# A Dynamical–Statistical Forecasting Model of the Western Pacific Subtropical High Area Index Based on an Improved Self-Memorization Principle

MEI HONG AND REN ZHANG

*Research Center of Ocean Environment Numerical Simulation, Institute of Meteorology and Oceanography, People's Liberation Army University of Science and Technology, Nanjing, China*

DONG WANG

*Key Laboratory of Surficial Geochemistry, Ministry of Education, and Department of Hydrosciences, School of Earth Sciences and Engineering, Collaborative Innovation Center of South China Sea Studies, State Key Laboratory of Pollution Control and Resource Reuse, Nanjing University, Nanjing, China*

MIN WANG AND KEFENG LIU

*Research Center of Ocean Environment Numerical Simulation, Institute of Meteorology and Oceanography, People's Liberation Army University of Science and Technology, Nanjing, China*

VIJAY P. SINGH

*Department of Biological and Agricultural Engineering, and Zachry Department of Civil Engineering, Texas A&M University, College Station, Texas*

(Manuscript received 9 May 2015, in final form 5 August 2015)

ABSTRACT

A new dynamical–statistical forecasting model of the western Pacific subtropical high (WPSH) area index (AI) was developed, based on dynamical model reconstruction and improved self-memorization, in order to address the inaccuracy of long-term WPSH forecasts. To overcome the problem of single initial prediction values, the self-memorization function was introduced to improve the traditional reconstruction model, thereby making it more effective for describing chaotic systems, such as WPSH. Processing actual data, the reconstruction equation was used as a dynamical core to overcome the problem of employing a simple core. The resulting dynamical–statistical forecasting model for AI was used to predict the strength of long-term WPSH forecasting. Based on 17 experiments with the WPSH during normal and abnormal years, forecast results for a period of 25 days were found to be good, with a correlation coefficient of ~0.80 and a mean absolute percentage error of <8%, showing that the improved model produced satisfactory long-term forecasting results. Additional experiments for predicting the ridgeline index (RI) and the west ridge-point index (WI) were also performed to demonstrate that the developed model was effective for the complete prediction of the WPSH. Compared with the authors' previous models and other established models of reasonable complexity, the current model shows better long-term WPSH forecasting ability than do other models, meaning that the aberrations of the subtropical high could be defined and forecast by the model.

## 1. Introduction

Western Pacific subtropical high (WPSH) is one of the most important components of the East Asian summer monsoon (EASM) system. The intensity and position of the WPSH result in complex seasonal evolutions and changes that significantly affect the rainy season in China, including floods, droughts, and heavy rains (Miyasaka and Nakamura 2005). For example, when WPSH reaches its northernmost position, especially in summer, it significantly influences the distribution of rainfall in China and Japan (Kurihara 1989).

Because WPSH plays an important role in the East Asian climate, many scientists have attempted to

*Corresponding author address*: Mei Hong, Research Center of Ocean Environment Numerical Simulation, Institute of Meteorology, PLA University of Science and Technology, Nanjing 211101, China. E-mail: flowerrainhm@126.com

DOI: 10.1175/MWR-D-15-0181.1

forecast its patterns and trends (Park et al. 2010; Chang et al. 2000a). Because the occurrences and aberrations of WPSH are complex processes (Miyasaka and Nakamura 2005), forecasting WPSH, especially abnormal WPSH, remains difficult (Grinsted et al. 2004). Current forecast models can be divided into two main categories: numerical and statistical. The European Centre for Medium-Range Weather Forecasts model (Wu et al. 2009), for example, requires significant numerical forecasting field boundaries and complex computations; it has low efficiency that makes the results unpredictable. Statistical forecasting using neural networks can incorporate historical data but cannot explain the physical mechanisms of the WPSH (Lu et al. 2007). Moreover, the reliability of these models is progressively reduced as forecast time increases; as a result, forecast results and dependability become unreliable beyond a 2-week lead time (Chang et al. 2000b). Both statistical and numerical forecasting models have some degree of bias. Specifically, errors occur in WPSH anomalies and long-term forecasting (He and Gong 2002). As a result, the prediction of unusual activities within a season and the long-term forecasting of WPSH are significantly complicated.

Because of its complex weather system, it is difficult to develop a specific dynamical model to forecast WPSH. A dynamical–statistical model of a weather system is constructed from actual data, which provides a feasible method for describing the physical mechanisms of such a complex system. For the local convergence of errors, Zhang et al. (2006) introduced a genetic algorithm (GA) to improve the estimation of the model parameters. Subsequently, Hong et al. (2013a,b) investigated the reconstruction of a nonlinear dynamical–statistical forecast model of the WPSH index and achieved satisfactory results.

However, the dynamical–statistical prediction equations derived by Zhang et al. (2006) and Hong et al. (2013a,b) depend significantly on the initial values; as a result, long-term forecasts for lead times greater than 15 days diverge significantly from observed values and the model therefore requires improvement. Cao (1993) proposed a self-memorization principle that transforms the dynamical equation into a memorization equation in a broader sense; it is called a differential–integral equation, whereby the memory coefficients can be determined by actual data. This method has been employed to address prediction problems in meteorological, hydrological, and environmental fields (Gu 1998; Feng et al. 2001; Chen et al. 2009). Because this principle avoids the problem of initial conditions in differential equations, its inclusion can hopefully improve the dynamical model forecasting.

The remainder of this paper is organized as follows. Section 2 introduces the data and factors. Three summer monsoon factors with a higher correlation with AI were chosen. The dynamical–statistical model of AI and these three factors are reconstructed in section 3. In section 4, self-memorization dynamics are introduced to improve the reconstructed model. The improved self-memorization functions are investigated and discussed. Forecast experiments from 2010 and further tests of nine additional abnormal WPSH years and eight normal years are described in section 5. The results and conclusions are summarized and discussed in section 6.

## 2. Research data and factors

### *a. Data*

Data from May to October 1962–2011 (50 yr) were obtained from the National Centers for Environmental Prediction Climate Forecast System Reanalysis (CFSR). This dataset included horizontal wind field and geopotential height fields at 850 and 200 hPa, respectively; geopotential height field at 500 hPa; sea level pressure field with a resolution of 0.5° × 0.5°; as well as sensible and latent heat fluxes and precipitation rate in the Gaussian grid. Outgoing longwave radiation (OLR) data from May to October 1962–2011 (50 yr) were obtained from National Oceanic and Atmospheric Administration (NOAA) satellites at a resolution of 0.5° × 0.5°. The unit of measurement was watts per meter squared. These data were primarily used to calculate AI and the three summer monsoon impact factors described in section 2c.

### *b. WPSH activity during the summer of 2010*

Changes in WPSH within seasons in various years were different from average ''abnormal'' activities, which often led to East Asian subtropical circulation anomalies and extreme weather events in China, such as during 1998, 2003, 2006, and 2010 (Lu 2001; Yu et al. 2007). To better describe the changes in WPSH, AI was used by the Central Meteorological Observatory Long-Term Forecasting Group (1976) to characterize the WPSH range and intensity, that is, the number of grid points with 500-hPa height >588 gpm in the range of 10°–90°N, 110°E–180°. The greater the value, the wider the range and the greater the intensity.

The AI of summer half-years (May–October) from 1982 to 2011 was calculated; the changes in AI in different years are shown in Fig. 1. The straight line represents the average AI. AI varied widely in different years, representing the irregular activities of the WPSH;

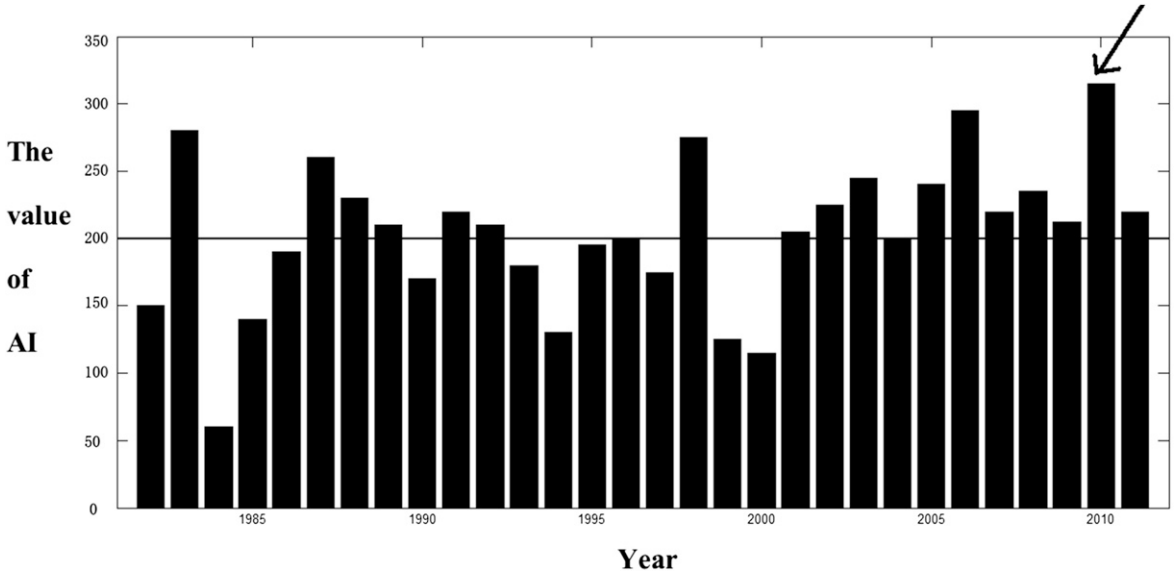


FIG. 1. The AI results of 30 summer half-years (May–October) from 1982 to 2011.

in some years, AI showed dramatic changes. For example, in 2010, the number was 32, while in 1984 it was just 60, indicating that WPSH exhibited an abnormal pattern of behavior during those years. Examining the averaged data from several years, these exceptions were not apparent. As a result, we chose an abnormal AI year for investigation.

As shown by the arrow in Fig. 1, 2010 was the most prominent year for the WPSH anomaly. From May to October 2010, AI was above the mean and was at a 30-yr peak. This anomaly in the WPSH strength resulted in a very unusual climate in China. Extreme heat and heavy precipitation events occurred frequently. The strength and range of those events had rarely been observed in the past. Specifically, the most powerful WPSH occurred in June of 2010, and is to date the strongest occurrence in meteorological records (Fig. 2). This particular WPSH directly resulted in rare flood disasters in the eastern part of south China, across the Yangtze River basin, and in northeastern and northwestern China. As a result, we chose the summer of 2010 for analyzing the relationship between high WPSH and countries affected by the monsoon system.

### *c. Selection of three factors*

According to Lu (2001), Lu and Dong (2001), and Yu et al. (2007), of the many countries affected by the EASM system, 21 are closely affected by the WPSH. Using all countries for modeling would make the equation too complex. Therefore, the correlation between three specific factors and AI was analyzed. Definitions of these factors can be found in Xue et al. (2003) and Yu et al. (2007). Values for these factors were calculated from CFSR data (50 yr), but not described in detail here. Based on correlation analysis, the following three factors were chosen for further study:

1) the Mascarene cold high strength index (MH), which represents the average grid points of sea level pressure within the 20°–10°S, 40°–60°E region;
2) the South China Sea monsoon trough (ST), which shows the average of OLR grid points within 7.5°–17.5°N, 110°–130°E; and
3) the monsoon circulation index for the Bay of Bengal (J1V), which represents the average grid point J1V = V850 − V200 within the 0°–20°N, 80°–100°E region.

From Table 1, it can be seen that the correlation coefficients between the three factors and AI were greater than 0.75. The Mascarene high in the Southern Hemisphere enhanced the WPSH early in the year, which represented a strong relationship with a positive correlation. This finding is consistent with previous research (Sui et al. 2007). The strong relationships between ST, J1V, and AI are also consistent with previous research (Zhou et al. 2009; Yu et al. 2007).

## 3. Reconstruction of the dynamical–statistical model based on GA

Takens (1981) hypothesized and tested the concept of reconstructing a dynamical system from a time series of observed data obtained from his phase-space

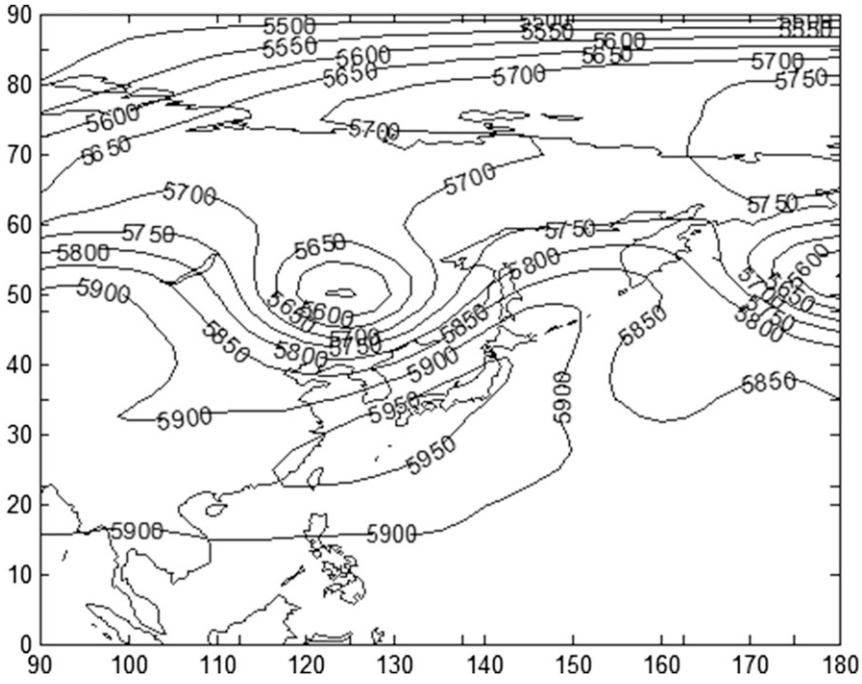


FIG. 2. Average monthly geopotential height field at 500 hPa during July–August 2010.

reconstruction theory. As a result, research into nonlinear dynamics entered a new stage in which the evolution of information from a dynamical system could be extracted from a time series [e.g., calculations of fractal dimensions and Lyapunov exponents (Palmer et al. 2004)]. However, these studies only helped us to understand what chaotic and periodic systems were. To date, dynamical models for practical problems have not yet been fully developed. Huang and Yi (1991) proposed a method for nonlinear dynamic model reconstruction from actual data. They tested it with the Lorenz system with satisfactory results. Therefore, the current study will examine this method and suggest improvements for reconstructing a dynamical–statistical model of AI using the three factors. The principles of dynamical–statistical model reconstruction have been introduced in previous studies (Zhang et al. 2006; Hong et al. 2013a,b) and are discussed in appendix A.

Existing parameter estimation methods, such as neighborhood searches and least squares estimation, are mostly one-way searches that need to travel the entire parameter space, resulting in low search efficiency. Because of the limitations of the error-gradient convergence and its dependence on the initial solutions, parameter estimation is prone to fall into a local optimum rather than a global optimum. GA is a recently developed method that has been used extensively in global optimization methods. It has been found to be excellent in global searches and parallel computing, and the error convergence rate has been greatly improved; therefore, it is helpful in parameter estimation.

We used a simplified second-order nonlinear dynamic model to illustrate the basic characteristics of the atmosphere and ocean interactions, as per Yeh et al. (2009). The time coefficient series of AI, MH, ST, and J1V were represented by $x_1$, $x_2$, $x_3$, and $x_4$, respectively. GA was then introduced for parameter optimization and to reconstruct the dynamical–statistical model.

The definitions of $S, \mathbf{D}, \mathbf{G}$, and $\mathbf{P}$ are given in appendix A. With $S = (\mathbf{D} - \mathbf{GP})^{\mathrm{T}}(\mathbf{D} - \mathbf{GP})$ as a minimum restriction, an optimization solution search was performed in the model parameter space using GA. Supposing that the following nonlinear second-order ordinary differential equations were taken as the dynamical–statistical model of reconstruction, we chose the time coefficient series of AI, MH, ST, and J1V from 1 May to 20 July 2010 as the data for optimizing and retrieving model parameters as follows:

TABLE 1. Correlation analysis between AI and three EASM factors.

| | Factors | | |
|---|---|---|---|
| | MH | ST | J1V |
| AI | 0.77 | −0.83 | 0.86 |

$$\begin{aligned}
\frac{dx_1}{dt} &= a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_1^2 + a_6x_2^2 + a_7x_3^2 + a_8x_4^2 + a_9x_1x_2 + a_{10}x_1x_3 + a_{11}x_1x_4 + a_{12}x_2x_3 + a_{13}x_2x_4 + a_{14}x_3x_4,\\
\frac{dx_2}{dt} &= b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_1^2 + b_6x_2^2 + b_7x_3^2 + b_8x_4^2 + b_9x_1x_2 + b_{10}x_1x_3 + b_{11}x_1x_4 + b_{12}x_2x_3 + b_{13}x_2x_4 + b_{14}x_3x_4,\\
\frac{dx_3}{dt} &= c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_1^2 + c_6x_2^2 + c_7x_3^2 + c_8x_4^2 + c_9x_1x_2 + c_{10}x_1x_3 + c_{11}x_1x_4 + c_{12}x_2x_3 + c_{13}x_2x_4 + c_{14}x_3x_4,\\
\frac{dx_4}{dt} &= d_1x_1 + d_2x_2 + d_3x_3 + d_4x_4 + d_5x_1^2 + d_6x_2^2 + d_7x_3^2 + d_8x_4^2 + d_9x_1x_2 + d_{10}x_1x_3 + d_{11}x_1x_4 + d_{12}x_2x_3 + d_{13}x_2x_4 + d_{14}x_3x_4.
\end{aligned} \tag{1}$$

Assuming that the parameter matrix $\mathbf{P} = [a_1, a_2, \ldots, a_9; b_1, b_2, \ldots, b_9; c_1, c_2, \ldots, c_9]$ in the above equation is the population, the minimal residual sum of squares $S$ is the objective function. The value of the objective function of the $i$th individual is $S_i$, so $l_i = 1/S_i$ is the individual fitness value and $L = \sum_{i=1}^{n} l_i$ is the total fitness value. The idiographic operating steps include coding and creating the initial population, calculating fitness (individually made choices), crossover and mutation, etc. Crossover and mutation are known as the main genetic operators. The calculation method and a detailed explanation can be seen in Goldberg et al. (1989). During calculations, the step length was chosen to be 1 day. After ~37 genetic operations and optimization searches, it was possible to rapidly converge to the target adaptive value and retrieve each optimized parameter of the dynamical–statistical equations. In Eqs. (1), $L_k = (G_kP_k)$ reflects the relative contribution of each item to the system evolution. After eliminating parameters of weak items with little contribution $L_k$, such as $a_5$, $b_3$, etc., the dynamical–statistical model of AI and the three factors were inverted as follows:

$$\begin{cases}
\dfrac{dx_1}{dt} = F_1 = 77.22x_1 - 0.4328x_2 - 260.7114x_3 + 55.1263x_4 + 2.334\times10^{-3}x_2^2 - 3.2139\times10^{-2}x_1x_2 \\
\qquad\qquad + 0.3159x_2x_3 - 1.2266\times10^{-3}x_2x_4,\\
\dfrac{dx_2}{dt} = F_2 = -18.0321x_1 - 1.442\times10^{2}x_3 + 217.418x_4 + 1.886\times10^{-3}x_1x_2 + 0.7651x_2x_3 - 0.0490x_2x_4,\\
\dfrac{dx_3}{dt} = F_3 = -4.2895x_1 - 8.2678x_2 + 56.2180x_3 + 7.1429\times10^{-4}x_2^2 + 3.9078\times10^{-3}x_1x_2 - 6.9302\times10^{-2}x_2x_3,\\
\dfrac{dx_4}{dt} = F_4 = -8.1288x_1 + 3.2854x_2 + 9.0111x_4 + 1.8256x_3^2 + 2.5342x_4^2 - 5.7102x_1x_3 - 8.2295x_3x_4.
\end{cases} \tag{2}$$

However, in dynamical systems "weak terms" can also play a crucial role in the dynamics. Therefore, we should analyze the dynamical characteristics of the model to show the rationality of the above elimination.

First, the Lyapunov exponent spectrum of the improved dynamical–statistical Eqs. (2) of SI and three factors can be obtained. The final Lyapunov exponents were (0.0415, 0.0016, and −0.1238). The largest Lyapunov exponent (LLE) is 0.0415 and the reciprocal of LLE was 24.09. Previous studies (Firdaus and Von Bremen 2002) show that the reciprocal of LLE can represent the maximum predicted time scale of the system, which for our model was about 24.09. This result is in line with the forecast results in section 5. We also calculate the Lyapunov exponent spectrum of Eqs. (1) of the original model in which weak items were not eliminated, which is basically the same as that of Eqs. (2). It indicates that the above elimination does little to the chaotic characteristics and the maximum predicted time scale of the model.

Second, we calculated the equilibrium roots of the original Eqs. (1) and Eqs. (2) and eliminated the weak items. There was little difference between the equilibrium roots of the original Eqs. (1) and those of the Eqs. (2), which indicates the above elimination does little to the dynamical characteristics of the model. Based on the above two-aspect analysis, we can see that the elimination of weak items is reasonable.

The model required testing, so we chose AI, MH, ST, and J1V data from 21 July 2010, which were not used in the modeling as initial forecast data. Next, the Runge–Kutta method was used to carry out numerical integration of the above equations, and every step of the integration was regarded as 1 day's worth of forecasting results. Here, forecast results of four time series over a

period of 25 days were obtained. The correlation coefficients between the forecast and actual results of AI, MH, TH, and J1V for the 25 days were, respectively, 0.5538, 0.5871, 0.6140, and 0.5128. The mean absolute percentage errors (MAPEs), expressed as $\mathrm{MAPE} = 1/n\sum_{i=1}^{n}\{[D_e(i) - D_0(i)]/D_0(i)\} \times 100$ (Hu et al. 2001; Wang et al. 2009), were 44.12%, 37.29%, 35.08%, and 40.55%, respectively. This indicated that the results of 25-day long-term forecast were unsatisfactory for use in the dynamical–statistical reconstruction model. This was because the integration results diverged significantly over time and our initial data for integration were relatively simple. Therefore, improvements in the model were required for better long-term forecasting.

In atmospheric and oceanic research, the correlation coefficient $r$ is often evaluated using a Student's $t$ test in which the test statistic follows a Student's $t$ distribution if the null hypothesis is supported. It can be used to determine if two sets of data are significantly different from each other (David and Gunnink 1997). As a result, for a given significance level $\alpha = 0.05$, our study obtained

$$r_c = \sqrt{\frac{t_\alpha^2}{n-2+t_\alpha^2}} = \sqrt{\frac{1.69^2}{35-2+1.69^2}} = \sqrt{\frac{2.8561}{33+2.8561}} = 0.2822. \tag{3}$$

This indicates that when the correlation coefficient was greater than 0.2822, results were acceptable at a 95% confidence interval. Our correlation coefficients were all >0.2822 and, thus, acceptable.

## 4. Introduction of self-memorization dynamics to improve the reconstructed model

We have shown that the accuracy of forecast results of a dynamical–statistical reconstruction model for 25 days is unsatisfactory. The literature suggests that introducing the principle of self-memorization into a mature model can improve long-term forecasting (Gu 1998; Chen et al. 2009). Hence, a self-memorization principle was introduced to improve our model. Following Cao (1993) and Feng et al. (2001), the mathematical principle of self-memorization dynamics of systems is described in appendix B.

For convenience of calculations, to discretize the self-memorization equation, the integration in Eq. (B9) is replaced by summation and the differential by difference, and the median $x_i^m$ is replaced by the mean of two values at adjoining times; that is, $x_i^m \approx (1/2)(x_{i+1} + x_i) \equiv y_i$.

By sampling an equal time interval, $\Delta t$, that is, $t_i = t_0 + i\Delta t$, where $i = 1, 0, -1, -2, \ldots, -p$, taking an equal time interval $\Delta t_i = t_{i+1} - t_i = 1$, where $t_0$ is the initial time, $t_0 + \Delta t$ becomes the forecast time, and $p$ is the retrospective order. Incorporating $\beta_i$ and $\beta_t$, a discretized self-memorization equation is obtained as

$$x_t = \sum_{i=-p-1}^{-1} \alpha_i y_i + \sum_{i=-p}^{0} \theta_i F(x, i), \tag{4}$$

where $F$ is the dynamic core of the self-memorization equation: $\alpha_i = [(\beta_{i+1} - \beta_i)/\beta_t]$; $\theta_i = \beta_i/\beta_t$.

We call the technique where one performs a forecast and computation based on Eq. (4) a self-memorization principle. Using Eq. (2) as the dynamical kernel, the improved model based on self-memorization Eq. (4) can be expressed as

$$\begin{cases} x_{1t} = \sum_{i=-p}^{0} \alpha_{1i} y_{1i} + \sum_{i=-p}^{0} \theta_{1i} F_1(x_{1i}, x_{2i}, x_{3i}, x_{4i}) \\ x_{2t} = \sum_{i=-p}^{0} \alpha_{2i} y_{2i} + \sum_{i=-p}^{0} \theta_{2i} F_2(x_{1i}, x_{2i}, x_{3i}, x_{4i}) \\ x_{3t} = \sum_{i=-p}^{0} \alpha_{3i} y_{3i} + \sum_{i=-p}^{0} \theta_{3i} F_3(x_{1i}, x_{2i}, x_{3i}, x_{4i}) \\ x_{4t} = \sum_{i=-p}^{0} \alpha_{4i} y_{4i} + \sum_{i=-p}^{0} \theta_{4i} F_4(x_{1i}, x_{2i}, x_{3i}, x_{4i}) \end{cases}. \tag{5}$$

Because we predicted the value of AI, the first formula in Eq. (5) was used to obtain the results of our final prediction as follows:

$$x_{1t} = \sum_{i=-p}^{0} \alpha_{1i} y_{1i} + \sum_{i=-p}^{0} \theta_{1i} F_1(x_{1i}, x_{2i}, x_{3i}, x_{4i}). \tag{6}$$

If we can get the value of $\alpha$, $\theta$, Eq. (6) can be used for prediction. The memory coefficients $\alpha$, $\theta$ in Eq. (6) were calibrated by the least squares method with the same data (1 May–20 July 2010) used in section 3. Equation (6) can be decomposed as follows (in the equations, $M$ is the length of the time series):

$$\mathbf{X} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1M} \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} \alpha_{-p-1} \\ \alpha_{-p} \\ \vdots \\ \alpha_{-1} \end{bmatrix},$$

$$\mathbf{Y} = \begin{bmatrix} y_{-p-1,1} & y_{-p,1} & \cdots & y_{-1,1} \\ y_{-p-1,2} & y_{-p,2} & \cdots & y_{-1,2} \\ \vdots & \vdots & & \vdots \\ y_{-p-1,M} & y_{-p,M} & \cdots & y_{-1,M} \end{bmatrix},$$

TABLE 2. The correlation coefficient (CC) and MAPE results of the fitting and prediction tests when the retrospective order $p$ is different.

| | | $p$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 20 |
| Fitting test | CC | 0.39 | 0.49 | 0.75 | 0.64 | 0.56 | 0.45 | 0.39 | 0.36 | 0.28 | 0.19 |
| | MAPE (%) | 37.40 | 24.34 | 16.22 | 22.07 | 34.70 | 40.34 | 46.71 | 49.22 | 53.21 | 64.21 |
| Prediction test | CC | 0.36 | 0.51 | 0.64 | 0.56 | 0.53 | 0.42 | 0.37 | 0.34 | 0.26 | 0.17 |
| | MAPE (%) | 39.27 | 21.61 | 24.22 | 28.61 | 37.89 | 46.27 | 53.18 | 51.33 | 58.86 | 69.01 |

$$\mathbf{\Theta} = \begin{bmatrix} \theta_{-p} \\ \theta_{-p+1} \\ \vdots \\ \theta_0 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} F_{-p,1} & F_{-p+1,1} & \cdots & F_{0,1} \\ F_{-p,2} & F_{-p+1,2} & \cdots & F_{0,2} \\ \vdots & \vdots & & \vdots \\ F_{-p,M} & F_{-p+1,M} & \cdots & F_{0,M} \end{bmatrix}.$$

The matrix equation is

$$\mathbf{X} = \mathbf{Y}\boldsymbol{\alpha} + \mathbf{F}\boldsymbol{\theta}, \tag{7}$$

where $\mathbf{Z} = [\mathbf{Y}\vdots\mathbf{F}]$ and $\mathbf{W} = \begin{bmatrix} \boldsymbol{\alpha} \\ \vdots \\ \mathbf{\Theta} \end{bmatrix}$.

Equation (7) can be written as

$$\mathbf{X} = \mathbf{Z}\mathbf{W}. \tag{8}$$

The memory coefficients vector $\mathbf{W}$ can be calibrated with the least squares method:

$$\mathbf{W} = (\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X}. \tag{9}$$

The memory coefficients $a$, $\theta$ can be obtained from Eq. (9). Then, we make prediction with the self-memorization in Eq. (6), which uses the $p$ values before $t_0$.

The physical basis for introducing the self-memorization principle is that the thermodynamic equation is one of atmospheric motion equations, which implies that the atmospheric motion is an irreversible process. An important contribution to the study of irreversible processes is the introduction of the memory concept in physics. So the atmospheric development in the future is not only related to the state at the initial time, but also to states in the past, which means the atmosphere does not forget its past.

## 5. Model prediction experiments

### *a. Forecasts of the 2010 experiment*

The period between 21 July and 24 August 2010, was selected for performing forecasts (35 days). In the period from 21 July to 25 August, the strength of AI was the strongest and had drastic changes, so the forecast of AI in this period was more difficult to make than in other periods. Thus, we chose this period to forecast to evaluate our forecasting model. The forecast was obtained by Eq. (6), which is called a step-by-step forecast. The specific procedure is described as follows.

Step-by-step forecasts were performed after the retrospective order $p$ was fixed. That is, when we forecasted the AI value on 21 July, we had to obtain $y_i$ based on the previous $p + 1$ time AI data and $F(x_{1i}, x_{2i}, x_{3i}, x_{4i})$ based on the previous $p$ times for AI, MH, ST, and J1V data, because $x_2$, $x_3$, and $x_4$ were MH, ST, and J1V, respectively. Using them in Eq. (6), we obtained the AI value of 21 July. Next, taking the AI value of 21 July as the initial value for the next prediction step, the AI values of 22 July and so on were generated.

#### 1) DETERMINATION OF $P$

When the principle of self-memorization is introduced, the retrospective order $p$ becomes related to the self-memorization of the system (Cao 1993). If the system slowly "forgets," meaning parameters $\alpha$ and $\theta$ become smaller, then a high value of $p$ should be used. The WPSH abnormal activities are on a 10-day scale (Xue et al. 2003; Yu et al. 2007), which is a slow process compared to large-scale atmospheric movements. As a result, parameters $\alpha$ and $\theta$ were smaller and generally $p$ was in the range of 4–15.

The retrospective order $p$ was determined by trial and error. That is, $p$ = 1, 4, 5, 6, 7, 8, 9, 10, 11, and 20 values were selected to construct the prediction model. The correlation coefficients and MAPE of fitting [from the average results of 20 fitting experiments from different starting points, which all belong to the training sample period (1 May–20 July 2010) and the forecast length was 35 days], as well as the multistep extrapolation prediction (from 21 July to 24 August, 35 days), were chosen as the criterion to filter the retrospective order $p$. The calculation results are shown in Table 2.

It can be observed from Table 2 that when $p$ = 5, the correlation coefficient of both fitting and prediction was the largest and MAPE was the least. Thus, the value of

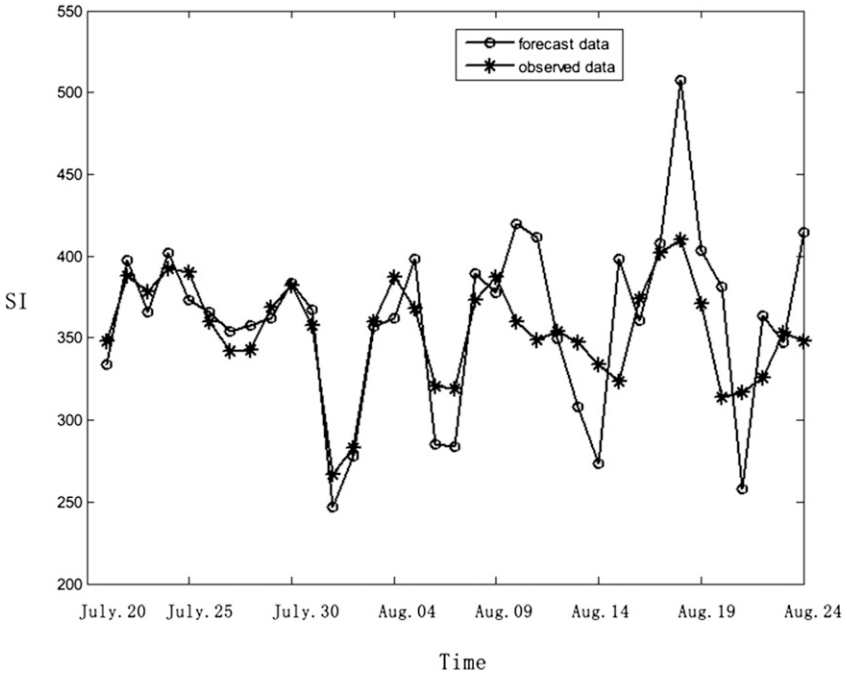


FIG. 3. Results of the subtropical high area index forecasted for 35 days by our model.

retrospective order was selected as $p = 5$. Once $p$ was determined, the improved self-memorization Eq. (6) was used for prediction experiments. If Eqs. (2) from section 3 were regarded as the dynamic kernel $F$, then $dx/dt = F$.

Long-term step-by-step forecasts for 35 days were carried out. A retrospective order of $p = 5$ implies that earlier observational data ($p + 1 = 6$) were used when the forecast began. The forecast results per day were preserved as preliminary data and used for predictions during the next period.

2) PREDICTION RESULTS OF 2010

The forecast results for the 35-day period, shown in Fig. 3, reveal that the forecast results for AI improved. As can be seen, the forecast performance of the first 15 days was better. The correlation coefficient reached 0.9322 and MAPE was 4.37%. Two peaks and one valley for AI were also forecasted accurately. The forecast time series from 15 to 25 days gradually diverged but the trend was still satisfactory. The correlation coefficient reached 0.8654 and MAPE was 8.25%, and the forecast rate remained accurate. The peak AI on 9 August was also accurately forecasted.

After nearly 20 days, MAPE began to increase: the forecast trend remained accurate with a correlation coefficient of 0.7429. The forecast curve from 25 to 35 days was accurate compared with the actual situation. The forecast results within 15 days were significantly better than those from 15 to 25 days. However, after 25 days, the error began to increase significantly. For example, the forecast value of 18 August was nearly 1.3 times greater than the actual value, resulting in a false peak. The MAPE error increased to 24.22% but was still <25%. We can also see from Fig. 3 that the forecast results after 25 days were obviously unsatisfactory, with the occurrence of greater oscillations. The effective forecast period (within 25 days) happened to be in accord with the maximum predicted time scale (24.09 units) calculated in section 3.

To summarize, Fig. 3 shows that short-term forecasts up to 15 days were accurate and the average MAPEs were <10%, which indicated that the reconstruction model could perform accurate predictions of the changing trends in indices.

### b. *Additional forecasting experiments of abnormal WPSH years*

To further test the forecasting performance of the improved model, cross testing of more experiments was performed. As seen in Fig. 1 (section 2b), we chose four other years (1998, 2006, 1987, and 1983) during which the WPSH intensity was abnormally strong (higher AI), and five years (1984, 2000, 1994, 1999, and 1985) during which the WPSH intensity was abnormally weak (lower AI), to carry out the integral forecast experiments of AI. As previously discussed, the reconstructed 2010 model was used to perform cross testing. The

TABLE 3. Correlation coefficients (CC) and MAPE results between forecast values and actual values of different events in abnormal WPSH years.

| Forecast events | Statistical tests | | | | | |
|---|---|---|---|---|---|---|
| | Short term (1–15 days) | | Medium term (16–25 days) | | Long term (26–35 days) | |
| | CC | MAPE (%) | CC | MAPE (%) | CC | MAPE (%) |
| AI bigger event 1 (The AI value of 21 Jun 1998 is selected as the initial value to forecast) | 0.933 | 4.72 | 0.803 | 5.88 | 0.689 | 18.77 |
| AI bigger event 2 (The AI value of 18 Jul 2006 is selected as the initial value to forecast) | 0.912 | 5.17 | 0.833 | 8.80 | 0.701 | 25.11 |
| AI bigger event 3 (The AI value of 8 Jul 1987 is selected as the initial value to forecast) | 0.907 | 4.89 | 0.792 | 9.01 | 0.715 | 21.38 |
| AI bigger event 4 (The AI value of 5 Aug 1983 is selected as the initial value to forecast) | 0.921 | 5.04 | 0.812 | 7.13 | 0.705 | 19.88 |
| AI smaller event 1 (The AI value of 28 Jul 1984 is selected as the initial value to forecast) | 0.901 | 6.55 | 0.788 | 9.06 | 0.688 | 23.70 |
| AI smaller event 2 (The AI value of 29 Jun 2000 is selected as the initial value to forecast) | 0.891 | 5.28 | 0.821 | 9.33 | 0.717 | 20.55 |
| AI smaller event 3 (The AI value of 17 Aug 1994 is selected as the initial value to forecast) | 0.905 | 4.64 | 0.761 | 7.22 | 0.680 | 20.54 |
| AI smaller event 4 (The AI value of 12 Jun 1999 is selected as the initial value to forecast) | 0.842 | 5.07 | 0.812 | 10.14 | 0.738 | 21.53 |
| AI smaller event 5 (The AI value of 11 Jul 1985 is selected as the initial value to forecast) | 0.951 | 2.11 | 0.892 | 6.28 | 0.803 | 15.22 |
| Avg | 0.907 | 4.83 | 0.813 | 8.09 | 0.715 | 20.74 |

forecast results of different periods (1–15 days for the short term, 16–25 days for the medium term, and 26–35 days for the long term) were compared with actual results. Cross-test results are shown in Table 3. It can be observed that the forecast results for the short- and medium-term periods were accurate and those of the long term (>26 days) were acceptable. Results for MH, ST, and J1V were similar to those of AI.

### c. *More forecasting experiments of WPSH normal years*

From section 5b, we can see accurate forecasting results of the WPSH abnormal years. To further test the forecasting performance of the reconstruction model, more experiments of the WPSH normal years were performed. From Fig. 1 discussed in section 2b, we chose eight WPSH normal years: 1986, 1989, 1992, 1995, 1996, 2001, 2004, and 2009. We carried out integral forecasting experiments of these eight WPSH normal years. As in section 5b, the common 2010 model was used to forecast and compare forecast results of different time periods (1–15 days for the short term, 16–25 days for the medium term, and 26–35 days for the long term) with actual observations. All forecast results of the eight years in the short term and the medium term were accurate and MAPE values were less than 15%. Although all forecast results of the eight years in the long term (>26 days) were not accurate, MAPE values were still under 30%, which can be accepted. The results of MH, ST, and J1V were similar to those of AI.

### d. *Comparison of three types of forecasting experiments*

From the tests in sections 3, 5b, and 5c, the correlation coefficients and MAPE results of the three types of tests were different. The average correlation coefficients and MAPE results between the forecast values and actual values of three types were changing with the forecasting time, as shown in Fig. 4. The values of correlation coefficients and MAPEs were the average of 20 different forecast experiments of three types. The forecasting results of the initial dynamical–statistical reconstruction model were unsatisfactory. That is why we introduced the improved self-memorization principle. The forecasting results of normal years were worse than those of abnormal years. The main reason is that the parameters of the forecasting model were calculated from the data of 2010, as was mentioned in sections 3 and 4. The year 2010 was a WPSH abnormal year, which meant the parameters of the forecasting model contained sufficient information on the abnormal WPSH activity, so our model was more sensitive to the WPSH abnormal years than normal years. But the average MAPE of the WPSH normal years forecasting test in the short term was less than 10% and that in the long term was less than 30%. The forecast

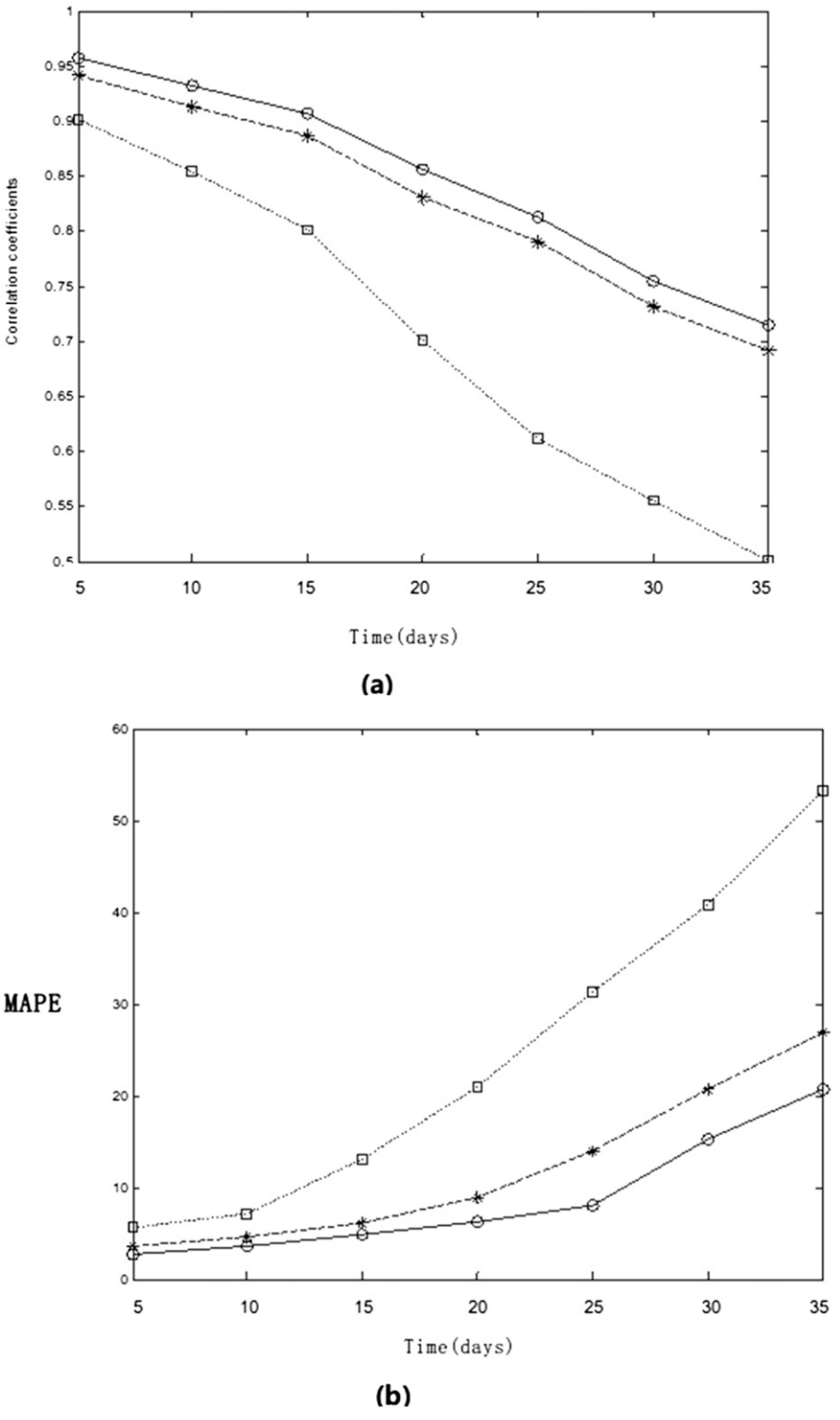


FIG. 4. The changes in (a) average correlation coefficients and (b) MAPE results between forecast values and actual values of three types of forecasting experiments with forecast time. (Note that the circle represents the experiments in abnormal years forecasted by the improved model, introducing the self-memorization principle; the star represents the experiments in normal years forecasted by the improved model introducing the self-memorization principle; and the square represents the experiments in abnormal years forecasted by the initial dynamical–statistical reconstruction model.)

TABLE 4. The average correlation coefficient (CC) and MAPE results for AI, RI, and WI predictions.

| Prediction events | Short term (1–15 days) | | Medium term (16–25 days) | | Long term (26–35 days) | |
|---|---|---|---|---|---|---|
| | Avg CC | Avg MAPE (%) | Avg CC | Avg MAPE (%) | Avg CC | Avg MAPE (%) |
| AI | 0.907 | 4.83 | 0.813 | 8.09 | 0.715 | 20.74 |
| RI | 0.805 | 7.62 | 0.712 | 12.66 | 0.616 | 28.51 |
| WI | 0.901 | 5.02 | 0.801 | 9.88 | 0.701 | 23.11 |

results still can be accepted, which shows our model had good overall utility.

*e. More forecasting experiments of the ridgeline index and west ridge-point index*

The complete prediction of WPSH was not only related to AI but also to the ridgeline index (RI) and the west ridge-point index (WI), because RI usually represents the northward and southward positions of the WPSH and WI usually represents the eastward and westward positions of the WPSH. Once the three indices were forecasted accurately, complete predictions of the WPSH could be performed. RI and WI were defined by the Central Meteorological Observatory Long-Term Forecasting Group (1976) as follows: the larger the value of RI, the more northward the position of WPSH, and the smaller the value of WI, the more westward the position.

Because the correlation coefficients among AI, RI, and WI were all less than 0.5, if we establish one model containing AI, RI, and WI, the forecast results may not be good. Therefore, we should establish three models to, respectively, forecast AI, RI, and WI. The same method can be used for RI and WI prediction during abnormal years based on historical data. The same method was used, but the EASM factors and the forecasting models of RI and WI were different from those of AI. Similar to forecasting AI, we established a model of RI and the Mascarene cold high (MH) index, the Somali low-level jet (SJ) index, and the Indian monsoon latent heat flux (FLH) index based on the historical data of 1998, which had the lowest value of RI in 30 yr. We chose four years (1992, 1999, 1986, and 2002) during which the WPSH ridge showed an abnormal southerly drop (smaller RI). We chose five additional years (2010, 1991, 1989, 1985, and 1982) during which the WPSH ridge showed an abnormal northerly jump (larger RI) to conduct our integral forecast experiments of RI.

We also established a model of WI and the Australian high (AH) index, the sensible heat in the northern Indochina Peninsula (IPS) index, and the cross-equatorial flow South China Sea (SSE) index based on the historical data of 2006, which had the lowest value of WI among 30 years. We chose four additional years (2010, 2003, 1989, and 1988) during which the ridge showed an abnormal westward extension (smaller WI) and five years (1984, 1985, 2008, 1994, and 1995) when the ridge showed an abnormal eastward retreat (larger WI) to conduct our integral forecast experiments of WI. The average correlation coefficients and MAPE for RI and WI were compared with those of AI, as shown in Table 4.

It can be seen that the forecast results for WI were similar to those of AI and results within 25 days were better, while the results for RI were significantly worse than those for AI and WI. MAPE for RI reached about 15% within 25 days. This was because the changes in RI were more complex, such as the phenomenon of ''double ridgeline'' (Qi et al. 2008). Therefore, it was more difficult to predict RI than the other two indices. Over 25 days, the overall forecast results of the three indices were good, indicating that our method was able to predict the total WPSH.

## 6. Summary and discussion

*a. Summary*

Although in recent years WPSH has become one of the leading topics of discussion in the atmospheric sciences, it is still difficult to forecast. Combining dynamical–statistical system reconstruction with an improved self-memorization principle, a new dynamical–statistical forecasting model of the AI index was developed. This paper used the following four-step approach:

1) We used correlation analysis to explore the contributions and impacts of abnormal WPSH activity in 2010 on key countries affected by the EASM and identified three significant factors: the Mascarene high, the Tibetan high (eastern type), and the monsoon circulation index for the Bay of Bengal.
2) We examined AI, MH, XZ, and J1V time series and considered them as trajectories of a set of four coupled quadratic differential equations based on the dynamic system reconstruction approach. Parameters of this dynamical–statistical model were estimated by a genetic algorithm.
3) We applied the "principle of self-memorization" to improve the forecasting results of the dynamical–statistical

TABLE 5. The minimum and the maximum from 40 different forecast experiments forecasted by two different models.

| Model | Short term (1–15 days) | | Medium term (16–25 days) | | Long term (26–35 days) | |
|---|---|---|---|---|---|---|
| | Min MAPE | Max MAPE | Min MAPE | Max MAPE | Min MAPE | Max MAPE |
| Improved model from this paper | 3.18 | 5.74 | 6.22 | 9.78 | 18.17 | 22.86 |
| Hong et al.'s (2015) model | 3.71 | 12.55 | 7.44 | 18.91 | 20.08 | 35.45 |

model. In addition, the dynamical characteristics of the improved model have been discussed, such as LLE and the equilibrium, which can show that our dynamical model is indeed a chaotic system and reflects the dynamical regularity of the subtropical high.

4) The improved model was used to forecast the AI index. Based on experiments during 2010 and nine other experiments on WPSH during abnormal years, we found that the forecast results for 25 days were good, which showed that our improved model had better long-term forecasting accuracy. Additional experiments for predicting RI and WI were also performed to demonstrate that our method was effective for the complete prediction of WPSH. Given the complexity of mechanisms driving the WPSH, the new dynamical–statistical forecasting model shows scientific significance and practical value.

We know that the forecast results and credibility of previous methods are very low after 2 weeks (Zhou et al. 2009). As a result, our improved dynamical–statistical model represents an exploration of and supplement to traditional numerical and statistical forecasting methods, as well as extending the prediction time.

### b. Discussion

#### 1) COMPARISON WITH OUR PREVIOUS STUDIES

Hong et al. (2013b) have established the nonlinear forecast model of the subtropical high activity and aberration and forecasted the 500-hPa geopotential field data of August 2010. Hong et al. (2015) also constructed a nonlinear statistical–dynamical model of WPSH and carried out the forecast experiments of AI. Compared with our previous studies, this paper introduces the principle of self-memorization dynamics to improve the traditional dynamical reconstruction theory. The outstanding improvements are as follows:

1) The long-term forecast results are better than those of our previous studies. The forecast of our previous studies only depends on the integral equation. With the increase in integral forecasting time, there will be a significant divergence, which will lead to ineffective long-term forecasts. To improve the forecast accuracy, we introduce the principle of self-memorization dynamics. Thus, the forecasts in our model not only depend on the integral equation, but also depend on the earlier $p + 1$ observed data, which leads to the more accurate forecast results of our model in the long term.

2) The forecast results are more stable than those of our previous studies. The forecast results of our previous studies depend on a single initial value. Choosing a different initial value will cause differing degrees of forecast accuracy, and the differences among the forecast accuracies are significant. We carried out 40 different forecast experiments from different forecast initial values using the improved model in this paper and the model of Hong et al. (2015), as shown in Table 5.

   It can be seen in Table 5 that the differences among the MAPE results from our improved model are much smaller than those of Hong et al. (2015). This indicates that the forecasts that depend on a single initial value will cause instability in the forecast results. The forecast of the improved model depends on the earlier $p + 1$ observed data instead of a single initial value by introducing the self-memorization principle, which makes the forecast results more stable.

3) The forecast results of the WPSH abnormal activities are better than those found by previous studies. For example, we carried out 40 different forecast experiments within 35 days using the improved model in this paper and the model of Hong et al. (2015). There are 126 peak values and 114 valley values in total. The improved model of this paper has forecasted 113 peak values and 102 valley values, while the model of Hong et al. (2015) forecasted 98 peak values and 84 valley values. This indicates that for the abnormal increased peak values and the abnormal decreased valley values, the forecast results of the improved model are better than those of the previous studies.

4) This paper has discussed the dynamical characteristics of the model, such as LLE and the equilibrium, which can reflect the chaotic characteristics and the maximum predicted time scale of the model.

5) The previous studies only forecasted the AI values. This paper has added the forecast experiments of the RI and WI values, which indicate that our method was able to predict the whole WPSH.

### 2) COMPARISON WITH OTHER PREVIOUS STUDIES

Current forecast models of WPSH can be divided into two main categories: numerical and statistical.

The numerical forecasts of the subtropical high mainly depend on the products of the numerical forecast models. Wang et al. (2007) analyzed the prediction results of subtropical high characteristic indices based on the products of the ECMWF model from June to August 2006. Results indicated that the errors of subtropical high prediction based on the ECMWF model within 120 h were small, while the error within 144–192 h increased noticeably. A number of researchers (Wang et al. (2002); Fu 2012; Zhang and Zhang 2011; Jiang and Cai 2011) also carried out performance verification simulations of the medium-range forecasting via the T639 and ECMWF, GFS, and the Japan Meteorological Agency (JMA) models. Results show that the ability of the ECMWF model to forecast the WPSH is the best among all schemes considered. For example, using the prediction results of the 500-hPa geopotential field to forecast the west ridge-point index (WI), the correlation coefficient between the forecast results and the observations based on the ECMWF model within 144 h (6 days) is over 0.8 and MAPE is under 10%. However, the error significantly increases after 240 h (10 days). The correlation coefficient is about 0.6 and the MAPE is around 28%.

From the above research, it can be seen that the numerical forecast results of the subtropical high in the short term (10 days) are commonly good, but the forecast results in the medium and long term (>10 days) are not very good. The main reasons are as follows. 1) The numerical model mainly forecasts the 500-hPa geopotential field and does not establish a model of the subtropical high index to forecast especially well, as does our paper. So the long-term forecast results of the subtropical high index transformed by the numerical prediction products are not good. 2) In practical applications, the forecast results of the numerical model are restricted by many aspects, such as the initial field, the boundary conditions, the physical process, the terrain, the vegetation, and the design of model, which inevitably cause some degree of error. The combination of these aspects also leads to the corresponding error of the subtropical high index forecast based on the numerical prediction model. 3) For forecasts of large-scale weather systems, we cannot completely rely on the numerical prediction product. The chaotic characteristics of the subtropical high should also be considered (Liu et al. 2013), which can lead to the correct judgment on the future change of the weather system. The dynamical characteristics of our model reflect the chaotic characteristics of the subtropical high (section 3), which may well represent the changes within the subtropical high system in the future.

For statistical forecasts of the WPSH, Liu et al. (2007) established a forecast model of the WPSH area index based on the method of wavelet decomposition and a support vector machine. The forecast within 10 days was carried out and the forecast results were good. The average correlation coefficient was around 0.88, but fails the long-term prediction test after 10 days. Liu et al. (2008) constructed a forecasting model of the subtropical high based on the combination of a radial basis function (RBF) neural network and hybrid-hierarchy-based algorithm, using the subtropical high area index data from 1995 and 1996. When the prediction experiment for AI in 1997 was carried out, the correlation coefficient within 20 days was 0.8134 and MAPE was 26.5%. The forecast results were better than those of T213. Xu et al. (2007) established a forecast model of the WPSH area index based on self-organizing feature map (SOFM) and artificial neural network (ANN) concepts. The forecast results were better than those of the T639 numerical model forecast. The correlation coefficient was 0.8349 within 10 days, while the long-term forecast results from 10 to 25 days turned bad quickly, and could only reach a correlation coefficient of about 0.65.

The forecast results of subtropical high statistical models are similar to those of numerical models. The forecast results are good within 10 days. But the prediction errors from 10 to 25 days increase quickly, which is around 30%. That is because the statistical algorithms, such as neural networks, mainly rely on the empirical means based on a large amount of data. The network structure and error convergence algorithm still have room for improvement. Especially, the statistical models cannot reflect the dynamical characteristics of the subtropical high system, so it cannot make an accurate judgment on the medium- and long-term changes of the subtropical high system.

Some researchers in recent years have also tried to use dynamical–statistical methods to forecast the subtropical high index. For example, Wang et al. (2006) used Kalman filtering to revise errors and optimize the forecasts made by T213 numerical forecast products. Experiments showed that the forecast results within 10 days were good and the average error was about 20%. However, the prediction had a significant divergence after 10 days. Yang et al. (2012) established an optimal dynamical multifactor scheme to predict the subtropical high in the western Pacific based on the basic principle of the dynamical–statistical analog prediction. The

forecast tests of AI, north ridge-point index, and WI of the subtropical high were carried out. But the correlation coefficient between the forecast results and actual data after eight prediction steps was only about 0.5. In recent years, dynamical–statistical forecasting methods have been mainly confined to using statistical methods to revise the numerical forecast products. Although some improvements have been made, these upgrades do not fundamentally solve the three basic problems of numerical prediction models forecasting the subtropical high. So the forecast results of the subtropical high index in the medium and long terms still need to be improved.

Based on the above discussion and the results in Tables 3 and 4, it can be seen that the forecast results of the subtropical high index in the medium and long terms (>10 days) based on our model are better than those of previous studies. One may ask, why are the forecasting results of our improved model particularly good? 1) From the above discussion, it can be seen that our dynamical–statistical model is mainly established to forecast AI, the target of which is clearer than in conventional numerical models. Based on the analysis in section 3, the dynamical characteristics of our model can satisfactorily reflect the chaotic regulation of the subtropical high. So our model can represent the future changes of the subtropical high system. From the analysis of LLE, the basic forecast length of our model can reach 20–25 days, which conforms to the forecast tests in section 5. 2) Previous studies have shown that long-term forecasting results of reconstructed dynamical–statistical system models are unsatisfactory; thus, we introduced the self-memorization principle to improve long-term forecasting results. Previous studies have shown that this approach is feasible. For example, Gu (1998) used the self-memorization principle to improve the traditional T42 model, and Wang et al. (2012) carried out dynamical predictions of building subsidence deformations with a self-memorization model. Their long-term forecasting models showed good results. Our study also has shown that introducing a self-memorization principle to improve a mature model can provide better long-term forecasting results. 3) While developing this improved model, we used parameters from historical data that contained additional information on abnormal WPSH processes. We combined statistical data with actual results from our improved dynamical–statistical model to show that forecasting can be made more reliable.

The final number of 27 parameters in Eq. (2) is a bit large from a sample size of 80. It is necessary to discuss whether our model has the overfitting problem or not. By comparing the fitting results of the training set and the testing set, and while noting that the results of the testing set are slightly worse than those of the training set, the difference is not big. Thus, we feel our model has no obvious overfitting phenomena.

Although the forecast results of our improved model were better than those of traditional models, there remain some issues that require further research. First, the physical factors in our model were not clearly defined; as a result, their dynamical characteristics should be further analyzed. Second, since forecast accuracy is closely related to self-memorization functions, the development of a better self-memorization function that improves forecast accuracy in the long term should be pursued. These will be the focus of our next study.

*Acknowledgments.* We thank the three anonymous reviewers and the editor Josh Hacker for their thoughtful comments and suggestion. This study was supported by the Chinese National Natural Science Fund (41375002, 41075045), the Chinese National Natural Science Fund for Young Scholars (41005025, 41301370), and the Chinese National Natural Science Fund (BK2011123) of Jiangsu Province; the Program for New Century Excellent Talents in University (NCET-12-0262); the China Doctoral Program of Higher Education (20120091110026); the Qing Lan Project; the Skeleton Young Teachers Program and the Excellent Disciplines Leaders in Midlife-Youth Program of Nanjing University; the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province of China under Grant 13KJB110022; and the Scientific Research Foundation of Nanjing University of Posts and Telecommunications under Grant NY213052.

## APPENDIX A

### Principle of Dynamical Model Reconstruction

Suppose that the finite-difference form of the physical law of any nonlinear system evolving with time can be expressed as

$$\frac{q_i^{(j+1)\Delta t} - q_i^{(j-1)\Delta t}}{2\Delta t} = f_i(q_1^{j\Delta t}, q_2^{j\Delta t}, \ldots, q_1^{j\Delta t}, \ldots, q_N^{j\Delta t})$$
$$j = 2, 3, \ldots, M-1, \tag{A1}$$

where $f_i$ is the generalized nonlinear function of $q_1, q_2, \ldots, q_i, \ldots, q_N$, $N$ is the number of state variables, and $M$ is the length of the time series of the observed data. We assume that $f_i(q_1^{j\Delta t}, q_2^{j\Delta t}, \ldots, q_i^{j\Delta t}, \ldots, q_N^{j\Delta t})$ contains two parts: $G_{jk}$, representing the expanding items containing variable $q_i$, and $P_{ik}$, which simply represents the corresponding parameters that are real numbers ($i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, M$, $k = 1, 2, \ldots, K$).

It can be supposed that

$$f_i(q_1, q_2, \ldots, q_n) = \sum_{k=1}^{K} G_{jk} P_{ik}. \tag{A2}$$

The matrix form of Eq. (A2) is $\mathbf{D} = \mathbf{GP}$, in which

$$\mathbf{D} = \begin{Bmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{Bmatrix} = \begin{Bmatrix} \dfrac{q_i^{3\Delta t} - q_i^{\Delta t}}{2\Delta t} \\ \dfrac{q_i^{4\Delta t} - q_i^{2\Delta t}}{2\Delta t} \\ \vdots \\ \dfrac{q_i^{M\Delta t} - q_i^{(M-2)\Delta t}}{2\Delta t} \end{Bmatrix},$$

$$\mathbf{G} = \begin{Bmatrix} G_{11}, G_{12}, \ldots, G_{1K} \\ G_{21}, G_{22}, \ldots, G_{2,K} \\ \vdots \\ G_{M1}, G_{M2}, \ldots, G_{M,K} \end{Bmatrix}, \quad \mathbf{P} = \begin{Bmatrix} P_{i1} \\ P_{i2} \\ \vdots \\ P_{iK} \end{Bmatrix}. \tag{A3}$$

Coefficients of the above generalized unknown equation can be identified by inverting the observed data. Given a vector $\mathbf{D}$, the vector $\mathbf{P}$ can be solved to satisfy the above equation. It is a nonlinear system with respect to $q$; however, it is a linear system with respect to $\mathbf{P}$ (assuming $\mathbf{P}$ is unknown). So the classical least squares method can be introduced to estimate the equation, and the regular equation $\mathbf{G}^{\mathrm{T}}\mathbf{GP} = \mathbf{G}^{\mathrm{T}}\mathbf{D}$ can be derived by making the residual sum of squares $S = (\mathbf{D} - \mathbf{GP})^{\mathrm{T}}(\mathbf{D} - \mathbf{GP})$ minimum.

Based on the above approach, coefficients of the nonlinear dynamical systems can be determined and the nonlinear dynamical equations of observed data can be established.

## APPENDIX B

### Mathematical Principle of Self-Memorization Dynamics of Systems

In general, dynamical equations of a system can be written as

$$\frac{\partial x_i}{\partial t} = F_i(x, \lambda, t), \quad i = 1, 2, \ldots, J, \tag{B1}$$

where $J$ is an integer, $x_i$ is the $i$th variable of the system state, and $\lambda$ is the parameter. Equation (B1) expresses the relationship between a local change in $x$ and a source function $F$. Obviously, $x$ is a scalar function at space $r_0$ and time $t$. Consider a set of time $T = [t_{-p}, \ldots, t_0, \ldots, t_q]$, where $t_0$ is an initial time, and a set of space $R = [r_a, \ldots, r_i, \ldots, r_\beta]$, where $r_i$ is a spatial point considered. An inner product in space $L^2$: $T \times R$ is defined by

$$(f, g) = \int_a^b f(\xi) g(\xi)\, d\xi, \quad \text{where} \quad f, g \in L^2. \tag{B2}$$

Accordingly, we define a norm

$$\|f\| = \left\{ \int_a^b [f(\xi)^2\, d\xi] \right\}^{1/2}.$$

Making a completion for $L^2$, it becomes a Hilbert space $H$. A solution to the multitime model can be regarded as a generalized one in $H$. Dropping $i$ in Eq. (B1) and applying an operation of the inner product defined in Eq. (B2) by introducing a memorization function $\beta(r, t)$, we obtain

$$\int_{t_0}^{t} \beta(\tau) \frac{\partial x}{\partial \tau} d\tau = \int_{t_0}^{t} \beta(\tau) F(x, \tau)\, d\tau, \tag{B3}$$

where $r$ in $\beta(r, t)$ is dropped by fixing on the spatial point $r_0$. Suppose that the variable $x$ and function $\beta(r, t)$ etc. are all continuous, differentiable, and integrable, then following the calculus and making an integration by parts for the left of Eq. (B3) yields

$$\int_{t_0}^{t} \beta(\tau) \frac{\partial x}{\partial \tau} d\tau = \beta(t)x(t) - \beta(t_0)x(t_0) - \int_{t_0}^{t} x(\tau)\beta'(\tau)\, d\tau, \tag{B4}$$

where $\beta'(t) = \partial\beta(t)/\partial t$. Applying the median theorem in calculus to the third term on the right-hand side of Eq. (B4), we obtain

$$-\int_{t_0}^{t} x(\tau)\beta'(\tau)\, d\tau = -x^m(t_0)[\beta(t) - \beta(t_0)], \tag{B5}$$

where the median $x^m(t_0) \equiv x(t_m)$, and $t_0 < t_m < t$. Substituting Eqs. (B4) and (B5) into Eq. (B3) and performing an algebraic operation, we obtain

$$x(t) = \frac{\beta(t_0)}{\beta(t)} x(t_0) + \frac{\beta(t) - \beta(t_0)}{\beta(t)} x^m(t_0) + \frac{1}{\beta(t)} \int_{t_0}^{t} \beta(\tau) F(x, \tau)\, d\tau, \tag{B6}$$

Since the first and second terms in Eq. (B6) relate to the $x$ value only on the fixed point $r_0$ itself at the initial time $t_0$ and middle time $t_m$, they are therefore called a self-memory term. Naturally, we call the third term an exogenous effect (i.e., the total effect contributed by other spatial points to point $r_0$ in an interval $[t_0, t]$).

For multitime $t_i$, $i = -p, -p+1, \ldots, 0, 1$, similar to Eq. (B4), we have

$$\int_{t_{-p}}^{t_{-p+1}} \beta(\tau)\frac{\partial x}{\partial \tau}d\tau + \int_{t_{-p+1}}^{t_{-p+2}} \beta(\tau)\frac{\partial x}{\partial \tau}d\tau + \cdots + \int_{t_0}^{t} \beta(\tau)\frac{\partial x}{\partial \tau}d\tau = \int_{t_{-p}}^{t} \beta(\tau)F(x,\tau)\,d\tau.$$

By eliminating the same term $\beta(t_i)x(t_i)$, $i = -p+1, -p+2, \ldots, 0$, it gives

$$\beta(t)x(t) - \beta(t_{-p})x(t_{-p}) - \sum_{i=-p}^{0}[\beta(t_{i+1}) - \beta(t_i)]x^m(t_i) - \int_{t_{-p}}^{t} \beta(\tau)F(x,\tau)\,d\tau = 0. \tag{B7}$$

For simplicity, we set $\beta_t \equiv \beta(t)$, $\beta_0 \equiv \beta(t_0)$, $x_t \equiv x(t)$, and $x_0 \equiv x(t_0)$; similar notation is used in the following context. Then, Eq. (B7) can be written as

$$\beta_t x_t - \beta_{-p}x_{-p} - \sum_{i=-p}^{0} x_i^m(\beta_{i+1} - \beta_i) - \int_{t_{-p}}^{t} \beta(\tau)F(x,\tau)\,d\tau = 0. \tag{B8}$$

Setting $x_{-p} \equiv x_{-p-1}^m$, $\beta_{-p-1} = 0$, we rewrite Eq. (B8) as

$$x_t = \frac{1}{\beta_t}\sum_{i=-p-1}^{0} x_i^m(\beta_{i+1} - \beta_i) + \frac{1}{\beta_t}\int_{t_{-p}}^{t} \beta(\tau)F(x,\tau)\,d\tau = S_1 + S_2. \tag{B9}$$

We call $S_1$ a self-memorization term and $S_2$ an exogenous effect term.

## REFERENCES


Cao, H.-X., 1993: Self-memorization equation in atmospheric motion. *Science China*, **36B**, 845–855.

Central Meteorological Station Long-Term Forecasting Group, 1976: The technology experience of the long-term weather forecast (appendix) (in Chinese). Central Meteorological Observatory, Beijing, China, 233 pp.

Chang, C.-P., Y. Zhang, and T. Li, 2000a: Interannual and interdecadal variations of the East Asian summer monsoon and tropical Pacific SSTs. Part I: Role of the subtropical ridge. *J. Climate*, **13**, 4310–4325, doi:10.1175/1520-0442(2000)013<4310:IAIVOT>2.0.CO;2.

——, ——, and ——, 2000b: Interannual and interdecadal variations of the East Asian summer monsoon and tropical Pacific SSTs. Part II: Meridional structure of the monsoon. *J. Climate*, **13**, 4326–4340, doi:10.1175/1520-0442(2000)013<4326:IAIVOT>2.0.CO;2.

Chen, X., J. Xia, and Q. Xu, 2009: Differential Hydrological Grey Model (DHGM) with self-memory function and its application to flood forecasting. *Sci. China Ser. E: Technol. Sci.*, **52**, 1039–1049, doi:10.1007/s11431-008-0320-5.

David, H. A., and J. L. Gunnink, 1997: The paired *t* test under artificial pairing. *Amer. Stat.*, **51**, 9–12, doi:10.1080/00031305.1997.10473578.

Feng, G., H. Cao, X. Gao, D. Wenjicet, and J. Chou, 2001: Prediction of precipitation during summer monsoon with self-memorial model. *Adv. Atmos. Sci.*, **18**, 701–709.

Firdaus, E. U., and H. F. von Bremen, 2002: Computation of Lyapunov characteristic exponents for continuous dynamical systems. *Z. Angew. Math. Phys.*, **53**, 123–146, doi:10.1007/s00033-002-8146-7.

Fu, J. L., 2012: The performance verification of the medium-range forecasting for T639, ECMWF and JAPAN models from September to November 2011 (in Chinese). *Meteor. Mon.*, **38**, 238–243.

Goldberg, D. E., B. Korb, and K. Deb, 1989: Messy genetic algorithms: Motivation, analysis, and first results. *Complex Syst.*, **5**, 493–530.

Grinsted, J., J. C. Moore, and S. Jevrajeva, 2004: Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes Geophys.*, **11**, 561–566, doi:10.5194/npg-11-561-2004.

Gu, X., 1998: A spectral model based on atmospheric self-memorization principle. *Chin. Sci. Bull.*, **43**, 1692–1702, doi:10.1007/BF02883967.

He, X.-Z., and D.-Y. Gong, 2002: Interdecadal change in western Pacific subtropical high and climatic effects. *J. Geogr. Sci.*, **12**, 202–209, doi:10.1007/BF02837475.

Hong, M., R. Zhang, J. X. Li, and K. F. Liu, 2013a: Inversion of the western Pacific subtropical high dynamic model and analysis of dynamic characteristics for its abnormality. *Nonlinear Processes Geophys.*, **20**, 131–142, doi:10.5194/npg-20-131-2013.

——, ——, and K. F. Liu, 2013b: Retrieving dynamic forecast model of the western Pacific subtropical high in abnormal years based on GA (in Chinese). *Wuli Xuebao*, **62**, 1–13.

——, D. Wang, R. Zhang, X. Chen, J.-J. Ge, and D. Yu, 2015: Reconstruction and forecast experiments of a statistical–dynamical model of the western Pacific subtropical high and East Asian summer monsoon factors. *Wea. Forecasting*, **30**, 206–216, doi:10.1175/WAF-D-14-00048.1.

Hu, T. S., K. C. Lam, and S. T. Ng, 2001: River flow time series prediction with a range-dependent neural network. *Hydrol. Sci. J.*, **46**, 729–745, doi:10.1080/02626660109492867.

Huang, J. P., and Y. H. Yi, 1991: A nonlinear dynamic system reconstruction of actual data (in Chinese). *Sci. China*, **3**, 331–336.

Jiang, X., and X. N. Cai, 2011: The performance verification of the medium-range forecasting for T639, ECMWF and JAPAN models from June to August 2011 (in Chinese). *Meteor. Mon.*, **37**, 1448–1452.

Kurihara, K., 1989: A climatological study on the relationship between the Japanese summer weather and the subtropical high in the western northern Pacific. *Geophys. Mag.*, **43**, 45–104.

Liu, K. F., R. Zhang, P. Yu, Y.-L. Wang, and D.-D. Yu, 2007: Area exponent of western Pacific subtropical high forecast model based on wavelet decomposition support vector machine (in Chinese). *J. Trop. Meteor.*, **23**, 491–496.

——, ——, M. Hong, D.-D. Yu, and Y.-L. Wang, 2008: RBFNN based on hybrid hierarchy genetic algorithm and its application in subtropical high forecast (in Chinese). *J. Trop. Meteor.*, **24**, 507–511.

Liu, M., R. Hang, and B. Zhang, 2013: Influence of subtropical high's variation period and structure on plum rain onset (in Chinese). *J. Meteor. Sci.*, **33**, 430–435.

Lu, R., 2001: Interannual variability of the summertime North Pacific subtropical high and its relation to atmospheric convection over the warm pool. *J. Meteor. Soc. Japan*, **79**, 771–783, doi:10.2151/jmsj.79.771.

——, and B. W. Dong, 2001: Westward extension of North Pacific subtropical high in summer. *J. Meteor. Soc. Japan*, **79**, 1229–1241, doi:10.2151/jmsj.79.1229.

——, H. Ding, C.-S. Ryu, Z. Lin, and H. Dong, 2007: Mid-latitude westward propagating disturbances preceding intraseasonal oscillations of convection over the subtropical western North Pacific during summer. *Geophys. Res. Lett.*, **34**, L21702, doi:10.1029/2007GL031277.

Miyasaka, T., and H. Nakamura, 2005: Structure and formation mechanisms of the Northern Hemisphere summertime subtropical highs. *J. Climate*, **18**, 5046–5065, doi:10.1175/JCLI3599.1.

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, doi:10.1175/BAMS-85-6-853.

Park, J.-Y., J.-G. Jhun, S.-Y. Yim, and W.-M. Kim, 2010: Decadal changes in two types of the western North Pacific subtropical high in boreal summer associated with Asian summer monsoon/El Niño–Southern Oscillation connections. *J. Geophys. Res.*, **115**, D21129, doi:10.1029/2009JD013642.

Qi, L., Z.-Q. Zhang, J.-H. He, and H. Chen, 2008: A probe into the maintaining mechanism of one type of the double-ridges processes of West Pacific subtropical high (in Chinese). *Chin. J. Geophys.*, **51**, 495–504, doi:10.1002/cjg2.1240.

Sui, C.-H., P.-H. Chung, and T. Li, 2007: Interannual and interdecadal variability of the summertime western North Pacific subtropical high. *Geophys. Res. Lett.*, **34**, L11701, doi:10.1029/2006GL029204.

Takens, F., 1981: Detecting strange attractors in fluid turbulence. *Detecting Strange Attractors in Fluid Turbulence: Proceedings of a Symposium Held at the University of Warwick 1979/80*, D. Rand and L.-S. Yang, Eds., Lecture Notes in Mathematics, Vol. 898, Springer, 361–381.

Wang, H. Z., R. Zhang, Y. L. Wang, and K.-F. Liu, 2006: Errors revised for numerical forecast products of subtropical high based on Kalman filtering (in Chinese). *J. Trop. Meteor.*, **22**, 661–666.

Wang, W., J. Y. Su, B. W. Hou, J. Tian, and D. H. Ma, 2012: Dynamic prediction of building subsidence deformation with data-based mechanistic self-memory model. *Chin. Sci. Bull.*, **57**, 3430–3435, doi:10.1007/s11434-012-5386-6.

Wang, W.-C., K.-W. Chau, C.-T. Cheng, and L. Qiu, 2009: A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *J. Hydrol.*, **374**, 294–306, doi:10.1016/j.jhydrol.2009.06.019.

Wang, X. R., Y. Q. Yao, Y. Shang, X. P. Chen, X. Q. Cheng, and A. M. Shuai, 2002: The error analyze of medium-term numerical forecast for subtropical high in 1998. (in Chinese). *J. Trop. Meteor.*, **18**, 351–358.

Wang, Y., Q. Wang, Y. Chen, and M. Wu, 2007: Application of subtropical high index from ECMWF model (in Chinese). *J. Meteor. Environ.*, **23**, 26–31.

Wu, M. L., H. Liang, Y. Wang, and Y.-M. Shen, 2009: Contrast tests of precipitation products between T213 and Germany numerical prediction in 2008 (in Chinese). *J. Meteor. Environ.*, **25**, 4–12.

Xu, H. B., R. Zhang, and K. F. Liu, 2007: Numerical forecast products optimization of the West-Pacific subtropical high based on the wavelet decomposition and SOFM-BP artificial neural networks (in Chinese). *J. Meteor. Environ.*, **23**, 265–270.

Xue, F., H. J. Wang, and J. H. He, 2003: The influence of Mask Lin high and Australia high interannual variation to the East Asian summer monsoon precipitation (in Chinese). *Chin. Sci. Bull.*, **3**, 19–27.

Yang, J., G. L. Feng, J. H. Zhao, and Z. H. A. Zhen, 2012: A study of objective and quantitative forecasting the western Pacific subtropical high and its indication for precipitation in summer over China (in Chinese). *Acta Meteor. Sin.*, **70**, 1032–1044.

Yeh, S. W., J.-S. Kug, B. Dewitte, M.-H. Kwon, B. P. Kirtman, and F.-F. Jin, 2009: El Niño in a changing climate. *Nature*, **461**, 511–514, doi:10.1038/nature08316.

Yu, D.-D., R. Zhang, M. Hong, T.-Z. Min, and P.-W. Guo, 2007: A characteristic correlation analysis between the Asia summer monsoon memberships and west Pacific subtropical high (in Chinese). *J. Trop. Meteor.*, **1**, 58–67.

Zhang, R., M. Hong, Z.-B. Sun, S.-J. Niu, W.-J. Zhu, J.-Z. Min, and Q.-L. Wan, 2006: Non-linear dynamic model retrieval of subtropical high based on empirical orthogonal function and genetic algorithm. *Appl. Math. Mech.*, **27**, 1645–1654, doi:10.1007/s10483-006-1207-z.

Zhang, Y. N., and J. Y. Zhang, 2011: The performance verification of the medium-range forecasting for T639, ECMWF and JAPAN models from December 2010 to February 2011 (in Chinese). *Meteor. Mon.*, **37**, 633–638.

Zhou, T. J., and Coauthors, 2009: Why the western Pacific subtropical high has extended westward since the late 1970s. *J. Climate*, **22**, 2199–2214, doi:10.1175/2008JCLI2527.1.