

## Research Article

# Set Pair Analysis Based on Phase Space Reconstruction Model and Its Application in Forecasting Extreme Temperature

Yin Zhang,<sup>1</sup> Xiao-hua Yang,<sup>2</sup> Ling Zhang,<sup>3</sup> Wan-ying Ma,<sup>1</sup> and Ling-xia Qiao<sup>1</sup>

<sup>1</sup> School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

<sup>2</sup> State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing 100875, China

<sup>3</sup> Department of Mathematics and Statistics, Auburn University, Auburn, AL 36832, USA

Correspondence should be addressed to Xiao-hua Yang; xiaohuayang@bnu.edu.cn

Received 3 June 2013; Revised 16 July 2013; Accepted 19 July 2013

Academic Editor: Ming Li

Copyright © 2013 Yin Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the precision of forecasting a time series, set pair analysis based on phase space reconstruction (SPA-PSR) model is established. In the new model, by using chaos analysis, we reconstruct the phase space with delay time and embedding dimension. Based on it, we rebuilt history sets and current sets in the SPA-PSR model. Two cases of forecasting extreme temperature in Mount Wutai and Datong are taken to examine the performance of SPA-PSR model. The results indicate that the mean relative error (MRE) of SPA-PSR model has decreased by 65.97%, 59.32%, and 7.79% in the case of Mount Wutai and 29.11%, 32.82%, and 9.03% in the case of Datong, respectively, compared with autoregression (AR) model, rank set pair analysis (R-SPA) model, and Back-Propagation (BP) neural network model. It gives a theoretical support for set pair analysis and improves precision of numerical forecasting.

## 1. Introduction

The global and regional climates have already begun changing [1], and as the important factor of climate change, temperature plays a significant role in human's daily life [2, 3]. It is important to forecast extreme temperature accurately [4]. Temperature change process is usually nonlinear, complex, and dynamic, so the accurate prediction of extreme temperature is faced with a high degree of scientific uncertainty, which traditional deterministic mathematical model cannot solve perfectly. And numerical simulation method can solve the problem better [4–6].

Auto regression (AR) model is the traditional method used to deal with the time series forecasting [7]; for example, Bańbura et al. used Bayesian vector autoregressions for commercial forecasting [8]. In recent years, artificial neural network (ANN) algorithms are widely used to deal with forecasting meteorological objects [9]. Based on the genetic algorithm (GA) and particle swarm algorithm, Yang designed the Back-Propagation (BP) neural networks to establish the multifactor time series forecasting model [10]. At the same

time, set pair analysis (SPA) model which is easy to operate and gives good prediction results is also popular in meteorological forecast field [5, 6]. Yang et al. gave the set pair analysis based on similarity forecast (SPA-SF) model for forecasting water resources changing process, and the application results showed that the statistic and physical concepts of SPA-SF were distinct and its precision was high [9]. Recently, Mei et al. used SPA to find an optimal choice of Bioretention media [11] and Guo et al. employed it to assess the ecoenvironment quality for uncertain problems [12]. Because SPA model does not provide a unified standard of quantifying the set element symbols, the rank set pair analysis (R-SPA) model is presented by using the rank set pair analysis [5, 9, 10], and the results are proved to be better.

However, for SPA (including R-SPA) [5, 9], there is no accurate method to determine the dimension of history sets and current sets, which is used to calculate the connection degree and affects prediction results. And by finding the most similar history set, SPA model uses subsequent value as the forecasting value of current set, while it does not give a satisfying theoretical proof. Phase space

reconstruction, based on chaos time series analysis, is the newest development to deal with nonlinear time series [13–17]. It has an excellent description of the system's dynamic behavior by applying the nonlinear dynamics theory and fractal theory. In particular chaos systems are dynamical systems that defy synchronization. They are ubiquitous in nature, and most of them do not have an explicit dynamical equation and can be only understood through the available time series [18]. Under some circumstances, such processes can create time series that appear to be completely random—the corollary of this is that some seemingly random series are in fact chaotic and thus to a certain extent predictable [14]. In 2012, Khatibi et al. studied the chaos in river time series [16] and Di et al. discussed the chaos control and synchronization of a nonlinear system [17]. Based on multiple criteria decision making (MCDM), Yang et al. presented using the chaotic Bayesian method for forecasting nonlinear hydrological time series [19]. Compared with the results of the add-weighted one-rank local-region method (AOLM), the method can improve the forecast accuracy of daily runoffs. And by analyzing the chaos of time series, She and Yang also used the new adaptive local linear prediction in hydrological time [20].

As discussed above, to have a more accurate numerical forecasting of extreme temperature, by using chaos time series analysis, a set pair analysis based on phase space reconstruction (SPA-PSR) model is proposed in the paper: we use the Takens embedding theory [21] to embed the rebuilt history sets and current sets in SPA so that the parameter (dimension of history sets) can be calculated and both kinds of sets have theoretical meanings for prediction. Two cases of forecasting extreme temperature in Mount Wutai and Datong stations are taken to examine the efficiency of SPA-PSR model.

## 2. The Method of Phase Space Reconstruction (PSR)

For a scalar time series,  $x_1, x_2, \dots, x_n$ , according to Takens embedding theory [21], the phase space can be reconstructed to a multidimensional space and the coordinate delay method is commonly used. The constructed  $m$  dimension state vector  $Y_i$  is

$$Y_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}), \quad (1)$$

where  $n$  is the length of time series,  $\tau$  is the delay time and  $m$  is the embedding dimension  $i = 1, 2, \dots, M$ ,  $M$  is the total points number of the phase-space, and  $M = n - (m - 1)\tau$ .

**2.1. The Determination of Delay Time.** In this study, we use the autocorrelation function to determine the delay time  $\tau$  [22–24]. As for time series  $x_1, x_2, \dots, x_n$ , the autocorrelation coefficient function with the lag time  $t$  is [20, 25, 26]

$$C(t) = \frac{1}{n-t} \frac{\sum_{i=1}^{n-t} (x_i - \mu)(x_{i+t} - \mu)}{\sigma^2(x)}, \quad (2)$$

where  $\mu$  and  $\sigma$  are the mean and standard variation of the time series respectively. Drawing the plot of  $t - C(t)$ , the delay

time  $\tau$  is selected when the autocorrelation coefficient has dropped to  $(1 - 1/e)$  [22] of its initial value ( $e$  is the base of natural logarithm).

**2.2. The Determination of Embedding Dimension.** There are also many ways to decide the embedding dimension [27–31]. In order to have a good predicting result, we choose the minimum prediction error method [27] to assure embedding dimension.

For a scalar time series  $x_1, x_2, \dots, x_n$ , reconstruct it with (1). According to Takens embedding theory, when  $\tau$  and  $m$  are the best delay time and best embedding dimension, respectively, there is a mapping that existed:  $F : R^m \rightarrow R^m$ , such that

$$Y_{i+1} = F(Y_i). \quad (3)$$

At this time, the average prediction error (also can be the maximum prediction error, or other prediction errors):

$$E(m, \tau) = \frac{1}{M-1} \sum_{i=1}^{M-1} |x_{i+1+(m-1)\tau} - x_{\eta(i)+1+(m-1)\tau}| \quad (4)$$

should be the minimum. In formula (4),  $\eta(i)$  is assured by finding the nearest neighbor spot  $Y_{\eta(i)}$  of  $Y_i$ , and the way is shown as follows:

$$\|Y_{\eta(i)} - Y_i\| = \min_{j=1,2,\dots,M; j \neq i} \|Y_j - Y_i\|, \quad (5)$$

where  $\|Y_{\eta(i)} - Y_i\|$  means the Euclidian distance (the norm in certain space) between  $Y_{\eta(i)}$  and  $Y_i$ , and we take the norm of  $L_\infty$  space:

$$\begin{aligned} \|Y_{\eta(i)} - Y_i\| &= \min_{j=1,2,\dots,M; j \neq i} \|Y_j - Y_i\| \\ &= \min_{j=1,2,\dots,M; j \neq i} \max_{1 \leq l \leq m-1} |x_{j+l\tau} - x_{i+l\tau}|. \end{aligned} \quad (6)$$

According to the existence of the largest Lyapunov exponents and noise, with embedding dimension  $m$  increasing, the value of  $E(m, \tau)$  will be heavily influenced. So we make  $m$  increase from two, taking the first minimum point in  $E(m, \tau) - m$  curve as the best embedding dimension [27].

The PSR algorithm is shown as follows.

- (1) For a scalar time series  $x_1, x_2, \dots, x_n$ , first, take a smaller value of embedding dimension  $m = 2$ ; then the corresponding phase space reconstruction state vectors are

$$Y_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}) \quad i = 1, 2, \dots \quad (7)$$

- (2) Using formula (5) to find the  $Y_i$ 's nearest neighbor spot,  $Y_{\eta(i)}$ ,  $i = 1, 2, \dots, N$ , and then taking them into formula (4), calculate the corresponding value of  $E(m, \tau)$ .
- (3) Increase the embedding dimension, take  $m + 1 \rightarrow m$ , and repeat Steps 1 and 2.
- (4) Find the first minimum value of  $E(m, \tau)$ , and the corresponding value of  $m$  is the embedding dimension.

**2.3. Verification of the Chaos of Time Series.** In order to verify whether time series is chaotic, we calculate the largest Lyapunov exponent of our studied time series by using the delay time and embedding dimension above [22]. The largest Lyapunov exponent is defined by the value of its nearest neighbor divergence rate on average

$$\ln D_i(k) = \ln C_i + \lambda \cdot (k \cdot \Delta t), \quad (8)$$

where  $D_i(k)$  represents the distance between  $i$ th point and its nearest neighbor after  $k$  time units, and

$$D_i(0) = \min_j \|Y_i - Y_j\|, \quad (9)$$

where  $|i - j| > w$  and  $C_i$  is the initial distance. So the slope  $\lambda$  can represent the largest Lyapunov exponent, which can be calculated by using the least square method.

If the largest Lyapunov exponent  $\lambda$  is greater than 0, the time series is chaotic. If not, the Takens embedding theory is not tenable, and the set pair analysis based on phase space reconstruction model (SPA-PSR) model could not be used.

### 3. Set Pair Analysis Based on Phase Space Reconstruction (SPA-PSR) Model

By analyzing the chaotic time series, the reconstructed phase space is applied to set pair analysis. The set pair analysis based on phase space reconstruction (SPA-PSR) model is shown as follows.

*Step 1* (the phase space reconstruction). According to the methods mentioned in Section 2, the reconstructed phase space is established, and the corresponding state vector  $Y_i$  is

$$Y_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}) \quad (10)$$

according to Takens embedding theory: when  $\tau$  and  $m$  are the best delay time and best embedding dimension, respectively; another corresponding smooth function existed,  $F_1 : R^m \rightarrow R$ :

$$x_{i+1+(m-1)\tau} = F_1(Y_i), \quad (11)$$

where  $F_1$  represents the state transformation of studied time series in  $m$  dimension space.

*Step 2* (rebuilt set pair  $(A_i, B)$ ). According to the similarity for development theory, the SPA model uses the history sets' subsequent value to predict future values. So based on the phase space reconstruction with delay time  $\tau$  and embedding dimension  $m$  in Step 1, we can obtain the dimension  $T$  and rebuilt history sets and the current set: the dimension  $T = m$ , the history set  $A_i$  is  $Y_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau})$ ,  $i = 1, 2, \dots, N$ , current set  $B = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau})$ , and all are shown in Table 1.

Then, we got the rank transformation of the rebuilt sets  $A_1, A_2, \dots, A_{M-1}$ . If some elements have the same rank, we mark them according to their average rank and round off the value. We could obtain the rank set  $A'_1, A'_2, \dots, A'_{M-1}$ .

By combining the set  $B'$  with  $M - 1$  set  $A'_i$  separately, we get the rank set pair  $(A'_i, B')$ . The sets  $(A'_i, B')$  are the rebuilt set pair in constructed phase space. Simply, we make  $(A_i, B)$  represent  $(A'_i, B')$ .

*Step 3* (the calculation of connection degree). According to the connection degree formula of  $A_i, B$  [5, 9],

$$U_{A_i-B} = \frac{S}{N} + \frac{F}{N}i + \frac{P}{N}j, \quad (12)$$

where  $N$  is the number of elements in set  $A_i$  or  $B$ ,  $S$  represents the number of identical elements,  $P$  represents the number of contrary elements,  $F$  represents the number of discrepant elements,  $i$  and  $j$  represent discrepancy degree and contrary degree, respectively.

Use  $d(a_k, b_k) = |a_k - b_k|$ ,  $(a_k, b_k) \in (A_i, B)$  to describe differences between sets  $(A_i, B)$ , and the codomain of  $d(a_k, b_k)$  is  $[0, M - 1]$ . So the  $U_{A_i-B}$  is calculated as follows.

- (1) When  $d(a_k, b_k) = 0$ , add 1 to the element number in set  $S$ .
- (2) When  $0 < d(a_k, b_k) \leq (M - 2)$ , add 1 to the element number in set  $F$ .
- (3) When  $d(a_k, b_k) = M - 1$ , add 1 to the element number in set  $P$ .

According the three principle above, take  $N = m$ ; using formula (12), the  $U_{A_i-B}$  can be obtained.

*Step 4* (the determination of similar set and prediction). According to the connection degree maximum principle [5, 9], some  $A_k$  similar to  $B$  are chosen from all the history sets. Thus, the forecast value of  $x_{M+(m-1)\tau}$ , namely, the  $x_{n+1}$ , is  $\hat{x}_{n+1}$  as follows:

$$\hat{x}_{n+1} = \frac{1}{m} \sum_{k=1}^m w_k x_{k+(m-1)\tau+1}, \quad (13)$$

where  $w_k$  denotes the ratio of mean of elements in  $B$  to mean of elements in  $A_k$ .

## 4. Applications in Extreme Temperature

### 4.1. Case 1

*(1) Study Area and Data Description.* The SPA-PSR model is applied to the highest temperature prediction of Mount Wutai in July 1956~2010, which is relatively more accurate in measurement, easy to get, and useful in practice [32, 33].

We use the temperatures in 1956~1999 as known data and thus obtain the time series  $x_1, x_2, \dots, x_{44}$ . Then we will forecast the highest temperature of Mount Wutai in July 2000~2010.

*(2) The Phase Space Reconstruction.* For the scalar time series of the highest temperature prediction of Mount Wutai in July  $x_1, x_2, \dots, x_{44}$ , we use the autocorrelation method and minimum prediction error method to calculate Mount Wutai highest temperature in July's corresponding delay time and embedding dimension, which is shown in Figure 1.

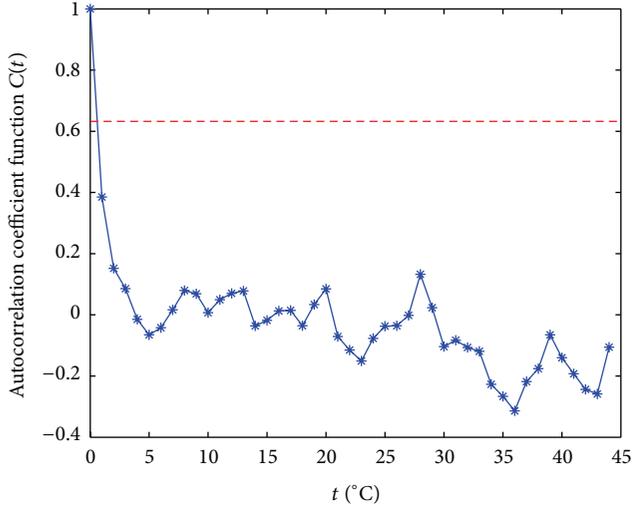


FIGURE 1: Autocorrelation coefficient in Case 1 (the dash line presents the situation where the autocorrelation function value is  $(1 - 1/e)$  of its initial value).

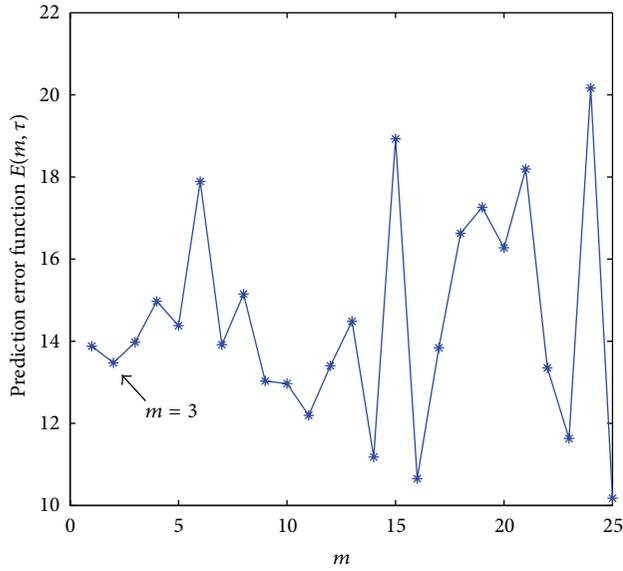


FIGURE 2: The corresponding prediction error function, when delay time  $\tau = 1$ .

In Figure 1, we calculate the autocorrelation of the highest temperature in July of Mount Wutai. When the autocorrelation coefficient has dropped to  $(1 - 1/e)$  of its initial value, the corresponding  $t < 1$ , it indicates that the best delay time is less than 1. However, in reality, the highest temperature in July appears in every other year, which means that  $\tau \geq 1$ ; thus, we take the nearest time to be the delay time, which is  $\tau = 1$ .

From Figure 2, it is obvious that when  $\tau = 1$ ,  $m = 3$ , the prediction error function  $E(m, \tau)$  attains its first minimum point, which indicates that  $m = 3$  is the best embedding dimension for the reconstruction of phase space. Taking  $w = 2$ , with the largest Lyapunov exponent  $\lambda = 0.00086 > 0$ ,

TABLE 1: The rebuilt history sets  $A_i$  and the current set  $B$ .

Sets	Elements in $A_i, B$				Subsequent value
$A_1$	$x_1$	$x_{1+\tau}$	$\cdots$	$x_{1+(m-1)\tau}$	$x_{1+(m-1)\tau+1}$
$A_2$	$x_2$	$x_{2+\tau}$	$\cdots$	$x_{2+(m-1)\tau}$	$x_{2+(m-1)\tau+1}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$A_i$	$x_i$	$x_{i+\tau}$	$\cdots$	$x_{i+(m-1)\tau}$	$x_{i+(m-1)\tau+1}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$A_{M-1}$	$x_{M-1}$	$x_{M-1+\tau}$	$\cdots$	$x_{M-1+(m-1)\tau}$	$x_{i+(m-1)\tau+1}$
$B$	$x_M$	$x_{M+\tau}$	$\cdots$	$x_{M+(m-1)\tau}$	$x_{M+(m-1)\tau+1}$

which indicates that the time series of Mount Wutai is chaotic and satisfies the Takens embedding series.

In sum, with the delay time  $\tau = 1$  and embedding dimension  $m = 3$ , the phase space of the highest temperature in July of Mount Wutai is reconstructed, and the corresponding state vector  $Y_i = (x_i, x_{i+1}, x_{i+2})$ .

(3) *Set Pair Analysis Based on Phase Space Reconstruction (SPA-PSR)*. Based on the reconstruction of phase space in formula (1), the number of reconstructed spots  $M = n - (m - 1) \cdot \tau = 44 - 2 = 42$  and the rebuilt set pairs of  $(A_i, B)$  of the highest temperature of Mount Wutai in July time series  $x_1, x_2, \dots, x_{44}$  is shown as follows:

$$\begin{aligned}
 A_1 &= (x_1, x_2, x_3), \\
 A_2 &= (x_2, x_3, x_4), \\
 &\vdots \\
 A_{41} &= (x_{41}, x_{42}, x_{43}),
 \end{aligned} \tag{14}$$

and the corresponding current set  $B$  is

$$B = (x_{42}, x_{43}, x_{44}) \tag{15}$$

then using the method mentioned in Section 3, We can have the  $x_{45}$  value, namely, the prediction of the highest temperature of Mount Wutai in July 2000.

(4) *The Prediction Results and Analysis*. Prediction results of the highest temperature (P) in Mount Wutai in July 2000~2010 and the corresponding relative error (RE) are shown in Table 1.

Moreover, to verify the prediction effect of SPA-PSR model, we take R-SPA model, AR model, and BP model to make comparisons. To be the same, in AR model [7], it uses the temperatures in 1946~1999 in every 6 years to obtain the regression linear equation of Mount Wutai, and by using the equation, we can predict the corresponding data in 2000~2010; in BP model [10], the temperatures in 1956~1999 are used as trained sets, and the data in 2000~2010 are tested sets, with 500 trained times and the study rate 0.3; in R-SPA model [9], it uses the pervious 6 years as the history sets to predict the temperatures in 2000~2010.

Compared with the results of AR model, R-SPA model, and BP model, SPA-PSR model has the smallest deviation (RE) with the measured value in those 11 years in general as

TABLE 2: Prediction results and the corresponding relative errors (REs).

Year	Measured value (0.1/°C)	AR model		BP model		R-SPA model		SPA-PSR model	
		P (0.1/°C)	RE (%)	P (0.1/°C)	RE (%)	P (0.1/°C)	RE (%)	P (0.1/°C)	RE (%)
2000	250	177.91	28.84	172.66	30.94	220.89	11.64	208.91	16.43
2001	239	182.30	23.72	182.74	23.54	208.08	12.94	258.14	8.01
2002	241	183.82	23.73	183.67	23.79	212.71	11.74	239.92	0.45
2003	233	173.99	25.33	202.02	13.30	236.92	1.68	254.56	9.25
2004	216	173.07	19.88	184.38	14.64	242.55	12.29	244.02	12.97
2005	241	178.01	26.14	210.23	12.77	253.00	4.98	228.12	5.35
2006	227	179.06	21.12	192.55	15.17	243.40	7.23	248.73	9.57
2007	234	173.69	25.77	192.24	17.85	247.74	5.87	214.79	8.21
2008	232	181.48	21.78	184.54	20.46	247.80	6.81	243.74	5.06
2009	236	178.04	24.56	179.79	23.82	274.66	16.38	234.01	0.84
2010	286	176.15	38.41	179.06	37.39	253.15	11.48	232.09	18.85

TABLE 3: The error analysis in four models.

Models	AR model	BP model	R-SPA model	SPA-PSR model
MRE (%)	25.39	21.24	9.37	8.64
MAE	61.59	51.92	22.57	21.12

shown in Table 2. AR model and R-SPA model does not have the prediction in which relative errors are below 10%, while in R-SPA model and SPA-PSR model are 5 and 8, respectively. According to those points, SPA-PSR model has a better forecasting result.

To make it clear, the precision of prediction is evaluated by two measurement indices in this paper, namely, the mean relative error (MRE) and Mean absolute error (MAE). The results are calculated in Table 3.

In Table 3, we can find that AR model and BP model do not have good forecasting results of Mount Wutai, for that their MREs are all above 20%, and MAEs are above 50, while for R-SPA and SPA-PSR, the prediction error is reduced a lot. Compared with AR model, BP model, and R-SPA model, the MRE of SPA-PSR model is relatively decreased by 65.97%, 59.32%, and 7.79%, respectively, and the MAE is also the lowest, which relatively decreased by 65.71%, 59.32%, and 6.42%. The two measurement indices both indicate that SPA-PSR model has the best forecasting results.

4.2. Case 2

(1) *Study Area and Data Description.* We also apply SPA-PSR model to predict results of the highest temperature in Datong in July 2000~2010, known data in 1955–2010.

(2) *The Phase Space Reconstruction.* For the scalar time series of the highest temperature prediction of Mount Wutai in July  $x_1, x_2, \dots, x_{45}$ , the same methods are used to reconstruct phase space. The corresponding delay time and embedding dimension are shown in Figures 3 and 4.

From Figure 3, we find that similar to Mount Wutai, the Datong corresponding autocorrelation coefficient sharpen decreased from  $t = 0$ ; when it has dropped to  $(1 - 1/e)$  of

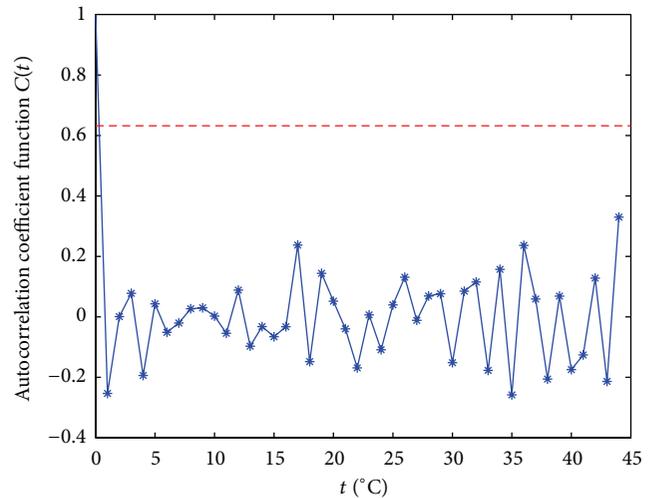


FIGURE 3: The autocorrelation coefficient in Case 2 (the dash line presents the situation where the autocorrelation function value is  $(1 - 1/e)$  of its initial value).

its initial value, the corresponding  $t < 1$ , in the same way, we take  $\tau = 1$  as the delay time. In Figure 4, when  $\tau = 1$  and  $m = 5$ , the prediction error function  $E(m, \tau)$  attains its first minimum point, so the best embedding dimension for Datong is  $m = 5$ . And taking  $w = 2$ , the largest Lyapunov exponent  $\lambda = 0.00024 > 0$ , which indicates that the Datong's time series is chaotic and satisfies the Takens embedding series.

In sum, with the delay time  $\tau = 1$  and embedding dimension  $m = 5$ , the phase space of the highest temperature in July of Datong is reconstructed, and the corresponding state vector is  $Y_i = (x_i, x_{i+1}, x_{i+2}, x_{i+3}, x_{i+4})$ . Being similar to Mount Wutai, the set pairs are rebuilt according to phase space reconstruction and used to forecast the highest temperature of Datong in July 2000–2010.

(3) *The Prediction Results and Analysis.* Prediction results of the highest temperature (P) in Datong in July 2000~2010 and the corresponding relative error (RE) are shown in Table 4.

TABLE 4: Prediction results and the corresponding relative errors (REs).

Year	Measured value (0.1/°C)	AR model		BP model		R-SPA model		SPA-PSR model	
		P (0.1/°C)	RE (%)	P (0.1/°C)	RE (%)	P (0.1/°C)	RE (%)	P (0.1/°C)	RE (%)
2000	361	325.44	9.85	338.18	6.32	313.79	13.08	331.21	8.25
2001	355	310.31	12.59	312.06	12.10	334.54	5.76	317.05	10.69
2002	357	320.83	10.13	324.77	9.03	338.44	5.20	354.39	0.73
2003	324	310.33	4.22	278.54	14.03	350.04	8.04	345.62	6.67
2004	333	323.09	2.98	323.39	2.89	351.28	5.49	345.68	3.81
2005	341	325.48	4.55	280.00	17.89	365.02	7.05	335.29	1.67
2006	335	320.68	4.27	322.46	3.74	327.98	2.10	344.97	2.98
2007	345	328.86	4.68	346.26	0.37	343.53	0.43	367.30	6.46
2008	379	327.16	13.68	321.26	15.23	328.10	13.43	336.75	11.15
2009	360	310.04	13.88	332.43	7.66	347.50	3.47	326.48	9.31
2010	392	310.39	20.82	321.35	18.02	332.37	15.21	351.54	10.32

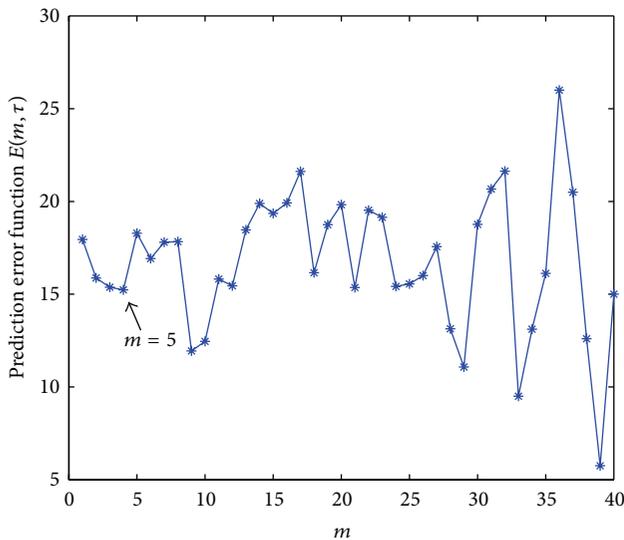


FIGURE 4: The corresponding prediction error function, when delay time  $\tau = 1$ .

The computational methods of the other three models are similar to Case 1.

Compared with the results of AR model, R-SPA model, and BP model, SPA-PSR model has the smallest deviation (RE) with the measured value in those 11 years in general. The numbers of relative errors below 10% in three models are, respectively, 6, 6, 8, and 8. And the majority of forecasting results in SPA-PSR model is below 5%. According to those points, SPA-PSR model is better to be used to forecast.

In Table 5, it indicates that AR model and BP model have better forecasting results for Datong than for Mount Wutai, due to that the corresponding MRE is within 10% and the MAE is also decreased. What is more, the BP model is not better than AR model; the reason is likely to be train function or other factors. At the same time, the two indices all indicate that SPA-PSR model has the best forecasting results. Compared with AR model, BP model, and R-SPA model, the MRE of SPA-PSR model is relatively decreased by 29.11%, 32.82%, and 9.03%, respectively, and the MAE is relatively decreased by 29.93%, 32.56%, and 9.53%.

TABLE 5: The error analysis in four models.

Models	AR model	BP model	R-SPA model	SPA-PSR model
MRE (%)	9.24	9.75	7.20	6.55
MAE	33.58	34.89	26.01	23.53

### 5. Conclusions

To obtain the optimum prediction results, set pair analysis based on phase space reconstruction (SPA-PSR) model is proposed, in which the phase space with embedding dimension and the delay time is reconstructed and prediction results are gotten by combing SPA and phase space reconstruct methods. Two cases of forecasting the highest temperature in July of Mount Wutai and Datong stations are studied by using the new model. The main conclusions are shown as follows.

- (1) By combining the chaotic time series analysis and SPA analysis in the SPA-PSR model, we have a more clearly depiction of the complexity nonlinear dynamical behavior of original forecasting system of extreme temperature, which is a good foundation to make nonlinear forecasting.
- (2) Cases applications indicate that the SPA-PSR model can have a better result for forecasting the highest temperature in July of Mount Wutai and Datong stations. In two cases, compared with AR model, BP model, and R-SPA model, the MRE of SPA-PSR model is relatively decreased by 65.97%, 59.32%, and 7.79% in the case of Mount Wutai and 29.11%, 32.82%, and 9.03% in the case of Datong, respectively, and the MAE is relatively decreased by 65.71%, 59.32%, and 6.42% in the case of Mount Wutai and 29.93%, 32.56%, and 9.53% in the case of Datong, respectively. The results indicate that SPA-PSR model has the better forecasting values.
- (3) Due to the widely use of SPA [5, 6] and ubiquity of chaotic time series [18], the SPA-PSR model can also be used in predicting other nonlinear time series in the future and it will be further studied.

## Acknowledgments

This work was supported by the Project of National Basic Research Program of China (no. 2010CB951104), the National Natural Foundation of China (no. 50939001, 51379013), the Funds for Creative Research Groups of China (no. 51121003), the Specialized Research Fund for the Doctoral Program of Higher Education (no. 20100003110024), and the Program for Changjiang Scholars and Innovative Research Team in University (no. IRT0809).

## References

- [1] G. B. Sahoo and S. G. Schladow, "Impacts of climate change on lakes and reservoirs dynamics and restoration policies," *Sustainability Science*, vol. 3, no. 2, pp. 189–199, 2008.
- [2] B. Bošnjaković, "Geopolitics of climate change: a review," *Thermal Science*, vol. 16, no. 3, pp. 629–654, 2012.
- [3] B. Matic, J. Tepić, S. Sremac et al., "Development and evaluation of the model for the surface pavement temperature prediction," *Metalurgija*, vol. 51, no. 3, pp. 329–332, 2012.
- [4] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Combining neural networks and ARIMA models for hourly temperature forecast," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '00)*, vol. 4, pp. 414–419, July 2000.
- [5] F. G. Liu and Y. Zuo, "The application of rank set pair analysis similar prediction model in futures price forecasting," in *Proceedings of the International Conference*, pp. 240–243, Beijing, China, 2009.
- [6] W. Li, L. Qiu, X. Chen, and Q. Huang, "Assessment model for river ecology health based on set pair analysis and variable fuzzy set," *Journal of Hydraulic Engineering*, vol. 42, no. 7, pp. 775–782, 2011.
- [7] K. Holden, "Vector auto regression modeling and forecasting," *Journal of Forecasting*, vol. 14, no. 3, pp. 159–166, 1995.
- [8] M. Bańbura, D. Giannone, and L. Reichlin, "Large Bayesian vector auto regressions," *Journal of Applied Econometrics*, vol. 25, no. 1, pp. 71–92, 2010.
- [9] S. S. Yang, X. H. Hua, R. Jiang, and Y. C. Zhang, "New optimal weight combination model for forecasting precipitation," *Mathematical Problems in Engineering*, vol. 2012, Article ID 376010, 13 pages, 2012.
- [10] S. Yang, "Neural network forecast under the organic hybrid model of genetic algorithm and particle swarm algorithm," in *Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR '08)*, vol. 1, 2, pp. 254–258, August 2008.
- [11] Y. Mei, X. H. Yang, C. C. Chang et al., "Set pair analysis for optimal selection of bioretention media," *Advanced Science Letters*, vol. 10, no. 1, pp. 687–689, 2012.
- [12] T. N. Guo, X. H. Yang, Y. Mei et al., "Set pair analysis of eco-environment quality using maximum entropy," *Advanced Science Letters*, vol. 10, no. 1, pp. 640–643, 2012.
- [13] L. Cao, Y. Hong, H. Fang, and G. He, "Predicting chaotic time series with wavelet networks," *Physica D*, vol. 85, no. 1-2, pp. 225–238, 1995.
- [14] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," *Physical Review Letters*, vol. 59, no. 8, pp. 845–848, 1987.
- [15] D. I. Jeong and Y. Kim, "Combining single-value streamflow forecasts—a review and guidelines for selecting techniques," *Journal of Hydrology*, vol. 377, no. 3-4, pp. 284–299, 2009.
- [16] R. Khatibi, B. Sivakumar, M. A. Ghorbani, O. Kisi, K. Koçak, and D. Farsadi Zadeh, "Investigating chaos in river stage and discharge time series," *Journal of Hydrology*, vol. 414-415, pp. 108–117, 2012.
- [17] C. L. Di, X. H. Yang, D. W. Huang et al., "Chaos control and synchronization of a nonlinear system," *Nonlinear Science Letters C*, vol. 2, no. 2, pp. 47–54, 2012.
- [18] Z. Liu, "Chaotic time series analysis," *Mathematical Problems in Engineering*, vol. 2010, Article ID 720190, 31 pages, 2010.
- [19] X. H. Yang, D. X. She, Z. F. Yang, Q. H. Tang, and J. Q. Li, "Chaotic bayesian method based on multiple criteria decision making (MCDM) for forecasting nonlinear hydrological time series," *International Journal of Nonlinear Sciences and Numerical Simulation*, vol. 10, no. 11-12, pp. 1595–1610, 2009.
- [20] D. She and X. Yang, "A new adaptive local linear prediction method and its application in hydrological time series," *Mathematical Problems in Engineering*, vol. 2010, Article ID 205438, 15 pages, 2010.
- [21] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, D. A. Rand and L. S. Young, Eds., vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381, Springer, Berlin, Germany, 1981.
- [22] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, "A practical method for calculating largest Lyapunov exponents from small data sets," *Physica D*, vol. 65, no. 1-2, pp. 117–134, 1993.
- [23] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A*, vol. 33, no. 2, pp. 1134–1140, 1985.
- [24] M. T. Rosenstein, J. J. Collins, and C. J. De Luca, "Reconstruction expansion as a geometry-based framework for choosing proper delay times," *Physica D*, vol. 73, no. 1, pp. 82–98, 1994.
- [25] M. Li and W. Zhao, "Smoothing the sample autocorrelation of long-range dependent traffic," *Mathematical Problems in Engineering*, vol. 2013, Article ID 631498, 10 pages, 2013.
- [26] M. Li and W. Zhao, "Quantitatively investigating locally weak stationarity of modified multifractional Gaussian noise," *Physica A*, vol. 391, no. 24, pp. 6268–6278, 2012.
- [27] Y. Lai and D. Lerner, "Effective scaling regime for computing the correlation dimension from chaotic time series," *Physica D*, vol. 115, no. 1-2, pp. 1–18, 1998.
- [28] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," *Physica D*, vol. 9, no. 1, pp. 189–208, 1983.
- [29] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A*, vol. 45, no. 6, pp. 3403–3411, 1992.
- [30] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring, "The analysis of observed chaotic data in physical systems," *Reviews of Modern Physics*, vol. 65, no. 4, pp. 1331–1392, 1993.
- [31] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D*, vol. 110, no. 1, pp. 43–50, 1997.
- [32] J. C. Price, "Land surface temperature measurements from the split window channels of the NOAA 7 advanced very high resolution radiometer," *Journal of Geophysical Research D*, vol. 89, no. 5, pp. 7231–7237, 1984.
- [33] D. R. Easterling, G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, "Climate extremes: observations, modeling, and impacts," *Science*, vol. 289, no. 5487, pp. 2068–2074, 2000.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

