



Enhanced time series predictability with well-defined structures

Yu Huang¹ · Zuntao Fu¹

Received: 31 December 2018 / Accepted: 5 March 2019
© Springer-Verlag GmbH Austria, part of Springer Nature 2019

Abstract

For any given time series, how to optimize its forecast strategies and what prediction model is adopted are of great importance. In order to reach this goal, insight from analyzing predictability of series with known structure information is necessary. Time series generated by theoretical models with four kinds of known predictive structures, i.e., short-term correlation, long-term correlation, and multifractal and chaotic patterns, are applied to demonstrate that there is a well-defined relation between series' intrinsic predictability and prediction accuracy of any specific prediction model. And results show that both intrinsic predictability and prediction accuracy are enhanced by these well-defined structures. There are different regimes in the relation between intrinsic predictability and prediction accuracy for series with different known deterministic or stochastic predictive structures. These regimes in the relation between intrinsic predictability and prediction accuracy can guide us to preselect a suitable prediction model and forecast strategies for any underlying series by only analyzing the permutation entropy of a given series. Results from three pieces of climate series further confirm that insights from theoretical series with known structure information indeed work well.

1 Introduction

Predictability studies are hot topics in time series analysis fields (Lorenz 1996; Boffetta et al. 2002). Predictability has been taken as a way to characterize the complexity of the time series' dynamics (Boffetta et al. 2002), so there is close relation between predictability and complexity of the time series. Currently, structural complexity (Li and Fu 2014; Dakos and Soler-Toscano 2017) and prediction accuracy (Bauer et al. 2015; Dietze 2017; Babu and Reddy 2014) of time series attract huge focuses in various fields, such as climate, ecology, economy, and social service. However, there is no definite conclusion about their association. It is generally thought (Garland et al. 2014) that stochastic processes possess the higher structural complexity and deterministic processes like chaotic outputs have a lower level of complexity. Different structural patterns in a given series may influence its prediction accuracy. Previous studies indeed show that ordinal pattern information, such as the stronger long-term memory in stochastic processes, can enhance prediction accuracy (Franzke and Woollings 2011; Yuan et al.

2018). And it was also found that the increased nonlinearity strength in chaotic series can improve the prediction accuracy of deterministic processes (Ye and Hsieh 2008). Some previous studies also conjectured there are some well-defined structures hidden in the real-world series, which can induce a different prediction accuracy (Patil et al. 2001; Yuan et al. 2013; Fu et al. 2019; Molgedey and Ebeling 2000), but there is no further study on whether the higher prediction accuracy is indeed induced by strengthening or weakening well-defined structures. There are different kinds of well-defined structures in the real-world time series, both stochastic and deterministic. Will different types of well-defined structures in the series contribute differentially to the prediction accuracy? Will it have its own regime or phase for each specific well-defined structure in the predictability and prediction accuracy plot? Conclusive answers to these questions will contribute greatly to the understanding and prediction of complex time series.

As a measure of the highest realizable prediction degree, time series intrinsic predictability (Lorenz 1996) also directly reflects complexity of the time series (Boffetta et al. 2002). Both of them can be quantified by permutation entropy (PE) or weighted permutation entropy (WPE) (Garland et al. 2014; Bandt and Pompe 2002; Fadlallah et al. 2013). Previous studies conjectured that there exist monotonous relation between WPE and prediction accuracy for certain data, and this relation was recommended to guide prediction of complex time series (Garland et al. 2014; Pennekamp et al. 2018). However, only

✉ Zuntao Fu
fuzt@pku.edu.cn

¹ Lab for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

limited kinds of well-defined structures in time series were considered in their studies; no definitive results about regime or phase of series with specific well-defined structure in the predictability and prediction accuracy plot are provided.

In our present work, we will first make clear what complexity corresponds to time series with known different types of well-defined structures, which has not been clearly revealed in the literature. Through these detailed studies, regime or phase of each specific well-defined structure can be determined in the predictability and prediction accuracy plot. To accomplish this goal, the theoretically modeled time series are considered with commonly existing well-defined structures in different fields, such as short-term memory, long-term memory, multifractal patterns, and nonlinearity in chaotic series (Graves et al. 2017; Schmitt et al. 2000; Sugihara et al. 2012). Then, two types of prediction strategies, linear or nonlinear modeling, are exploited to forecast these time series and to check to what level of prediction accuracy these time series with different well-defined structures correspond. Finally, we will reach the answer to what regime or phase of series with specific well-defined structures in the predictability and prediction accuracy plot. And this conclusive result will be tested to three climate series to validate its guiding in prediction modeling of real-world time series.

In the following, the methodology used in this work will be introduced in Section 2, then Section 3 reveals the influence of four kinds of well-defined structures on predictability and prediction accuracy, and lastly, Section 4 concludes this work with some conclusions and discussions.

2 Methodology

2.1 Synthetic time series with well-defined structures

2.1.1 Short-term memory

The first considered well-defined ordinal pattern is short-term correlation or short-term memory commonly found in the real-world time series. The autocorrelation function of time series with short-term memory will rapidly decrease exponentially with time delay (Höll and Kantz 2015), and correlation can only exist in neighboring data points for this kind of structural time series. We employ the first order of the autoregressive process (AR(1)) $x_i = ax_{i-1} + \varepsilon_i$ to simulate time series with this kind of structure. Here, a represents the strength of short-term memory, and it can be strengthened from 0 to 1, and $\{\varepsilon_i\}$ is Gaussian white noise with zero mean and unit standard deviation.

2.1.2 Long-term memory

For a long-term correlated time series, its autocorrelation function decreases in the form of a power function with time delay (Höll and Kantz 2015). Here, its simulation is accomplished by an autoregressive fractionally integrated moving average (ARFIMA (p, d, q) , Granger and Joyeux 1980; Massah and Kantz 2016) process with p and q representing the intensities of the autoregressive process and moving average process, respectively. Since only the long-term memory is needed here, the model ARFIMA $(0, d, 0)$ is adopted. We can control long-term memory intensity of the wanted time series by the parameter d . The famous Hurst exponent (Graves et al. 2017) H can be computed by $H = d + 0.5$, and for positive long-term correlation, H is in the interval between 0.5 and 1.

2.1.3 Multifractal patterns

In multifractal time series, there are different autocorrelation intensities over different magnitude levels. The multifractal strength can be quantified by the width of a singular spectrum $d\tau$ (Kantelhardt et al. 2002). We simulated multifractal time series with the binomial multifractal model $x_i = a^{B_{i-1}}(1-a)^{\log_2 N - B_{i-1}}$; here, N is the total number of data points, and for the i th data point x_i , B_{i-1} is the number of digits equal to 1 in the binary representation of the index $i-1$ (this means that the index $i-1$ will be first transformed into binary digits). The parameter a will modulate the strength of the modeled multifractal time series (a could be changed from 0.5 to 1), since the relation between a and $d\tau$ is $d\tau = \frac{\ln a - \ln(1-a)}{\ln 2}$. This model has been widely employed to simulate the characteristics of finance, turbulence, precipitation, and runoff temporal data (Kantelhardt et al. 2002; Rybski et al. 2011; Nian and Fu 2019).

2.1.4 Nonlinearity in chaotic series

For a chaotic system, we take the Lorenz63 system (Lorenz 1963) as an example; the system reads

$$\begin{aligned} dx/dt &= -ax + ay \\ dy/dt &= bx - y - xz \\ dz/dt &= xy - cz \end{aligned} \quad (1)$$

The nonlinearity of chaotic series from these system outputs can be controlled, and the integral chaotic regime will also vary with this controlling (Ye and Hsieh 2008; Basu and Foufoula-Georgiou 2002; Elsner and Tsonis 1992; Ing and Wei 2003; Provenzale et al. 1992). In the previous studies, the output time series were demonstrated to behave differently for some choices of parameters (a, b, c) in Eq. (1) (Basu and Foufoula-Georgiou 2002; Elsner and Tsonis 1992; Ing and Wei 2003). Among them, the most important findings are

those that increased nonlinearity can enhance the predictability of the output time series (Ye and Hsieh 2008).

We first numerically solve Eq. (1) by using the fourth order of the Runge-Kutta method to get the time series of variables X , Y , and Z (initial values are set as (2.85, -4.77, 30.85) for X , Y , and Z , respectively, and the time step is 0.01). Then, we compute the ratio of the nonlinear term and linear term in the second equation and third equation in Eq. (1) as $\beta_y = \frac{\langle |xz| \rangle}{\langle |bx| + |y| \rangle}$ and $\beta_z = \frac{\langle |xy| \rangle}{\langle |cz| \rangle}$. Both β_y and β_z represent the nonlinearity degree of the Lorenz63 system (“ $\langle \rangle$ ” and “ $| \cdot |$ ” denote temporal average and absolute value, respectively). The detailed results about parameters and computed nonlinearity degrees are listed in Table 1. It should be pointed out that five cases for both β_y and β_z ensure that nonlinearity influence on both predictability and prediction accuracy can be quantified.

Up to now, we have constructed the required time series for our present study, and we call short-term memory time series as $S(t)$, long-term memory time series as $L(t)$, multifractal time series as $M(t)$, and chaotic time series as $Z(t)$ (because we only employ the variable Z in the outputs of the Lorenz63 system). The analyzed lengths of time series are 10,000 for all different cases, where the first 8000 points are taken as training series and the last 2000 points as testing series.

2.2 Model-free predictability

2.2.1 Permutation entropy

Permutation entropy (Bandt and Pompe 2002) is widely employed to quantify the complexity of a time series $\{x_t, t = 1, 2, \dots, T\}$. Firstly, reconstructed phase space of this time series and the subvector in the phase space are $X_j^{m,\tau} = \{x_j, x_{j+\tau}, \dots, x_{j+(m-1)\tau}\}$; here, m and τ denote the embedding dimension and time delay. The subscript satisfies $j = 1, 2, \dots, T - (m - 1)\tau$ and it represents the j th subvector. In each subvector, there are $m!$ possible permutations of the elements, and every possible permutation is denoted as π_i . Then, the permutation entropy can be defined as

$$\begin{aligned}
 p(\pi_i^{m,\tau}) &= \frac{\sum_{j \leq N} I_{u:\text{type}(u)=\pi_i} (X_j^{m,\tau})}{\sum_{j \leq N} I_{u:\text{type}(u)=\Pi} (X_j^{m,\tau})}, \text{PE}(m, \tau) \\
 &= - \sum_{i:\pi_i^{m,\tau} \in \Pi} p(\pi_i^{m,\tau}) \ln p(\pi_i^{m,\tau}) \quad (2)
 \end{aligned}$$

2.2.2 Weighted permutation entropy

The complexity of the time series sometimes does not only depend on the permutation but also depend on the amplitude. In this case, weighted permutation entropy (WPE) (Fadlallah et al. 2013) is found to perform better in quantifying

complexity, since it takes the amplitude information in the time series into account. The algorithm of WPE is just like PE, after reconstructing the phase space; it is needed to give weights for every permutation in advance by computing the variance of subvectors $\bar{X}_j^{m,\tau} = \frac{1}{m} \sum_{k=1}^m x_{j+(k-1)\tau}$ and $w_j = \frac{1}{m} \sum_{k=1}^m [x_{j+(k-1)\tau} - \bar{X}_j^{m,\tau}]^2$, then the weight w_i will be taken into the calculation of WPE

$$\begin{aligned}
 p_w(\pi_i^{m,\tau}) &= \frac{\sum_{j \leq N} I_{u:\text{type}(u)=\pi_i} (X_j^{m,\tau}) w_j}{\sum_{j \leq N} I_{u:\text{type}(u)=\Pi} (X_j^{m,\tau}) w_j}, \text{WPE}(m, \tau) \\
 &= - \sum_{i:\pi_i^{m,\tau} \in \Pi} p_w(\pi_i^{m,\tau}) \ln p_w(\pi_i^{m,\tau}) \quad (3)
 \end{aligned}$$

To avoid the finite size effect on PE/WPE analysis, it is necessary to ensure that the data length is larger than $10m!$ for the analyzed time series (Riedl et al. 2013). In our present work, the data lengths of underlying time series are 15,000, 10,000, and 4000, respectively. We should use the same value of m for all of them so that 5 is a solution for m . And then τ is set as 1, which has been suggested to be suitable for quantifying permutation complexity (Bandt 2005; Riedl et al. 2013; Pennekamp et al. 2018).

2.3 Prediction model

Since there are both linear and nonlinear ordinal patterns hidden in the series generated by the four theoretical models, correspondingly, both linear and nonlinear methods should be chosen to evaluate the prediction accuracy. At the same time, the main objective in this study is not to optimize the best model to minimize the predictive error but to show that the increased predictive structure strength in the series can improve the predictability and prediction accuracy, which can provide insight to choose a suitable prediction model. So in this study, only one linear strategy and one nonlinear strategy are considered.

2.3.1 Linear prediction strategy

As a representative linear model (Ing and Wei 2003), the fourth-order autoregressive (AR (4)) model is employed to fit a hyperplane to the given points and then use it to make prediction in our work. With the help of $x_{i+1} = k_0 + k_1 x_i + k_2 x_{i-1} + k_3 x_{i-2} + k_4 x_{i-3}$, we first fit the training time series to acquire model’s parameters (k_0, k_1, k_2, k_3, k_4) by the least square method and then make one-step ahead prediction for the testing time series.

Table 1 Detailed information of chaotic time series

Reference	Chaotic case				
	I (Elsner and Tsonis 1992)	II (Basu and Fofoula-Georgiou 2002)	III	IV	V (Lorenz 1963)
(a, b, c)	(16, 120.1, 4)	(16, 45.92, 4)	(13, 36.4, 3.47)	(10, 31.6, 8/3)	(10, 28, 8/3)
β_y	0.008	0.019	0.026	0.030	0.033
β_z	0.26	0.27	0.31	0.39	0.40

2.3.2 Nonlinear prediction strategy

For nonlinear and dimensional time series (Provenzale et al. 1992; Lorenz 1969), nonlinear prediction strategies could outperform the linear strategies. Here, we employ a classical nonlinear method, Lorenz method analogues (LMA) (Elsner and Tsonis 1992; Lorenz 1969; Fraser and Swinney 1986), to achieve the nonlinear prediction strategy. This method first reconstructs the phase space of the training time series and gets the subvector $X_j^{m,\tau} = \{x_j, x_{j+\tau}, \dots, x_{j+(m-1)\tau}\}$. The choices of the embedding dimension and time delay need some special handlings, which can be found in the studies of Fraser and Swinney (1986) and Kennel (1992). Then, the one-step ahead prediction for the testing time series is carried out in the reconstructed phase space. Based on the current subvector (corresponding to a point in the phase space), it is generally to choose the closest $m + 1$ points (the distances are computed from the Euler distances of the points/vectors) to get the weighted mean of these vectors (weights are counted by the distances) as the forecast of the vector in next step.

It should be noted here, before fitting and prediction, that the original time series are all normalized by means of $x'_i = (x_i - \langle \{x_i\} \rangle) / std(\{x_i\})$ (" $\langle \rangle$ " and " std " denote the temporal average and standard deviation, respectively), and the lengths of the training time series in this work are 8000 and 2000.

2.4 Prediction accuracy

To quantify the prediction accuracy and realizable predictability, here, we employ two important metrics to depict the predicting residuals. The first one is the forecast error (FE), and it quantifies the levels of the residual series' variance relative to the normalized white noise's variance (Hyndman and Koehler 2006). So, FE is defined as

$$FE = \frac{\sum_{i=1}^n (p_i - x_i)^2}{\sum_{i=1}^n \varepsilon_i^2}, \tag{4}$$

where p_i denotes the predicted value, x_i denotes the true value in the testing time series, and $\{\varepsilon_i\}$ is the Gaussian white noise.

The smaller the FE, the better the prediction accuracy is. When FE is less than 1, the forecast skill is acceptable.

The second metric is the mean absolute scaled error (MASE) between the true and predicted data (Hyndman and Koehler 2006), and it can evaluate the match degree between the time series and model. MASE quantifies the residual relative to the one-step variability (OSV) in the training time series, so the prediction accuracy can be compared with random walk prediction based on the training data. MASE is defined as

$$MASE = \frac{\sum_{j=tr+1}^{N_{te}+N_{tr}+1} |p_j - x_j|}{\frac{N_{te}}{N_{tr}} \sum_{i=2}^{N_{tr}} |x_i - x_{i-1}|}, \tag{5}$$

where N_{te} and N_{tr} represent the lengths of the testing and training time series, respectively. $MASE > 1$ means that on average the prediction model does perform worse than the random walk forecast on the training data, but for $MASE < 1$, it performs better.

In addition to the above two metrics, the averaged relative OSV can be taken to evaluate the variability between neighboring data points in the time series, which can provide intuitively the variation in time series when their ordinal structures are strengthened or weakened. The averaged relative OSV is defined as

$$OSV = \frac{\sum_{i=2}^n |x_i - x_{i-1}|}{\sum_{i=2}^n |\varepsilon_i - \varepsilon_{i-1}|}, \tag{6}$$

where $\{\varepsilon_i\}$ is the Gaussian white noise with zero mean and unit standard deviation.

3 Results

3.1 Influence of enhanced ordinal structures on predictability

For complex time series, four kinds of well-defined ordinal patterns, i.e., short-term memory, long-term memory, multifractal patterns, and nonlinearity in chaotic series, are

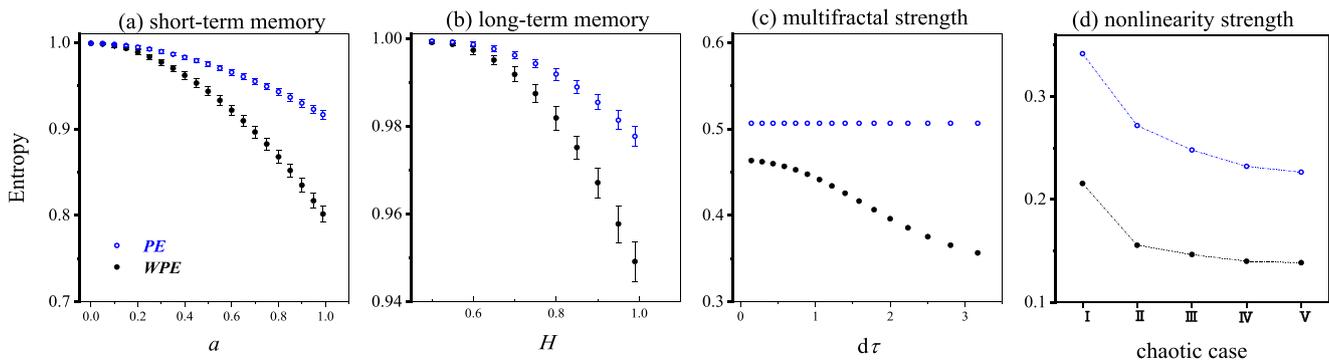


Fig. 1 Scatter plots for the monotonous relation between PE/WPE (hollow blue/solid black dots) and the strengths of ordinal structures with (a) short-term memory, (b) long-term memory, (c) multifractal patterns, and

(d) chaotic attractors, respectively. The error bars in (a) and (b) are the intervals of 2.5 standard deviation

common ordinal structures. Among them, the short-term memory and long-term memory are linear ordinal structures, but the multifractal patterns and nonlinearity in chaotic series are nonlinear ordinal structures. Both linear and nonlinear ordinal structures may play a differential role in adjusting the corresponding series' intrinsic predictability. Actually, the increased strength of both linear and nonlinear ordinal structures can lower the time series' complexity (PE/WPE) and enhance time series' intrinsic predictability. As suggested by Garland et al. 2014, the time series' intrinsic predictability can be quantified by $1 - \text{WPE}$ or $1 - \text{PE}$ (see Fig. 1 for details). Besides this uniform monotonous association between the time series' intrinsic predictability and the strength of ordinal structures, more information can be revealed in the association between the time series' intrinsic predictability and the strength of ordinal structures. Most importantly, linear and nonlinear ordinal structures may play a differential role in adjusting the corresponding series' intrinsic predictability. There are different regimes or phases for linear and nonlinear ordinal structures in the intrinsic predictability and ordinal structure plot. Time series with different types of ordinal structures admit different levels of WPE's value, where the linear stochastic process's ordinal structures, such as short-term memory and long-term memory, have higher WPE (see Fig. 1a and b) and where the deterministic nonlinear process's ordinal structures, such as multifractal patterns and chaotic attractors, have lower WPE (see Fig. 1c and d). This well-defined distinguishable regimes or phases for linear and nonlinear ordinal structures in the intrinsic predictability and ordinal structure plot can be taken as an indicator to preselect a corresponding suitable model to model and predict the underlying series (Garland et al. 2014). And lastly, we should point out that PE may not work well for all kinds of time series, just like multifractal series; it cannot differentiate the multifractal series of different multifractal strengths (see Fig. 1c). This may be caused by the fact that multifractal structures in multifractal series are induced by different amplitudes rather than temporal correlations.

3.2 Influence of enhanced ordinal structures on prediction accuracy

Just as we mentioned in the previous sections, the main objective in this study is not to optimize the best model to minimize the predictive error but to show that the increased predictive structure strength in the series can improve prediction accuracy, which can provide insights to choose a suitable prediction model. So, in this subsection, results from only AR(4) and LMA are compared.

3.2.1 Linear structures

First of all, let us demonstrate how different strengths of ordinal structures influence practical prediction from the AR and LMA methods for time series with short-term memory. Figure 2 presents the comparison among the testing series, predicted series from AR, and predicted series from LMA under three typical cases with $a = 0.2$, $a = 0.55$, and $a = 0.9$. The most important common finding is that the match degree between the testing series and predicted ones increases with the increasing strengths of ordinal structures, and the more the predictive patterns, the better the prediction model performs. Another finding is that both AR and LMA perform almost equally for the short-term correlated series; when the strength of ordinal structures is weaker (such as $a = 0.2$), both methods cannot capture the extreme variations in the testing series (see Fig. 2a), but for the higher strength of ordinal structures (such as $a = 0.9$), both methods can capture the detailed variations in the testing series (see Fig. 2c). This finding is consistent with previous studies, since LMA predictor is practicable for both linear and nonlinear series (Garland et al. 2014).

More quantitative results can be provided by the two prediction accuracy metrics. For all cases under different strengths of ordinal structures, both the AR and LMA methods reach the same results. The mean of FE is not larger than 1, and FE monotonously decreases to 0.05 when short-term memory is enhanced to $a = 1.0$, which means that there exists a forecast skill and that the prediction accuracy is

becoming better. The standard deviations of FE from both the AR and LMA methods decrease with the increasing strengths of ordinal structures (see Fig. 3a). Similarly, MASE shows the same results for both the AR and LMA methods. Since MASE reflects the match degree between the predicted series and the testing series compared with those from random walk prediction, whether it is less than 1 is the rule for the match degree. We can see for the AR and LMA methods that MASE is less than 1 and MASE increases with increasing strengths of ordinal structures (see Fig. 3b). The standard deviations of MASE from both the AR and LMA methods decrease with the increasing strengths of ordinal structures (see Fig. 3b). It should be noted that the behavior of MASE is contrary to that of OSV of the time series itself. Figure 3a shows the changing OSV of the time series with short-term memory and the averaged OSV is becoming weak when the short-term memory is enhanced, which coincides with the results from WPE.

Since both series with short-term memory and series with long-term memory are linear stochastic series, the results are similar for both kinds of series (see Fig. 2 and Fig. 4). When long-term memory is strengthened, some local patterns like trends in sequences become more persistent (Franzke and Woollings 2011), and the averaged OSV in Fig. 5a decreases. However, still minor differences can be found for both kinds of series (see Fig. 3 and Fig. 5). The first difference between them is that the range of OSV, FE, and MASE is narrower for series with long-term memory than with short-term memory. At the same time, the standard deviation for both FE and MASE is almost unchanged with different strengths of ordinal structures in series with long-term memory, and this feature is totally different from that in series with short-term memory.

Fig. 2 Comparison between predicted time series (black lines from AR(4) and green lines from LMA) and testing time series with different long-term memory (gray lines): (a) 0.2, (b) 0.55, and (c) 0.9

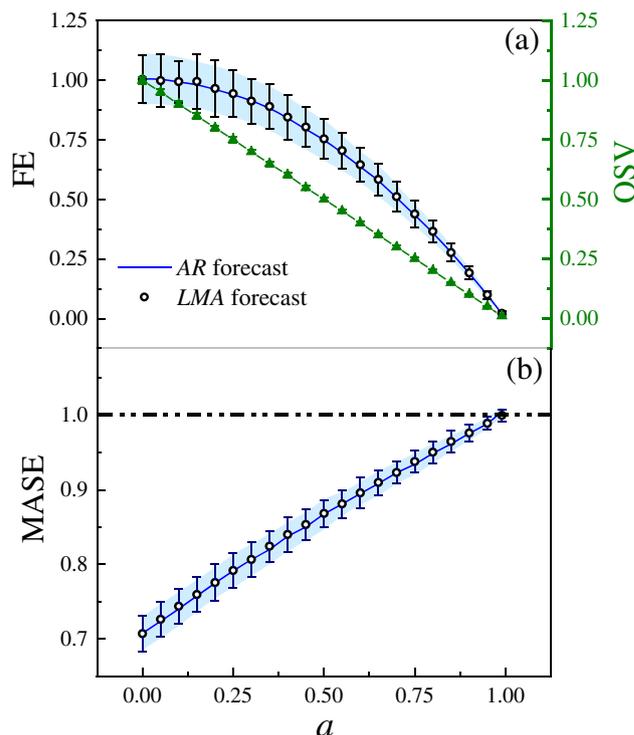
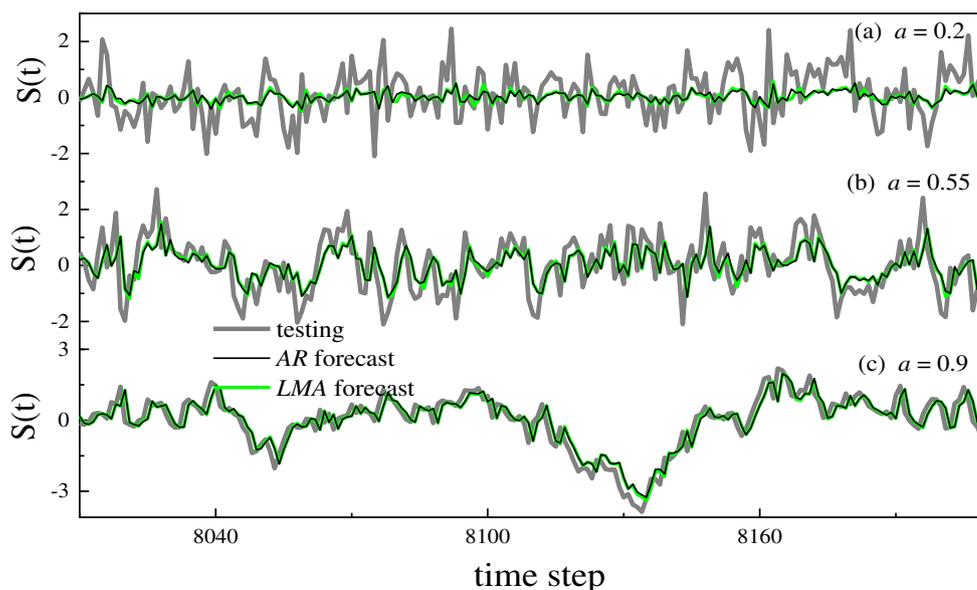


Fig. 3 Evaluation of predicted time series with short-term memory: (a) FE and (b) MASE. The solid blue lines present AR predictions (blue shadows behind are intervals of 2.5 standard deviation) and hollow black dots represent LMA predictions (error bars are intervals of 2.5 standard deviation). In particular, the solid green up-triangles represent the variation of one-step variability in chaotic time series

And lastly, we want to stress that the computation cost in LMA is huger than that in AR; so if we can learn from the WPE information for any given series that both LMA and AR perform equally, we do not need to repeat the computation in LMA.

3.2.2 Nonlinear structures

The aforementioned results are about linear structures in the time series; will enhancing nonlinear structures make FE, MASE, and OSV have the same responses or more specific features that can be revealed? Among the nonlinear structures, multifractal patterns and chaotic attractors are two typical nonlinear structures in real-world time series. Both multifractal series and chaotic series share more features, such as the dimension is all fractal. However, more different behaviors are also revealed in both kinds of series, for example, chaotic series have no marked magnitude differences with sharp transition commonly found in the multifractal series (see Fig. 6). These marked magnitude differences may result in distinguished predictability and prediction accuracy.

First of all, we can learn from the multifractal series that the peaks with sharp transition are more dominated when the multifractal strength is increased (see Fig. 6). At the same time, the temporal distribution of peaks becomes uniform when the multifractal strength is stronger, and the differences between large and small fluctuations are also magnified. However, the averaged OSV decreases with strengthening multifractal structures (Fig. 7a), which coincides with results given by WPE (Fig. 1c). And these features are markedly different from those revealed in linear series shown in the aforementioned results. So, the most marked difference is that AR and LMA perform totally differently (Fig. 6 compared with Fig. 2 and Fig. 5). AR cannot capture the detailed variations in the testing series, especially for the information related to the larger magnitudes, but LMA can. With the increasing multifractal strength, the performance of LMA is nearly perfect, which can be quantitatively found in two metric results given in Fig. 7b. We can see that for most of cases, FE from

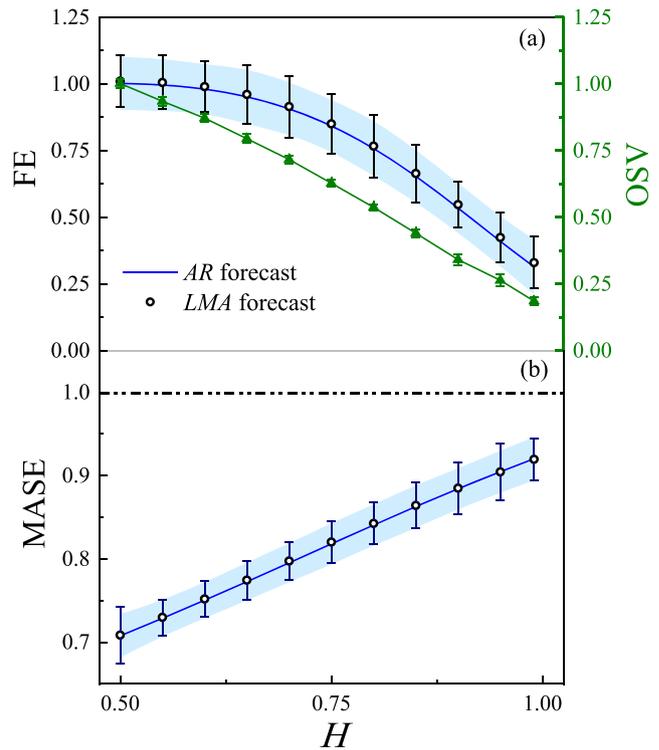


Fig. 5 Evaluation of predicted time series with different long-term memory: (a) FE and (b) MASE. The solid blue lines present AR predictions (blue shadows behind are intervals of 2.5 standard deviation) and hollow black dots represent LMA predictions (error bars are intervals of 2.5 standard deviation). In particular, the solid green up-triangles represent the variation of one-step variability in chaotic time series

LMA is much smaller than that from AR, and its value is limited below 0.125 for all cases. At the same time, MASE is increasing slowly with its value below 0.75 for all cases. However, the value of MASE is increasing slowly with its

Fig. 4 Comparison between predicted time series (black lines from AR(4) and green lines from LMA) and testing time series with different long-term memory (gray lines): (a) 0.6, (b) 0.9, and (c) 0.99

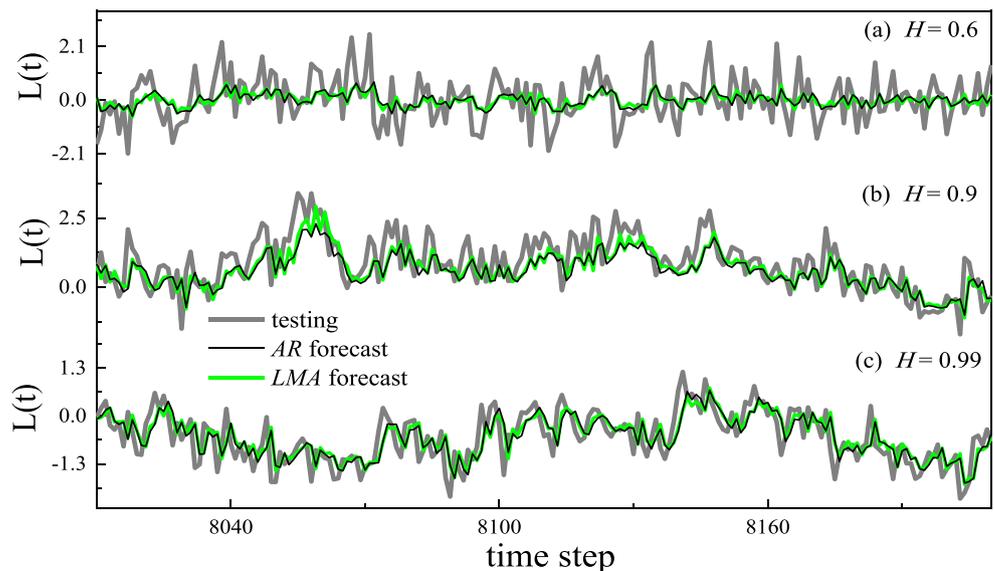
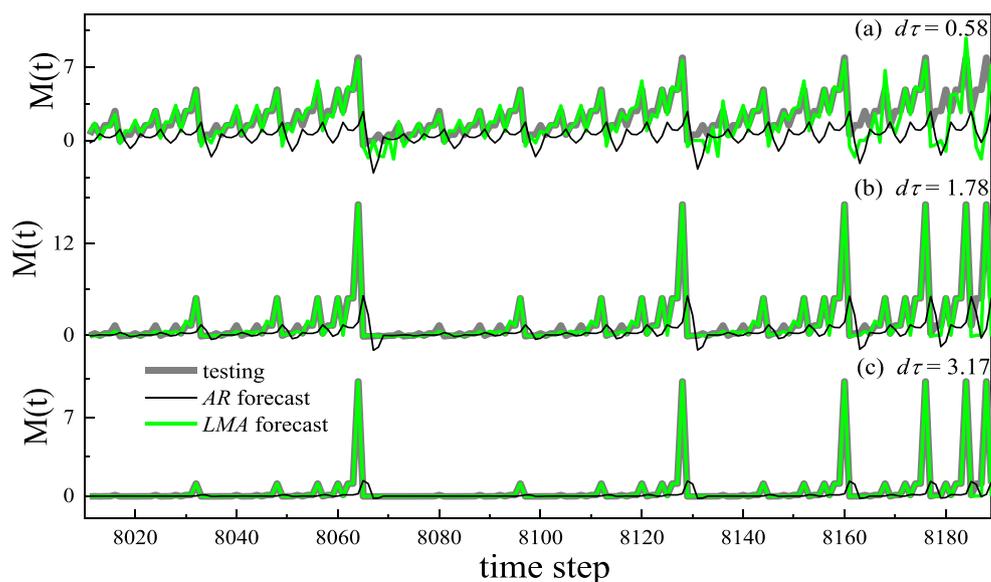


Fig. 6 Comparison between predicted time series (black lines from AR(4) and green lines from LMA) and testing multifractal time series (gray lines) with different multifractal strengths: (a) 0.58, (b) 1.78, and (c) 3.17



value above 1.25 for all cases, which indicates that the prediction strategy from linear methods is unsuitable to predict the multifractal series.

Whereas for the chaotic series with different nonlinear strengths, the results are a little different from those given for the multifractal series, where LMA works well but AR

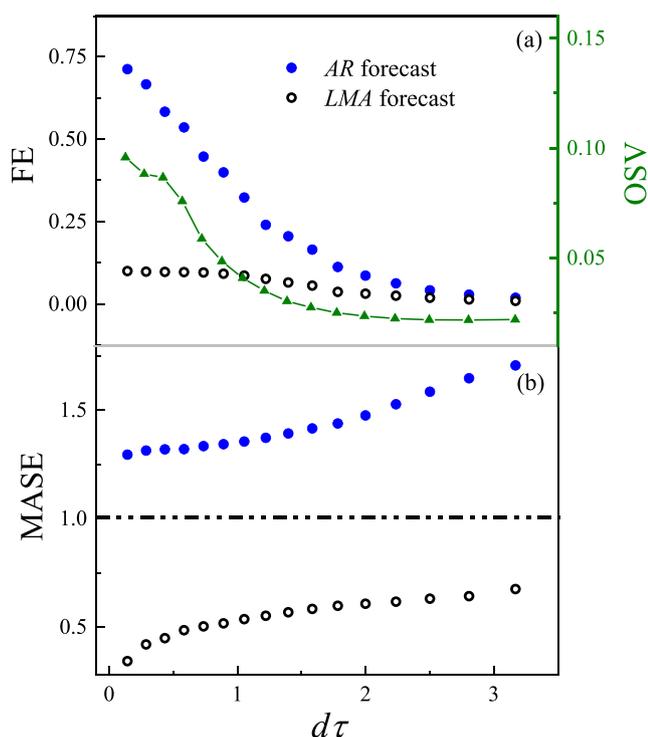


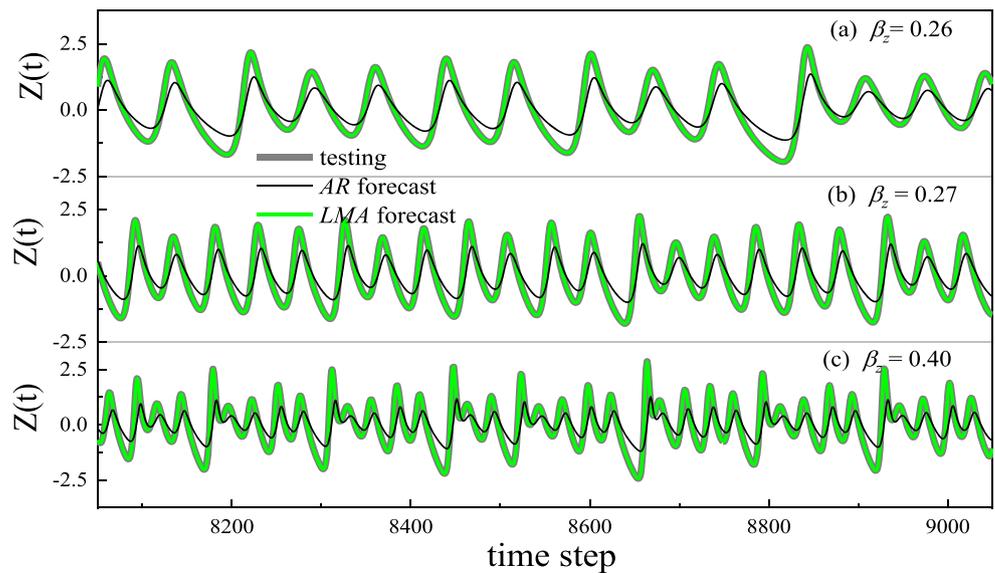
Fig. 7 Evaluation of predicted multifractal time series under different multifractal strengths: (a) FE and (b) MASE. The solid blue dots represent AR predictions and hollow black dots represent LMA predictions. In particular, the solid green up-triangles represent the variation of one-step variability in chaotic time series

fails for all cases (Fig. 8). The variations in the chaotic series are smooth without peaks with sharp transition, which makes the averaged OSV be below the minimum value given in the multifractal series for all cases but one (see Fig. 9a). Quantitatively, both FE and MASE from LMA are the lowest among four kinds of well-defined ordinal structures, where FE from LMA nearly collapses to 0 (Fig. 9a) and where MASE from LMA is below 0.03, which is two orders smaller than that from AR (Fig. 9b). For AR, MASE is larger than 3 for all cases. The results show that the nonlinear model such as LMA performs very well for the prediction of the chaotic time series, but the linear model such as AR works even much worse. The reason for this is that trajectory becomes denser with enhancing nonlinearity in the chaotic series (Ye and Hsieh 2008; Sugihara et al. 2012; Elsner and Tsonis 1992; Ing and Wei 2003; Provenzale et al. 1992) and that the variability in neighboring points will decrease. And all of these facts will make the variations in the chaotic series more ordered with the lowest WPE (see Fig. 1).

3.3 Association between prediction accuracy and predictability

Previous studies have conjectured there is a well association between the intrinsic predictability (WPE/PE) and the realizable predictability or prediction accuracy (FE/MASE) in any given series (Garland et al. 2014). And this conjecture has been validated in several series (Fu et al. 2019; Pennekamp et al. 2018). However, there is no further study to exploit the deep association between the intrinsic predictability (WPE/PE) and the realizable predictability or prediction accuracy (FE/MASE). For the aforementioned four kinds of theoretical series, we have a chance to achieve this goal. If we show the results of $(1 - WPE)$ and MASE from these four kinds of

Fig. 8 Comparison between predicted time series (black lines from AR(4) and green lines from LMA) and testing chaotic time series (gray lines) with different nonlinearity: (a) 0.26, (b) 0.27, and (c) 0.40



theoretical series in a plot, a clearer association is reached between the intrinsic predictability and the realizable predictability or prediction accuracy (see Fig. 10). There are different regimes or phases for linear and nonlinear time series in the (1 - WPE)-MASE plot. The regime with the highest 1 - WPE and the lowest MASE corresponds to the case of the chaotic series, whereas the middle regime is for the case of the multifractal series, and the regime with the lowest 1 - WPE and the highest MASE is for the linear series. There is a distinct regime separation between linear series and nonlinear series. The regime in the (1 - WPE)-MASE plot can be taken as a benchmark to guide the choice of the prediction strategy for any given series from the real world. Since 1 - WPE for any given series is easier to compute and compare the estimated 1 - WPE with the results given in Fig. 10, we can decide whether a linear or nonlinear strategy is chosen to model or predict this given series.

3.4 Application in predicting real-world time series

To illustrate the power of the regime revealed in the (1 - WPE)-MASE plot in guiding the suitable modeling or predicting strategy to some real-world complex time series, three climatic records are studied here. All climatic records, including daily air temperature anomaly in Valkenburg (TEM) from 1976 to 2017, daily indices of El Niño-Southern Oscillation (ENSO) from 1980 to 2017, and daily indices of Atlantic Meridional Overturning Circulation (AMOC) from 2004 to 2017, were downloaded from the site (<https://climexp.knmi.nl/start.cgi>). First of all, we can compute the intrinsic predictability (1 - WPE) for each series, and their values are 0.18 for TEM, 0.43 for ENSO, and 0.73 for AMOC, detailed results are summarized in Table 2.

Comparing the 1 - WPE results with the regimes shown in the (1 - WPE)-MASE plot (Fig. 10), the suggested modeling or predicting strategy is totally different. Firstly, 1 - WPE = 0.18 indicates that the daily air temperature anomaly in Valkenburg should be modeled and predicted by the linear model such as AR, and the nonlinear method such as LMA

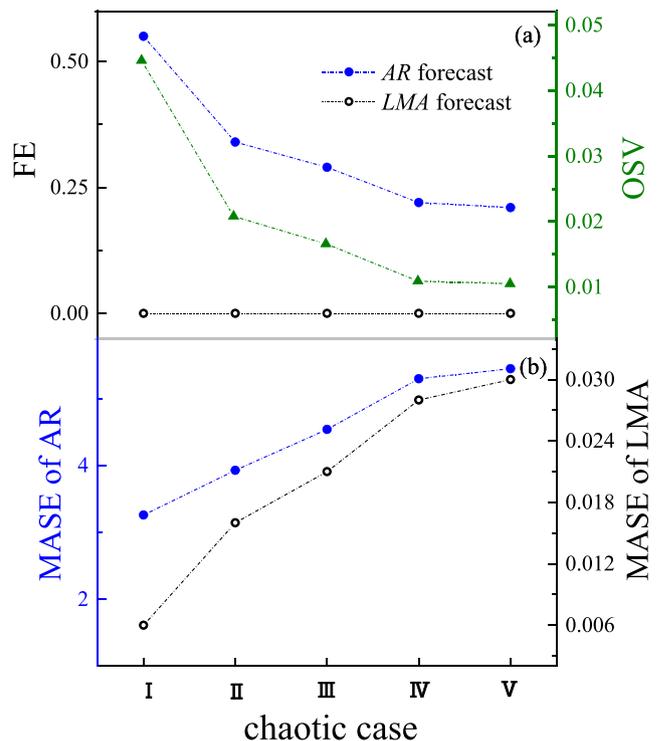
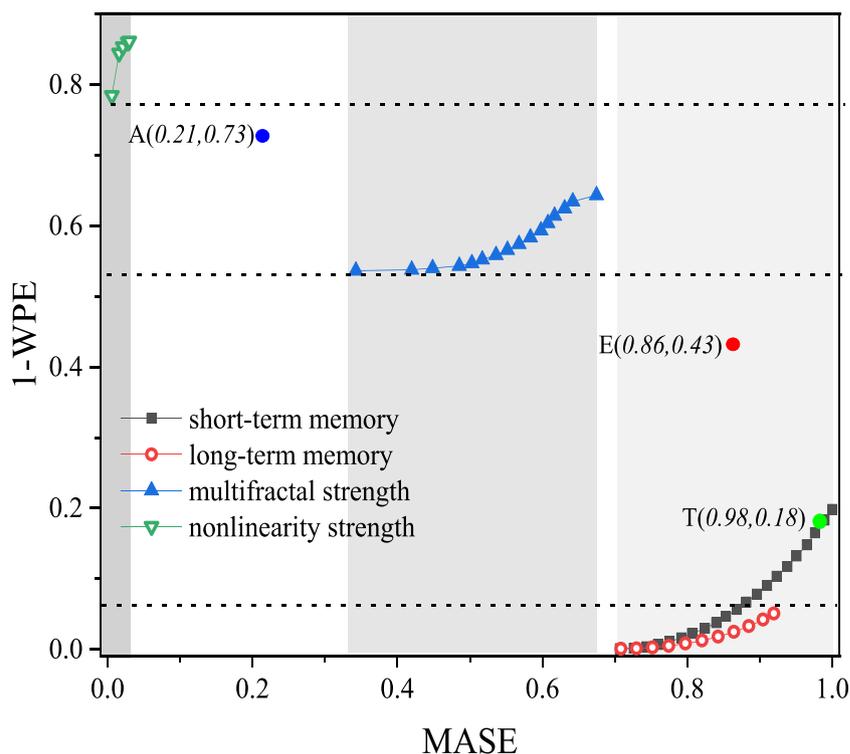


Fig. 9 Evaluation of predicted chaotic time series under different nonlinearity: (a) FE and (b) MASE. The solid blue dots represent AR predictions and hollow black dots represent LMA predictions. In particular, the solid green up-triangles represent the variation of one-step variability in chaotic time series

Fig. 10 Scatter plot MASE (LMA) versus $1 - \text{WPE}$ for series with different ordinal structures (solid black squares for short-term memory, hollow red dots for long-term memory, solid blue up-triangles for multifractal patterns, and hollow green down-triangles for chaotic attractors). There are different regimes for series with different ordinal patterns. Solid blue dot denotes state $A(0.21, 0.73)$ from the daily AMOC index, solid red dot denotes state $E(0.86, 0.43)$ from the daily ENSO index, and solid green dot denotes state $T(0.98, 0.18)$ from the daily air temperature anomaly



will reach similar results. The prediction accuracy quantified by MASE for LMA is 0.98 and for AR is 0.96 (see Table 2), and they are all below 1, which indicates both methods work well. The state $(0.18, 0.98)$ corresponds well to the output with short-term memory (see green dot $T(0.18, 0.98)$ in Fig. 10). The well-matched degree between testing and predicted series can be found in Fig. 11a and d. Secondly, for the ENSO index, $1 - \text{WPE} = 0.43$, which lies between the multifractal regime and the regime with short-term memory, and it is much close to the multifractal regime (see red dot $E(0.43, 0.86)$ in Fig. 10). So, nonlinear methods should be adopted to model and predict the daily ENSO index variations; further computation confirms that LMA indeed performs better (with $\text{MASE} = 0.86$, which is below 1) than AR does (with $\text{MASE} = 2.19$, which indicates the AR model fails to capture the detailed ENSO index variations, see Fig. 11b and e). The well-matched degree between testing and predicted series from LMA can be found in Fig. 11b and e. And lastly, for the AMOC index, $1 - \text{WPE} = 0.73$, the intrinsic predictability is really high. The state in the $(1 - \text{WPE})$ -MASE plot belongs to the nonlinear regime

between the chaotic regime and the multifractal regime (see blue dot $A(0.73, 0.21)$ in Fig. 10), which indicates that the linear model cannot model this series well (with $\text{MASE} = 2.79$, which indicates the AR model fails to capture the detailed AMOC index variations, see Fig. 11c and f) and that the nonlinear methods must be chosen to model and predict the daily AMOC index variations. In fact, from the daily AMOC index series, we can find there are regime shifts (see Fig. 11f) just like what we find in the chaotic series (see Fig. 8c).

4 Conclusion and discussion

This article reveals that predictability is enhanced by the increasing strength of the ordinal structures, such as the short-term memory, long-term memory, multifractal patterns, and chaotic attractors, which commonly exist in real-world time series. Since the time series' complexity and one-step variability are reduced with the increasing strength of these well-defined structures, the prediction models and methods can

Table 2 Details for real-world time series

Variable	Time range	Data length	Prediction length	$1 - \text{WPE}$	MASE (LMA)	MASE (AR)
TEM in Valkenburg	1976–2017 (daily)	15,000	2000	0.18	0.98	0.96
ENSO	1980–2017 (daily)	13,605	2000	0.43	0.86	2.19
AMOC	2004–2017 (daily)	4706	2000	0.73	0.21	2.79

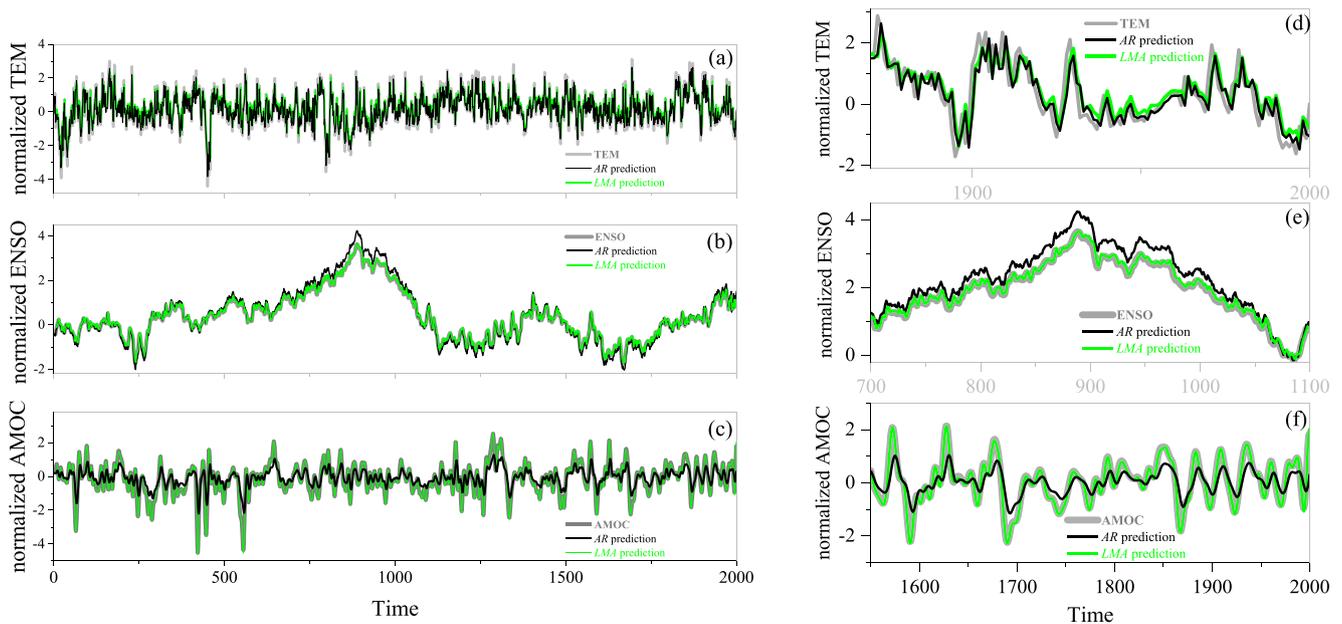


Fig. 11 Comparison between predicted time series (black lines from AR(4) and green lines from LMA) and testing real-world time series: (a) TEM, (b) ENSO, and (c) AMOC. (d)–(f) Locally enlarged version for (a)–(c), respectively

depict and predict the temporal variations of strengthening ordinal structures better. Detailed studies on the intrinsic predictability (quantified by $1 - \text{WPE}$) and prediction accuracy (by FE or MASE) for these four kinds of theoretical series with known differential ordinal structures show that there is an own specific regime or phase in the intrinsic predictability or prediction accuracy for each kind of theoretical series. Deterministic and nonlinear series take the higher intrinsic predictability or prediction accuracy (the lower forecast error), whereas linear and stochastic series have the lower intrinsic predictability or prediction accuracy (the higher forecast error).

The well corresponding relation between the intrinsic predictability and the prediction accuracy for each specific series with its own ordinal structures indicates there is specific regime in the $(1 - \text{WPE})$ -MASE plot for each kind of series with specific ordinal structures. Only from the estimated $1 - \text{WPE}$ for any given series can one determine which regime this underanalyzing series belongs to. This can guide us to preselect and optimize the suitable model or method to model or predict this series with unknown ordinal structures from the real world. Taking this insight into account, we analyze three climate series, i.e., daily air temperature anomaly in Valkenburg (TEM) from 1976 to 2017, daily indices of El Niño-Southern Oscillation (ENSO) from 1980 to 2017, and daily indices of Atlantic Meridional Overturning Circulation (AMOC) from 2004 to 2017. From the estimated $1 - \text{WPE}$ for these three different series, 0.18 for the temperature anomaly, 0.43 for the ENSO index, and 0.73 for the AMOC index, we can easily classify the daily air temperature anomaly in Valkenburg as series with the short-term memory and the daily ENSO index and daily AMOC index as nonlinear series.

Further prediction studies on these series confirm that the AR model is enough to the daily air temperature anomaly in Valkenburg, and this result is consistent with the previous findings that higher-frequency daily surface temperature fluctuations can be well modeled by the AR model after proper detrending (von Storch and Zwiers 1999; Bartos and Janosi 2005). However, the AR model fails to model and predict the daily ENSO index and daily AMOC index. Especially, there is substantial variability on short time scales of a few days (Balan Sarojini et al. 2011; Cunningham et al. 2007) in the AMOC index taking the chaos-like behaviors, so a model or method that can address chaotic series is required. Whereas for the high-frequency ENSO index, there are more complicated features with multiple periods and there is good memory and no single scaling (Petroni and Ausloos 2008), which are certainly different from those of the linear stochastic processes. Since the nonlinear strategy like LMA is computationally expensive, the estimation of $1 - \text{WPE}$ for any given series is simple with a little computation cost. Only from the estimated $1 - \text{WPE}$ for any given series can we optimize our modeling or prediction strategy in advance on which modeling or prediction strategy we should choose.

It should be pointed out that there are many other methods to infer the time series' intrinsic predictability, such as mean prediction time (Salvino et al. 1995), fractal dimension (Rangarajan and Sant 1997), memory or persistence (Franzke and Woollings 2011), Lyapunov exponents, and improved Lyapunov exponents (Patil et al. 2001; Ding et al. 2010; Ding et al. 2011). Here, the choice of WPE (Garland et al. 2014; Fu et al. 2019; Pennekamp et al. 2018) in this study is due to its sensitiveness to different structures and robustness to different transformation (Garland et al. 2014;

Fu et al. 2019; Pennekamp et al. 2018). In addition, although one-step prediction is investigated here, the relevant results are qualitatively similar for multistep prediction. But there may be more marked differences between linear and nonlinear methods for series with nonlinear ordinal patterns, since nonlinear behaviors will be dominant for multistep prediction (Sugihara 1990).

Funding information This research was supported by the National Natural Science Foundation of China through grants (No. 41675049 and No. 41475048).

References

- Babu CN, Reddy BE (2014) A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data. *Appl Soft Comput* 23:27–38
- Balan Sarojini B, Gregory JM, Tailleur R, Bigg GR, Blaker AT, Cameron DR, Edwards NR, Megann AP, Shaffrey LC, Sinha B (2011) High frequency variability of the Atlantic meridional overturning circulation. *Ocean Sci* 7:471–486
- Bandt C (2005) Ordinal time series analysis. *Ecol Model* 182:229–238
- Bandt C, Pompe B (2002) Permutation entropy: a natural complexity measure for time series. *Phys Rev Lett* 88:174102
- Bartos I, Janosi IM (2005) Atmospheric response function over land: strong asymmetries in daily temperature fluctuations. *Geophys Res Lett* 32:L23820
- Basu S, Foufoula-Georgiou E (2002) Detection of nonlinearity and chaoticity in time series using the transportation distance function. *Phys Lett A* 301:413–423
- Bauer P, Thorpe A, Brunet G (2015) The quiet revolution of numerical weather prediction. *Nature* 525:47–55
- Boffetta G, Cencini M, Falcioni M, Vulpiani A (2002) Predictability: a way to characterize complexity. *Phys Rep* 356:367–474
- Cunningham SA, Kanzow T, Rayner D, Baringer MO, Johns WE, Marotzke J, Longworth HR, Grant EM, Hirschi JJM, Beal LM, Meinen CS, Bryden H (2007) Temporal variability of the Atlantic meridional overturning circulation at 26.5N. *Science* 317:935–938
- Dakos V, Soler-Toscano F (2017) Easuring complexity to infer changes in the dynamics of ecological systems under stress. *Ecol Complex* 32: 44–155
- Dietze MC (2017) Prediction in ecology: a first-principles framework. *Ecol Appl* 27:048–2060
- Ding R, Li J, Seo KH (2010) Predictability of the Madden-Julian oscillation estimated using observational data. *Mon Weather Rev* 138: 1004–1013
- Ding R, Li J, Seo KH (2011) Estimate of the predictability of boreal summer and winter intra-seasonal oscillations from observations. *Mon Weather Rev* 139:2421–2438
- Elsner JB, Tsonis AA (1992) Nonlinear prediction, chaos, and noise. *Bull Am Meteorol Soc* 73:49–60
- Fadlallah B, Chen B, Keil A, Principe J (2013) Weighted-permutation entropy: a complexity measure for time series incorporating amplitude information. *Phys Rev E* 87:022911
- Franzke C, Woollings T (2011) On the persistence and predictability properties of North Atlantic climate variability. *J Clim* 24:466–472
- Fraser AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. *Phys Rev A* 33:1134–1140
- Fu S, Huang Y, Feng T, Nian D, Fu Z (2019) Regional contrasting DTR's predictability over China. *Phys A* 521:282–292
- Garland J, James R, Bradley E (2014) Model-free quantification of time-series predictability. *Phys Rev E* 90:052910
- Granger CW, Joyeux R (1980) An introduction to long-memory time series models and fractional differencing. *J Time Ser Anal* 1:15–29
- Graves T, Gramacy R, Watkins N, Franzke C (2017) A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA, 1951–1980. *Entropy* 19:437
- Höll M, Kantz H (2015) The fluctuation function of the detrended fluctuation analysis: investigation on the AR(1) process. *Eur Phys J* 88: 1–9
- Hyndman RJ, Koehler AB (2006) Another look at measures of forecast accuracy. *Int J Forecast* 22:679–688
- Ing CK, Wei CZ (2003) On same-realization prediction in an infinite-order autoregressive process. *J Multivar Anal* 85:130–155
- Kantelhardt JW, Zschiegner SA, Koscielny-Bunde E, Havlin S, Bunde A, Stanley HE (2002) Multifractal detrended fluctuation analysis of nonstationary time series. *Phys A* 316:87–114
- Kennel MB (1992) Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys Rev A* 45: 3403–3411
- Li Q, Fu Z (2014) Permutation entropy and statistical complexity quantifier of non-stationarity effect in the vertical velocity records. *Phys Rev E* 89:012905
- Lorenz EN (1963) Deterministic non-periodic flow. *J Atmos Sci* 20:130–141
- Lorenz EN (1969) Atmospheric predictability as revealed by naturally occurring analogues. *J Atmos Sci* 26:636–646
- Lorenz EN (1996) Predictability: a problem partly solved. *Proc. ECMWF Seminar on Predictability, vol I, Reading, United Kingdom, ECMWF, pp 40–58*
- Massah M, Kantz H (2016) Confidence intervals for time averages in the presence of long-range correlations: a case study on earth surface temperature anomalies. *Geophys Res Lett* 43:9243–9249
- Molgedey L, Ebeling W (2000) Local order, entropy and predictability of financial time series. *Eur Phys J* 15:733–737
- Nian D, Fu Z (2019) Extended self-similarity based multi-fractal detrended fluctuation analysis: a novel multi-fractal quantifying method. *Commun Nonlinear Sci Numer Simul* 67:568–576
- Patil DJ, Hunt BR, Kalnay E, Yorke JA, Ott E (2001) Local low dimensionality of atmospheric dynamics. *Phys Rev Lett* 86:5878–5881
- Pennekamp F, Iles A, Garland J et al (2018) The intrinsic predictability of ecological time series and its potential to guide forecasting. <https://doi.org/10.1101/350017>
- Petroni F, Ausloos M (2008) High frequency intrinsic modes in El Niño-Southern Oscillation Index. *Phys. A* 387:5246–5254
- Provenzale A, Smith LA, Vio R, Murant R (1992) Distinguishing between low-dimensional dynamics and randomness in measured time series. *Phys D* 58:31–49
- Rangarajan G, Sant DA (1997) A climate predictability index and its applications. *Geophys Res Lett* 24:1239–1242
- Riedl M, Müller A, Wessel N (2013) Practical considerations of permutation entropy. *Eur Phys J* 222:249–262
- Rybski D, Bunde A, Havlin S, Kantelhardt JW, Koscielny-Bunde E (2011) Detrended fluctuation studies of long-term persistence and multifractality of precipitation and river runoff records. *J Hydrol* 111:216–248
- Salvino LW, Cawley R, Grebogi C, Yorke JA (1995) Predictability in time series. *Phys Lett A* 209:327–332
- Schmitt F, Schertzer D, Lovejoy S (2000) Multi-fractal fluctuations in finance. *Int J Theor Appl Fin* 3:361–364
- Sugihara G (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series' nature. *Nature* 344:734–741

- Sugihara G, May R, Ye H, Hsieh CH, Deyle E, Fogarty M, Munch S (2012) Detecting causality in complex ecosystems. *Science* 338: 496–500
- von Storch H, Zwiers FW (1999) *Statistical analysis in climate research*. Cambridge Univ Press, Cambridge
- Ye Z, Hsieh WW (2008) Enhancing predictability by increasing nonlinearity in ENSO and Lorenz systems. *Nonlinear Process Geophys* 15: 793–801
- Yuan N, Fu Z, Mao J (2013) Different multi-fractal behaviors of diurnal temperature range over the north and the south of China. *Theor Appl Climatol* 112:673–682
- Yuan N, Huang Y, Duan J, Zhu C, Xoplaki E, Luterbacher J (2018) On climate prediction: how much can we expect from climate memory? *Clim Dyn* 52:855–864. <https://doi.org/10.1007/s00382-018-4168-5>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.