCM models, we interpolated the elevation variable to the common $0.5^\circ \times 0.5^\circ$ grid. Further details on the Bayesian model set up and computational techniques are presented in the Appendix.

3. Analysis results

a. Results for JJA

Figure 3 shows posterior means for μ, α, β, and $(\alpha\beta)$, which represent the grand mean, regional model, period, and interaction, respectively [Eq. (1)]. This figure shows that the interaction term [bottom right, $(\alpha\beta)$], is very small, compared to the other main effects. It also suggests that the temperature from the RegCM4 model is lower than the temperature from the other RCM (top right) and that there is warming over time, across the entire domain (bottom left). Finally, it shows that the magnitude of the temperature increase depends on the latitude (bottom left).

Using posterior samples of $(\alpha\beta)$, we form a 95% credible set on the interaction term. If the lower bound of the 95% credible set exceeds 0 at a location (grid pixel), we say that $(\alpha\beta)$ is "credibly positive". Similarly, if the upper bound of the 95% credible set is smaller than 0, we say that $(\alpha\beta)$ is "credibly negative". Black grid boxes in Fig. 4a indicate a credibly positive $(\alpha\beta)$, whereas white grid boxes indicate a credibly negative $(\alpha\beta)$. There is a clear coastline effect, although similarly to what Fig. 3 suggests, overall, the size of the interaction is small.

Turning our attention to the main effects, Fig. 4b shows results for the main effect of the regional model, α. From Eq.

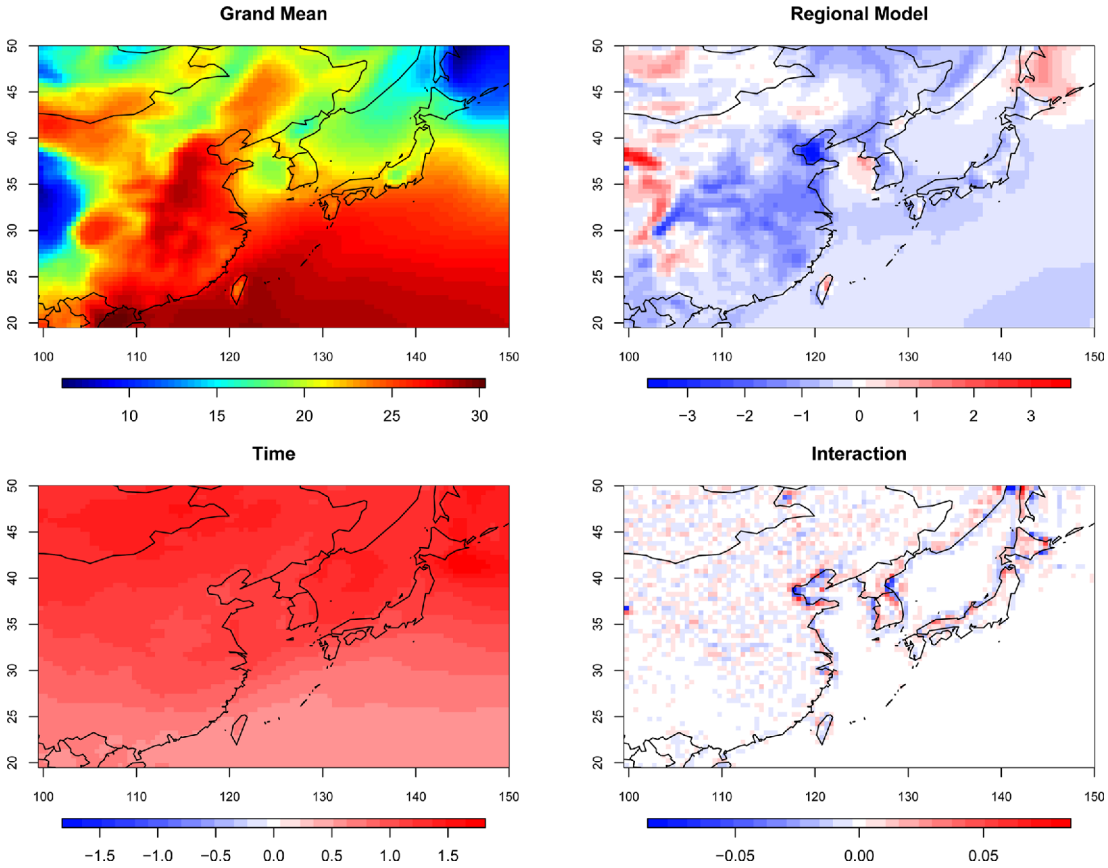

Fig. 3. Posterior means of μ ("Grand Mean"), α ("Regional Model"), β ("Time"), and $(\alpha\beta)$ ("Interaction"), for JJA mean temperatures (unit: ^{o}C).

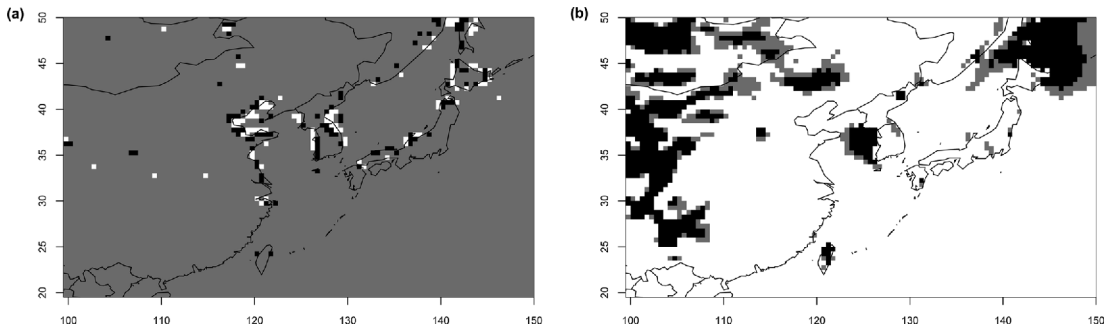

Fig. 4. JJA case: (a) shows the regions where $(\alpha\beta)$ is credibly different from zero (black regions suggest positive $(\alpha\beta)$ and white regions suggest negative $(\alpha\beta)$). (b) shows the regions where α is credibly different from zero (black regions suggest positive α and white regions suggest negative α).

(1), with $i = -1$ for HadGEM3-RA and $i = 1$ for RegCM4, 2α represents the difference in temperature between the two RCMs. Black grid boxes indicate a credibly positive α, whereas white grid boxes indicate a credibly negative α. In other words, the output from the RegCM4 model gives higher temperatures in black regions and lower temperatures in white regions, compared to the HadGEM3-RA model. In Section 2, we discussed the differences in the outputs from the two RCMs. Figure 4b gives similar results. It shows that RegCM4 tends to produce higher temperatures in inland areas (over central China and Mongolia) and lower temperatures in other areas, compared to HadGEM3-RA. This agrees with the difference in model biases shown in Fig. 2, although the biases in Fig. 1 only concern the current time period, whereas α in Fig. 4 concerns both the current and future time periods.

We have seen from Fig. 3 that the posterior means of the main effect regarding the period, β, is positive over the whole domain, implying warming across the entire domain. The term 2β represents the increases in temperature by the period 2024-2050 compared to the period 1979-2005. More precisely, $2\beta +$

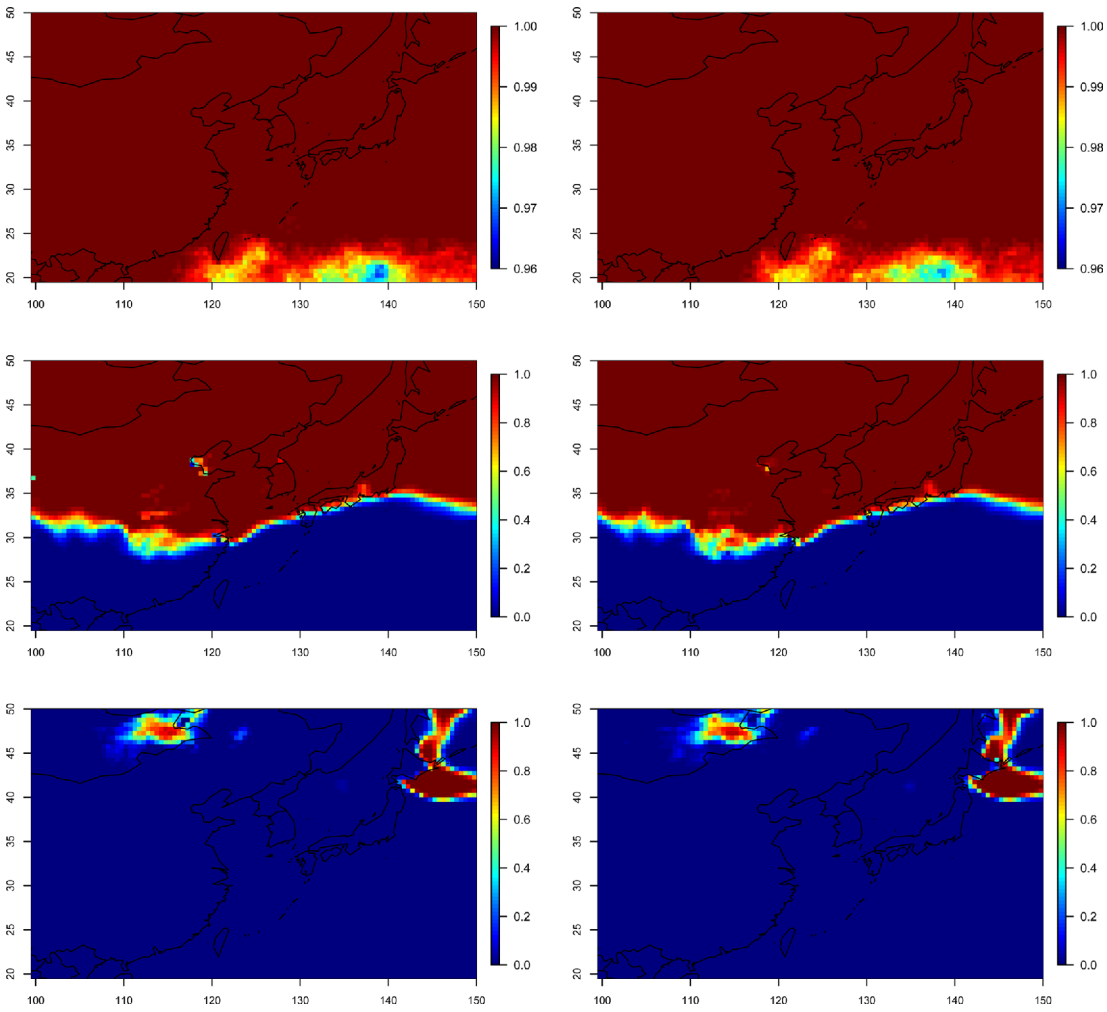

Fig. 5. Posterior probabilities that temperature (JJA case) increases more than 1°C (top), 2°C (middle), and 3°C (bottom) from RegCM4 model (left) and HadGEM3-RA model (right).

$2(\alpha\beta)$ represents the temperature increases predicted by the RegCM4 regional model, and $2\beta - 2(\alpha\beta)$ represents the temperature increases predicted by the HadGEM3-RA regional model. Thus, in this sense Fig. 3 implies a temperature increase of about 1-3°C in 45 years. Figure 5 gives more detailed information about β. It displays the posterior probabilities that the temperature increases more than 1°C, 2°C, and 3°C in 45 years, respectively (row-wise) from the two RCM models (column-wise). The first column displays $2\beta + 2(\alpha\beta)$ (for RegCM4), and the second $2\beta - 2(\alpha\beta)$ (for HadGEM3-RA). The two columns of Fig. 5 look similar, though this is due to a small interaction term $(\alpha\beta)$. The first row suggests that the probability that the temperature increases more than 1°C is higher than 0.96 over the entire domain (from both models). The second row clearly demonstrates that the northern part of the domain is expected to suffer from more serious warming. Note that the area-averaged value of the temperature increase by 2024-2050 under the RCP8.5 scenario is estimated to be 2.11 (this value is obtained from posterior samples of 2β, averaged over the domain).

Our results are comparable to results in the literature for similar regions. For instance, Baek et al. (2013) analyzed HadGEM2-AO, which is used to derive the RCM results. Table 2 of Baek et al. (2013) shows that the estimated area-averaged temperature increase over the East Asia region (20-50°N, 100-150°E) is greater than the area-averaged temperature increase over the global region. Specifically, using RCP8.5, the area-averaged JJA temperature change for 2081-2100 relative to 1986-2005 is estimated to be 4.7°C over the East Asia region, whereas the temperature change is estimated to be 4.0°C over the global domain. The annual temperature is estimated to increase by 4.6°C over the East Asia region and 4.1°C over the global domain. According to Baek et al. (2013) (using the RCP 8.5 scenario), the global area-averaged surface air temperature will increase by 1-2°C by 2050 from the temperature for 1986-2005. These results imply that the area-averaged temperature over the East Asia region will increase slightly more than 1-2°C by 2050 from the temperature for 1986-2005, which agrees with our results.

Figure 5 (particularly the second row) suggests that the

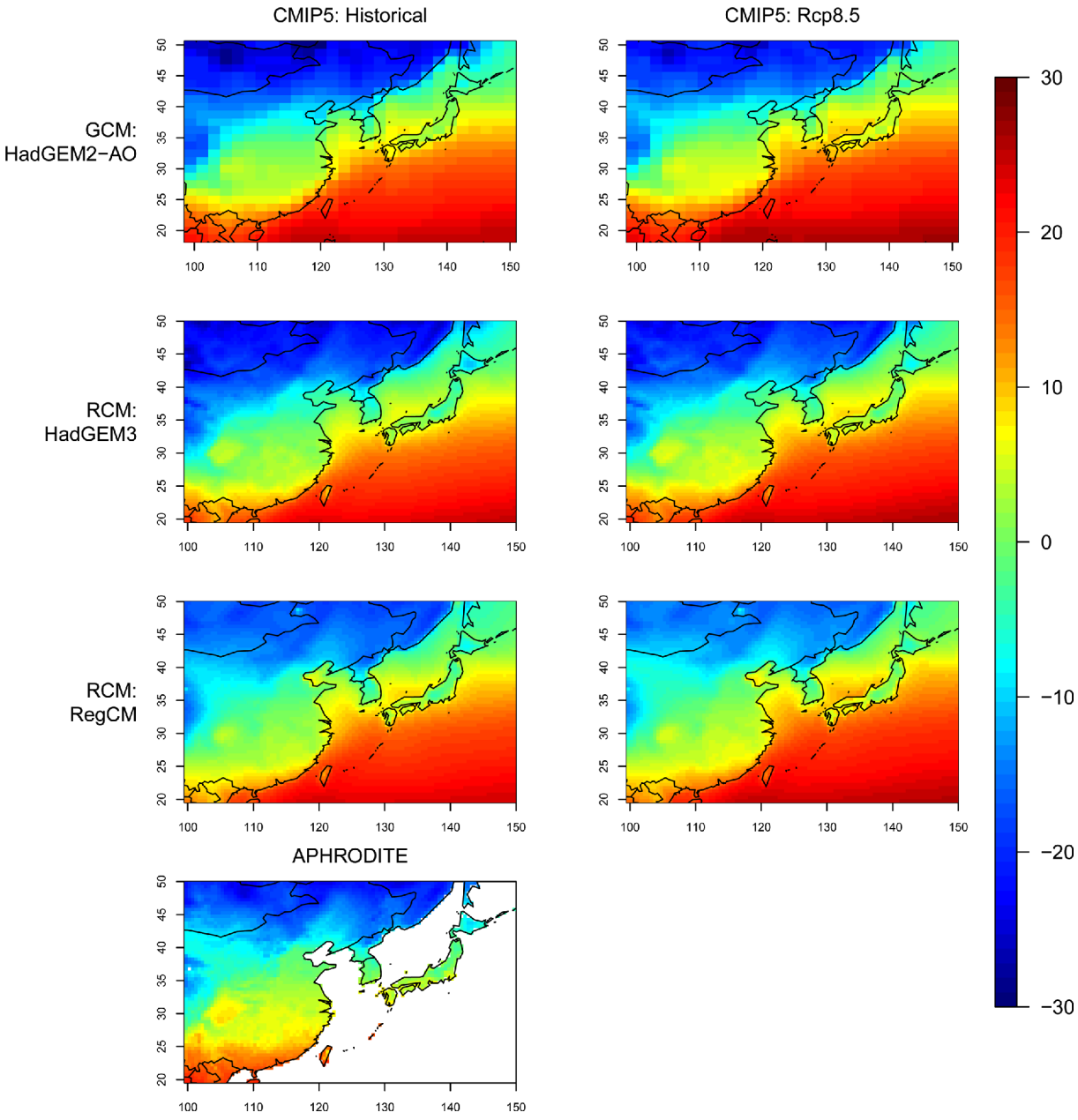

Fig. 6. Spatial map of surface air temperatures (unit: °C) by GCM (HadGEM2-AO) and two RCMs (HadGEM3-RA and RegCM4) for the historical (left) and RCP8.5 (right) scenario experiments. Each figure shows DJF averages, averaged over 27 years. APHRODITE data for the years 1979-2005 (unit: °C) is shown in the last row.

latitude plays an important role in temperature increases. We find similar results in previous studies of the East Asian region (Min et al., 2004; Baek et al., 2013). Min et al. (2004) used CMIP3 multi-model ensembles to predict the future climate change over the East Asia region [20-50°N, 100-145°E, based on Special Report on Emissions Scenarios (SRES) scenarios]. Figure 5 in Min et al. (2004) clearly shows the effect of latitude on the temperature changes across the region, suggesting a larger temperature increase in the region with the higher latitude. Min et al. (2004) also suggest that the area-averaged temperature will increase by 2.49°C by the 2050s, compared to 1961-1990. Considering the fact that our estimate is relative to 1979-2005 and that different scenarios are used, our estimate of 2.11°C across the East Asia domain seems comparable.

b. Results for DJF

To compare our results for JJA to another season, we considered the seasonal average over the Boreal winter (DJF) and conducted the same spatial analysis. Figure 6 shows a spatial map of DJF mean temperatures over the domain.

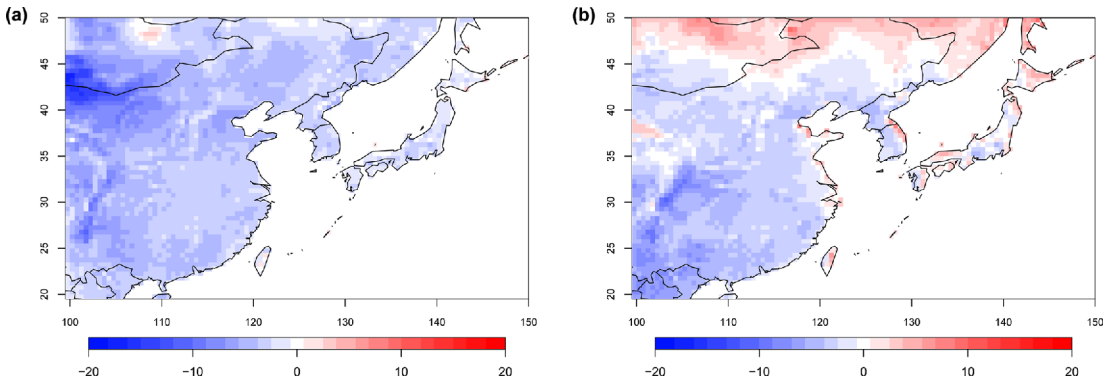

Fig. 7. Biases from two RCMs (unit: °C). (a) HadGEM3-RA, (b) RegCM4 for DJF temperatures (1979-2005).

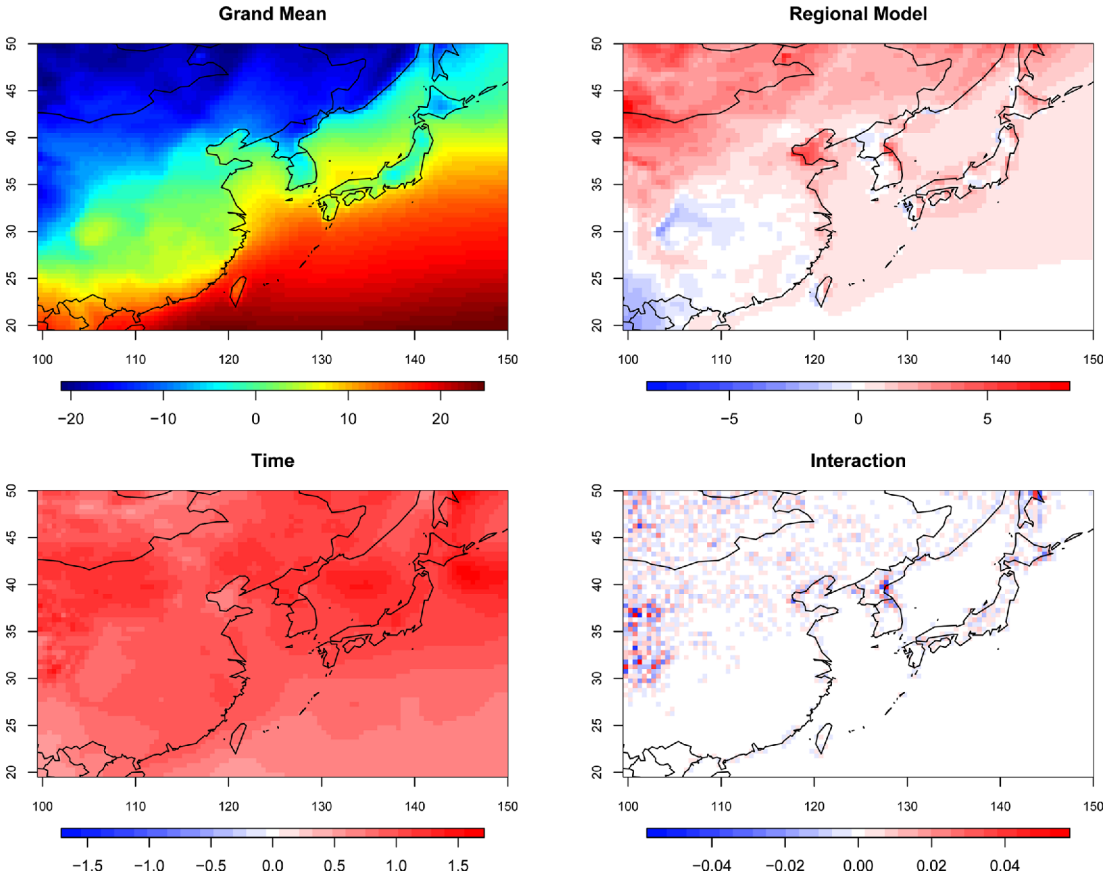

Fig. 8. Posterior means of μ ("Grand Mean"), α ("Regional Model"), β ("Time"), and $(\alpha\beta)$ ("Interaction"), for DJF mean temperatures (unit: °C).

Unlike the Boreal summer, we see that the RegCM4 model gives slightly higher temperatures for the winter than the HadGEM3-RA model. The RCM outputs are also compared to the observation data (area-averaged value for 1979-2005). Figure 7 shows the model biases of the two RCM outputs. The two RCMs produce lower temperatures than the observed temperature over most of the domain. Comparing Figs. 2 and 7, we see that the cold biases for the winter temperature are about twice as large as those for the summer.

For a Bayesian functional ANOVA analysis, we fit the DJF data to a similar model, using Eq. (1). We used historical surface temperature data from December 1979 to February 2005, as well as a future projection with the RCP8.5 scenario from December 2024 to February 2050: 26 years of data each for both the current and the future climate. We consider the DJF average over each year. Figure 8 shows posterior means for μ, α, β, and $(\alpha\beta)$ for Boreal winter temperatures. Unlike the JJA case, temperatures produced from the RegCM4 model

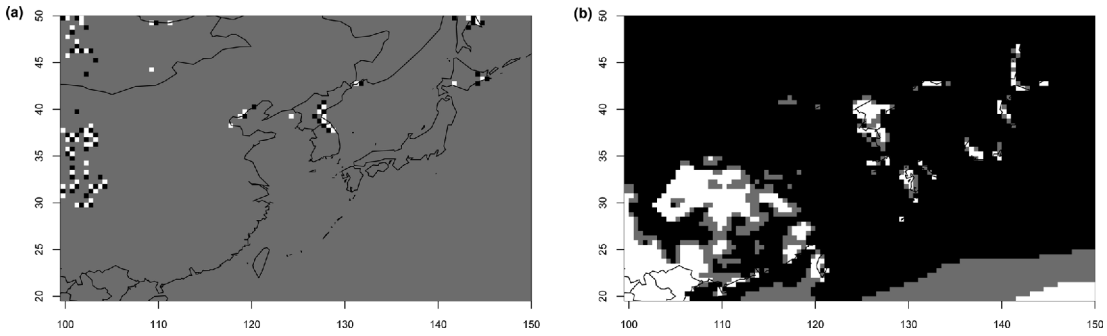

Fig. 9. DJF case: (a) shows the regions where ($\alpha\beta$) is credibly different from zero (black regions suggest positive ($\alpha\beta$) and white regions suggest negative ($\alpha\beta$). (b) shows the regions where α is credibly different from zero (black regions suggest positive α and white regions suggest negative α).

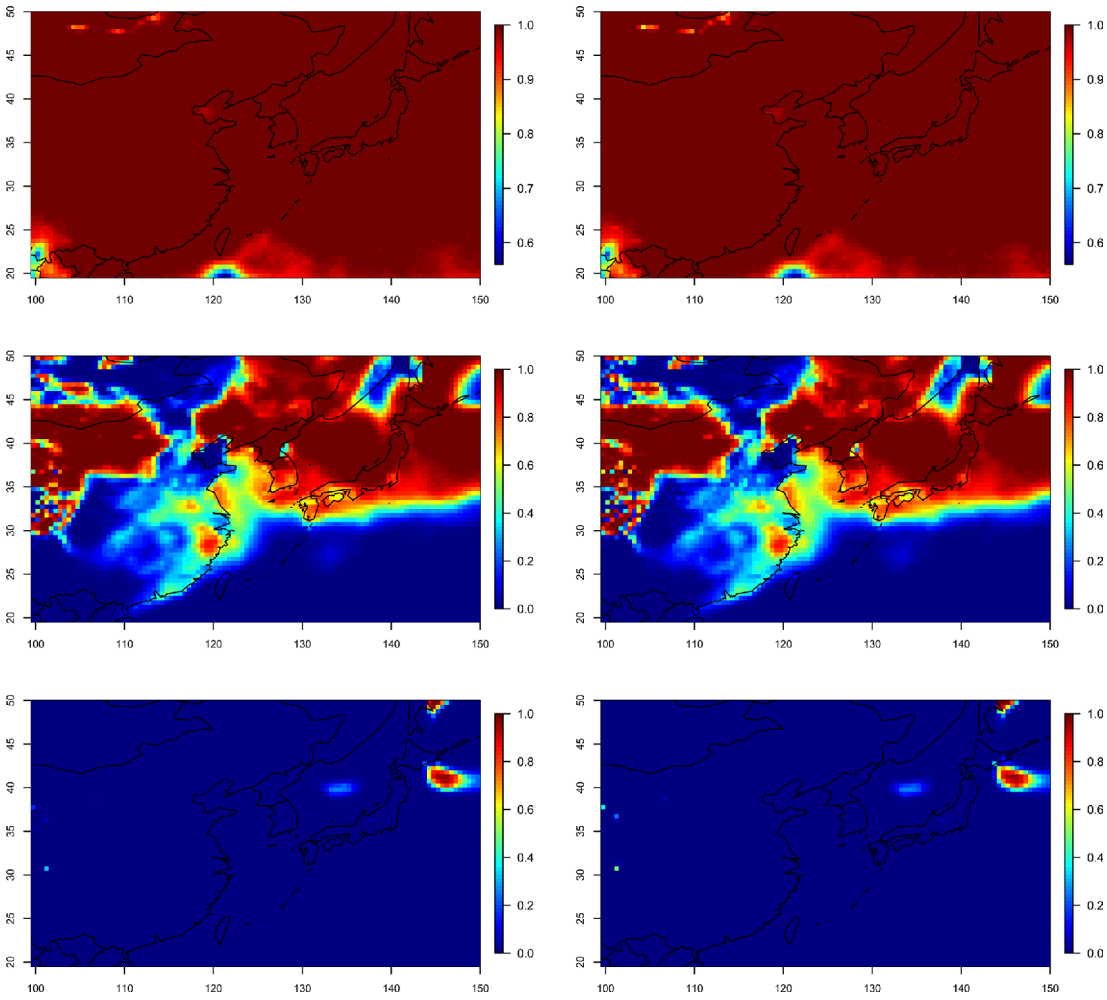

Fig. 10. Posterior probabilities that temperature (DJF case) increases more than 1°C (top), 2°C (middle), and 3°C (bottom) from RegCM4 model (left) and HadGEM3-RA model (right).

are higher. This agrees with what we see in Figs. 6 and 7. Figure 9a shows that the interaction over the domain is credibly close to 0. Figure 9b, which gives the main effect of the regional model, α, suggests that the output from the RegCM4 model tends to produce higher temperatures over a large part of the domain except for the lower left corner region, which agrees with Figs. 7 and 8. Figure 8 (bottom left) shows the posterior means of β, the main effect regarding time. It suggests warming over the domain. Furthermore, the first row of Fig. 10 suggests that the probability of a temperature

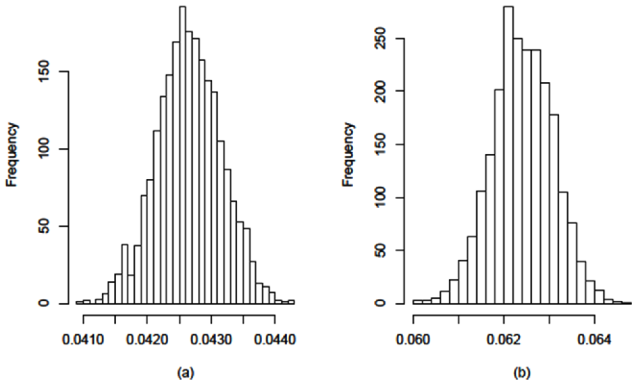

Fig. 11. Histogram of the posterior samples of γ (a) JJA, (b) DJF.

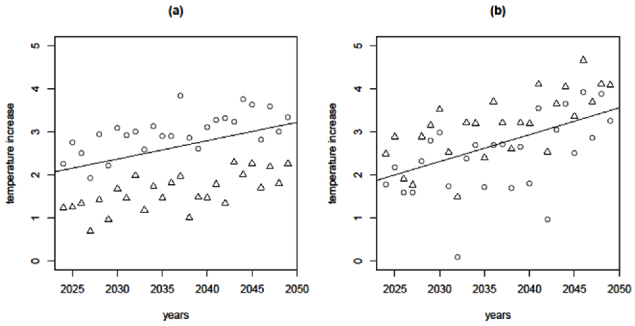

Fig. 12. Estimated temperature increases (°C) relative to the year 1979 in (a) JJA, (b) DJF. Circles indicate the data from HadGEM3-RA regional model and triangles indicate the data from RegCM4 model.

increase greater than 1°C is greater than 0.95 over most of the domain (similarly to JJA case). It suggests that the temperature change in 2024-2050 relative to 1979-2005 for DJF is about 1-2°C over the lower latitude regions and 2-3°C over the higher latitude regions. The area-averaged temperature increase in winter over East Asia by 2024-2050 from 1979 to 2005 is about 1.94°C.

c. Results regarding the trend

To see the seasonal effect on temperature increases, we check the estimated values for both β and γ in Eq. (1). Here, 2β indicates the temperature increases between the years 2024-2050 and 1979-2005. In contrast, γ indicates the linear trend of temperature increases in a single year. Also, β is a spatial field, but γ is a fixed constant over the domain.

Figure 11 shows the distributions of the posterior samples of γs for summer and winter. Even though the area-averaged value of β for the summer (1.05671) is slightly greater than the area-averaged value of β for the winter (0.96769), this may not mean that the warming in the summer will be more intense than in the winter. Since the estimate for γ for winter is larger than the estimate for summer, the temperature increase for consecutive years is estimated to be larger for winter.

Figure 12a shows the estimated area-averaged temperature increase relative to the year 1979 in the Boreal summer. The line shows the estimated temperature increases compared to the summer of 1979, i.e.,

$$\gamma(x - 2024) + 2\beta, x \in (2024, 2050) \quad (5)$$

Remember that 2β gives the temperature increase over 45 years (temperature increase from 1979 to 2024) and γ indicates the linear trend of temperature increases in single consecutive years. Figure 12b shows the same information as Fig. 12a, but for the Boreal winter. Since the estimated value of β (area-averaged value) is higher for the summer, the line intercept in Fig. 12a is larger than in Fig. 12b. However, since the estimated value of γ (the mean value) is larger for the winter, the slope of the line in Fig. 12b (0.06241) is greater than the slope in Fig. 12a (0.04265). Figure 12 shows a clear distinction between the two RCM models in terms of a temperature increase relative to 1979 for JJA, but not for DJF. Priors, posterior means, and credible intervals for all the parameters considered for the two seasons, JJA, and DJF, are shown in Tables 1 and 2 in the Appendix.

4. Conclusions

In this study, the spatial patterns of temperature changes are analyzed using two regional climate model (RCM) simulations participating in the CORDEX-East Asia experiment-HadGEM3-RA and RegCM4. The spatial covariability of temperatures is rigorously examined using the Bayesian functional Analysis of Variance (ANOVA) approach, which simultaneously models space-time structures of temperature change patterns. This method enables examination of similarities and differences in temperatures between two RCMs over the present and future periods. The spatial dependence structure of the temperature is allowed to depend on latitude and altitude. We consider both summer and winter for a seasonal comparison.

Temperature patterns from the two RCMs are similar to those from GCMs, which represent dominant force of the driving global climate models (GCMs), as indicated by previous studies. The analysis results show that HadGEM3-RA tends to produce warmer temperatures than RegCM4 during the summer, while the opposite is seen during the winter - the HadGEM3-RA projection tends to be colder than the RegCM4 projection. Comparison of the model biases suggests that this difference in the projected temperature change is mainly due to different model climatologies for the seasons. This implies that Bayesian functional ANOVA is quite sensitive to model bias and that model skills are one of the important factors for this analysis.

The results also suggest that the average temperature increase over East Asia will be 1-3°C by the mid-21st century, which is similar to previous studies based on different models and scenarios. This seems to be due to the fact that there are not significant differences in radiative forcing among different emission scenarios for SRES or RCP for the near future

period, which results in relatively small differences in temperature changes.

Our spatial analysis clearly identifies a north-south difference in temperature changes that dominates the spatial pattern of warming, i.e. a latitudinal dependence. This implies that stronger and faster temperature increases are expected to occur at higher latitudes than at lower latitudes.

The main objective of this study was to exploit a spatial analysis approach to East Asian temperatures in a relatively simple setting. Our functional ANOVA model assesses the significance of the differences in temperature projected by two regional climate models, differences in current and future temperatures, differences in the rate of temperature increase, and linear trends, accounting for the spatial dependence of these factors as well as the error in the ANOVA model. Our method has many caveats. First, we use limited experiment datasets from one GCM combination of two RCMs. Second, we only considered one scenario, RCP8.5. It is well known that projected future climate changes at regional scales are characterized by high uncertainties, largely due to model differences and scenario uncertainties (e.g., Hawkins and Sutton, 2009; Salazar et al., 2011). In the future, we plan to carry out an extended analysis, considering multi-models and multi-scenarios in our statistical model.

Acknowledgments. This research was supported by NSF grant DMS-1208421. S.-K. Min was funded by the Korea Meteorological Administration Research and Development Program under Grant KMIPA 2015-2082.

Edited by: Song-You Hong, Kim and Yeh

Appendix: Computational details

In this appendix, we offer details about the Bayesian implementation of our model fitting.

1. Priors:

We put a Gaussian distribution prior to each effect. In particular, we let $\alpha \sim N(0, \sigma_\alpha^2 C(\cdot;\phi_\alpha, 2))$, $\beta \sim N(\mu_\beta, \sigma_\beta^2 C(\cdot;\phi_\beta, 2))$, and $(\alpha\beta) \sim N(0, \sigma_{int}^2 C(\cdot;\phi_{int}, 2))$. We assume that $\in_{ijt} \sim N(0, \sigma_\in^2 C(\cdot;\phi_\in, 2))$. The covariance function C is defined in Eq. (3).

We find that for the JJA temperatures, μ_β is well explained by latitude, since latitude seems to be the dominant factor in describing the term β. Thus, for the JJA temperature analysis, we write μ_β as a linear combination of Legendre polynomials of order 3 with the sine of the latitude as their arguments:

$$\mu_\beta(\boldsymbol{s}) = \mu_{20} + \mu_{21}P_1[\sin\{L(\boldsymbol{s})\}] + \mu_{22}P_2[\sin\{L(\mathrm{s})\}] + \mu_{23}P_3[\sin\{\mathrm{L}(\boldsymbol{s})\}], \quad (6)$$

where L denotes the latitude and $P_n(x)$ is the nth Legendre polynomial. We put a uniform prior $U(-90, 90)$ on the coefficient terms for μ_β, i.e., μ_{20}, μ_{21}, μ_{22}, and μ_{23}.

For the DJF temperatures, β is more complex and cannot be fully explained by latitude alone. Thus, we sample μ_β directly from the vector space. More specifically, we put a uniform prior on $\mu_\beta \in \mathbb{R}^{6161}$. This prior distribution, $U(\mathbb{R}^{6161})$, is an improper prior, but it yields a proper posterior distribution, which is multivariate normal in our setting.

We put a non-informative prior distribution on the remaining parameters including the coefficient terms in the linear combinations for μ_0. Here, the coefficient terms for μ_0 are μ_{00}, μ_{01}, μ_{02}, and μ_{03}. We let $\gamma \sim U(-5,5)$, which reflects our belief that the average temperature change in a given year is less than 5°C. For covariance parameters, we let $\sigma^2 \sim U(0,1250)$, $\eta \sim U(0,3)$, and $\phi \sim U(0,1000)$. For σ, we followed the same rational as Kaufman and Sain (2010), and used $\sigma^2 \sim U(0,1250)$. Due to the greater heat capacity of the sea, the rate of the temperature increase over land is about twice the rate of the temperature increase over the sea and this suggests that the sign of η should be positive, so we used $\eta \sim U(0,3)$. The prior, $\phi \sim U(0,1000)$, implies a restriction of the maximum correlation between neighboring points to 0.9993.

2. MCMC:

If there are highly correlated parameters, it is recommended that each pair of correlated parameters be sampled as a block.

Table 1. Estimates of hyper-parameters for JJA mean temperatures data.

	Prior Used	Estimate	95% Credible Set
μ_{00}	U(−90, 90)	54.47	(52.15, 57.06)
μ_{01}	U(−90, 90)	−59.53	(−64.01, −55.71)
μ_{02}	U(−90, 90)	−0.17	(−0.30, −0.07)
μ_{03}	U(−90, 90)	−0.22	(−0.26, −0.19)
μ_{20}	U(−90, 90)	21.36	(16.82, 25.33)
μ_{21}	U(−90, 90)	−43.63	(−52.05, −34.09)
μ_{22}	U(−90, 90)	34.38	(27.33, 40.68)
μ_{23}	U(−90, 90)	−12.88	(−15.18, −10.33)
γ	U(−5, 5)	0.043	(0.042, 0.044)
η	U(0, 3)	0.093	(0.079, 0.105)
ϕ_μ	U(0, 1000)	153.46	(144.83, 163.08)
σ_μ^2	U(0, 1250)	10.33	(8.22, 13.00)
ϕ_α	U(0, 1000)	56.06	(54.45, 57.55)
σ_α^2	U(0, 1250)	0.26	(0.24, 0.28)
ϕ_β	U(0, 1000)	70.75	(66.77, 74.59)
σ_β^2	U(0, 1250)	0.0037	(0.0030, 0.0044)
ϕ_{int}	U(0, 1000)	21.54	(20.07, 23.21)
σ_{int}^2	U(0, 1250)	0.00012	(0.00011, 0.00014)
$\phi_\in$	U(0, 1000)	82.01	(81.75, 82.24)
$\sigma_\in^2$	U(0, 1250)	0.178	(0.176, 0.180)

Table 2. Estimates of hyper-parameters for DJF mean temperatures data. Omitted estimate for μ_β which is a vector.

	Prior Used	Estimate	95% Credible Set
μ_{00}	U(−90, 90)	49.42	(46.73, 52.01)
μ_{01}	U(−90, 90)	−81.99	(−86.66, −77.16)
μ_{02}	U(−90, 90)	−1.41	(−1.54, −1.29)
μ_{03}	U(−90, 90)	−0.28	(−0.36, −0.21)
γ	U(−5 ,5)	0.062	(0.061, 0.064)
η	U(0, 3)	0.091	(0.067, 0.115)
ϕ_μ	U(0, 1000)	97.83	(94.62, 101.55)
σ^2_μ	U(0, 1250)	9.37	(8.29, 10.78)
ϕ_α	U(0, 1000)	49.88	(48.72, 51.21)
σ^2_α	U(0, 1250)	1.08	(1.01, 1.17)
ϕ_β	U(0, 1000)	2.86	(0.21, 6.22)
σ^2_β	U(0, 1250)	1.027	(0.992, 1.062)
ϕ_{int}	U(0, 1000)	2.48	(0.11, 5.84)
σ^2_{int}	U(0 1250)	8.8e-05	(7.8e-5 ,9.9e-5)
ϕ_ϵ	U(0, 1000)	55.14	(55.00, 55.29)
σ^2_ϵ	U(0, 1250)	0.625	(0.619, 0.630)

We paired the mean parameters of μ, the mean parameters of β, and each (σ^2, ϕ). We can easily compute the normal posterior distributions for the parameters except for the covariance parameters. We used a Gibbs sampler to sample from the normal conditional distributions for the mean parameters μ, α, β, ($\alpha\beta$), and γ. For each pair (σ^2, ϕ), we first sampled ϕ from its distribution conditional on everything except σ^2, using Metropolis-Hastings steps with Gaussian proposal distributions. We then sampled each σ^2 with the sampled value of *hi* using an inverse gamma conditional distribution.

For each iteration, we randomly chose the order of parameters to be updated, as Roberts and Sahu (1997) suggested. Additionally, during the first 2000 iterations, we updated proposal variances for the Metropolis-Hastings steps in every 100 iteration. We carried out 5000 iterations, discarded the first 500 iterations for burn-in, and used every second sample. We checked convergence visually with ACF and PACF plots (not shown). Tables 1 and 2 show posterior estimates of the hyper-parameters for JJA and DJF, respectively.

3. Covariance approximation for a large data set

There are a large number of pixels even in the limited spatial domain, and computation for statistical estimation using the Bayesian method as described above may be challenging. The total number of the grid points on the domain is 6161, which involves inverting of matrices of size 6161×6161 numerous times for the Bayesian implementation.

Various computational techniques, including covariance approximation and likelihood approximation, have recently been developed in the statistics literature (Stein et al., 2004; Furrer et al., 2006; Kaufman et al., 2008). Of these, we use the *full scale approximation* developed in Sang et al. (2011) and Sang and Huang (2012).

Full scale approximation consists of two parts: reduced rank approximation and sparse approximation of the residual (the difference between the original covariance matrix and the reduced rank approximation). For the reduced rank approximation, we used 375 ($m = 375$) evenly distributed points for knots on the domain. For the sparse approximation of the residual part, we divided the entire domain into 6 disjoint subdomains and assumed that there is no spatial dependence across different subdomains. As a result, the sparse approximation produced a block diagonal matrix with 6 non-diagonal blocks. We chose 6 non-overlapping subdomains arbitrarily with the *i*th subdomain consisting of n_i grid points: $n_1 = 1054$, $n_2 = 1020$, $n_3 = 1023$, $n_4 = 990$, $n_5 = 1054$, and $n_6 = 1020$. See Sang and Huang (2012) for full details on the full scale approximation method.

References

Baek, H.-J., and Coauthors, 2013: Climate change in the 21st century simulated by hadgen2ao under representative concentration pathways. *Asia-Pac. J. Atmos. Sci.,* **49**, 603-618.

Christensen, J. H., and Coauthors, 2013: Climate phenomena and their relevance for future regional climate change. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change,* T. S. Stocker et al. Eds., Cambridge University Press, 1217-1308.

Davies, T., M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood, 2005: A new dynamical core for the met office's global and regional modeling of the atmosphere. *Quart. J. Roy. Meteorol. Soc*., **131**, 1759-1782.

Furrer, R., M. Genton, and D. Nychka, 2006: Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Stat.*, **15**, 502-523.

Giorgi, F., C. Jones, and A. G. R., 2009: Addressing climate information needs at the regional level: the cordex framework. *WMO Bull.*, **58**, 175-183.

_____, and Coauthors, 2012: Regcm4: model description and preliminary tests over multiple CORDEX domains. *Clim. Res*., **52**, 7-29.

Greasby, T. A., and S. R. Sain, 2011: Multivariate spatial analysis of climate change projections. *J. Agr. Biol. Environ. Stat.*, **16**, 571-585.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc*., **90**, 1095-1107.

Kang, E., N. Cressie, and S. Sain, 2011: Combining outputs from the NARCCAP regional climate models using a Bayesian hierarchical model. *Appl. Stat.*, **61**, 291-313.

Kaufman, C., and S. R. Sain, 2010: Bayesian functional ANOVA modeling using Gaussian process prior distributions. *Bayesian Anal.*, **5**, 123-150.

_____, M. Schervish, and D. Nychka, 2008: Covariance tapering for likelihood-based estimation in large spatial data sets. *J. Am. Stat. Assoc.*, **103**, 1545-1555.

Min, S.-K., E.-H. Park, and W.-T. Kwon, 2004: Future projections of East Asian climate change from multi-AOGCM ensembles of IPCC SRES scenario simulations. *J. Meteor. Soc. Japan*, **82**, 1187-1211.

_____, and Coauthors, 2015: Changes in weather and climate extremes over Korea and possible causes: A review. *Asia-Pac. J. Atmos. Sci.*, **51**,

103-121.

Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747-756.

Park, C., and Coauthors, 2016: Evaluation of multiple regional climate models for summer climate extremes over East Asia. *Clim. Dynam.*, **46**, 2469-2486.

Park, J.-H., S.-G. Oh, and M.-S. Suh, 2013: Impacts of boundary conditions on the precipitation simulation of regcm4 in the CORDEX East Asia domain. *J. Geophys. Res.*, **118**, 1652-1667.

Roberts, G., and S. Sahu, 1997: Updating schemes, correlation structure, blocking and parameterization for Gibbs sampler. *J. Roy. Stat. Soc.*, **59**, 291-317.

Sain, S. R., D. Nychka, and L. Mearns, 2011: Functional ANOVA and regional climate experiments: a statistical analysis of dynamic downscaling. *Environmetrics*, **22**, 700-711.

Salazar, E., B. Sanò, A. O. Finley, D. Hammerling, I. Steinsland, X. Wang, and P. Delamater, 2011: Comparing and blending regional climate model predictions for the American southwest. *J. Agr. Biol. Environ. Stat.*, **16**, 586-605.

Sang, H., and J. Z. Huang, 2012: A full-scale approximation of covariance functions for large spatial data sets. *J. Roy. Stat. Soc. Series B*, **74**, 111-132.

_____, M. Jun, and J. Z. Huang, 2011: Covariance approximation for large multivariate spatial datasets with an application to multiple climate model errors. *Ann. Appl. Stat.*, **5**, 2519-2548.

Stein, M. L., 1999: *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 247 pp.

_____, Z. Chi, and L. Welty, 2004: Approximating likelihoods for large spatial data sets. *J. Roy. Stat. Soc. Series B*, 275-296.

Suh, M.-S., and S.-G. Oh, 2012: Development of new ensemble methods based on the performance skills of regional climate models over South Korea. *J. Climate*, **25**, 7067-7082.

Yasutomi, N., A. Hamada, and A. Yatagai, 2011: Development of a long-term daily gridded temperature dataset and its application to rain/snow discrimination of daily precipitation. *Global Environ. Res.*, **15**, 165-172.

Yatagai, A., K. Kaminguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 2012: Aphrodite: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull. Am. Meteorol. Soc.*, **93**, 1401-1415.