Dear Prof. Lucarini

We are resubmitting you our revised manuscript entitled “Nonstationary time series prediction combined with slow feature analysis”, which number is npg-2014-82. We would like to thank you and the three anonymous reviewers for the careful reading of the manuscript and constructive comments during the discussion phase in NPGD which are very important and helpful on improving the quality of this manuscript.

We have considered all the comments and revised the manuscript accordingly with red font on the revised version. Following is a point-by-point reply to the comments and the marked-up manuscript version.

Referee #1

1. The method is outlined …as this approach will see more and more applications, given that in climate most of the signals are nonstationary.

As a data driven attempt to make progress in prediction of nonstationary climatic time series, this approach will be applied to other climatic signals.

2. A minor comment: In the abstract line 5 the word “combining” should be replaced by “recovering” or “extracting”.

Corrected.

Referee #2

1. The only weakness of this paper is that the authors fail to demonstrate that the methodology they propose performs better than other competing forecasting methodologies.

To make an additional comparison, the best prediction skill was achieved when four climate modes were used as nonlinearly interacting inputs to the prediction model to predict global temperature in the former paper by Wang et al, 2012 (Wang et al :Directional influences on global temperature prediction. Geophys. Res. Lett., 39, L13704, 2012), where the results of the statistical prediction method of persistence have been shown.

2. I suggest the authors to make an additional comparison with other methodologies also because from their figures 2 and 4 the forcing model works better for just a very few steps, just 1-4 steps. This result may be important, but it is not well explained in the paper why it may be important.

As for the prediction experiments shown in Figure 2 and 4 present the forcing model works better for just a very few steps, this may due to the chaotic or nonstationary nature of the two signals (both theoretical and observed data), predictability is lost fast.

Referee #3

The paper is well structured, the results are convincing, and the conclusions are well supported by the data. Therefore I recommend its publication pending technical corrections. Please find specific comments/edits/suggestions in the attached .pdf.

The specific technical edits and suggestions are very helpful for improving this manuscript. We have considered all comments/edits/suggestions and revised the manuscript accordingly.

If you have any other questions, please let us know.

Thanks again for your help on this paper!

Yours sincerely

Geli Wang

Nonstationary time series prediction combined with slow feature analysis

G. Wang [1] and X. Chen[1,2]

1. Key Laboratory of Middle Atmosphere and Global Environment Observations, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029

2. Pingtan Meteorological Bureau of Fujian Province, Pingtan 350400

Correspondence to: G. Wang (wgl@mail.iap.ac.cn)

Abstract. Almost all climate time series have some degree of nonstationarity due to external driving forces perturbing the observed system. Therefore, these external driving forces should be taken into account when constructing the climate dynamics. This paper presents a new technique of obtaining the driving forces of a time series from the Slow Feature Analysis (SFA) approach, then introducing them into a predictive model to predict non-stationary time series. The basic theory of the technique is to consider the driving forces as state variables and incorporate them into the predictive model. Experiments using a modified logistic time series and winter ozone data in Arosa, Switzerland, were conducted to test the model. The results showed improved prediction skills.

1 Introduction

Many previous studies have concluded that the climate system is essentially non-stationary (Trenberth, 1990; Tsonis, 1996; Yang and Zhou, 2005; Boucharel et al., 2009). However, lacking of any general theory for predicting non-stationary processes has become one of the main barriers in the field of climate prediction. To unravel this issue, in recent years, increasing effort has been devoted to devising methods to analyze and predict nonstationary time series.（e.g. Hegger et al., 2000; Verdes et al., 2000, Wan et al, 2005; Wang and Yang, 2005; Yang et al., 2010）. The mostly used method in such studies was to remove or reduce the nonstationarity of the predicted system using some mathematical techniques, thereby improving the prediction skills.

The nonstationarity exists due to the fact that the driving forces that perturb the observed system change with time (Manuca and Savit, 1996). Consequently, the most effective way to remove the nonstationarity may be to incorporate all the driving forces in the constructed dynamical system, and to consider them as the state variables of that system when establishing a prediction equation within a general circulation model (GCM). Based on this principle, lately a data-driven modeling path that compatible with GCM was proposed to predict several artificial non-stationary time series with known external forces. It has achieved success in improving predictions when driving forces were included in some ideal or climate systems, such as the Lorenz system, a logistic model, or global temperature over seasonal timescales including the North Atlantic Oscillation (NAO), the Pacific Decadal Oscillation (PDO), the El Niño/Southern Oscillation (ENSO), and the North Pacific Index (NPI) variability (Wang et. al., 2012, 2013). However, a disadvantage of this technique is that it can not differentiate the assumed driving forces from the predictive model. Therefore, in the present study we considered the extraction of driving forces from the time series itself and established a predictive model by incorporating the constructed driving forces. As a result, the extraction of driving forces became the focus of this study.

Wiscott (2003) developed a technique called Slow Feature Analysis (SFA) to extract driving forces from time series. This technique has been applied to nonstationary time series with some success (Wiskott, 2003; Berkes and Wiskott, 2005; Gunturkun, 2010; Konen and Koch, 2011). In this paper, we used SFA to construct the driving forces from a testing time series, and then established a predictive model that incorporated the driving forces. The paper is organized as follows: A brief description of the predictive technique is presented in section 2. In section 3, results are reported from applying the approach to a modified logistic time series and the total ozone data of Arosa, Switzerland. A summary is provided in section 4.

2 Methodology

SFA is a method that extracts slowly varying driving forces from a quickly varying non-stationary time series. In this section we provide a brief overview of SFA and its application on the extraction and construction of the driving forces from the time series. The details of SFA is presented in Wiscott (2003), but the basic steps of the technique are provided here for convenience and completeness. Let us assume that we have a single variable time series $\{x(t)\}_{t=1,2,\ldots n}$ from a dynamical system:

1) Embed the above time series into an m-dimensional space (also named the length of the m window), a phase trajectory in the m-dimensional space denoted as

$$X(t) = \{x(t), x(t-1), \ldots, x(t-(m-1))\}_{t=1,\ldots,N} \quad \text{or}$$

$$X(t) = \{x_1(t), x_2(t), ..., x_m(t)\}_{t=1,...,N} \quad (1)$$

where $N = n - m + 1$.

2) Generate an expanded signal H(t) for a quadratic expansion, all monomials of degree one and two including mixed terms are used:

$$H(t) = \{x_1(t), ..., x_m(t), x_1^2(t), ..., x_1(t)x_m(t), ...x_{m-1}^2(t), x_{m-1}(t)x_m(t), x_m^2(t)\}_{t=1,...,N}, \quad (2)$$

where *H(t)* is an $k \times N$ matrix and k = m + m (m + 1)/2.

To simplify (2) as

$$H(t) = \{h_1(t), h_2(t)..., h_k(t)\}_{t=1,...,N}. \quad (3)$$

The general objective of SFA is to extract slowly varying features from the time series $\{x(t)\}_{t=1,2,...n}$, in other words, to find a set of coefficients, $W^* = (w_1^*, w_2^*, ..., w_K^*)$, to make the output signal $y^*(t) = W^* \bullet H(t)$ satisfy

$$(\dot{y}^* \dot{y}^{*T}) = \min_k \{(\dot{y}_k \dot{y}_k^T)\}. \quad (4)$$

Here, $\dot{y}_k$ is first-order derivative, calculated by $\Delta y_k(t_i) = y_k(t_{i+1}) - y_k(t_i)$.

3) Normalize the expanded signal H(t), by an affine transformation to generate H'(t) with zero mean and unit covariance matrix:

$$H'(t) = \{h_1'(t), h_2'(t)..., h_k'(t)\}_{t=1,...,N} \quad (5)$$

Where $\bar{h}_j' = 0$, $h_j' h_j'^T = 1$, $h_j'(t) = (h_j(t) - \bar{h}_j)/S$, and $S = \frac{1}{k}\sqrt{\sum_{j=1}^{k}\left(h_j(\mathrm{t}) - \bar{h}\right)^2}$

4) By means of the Schmidt algorithm, the function space (5) is orthogonalized as

$$z_1(t) = h_1'(t),$$

$$z_j(t) = h_j'(t) - \sum_{i=1}^{j-1} \frac{h_{i+1}'(t) \bullet z_i(t)}{\|z_i\|} z_i(t) \qquad (j = 2, ..., K) \quad (6)$$

which is also denoted as $Z(t) = \{z_1(t), z_2(t)..., z_k(t)\}_{t=1,...N}$. Here, $z_i(t) \bullet z_j(t) = 0$ (i≠j) and it guarantees that every variable of the output is uncorrelated

5) Establish the covariance matrix of $Z(t)$, denoted as $B=(\dot{Z}\dot{Z}^T)_{K\times K}$.The k eigenvectors with smallest eigenvalues, λ_k, yield the normalized weight vectors with $\lambda_1 \le \lambda_2 \le .... \le \lambda_k$, which can be easily found by principle component analysis. The smallest eigenvalue, λ_1, corresponding to the eigenvector W_1 can satisfy equation (4), which represents the weight coefficient of the slowest varying component. Here, W_1 has a free scale factor (presented as r), and then the slowest varying variable, or the driving forcing, can be obtained by the following equation:

$$y_1(t) = rW_1 \bullet Z(t) + c, \tag{7}$$

Where c is a given constant and $\{y_1(t)\}$ is the output signal of the slowest driving force obtained by equation (7).

In this study, the SFA was tested on a logistic map

$$s_{t+1} = \mu_t s_t (1 - s_t) \tag{8}$$

with a given driving force parameter

$$\mu_t = 3.5 - 0.45\cos(3\pi t / 1600)\exp(-t / 2500) \tag{9}$$

To test the ability to construct the driving force from this modified logistic map, we took a time series that consists of 5000 data points from this map. Applying the SFA algorithm on this time series with the embedded dimension chosen as 3, we constructed the driving force shown in Figure 1, in which the dotted line represents the true driving force given by (9) and the solid line the constructed driving force by the SFA approach. There is a good agreement between the constructed and the true driving forces with a correlation coefficient of 0.998. This suggests that SFA was able to extract the driving force from the observed time series in an unsupervised manner.

By far we have two time series, one is the original time series $\{x(t)\}$, the another is the slowest driving force $\{y_1(t)\}$. Next we demonstrate how to establish a predictive model that includes the driving force constructed by the SFA procedure described above. We present the basic principle to build the prediction model, for convenience, we assume a nonstationary process composes of two series, $\{x(t)\}_{t=1,2,\ldots n}$ and $\{y_1(t)\}$, with the former being the state variable time series and the latter as the constructed external driving force obtained through the SFA approach. The two time series were embedded in an $m_1 + m_2$ dimensional phase space with a selected time lag τ. The constructed phase trajectory using the embedding theorem of Takens (1981) is shown as:

$$\vec{E}(t) = \{x(t), x(t-\tau), \ldots, x(t-(m_1-1)\tau); y_1(t), y_1(t-\tau), \ldots, y_1(t-(m_2-1)\tau)\}_{t=1,2,\ldots,N} \quad (10)$$

Here, m_1 and m_2 are the given embedding dimensions for $\{x(t)\}$ and $\{y_1(t)\}$, respectively, and $N = n - (\max(m_1, m_2) - 1)\tau$ is the number of phase points on the trajectory.

Based on this trajectory, a predictive model to predict the future state of the system can be established as:

$$x(t+p) = \hat{f}_p(\vec{x}(t); \vec{y}_1(t)) + \varepsilon(t) \quad (11)$$

Where p is the prediction time step (considered as 1 in the present study), $\varepsilon(t)$ is the fitting error, and $\hat{f}$ is assumed to be a quadratic polynomial in this study. The Takens embedding theorem is only appropriate for an autonomous dynamical system, therefore we followed the method of Stark (1999) to embed the driving forces in the same state space for a nonstationary system. The next task is to find the cost function

$$\eta = \sum_{t=1}^{N} [f(x(t), y_1(t)) - x(t+1)]^2 \quad (12)$$

when it reaches its minimum value. For more details, refer to the studies of Farmer and Sidorowich (1987) and Casdagli (1989).

3 Experiments

We applied the prediction technique described above to perform some prediction experiments using several given non-stationary time series. The experiment presented in Section 3.1 was performed with data from the modified logistic model given above.

3.1 Prediction experiments for ideal time series

The prediction experiments were based on 5000 data points from the above verified logistic map (8) with the assumed driving force (9). The first 4800 data points were applied to establish the predictive model, and the remaining 200 data points were used to test the prediction and estimate the correlation coefficient between the actual and predicted values as a function of the prediction time step. The embedding dimension of the verified logistic time series, namely m_1, took values from 2 to 3, and the embedding dimension of the driving force time series, namely m_2, was set to either 0 (the driving force was not taken into account, and is referred to as the 'stationary model' hereinafter) or 1 (the driving force extracted from the verified logistic map by SFA was taken into account, and is referred to as the 'forcing model' hereinafter).The time lag τ was always set to be 1. Figure 2 shows the prediction skill with and without the influence of the driving force, which was constructed by the SFA approach. The forcing model excelled over the stationary model. In particular, at the fourth prediction step, the correlation coefficients were below 0.2 in the stationary model compared to above 0.6 in the forcing model. The average correlation across the prediction time steps was improved, indicating that introducing the driving force extracted through the SFA approach into the prediction model can yield a significant improvement in accuracy.

3.2 Prediction experiment for total ozone

Many studies have sought to explain the variables involved in ozone dynamics, such as the Quasi-Biennial Oscillation (QBO), the 11-year solar cycle, volcanic eruptions, the El Niño Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) (e.g., Brasseur and Granier, 1992; Hood, 1997; Schmidt et al., 2010; Rieder et al., 2010). In this paper we focused on prediction experiments with total ozone data. The total ozone data were from Arosa, Switzerland, and were the world's longest total ozone record. Homogenized total ozone data from 1927 to 2007 were obtained from the World Ozone and Ultraviolet Radiation Data Centre (WOUDC; http://www.woudc.org).

By using the SFA technique on Arosa's daily total ozone data in winter (from January to March) for the period 1927 to 2007, we obtained the first output of the driving force $\{y_1\}$ when the embedding dimension was chosen as 3,5,7,9,11, respectively (shown in Figure 3). Note that the result did not change significantly with different embedding dimension values.

We established a prediction model for winter ozone data by incorporating the driving force constructed by SFA. The prediction was based on 7305 data points. Out of the 7305 data points, the first 7125 data points were used to build the predictive model, and the remaining 180 data points were used to test the prediction using root-mean-square error (RMSE) and the correlation coefficient between observed and predicted values. The time lag τ was taken to be 1, the embedding dimension of the total ozone data m_1 took values from 3 to 5, and the embedding dimension of the driving force time series m_2 was set to either 0 for the stationary model or 3 to 5 for the forcing model.

The experimental results for this case are listed in Table 1, also shown in Figure 4 and Figure 5. From Table 1, it can be seen that all RMSE values given by the forcing model were much lower than those by the stationary model. Figure 4 presents the correlation coefficients between the observed and predicted values. The forcing model outperformed the stationary model, especially at the first two steps. At the first prediction step, the correlation coefficients reached 0.61 for the stationary model but 0.91 for the forcing model. At the 8th prediction step, the correlation coefficients reduced to 0.39 for the stationary model, but still maintained at 0.45 for the forcing model. At the 12th prediction step, the correlation coefficients were 0.22 and 0.33 for the stationary model, and the forcing model respectively. This has clearly shown that, when the constructed driving force is introduced, the accuracy of prediction is dramatically improved. The average correlation over the prediction time steps is improved by 50% when the driving force extracted through SFA technique is included. Figure 5 illustrates the error between the prediction and observation. The prediction errors for every time step is lower for the forcing model than the stationary model. All these results indicate that the inclusion of the driving force constructed by the SFA approach into the prediction model largely improve the predictive skill of winter total ozone in Arosa. Some sensitivity analysis with different training/verifying lengths do not alter this conclusion.

4 Discussion

In this study, we first constructed the driving forces of a time series based on the SFA approach, and then the driving forces were introduced into a predictive model. By doing so, we extend the study by Wang et. al. (2012, 2013) and present a novel technique to predict non-stationary time series. Unlike the former works by Wang et. al. (2012, 2013) with assumed driving forces, in this study the driving force was extracted from original time series. The experimental results obtained from a modified logistic time series and winter ozone data in Arosa confirmed the effectiveness of the model.

The driving force construction technique based on SFA represents a progress for climate causal relations. Such an approach may provide a compatible and direct window for studying causality using external driving forces. We constructed the driving forces with SFA and then combined these driving forces to establish the predictive model. Although we found this approach was able to effectively improve the predictive ability, the constructed driving force time series still lacks of physical explanation. In order to understand the real background of these, one has to further explore the physical processes behind it. One recommended method, provided by Verdes (2005), suggests using a measure called 'transfer entropy' to analyze the causality; another recommended method is named 'convergent cross mapping' provided by Sugihara et. al. (2012), which measures causality in nonlinear dynamic systems. Work in this area is in progress and will be reported in future publications.

Acknowledgments. This research was supported by the National Natural Science Foundation of China under Grant 41275087. The authors are grateful to the editor and three anonymous referees for useful comments during the discussion phase in NPGD.

References

Berkes, P., Wiskott, L.: Slow feature analysis yields a rich repertoire of complex cell properties, Journal of Vision, 5, 579-602, 2005.

Boucharel, J., Dewitte, B., Garel, B., Penhoat, Y.: ENSO's non-stationary and non-Gaussian character: the role of climate shifts, Nonlin. Processes Geophys., 16, 453–473, 2009.

Brasseur, G., Granier, C.: Mount Pinatubo aerosols, chlorofluorocarbons and ozone depletion, Science, 257, 1239–1242, 1992.

Casdagli, M.: Nonlinear prediction of chaotic time series, Phys. D., 35, 335–356, 1989.

Farmer, J. D., Sidorowich, J.: Predicting chaotic time series, Phys. Rev. E., 59, 845–848, 1989.

Gunturkun, U.: Sequential reconstruction of driving-forces from nonlinear nonstationary dynamics, Phys. D., 239, 1095-1107, 2010.

Hegger, R., Kantz, H., Matassini, L., Schreiber, T.: Coping with non-stationarity by over-embedding, Phys. Rev. E., 84, 4092–4101, 2000.

Hood, L.: The solar cycle variation of total ozone: Dynamical forcing in the lower stratosphere, J Geophys Res., 102, 1355-1370, 1997.

Konen, W., Koch, P.: The slowness principle: SFA can detect different slow components in non-stationary time series, Int. J. Innovative Computing and Applications, 3, 3-10, 2011.

Manuca, R., Savit, R.: Stationarity and nonstationarity in time series analysis, Phys. D., 99, 134–161, 1996.

Rieder, H. E., Staehelin, J., Maeder, J. A., Peter, T., Ribatet, M., Davison, A. C., Stubi, R., Weihs, P., Holawe, F.,: Extreme events in total ozone over Arosa – Part 2: Fingerprints of atmospheric dynamics and chemistry and effects on mean values and long-term changes, Atmos Chem Phys., 10,10033--10045, 2010.

Schmidt, H., Brasseur, G. Giorgetta, M., : Solar cycle signal in a general circulation and chemistry model with internally generated quasi‐biennial oscillation, J Geophys Res., 115,D00I14, doi:10.1029/2009JD012542, 2010.

Stark. J,: Delay embeddings for forced systems: deterministic forcing, Journal of Nonlinear Science, 9, 255–332,1999.

Sugihara, G., May,R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., Munch, S.: Detecting causality in complex ecosystems, Science, 338, 496-500, Doi:10.1126/ Science.1227079, 2012.

Takens, F.: Detecting Strange Attractors in Turbulence, Dynamical Systems and Turbulence, Heidelberg, Springer-Verlag, 366–381, 1981.

Trenberth, K. E.: Recent observed inter-decadal climate changes in the northern hemisphere, Bull. Amer. Meteor. Soc., 7, 988–993, 1990.

Tsonis, A. A.: Widespread increases in low-frequency variability of precipitation over the past century, Nature, 382, 700–702, 1996.

Verdes, P. F., Parodi, M. A., Granitto, P. M., Navone, H. D., Piacentini, R. D., Ceccatto, H. A. : Predictions of the maximum amplitude for solar cycle 23 and its subsequent behavior using nonlinear methods, Sol. Phys., 191(2), 419 – 425, 2000.

Verdes, P. F.: Assessing causality from multivariate time series, Phys. Rev. E., 72, 026222, 2005.

Wan, S., Feng, G., Zhou, G., Wan, B., Qin, M., Xu, X. : Extracting useful information from the observations for the prediction based on EMD method, Acta Meteorologica Sinica, 63, 516–525, 2005.

Wang, G., Yang, P.: A compound reconstructed prediction model for nonstationary climate process, Inter. J. Climatol, 25,1265–1277, 2005.

Wang, G., Yang, P., Zhou, X., Swanson, K., Tsonis, A.: Directional influences on global temperature prediction, Geophys. Res. Lett., 39, L13704, doi:10.1029/ 2012GL052149, 2012.

Wang, G., Yang, P., Zhou, X.: Nonstationary time series prediction by incorporating external forces, Adv. Atmos. Sci., 30, 1601-1607, 2013.

Wiskott, L.: Estimating driving forces of nonstationary time series with slow feature analysis, arXiv.org e-Print archive, http://arxiv.org/ abs/cond-mat/ 0312317/, 2003.

Yang, P., Zhou, X.: On nonstationary behaviors and prediction theory of climate systems, Acta Meteorologica Sinica, 63, 556-570, 2005

Yang, P., Wang, G., Bian, J., Zhou, X.: The prediction of non-stationary climate series based on empirical mode decomposition, Adv. Atmos. Sci., 27(4), 845–854, doi: 10.1007/s00376-009-9128-x, 2010.

Table 1 RMSE comparison of the prediction experiments (unit: Dobson units)

	1	2	3	4	5	6	7	8	9	10
Stationary model	0.80	0.88	0.90	0.94	0.96	0.99	1.03	1.02	1.04	1.05
Forcing model	0.62	0.55	0.62	0.74	0.87	0.93	0.97	0.98	1.01	1.01

Figure Captions

Figure 1 The true and constructed driving force.

Figure 2 The comparison of prediction skills between models combined with or without driving force.

Figure 3 The slowest driving force with different embedding dimension for total ozone data.

Figure 4 The comparison of prediction skills between models combined with or without driving force.

Figure 5 Errors (Dobson Units) at prediction steps with or without forcing input.

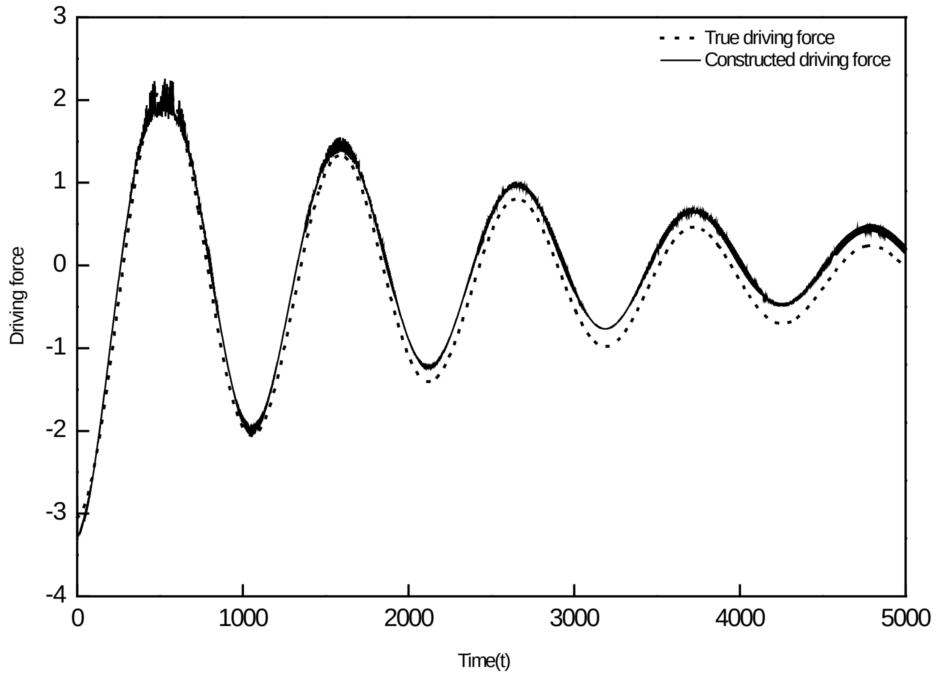

Figure 1　The true and constructed driving force.

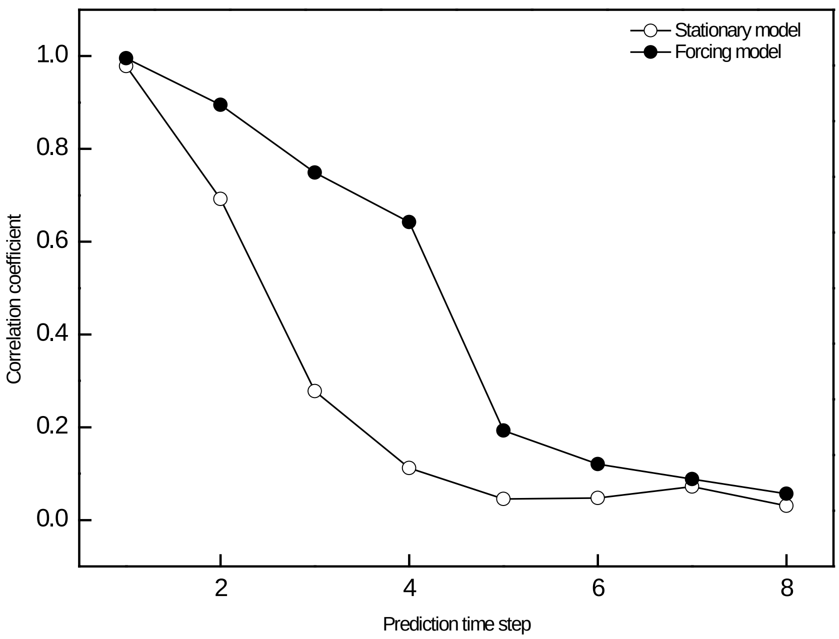

Figure 2　The comparison of prediction skills between models combined with or without driving force.

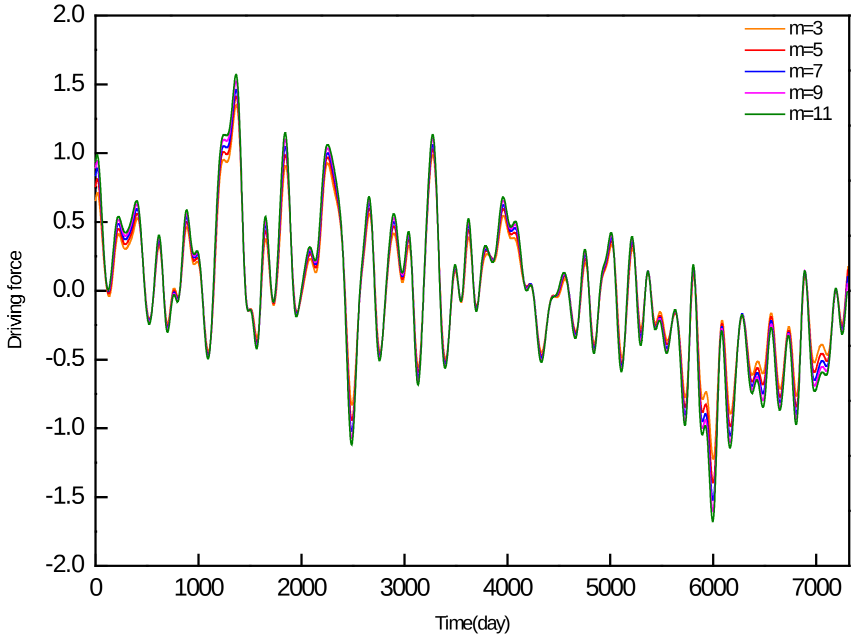

Figure 3 The slowest driving force with different embedding dimension for total ozone data.

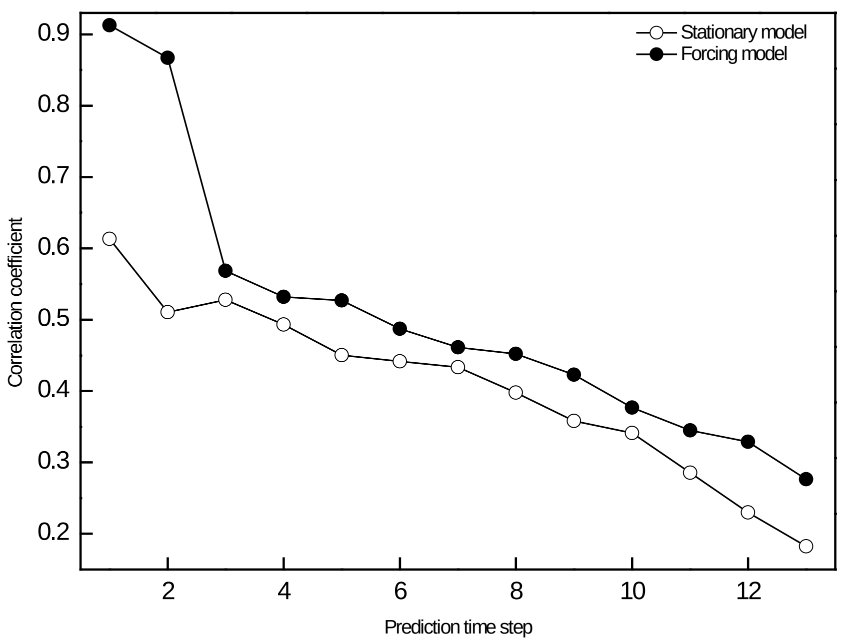

Figure 4 The comparison of prediction skills between models combined with or without driving force.

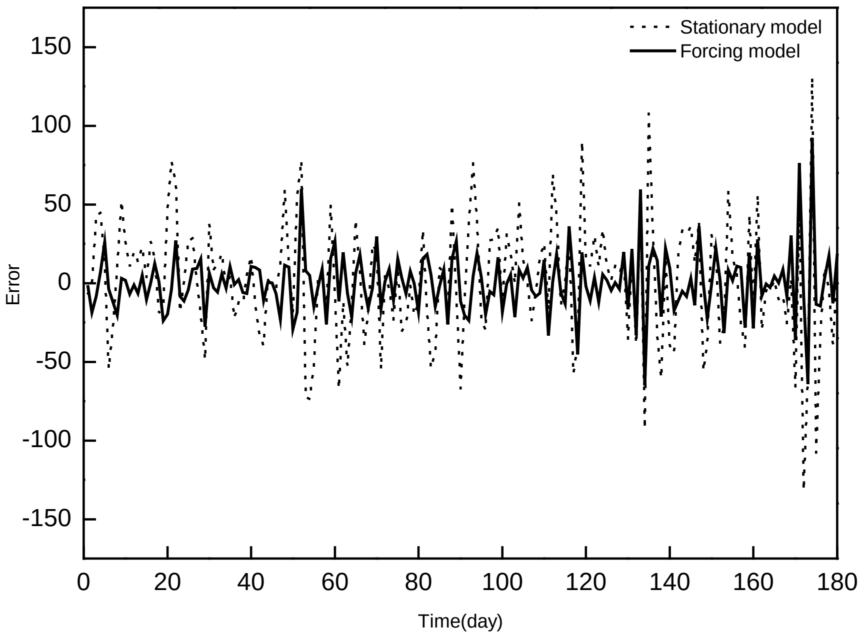

Figure 5 Errors (Dobson Units) at prediction steps with or without forcing input.