Theor Appl Climatol (2018) 134:37–50
DOI 10.1007/s00704-017-2250-2

ORIGINAL PAPER

Simulation of tropical cyclone activity over the western North Pacific based on CMIP5 models

Haibo Shen[1] · **Weican Zhou**[1,2,3] · **Haikun Zhao**[4]

Received: 15 June 2016 / Accepted: 14 August 2017 / Published online: 15 September 2017

Abstract Based on the Coupled Model Inter-comparison Project 5 (CMIP5) models, the tropical cyclone (TC) activity in the summers of 1965–2005 over the western North Pacific (WNP) is simulated by a TC dynamically downscaling system. In consideration of diversity among climate models, Bayesian model averaging (BMA) and equal-weighed model averaging (EMA) methods are applied to produce the ensemble large-scale environmental factors of the CMIP5 model outputs. The environmental factors generated by BMA and EMA methods are compared, as well as the corresponding TC simulations by the downscaling system. Results indicate that BMA method shows a significant advantage over the EMA. In addition, impacts of model selections on BMA method are examined. To each factor, ten models with better performance are selected from 30 CMIP5 models and then conduct BMA, respectively. As a consequence, the ensemble environmental factors and simulated TC activity are similar with the results from the 30 models' BMA, which verifies the BMA method can afford corresponding weight for each model in the ensemble based on the model's predictive skill. Thereby, the existence of poor performance models will not particularly affect the BMA effectiveness and the ensemble outcomes are improved. Finally, based upon the BMA method and downscaling system, we analyze the sensitivity of TC activity to three important environmental factors, i.e., sea surface temperature (SST), large-scale steering flow, and vertical wind shear. Among three factors, SST and large-scale steering flow greatly affect TC tracks, while average intensity distribution is sensitive to all three environmental factors. Moreover, SST and vertical wind shear jointly play a critical role in the inter-annual variability of TC lifetime maximum intensity and frequency of intense TCs.

1 Introduction

The western North Pacific (WNP) is the region where tropical cyclones (TCs) occur most frequently in the world. TC activity in this basin often causes serious casualties and economic losses in China and other countries in Southeast Asia. In recent years, such losses are particularly severe in China due to the high population density and rapid development in the coastal region of China (Zhang et al. 2009; Zhang et al. 2013). Thereby, modeling simulations and projections of TC activity in the western North Pacific are an important research topic that has significant societal impact.

In the context of global warming, simulations and projections of TC activity have been conducted progressively. Presently, the research mainly focuses on the following three aspects: (1) to deduce present TC activity and its future projection based on changes of large-scale environmental factors (Vecchi and Soden 2007; Yokoi and Takayabu 2009; Zhang

✉ Weican Zhou
zhwnim@nuist.edu.cn

[1] Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing 210044, China

[2] Joint International Research Laboratory of Climate and Environment Change (ILCEC), Nanjing University of Information Science and Technology, Nanjing 210044, China

[3] Key Laboratory of Meteorological Disaster, Ministry of Education (KLME)/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing 210044, China

[4] Pacific Typhoon Research Center, Nanjing University of Information Science and Technology, Nanjing 210044, China

et al. 2010). Generally, the TC genesis potential index (GPI) and potential intensity (PI), which are directly linked with large-scale environmental factors, are calculated and applied to project TC activity. This approach can roughly estimate TC activity and changes, but the results are often less accurate and incomplete, which cannot provide detailed features of TC activity; (2) to simulate TC activity directly using high-resolution numerical models or algorithm (Oouchi et al. 2006; LaRow et al. 2008; Sugi et al. 2009; Zhao et al. 2009; Chen and Lin 2011; Camargo 2013). This method is straightforward and provides clear information of TC activity. However, most studies use high-resolution atmospheric models to detect TC forced by prescribed sea surface temperature from observations or predictions, rarely employing ocean-atmosphere coupling models (Camargo 2013). Emanuel (2013) still pointed out that this approach severely underresolves tropical cyclones, resulting in a substantial truncation of the intensity spectrum of simulated storms, even at 50-km grid spacing (Zhao et al. 2009), and usually produces fewer events than observed; (3) to simulate and predict TC activity based on coarse-resolution general circulation model outputs and/or global reanalysis datasets using statistically (Yu and Wang 2009; Zhao and Held 2010; Villarini and Vecchi 2012,2013) or dynamically downscaling methods (Knutson et al. 2008; Bender et al. 2010; Lavender and Walsh 2011). The statistically downscaling method is based on statistical skill and theory without representation of clear physical processes. Meanwhile, the dynamically downscaling method can better describe various physical processes and obtain details related to TC activity with high-resolution, but consumes much time and calculation resource. On the basis of the statistical/dynamic techniques proposed by Emanuel et al. (2008), a downscaling system has been developed specifically for the study of TC activity in the WNP basin by Zhao et al. (2010). Using this downscaling system, possible physical mechanisms for the TC activity change in the WNP basin on various time scales have been explored (Zhao et al. 2010, 2011, 2014; Zhao and Wu 2014; Zhao 2016) and the future changes in the TC activity have been predicted (Wang and Wu 2012, 2015).

Simulations of TC activity are highly dependent on the accuracy of large-scale environmental factors. In previous study, multi-models average of environmental factors is utilized to make light of the uncertainty between climate models(Knutson et al. 2007, 2008, 2010; Emanuel et al. 2008; Bender et al. 2010; Villarini and Vecchi 2012, 2013; Camargo 2013). However, the existence of poor skill models can lessen the reliability of the ensemble mean. In addition, more complex physical and chemical parameterization schemes employed in models, as well as the increased quantity of models, the multi-models average behaves unsatisfactorily. It is necessary to carefully evaluate the performance of each individual model, and then conduct differently weighted ensemble to obtain the optimum results. By this, Bayesian model averaging (BMA) may achieve the desired effect. It is a statistical analysis method based on Bayes' theorem and accounts for uncertainties in the models. BMA can combine information from different sources and maximize the use of each individual model result (Raftery et al. 2005). The method performs well in the model ensemble prediction of temperature, precipitation, wind vectors etc. on synoptic-scale (Raftery et al. 2005; Tebaldi et al. 2005; Sloughter et al. 2007, 2013; Liu et al. 2013; Zhi et al. 2014).

In the present study, we propose an approach to simulate climatic features of TC activity over the WNP basin. In consideration of the uncertainties among CMIP5 models, other than multi-model average with or without evaluation and selection (e.g., Emanuel et al. 2008; Bender et al. 2010), BMA Method is applied to produce superior ensemble results of environmental factors influencing TC activity and is also compared with multi-model average. Furthermore, a TC dynamically downscaling system is adopted to enhance the regional resolution. The data and method used in this study are introduced in Section 2. In Section 3, the ensemble effectiveness of BMA and EMA methods are compared based upon two aspects: large-scale environment features and simulations of TC activity. In Section 4, ten models that perform better are selected in terms of the models' evaluation, and then the ten models' BMA results are compared with the 30 models' to detect possible effect of model selection in the BMA method. Later, three sensitive experiments are designed to explore the sensitivity of TC activity to SST, large-scale steering flow, and vertical wind shear using the BMA method and the downscaling system in Section 5. Lastly, the conclusion is presented in Section 6.

2 Data and methodology

2.1 Data

The study period covers the summers (July–September) from 1965 to 2005. TC observations are derived from the Joint Typhoon Warning Center (JTWC) best-track dataset, and SST observation is from NOAA Extended Reconstructed SST V3.0 in $2^{\circ} \times 2^{\circ}$ grid (Smith et al. 2008). Monthly mean winds are extracted from NCEP Reanalysis with a resolution of $2.5^{\circ} \times 2.5^{\circ}$ in latitude and longitude (Kalnay et al. 1996).

The historical experiment output of 30 CMIP5 models is used in this study (http://cmip-pcmdi.llnl.gov/cmip5/availability.html). Details of the 30 CMIP5 models are listed in Table 1. Due to differences in resolution among these models and observations, all data are interpolated to a $2.5^{\circ} \times 2.5^{\circ}$ grid for the convenience and unification of analysis.

Table 1 Description of 30 CMIP5 models

Model	Institution	Country	Horizontal resolution
ACCESS1-0	CSIRO-BOM	Australia	1.88 × 1.25
ACCESS1-3	CSIRO-BOM	Australia	1.88 × 1.25
BCC-CSM1-1	BCC	China	2.80 × 2.80
BCC-CSM1-1-m	BCC	China	1.13 × 1.12
CanESM2	CCCMA	Canada	2.80 × 2.80
CCSM4	NCAR	USA	1.25 × 0.94
CESM1-BGC	NSF-DOE-NCAR	USA	1.25 × 0.94
CESM1-CAM5	NSF-DOE-NCAR	USA	1.25 × 0.94
CESM1-FASTCHEM	NSF-DOE-NCAR	USA	1.25 × 0.94
CESM1-WACCM	NSF-DOE-NCAR	USA	2.50 × 1.89
CMCC-CESM	CMCC	Italy	3.76 × 3.76
CMCC-CMS	CMCC	Italy	1.88 × 1.88
CNRM-CM5	CNRM-CERFACS	France	1.41 × 1.40
CNRM-CM5-2	CNRM-CERFACS	France	1.41 × 1.40
FGOALS-g2	LASG-CESS	China	2.81 × 3.05
FIO-ESM	FIO	China	2.81 × 2.81
GFDL-CM2p1	NOAA GFDL	USA	2.50 × 2.00
GFDL-CM3	NOAA GFDL	USA	2.50 × 2.00
GFDL-ESM2G	NOAA GFDL	USA	2.50 × 2.00
GFDL-ESM2M	NOAA GFDL	USA	2.50 × 2.00
GISS-E2-H	NASA GISS	USA	2.50 × 2.00
GISS-E2-H-CC	NASA GISS	USA	2.50 × 2.00
GISS-E2-R	NASA GISS	USA	2.50 × 2.00
GISS-E2-R-CC	NASA GISS	USA	2.50 × 2.00
HadGEM2-AO	NIMR/KMA	Korea/UK	1.88 × 1.25
IPSL-CM5A-LR	IPSL	France	3.75 × 1.88
IPSL-CM5A-MR	IPSL	France	2.50 × 1.25
IPSL-CM5B-LR	IPSL	France	3.75 × 1.88
NorESM1-M	NCC	Norway	2.5 × 1.88
NorESM1-ME	NCC	Norway	2.5 × 1.88

2.2 Methodology

BMA is a statistical analysis method that is based on Bayes' theorem and accounts for uncertainties in the model ensemble. It combines information from different sources and maximize the use of each individual model result by assigning different weighting coefficient. The weighting coefficients provided by BMA are the possible posterior probabilities of each model, which reflect the models' skill over the training period. Detailed introduction of the BMA method can be found in Raftery et al. (2005) and Liu et al. (2013). In this study, we focus on the BMA ensemble results of environmental factors from CMIP5 models.

The downscaling system used in this study is similar to that in Emanuel (2006) and Emanuel et al. (2008). The system consists of TC track module and TC intensity module, which are used to simulate TC track and TC intensity, respectively, and the TC generation module is replaced by observed generation of storms. The TC track module is developed based on the trajectory model proposed by Wu and Wang (2004). The generated storm is described by equations of motion for particles and moves following the large-scale steering flow (i.e., vertically averaged winds between 850 and 300 hPa) plus mean beta drift. In addition, random synoptic-scale disturbances are put into the translation vectors. TC intensity is described following the coupled hurricane intensity prediction model proposed by Emanuel et al. (2008), which is developed based on the Carnot heat engine and TC maximum potential intensity theory. This intensity module is designed for the simulations of intensity of an ideal symmetrical hurricane. It mainly considers effects of latent heat flux transport from tropical oceans to tropical cyclones, radiative cooling at the outflowing layer of the tropical cyclone, and vertical wind shear. This model has been applied in operational forecast due to its satisfactory performance in TC intensity simulations. In the TC intensity module, TC intensity is calculated

along the simulated tracks based on surrounding environmental factors, including SST and vertical wind shear. Some parameters used in the modules are the same as that in Emanuel et al. (2008).

3 Comparison of two multi-model ensemble methods

3.1 Comparison of ensemble environmental factors by the two methods

Multi-model ensemble can somewhat reduce uncertainty and increase the accuracy of large-scale environmental factors simulated by global climate models. However, the weighting coefficient assigned to each model for ensemble analysis must be considered carefully. The climatologically mean SST distributions from ensemble results using BMA method and EMA method are shown in Fig. 1. It shows clearly that SST distribution from the BMA is consistent with that of Reanalysis, realistically reproduces the spatial pattern of the observation. While the results from the EMA method are not satisfactory, SST is underestimated in much of the area. Over the region around 20° N and lower latitudes where TCs are active, SST produced by EMA method is lower than observations by more than 0.5 °C. Besides, the correlation coefficient of time series of SST averaged over the region (0 ~ 40° N, 100° E ~ 180° E) between BMA ensemble result and Reanalysis is 0.79, and the value is only 0.14 between the EMA result and Reanalysis. In addition, meridional and zonal wind speed (figure not shown) from the BMA method are also more accurate than those from the EMA method.

Thereby, it is concluded that the BMA can obtain environmental factors more consistent with observations. Furthermore, are the simulations of TC using the environmental factors from the BMA ensemble can also be satisfactory?

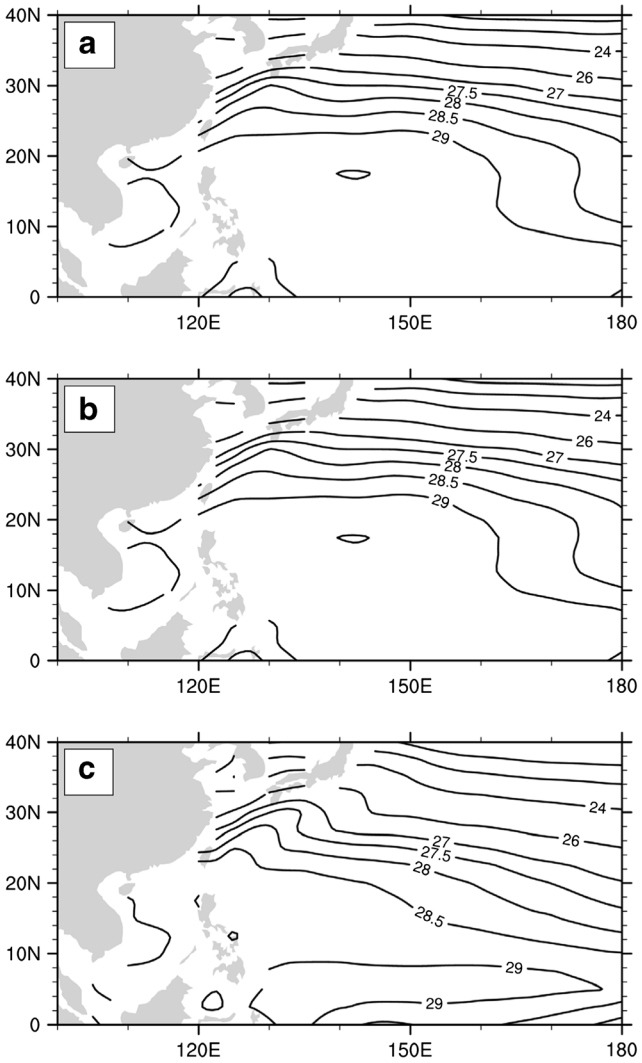

Fig. 1 Average SST distributions in July–September during the period 1965–2005 (°C). **a** NOAA Extended Reconstructed SST. **b** SST of 30 models BMA ensemble. **c** SST of 30 models EMA ensemble

3.2 Comparison of TC simulations using the two types of ensemble environmental factors

Input environmental factors for the downscaling system include SST, large-scale steering flow and vertical wind shear. Characteristics of the TC activity are simulated using the downscaling system. First of all, simulate TC characteristics using the NOAA and NCEP Reanalysis data as input, viewing the simulations as the best performance of the downscaling system, and compare the simulations with JTWC data to verify the reliability of the downscaling technique. Then, simulate TC from BMA and EMA ensemble environmental factors, Reanalysis downscaling simulations as reference, in order to exclude the influence of downscaling system.

The frequency of TC occurrence at each 2.5° × 2.5° grid is calculated to quantify TC tracks (Wu and Wang 2004), and Fig. 2 depicts the annual frequency distribution of TC occurrence from July to September. Fig. 2a, b shows that the TC tracks simulated with environment input of Reanalysis is consistent with TC observation, from one hand, verifying the reliability of the downscaling technique (Wang and Wu 2015). From Fig. 2b–d, the TC simulations using the environmental factors obtained by BMA method perform better than those by the EMA method. The former's spatial distribution is similar to Fig. 2b, while the latter underestimates the frequency of TC occurrence over the region east to Taiwan and the extratropics. However, both the two cannot realistically reproduce the two maximum area of TC occurrence. Figure 3 illustrates the spatial pattern of TC intensities in July–September during the period 1965–2005. The simulations with NOAA and NCEP Reanalysis as large-scale forcing can approximately reproduce the spatial features of TC intensities from observations. However, the strong intensity band to the east of Taiwan Island is overestimated, while the simulated TC intensities

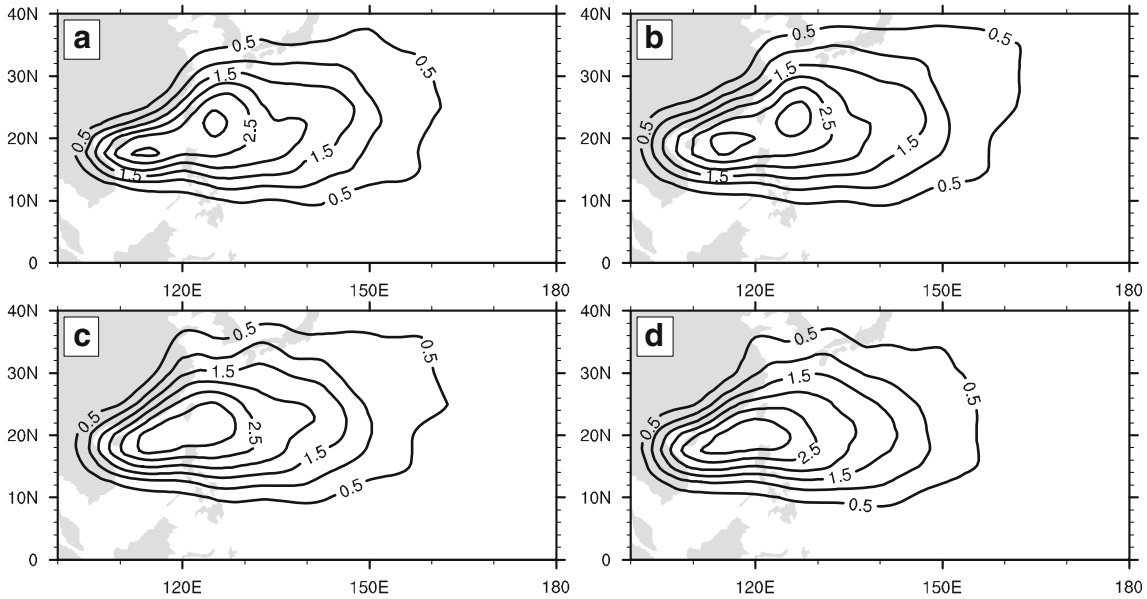

Fig. 2 Annual frequency of TC occurrence in July–September during the period 1965–2005 derived from **a** the JTWC best-track data, **b** the NCEP Reanalysis simulation, **c** the BMA ensemble simulation, and **d** the EMA ensemble simulation

are slightly lower than JTWC observations in the Philippine Islands and east of WNP. The discrepancy reflects the defect of the downscaling system (Wang and Wu 2015), and some studies also pointed out the possible uncertainty in the TC intensity of the JTWC dataset (Kossin et al. 2007). The spatial pattern of TC intensity simulated with ensemble environment factors by BMA as input is highly consistent with simulations with Reanalysis as input (Fig. 3b, c. Meanwhile, the simulations with ensemble environmental factors by EMA method as input underestimate in a great amount of the area and the simulated strong intensity region shifts to west (Fig. 3b, d).

In addition to the spatial pattern of basin-wide TC intensity, realistic simulations of inter-annual variability of TC intensity are also important. Recent studies (Wu and Zhao 2012; Zhao et al. 2014) suggested that the frequency of intense TCs (categories 4 and 5 on the Saffir-Simpson scale) and the lifetime maximum intensity are more sensitive to large-scale environmental conditions and changes. Thereby, the two metrics are employed to investigate the inter-annual variability of TC

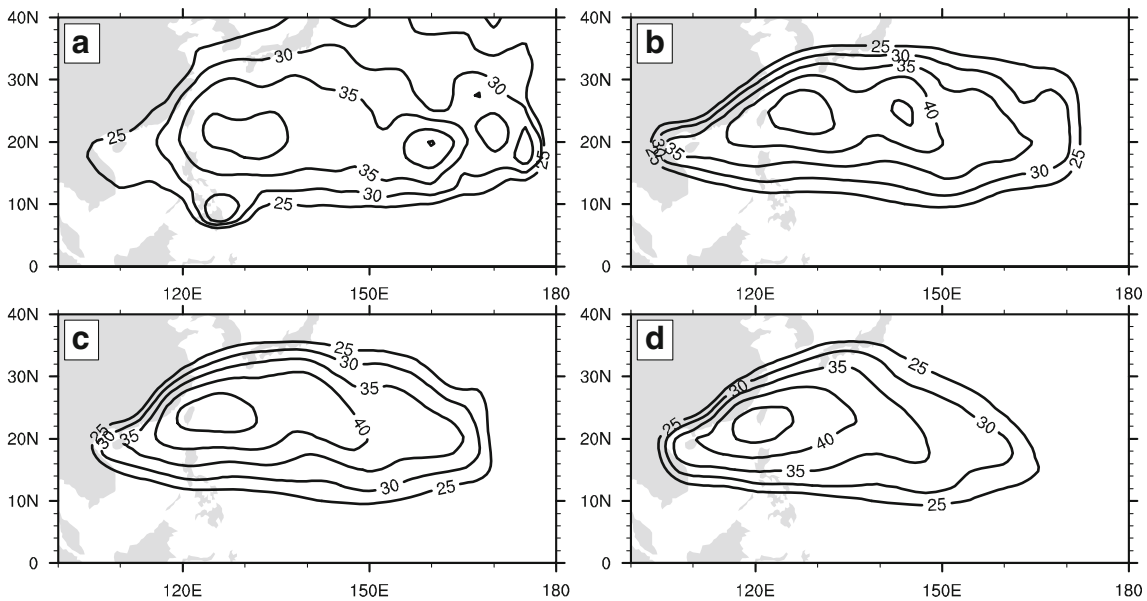

Fig. 3 As in Fig. 2, but of spatial distribution of TC intensity (units m s^{-1})

intensity. The annual lifetime maximum intensity is calculated by averaging the lifetime maximum intensities of individual TCs in each year (Wu 2007). As shown in Fig. 4, the simulated annual TC lifetime maximum intensity and frequency of cat-45 TCs with Reanalysis as input are close to JTWC data, particularly after 1975 when Dvorak technique is applied for TC observations. The temporal correlation coefficients between the simulations and observations are 0.75 and 0.67, respectively. Compared the simulations with environmental factors from the two ensemble methods with the simulations using Reanalysis as input, it is clear that the simulations with environmental factors from BMA method are more consistent with the simulations from Reanalysis than the simulations from EMA method. Correlation coefficients for lifetime maximum intensity and frequency of cat-45 TCs between the simulations with BMA method and from Reanalysis are 0.96 and 0.86, respectively. Biases in the simulations with EMA method are relatively large, and the correlation coefficients are 0.83 and 0.74 for the lifetime maximum intensity and frequency of cat-45 TCs, respectively. The root mean square errors (RMSE) for are also larger in the simulations with environmental factors from EMA than those from BMA method.

Generally, the BMA method can provide more accurate large-scale environmental factors than the EMA method and simulate TC activity more realistically with the downscaling system. The BMA method shows significant advantages in providing environmental factors for TC simulations. However, the above results are based on direct application of BMA method to all the 30 models. There may exist some models that have bad skills in simulating environmental factors. Effects of those models with unsatisfactory performance on the ensemble result should be assessed.

4 Impacts of model selections on BMA ensemble results

The BMA ensemble results might be affected by those among the 30 models that perform poorly in the simulations of environmental factors for TC activity. For this reason, we first

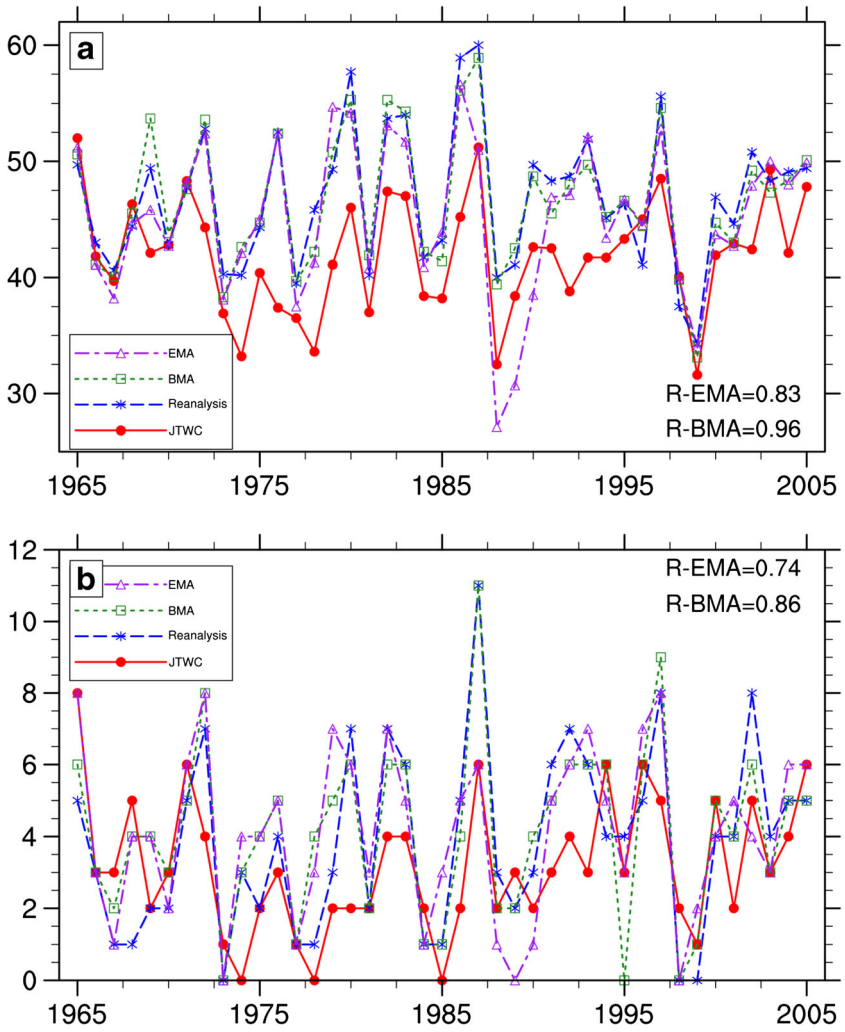

Fig. 4 Time series of **a** lifetime maximum intensity (units m s^{-1}) and **b** the number of intense TCs from the JTWC dataset and three simulations in July–September during the period 1965–2005. (R-EMA represents the correlation coefficient of NCEP Reanalysis simulation and EMA simulation, while R-BMA represents the correlation coefficient of NCEP Reanalysis simulation and BMA simulation)

evaluate the performance of each individual model for the simulations of SST and wind fields, and identify those models that can well simulate SST or wind fields. BMA are then conducted to the selected models and compare the ensemble results to 30 models BMA ensemble.

4.1 Evaluations of model performance for the simulation of environmental factors

Evaluations of the environmental factors from global climate models focus on two aspects: the mean climate state and inter-annual variability. The simulations of mean climate state reflect the models' capability of describing climate conditions on certain period, while the capability to realistically simulate the inter-annual climate variability is important for reliable prediction of future climate change scenario. NOAA and NCEP Reanalysis datasets are used as observations. The SST and wind fields (at 850 hPa) are evaluated, which are chosen based on the structure of the downscaling system.

The NOAA Extended Reconstructed SST and NCEP Reanalysis data are taken as observations for SST and wind speed. Pattern correlation coefficients (PCC) and root mean square errors (RMSE) between the factors' climate state of models output and observations are calculated and applied for evaluation of the model performance. Results are shown in Fig. 5a–e. The 30 models can generally well simulate the spatial pattern of climatic SST, and the correlation coefficients between all the model results and observations are greater than 0.9. In general, CMIP5 models' simulations of SST have been improved compared to that of IPCC AR4 (Yu et al. 2011). Large differences are found in the simulations of climate state of 850 hPa wind fields. For zonal wind, models that achieve high correlation coefficients also obtain small RMSE. CESM1-CAM5, CNRM-CM5, CNRM-CM5–2 can better simulate zonal wind at 850 hPa than other models. Correlation coefficients between these three model simulations and observations of zonal wind are greater than 0.95, and the RMSEs are around 1 m/s. The meridional wind simulations are generally less reasonable than the zonal wind simulations. Compared to models of IPCC AR4 (Yu et al. 2014), there is no significant improvement in the simulations of climatic 850 hPa wind fields in CMIP5.

The capability of the CMIP5 models for the simulations of inter-annual variability of environmental factors is evaluated based on mean SST and 850 hPa winds for the 41 summers averaged over 100° E ~ 180°, 0 ~ 40° N. We calculate the RMSE and time correlation coefficient (TCC) between the observation and outputs of each climate model. As shown in the right panels of Fig. 5b, d, f, the CMIP5 models' capability for SST simulation is still limited and large differences exist among results of CMIP5 models, but better than that of IPCC AR4 (Yu et al. 2011). The simulations of inter-annual variability of low-level winds are even worse. For the model simulations of inter-annual variability, further optimization and improvement are still necessary.

4.2 Selected models ensemble

The above evaluations have shown that the climate state of environmental factors can be better simulated than their inter-annual variability by CMIP5 models, and SST simulations are better than wind simulations for both climate state and inter-annual variability. It is worth noting that for simulations for the climate state and inter-annual variability of both SST and wind, the performance of each individual model changes significantly (Fig. 5). Considering the general performance, to each factor, those models are selected if their TCC, PCC, and RMSE of each environmental factor reach certain thresholds (denoted by blue and red lines in Fig. 5). Eventually, ten models are selected to each factor, respectively.

The BMA method is applied to SST and wind speed based on the ten selected models separately. Results are compared with that from the 30 CMIP5 models simulations. It is found that the difference between the two is negligible, and both results are almost identical to reanalysis (please refer to Fig. 1). For the description of inter-annual variability of environmental factors, the ensemble results of the 30 models' simulations are slightly better than that of the ten selected models' simulations. The correlation coefficients between the two ensemble results and Reanalysis are 0.52 (30 models ensemble) and 0.41 (10 selected models ensemble), respectively.

The above analysis indicates that both ensemble results can well reproduce the realistic environmental factors. Furthermore, we evaluate the impacts on TC simulations. The two BMA ensemble results of environmental factors are used respectively as input for the downscaling system to simulate TC activity. The simulated frequency of TC occurrence and TC intensity are shown in Fig. 6. It is found that the spatial pattern of TC occurrence is well simulated in two BMA results. Little difference can be found between the simulated TC occurrences driven by the two BMA ensemble results. However, the areas of maximum frequency are not correctly represented in both two simulations. The spatial pattern of TC intensity is also similar in the two simulations. TC lifetime maximum intensity in the two simulations is consistent to that from Reanalysis with correlation coefficients up to 0.96 and 0.95, respectively. The simulations of frequency of intense TCs are slightly worse, with correlation coefficients of 0.86 and 0.87, respectively between the two simulations and Reanalysis (Fig. 7). In general, it is illustrated that TC simulations driven by the BMA ensemble results of ten selected models and 30 models both present similar and realistic results.

The above results indicate that the BMA ensemble results of the ten selected models and 30 models produce similar large-scale environmental fields. The simulations of TC

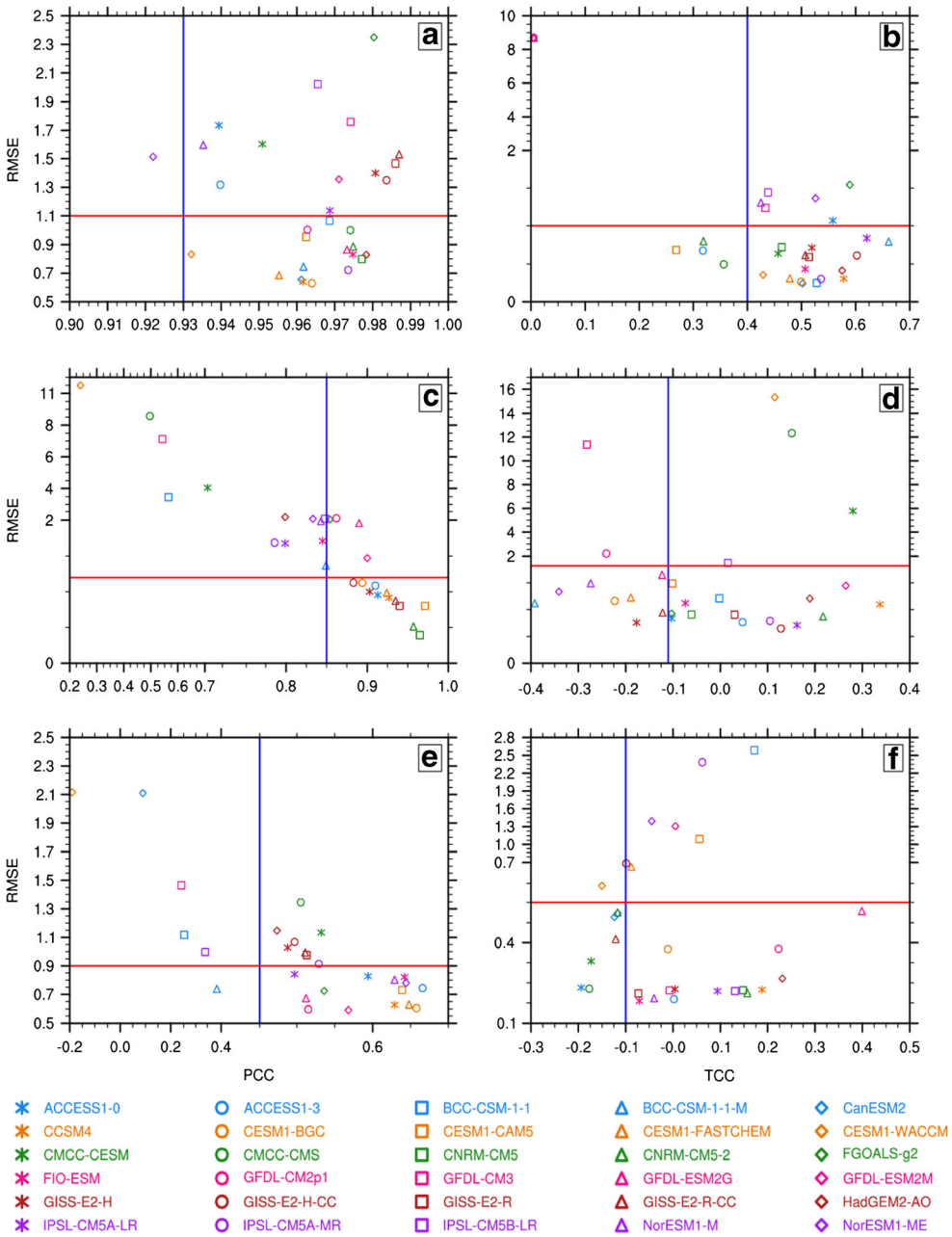

Fig. 5 Scatter plots of RMSE, PCC, and TCC derived from 30 individual GCMs. (**a**, **b**) SST; (**c**, **d**) zonal wind speed at 850 hPa; (**e**, **f**) meridional wind speed at 850 hPa. (**a**, **c**, **e**) RMSE and PCC of factors' climate state, (**b**, **d**, **f**) RMSE and TCC of factors' inter-annual variability. Red (blue) lines indicate the selection thresholds for RSME (PCC, TCC)

activity driven by the large-scale environmental fields from the two BMA results are also quite close. Apparently, the ensemble result of the 30 models is still reasonable, despite the poor performance of some models among the 30 CMIP5 models. This might be attributed to two reasons. First, increases in the number of models used for ensemble study may effectively reduce the uncertainties in various models and improve the general representation of the model results. Second, the BMA method has an advantage in ensemble. It uses the posterior probabilities of the models as the weights in the ensemble. The posterior probability of a model actually reflects the model's relative contribution to predictive skill. So theoretically, the BMA method can greatly reduce the impacts of poor performance models on the ensemble result. We can roughly verify the advantage of BMA. The BMA method is applied to each individual grid, and the domain average of weights can generally show the weights of the model in the ensemble result. We compare the rankings of 30 models' weights generated by BMA and sequences of evaluation of their performances to verify its weights allocation. It turns out that the orders of weights correspond well to the orders of evaluation performance for the 30 models, which demonstrates that the BMA method indeed determines the weight of each individual model based on the model skill. Thereby, the impacts of poor performance models on the ensemble result are greatly reduced. For the two reasons, we suggest

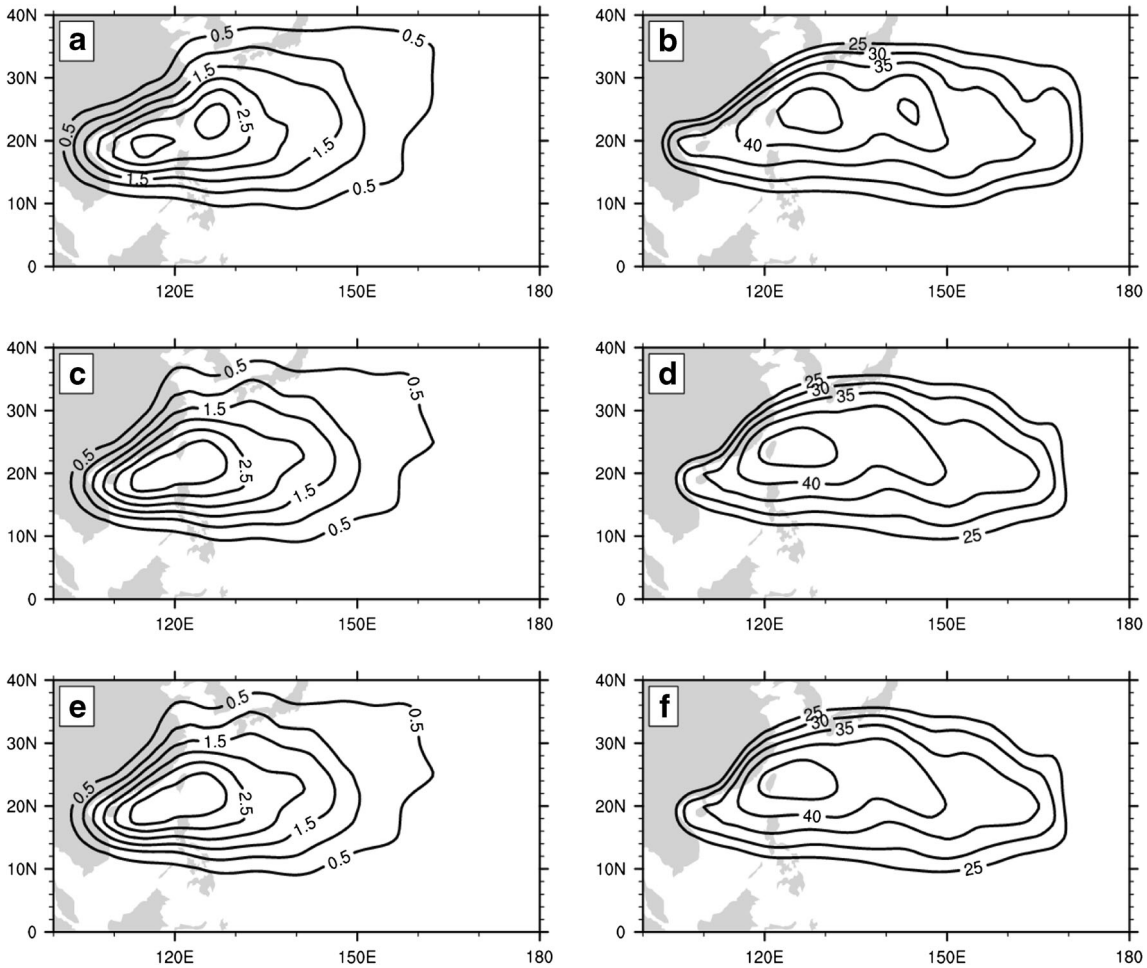

Fig. 6 Annual frequency of TC occurrence and spatial distribution of TC intensity (units m s^{-1}) in July–September during the period 1965–2005 derived from (**a**, **b**) the NCEP Reanalysis simulation, (**c**, **d**) the 30-models BMA ensemble simulation, (**e**, **f**) the 10-selected models BMA ensemble simulation

that the BMA method can be applied in multi-model ensemble while the evaluations and selection of models are unnecessary.

5 Sensitivity of the simulated TC activity to various environmental factors

Investigating the contribution of each individual environmental factor to TC activity helps better understand the predictability of TC, which is important for the projections of TC using various global climate models. Previous studies have shown that TC activity especially TC intensity is sensitive to SST anomaly (Knutson and Tuleya 2004; Emanuel 2005; Webster et al. 2005). However, most of these studies focused on the impact of local SST anomaly on TC activity and rarely considered changes in large-scale circulation induced by SST anomaly. In the studies of Wang and Wu (2012, 2015), the downscaling system was applied to investigate the impacts of environmental factors on TC intensity. In each experiment of their studies, only one single environmental factor was changed with global warming while all other factors remained historical conditions. By this method, they quantified the relative importance of each individual factor for changes of TC intensity. However, one weakness in their study is that the balance and interactions between these environmental factors are ignored. In the present study, sensitivity experiments are conducted using the downscaling model system and BMA method. The inherent physical and dynamical interactions among the environmental factors are considered and their impacts on TC track and intensity are investigated.

We focus on the impacts of SST, large-scale steering flow, and vertical wind shear on TC track and TC intensity. The control experiment conducts BMA to three factors respectively, which has demonstrated and compared to the observation in Section 3. The sensitivity experiment applies BMA first to one environmental factor from 30 models to obtain the weights, which are then employed to the other two factors. Hence, three factors ensemble results are obtained and comply one weight. This approach is repeated for the three environmental factors, respectively, and here comes three series of

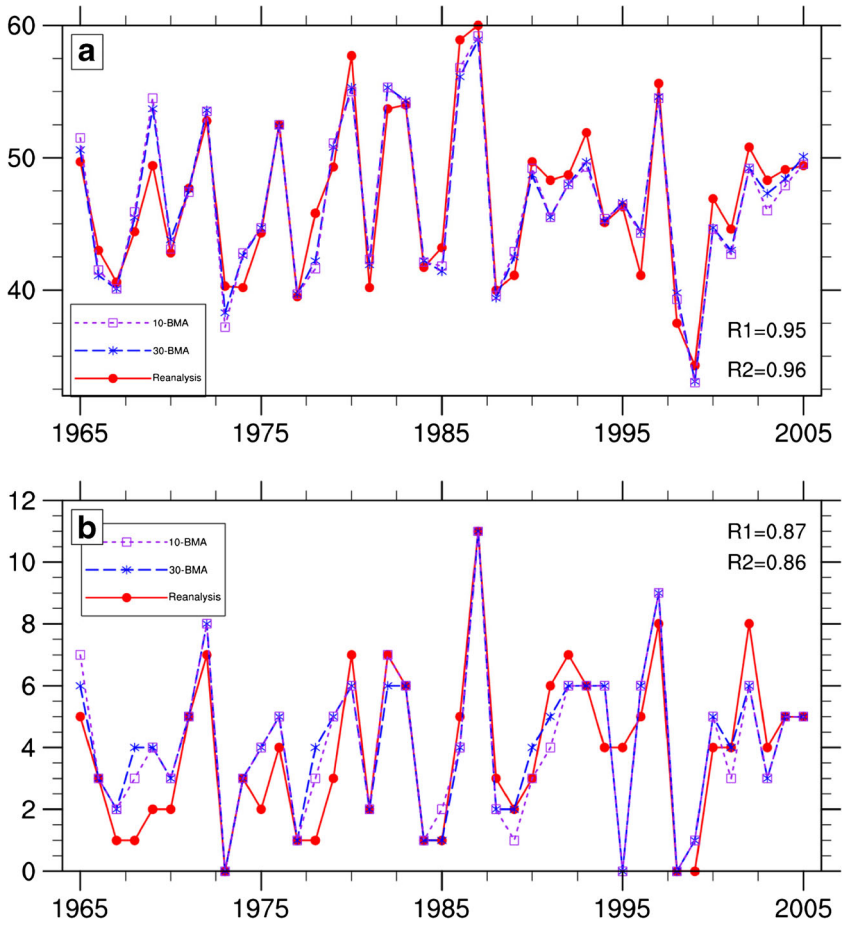

Fig. 7 Time series of **a** lifetime maximum intensity (units m s^{-1}) and **b** the number of intense TCs from NCEP Reanalysis and two BMA ensemble simulations in July–September during the period 1965–2005. (R1 represents the correlation coefficient of Reanalysis simulation and 10-models BMA simulation, while R2 represents the correlation coefficient of Reanalysis simulation and 30-models BMA simulation)

SST, large-scale steering flow, and vertical wind shear. The ensemble results are then used as input for the downscaling system to simulate TC activity. The characteristics of TC activity driven by the three series are compared with control experiment results to explore the TC sensitivity to SST, large-scale steering flow, and vertical wind shear. The three sensitive experiments are named as SST-BMA, LSF-BMA (large-scale steering flow), VWS-BMA (vertical wind shear), respectively. The weights of the models in the ensemble are determined by one single environmental factor, which makes this factor to be the optimum ensemble result while the other two factors are not. Since the same weights are used for all three environmental factors, the relationship among these factors in each climate model remains unbroken. For example, the SST-BMA sensitivity experiment applies BMA to get optimum SST ensemble and corresponding weights are then applied to retrieve other two factors, which can be considered that the other two factors are determined by SST based on the physical and chemical processes in individual coupled ocean-atmosphere climate model. SST is the key factor that dominates the process. The sensitivity experiment results closer to control experiment results mean that its corresponding factor plays a more significant role in TC activity. Otherwise, if the bias between the sensitivity experiment results and control experiment results is relatively large, then TC activity is more sensitive to the other two factors.

Figure 8 shows annual frequency of TC occurrence from control and three sensitivity experiments. Compared with the results of control experiment, the TC occurrence of SST-BMA experiment is generally similar in holistic pattern except the maximum area; LSF-BMA experiment better reproduces the spatial pattern but overestimates the maximum area; the bias between VWS-BMA experiment result and control experiment result is much larger, which failed to reproduce the TC occurrence pattern from control experiment. Large discrepancies exist among the simulations of TC intensity by the three experiments. As shown in Fig. 9, SST-BMA experiment underestimates TC intensity in the east of the region but overestimates TC intensity in the area west to 150° E; LSF-BMA experiment underestimates the TC intensity over most of the area; VWS-BMA experiment also generally underestimates the intensity and mistakes the position of TC intensity maximum. The three sensitivity experiments all perform poorly in the spatial distribution of TC intensity, so the influence of each factor to annual intensity cannot be neglected. Results shown in Fig. 10 and Table 2 indicate that the SST-BMA experiment can well reproduce the inter-annual variability of TC lifetime maximum intensity and frequency of intense TCs. The

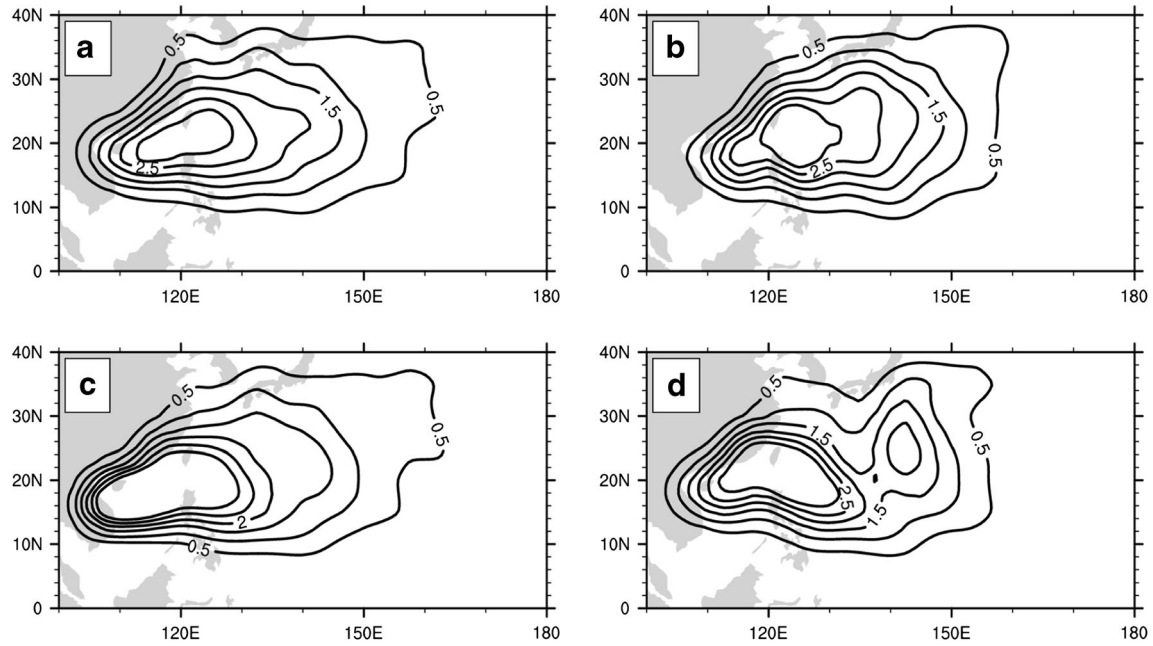

Fig. 8 Annual frequency of TC occurrence in July–September during the period 1965–2005 derived from **a** control experiment simulation, **b** the SST-BMA experiment simulation, **c** the LSF-BMA experiment simulation, **d** the VWS-BMA experiment simulation

correlation coefficients between the SST-BMA experiment results and control experiment results are 0.96 and 0.86 for TC lifetime maximum intensity and frequency of intense TCs, respectively. The results of VWS-BMA experiments are also coincide with these of control experiment. The correlation coefficients between LSF-BMA results and control experiment results are even less, both in TC lifetime maximum intensity and frequency of intense TCs. Thereby, SST and vertical wind shear jointly have a great impact on TC lifetime maximum intensity and frequency of intense TCs.

In summary, the TC track is under major influences of SST and large-scale steering flow, while TC average intensity distribution is sensitive to all three environmental factors. SST and vertical wind shear jointly play a critical role in the inter-annual variability of TC lifetime maximum intensity and frequency of intense TCs.

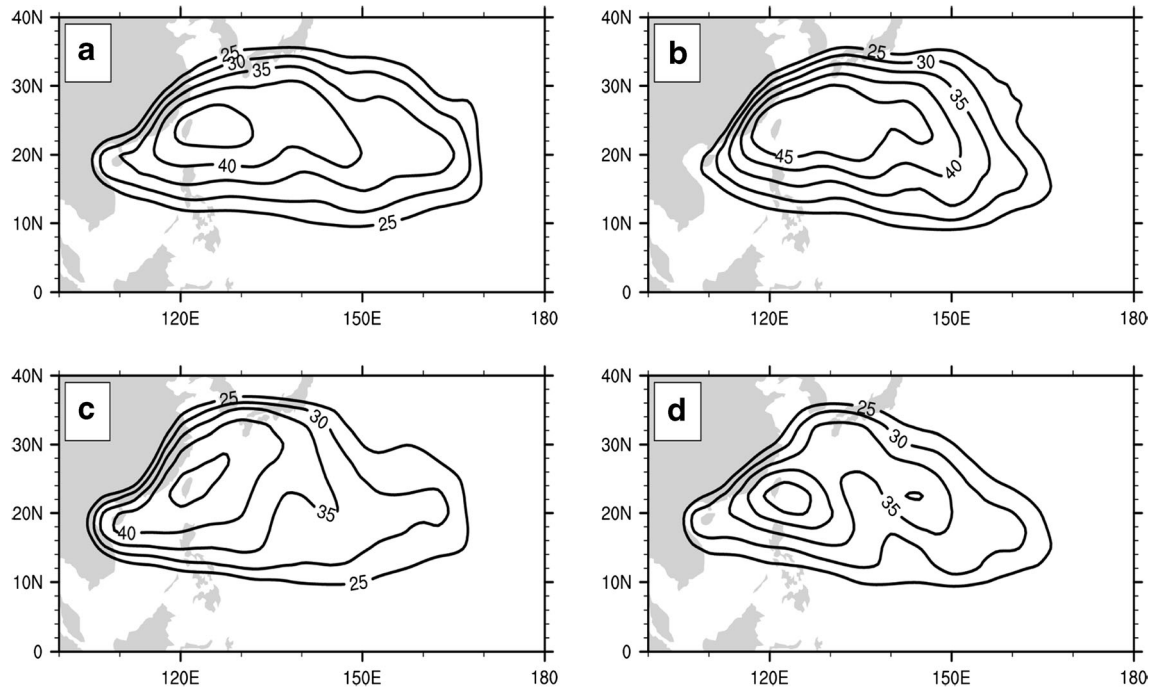

Fig. 9 As in Fig.8, but of spatial distribution of TC intensity (units m s^{-1})

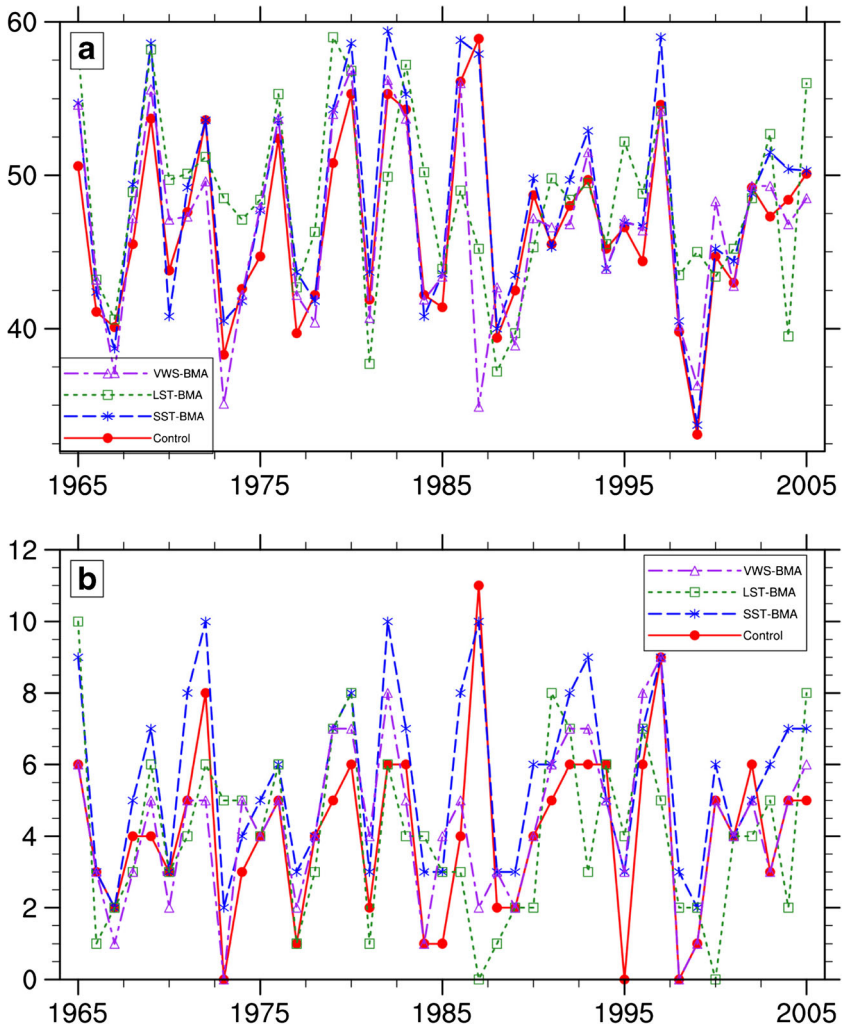

Fig. 10 Time series of **a** lifetime maximum intensity (units m s^{-1}) and **b** the number of intense TCs from control experiment simulation and three sensitivity experiment simulations in July–September during the period 1965–2005

6 Conclusions

With the development of the global climate models, more and more researchers apply them to TC investigation, mainly on TC simulations and projections. Other than qualitative analysis based on large-scale environmental factors (Vecchi and Soden 2007; Yokoi and Takayabu 2009; Zhang et al. 2010) and high-resolution numerical climate models (Oouchi et al. 2006; LaRow et al. 2008; Sugi et al. 2009; Zhao et al. 2009; Chen and Lin 2011), we adopt the dynamically downscaling method combining the CMIP5 outputs. In this way, the reliability of TC simulation highly depends on the large-scale environmental factors. Some previous study utilize the multi-models average of the environmental factors from different climate models to reduce the uncertainty (Knutson et al. 2008, 2010; Emanuel et al. 2008; Bender et al. 2010). However, with the existence of poor skill models as well as the increased quantity and complexity of models, the multi-models average behaves unsatisfactorily. In our paper, we employ the BMA method to assign corresponding weight for each model in the ensemble based on the model's predictive skill and compare it with the EMA (i.e., multi-models average). Furthermore, the advantage of the BMA is verified. Finally, with the BMA and downscaling system, three sensitive experiments are designed to explore the sensitivity of TC activity to SST, large-scale steering flow, and vertical wind shear. Based upon the above analysis, the major conclusions in the present study are as follows:

Table 2 Correlation coefficient of three sensitivity experiment and control experiment in TC lifetime intensity and numbers of intense TC

Sensitive experiment	Lifetime intensity	Frequency of intense TCs
SST-BMA	0.96	0.86
LSF-BMA	0.61	0.31
VWS-BMA	0.73	0.67

1. The BMA and EMA methods are applied to the ensemble of environmental factors simulated by the 30 CMIP5 models. Compared to those results using the EMA method, the ensemble factors with the BMA method are more

consistent with observation in both the mean climate state and the inter-annual variability of these environmental factors. The ensemble factors produced by the two methods are then used as input for the downscaling system to simulate TC activity. Results further indicate that the simulations when using the ensemble factors generated by the BMA method as inputs are more consistent with the downscaling simulations using the inputs from Reanalysis datasets and JTWC data. It is concluded that the ensemble factors from the BMA method are more realistic and reliable than those from the EMA method.

2. The environmental factors from the 30 CMIP5 models are evaluated and top ten models with better performance are selected for BMA ensemble. Results are compared with the ensemble factors of all the 30 models using the same BMA method. It is found that both ensemble results are consistent with observations, and little difference can be found between the two ensemble results, suggesting the advantages of weights allocation BMA. This effectively reduces the impacts of poor model performance and uncertainties on the ensemble results. Thereby, multiple models can be applied in ensemble using the BMA method while evaluations and selection of models are not necessary.
3. Among the three environmental factors (i.e., SST, large-scale steering flow, and vertical wind shear), SST and large-scale steering flow can greatly affect TC tracks in the WNP, while TC average intensity distribution is sensitive to all three environmental factors. SST and vertical wind shear jointly play a critical role in the inter-annual variability of TC lifetime maximum intensity and frequency of intense TCs.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. 41475091 and 41675072).

References

Bender MA, Knutson TR, Tuleya RE, Sirutis JJ, Vecchi GA, Garner ST, Held IM (2010) Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes. Science 327:454–458

Camargo SJ (2013) Global and regional aspects of tropical cyclone activity in the CMIP5 models. J Clim 26:9880–9902

Chen J, Lin S (2011) The remarkable predictability of interannual variability of atlantic hurricanes during the past decade. Geophys Res Lett 38:103–110

Emanuel K (2005) Increasing destructiveness of tropical cyclones over the past 30 years. Nature 436:686–688

Emanuel K (2006) Climate and tropical cyclone activity: a new model downscaling approach. J Clim 19:4797–4802

Emanuel K (2013) Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. PANS 110:12219–12224

Emanuel K, Sundararajan R, Williams J (2008) Hurricanes and global warming: results from downscaling IPCC AR4 simulations. Bull Am Meteorol Soc 89:347–367

Kalnay E et al (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77:437–471

Kossin JP, Knapp KR, Vimont DJ, Murnane RJ, Harper BA (2007) A globally consistent reanalysis of hurricane variability and trends. Geophys Res Lett 34:L04815

Knutson TR, Tuleya RE (2004) Impact of CO2-induced warming on simulated hurricane intensity and precipitation: sensitivity to the choice of climate model and convective parameterization. J Clim 17:3477–3495

Knutson TR, Sirutis JJ, Garner ST, Held IM, Tuleya RE (2007) Simulation of the recent multidecadal increase of atlantic hurricane activity using an 18-km-grid regional model. Bull Am Meteorol Soc 88:1549

Knutson TR, Sirutis JJ, Garner ST, Vecchi GA, Held IM (2008) Simulated reduction in Atlantic hurricane frequency under twenty-first-century warming conditions. Nat Geosci 1:359–364

Knutson TR, McBride JL, Chan J, Emanuel K, Holland G, Landsea C, Held I, Kossin JP, Srivastava AK, Sugi M (2010) Tropical cyclone and climate change. Nat Geosci 3:157–163

Larow TE, Lim YK, Shin DW, Chassignet EP, Cocke S (2008) Atlantic Basin seasonal hurricane simulations. J Clim 21:3191–3206

Lavender SL, Walsh KJE (2011) Dynamically downscaled simulations of Australian region tropical cyclones in current and future climates. Geophys Res Lett 38:264–265

Liu J, Xie Z, Linna Z, Binghao J (2013) BMA probabilistic forecasting for the 24-h TIGGE multi-model ensemble forecasts of surface air temperature. Chin J Atmos Sci 37:43–53 (in Chinese)

Oouchi K, Yoshimura J, Yoshimura H, Mizuta R, Kusunoki S, Noda A (2006) Tropical cyclone climatology in a global-warming climate as simulated in a 20 km-mesh global atmospheric model: frequency and wind intensity analyses. J Meteor Soc Jpn 84:259–276

Raftery AE, Gneiting T, Balabdaoui F, Polakowski M (2005) Using Bayesian model averaging to calibrate forecast ensembles. Mon Weather Rev 133:1155–1174

Smith TM, Reynolds RW, Peterson TC, Lawrimore J (2008) Improvements NOAAs historical merged land–ocean temp analysis (1880–2006). J Clim 21:2283–2296

Sloughter ML, Gneiting T, Raftery AE (2013) Probabilistic wind vector forecasting using ensembles and Bayesian model averaging. Mon Weather Rev 141:2107–2119

Sloughter JML, Raftery AE, Gneiting T, Fraley C (2007) Probabilistic quantitative precipitation forecasting using Bayesian model averaging. Mon Weather Rev 135:3209–3220

Sugi M, Murakami H, Yoshimura J (2009) A reduction in global tropical cyclone frequency due to global warming. Sola 5:164–167

Tebaldi C, Smith RL, Nychka D, Mearns LO (2005) Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multimodel ensembles. J Clim 18:1524–1540

Vecchi GA, Soden BJ (2007) Effect of remote sea surface temperature change on tropical cyclone potential intensity. Nature 450:1066–1070

Villarini G, Vecchi GA (2012) Twenty-first-century projections of north Atlantic tropical storms from cmip5 models. Nat Clim Chang 2:604–607

Villarini G, Vecchi GA (2013) Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. J Clim 26:3231–3240

Villarini G, Vecchi GA (2013) Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. J Clim 26:3231–3240

Wang C, Wu L (2012) Tropical cyclone intensity change in the western North Pacific: downscaling from IPCC AR4 experiments. J Meteor Soc Jpn 90:223–233

Wang C, Wu L (2015) Influence of future tropical cyclone track changes on their basin-wide intensity over the western North Pacific: downscaled CMIP5 projections. Adv Atmos Sci 32: 613–623

Webster PJ, Holland GJ, Curry JA, Chang HR (2005) Changes in tropical cyclone number, duration, and intensity in a warming environment. Science 309:1844–1846

Wu L (2007) Impact of Saharan air layer on hurricane peak intensity. Geophys Res Lett 34:139–158

Wu L, Wang B (2004) Assessing impacts of global warming on tropical cyclone tracks. J Clim 17:1686–1698

Wu L, Zhao H (2012) Dynamically derived tropical cyclone intensity changes over the western North Pacific. J Clim 25:89–98

Yokoi S, Takayabu YN (2009) Multi-model projection of global warming impact on tropical cyclone genesis frequency over the western north pacific. J Meteor Soc Jpn 87:525–538

Yu JH, Wang Y (2009) Response of tropical cyclone potential intensity over the North Indian Ocean to global warming. Geophys Res Lett 36:221–228

Yu JH, Tang S, Wu L, Shi N (2011) Assessment on simulation of thermodynamic parameters of tropical cyclone in IPCC-AR4 models. Acta Oceanol Sin 33:39–54 (in Chinese)

Yu JH, Zhao X, Chen C (2014) Assessment on climatologic simulation of atmospheric dynamic environment of tropical cyclone over western North Pacific in IPCC-AR4 models. Acta Oceanol Sin 36:61–72 (in Chinese)

Zhao H (2016) A downscaling technique to simulate changes in western North Pacific tropical cyclone activity between two types of El Nino events. Theor Appl Clim 123:487–501

Zhao H, Wu L (2014) Inter-decadal shift of the prevailing tropical cyclone tracks over the western North Pacific and its mechanism study. Meteorol Atmos Phys 125:89–101

Zhao H, Wu L, Wang R (2014) Decadal variations of intense tropical cyclones over the western North Pacific during 1948–2010. Adv Atmos Sci 31:57–65

Zhao H, Wu L, Zhou W (2010) Assessing the influence of the ENSO on tropical cyclone prevailing tracks in the western North Pacific. Adv Atmos Sci 27:1361–1371

Zhao H, Wu L, Zhou W (2011) Interannual changes of tropical cyclone intensity in the western North Pacific. J Meteor Soc Jpn 89:245–255

Zhao M, Held IM (2010) An analysis of the effect of global warming on the intensity of Atlantic hurricanes using a GCM with statistical refinement. J Clim 23:6382–6393

Zhao M, Held IM, Lin SJ, Vecchi GA (2009) Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. J Clim 22:6653–6678

Zhang Q, Wu L, Liu Q (2009) Tropical cyclone damages in China 1983-2006. Bull Am Meteorol Soc 90:485–495

Zhang J, Wu L, Ren F (2013) Changes in tropical cyclone rainfall in China. J Meteor Soc Jpn 91:585–595

Zhang Y, Wang HJ, Sun JQ, Helge D (2010) Changes in the tropical cyclone genesis potential index over the western North Pacific in the SRES A2 Scenario. Adv Atmos Sci 27:1246–1128

Zhi X, Li G, Peng T (2014) On the probabilistic forecast of 2 meter temperature of a single station based on Bayesian theory. Trans Atmos Sci 37:740–748 (in Chinese)