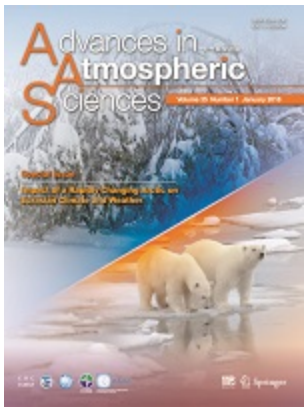

Application of breeding method in Beijing Climate Center Climate System Model for ensemble sub-seasonal forecast

Journal:	*Advances in Atmospheric Sciences*
Manuscript ID	AAS-2019-0120
Manuscript Type:	Original Article
Date Submitted by the Author:	31-May-2019
Complete List of Authors:	Wu, Li; Tsinghua University, Computer Science and Technology XIN, Xiaoge; Beijing Climate Center Xue, Wei; Tsinghua University, Computer Science and Technology Cheng, Yanjie; National Climate Center, China Meteorological Administration Liu, Xiangwen; National Climate Center, China Meteorological Administration Wu, Tongwen; National Climate Center, China Meteorological Administration,

Application of breeding method in Beijing Climate Center Climate System Model for ensemble subseasonal forecast

Li Wu[1], Xiaoge Xin*[2], Wei Xue[1], Yanjie Cheng[2], Xiangwen Liu[2], and Tongwen Wu[2]

1 Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

2 Beijing Climate Center, China Meteorological Administration, Beijing 100081, China

ABSTRACT

The ensemble method for the initial conditions used in the subseasonal and seasonal forecast system of Beijing Climate Center (BCC) is Lagged Average Forecasting (LAF). This study attempts to apply the Breeding of Growing Mode (BGM) method to generate perturbations for the ensemble forecast of BCC. The new ensemble hindcast with nine members including two pairs of BGM-perturbed initials and the others of LAF initials was carried out starting from May 1st and predicting for four months during each year of 2000-2014, namely LAF&BGM. Results are compared with the nine-member ensemble hindcast generated with the LAF. Comparisons show that the BGM&LAF ensemble decreases the root mean square error (RMSE) of zonal wind in the lower and upper troposphere and geopotential height in the middle troposphere with the largest amount in the forecast of the first month. Obvious improvement is also found for temporal correlation coefficient (TCC)

* Corresponding author: Xiaoge Xin
Email: xinxg@cma.gov.cn

skills of the atmospheric circulation for the forecast of one month. The LAF&BGM ensemble is better than the LAF ensemble to predict the precipitation and surface air temperature in North China. These improvements obtained by using the BGM method indicate its success in providing better skillful subseasonal forecasts, which has potential value to be used in the operational subseasonal forecasting.

Key words: Breeding of Growing Mode, Lagged Average Forecasting, ensemble, seasonal forecast, Beijing Climate Center.

1. Introduction

Seasonal forecasts based on Coupled General Circulation Models (CGCMs) are superior to those based on statistical models or uncoupled atmospheric general circulation models (AGCMs) (Graham et al., 2005； Beraki et al., 2015). However, the predictability of the forecast system based on CGCMs is primarily limited by the initial condition uncertainty (Stan and Kirtman, 2008). Ensemble forecasting with perturbed initial conditions is able to reduce the uncertainty from initial conditions and improve the prediction performance of the forecast system (Toth and Kalnay, 1993). One critical aspect of the ensemble strategy is how to generate the initial perturbations. According to previous studies (Leith, 1974; Toth and Kalnay, 1993), the initial perturbation should span the possible errors in the control analysis. As long as the ensemble perturbations represent the initial probability distribution around the control analysis, the ensemble forecast is better than the control forecast.

The ensemble method usually used for initialization of weather forecasts includes singular vectors (SVs), and the Breeding of Growing Mode (BGM). The SVs method was developed by Buizza and Palmer (1995), which computes the singular vector using a linear tangent model and its adjoint. The singular vector corresponding to the largest singular value is used as the fastest growing perturbation. The SVs method was implemented and used in European Centre for Medium-Range Weather Forecasts (ECMWF, Molteni et al., 1996). The BGM method provides an estimate of the fastest

sustainable growth error and represents a possible evolution of analysis errors (Toth and Kalnay, 1997). This method has been used to generate perturbations for ensemble forecasting at the National Centers for Environmental Prediction (NCEP) since the 1990s. The SV and BGM methods have been maturely applied in mesoscale probabilistic forecasting (Gneiting and Raftery, 2005).

Although both BGM and SVs ensemble methods are often used in the weather and extended weather forecasting, they are rarely used in subseasonaland seasonal forecasting. The main reasons are that generating the initial perturbation of SVs method is complicated and requires high computational resources, and the growth rate of BGM method is saturated too fast to be fully captured because of the high-frequency weather noise. So the ensemble strategy of many seasonal forecasting systems still uses the Lagged Average Forecasting (LAF), including the NCEP Climate Forecast System version 2 (CFSv2) (Saha et al., 2014), and the seasonal forecasts of Beijing Climate Center (Liu et al. 2014). The LAF method is started from the initial conditions at a time lagging the start of the forecast time by a different amount (Hoffman and Kalnay, 1983). This method is operationally feasible for subseasonal and seasonal forecasts, especially for the hindcast of several decades. However, using information from just before a prediction period may not be adequate because the atmosphere changes constantly.

Recently, improvements of the ensemble methods were proposed by some studies for predictions from subseasonal to seasonal time scale. Kug et al. (2010) applied an

empirical SVs method to the ENSO prediction of a hybrid coupled model and showed that the method was of potential application to seasonal prediction in the CGCM. Johanna and Piontek (2014) showed that an implementation of BGM in the ocean component within simulations of the ECHAM5/MPIOM coupled climate system model could improve the ensemble spread for both temperature and salinity compared with the LAF method. In studies of Hudson et al. (2013) and Kang et al. (2014), the prediction skill of Madden-Julian Oscillation (MJO) was improved by using the breeding method in the seasonal forecasts system of climate system models.

Currently there are still deficiencies for the subseasonal and seasonal prediction of Beijing Climate Center based on the LAF ensemble, especially in the Asian monsoon region (Liu et al., 2014, 2015). Application of a different ensemble method in the BCC forecast system is one of the solutions to improve its predictive skill. The SVs method is the fastest linear initial error growth perturbation method (Cheng et al., 2010), but it is difficult to be used in operational seasonal forecast system. Compared to the SVs, the BGM method is more economical on computational cost and easier to be implemented (Chikamoto et al., 2007; Saito et al., 2011). So this study tries to apply the BGM ensemble method in the subseasonal and seasonal forecasting of Beijing Climate Center. We will explore whether this method can improve the predictive skill in comparison with the LAF ensemble already used in BCC.

This article is organized as follows. Section 2 introduces the BCC climate system model, the BGM ensemble method and the experiments used in this study. Section 3

compares the hindcast skills predicted with the BGM and LAF ensemble methods. Section 4 presents the summary.

2. Model, Method and experiments

2.1 Model

The model used in this study is the version 2.0 of the BCC Climate System Model (BCC-CSM2-MR, Wu et al., 2019), which precipitates the sixth phase of Climate Model Intercomparison Project (CMIP6). This model is the updated version of BCC-CSM1.1m used in CMIP5 (Wu et al., 2013; Xin et al., 2013) and earlier seasonal forecasts (Liu et al., 2014, 2015). The main differences between BCC-CSM2-MR and BCC-CSM1.1m are the model physical processes in the atmosphere and land model components. In the atmospheric model, the vertical layers are increased from 26 to 46, and the gravity wave drag generated from convective sources is introduced in the stratosphere. In addition, the parameterization schemes of deep convection and cloud fraction in the atmosphere are both updated, and the indirect effects of aerosols are included in BCC-CSM2-MR (Wu et al., 2019).

The land component of BCC-CSM2-MR is BCC-AVIM2.0, the improvement of which from the earlier version BCC-AVIM1.0 includes the snow cover, plant function types and soil freezing/thawing processes. The ocean and sea ice model components of BCC-CSM2-MR are the same as BCC-CSM1.1m. The ocean model is Modular Ocean Model version 4.0 (MOM4) with 40 levels in the vertical (Griffies et al., 2005),

developed by the Geophysical Fluid Dynamics Laboratory (GFDL). The sea ice component is the Sea Ice Simulator (SIS) developed by GFDL. The horizontal resolution of the ocean and sea ice models is 1° longitude by 1/3° latitude with tripolar grid. Compared to BCC-CSM1.1m, the new model BCC-CSM2-MR shows significant improvements in simulating stratospheric quasi-biennial oscillation (QBO), precipitation in East Asia, and long-term trend of surface air temperature (Wu et al., 2019). BCC-CSM2-MR is currently used in the investigation of the prediction skill of seasonal forecasts, prepared for the new generation of operational forecast system of BCC.

2.2 Method

In the ensemble weather forecast, the breeding cycle adopted in the BGM method is 6h or 12 hours. Then the growth rate is generally saturated after 3-4 days (Toth and Kalnay 1997; Zhang and Pu, 2010). The fast saturation is not appropriate for seasonal forecasting, because the deviations in the slow changing physical processes have not grown yet. Yang et al. (2006) showed that the length of the breeding cycle should be chosen to separate the slowly varying dynamical modes. Cai et al. (2003) and Baehr et al. (2014) indicated that the length of the breeding cycle in coupled models needs to be longer than weather instabilities. After tuning the time of the breeding cycles, we set it to be 5 days. In addition, rescaling factors also need to be tuned to fit for the seasonal scale prediction. The procedure of the BGM method used in this study is described as follows.

1. The initial perturbation is added to a control forecast. Here we use a random perturbation of 1% on the initial field for the four atmospheric variables, including zonal wind, meridional wind, temperature and specific humidity as initial perturbation.

2. Both the control forecast and perturbation forecasts propagate a cycle (5 days) forward, then the unperturbed control forecast is subtracted from the perturbed forecasts to obtain a new perturbation.

3. The new perturbation obtained in step 2 is rescaled to the magnitude of the initial perturbation. Here the rescaling factor is the ratio of the initial perturbation and the new perturbation on the root mean square form.

4. The rescaled perturbations are added to the control forecast, providing the initial conditions for the new perturbed forecast.

5. The following cycles are repeated in steps 2-4 until the growth rate reaches saturation. And the final perturbations of 8 members in the BGM ensemble come from the fastest growing modes produced by four pairs of positive and negative perturbations.

As can be seen in Fig.1, the growth rate of global mean temperature slows after 4 breeding cycles (20 days). This is appropriate for the subseasonal forecast, because the short-scale atmospheric noise could be decreased. Since the hindcast starts on May 1st in this study, and the breeding process is designed to start from April 11th.

2.3 Experiments

The initial conditions of the hindcast in this study are produced by a coupled initialization run with BCC-CSM2-MR during 2000-2014. In the coupled initialization, the atmospheric component is nudged to the NCEP FNL Operational Model Global Tropospheric Analyses. The ocean component assimilates the sea surface temperature data from the Advanced Very High Resolution Radiometer (AVHRR) merged product, and temperature and salinity profiles from the Array for Real-time Geostrophic Oceanography (ARGO). Two sets of hindcasts with BCC-CSM2-MR were carried out with nine members. One includes two-pairs of BGM-generated initial conditions and the others of LAF initial conditions (LAF&BGM). The members of the other hindcast are all generated with the LAF methods. The hindcasts are initiated from 1 May of each year from 2000 to 2014 with four months' integration.

Among the nine members of the LAF ensemble, eight initial conditions are chosen lagged every three days with the control forecast. Among the nine LAF&BGM ensemble members, two pairs of initial conditions are got with the initial perturbations generated by the BGM method, and four LAF initial conditions are the first, third, fifth and seventh of the LAF ensemble members. So the difference between the two sets of hindcasts is the initial perturbations method. Hereinafter, these two sets of ensemble hindcasts are named LAF&BGM ensemble and LAF ensemble, respectively.

Results of the hindcasts are verified against ERA-Interim for atmospheric fields produced by the ECMWF (Dee et al. 2011), and observed precipitation data from the Global Precipitation Climatology Project (GPCP) version 2.2 (Adler et al. 2003; Huffman et al. 2009). Two different measures of forecast quality are adopted to assess the ensemble forecasts, including the temporal correlation coefficient (TCC) and root mean square error (RMSE). The prediction skills are computed based on the data from 2000 to 2014.

3. Results

The RMSEs between the hindcast and the observation are calculated separately for the LAF ensemble and BGM ensemble. Figure 2 shows the difference of RMSEs between BGM and LAF ensembles 850hPa meridional wind (V850), 850hPa zonal wind (U850), 500hPa geopotential height (Z500), and 200hPa zonal wind (U200) predicted for the first month (May). The value below zero (blue) indicates that the LAF&BGM ensemble forecasts have less error than the LAF ensemble forecasts. As shown in Fig. 2a, discrepancies of the RMSEs of V850 predicted the LAF&BGM and LAF ensembles mainly appear in the middle and high latitudes of the globe. Compared to the LAF ensemble, the RMSE predicted by the BGM ensemble shows prominent decrease for U850 in most area of the Northern Hemisphere (NH) (Fig. 2b), for Z500 in the middle and high latitudes of the globe, and for U200 in the middle latitudes of the NH (Fig. 2c, 2d). So the LAF&BGM ensemble is effective relative to the LAF ensemble in reducing the modeled errors of zonal wind in the

lower and upper troposphere and geopotential height in the middle troposphere for the prediction of the first month.

We further explore whether the LAF&BGM ensemble method is effective in reducing the RMSE beyond one month. Changes of RMSE from May to August (0 lead month to 3 lead month) for the four atmospheric variables averaged over the globe predicted by LAF&BGM ensemble relative to the LAF ensemble are calculated and shown in Fig. 3. It shows that the RMSE decreases most prominently in May for each variable. The RMSE of Z500 has the largest decrease among the four variables and the decrease extends through August. The globally averaged RMSE of Z500 is reduced by 7.9% in May, and 0.82% in summer. The reduction of the RMSE for V850 appears from May (2.3%) to July (0.4%). The RMSE of U200 is reduced by 6.2% in May, and 0.5% in June. The RMSE of U850 predicted by BGM ensemble decreases in May (6.1%) and June (0.4%). So the LAF&BGM ensemble method improves the prediction skill of Z500, V850 and U200 mainly in the forecast of the first month (May) measured by the global mean RMSEs.

The TCC skills of the four atmospheric variables (V850, U850, Z500, U200) in May predicted by BGM and LAF ensembles are compared in Fig. 4. There is more area with significant TCC skill over the mainland of the NH predicted by the LAF&BGM ensemble than by the LAF ensemble for all these variables. Significant TCC for V850 is found in North China and tropical western Pacific predicted by the LAF&BGM ensembles, but not by the LAF ensembles (Figs. 4a-4b). The

improvement of the TCC skill of U850 predicted by the LAF&BGM ensemble is found in Europe and East Asia (Figs. 4d-4f). Increased TCC skill in the BGM ensemble is more evident for the variables Z500 and U200. In most area of the middle-high latitudes of the NH, the TCC skill of Z500 and U200 of the LAF&BGM ensemble is higher than the LAF ensemble, especially over the Eurasian continent (Figs. 4i-4l).

The improved skills of the atmospheric variables predicted by the LAF&BGM ensemble indicate the potential better predictive skill of surface climate. Figure 5 shows the TCC skills of precipitation and SAT in May predicted by the LAF&BGM and LAF ensembles. The LAF&BGM ensemble has higher TCC skill for precipitation over North China, and parts of western China, Northeast China and southwestern China (Fig. 5a-5c). The TCC skill of the SAT predicted by the LAF&BGM ensemble is improved in North China and the western China compared to the LAF ensemble (Fig. 5d-5e).

We choose North China (105°-120°E, 32°-42°N) to calculate the regional mean precipitation and SAT time series in May from 2000 to 2014 predicted by the LAF&BGM and LAF ensembles. As shown in Fig. 6, the interannual variations of May precipitation in North China predicted by the LAF&BGM ensemble are more consistent with the observation than by the LAF ensemble (Fig. 6a). The temporal correlation coefficients of precipitation between the BGM ensemble prediction and the observation are 0.47, which is higher than that between the LAF ensemble and the

observation (0.37). The LAF&BGM ensemble also reproduces more reasonable interannual variation of the SAT in North China correlation coefficients of 0.86 and (Fig. 6b), which higher than the TCC between the LAF ensemble and the observation (0.52). The increased skill for precipitation and SAT in North China predicted by the LAF&BGM ensemble is consistent with the improvements in the atmospheric circulations including the Z500 and U200 in the Eurasian continent, as well as U850 in East Asia. However, the LAF&BGM ensemble cannot improve the predictive skills of the precipitation and SAT in summer over China compared with the LAF ensemble (Figure omitted).

It can be concluded that the subseasonal predictive skill in North China be improved through LAF&BGM ensemble method with the leading time of one month. Hudson et al. (2013) also found that the improvement in forecast performance using the breeding approach is primarily in the first month of the forecasts, although the breeding approach was used in both of the atmosphere and the ocean. Therefore, the BGM ensemble method is applicable for the subseasonal forecasts of climate models. The combined use of the breeding method and other ensemble strategy should be adopted for the operational seasonal forecasting. The operational seasonal forecasting of Japan Meteorological Agency (JMA) used both BGM and LAF ensemble methods to generate 51 ensemble members, which was shown to be able to well predict the Asian Summer Monsoon (Takaya et al. 2017). The combined use of BGM and other

methods in the seasonal forecast of BCC to generate large ensemble members will be investigated in the future.

4. Conclusions

The BGM ensemble method is used in the subseasonal forecasting of BCC to generate initial ensemble members. We choose to rescale the analysis error every five days and the growth rate of the analysis error saturates after four breeding cycle. The perturbed variables include zonal wind, meridional wind, temperature and specific humidity. Two pairs of initial conditions are generated with the BGM method, associated with the LAF ensembles. Hindcasts with nine members were carried out starting from May 1st and predicting for four months during each year of 2000-2014. Results are compared with the nine-member ensemble of the LAF method generated with different initial time lagging every 3 days.

Results show that the LAF&BGM ensemble can reduce the RMSE of U200, V850, U850, Z500, and U200 for the forecast of the first month (May) over most areas of the globe relative to the LAF ensemble. The RMSE of 500hPa geopotential height predicted by the BGM ensemble averaged over the globe is reduced by 7.9% in May and 0.82% in summer relative to the LAF ensemble. From the perspective of TCC, LAF&BGM ensemble method achieves higher predictive skills of U200 and Z500 over Eurasia, and U850 in East Asia. The TCC skills of Precipitation and SAT predicted by the LAF&BGM ensemble are better than LAF method over North China. This indicates that the LAF&BGM method is better than the LAF method for the

subseasonal prediction of BCC. Further work will focus on improving the seasonal predictive skill of the BCC-CSM model by using larger ensemble members with the combination of BGM and other ensemble methods.

Acknowledgments. This work was jointly supported by the National Key R&D Program of China (Grant No. 2016YFA0602100, 2016YFE0102400 and 2017YFA0604500), National Natural Science Foundation of China (Grant No. 41530961 and 41776010) and National Programme on Global Change and Air-Sea Interaction (Grant No. GASI-IPOVAI-06).

REFERENCES

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.P., Janowiak, J., and A. Gruber, 2003: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrol*., **4**, 1147-1167.

Baehr, J., and R. Piontek, 2014: Ensemble initialization of the oceanic component of a coupled model through bred vectors at seasonal-to-interannual timescales. *Geosci. Model Dev*., **7**, 453-461.

Beraki, A.F., Landman, W.A., and D. DeWitt, 2015: On the comparison between seasonal predictive skill of global circulation models: Coupled versus uncoupled. *J. Geophys. Res. Atmos*., **120**, 11-151.

Buizza, R., and T. N. Palmer,1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci*., **52**, 1434-1456.

Cai, M., Kalnay, E., and Z. Toth,2003: Bred Vectors of the Zebiak-Cane Model and their potential application to ENSO Predictions. *J. Clim*., **16**, 40-56.

Cheng, Y., Tang, Y., Zhou, X., Jackson, P., and D. Chen, 2010: Further analysis of singular vector and ENSO predictability in the Lamont model—Part I: singular vector and the control factors. *Clim. Dyn*., **35**, 807-826.

Chikamoto, Y., Mukougawa, H., Kubota, T., Sato, H., Ito, A., and S. Maeda, 2007: Evidence of growing bred vector associated with the tropical intraseasonal oscillation. *Geophys. Res. Lett*., https://doi.org/10.1029%2F2006GL028450

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., and P.

Bechtold, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc*., **137**, 553-597.

Gneiting, T., and A. E. Raftery, 2005: Weather forecasting with ensemble methods. *Sci*., **310**, 248-249.

Graham, R.J., Gordon, M., McLean, P.J., Ineson, S., Huddleston, M.R., Davey, M.K., Brookshaw, A., and R.T.H. Barnes, 2005: A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *TELLUS A*., **57**, 320-339.

Hoffman, R.N., and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *TELLUS A*., **35**, 100-118.

Huffman, G.J., Adler, R.F., Bolvin, D.T., and G. Gu, 2009: Improving the global precipitation record: GPCP Version 2.1. *Geophys. Res. Lett.*, **36**, L17808, doi:10.1029/2009GL040000.

Hudson, D., Marshall, A.G., Yin, Y., Alves, O., and H. H. Hendon, 2013: Improving intraseasonal prediction with a new ensemble generation strategy. *Mon. Weather. Rev*., **141**, 4429-4449.

Kang, I.S., Jang, P.-H., and M. Almazroui, 2014: Examination of multi perturbation methods for ensemble prediction of the MJO during boreal summer. *Clim. Dyn*., **42**, 2627-2637.

Kug, J.S., Ham, Y.G., Kimoto, M., Jin, F.F., and I. S. Kang, 2010: New approach for optimal perturbation method in ensemble climate prediction with empirical

singular vector. *Clim. Dyn*., **35**, 331-340.

Leith, C.E, 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Weather Rev*., **102**, 409-418.

Liu, X.W., Wu, T.W., Yang, S., Li, Q.P., Cheng, Y.J., Liang, X.Y., Fang, Y.J., Jie, W.H., and S.P. Nie, 2014: Relationships between interannual and intraseasonal variations of the Asian-western Pacific summer monsoon hindcasted by the BCC_CSM1.1(m). *Adv. Atmos. Sci.*, **31**, 1051-1064.

Liu, X.W., Wu, T.W., Yang, S., Jie, W.H., Nie, S.P., Li, Q.P., Cheng, Y.J., and X.Y. Liang, 2015: Performance of the Seasonal Forecasting of the Asian Summer Monsoon by BCC_CSM1.1(m). *Adv. Atmos. Sci*., **32**, 1156-1172.

Molteni, F., Buizza, R., Palmer, T.N., and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Q. J. R. Meteorol. Soc*., **122**, 73-119.

Saito, K., Hara, M., Kunii, M., Seko, H., and M. Yamaguchi, 2011: Comparison of initial perturbation methods for the mesoscale ensemble prediction system of the Meteorological Research Institute for the WWRP Beijing 2008 Olympics Research and Development Project (B08RDP). *TELLUS A*., **63**, 445-467.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y-T., Chuang, H.Y., Iredell, M., EK, M., Meng, J., Yang, R.Q., Mendez, M.P., Dool, H.V.D., Zhang, Q., Wang, W.Q., Chen, M.Y., and E. Becker, 2014: The NCEP climate forecast system version 2. *J. Clim*., **27**, 2185-2208.

Stan, C., and B. P. Kirtman, 2008: The influence of atmospheric noise and uncertainty in ocean initial conditions on the limit of predictability in a Coupled GCM. *J. Clim*., **21**, 3487-3503.

Takaya, Y., Yasuda, T., Fujii, Y., Matsumoto, S., Soga, T., Mori, H., Hirai, M., Ishikawa, I., Sato H., Shimpo, A., Kamachi, M., and T. Ose, 2017: Japan Meteorological Agency/Meteorological Research Institute-Coupled Prediction System version 1 (JMA/MRI-CPS1) for operational seasonal forecasting. *Clim. Dyn*., **48**, 313-333.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Weather Rev*., **125**, 297-3319.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *B.AM METEOROL.SOC*., **74**, 2317-2330.

Wu, T.W., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., Zhang, Y., Li, J., Zhang, F., Wei, M., Shi,X., Wu, F., Zhang, L., Chu, M., Jie, W., Liu, Y., Wang, F., Liu, X., Li, Q., Dong, M., Liang,X., Gao, Y., and J. Zhang, 2013: Global carbon budgets simulated by the Beijing climate center climate system model for the last century. *J. Geophys. Res. Atmos*., **118**, 4326-4347.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., Zhang, F., Zhang, Y., Wu, F., Li, J., Chu, M., Wang, Z., Shi, X., Liu, X., Wei, M., Huang, A., Zhang, Y., and X. Liu, 2019: The Beijing Climate Center Climate

System Model (BCC-CSM): Main Progress from CMIP5 to CMIP6, *Geosci. Model Dev.*, **12**, 1573-1600.

Xin, X. G., Wu, T. W., and J. Zhang, 2013: Introduction of CMIP5 experiments carried out with the climate system models of Beijing Climate Center. *Adv. Clim. Change Res.*, **4**, 41-49.

Yang, S.-C., Cai, M., Kalnay, E., Rienecker, M., Yuan, G., and Z. Toth, 2006: ENSO bred vectors in coupled ocean atmosphere general circulation models. *J. Clim.*, **19**, 1422-1436.

Zhang, H., and Z. Pu, 2010: Beating the uncertainties: ensemble forecasting and ensemble-based data assimilation in modern numerical weather prediction. *ADV.METEOROL.*, **2010**, doi:10.1155/2010/432160.

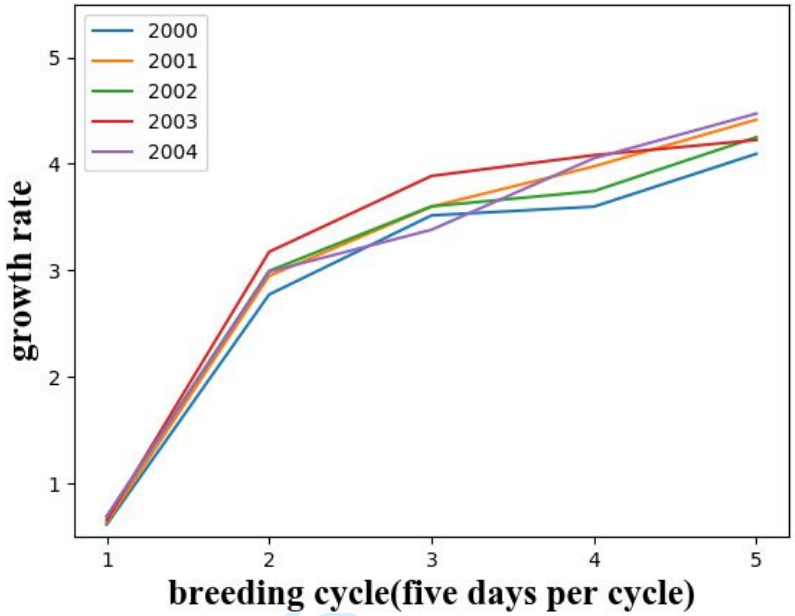

Fig. 1. Growth rates of global mean temperature's perturbations with breeding cycle initiated in April 11 of years 2000-2004.

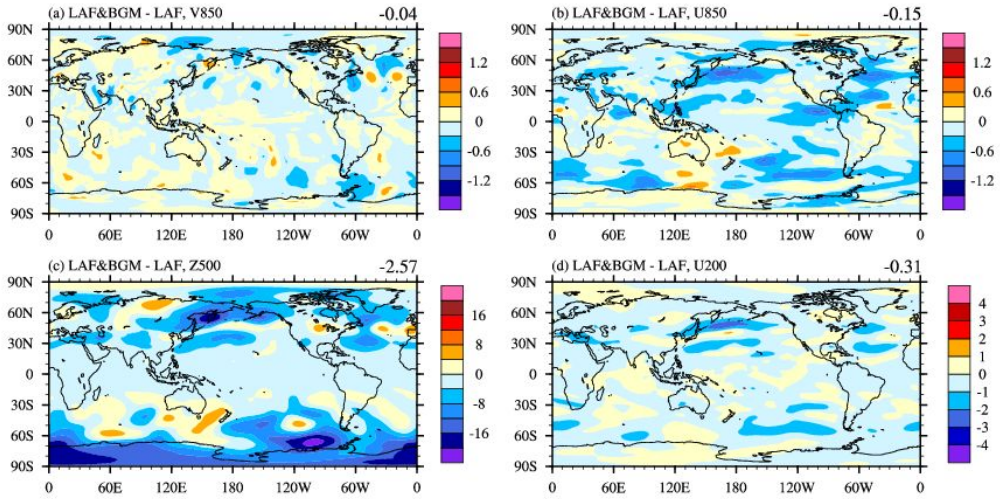

Fig. 2. Difference of RMSEs for 850hPa meridional wind (a), 850hPa zonal wind (b), 500hPa geopotential height (c), and 200hPa zonal wind (d) between LAF&BGM and LAF ensemble hindcast in May.

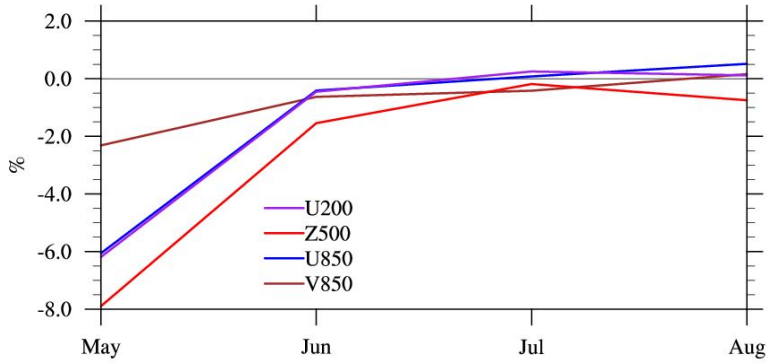

Fig. 3. Change (%) of RMSE for U200, Z500, U850, and V850 averaged over the globe predicted by LAF&BGM ensemble relative to LAF ensemble.

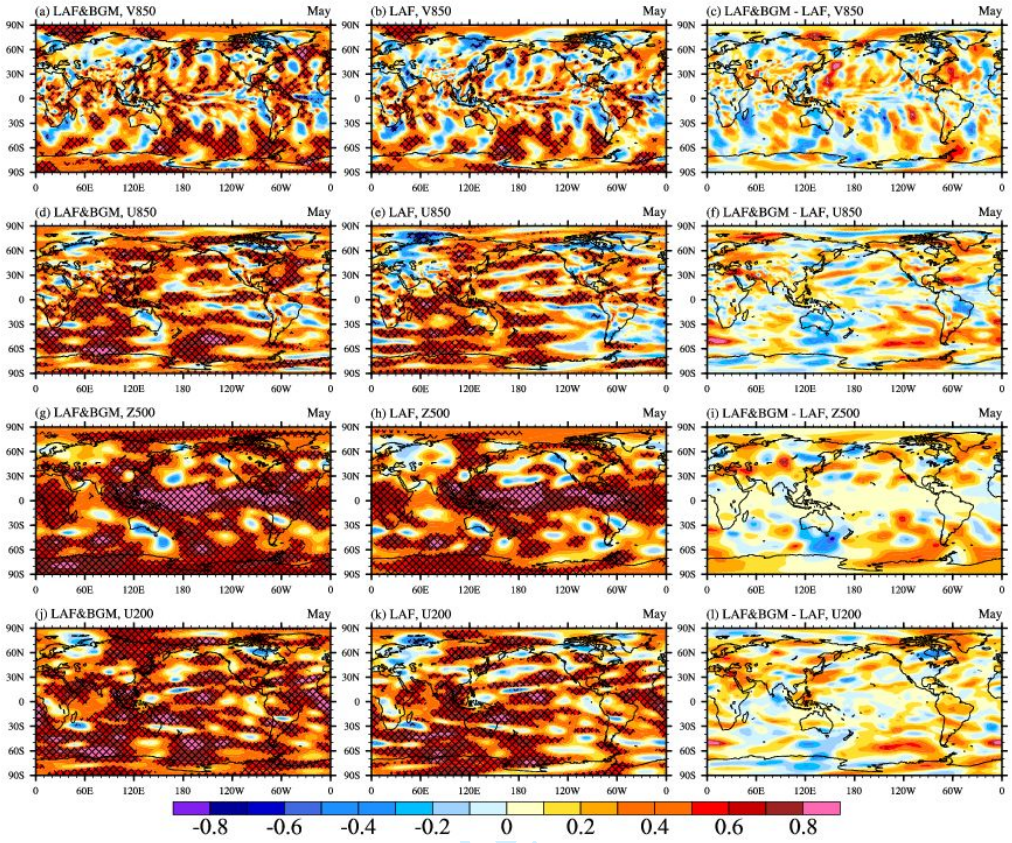

Fig. 4. Temporal Correlation Coefficients (TCCs) between the LAF&BGM ensemble hindcast and the observation (left panel), between the LAF ensemble hindcast and the observation (middle panel) for 850hPa meridional wind, 850hPa zonal wind and 500hPa geopotential height and 200hPa zonal wind in May, and the difference the LAF&BGM and LAF (right panel). The hatched regions are above the 90% confidence level.

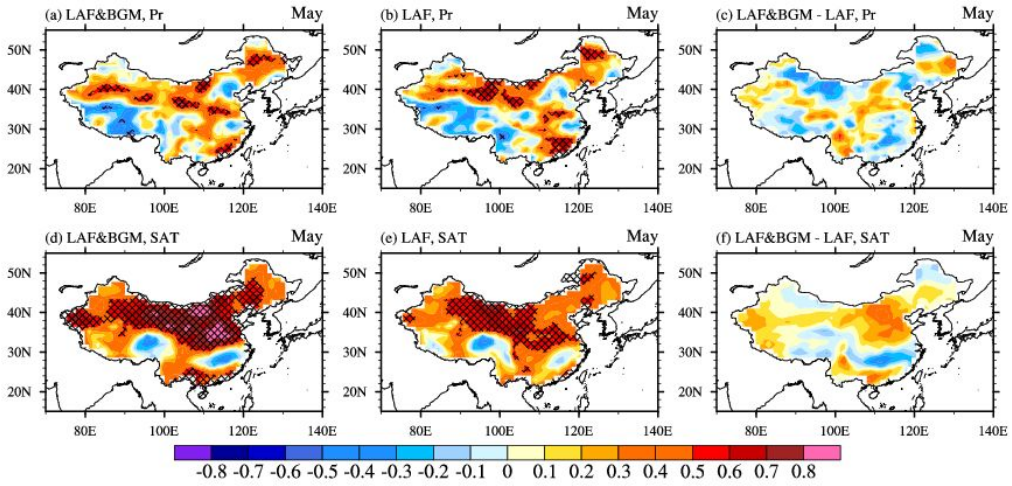

Fig. 5. TCC skills of precipitation (upper panel) and surface air temperature (lower panel) for the LAF&BGM ensemble and LAF ensemble hindcast, and the difference between them.

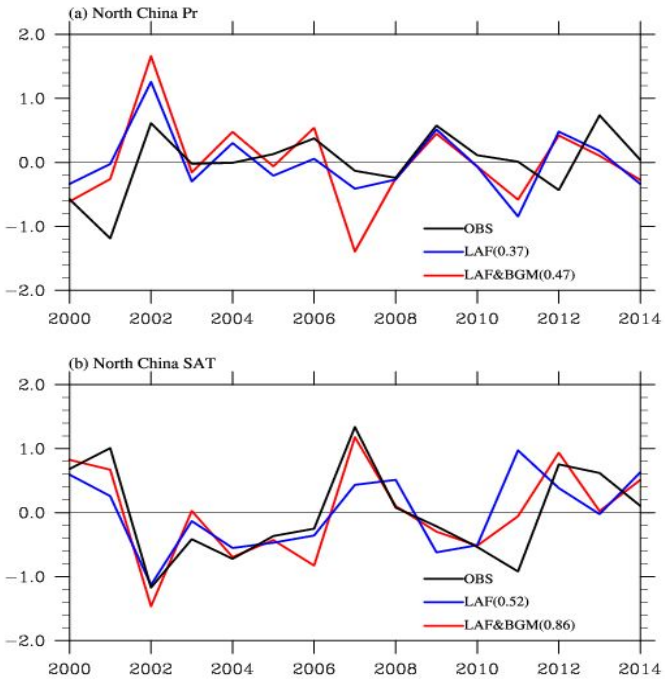

Fig. 6. Precipitation (a) and SAT (b) anomalies in May averaged over North China (105-120°E,32-42°N) during 2000-2014 predicted by LAF&BGM and LAF ensembles. The black line denotes the observation.