

Article

Forecasting for Chaotic Time Series Based on GRP-lstmGAN Model: Application to Temperature Series of Rotary Kiln

Wenyu Hu and Zhizhong Mao *

School of Information and Sciences, Northeastern University, Shenyang 110819, China

* Correspondence: maozhizhong@ise.neu.edu.cn

Abstract: Rotary kiln temperature forecasting plays a significant part of the automatic control of the sintering process. However, accurate forecasts are difficult owing to the complex nonlinear characteristics of rotary kiln temperature time series. With the development of chaos theory, the prediction accuracy is improved by analyzing the essential characteristics of time series. However, the existing prediction methods of chaotic time series cannot fully consider the local and global characteristics of time series at the same time. Therefore, in this study, the global recurrence plot (GRP)-based generative adversarial network (GAN) and the long short-term memory (LSTM) combination method, named GRP-lstmGAN, are proposed, which can effectively display important information about time scales. First, the data is subjected to a series of pre-processing operations, including data smoothing. Then, transforming one-dimensional time series into two-dimensional images by GRP makes full use of the global and local information of time series. Finally, the combination of LSTM and improves GAN models for temperature time series prediction. The experimental results show that our model is better than comparison models.

Keywords: chaotic time series; global recurrence plot; generative adversarial network; long short-term memory



Citation: Hu, W.; Mao, Z. Forecasting for Chaotic Time Series Based on GRP-lstmGAN Model: Application to Temperature Series of Rotary Kiln. *Entropy* **2023**, *25*, 52. <https://doi.org/10.3390/e25010052>

Academic Editors: Jaroslaw Krzywanski, Yunfei Gao, Marcin Sosnowski, Karolina Grabowska, Dorian Skrobek, Ghulam Moeen Uddin, Anna Kulakowska, Anna Zylka and Bachil El Fil

Received: 3 November 2022

Revised: 16 December 2022

Accepted: 19 December 2022

Published: 27 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rotary kiln sintering temperature (ST) is one of the basic parameters in the system, which is a key and global controlled parameter to the sintering technology and determines the quality of the production and energy consumption. Therefore, research on ST is highly valuable for industrial applications.

ST influenced by many uncertain factors, such as the composition of the raw material and the structure of the rotary kiln. However, it also has complex nonlinear characteristics caused by physical and chemical changes. Therefore, the high-precision prediction method of the ST is a hot topic in academic circles. Early investigations used the conservations of mass and energy or the physicochemical equation to simulate ST [1]. With the development and application of soft sensor methods in the process industries, numerous cutting-edge techniques have been used to anticipate and measure ST, such as data-driven methods [2] and sintering image analysis [3,4].

Nevertheless, these techniques disregard the inherent nonlinear pattern in ST. The concepts of chaos theory and nonlinear dynamic analysis can not only enhance the understanding of the fundamental characteristic of time series but also help to extend the predictive capability. Lv made the first mention of the sintering system of rotary kiln is a fifth-order chaotic system [5,6]. The traditional prediction method of chaotic time series is based on the assumption that real systems are recursive, that is, the future state is the same or similar to a previous state. For example, in [7], two multivariate time series phase space reconstruction methods are established; the authors implement wind power forecasting using the least-square support vector machine algorithm. In [8], a weighted largest Lyapunov exponent method was presented to enhance prediction accuracy.

As opposed to the standard chaotic time series forecasting process, which involves fitting a model to the historical data and simulating future data utilizing fitted model, selecting the most appropriate forecasting model or averaging several models has been a popular alternative approach. For example, in [9], researchers proposed a hybrid model of a stacked autoencoder and modified particle swarm optimization for multivariate chaotic time series forecasting. In [10], the author developed the recurrent predictor neural network for predicting chaotic time series, which is more accurate than the universal learning network. In [11], the authors presented a gate recurrent unit-based deep recurrent neural network to forecast the time series of Lorenz, Rabinovich–Fabrikant, and the Rössler chaotic system, which presented better performance than the long short-term memory (LSTM) model.

Reviewing the literature on chaotic time series forecasting, we find that (i) most of the present approaches depend on the manual choice of an appropriate set of features, which determines the prediction accuracy that relies on the feature-extraction method and characteristics of data. More significantly, (ii) the current literature focuses on the local features of the time series, leaving global characteristics under-emphasized. For the ST with a large time lag, the global dynamics of the time series contain important information such as irregular physical and chemical changes of sintered materials. Inspired by the recent work of [12–14], this paper aims to explore chaotic time series forecasting based on the idea of GRP, from which the global and local features of the time series can be automatically extracted using computer vision algorithms. The main contributions of this paper are as follows:

- This paper proposes to use GRP to transform the series into a two-dimensional image, which gives a priori knowledge about similarity and predictability by explaining the global and local information of the time series.
- Compared with WGAN-gp, our GAN model employs one generator and two discriminators, which enhances the convergence speed of the GAN and generates a more realistic GRP.
- A new chaotic time series forecasting model is put forward, which combines GAN and LSTM to automatically extract the features of the time series by the convolutional neural network (CNN). It achieved the highest prediction accuracy compared to the other models.
- Most existing models do not consider the time delay between the model input and output variables; however, when using multiple variables for prediction, it is important to consider the correlation between time series. In this paper, we find that choosing the appropriate time delay can improve the prediction accuracy through experiments.
- Images cannot be used as an input to LSTM. Therefore, a suitable image feature-extraction method is needed to not only reduce the input vector dimension but also to reserve as much information as possible. This paper analyzes the impact of different feature-extraction methods on the prediction accuracy.

The structure of the rest of this paper is as follows: Section 2 formulates the temperature of rotary kiln predicting problem and introduces the background of our model. Section 3 describes our forecasting model. Section 4 describes various evaluation experiments and gives a qualitative discussion of the experimental results. Section 5 is conclusion. List of abbreviations/acronyms used in this article.

2. Preliminary Knowledge

2.1. Formulation of Temperature of Rotary Kiln Predicting Problem

The goal of predicting the problem of the ST is to use previous data to predict the future temperature. t_n denotes the temperature data at moment n . Then, the control variable at moment n is represented by x_n . We document a group of consecutive observed data as

t_0, t_1, t_2, \dots . The problem is to forecast the temperature value at a certain point in the future with j previous temperature data and control variables by continuous observation:

$$\hat{t}_{n+\tau} = \arg \max_{t_{n+\tau}} p(t_{n+\tau} | f(\hat{t}_{n-j+1}, \hat{t}_{n-j+2}, \dots, \hat{t}_n), \hat{x}_{n-j+1}, \hat{x}_{n-j+2}). \quad (1)$$

where $p()$ denotes predictive model and $f()$ stands for pre-processing of previous temperature data. Before describing our approach in detail, a brief introduction to the basics is needed.

2.2. Phase Space Reconstruction

In our model, the PSR is the first step to the pre-processing of a rotary kiln temperature time series, which can reflect the richer information of the chaotic time series. For the one-dimensional scalar time series $x(i) : 1 \leq i \leq N$ with infinite length, no noise, and a d dimensional chaotic attractor, an m dimensional embedded phase space can be found in the sense of topology invariance. Takens theoretically proved that the delay coordinate vector space is the differential homeomorphism of the attractor of the original dynamical system in the Euclidean space \mathbb{R}^m if the dimension m of the delay coordinate vector is greater than twice the attraction dimension d of the original dynamical system [15].

The delay-coordinate reconstruction of the high-dimension dynamic system from time series by two important parameters, time delay τ and embedding dimension m . The State y on phase space trajectory.

$$y_1, y_2, \dots, y_l \quad (2)$$

where $l = N - (m - 1)\tau$ and

$$y_j = (x_j, x_{j+\tau}, \dots, x_{j+(m-1)\tau}). \quad (3)$$

There are two dominant views on finding the embedding dimension m and the delay time τ . The first view is that d and m are mutually uncorrelated and that the appropriate embedding dimension should be found after finding the delay time. Typical methods for finding time delays are the autocorrelation method [16] and the mutual information method [17]. The method for finding the optimal embedding dimension is the false nearest neighbour (FNN) [18].

2.3. Recurrence Plot

Based on the PSR, RP can convert one-dimensional chaotic time series into two-dimensional matrices to extract more enrichment information using convolutional neural networks. The recurrence of the States is a fundamental property for deterministic dynamic systems with a certain degree of nonlinearity. Such a new graphical tool was first proposed for time series analysis by [19]. The recurrence plot was used to visualize the recurring patterns, which display important interpretable information about time scales. It interprets the internal structure of time series, gives a priori knowledge of similarity and prediction, and is an important method for analyzing the recurrence, chaos, and non-stationarity of time series. With the help of CNN, we can extract local and global information of chaotic time series from RP. The RP can be expressed by the following equation:

$$R_{ij} = \Theta(\varepsilon - \|x_i - x_j\|). \quad (4)$$

For time series $x(i) : 1 \leq i \leq N$. After the relevant theoretical calculations to determine the appropriate embedding dimension and delay time and then reconstruct the time series, the reconstructed dynamical system is:

$$u_i = [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}]. \quad (5)$$

where $i = 1, 2, \dots, N - (m - 1)\tau$. Calculate the distance between x_i and x_j in the reconstructed phase space.

$$S_{ij} = \|u_i - u_j\|. \quad (6)$$

where $i = 1, 2, \dots, N - (m - 1)\tau, j = 1, 2, \dots, N - (m - 1)\tau$. $\|\cdot\|$ stand for the norm. Calculate recurrence values.

$$R(i, j) = \Theta(\varepsilon_i - S_{ij}). \quad (7)$$

Select a suitable threshold. $\Theta(\cdot)$ stand for the Heaviside function:

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (8)$$

However, in practical applications, it is difficult to find a suitable threshold; the unthreshold recursive graph is intrinsically a direct expression of the distance matrix, which measures the similarity between two-phase space vectors.

$$GRP_{ij} = \|u_i - u_j\|. \quad (9)$$

2.4. Generative Adversarial Networks

In our method, GAN is used to generate the GRP ($rp \in R^{N \times N}$) of the temperature data from a set of random noise. We use $z \sim p_z$, which follows a uniform distribution with a maximum value of 1 and a minimum value of -1 , as the hidden variable. Then generator is defined as $G(z; \theta_g)$. It maps the random noise to data space ($rp \in R^{N \times N}$), where $rp = G(z)$. The discriminator is described as $D(x; \theta_d)$. The input of D is a real data x sample from P_{data} , a fake sample generated by G , or a fake sample generated by G . The output of D represents the probability of its considering the picture to be real or unreal, expressed as a scalar from 0 to 1. The basic structure of GANs is shown in Figure 1. GAN is essentially a transformation or mapping from hidden space to sample space.

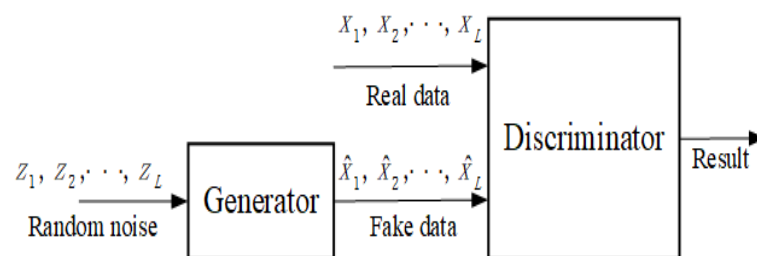


Figure 1. The basic structure of the GANs.

3. GRP-lstmGAN Model for Temperature Predict

We now present our GRP-LBP-lstmGAN model for ST forecast (see Figure 2). Firstly, the GRP of the ST is used for prediction rather than the temperature itself or the state after PSR. Since ST is a complex and chaotic time series, it is difficult to interpret and acquire complex nonlinear information from the data itself. Using PSR, we can reconstruct an equivalent system in state space by using the observed chaotic time series. The time delay and dimension embedding can be used to reconstruct the influence of other variables on the system, and GRP can show significant and easily explained local and global information about the time scale [19–21]. In addition to when a one-dimensional time series is transformed into an image, the noise is considered to be the illumination of the image by light, thus eliminating the effect of noise on the signal [22]. Second, We use the local binary pattern (LBP) to preprocess the GRP images before sending them into the network. Microscopic patterns such as the lines parallel to the main diagonal, the lines perpendicular to the main diagonal, the vertical (horizontal) lines, and the isolated points in the GRP reflect the characteristics of the time series, e.g., when two neighboring orbits in phase

space are continuously approaching one another during a period of evolutionary time, the lines parallel to the main diagonal appear, and the length of the lines reflects the speed of the separation of the neighboring orbits. LBP can extract these microscopic patterns easily and rapidly.

Finally, the LBP of the GRP used as the input to lstmGAN, which combines the image generation capability of the GAN and the time series memory capability of LSTM. The LSTM model extracts the features of the GRP pattern moving along the main diagonal. The generator of the GAN transforms the output of LSTM into the new RP, and the discriminator of the GAN automatically extracts the features of the RP for predicting the ST. The training procedure can be separated into two steps. At first, the real GRP of the ST is used to train the GAN. Secondly, when the GAN is well trained, the LSTM and the GAN are connected together. However, the parameters of the generator of the GAN will be fixed. Because the number of parameters of the LSTM is smaller than the generator network, if we train them at once, the generators will play a dominant role. During the training process, the neural network will tend to discriminate whether the texture features of the input image are real or not rather than predicting the future temperature [14].

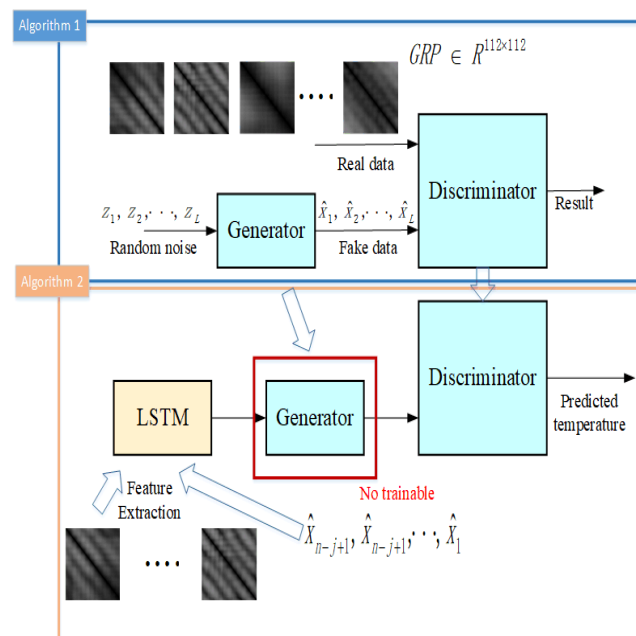


Figure 2. The structure of our proposed model.

4. Data Source

The data comes from the thermocouple data of No.5 kiln of Zhong-Zhou Aluminum Plant in Jiaozuo, Henan Province. Data length 205 h, sampling frequency 60 hz. Table 1 exhibits the statistical of the data set.

Table 1. Descriptive statistical of the data set.

	Mean	Std	Min	Max
temperature data	610.93	42.86	479.99	739.51

Figure 3 shows an example of the datasets for the model. Firstly, every 60 pieces of temperature data are averaged and smoothed (e.g., Data[0]). Secondly, one GRP is calculated for each 157 data (e.g., GRP[0]). Finally, the input x to the neural network consists of the texture features of the GRP (LBP for GRP, e.g., $X_{train}[0]$) and the control variables (the fuel consumption of rotary kiln, e.g., $Y_{train}[0]$).

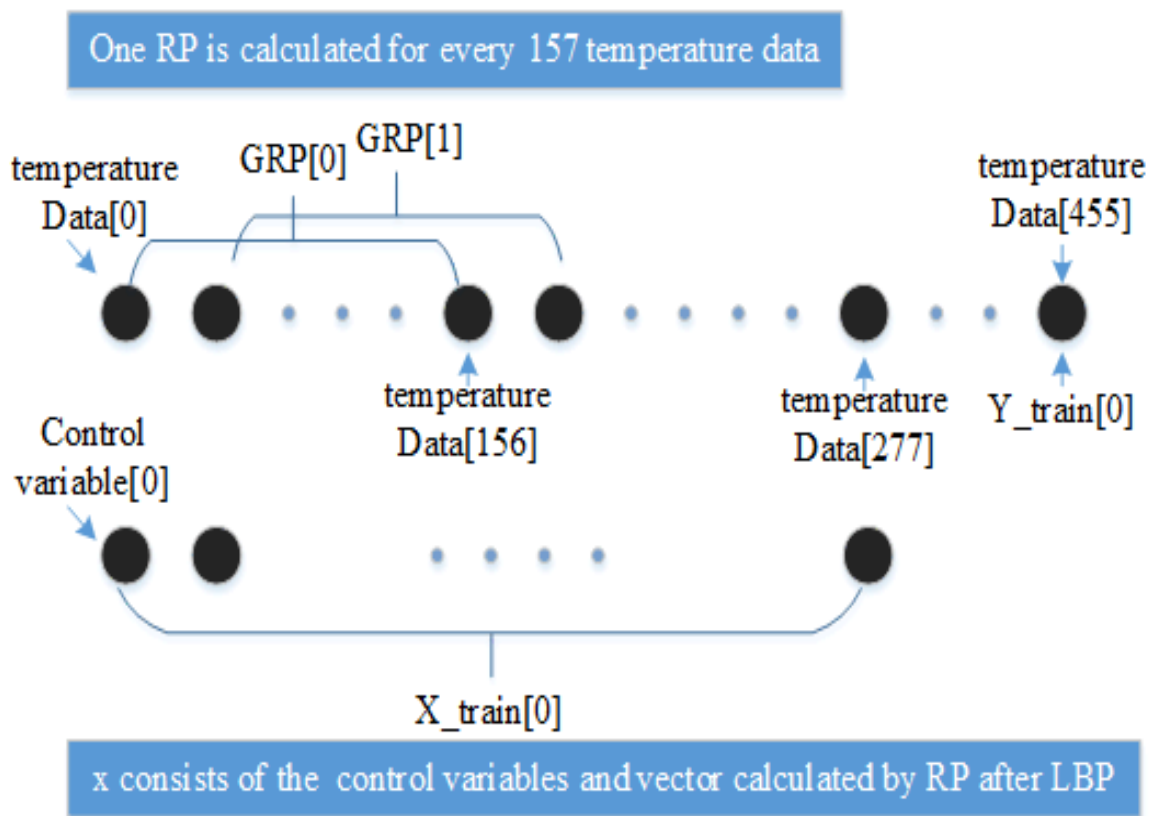


Figure 3. Making datasets for neural networks.

5. Experiments

In this section, we designed experiments to fulfill the following goals:

- To allow the WGAN-gp to generate a more realistic method than the original method.
- To compare the accuracy of our model with other prediction models for ST prediction.
- To improve the prediction accuracy by analyzing the time delay characteristics of the input and output.
- To improve the prediction accuracy by analyzing the features of different feature-extraction methods.

5.1. Compare of the GANs

The convergence problem of the GAN is still a pressing issue. The same model performs differently on varied datasets. The WGAN-gp trains faster and generates sharper images than the original GAN. However, the GRP is symmetrical strictly according to the main diagonal. To solve this problem, we propose to improve the WGAN-gp for RP (see Algorithm 1). The difference compared with the original one is that the loss function of the generator is divided into two parts. In the first part, the discriminator is the same. In the second part, the discriminator is to judge whether the output picture of the generator is symmetrical along the main diagonal.

$$\theta_G \leftarrow \text{RMSprop}(\nabla_{\theta_G} \frac{1}{m} \sum_{i=1}^m -D_{\theta_D}(G_{\theta_G}(z)) - \text{myloss}, \theta_G) \quad (10)$$

$$\text{where } \text{myloss} = 1 - \frac{i}{\text{Iteration}} * \frac{1}{n * n} * (|G_{\theta_G}(z) - G_{\theta_G}(z)^T|)$$

where $n * n$ indicates the number of GRP pixel points and $G_{\theta_G}(z) - G_{\theta_G}(z)^T$ represents the difference between the image output by the generator and the image transposition.

Algorithm 1 Improve WGAN-gp for RP Training Algorithm

Require: The batch size $m = 32$, the number of discriminator iterations per generator iteration $n_D = 5$, the gradient penalty coefficient $\lambda = 10$, the learning rate $lr = 0.0001$

Require: Initial generator parameters θ_G , initial discriminator parameters θ_D

```

while  $\theta_G$  has not converged do
  for  $t = 1, \dots, n_D$  do
    for  $i = 1, \dots, m$  do
      Sample real data  $x \sim P_{data}$ , a random number.
      Sample latent variable  $z \sim P_z$ .
      Sample  $\epsilon \sim U[0, 1]$ .
       $\tilde{x} \leftarrow G_{\theta_G}(z)$ 
       $\hat{x} \leftarrow \epsilon x + (1 - \epsilon)\tilde{x}$ 
       $L^{(i)} \leftarrow D_{\theta_D}(\tilde{x}) - D_{\theta_D}(x) + \lambda(\|\nabla_{\hat{x}} D_{\theta_D}(\tilde{x})\|_2 - 1)^2$ 
    end for
     $\theta_D \leftarrow Adam(\nabla_{\theta_D} \frac{1}{m} \sum_{i=1}^m L^{(i)}, \theta_D)$ 
  end for
  Sample a batch of latent variables  $\{z^{(i)}\}_{i=1}^m \sim P(z)$ .
   $\theta_G \leftarrow Adam(\nabla_{\theta_G} \frac{1}{m} \sum_{i=1}^m -D_{\theta_D}(G_{\theta_G}(z)) - myloss, \theta_G)$ 
  where  $myloss = 1 - \frac{i}{Iteration} \frac{1}{n*n} * (\|G_{\theta_G}(z) - G_{\theta_G}(z)^T\|)$ 
end while

```

To quantitatively describe whether the GRP generated by generator is symmetric or not, we define the symmetry degree:

$$SD = \frac{1}{n * n} * (\|G_{\theta_G}(z) - G_{\theta_G}(z)^T\|) \quad (11)$$

The smaller the SD, the better the symmetry of GRP.

According to the result shown in Figure 4, we can draw the following conclusions.

The Improved WGAN-gp for GRP has superior performance. Our proposed model can generate more symmetric GRP graphics. Firstly, the SD of our model is less than 0.1 after 100 iterations, but the original WGAN-gp requires 700 iterations. Secondly, the minimum value of SD for WGAN-gp is greater than our model. Finally, the training process of our model is more stable, without significant large fluctuations in SD.

It is not enough for GAN to generate symmetric graphs; in our task a good GAN should generate the most realistic GRP. Figure 5 shows the GRP generated by WGAN-gp and our model in different iterations. First of all, our model can generate clear images at 400 iterations; however, WGAN-gp needs 800 iterations. Moreover, in 5000 iterations, the GRP generated by our model is more realistic.

5.2. Accuracy of GRP-LSTMGAN Prediction

The objective in this part is to compare the accuracy of our proposed prediction model with others. In order to adequately test the superiority of our model, three algorithms are considered to compare the temperature-forecasting accuracy. These algorithms are the PSR-LSTM, the RP-LSTM, and the RP-LSTM-autoencoder, respectively. Based on the results shown in Table 2, the following conclusions can be drawn.

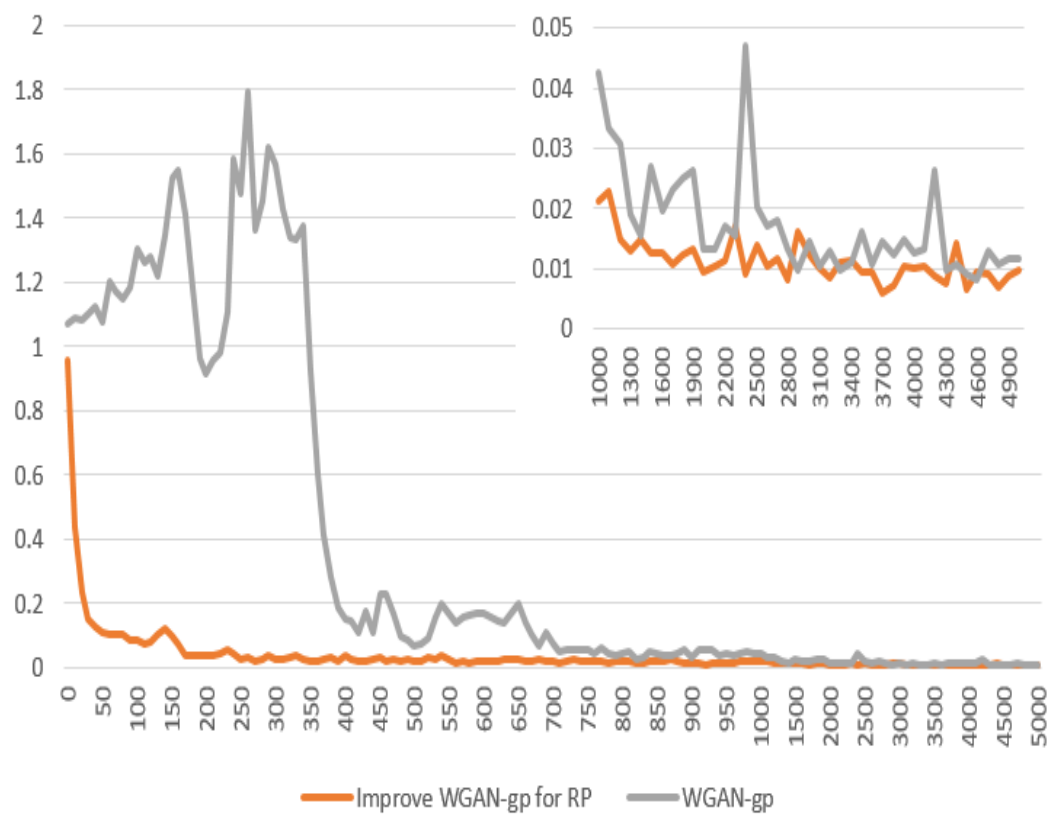


Figure 4. SD for WGAN-gp and Improved WGAN-gp for GRP. (X-axis is iteration, Y-axis is SD).

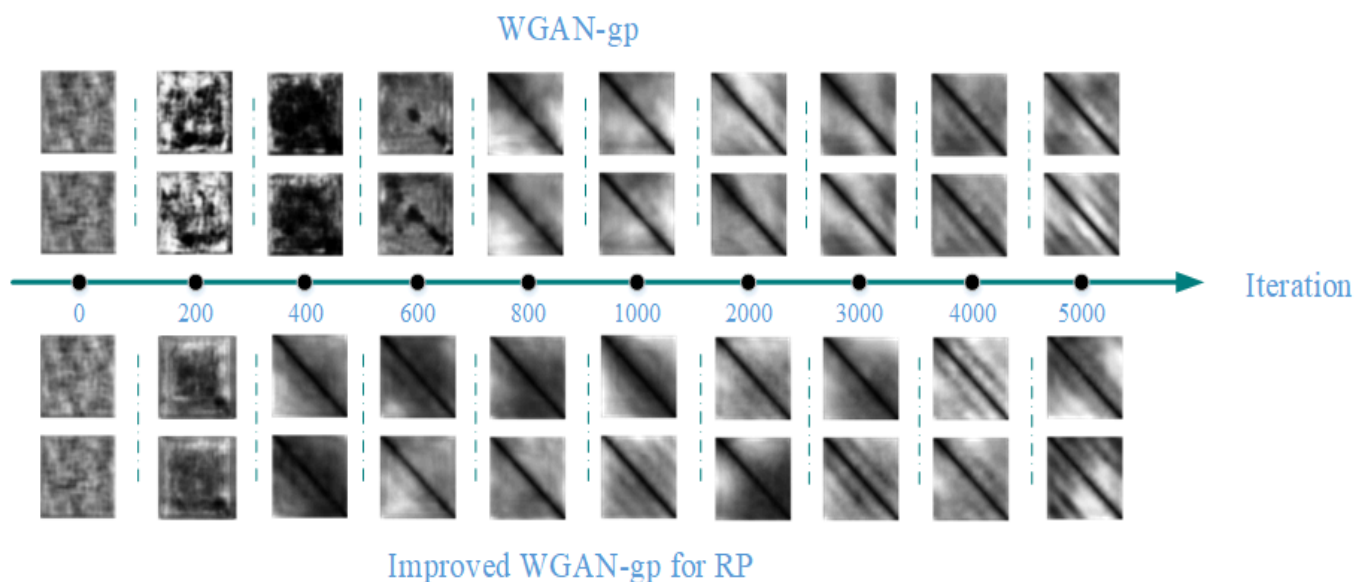


Figure 5. Samples generated by WGAN-gp and Improved WGAN-gp for RP. The moments for which the images are generated are the 0th, 200th, 400th, 600th, 800th, 1000th, 2000th, 3000th, 4000th, and 5000th iteration.

Table 2. Comparison of prediction performances of other models and proposed model.

Method	Logcosh	Mean Square Error	Absolute Temperature Error (°C)
PSR-LSTM	1.48×10^{-2}	3.02×10^{-2}	43.45
GRP-LSTM	1.78×10^{-4}	3.57×10^{-4}	4.73
GRP-LSTM -Autoencoder	9.83×10^{-5}	1.96×10^{-4}	3.51
Proposed model	8.03×10^{-5}	1.61×10^{-4}	3.16

1. By comparison with other methods, the proposed model has the highest prediction accuracy, and all of the indicators are optimal.
2. The precision of all of the methods using GRP is significantly higher than that using PSR. Further comparison of the precision of PSR-LSTM and GRP-LSTM shows that GRP provides more local and global information about the time series to improve the precision of the predictor compared to PSR (LSTM uses the same architecture, hyperparameters, and optimizer).
3. The GRP-LSTM-Autoencoder and the proposed model use the same architecture. For comparison, our proposed model is trained using Algorithm 2, and the GAN of GRP-LSTM will not be pre-trained. The lower error shows that our proposed training method can take advantage of the GAN image generation to strengthen the model-prediction accuracy.

Algorithm 2 GRP-LSTMGAN Training Algorithm

Require: The batch size $m = 16$, the learning rate $lr = 0.0005$

Require: Initial LSTM parameters θ_{lstm}

Train GAN with Algorithm 1.

Connecting GAN and LSTM, freeze θ_G parameters.

for number of training iterations **do**

 Sample $T_{i=1}^{(i)m}$ a batch from the pre-processing data.

$T^{(i)} = t_{i-k+1}, t_{i-k+2}, \dots, t_i$

$t_i = f(\hat{t}_{n-j+1}, \hat{t}_{n-j+2}, \dots, \hat{t}_n),$

$\hat{x}_{n-j+1}, \hat{x}_{n-j+2}, \dots, \hat{x}_n$

$D_{\theta_{lstm}} \leftarrow \nabla_{\theta_{lstm}} [\frac{1}{m} \sum_{i=1}^m \logcosh(D(G(f_{lstm}(T^{(i)}))), \hat{t}_{n+\tau})]$

$\theta_{lstm} \leftarrow \theta_{lstm} - lr \cdot RMSprop(\theta_{lstm}, D_{\theta_{lstm}})$

end for

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \quad (12)$$

By analyzing the data in Table 3, it can be concluded that our model has the highest accuracy compared to other delay times. On the one hand, the input data do not contain information about the high-delay-time temperature, which leads to a higher error in the high-delay-time prediction. On the other hand, the input data with low time delay contains redundant information, which will affect the prediction.

Table 3. Comparing the effect of different time delays on prediction accuracy.

Time Delay	188	Our Model	168	158	98	38
Logcosh	1.94×10^{-4}	8.03×10^{-5}	1.39×10^{-4}	9.88×10^{-5}	1.54×10^{-4}	1.26×10^{-4}
Mean Square Error	3.87×10^{-4}	1.61×10^{-4}	2.78×10^{-4}	1.98×10^{-4}	3.07×10^{-4}	2.52×10^{-4}
Absolute Temperature Error (C°)	4.92	3.17	4.16	3.51	4.38	3.97

5.3. The Effect of Time Delay

When using coal consumption as control variable to predict temperature, to obtain better prediction results, the temperature time series should have an appropriate time delay, which is important for the accuracy of the degree prediction and can improve the time scale of the prediction. The temperature time series and the coal-consumption time series generally have a strong correlation as it is demonstrated that the temperature time series will be very similar to the present coal-use time series after a period of delay. The Spearman correlation coefficient was used to quantitatively measure the correlation (see Figure 6). The Spearman correlation coefficient is defined as the Pearson correlation coefficient ρ between the rank variables.

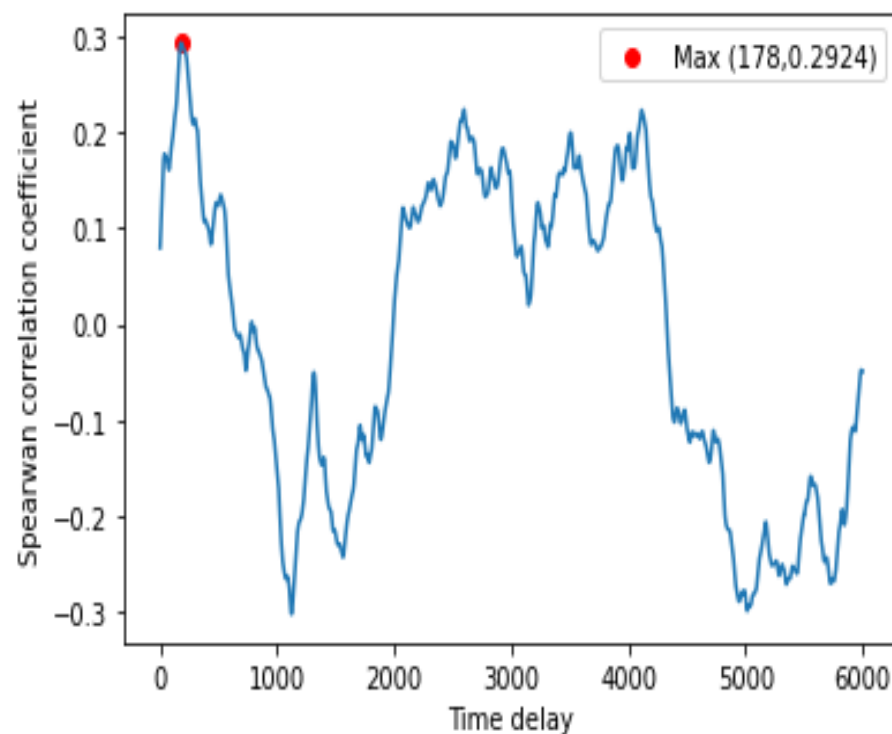


Figure 6. Spearman coefficients of temperature series with different time delays and levels of coal consumption (when the time delay is taken as 178, the a factor obtains the maximum value 0.2924).

5.4. Comparing The Effects of Different GRP Feature-Extraction Methods on Forecast Precision

We dimensionalize the image before sending it into the LSTM. If the input vector is too large, the convergence of the model will be very slow. In this section, we compare the impact of several LBP models Table 4 on prediction accuracy.

Table 4. Characteristics of different LBP models.

Default	Original LBP, which is gray-scale but not rotation-invariant [23].
Ror	Extension of default implementation, which is rotation-invariant and gray-scale.
Uniform	Improved rotation invariance with uniform patterns and finer quantization of the angular space which, is gray-scale and rotation-invariant [23].
Nri-uniform	Non rotation-invariant uniform patterns variant which is only gray-scale invariant [24,25].
Var	rotation-invariant variance measures of the contrast of local image texture which is rotation but not gray scale invariant.

The comparison results are shown in Table 5, and the following conclusions can be drawn. First of all, the uniform pattern is the most widely used; however, it has not obtained the best prediction accuracy. In the second place, the prediction accuracy of patterns with rotational invariance are low, probably because they cannot distinguish lines parallel to the main diagonal and horizontal (vertical) lines in the GRP. In the end, the default pattern achieves the highest accuracy, although the 256-dimensional feature vector may lead to excessive differences in the values of different features and cause errors. However, such a pattern can still extract more information from the GRP.

Table 5. Comparing the effects of different GRP feature-extraction methods on forecast precision.

Pattern	Logcosh	Mean Square Error	Absolute Temperature Error (C°)	Feature Vector Dimension
Ror	1.63×10^{-4}	3.25×10^{-4}	4.92	256
Default	8.03×10^{-5}	1.61×10^{-4}	3.17	256
Uniform	1.15×10^{-4}	2.31×10^{-4}	4.16	10
Nri-uniform	1.07×10^{-4}	2.14×10^{-4}	3.51	59
Var	1.84×10^{-4}	3.69×10^{-4}	4.38	59

6. Conclusions

In this paper, we have proposed to use GRP to transform the 1D chaotic time series into 2D images to obtain the global and local features and construct a GRP-lstmGAN prediction model with higher prediction accuracy than the existing methods. To improve the forecast accuracy, firstly, with respect to the traditional GAN with only one discriminator, the GAN was improved by adding a non-trainable discriminator, which determined whether the GRPs generated by generator are symmetric. The experiments showed that this method improves the model convergence speed and generates more realistic GRPs. Secondly, the correlation between the input variables coal consumption and temperature series was investigated; the highest correlation with the current coal consumption was obtained when the temperature series lagged 178 units, i.e., the model achieved the highest accuracy when predicting the temperature after 178 units of time by using the coal consumption and temperature data at the current moment. Finally, directly using the GRP image as an input to LSTM, the large input dimension will prevent the model from easily converging. In this paper, the experiments have compared different feature-extraction methods; LBP-default reduced the input from 112×112 of 2D image to 256 dimensions and had the highest prediction accuracy compared with other methods.

In our model, the performance of the GAN will affect the prediction accuracy. Knowing how to avoid the pattern collapse of the GAN is the future research direction.

Author Contributions: Formal analysis, W.H.; writing—review and editing, W.H. and Z.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

RP	Recurrence Plot
GRP	Global Recurrence Plot
GAN	Generative Adversarial Network
LSTM	Long Short-Term Memory
ST	Sintering Temperature
WGAN-gp	Wasserstein Generative Adversarial Network-Gradient Penalty
CNN	Convolutional Neural Network
LBP	Local Binary Pattern
SD	Symmetry Degree
PSR	Phase Space Reconstruction

References

- Boateng, A.; Barr, P. A thermal model for the rotary kiln including heat transfer within the bed. *Inte. J. Heat Mass Trans.* **1996**, *39*, 2131–2147. [\[CrossRef\]](#)
- Xu, J.; Fu, D.; Shao, L.; Zhang, X.; Liu, G. A soft sensor modeling of cement rotary kiln temperature field based on model-driven and data-driven methods. *IEEE Sens. J.* **2022**, *21*, 27632–27639. [\[CrossRef\]](#)
- Chen, H.; Zhang, X.; Hong, P.; Hu, H.; Yin, X. Recognition of the temperature condition of a rotary kiln using dynamic features of a series of blurry flame images. *IEEE Trans. Ind. Inform.* **2016**, *12*, 148–157. [\[CrossRef\]](#)
- Chen, H.; Yan, T.; Zhang, X. Burning condition recognition of rotary kiln based on spatiotemporal features of flame video. *Energy* **2020**, *211*, 118656. [\[CrossRef\]](#)
- Lv, M.; Zhang, X.; Chen, H.; Zhou, Y.; Li, J. Chaotic and multifractal characteristic analysis of noise of thermal variables from rotary kiln. *Non. Dyn.* **2020**, *99*, 3089–3111. [\[CrossRef\]](#)
- Zhang, X.; Lv, M.; Chen, H.; Dai, B.; Xu, Y. Chaotic characteristics analysis of the sintering process system with unknown dynamic functions based on phase space reconstruction and chaotic invariables. *Non. Dyn.* **2018**, *93*, 395–412. [\[CrossRef\]](#)
- Tian, J.; Jin, W.; Meng, L.; Qing, W. Short-term wind power forecast based on chaotic analysis and multivariate phase space reconstruction. *Energy Convers. Manag.* **2022**, *254*, 115196. [\[CrossRef\]](#)
- Wang, J.; Chi, D.; Wu, J.; Lu, H.Y. Chaotic time series method combined with particle swarm optimization and trend adjustment for electricity demand forecasting. *Expert Sys. Appl.* **2011**, *38*, 8419–8429. [\[CrossRef\]](#)
- Xu, X.; Ren, W. A hybrid model of the STacked autoencoder and modified particle swarm optimization for multivariate chaotic time series forecasting. *Appl. Soft Comput.* **2022**, *116*, 108321. [\[CrossRef\]](#)
- Han, M.; Xi, J.; Xu, S.; Yin, F.L. Prediction of chaotic time series based on the recurrent predictor neural network. *IEEE Trans. Sign. Proc.* **2004**, *52*, 3409–3416. [\[CrossRef\]](#)
- Serrano-Pérez, J.D.J.; Fernández-Anaya, G.; Carrillo-Moreno, S.; Yu, W. New result for prediction of chaotic systems using deep recurrent neural networks. *Neural Proc. Lett.* **2021**, *53*, 1579–1596. [\[CrossRef\]](#)
- Li, X.; Kang, Y.; Li, F. Forecasting with time series imaging. *Expert Syst. Appl.* **2020**, *160*, 113680. [\[CrossRef\]](#)
- Ben Said, A.; Erradi, A. Deep-Gap: A deep learning framework for forecasting crowdsourcing supply-demand gap based on imaging time series and residual learning. In Proceedings of the 2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom), Sydney, Australia, 11–13 December 2019; pp. 279–286.
- Xu, Z.; Du, J.; Wang, J.; Jiang, C.; Ren, Y. Satellite image prediction relying on GAN and LSTM neural networks. In Proceedings of the ICC 2019-2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019.
- Takens, F. Detecting strange attractors in fluid turbulence. In *Dynamical Systems and Turbulence*; Springer: Berlin/Heidelberg, Germany, 1981; Volume 898, pp. 366–381.
- Mackay, R.S. Nonlinear time series analysis. *Trends Biotech.* **1997**, *15*, 531–532. [\[CrossRef\]](#)
- Fraser, A.M. Information and entropy in strange attractors. *IEEE Trans. Inf. Theory* **2002**, *35*, 245–262. [\[CrossRef\]](#)
- Kennel, M.B.; Brown, R.; Abarbanel, H.D. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys Rev A.* **1992**, *15*, 3403–3411. [\[CrossRef\]](#)
- Eckmann, J.-P.; Kamphorst, S.O.; Ruelle, D. Recurrence plots of dynamical systems. *Europhys. Lett.* **1987**, *4*, 973–977. [\[CrossRef\]](#)
- Norbert, M.; Carmen, R.M.; Marco, T.; Jürgen, K. Recurrence plots for the analysis of complex systems. *Phys. Rep.* **2007**, *438*, 237–329.
- Pan, Z.; Fan, H.; Zhang, L. Texture classification using local pattern based on vector quantization. *IEEE Tran. Image Proc.* **2015**, *24*, 5379–5388. [\[CrossRef\]](#)
- Ruiz, M.; Mujica, L.E.; Alferez, S.; Acho, L.; Tutiven, C.; Vidal, Y.; Pozo, F. Wind turbine fault detection and classification by means of image texture analysis. *Mech. Syst. Signal Process.* **2018**, *107*, 149–167. [\[CrossRef\]](#)

23. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation-invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [[CrossRef](#)]
24. Ahonen, T.; Hadid, A.; Pietikainen, M. Face recognition with local binary patterns. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 469–481.
25. Ahonen, T.; Hadid, A.; Pietikainen, M. Face description with local binary patterns: Application to face recognition. *IEEE Trans. Patt. Anal. Mach. Intell.* **2006**, *28*, 2037–2041. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.