ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 30, NO. 4, 2013, 1129–1142

Probabilistic Multimodel Ensemble Prediction of Decadal Variability of East Asian Surface Air Temperature Based on IPCC-AR5 Near-term Climate Simulations

WANG Jia*[1,2] (王 佳), ZHI Xiefei[1] (智协飞), and CHEN Yuwen[2] (陈钰文)

[1] *Key Laboratory of Meteorological Disaster of Ministry of Education, Nanjing University of Information Science and Technology, Nanjing* 210044

[2] *Climate Center of Jiangsu Province, Nanjing* 210008

(Received 26 July 2012; revised 11 December 2012; accepted 17 December 2012)

ABSTRACT

Based on near-term climate simulations for IPCC-AR5 (The Fifth Assessment Report), probabilistic multimodel ensemble prediction (PMME) of decadal variability of surface air temperature in East Asia (20°–50°N, 100°–145°E) was conducted using the multivariate Gaussian ensemble kernel dressing (GED) methodology. The ensemble system exhibited high performance in hindcasting the decadal (1981–2010) mean and trend of temperature anomalies with respect to 1961–90, with a RPS of 0.94 and 0.88 respectively. The interpretation of PMME for future decades (2006–35) over East Asia was made on the basis of the bivariate probability density of the mean and trend. The results showed that, under the RCP4.5 (Representative Concentration Pathway 4.5 W m^{-2}) scenario, the annual mean temperature increases on average by about 1.1–1.2 K and the temperature trend reaches 0.6–0.7 K (30 yr)$^{-1}$. The pattern for both quantities was found to be that the temperature increase will be less intense in the south. While the temperature increase in terms of the 30-yr mean was found to be virtually certain, the results for the 30-yr trend showed an almost 25% chance of a negative value. This indicated that, using a multimodel ensemble system, even if a longer-term warming exists for 2006–35 over East Asia, the trend for temperature may produce a negative value. Temperature was found to be more affected by seasonal variability, with the increase in temperature over East Asia more intense in autumn (mainly), faster in summer to the west of 115°E, and faster still in autumn to the east of 115°E.

Key words: decadal climate prediction, PMME, GED, surface air temperature, East Asia

Citation: Wang, J., X. F. Zhi, and Y. W. Chen, 2013: Probabilistic multimodel ensemble prediction of decadal variability of East Asian surface air temperature based on IPCC-AR5 near-term climate simulations. *Adv. Atmos. Sci.*, **30**(4), 1129–1142, doi: 10.1007/s00376-012-2182-9.

1. Introduction

Climate change research today no longer focuses merely on global trends over hundreds of years. Instead, decadal (10–30 years) climate prediction is becoming a new priority (Hurrell et al., 2009; Latif et al., 2009; Meehl et al., 2009). In order to meet the actual needs of sustainable economic and social development, near-term decadal changes over regions of interest have gained more attention from both governments and society.

Prediction by fully coupled climate models is the most promising method for decadal climate prediction. An experiment testing the decadal hindcast/prediction of coupled models has recently been included in phase five of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al., 2009) and conducted by the main international institutions for climate modeling (Smith et al., 2007; Sugiura et al., 2009). Yoshimitsu et al. (2012) examined the predictability of major climate variability on decadal time scales in terms of the Atlantic Multidecadal Oscillation (AMO) and Pacific Decadal Oscillation (PDO) using the Model for Interdisciplinary Research on Climate (MIROC).

*Corresponding author: WANG Jia, nanjingwangjia2008@yahoo.com.cn

Wu and Zhou (2011) used the Flexible Global Ocean-Atmosphere-Land System model (FGOALS_gl), developed by the State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics (IAP) , to carry out predictions of SST decadal variability. Meehl et al. (2010) conducted decadal predictions of SST in the Pacific using version three of the Community Climate System Model (CCSM3). Du et al. (2012) employed a global atmosphere–ocean model to investigate the sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. Results from many experiments have shown that the dominant error sources in decadal model predictions are the individual model formulations and errors relating to the initial conditions, with the latter resulting especially from the lack of uniformity between schemes of model initialization. For example, the Max Planck Institute for Meteorology in Hamburg (MPI) applied the SST nudging approach (Keenlyside et al., 2008) and the MIROC adopted the Incremental Analysis Update (IAU) scheme (Mochizuki et al., 2010). Therefore, how to reduce uncertainty in model predictions has become a key problem in terms of decadal climate prediction.

Ensemble prediction technology has been developed to help address the problem of uncertainty in climate model prediction. Probabilistic ensemble prediction in particular is useful because it can estimate uncertainty quantitatively using probabilities, thus providing greater value than deterministic predictions in climate change risk analysis and decision-making (Thompson, 1962; Murphy, 1973; Krzysztofowicz, 1983; Min et al., 2009). To date, single-model probabilistic ensemble prediction has been widely applied in climate projection in order to reduce the projection errors caused by initial condition errors. However, a multimodel ensemble technique has also been applied recently to reduce the systematic and random errors from model formulation biases (Fritsch et al., 2000; Stephenson and Doblas-Reyes, 2000; Kharin and Zwiers, 2002; Peng et al., 2002; Palmer et al., 2004; Greene et al., 2006; Räisänen and Ruokolainen, 2006; Zhi et al., 2010; Furtado et al., 2011). Doblas-Reyes et al. (2000) studied the skill of three AGCMs in climate prediction, and found that the multimodel approach provided a systematic improvement in prediction. Tippett and Barnston (2008) discussed the skill of multimodel ENSO probability prediction based on forecast data from seven individual models of the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) project, and pointed out that multimodel ensemble predictions generally have higher skill than single model predictions. Min et al. (2009) constructed an operational seasonal forecast system using a probabilistic multimodel ensemble prediction (PMME) system at the Asia–Pacific Economic Cooperation (APEC) Climate Center (APCC). Zhi et al. (2010) examined the hindcasting skills of eight individual models for the IPCC AR4 (Fourth Assessment Report) scenario runs and found that the multimodel superensemble hindcasting skill for surface temperature was higher than that of the ensemble mean, as well as individual models, and performed a multimodel superensemble prediction of surface temperature under the SRES (Special Report on Emissions Scenarios) A1B scenario for the period 2010–30. Hoerling et al. (2011) diagnosed the predictability of North American decadal climate for the period 2011–20 in terms of surface air temperature and precipitation. However, decadal climate prediction is very much in the early stages of development. There have been very few studies in which multimodel ensemble prediction has been used on the decadal scale. Furthermore, the use of PMME for decadal climate prediction, especially in terms of using it to obtain and interpret the characteristics of decadal climate change, is a very new research field (Frame et al., 2007).

In this paper, we report on work in which up-to-date predictions of surface air temperature for the next 30 years from eight global climate system models for IPCC-AR5 (Meehl et al., 2009) were used to predict the decadal variability of surface air temperature over East Asia during the period 2006–35 based on Gaussian Kernel Ensemble Dressing (GED). The purpose of the study was not only to apply the GED method for grasping the spatiotemporal dependence of ensemble members by multivariate extension, but also for extracting the mean value and trend of surface air temperature anomalies in future decades by linear transformation and the interpretation of spatiotemporal characteristics in both quantities.

2. Data

2.1 *Data source*

The study was based on up-to-date near-term (30-yr) hindcasts/predictions of monthly mean surface air temperature from eight global climate system models for IPCC-AR5. The eight models were taken from China, Japan, Germany, France and England. Every model was initialized with climate system observation data near the end of 1960, 1980 and 2005 according to at least three schemes. Altogether, there were 45 results from eight models with three integrations of 1961–90, 1981–2010 and 2006–35. The external forcing factors, such as atmospheric composition (includ-

Table 1. Basic information of the eight models used in the study.

Model	Organization	Country	Resolution	Number of initialization scheme
Bcc-csm1.1	Beijing Climate Center China Meteorological Administration	China	$\sim 2.8^\circ \times 2.8^\circ$	4
MIROC5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, Japan Agency for Marine-Earth Science and Technology	Japan	$\sim 1.4^\circ \times 1.4^\circ$	6
MPI-ESM-LR	Max Planck Institute for Meteorology	Germany	$\sim 1.875^\circ \times 1.875^\circ$	3
MRI-CGCM3	Meteorological Research Institute (Tokyo)	Japan	$\sim 1.12^\circ \times 1.12^\circ$	3
MIROC4h	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, Japan Agency for Marine-Earth Science and Technology	Japan	$\sim 0.56^\circ \times 0.56^\circ$	3
IPSL-CM5A-LR	Institute Pierre-Simon Laplace	France	$\sim 3.75^\circ$(Lon)$\times 1.9^\circ$(Lat)	6
HadCM3	Met Office Hadley Centre	England	$\sim 3.75^\circ$(Lon)$\times 2.5^\circ$(Lat)	10
CNRM-CM5	Centre National de Recherches Meteorologiques /Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	France	$\sim 1.4^\circ \times 1.4^\circ$	10

ing CO_2) and solar forcing, were considered in three integrations. For past decades (1961–2010) the forcing used observed values. For future decades, the RCP4.5 (Representative Concentration Pathway 4.5 W m^{-2}) scenario was used. Table 1 shows the basic information of the models used in the study.

To estimate the skill of PMME, NCEP (National Centers for Environmental Prediction)–NCAR (National Center for Atmospheric Research) $2.5^\circ \times 2.5^\circ$ reanalysis data of monthly mean air temperature at 2 m were used as observed values of surface air temperature.

2.2 *Data preprocessing*

The preprocessing of the data was divided into three steps, as follows:

(1) Monthly mean surface air temperature grid data for East Asia (20°–50°N, 100°–145°E) were extracted from model predictions and actual observations.

(2) Annual and seasonal means of surface air temperature from the monthly mean data were computed. Then, the full field of annual and seasonal means was handled in the anomaly field with respect to the 30-yr average of the annual mean temperature for the period 1961–90. Finally, the anomaly field was interpolated into the $2.5^\circ \times 2.5^\circ$ grid.

(3) With the surface air temperature anomaly field data established, the annual series for every grid was constructed. The time series was also divided into three parts, representing the training, reference and prediction periods: 1961–90, 1981–2010, and 2006–35, respectively.

2.3 *Uncertainty of data after processing*

According to Hawkins and Sutton (2009), within a 30-yr period the main sources of prediction uncertainty for surface air temperature in climate models are the individual model formulations and initial condition errors, with the former of the two being the source of maximum error. Figure 1 (left) shows the mean RMSE of annual mean temperature over East Asia based on the eight models' hindcasts for the periods 1961–90 and 1981–2010, which reached 1.5–2.9 K and 1.5–3.1 K, respectively. The formulation for RMSE is

$$T_{\mathrm{RMSE}} = \left[\frac{1}{N} \sum_{i=1}^{N} (F_i - O_i)^2 \right]^{1/2}, \qquad (1)$$

where F_i is the value of prediction and O_i is the observed value, N is the number of sample forecast and observed values.

After data preprocessing, the prediction of decadal variability of surface air temperature was transformed into the prediction of temperature anomalies, so as to avoid the considerable model errors resulting from "drift". Figure 1 (right) shows the average RMSE of

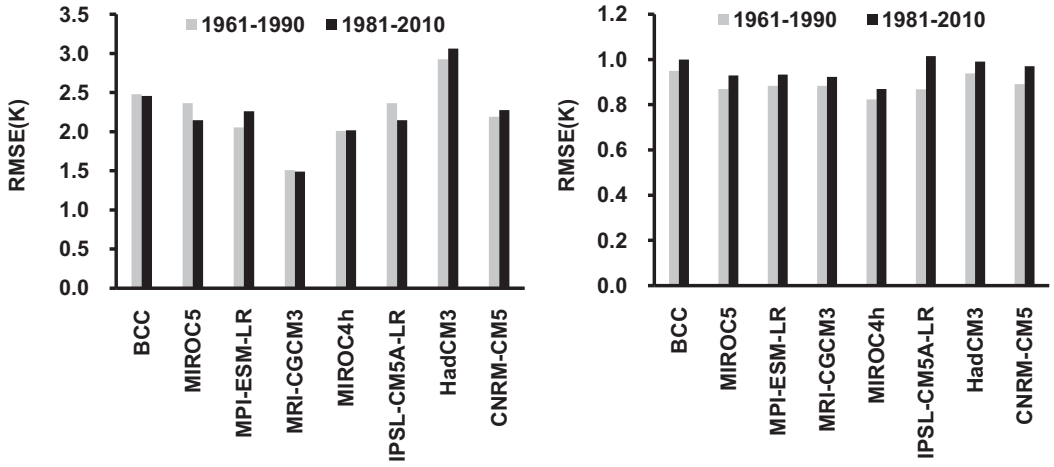

Fig. 1. The mean RMSE (K) of annual mean temperature (left) and its anomalies (right) with respect to 1961–90 from the eight models' hindcasts during the periods 1961–90 and 1981–2010 over East Asia (20°–50°N, 100°–140°E).

temperature anomalies over East Asia with respect to the period 1961–90 based on the eight models' hindcasts for the periods 1961–90 and 1981–2010. As can be seen, this reduced the results to 0.8–1.0 K and 0.9–1.0 K, respectively. However, the uncertainty from initialization and stochastic model errors still existed in the prediction of temperature anomaly decadal variability. Therefore, we used the 45 results from the different models and initial conditions, as described in section 2.1, as ensemble members to carry out PMME, so as to reduce the prediction model errors and, in particular, to estimate the uncertainty of prediction.

3. Methods

The GED method was used to perform the PMME of surface air temperature decadal variability over East Asia based on temperature anomalies from predictions by different models and initial conditions. The method treats prediction of annual series of temperature anomalies in subsequent decadal years as random events based on prediction results, in this case from eight individual models. The random events were represented by a random vector $\boldsymbol{X}$ of dimension m, which was the number of the ensemble size. The entry x_m of this vector was the n-year series of surface air temperature anomaly from each ensemble member. The probabilistic information of vector $\boldsymbol{X}$ could be derived by the m-dimension joint probability density for vector $\boldsymbol{X}$. Firstly, the probability density was determined by estimation of temporal autocovariance of the column vector $\boldsymbol{X}$. Then, the marginal probability density distribution of the 30-yr mean and trend of temperature anomalies (hereafter referred to simply as 'mean' and 'trend') was extracted using linear transformation. Finally, the spatiotemporal characteristics of surface air temperature decadal variability were analyzed comprehensively.

3.1 *Joint probability density of PMME*

The determination of joint probability density $f_x(\boldsymbol{x}, \theta)$ of the random vector $\boldsymbol{X}$ is the key to PMME. Here, θ represents the characteristic parameters of the probability density f. Firstly, the probability density for a deterministic prediction is Dirac's delta function at the realization $\boldsymbol{x}_1$:

$$f_x(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_1) \, . \tag{2}$$

δ represents Dirac's delta function.

Similarly, the probability density for a set of model predictions is equivalent to the sum of Dirac delta functions at several distinguishable realizations $\boldsymbol{x}_i, i = 1, \ldots, m$:

$$f_x(\boldsymbol{x}) = \sum_{i=1}^{m} \delta(\boldsymbol{x} - \boldsymbol{x}_i) \, . \tag{3}$$

However, the ensemble predictions are indistinguishable, especially for single-model ensemble predictions from different initial conditions. So, we needed to introduce the internal noise component $\boldsymbol{\varepsilon}_i$ as additive noise to the "true" prediction $\boldsymbol{x}$:

$$\boldsymbol{x} = \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i \, , \tag{4}$$

where $\boldsymbol{\varepsilon}_i$ is the component of noise vector $\boldsymbol{\varepsilon}$ with expectation 0. The density function of vector $\boldsymbol{\varepsilon}$ is defined as $f_\varepsilon(\boldsymbol{\varepsilon})$. Correspondingly, the probability density $f_x(\boldsymbol{x})$ of ensemble predictions is derived as the noise expectation of the average of delta functions containing noises

(Honerkamp, 2002):

$$\begin{aligned} f_x(\boldsymbol{x}) =& E_\varepsilon \left[\frac{1}{m} \sum_{i=1}^{m} \delta(\boldsymbol{x} - \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i) \right] \\ =& \int \frac{1}{m} \sum_{i=1}^{m} \delta(\boldsymbol{x} - \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i) f_\varepsilon(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon} \\ =& \frac{1}{m} \sum_{i=1}^{m} f_\varepsilon(\boldsymbol{x} - \boldsymbol{x}_i) \, . \end{aligned} \tag{5}$$

The above formula derivation shows that the $f_x(\boldsymbol{x})$ of ensemble prediction is the average of the noise density at indistinguishable predictions. Namely, dressing with the noise density $f_\varepsilon(\boldsymbol{\varepsilon})$, the probabilistic information between the ensemble prediction $\boldsymbol{x}_i$ is interpolated to the point $\boldsymbol{x}$. Equation 5 also shows that there is strong connection between the determination of $f_x(\boldsymbol{x})$ and kernel density estimation. In the present work, we mainly investigated annual and seasonal surface air temperature anomalies, so the distribution of the noise vector $\boldsymbol{\varepsilon}$ was assumed to be a multivariate Gaussian distribution with covariance matrix $\sum_D$. Correspondingly, the joint probability density of surface air temperature anomaly PMME was the multivariate Gaussian ensemble kernel dressing:

$$\begin{aligned} & f_x(\boldsymbol{x}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) \\ =& \frac{1}{m} \sum_{i=1}^{m} \frac{1}{2\pi^{m/2} |\sum_D|^{1/2}} \cdot \\ & \exp\left[-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{x}_i)^{\mathrm{T}} \sum_D^{-1} (\boldsymbol{x} - \boldsymbol{x}_i) \right] . \end{aligned} \tag{6}$$

The uncertainty of ensemble prediction is reflected by the temporal autocovariance matrix $\sum_D$.

However, as the multimodel ensemble hindcasts/predictions were distinguishable, we needed to comprise the weight ω_i characterizing the probability of $\boldsymbol{x}_i$ as the "best" prediction. Consequently, the joint probability density of surface air temperature PMME contains the weight:

$$\begin{aligned} & f_x(\boldsymbol{x}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) \\ =& \omega_i \sum_{i=1}^{m} \frac{1}{2\pi^{m/2} |\sum_D|^{1/2}} \cdot \\ & \exp\left[-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{x}_i)^{\mathrm{T}} \sum_D^{-1} (\boldsymbol{x} - \boldsymbol{x}_i) \right] . \end{aligned} \tag{7}$$

3.2 *Parameter estimation of joint probability density*

Parameter estimations including noise/dressing covariance and weights ω_i are the main ways to determine joint probability density of surface air temperature PMME.

3.2.1 *Dressing covariance*

The method of multivariate kernel density estimation (Silverman, 1986) was adopted to estimate the dressing covariance $\sum_D$. The results showed that there is a proportional relationship between $\sum_D$ and estimated raw error covariance $\sum_{\text{raw}}$: $\sum_D = h_{\text{opt}} \sum_{\text{raw}}$. The proportional coefficient h_{opt} is defined as (Silverman, 1986)

$$h_{\text{opt}} = \left[\frac{4}{m(n+2)} \right]^{\frac{1}{n+4}} , \tag{8}$$

where m is the number of ensemble members and n is the time dimension of entry $\boldsymbol{x}_m$.

In order to estimate $\sum_{\text{raw}}$, the distance vector of two arbitrary entries $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is defined as $\boldsymbol{d}_{ij} = 1/\sqrt{2}(\boldsymbol{x}_i - \boldsymbol{x}_j)$. The expectation of the square distance vector module is derived with the independent noise vector $\boldsymbol{\varepsilon}$. The derivation is expressed as follows:

$$\begin{aligned} E_\varepsilon \left(\boldsymbol{d}_{ij} \boldsymbol{d}_{ij}^{\mathrm{T}} \right) &= \frac{1}{2} E_\varepsilon \left(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^{\mathrm{T}} + \boldsymbol{\varepsilon}_j \boldsymbol{\varepsilon}_j^{\mathrm{T}} - \boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_j^{\mathrm{T}} - \boldsymbol{\varepsilon}_j \boldsymbol{\varepsilon}_i^{\mathrm{T}} \right) \\ &= \frac{1}{2} E_\varepsilon \left(\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^{\mathrm{T}} \right) + \frac{1}{2} E_\varepsilon \left(\boldsymbol{\varepsilon}_j \boldsymbol{\varepsilon}_j^{\mathrm{T}} \right) \\ &= \sum_{\text{raw}} . \end{aligned} \tag{9}$$

We find that $\sum_{\text{raw}}$ is equal to the average of the differences of all possible pairs of individual simulations

$$\sum_{\text{raw}} = \frac{1}{2N_{\text{tot}}} \sum_{i,j=1, i>j}^{m} (\boldsymbol{x}_i - \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j)^{\mathrm{T}} , \tag{10}$$

where N_{tot} is the pair number in the collection $\{(\boldsymbol{x}_i, \boldsymbol{x}_j) | i, j = 1, \ldots, m, i > j\}$.

Supposing the dressing covariance is stationary in time, $\sum_{\text{raw}}$ converges to a Toeplitz structure.

$$\sum_{\text{raw}} = \begin{pmatrix} \sigma_\varepsilon(0) & \sigma_\varepsilon(-1) & \cdots & \sigma_\varepsilon(-n+1) \\ \sigma_\varepsilon(1) & \sigma_\varepsilon(0) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \sigma_\varepsilon(-1) \\ \sigma_\varepsilon(n-1) & \cdots & \sigma_\varepsilon(-1) & \sigma_\varepsilon(0) \end{pmatrix} , \tag{11}$$

where the function σ_ε is

$$\begin{aligned} \sigma_\varepsilon = & \sum_{t=|\tau|+1}^{n} \sum_{i,j=1, i<j}^{m} \times \\ & \frac{(x_{i,\,t} - x_{j,\,t})(x_{i,\,(t-|\tau|)} - x_{j,\,(t-|\tau|)})}{2nN_{\text{tot}}} . \end{aligned} \tag{12}$$

trepresents the time dimension.

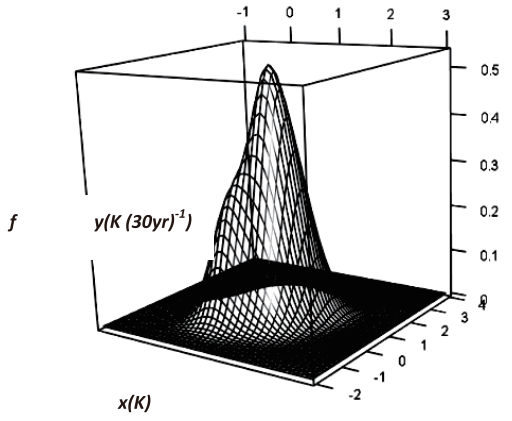

Fig. 2. Example of a bivariate probability density function of a random vector $\boldsymbol{X}$ (f is the probability density function; x and y are the mean and trend of the surface air temperature anomaly, respectively).

Finally, $\sum_{\text{raw}}$ could be derived from the training period 1961–90.

3.2.2 *Weight*

The weights ω_i were estimated at each grid point by the expectation maximization (EM) algorithm (Raftery et al., 2005), a method that takes observations into account. We applied this method to the training period 1961–90 with the predictions from the different models and observations to determine ω_i. For details of the EM algorithm, please refer to Raftery et al. (2005).

3.3 *Linear transformation*

$\sum_{\text{raw}}$ and ω_i derived from the training period 1961–90 were respectively applied in the construction of the joint probability density for the periods 1981–2010 and 2006–35, with the corresponding multimodel ensemble hindcasts/predictions. In order to derive the spatiotemporal characteristics of surface air temperature decadal variability, we took two steps into consideration. Firstly, the spatial information was disregarded. We concentrated on the individual grids from multimodel ensemble predictions. Then, the random vector $\boldsymbol{X}$ only contained the temporal information, and $\boldsymbol{x}_i$ was the 30-yr series of surface air temperature anomalies from each ensemble member. In order to extract useful information regarding decadal climate change, we translated the random vector $\boldsymbol{X}$ into a 2D vector $\boldsymbol{Z} = (\boldsymbol{Z}_0, \boldsymbol{Z}_1)^{\mathrm{T}}$ by linear transformation $\boldsymbol{X} \xrightarrow{\boldsymbol{L}} \boldsymbol{Z}$. $\boldsymbol{Z}_0$ and $\boldsymbol{Z}_1$ are, respectively, the 30-yr mean and trend of temperature.

$$\boldsymbol{Z} = \underbrace{(\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}})^{-1}\boldsymbol{P}}_{\boldsymbol{L}}\boldsymbol{X}$$

$$\boldsymbol{P} = \frac{1}{n}\begin{pmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 \\ 1 & 2 & 3 & \cdots & \cdots & n \end{pmatrix}, \tag{13}$$

where n is the number of time series.

Then, the m-dimension joint probability of surface air temperature density was transformed into the 2D joint probability of density:

$$\begin{aligned} & f_z(\boldsymbol{z}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) \\ &= \omega_i \sum_{i=1}^{m} \frac{1}{2\pi^{m/2}|\boldsymbol{L}\sum_D^{-1}\boldsymbol{L}^{\mathrm{T}}|^{1/2}} \cdot \\ & \exp\left[-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{L}\boldsymbol{x}_i)^{\mathrm{T}}\left(\boldsymbol{L}\sum_D^{-1}\boldsymbol{L}^{\mathrm{T}}\right)(\boldsymbol{z}-\boldsymbol{L}\boldsymbol{x}_i)\right] . \end{aligned} \tag{14}$$

By the above formulation, the joint probability density of mean and trend were computed for the periods 1981–2010 and 2006–35, respectively (Fig. 2). The marginal probability density of mean and trend contained characteristics of surface air temperature decadal variability. Finally, according to the grids' marginal probability density, the spatial and seasonal characteristics of surface air temperature decadal variability over East Asia were analyzed in depth using the probability maps.

4. Results

In this section we report the results from applying the above ensemble prediction method in PMME of decadal variability of 30-yr mean and trend of surface air temperature over East Asia based on the temperature anomalies from the eight ocean–atmosphere coupled models' predictions.

4.1 *Evaluation of PMME skill*

Before ensemble prediction, the skill of surface air temperature PMME over East Asia was evaluated by rank probability score (RPS) based on model hindcasts and observed data during the period 1981–2010. The formulation of RPS is as follows:

$$\mathrm{RPS}(\boldsymbol{p},\ \boldsymbol{d}) = 1 - \frac{1}{k-1}\left[\sum_{i=1}^{k}\left(\sum_{n=1}^{i} p_n - \sum_{n=1}^{i} d_n\right)^2\right], \tag{15}$$

where $\boldsymbol{p} : (p_1, p_2, \ldots, p_k)$ and $\boldsymbol{d} : (d_1, d_2, \ldots, d_k)$ are the forecast and observed probabilities of different ranks k, respectively. The member of d set is 1 or 0, and if $d_i (i = 1, 2, \ldots, k)$ is 1, all the other d_j $(j \neq i, j = 1, 2, \ldots, k)$ is 0. The probability prediction of annual mean temperature T_{ANN} was divided into 10 ranks: $T_{\text{ANN}} < -1.5$ K; -1.5 K $\leqslant T_{\text{ANN}} < -1.0$ K;

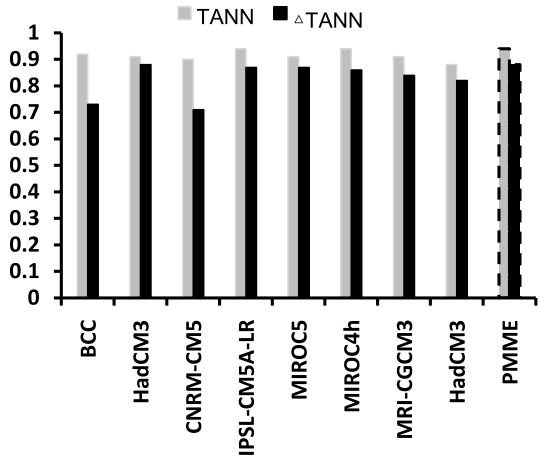

Fig. 3. Ranked probability scores for annual mean temperature ($T_{\rm ANN}$) and its trend ($T_{\rm ANN}$) for the eight models' hindcasts and PMME (dashed frame) during the period 1981–2010.

-1.0 K$\leqslant T_{\rm ANN} < -0.5$ K, -0.5 K$\leqslant T_{\rm ANN} < 0$ K; 0 K$\leqslant T_{\rm ANN} < 0.5$ K; 0.5 K$\leqslant T_{\rm ANN} < 1.0$ K; 1.0 K$\leqslant T_{\rm ANN} < 1.5$ K; 1.5 K$\leqslant T_{\rm ANN} < 2.0$ K; 2.0 K$\leqslant T_{\rm ANN} < 2.5$ K; $T_{\rm ANN} \geqslant 2.5$ K. The probability prediction of trend $\Delta T_{\rm ANN}$ was also divided into 10 ranks: $\Delta T_{\rm ANN} < -1.2$ K (30 yr)$^{-1}$; -1.2 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < -0.9$ K (30 yr)$^{-1}$; -0.9 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < -0.6$ K (30 yr)$^{-1}$; -0.6 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < -0.3$ K (30 yr)$^{-1}$; -0.3 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < 0$ K (30 yr)$^{-1}$; 0 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < 0.3$ K (30 yr)$^{-1}$; 0.3 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < 0.6$ K (30 yr)$^{-1}$; 0.6 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < 0.9$ K (30 yr)$^{-1}$; 0.9 K (30 yr)$^{-1}$ $\leqslant \Delta T_{\rm ANN} < 1.2$ K (30 yr)$^{-1}$; $\Delta T_{\rm ANN} \geqslant 1.2$ K (30 yr)$^{-1}$. Correspondingly, we were able to obtain the average RPS value of mean and trend in different ranks. The average was between 0 and 1. The more accurate the ensemble hindcast, the higher the value of average RPS was. The results showed that PMME performs better at predicting the mean and trend of surface air temperature for the next 30 years based on the results from the eight ocean-atmosphere models than it does based on the individual model (Fig. 3). The RPS of probability prediction of mean temperature was 0.94, and the RPS of the trend was 0.88.

4.2 *Results for a single grid point*

Figure 4 shows the results for the single grid point (45°N, 122.5°E). In the upper row the ensemble mean and spread from the eight individual models' predictions are plotted. The lower panel shows the marginal distribution of the mean Z_0 (left), trend Z_1 (middle), and bivariate distribution $\boldsymbol{Z} = (Z_0, Z_1)$ for the annual mean of surface air temperature respectively in the reference period 1981–2010 and prediction period 2006–35. All values are the anomaly field with respect to the 30-yr average of the annual mean temperature for the period 1961–90.

As seen in the lower panel of Fig. 4 (left), for 1981–2010 the probability density of the mean of 0.3 K was maximum; for 2006–35 the maximum was about 1.4 K with respect to 1961–90, and the likelihood of a positive mean reached 100%. This shows that there was a clear increase in annual mean temperature for the next 30 years of 2006–35 and the magnitude of increase was larger than that for the period 1981–2010.

For the trend over 30 years (Fig. 4, middle), there was a larger variability than the mean both in the period 1981–2010 and 2006–35. This shows that the trend was more affected by natural variability of the climate system and less determined by external forcings than the mean within 30 years. For the period 2006–35, the maximum trend was around 1.1 K (30 yr^{-1}) and around 0.8 K (30 yr^{-1}) for 1981–2010, but we can see that there was an almost 25% chance of obtaining a negative trend in the two periods. Figure 5 shows the probability density distributions of the surface air temperature 30-yr trend in 2006–35 from the eight models at the grid point (45°N, 122.5°E). For the BCC model, the probability density of the trend of -0.4 K (30 yr^{-1}) was the maximum. The chance of a 30-yr trend for negative values reached 76%, 47% and 27% based on model results from BCC (Beijing Climate Center), CNRM-CM5 (Centre National de Recherches Meteorologiques coupled model 5) and HadCM3 (Hadly Centre coupled model 3), respectively (Fig. 5, black line). Furthermore, we checked 45 runs from the eight models and found 10 runs that showed a negative trend (Table 2). Detecting a negative trend was part of the uncertainty from internal fluctuations of the climate system reflected by the ensemble system. This shows that in the presence of longer-term anthropogenic forced warming, the natural variability of the climate system is likely to produce multi-year periods with no positive trend. According to IPCC-AR4, the probability of a future climate event is divided into three categories: >99% (virtually certain), >90% (very likely), and >66% (likely). We found that in the next 30 years of 2006–35 an increase of mean temperature is 'virtually certain', but a trend with a positive value is only 'likely' at this grid point. Owing to little spatial difference in the trend, the situation that a positive trend is only 'likely' can be considered universal over East Asia.

4.3 *Spatial characteristics*

Results for a single grid point showed detailed local probabilistic information for surface air temperature decadal variability, but another main aim of the

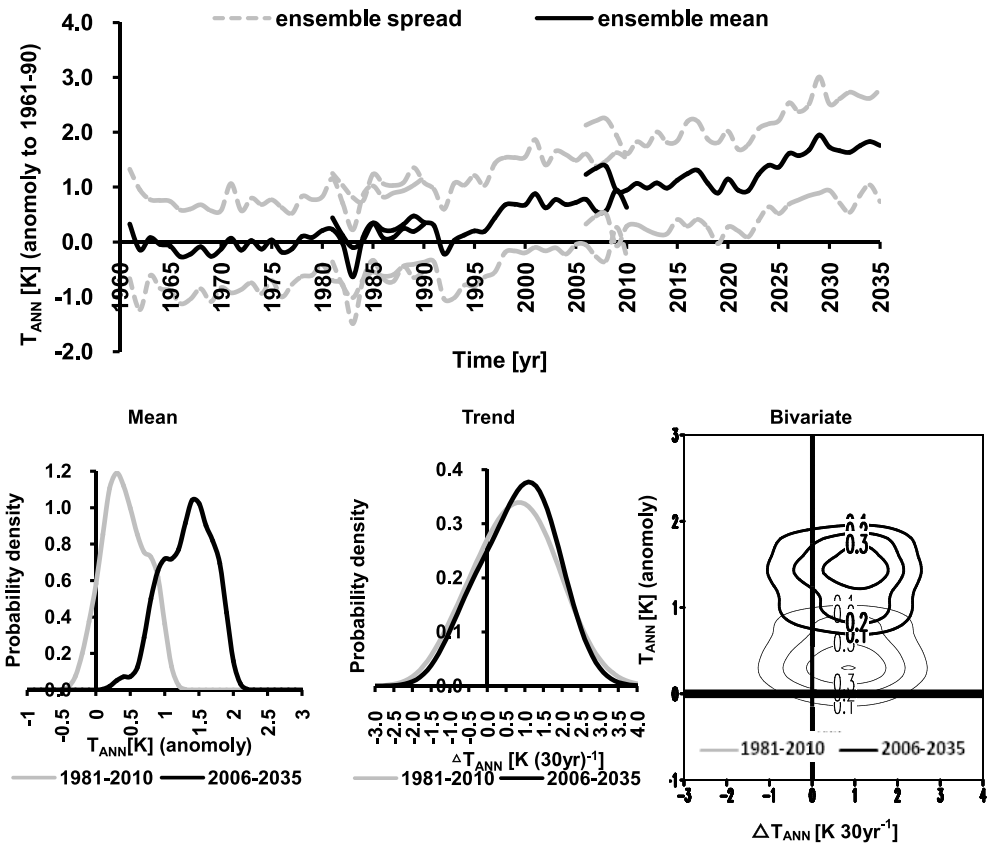

Fig. 4. Results of PMME of surface air temperature at the grid point (45°N, 122.5°E). All values are anomalies with respect to 1961–90. Upper: time series of ensemble mean (black solid line) and ensemble spread (two dashed lines). Bottom: marginal distribution of the mean (left) and trend (middle) in 1981–2010 (gray line) and 2006–35 (black line); corresponding bivariate distribution (right).

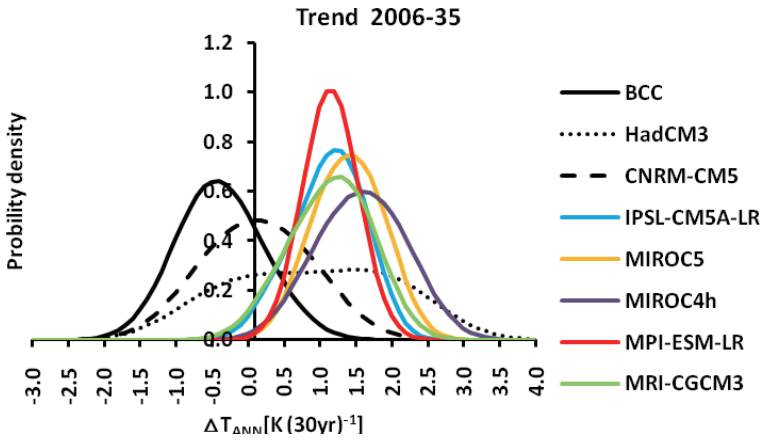

Fig. 5. Marginal distributions of the trend of surface air temperature in 2006–35 from the eight models at the grid point (45°N, 122.5°E).

Table 2. Statistics of negative trends of surface air temperature in 2006–35 from 45 runs of the eight models at grid point (45°N, 122.5°E).

Models	Number of ensemble member with different initialization schemes	Number of negative trend in different ensemble members
BCC	4	3
CNRM-CM5	10	5
HadCM3	10	2
IPSL-CM5A-LR	6	0
MIROC5	6	0
MIROC4h	3	0
MPI-ESM-LR	3	0
MRI-CGCM3	3	0
Total	45	10

study was to investigate the spatial characteristics of decadal temperature changes over East Asia. A probability map was used for this purpose.

The probability maps reflected the spatial patterns of mean temperature and temperature trend, exceeding prescribed values. Figure 6 shows the probability maps for the 30-yr mean in the period 2006–35 for exceeding 1.4 K, 1.3 K, 1.0 K and 0.9 K, as well as the trend for exceeding 0.8 K (30 yr)$^{-1}$, 0.7 K (30 yr)$^{-1}$, 0.6 K (30 yr)$^{-1}$ and 0.5 K (30 yr)$^{-1}$. The thresholds were selected according to an average expectation of 1.1–1.2 K for mean temperature over East Asia and 0.6–0.7 K (30 yr)$^{-1}$ for trend.

In Figs. 6a–d, a north–south pattern can be observed, where in the southern part of East Asia it is less likely that a specific threshold is exceeded than in the northern part. Moreover, in the northern part of East Asia, the probability inland was slightly higher than that in coastal areas and the surrounding sea, and in the southern part of East Asia the probability inland was also a little bit higher than in the western Pacific. From the above analysis we can infer that in the next 30 years the increase in mean temperature will be most significant in northwestern inland areas, and least significant in the western Pacific. Furthermore, the probabilities for $T_{\rm ANN}\geqslant$1.4 K (Fig. 6a) varied between 30% and 60% in northern inland areas, which approximately reflected the expected value of mean temperature. In other words, the surface air temperature in northern inland areas may increase by 1.4 K during the period 2006–35 with respect to the period 1960–90. Meanwhile, the magnitude of mean temperature increase may reach 1.3 K, 1.0 K and 0.9 K respectively in northeastern (Fig. 6b), southwestern (Fig. 6c) and southeastern (Fig. 6d) parts of East Asia.

The trends within the period 2006–35 (Figs. 6e–h) also showed a north–south pattern, but the spatial difference was significantly less than the mean value. Similar to the mean value, the probability of the trend over the western Pacific was the least. However, the most inconsistent feature in terms of the mean was that in the northern part of East Asia the probability around coastal areas and its surrounding sea was higher than that in inland areas. The average expectation of the trend in the northeastern part of East Asia was approximately 0.8 K (30 yr)$^{-1}$. Correspondingly, the probabilities of $T_{\rm ANN}\geqslant$0.8 K (30 yr)$^{-1}$ varied between 50% and 60% (Fig. 6e). Meanwhile, the magnitude of the trend may reach 0.7 K (30 yr)$^{-1}$, 0.6 K (30 yr)$^{-1}$ and 0.5 K (30 yr)$^{-1}$ in northwestern (Fig. 6f), southwestern (Fig. 6g) and southeastern (Fig. 6h) parts of East Asia, respectively.

Furthermore, results given by even-weighted and uneven-weighted PPME (probabilistic multimodel ensemble prediction) were compared. Figure 7 shows the probability maps exceeding 0.8 K (30 yr)$^{-1}$ for the trend of annual mean temperature over East Asia during the period 2006–35 by these two methods. As seen in the upper panel of Fig. 7, the north–south pattern was also demonstrated by even-weighted PPME, but there were still some differences in some areas, such as southwestern and southeastern parts of East Asia. For example, in the north part of Sichuan Province (32.5°N, 105°E, marked by two arrows in the upper panel of Fig. 7), the probability given by even-weighted PPME was slightly higher than by the uneven-weighted method. Therefore, we found that the trend of annual temperature increase around this area for the period 2006–35 was slower according to the uneven-weighted method compared to the even-weighted method (Fig. 7, bottom). Meanwhile, RPS (Fig. 7, column) for hindcasts during 1981–2010 in this area by the uneven-weighted method was higher than by the even-weighted method, which adds confidence to the result given by the uneven-weighted method. Owing to near independence between the mean and trend, the joint probability density of mean and trend can be expressed as the multiple of their marginal probability densities. Therefore, we were able to obtain joint probability information with high anomaly

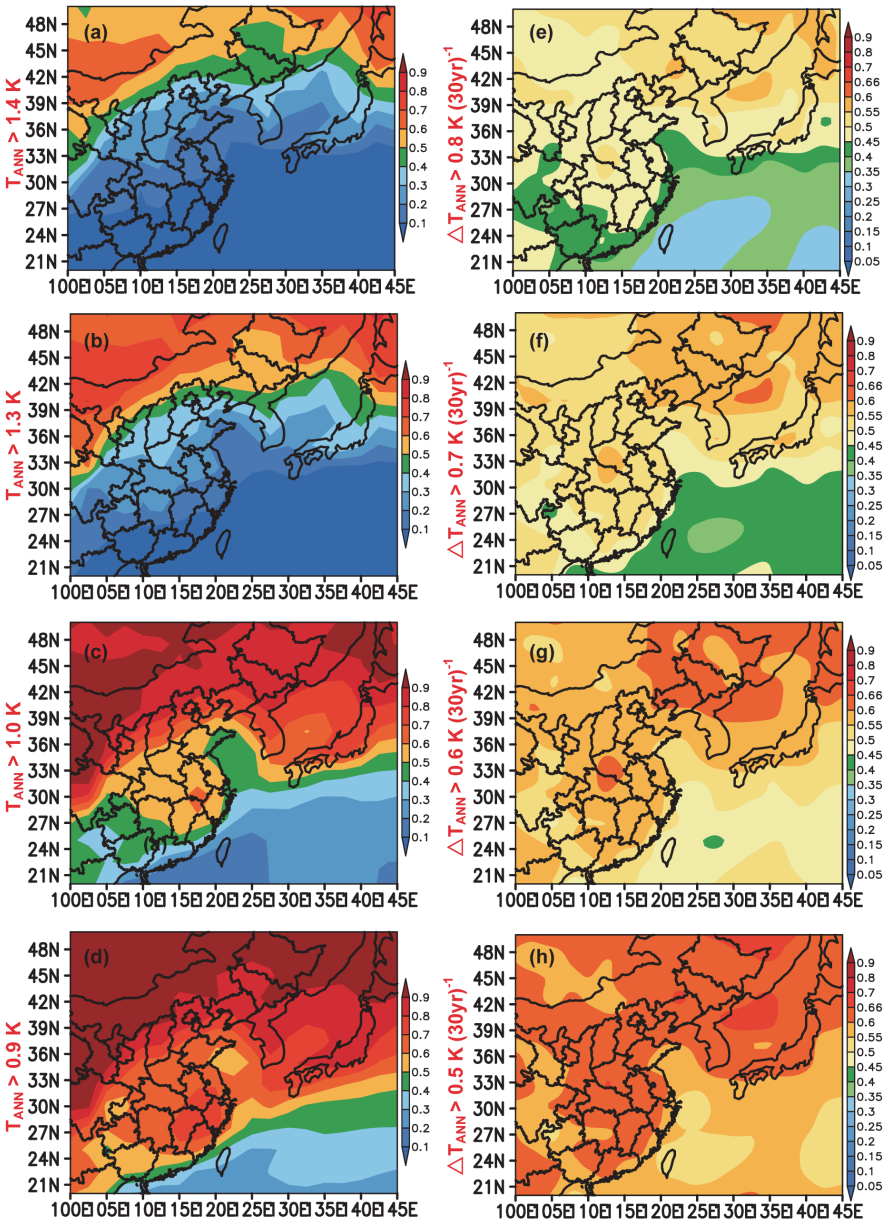

Fig. 6. Probability maps exceeding prescribed values for mean temperature (left column) and trend (right column) of annual mean temperature (ANN) over East Asia during 2006–35 given by the 45 simulations of the ensemble.

(>1.0 K) and high trend [>0.8 K (30 yr)$^{-1}$] according to the probability maps of mean and trend.

4.4 *Seasonality*

The seasonality of surface air temperature decadal variability over East Asia was investigated in terms of seasonal mean temperature anomalies. Figures 8a–d show the probability maps for $T_{\rm ANN}\geqslant$1.3 K in spring, summer, autumn and winter, respectively. The probability map for $T_{\rm ANN}\geqslant$1.3 K in autumn was the highest, especially in the northwestern part of East Asia (Fig. 8c). Except for some regions where the probability for $T_{\rm ANN}\geqslant$1.3 K was highest in winter (Fig. 8d, rectangle), such as east to 135°E and north to 42°N, and around the eastern part of the Tibetan Plateau, the probability for $T_{\rm ANN}\geqslant$1.3 K was the highest one in autumn from northwestern inland areas, east to 135°E and southeast to the West Pacific (Fig. 8c, arrows). Thus it can be seen that the average expectation of mean temperature over East Asia was the highest one in autumn, corresponding to a larger magnitude of temperature increase than in other seasons. Similar

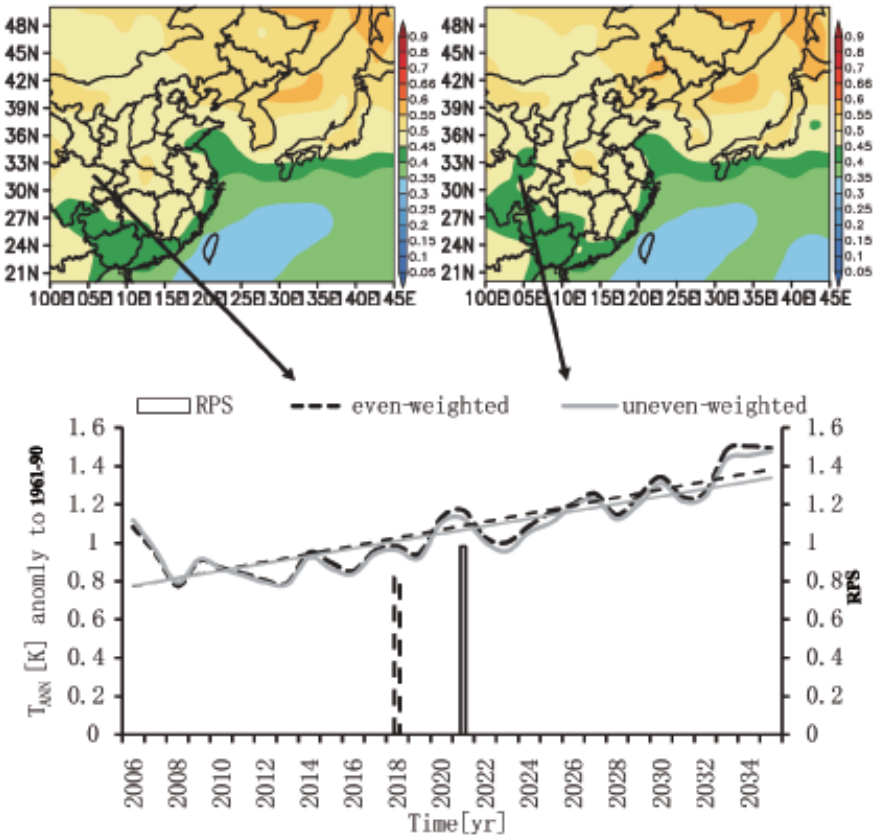

Fig. 7. Probability maps (upper) exceeding 0.8 K $(30\ \mathrm{yr})^{-1}$ for the trend of annual mean temperature (ANN) over East Asia during 2006–35 given by even-weighted (left) and uneven-weighted (right) PPME. Bottom: time series (curve) of ensemble mean and its trend (straight line) at the grid point (32.5°N, 105°E) (marked by two arrows in the panel). The column is the RPS for hindcasts during 1981–2010 by the even-weighted (dashed) and uneven-weighted (solid) PPME.

to the annual mean, the spatial structures of mean temperature in different seasons also exhibited higher values in the north and lower values in the south.

In contrast to the mean value, the trend within the period 2006–35 between different seasons was not so significant. However, the trend still possessed seasonality. As can be seen from Figs. 8e–h, the probabilities of $\Delta T_{\mathrm{ANN}}>0.7$ K $(30\ \mathrm{yr})^{-1}$ in summer and autumn were higher than those in winter and spring. In summer (Fig. 8f, arrows), the probability of $\Delta T_{\mathrm{ANN}}>0.7$ K $(30\ \mathrm{yr})^{-1}$ was the highest one from northwestern inland areas to the southeast of coastal areas, and in autumn (Fig. 8g, arrows) the highest probability was mainly located in the region from the northeastern part of East Asia, west to 115°E and south to the West Pacific. In the west, to 115°E, the expected value of the trend was higher; namely, with a more rapid temperature increase in summer. In the east, to 115°E, the temperature increase was more rapid in autumn. The spatial patterns of the trend in different seasons were similar to those of the annual trend.

5. Concluding remarks

The GED methodology was extended into a multivariate version. The additional dimensions were introduced to capture the temporal autocorrelation of the multimodel ensemble system. The uncertainty of the ensemble system was represented by the temporal autocovariance. According to the autocovariance derived in the training period, an uneven-weighted multivariate Gaussian kernel functions was identified and applied in the prediction of surface air temperature over East Asia for the next 30 years. After handling the full field of temperature predictions from the eight ocean–atmosphere models for IPCC-AR5 in the anomaly field, the parameters of the mixture model were estimated in the training period 1961–90. Then, by use of linear transformation, the PMME of mean and trend surface air temperature in the reference period 1981–2010 and prediction period 2006–35 was conducted by the above mixture model derived in the training period. The following results were obtained

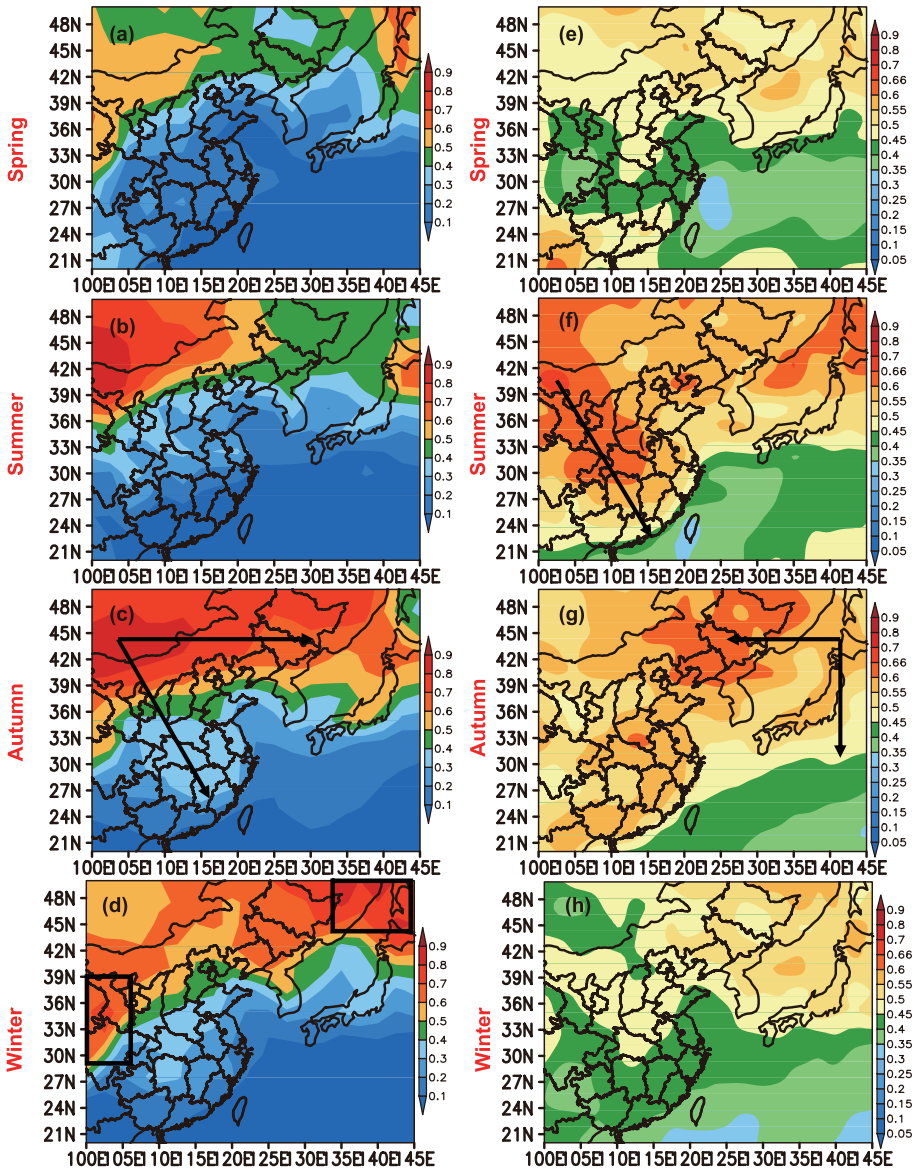

Fig. 8. Probability maps exceeding prescribed values for the temporal mean (left column, >1.3 K) and trend [right column, >0.7 K (30 yr)$^{-1}$] of spring (a, e), summer (b, f), autumn (c, g), winter (d, h) temperature over East Asia during 2006–35 given by the 45 simulations of the ensemble, equivalent to Figs. 6b and f.

under the RCP4.5 scenario:

(1) The analysis of prediction uncertainty showed that the average RMSEs of surface air temperature prediction over East Asia from the eight individual models in the periods 1961–90 and 1981–2010 were 1.5–2.9 K and 1.5–3.1 K, respectively. After anomaly transformation of the temperature was performed, the average RMSEs of temperature anomaly prediction were only 0.8–1.0 K and 0.9–1.0 K, and the range of average RMSE became significantly narrower to avoid model errors resulting from model "drift". However, the initial condition errors and model random errors still existed. It was possible to apply PMME to estimate the uncertainty of model prediction of surface air temperature and reduce prediction errors.

(2) According to the autocovariance and weights derived in the training period 1961–90, the uneven-weight PMME of mean and surface air temperature during the reference period (1981–2010) were reasonable with a RPS of 0.94 and 0.88, respectively.

(3) Through detailed interpretation of the probability density of surface air temperature PMME at the single grid point (45°N, 122.5°E), we found that in the next 30 years the probability of $T_{\mathrm{ANN}} \geqslant 1.4$ K

will be the maximum, the temperature increase is certain, and the magnitude of the temperature increase will be larger than that in the period 1981–2010 at this grid point. However, even if in the next 30 years the probability of $\Delta T_{\rm ANN} \geqslant 1.1$ K is the maximum, the negative trend is unlikely (with some confidence), but not impossible, which demonstrates that as far as the ensemble system shows, even if longer-term warming exists during the period 2006–35 over East Asia, the trend of surface air temperature may produce a negative value in multi-year periods.

(4) The average temperature anomaly over East Asia for the period 2006–35 was found to be approximately 1.1–1.2 K with respect to 1960–90, and its trend about 0.6–0.7 K $(30\ {\rm yr})^{-1}$. Both quantities showed spatial differences. The mean and trend of the temperature anomaly was higher in the north, but became smaller in the south, and the minimum was located in the West Pacific. Inconsistently, the maximum probability of the mean was in northwestern inland areas, but the maximum probability of the trend was located in the coastal area and its surrounding sea in the northeast of East Asia. Through preliminary analysis, the increases of the mean in northwestern, northeastern, southwestern and southeastern parts of East Asia were around 1.4 K, 1.3 K, 1.0 K and 0.9 K, respectively, while the increases of trend were around 0.8 K $(30\ {\rm yr})^{-1}$, 0.7 K $(30\ {\rm yr})^{-1}$, 0.6 K $(30\ {\rm yr})^{-1}$ and 0.5 K $(30\ {\rm yr})^{-1}$, respectively.

(5) The seasonality of surface air temperature decadal variability over East Asia in the period 2006–35 was found to be significantly distinct, except over some regions east to 135°E and north to 42°N, and around the eastern part of the Tibetan Plateau. The magnitude of temperature increase was the highest one in autumn in East Asia. Roughly, in the west to 115°E, the temperature increased more rapidly in summer, but in the east to the 115°E it increased more rapidly in autumn.

In this study, bivariate uneven-weighted GED was applied to PMME of mean and trend of surface air temperature anomalies over East Asia by analyzing the temporal uncertainty of temperature anomaly prediction. However, while the meteorological fields were found to have spatial uncertainty, conversely, some meteorological fields (e.g. precipitation and cloud, among others) did not satisfy the symmetrical Gaussian distribution. Therefore, how to estimate the uncertainty of climate model prediction for some other meteorological fields by higher-dimension extension of GED needs to be further investigated.

Acknowledgements. This study was supported by the National Key Basic Research and Development (973) Program of China (Grant No. 2012CB955204), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), and the Research openfund of Jiangsu Meteorology Bureau (Grant Nos. Q201205, KM201107, and K201009).

REFERENCES

Doblas-Reyes, F. J., M. Deque, and J.-P. Piedelievre, 2000: Multimodel spread and probabilistic seasonal forecasts in PROVOST. *Quart. J. Roy. Meteor. Soc.*, **126**, 2069–2087.

Du, H., F. J. Doblas-Reyes, J. García-Serrano, V. Guemas, Y. Soufflet, and B. Wouters, 2012: Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. *Climate Dyn.*, **39**(7–8), 2013-2023, doi: 10.1007/s00382-011-1285-9.

Frame, D. J., N. E. Faull, M. M. Joshi, and M. R. Allen, 2007: Probabilistic climate forecasts and inductive problems. *Philosophical Transactions of the Royal Society*, **365**(1857), 1971–1992, doi: 10.1098/rsta.2007.2069.

Fritsch, J. M., J. Hilliker, J. Ross, and R. L. Vislocky, 2000: Model consensus. *Wea. Forecasting*, **15**, 571–582.

Furtado, J. C., E. D. Lorenzo, N. Schneider, and N. A. Bond, 2011: North Pacific decadal variability and climate change in the IPCC AR4 models. *J. Climate*, **24**, 3049–3067.

Greene, A. M., L. Goddard, and U. Lall, 2006: Robabilistic multimodel regional temperature change projections. *J. Climate*, **19**, 4326–4343.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**, 1095–1107, doi: http://dx.doi.org/10.1175/2009BAMS2607.1.

Hoerling, M., and Coauthors, 2011: On North American decadal climate for 2011–20. *J. Climate*, **24**, 4519–4528, doi: http://dx.doi.org/10.1175/2011JCLI4137.1.

Honerkamp, J., 2002: Statistical physics: An Advanced Approach with Applications. 2nd ed. Springer, Berlin, 536.

Hurrell, J., G. Meehl, D. Bader, T. Delworth, B. Kirtman, and B. Wielicki, 2009: A unified modeling approach to climate system prediction. *Bull. Amer. Meteor.* Soc., **90**, 1819–1832.

Keenlyside, N., M. Latif, J. Jungclaus, E. Roeckner, V. Semenov, and W. Parket, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, **453**, 84–88.

Kharin, V. V., and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate*, **15**, 793–799.

Krzysztofowicz, R., 1983: Why should a forecaster and a decision maker use Bayes theorem. *Water Resour. Res.*, **19**, 327–336.

Latif, M., and Coauthors, 2009: Dynamics of decadal

climate variability and implication for its prediction. [Available on line at http://www.oceanobs09.net/blog/?p=104.]

Meehl, G. A., and Coauthors, 2009: Decadal prediction: can it be skillful? *Bull. Amer. Meteor. Soc.*, **90**, 1467–1484.

Meehl, G. A., A. Hu, and C. Tebaldi, 2010: Decadal prediction in the Pacific region. *J. Climate*, **23**, 2959–2973, doi: http://dx.doi.org/10.1175/2010JCLI3296.1.

Min, Y. M., V. N. Kryjov, and C. K. Park, 2009: A probabilistic multimodel ensemble approach to seasonal prediction. *Wea. Forecasting,* **24**, 812–828, doi: http://dx.doi.org/10.1175/ 2008WAF2222140.1.

Mochizuki, and Coauthors, 2010: Pacific decadal oscillation hindcasts relevant to near-term climate prediction. *Proc. the National Academy of Sciences*, **107**, 1833–1837.

Murphy, A. H., 1973: A new vector partition of the probability score. *J. Appl. Meteor.*, **12**, 595–600.

Palmer, T. N., and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal to Interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853–872.

Peng, P., A. Kumar, H. van den Dool, and A. G. Barnston, 2002: An analysis of multimodel ensemble predictions for seasonal climate anomalies. *J. Geophys. Res.*, **107**, 4710, doi: 10.1029/2002JD002712.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.*, **133**(5), 1155–1174.

Räisänen, J., and L. Ruokolainen, 2006: Probabilistic forecasts of near-term climate change based on a resampling ensemble technique. *Tellus*, **58A**, 461–472.

Silverman, B. W., 1986: *Density Estimation for Statistics and Data Analysis*. Chapman and Hall/CRC, London/Boca Raton, 75–93.

Smith, D., S. Cusack, A. Colman, C. Folland, G. Harris, and J. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global circulation model. *Science*, **317**, 796–799.

Stephenson, D. B., and F. J. Doblas-Reyes, 2000: Statistical methods for interpreting Monte Carlo ensemble forecasts. *Tellus A*, **52**, 300–322.

Sugiura, N., T. Awaji, S. Masuda, T. Toyoda, H. Igarashi, Y. Ishikawa, M. Ishii, and M. Kimoto, 2009: Potential for decadal predictability in the North Pacific region. *Geophys. Res. Lett.*, **36**, L20701, doi: 10.1029/2009GL039787.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, cited 2009: A Summary of the CMIP5 Experiment Design. [Available on line at http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf.]

Thompson, J. C., 1962: Economic gains from scientific advances and operational improvements in meteorological prediction. *J. Appl. Meteor.*, **1**, 13–17.

Tippett, M. K., and A. G. Barnston, 2008: Skill of Multimodel ENSO Probability Forecasts. *Mon. Wea. Rev.*, **136**(10), 3933–3946.

Wu, B., and T. J. Zhou, 2011: Prediction of decadal variability of sea surface temperature by a coupled global climate model FGOALS_gl developed in LASG/IAP. *Chinese Science Bulletin*, **57**(19), 2453–2459, doi: 10.1007/s11434-012-5134-y.

Yoshimitsu, C., and Coauthors, 2012: An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC. *Climate Dyn.*, **23**, 1–22, doi: 10.1007/s00382-012-1351-y.

Zhi, X. F., Q. Wu, Y. Q. Bai, and H. X. Qi, 2010: The multimodel superensemble prediction of the surface temperature using the IPCC AR4 scenario runs. *Scientia Meteorologica Sinica,* **30**(5), 708–714. (in Chinese)