Wind Power Time Series Missing Data Imputation Based on Generative Adversarial Network

Hang Fan
Tsinghua University
Department of Electrical Engineering
Beijing, China

Xuemin Zhang
Tsinghua University
Department of Electrical Engineering
Beijing, China
zhangxuemin@tsinghua.edu.cn

Shengwei Mei
Tsinghua University
Department of Electrical Engineering
Beijing, China

2021 IEEE 4th International Electrical and Energy Conference (CIEEC) | 978-1-7281-7149-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/CIEEC50170.2021.9510923

***Abstract*—With the large scale integration of wind farms, wind power prediction with high accuracy becomes more important. Many research focuses on developing advanced and complex methods to improve the prediction accuracy based on statistical learning, which is highly dependent on the data set. But due to the wind curtailment or the equipment fault, there are some missing data in the wind power time series and enhancing the data quality of training data is crucial. Most research about wind power imputation is based on the power and wind speed curve. However, for some wind farms and the dispatching center, there are no complete wind speed data available. Therefore, we developed a generative adversarial network (GAN) based method for the wind power time series correction. When there is no wind speed measurement, it still can impute wind power. Besides, it can take into consideration the wind power distribution when imputing the wind power time series. Compare to the traditional way such as mean value filled method and other machine learning methods, it has higher imputation accuracy. Case studies on the field wind farm data showed that it outperformed other classical imputation methods.**

Keywords—Wind Power Imputation; wind power prediction; generativee adversarial network; machine learning;

I. INTRODUCTION

In order to cope with climate change, renewable energy especially the wind energy is developing rapidly [1]. The new installed capacity of wind farms increased year by year, which not only brings clean energy but also leads to a severe challenge to the power grid due to the stochastic characteristics of wind energy. Wind power prediction is a very important issue to maintain the stable and economic operation of the power system [2]. Many researches concentrate on using complex method to improve the performance of the wind power prediction system. But the wind power data of high quality also plays a significant role in the improvement of the wind power prediction accuracy.

However, in practice missing data in the wind power time series is a kind of common phenomenon due to the wind curtailment policy or the equipment fault. Some researches dedicated to solve this problem especially on the identification and reconstruction of missing data in wind farms. Reference [3] used the quartile method to find out the abnormal wind power data and reconstructed them by using the cubic spline method. Zhou [4] developed a semi-supervised anomaly detection method for wind farm power data preprocess based on the power and wind speed curve. Reference [5] put forward the change point analysis method of abnormal data, which can consider the time series information of the wind power data. Reference [6] used the local outlier factor algorithm to assess the anomaly of the wind power data. Research [7] reconstructed the wind power based on the time delay of the nearby wind farms. Ye [8] used the copula function to draw the confidence interval of wind power under the condition of given wind speed and used the output correlation among multiple wind farms to identify and correct the abnormal data of wind farms. The time series information of wind farm power was also considered in this model. Zhao [9] combined the DBSCAN method and quantile method to detect the anomaly and used the fitted wind power and speed curve to impute the abnormal wind power data. The Extreme Learning Machine (ELM) and granger test were adopted in [10] for the missing data imputation.

However, most of those methods rely on the wind speed for the fitting of the power and wind speed curve in the imputation process. But in some cases, there is no wind speed measurement data and it is hard to use the wind power curve to correct those data. The traditional way is to impute those missing values according to the mean value of the nearest time or the linear fitting of the spatiotemporal relationship. But it does not take into consideration the nonlinear relation of nearby wind farms as well as the bidirectional mechanism of wind power time series. Besides, the missing data can be divided into two patterns, namely the random pattern and the continual missing pattern. In the continual missing pattern, the missing data can last for a long time due to equipment malfunction or regular maintenance. We can directly delete this period in the downstream prediction task. However, the fluctuation of the power grid and the loss of transmission packages can lead to random missing scenarios. It can also last for a long time but it is not convenient and unnecessary for us to directly delete them. This paper mainly focused on the imputation of the random missing pattern when there is no wind speed measurement. In this case, the wind power spatiotemporal relationship of multiple wind farms should be utilized effectively to impute the missing values.

The Recurrent Neural Network (RNN) especially the

This work is supported by National Key R&D Program of China (Technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDKOOKJJS1800266).

Gated Recurrent Unit (GRU) is a classic structure to model the temporal relationship. The bidirectional GRU can even extract the bidirectional feature of the sequence which is very suitable for the imputation problem. The generative adversarial network (GAN) is also an effective unsupervised way to model the distribution of the data. Therefore, we proposed our method to impute the missing data by using the combination of GAN and refined bidirectional GRU model when there is no wind speed measurement. Our contribution is twofold:

1) We use the Bi-GRU for the time series imputation which can take into consideration the bidirectional spatiotemporal relationship of wind farms compared to other linear regression methods or neural network methods.

2) We adopted the generative adversarial network to learn the distribution of the wind power time series. Therefore, the imputed wind power time series curve can be more similar to the sampling results from the real wind power process.

The rest part of the paper is organized as follows. Section II is the mathematical formulation of the missing data imputation. Section III presents the Generative adversarial network for the missing data imputation. Case studies are conducted in section IV. Finally the conclusion is reached in section V.

II. The Formulation of Missing Data Imputation

For the wind farm cluster with d wind farms, the wind power time series X observed in $T=\left(t_0,\ldots,t_{n-1}\right)$ is represented as $X=\left(x_{t_0},\ldots,x_{t_i},\ldots,x_{t_{n-1}}\right)^T \in R^{n\times d}$, where x_{t_i} is the t_i th observation about X and $x_{t_i}^j$ is the j th variable of x_{t_i}. The missing data can be divided into random pattern and continual missing pattern. In the continual pattern, we can use the wind power time series of nearby wind farm to approximate or use the multi-step prediction method to impute. This paper we just focus on the random missing pattern which is as follows.

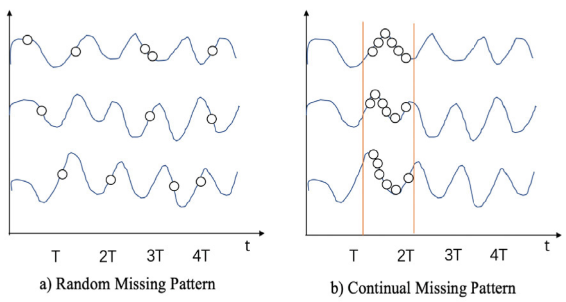

Fig. 1. The Missing Pattern of Wind Power Time Series

In the following example, there are 4 wind farms and 3 observations. “none” is used to represent the missing value.

$$X=\begin{bmatrix} x_{t_0}^1 & x_{t_0}^2 & none & x_{t_0}^4 \\ x_{t_1}^1 & none & x_{t_1}^3 & none \\ x_{t_2}^1 & none & x_{t_2}^3 & none \end{bmatrix} \tag{1}$$

$$T=\begin{bmatrix} t_0 \\ t_1 \\ t_2 \end{bmatrix} \tag{2}$$

The mask matrix of sample $M\in R^{n\times d}$ is introduced to represent the location of missing values. When $M_{t_i}^j=0$, there exits missing values, otherwise $M_{t_i}^j=1$. The imputation of wind power time series is filling the missing values according to the spatiotemporal relationship.

III. The Generative Adversarial Network For Data Imputation

A. General GAN Architecture

Generative adversarial network[11] is a classical machine learning model. It consists of two parts, which are generator network (G) and discriminator network (D). The generator network G learns a function $G(Z)$ which can change the random noise vector Z into a realistic time series. The discriminator network D is used to compute the input data's probability of being real. The input of D includes both the real but incomplete samples and generated but complete samples generated by G. WGAN which uses the Wasserstein distance is used to avoid the mode collapse problem in the traditional GAN model and we also use WGAN in this research. The training of WGAN can be formulated as a min-max game problem which is as follows[12].

$$\min_{\theta_G}\max_{\theta_D} E_{X\sim P_d}\left(D(X)\right)-E_{Z\sim P_g}\left(D\left(G(Z)\right)\right) \tag{3}$$

Where P_d is the distribution of real samples and P_g describes the distribution of generated samples. In equation (3), the discriminator increase the ability to discern the real samples and the generated samples by adjusting parameter θ_D. The generator pushes the generated distribution as close as the real distribution by adjusting parameter θ_G.

B. The Bidirectional Gate Reccurrent Unit for Imputation

Since there exists apparent temporal relationship in the input of the samples , the GRU is used to construct the basic structure of G and D. The GRU can be replaced by other recurrent neural networks. But because of the data incompleteness, two consecutive valid observations vary a lot and the traditional GRU and LSTM cannot handle this situation. They cannot take into consideration of the bidirectional mechanism of the wind power time series. Therefore, the Bidirectional Gated Recurrent Unit for data Imputation(BGRUI) is designed to replace GRU.

In our case, we mainly deal with the anomaly with missing values which are filled by "none" values traditionally. The time lag of two consecutive valid observations is always varying due to the uncertain location of those "none" values. The time lag between two observations are very important because those lags which represent the influence of the past observations is from an unknown distribution actually. When there exists missing value, those influence of past values will decay. Therefore, the GRU for imputation (GRUI) is utilized to fit the decayed influence of the past observations [13]. The time lag matrix $\delta \in R^{n\times d}$ is introduced as follows to record the time lag between current value and last valid value [14].

$$\delta_{t_i}^j = \begin{cases} t_i - t_{i-1}, & i > 0 \,\&\, M_{t_{i-1}}^j = 1 \\ \delta_{t_{i-1}}^j + t_i - t_{i-1}, & i > 0 \,\&\, M_{t_{i-1}}^j = 0 \\ 0, & i = 0 \end{cases} \tag{4}$$

Therefore, in the example in equation (1), the δ can be represented as follows.

$$\begin{aligned} \delta &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ t_1 - t_0 & t_1 - t_0 & t_1 - t_0 + \delta_0^3 & t_1 - t_0 \\ t_2 - t_1 & t_2 - t_1 + \delta_2^2 & t_2 - t_1 & t_2 - t_1 + \delta_2^4 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ t_1 - t_0 & t_1 - t_0 & t_1 - t_0 & t_1 - t_0 \\ t_2 - t_1 & t_2 - t_0 & t_2 - t_1 & t_2 - t_0 \end{bmatrix} \end{aligned} \tag{5}$$

The decay vector β is used to control the influence of the past observations. When δ is larger, the β is smaller. Decay vector β is a value bigger than zero and smaller than one. The decay vector β can be described as follows.

$$\beta_{t_i} = 1 \Big/ e^{\max\left(0, W_\beta \delta_{t_i} + b_\beta\right)} \tag{6}$$

Where W_β and b_β are parameters that need to learn. According to the definition in (6), the decay vector β_{t_i} can be a real number that smaller than 1 and bigger than 0. The hidden state h_{t_i-1} in the GRU can be updated by element-wise multiplying the decay vector β_{t_i} which is as follows.

$$h_{t_i-1}^{'(1)} = \beta_{t_i} \odot h_{t_i-1} \tag{7}$$

Where $\odot$ means the element-wise multiplication. $h_{t_i-1}^{'(1)}$ is the clockwise hidden state. But due to the bidirectional mechanism in wind power [10], we designed the bidirectional GRUI. Due to the directional mechanism, in the imputation process, the future value of the wind power fragment can be used to repair the missing value. The hidden state in the bidirectional GRUI is in equation (8).

$$h_{t_i-1}^{'} = \left(h_{t_i-1}^{'(1)} + h_{t_i-1}^{'(2)}\right) \Big/ 2 \tag{8}$$

For each direction, the state update function of GRUI for the time series imputation is as follows.

$$\begin{aligned} \mu_{t_i}^{(1)} &= \sigma\left(W_\mu^{(1)}\left[h_{t_i-1}^{'(1)}, x_{t_i}\right] + b_\mu^{(1)}\right) \\ r_{t_i}^{(1)} &= \sigma\left(W_r^{(1)}\left[h_{t_i-1}^{'(1)}, x_{t_i}\right] + b_r^{(1)}\right) \\ \tilde{h}_{t_i}^{(1)} &= \tanh\left(W_{\tilde{h}}^{(1)}\left[r_{t_i}^{(1)} \odot h_{t_i-1}^{'(1)}, x_{t_i}\right] + b_{\tilde{h}}^{(1)}\right) \\ h_{t_i}^{(1)} &= \left(1 - \mu_{t_i}^{(1)}\right) \odot h_{t_i-1}^{'(1)} + \mu_{t_i}^{(1)} \odot \tilde{h}_{t_i}^{(1)} \\ \mu_{t_i}^{(2)} &= \sigma\left(W_\mu^{(2)}\left[h_{t_i+1}^{'(2)}, x_{t_i}\right] + b_\mu^{(2)}\right) \\ r_{t_i}^{(2)} &= \sigma\left(W_r^{(2)}\left[h_{t_i+1}^{'(2)}, x_{t_i}\right] + b_r^{(2)}\right) \\ \tilde{h}_{t_i}^{(2)} &= \tanh\left(W_{\tilde{h}}^{(2)}\left[r_{t_i}^{(2)} \odot h_{t_i+1}^{'(2)}, x_{t_i}\right] + b_{\tilde{h}}^{(2)}\right) \\ h_{t_i}^{(2)} &= \left(1 - \mu_{t_i}^{(2)}\right) \odot h_{t_i+1}^{'(2)} + \mu_{t_i}^{(2)} \odot \tilde{h}_{t_i}^{(2)} \end{aligned} \tag{9}$$

Where $\mu_{t_i}^{(1)}$ and $\mu_{t_i}^{(2)}$ are the update gate, $r_{t_i}^{(1)}$ and $r_{t_i}^{(2)}$ are the reset gate, $\tilde{h}_{t_i}^{(1)}$, $\tilde{h}_{t_i}^{(2)}$, $h_{t_i}^{(1)}$ and $h_{t_i}^{(2)}$ are candidate hidden unit and σ is activation function. $W_{\tilde{h}}^{(1)}, W_{\tilde{h}}^{(2)}, W_r^{(1)}, W_r^{(2)}, W_\mu^{(1)}, W_\mu^{(2)}, b_\mu^{(1)}, b_\mu^{(2)}$, $b_r^{(1)}$ and $b_r^{(2)}$ is the model parameter. According to the update function, the spatiotemporal relationship of the wind power is considered in the imputation process.

C. Proposed GAN for Wind Power Imputation

For the traditional GAN, the random noise are added into the G (Generator) to generated synthesized data. But in our proposed model, the real samples which are the incomplete wind power time series and the random noise are added together to improve the ability of the G. D (Discriminator) is used to discern the synthesized samples from the real samples. Therefore, the objective function in WGAN is formulated in equation (10).

$$\min_{\theta_G} \max_{\theta_D} E_{X\sim P_d}\left(D(X)\right) - E_{Z\sim P_g}\left(D\left(G\left(X+Z\right)\right)\right) \tag{10}$$

In equation (10), G and D are the generator and discriminator in the GAN. The generator and discriminator are trained alternately to solve the model. The loss function of discriminator in WGAN is

$$\max_{\theta_D} E_{x\sim P_d}\left[D(x)\right] - E_{z\sim P_g}\left[D\left(G\left(X+Z\right)\right)\right] \tag{11}$$

The loss function of generator in WGAN is

$$\min_{\theta_G} E_{z\sim P_g}\left[D\left(G\left(X+Z\right)\right)\right] \quad (12)$$

The structure of the proposed GAN is as follows.

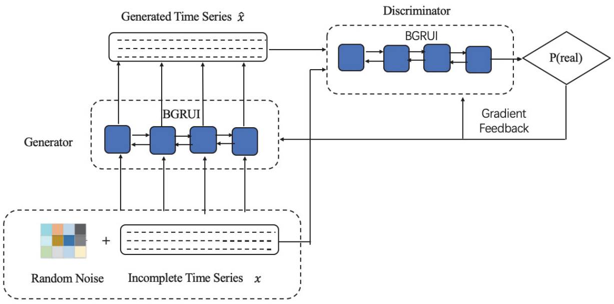

Fig. 2. The Generative Adversarial Network for Imputation

Where the generator and discriminator both are constructed by 1 layer bidirectional GRUI to capture the temporal relationship and 1 fully-connected layer for the sequence generation. The generator uses the incomplete sequence x and random noise to impute the sample and the imputed sequence $\hat{x}$ can be worked out. The training process is in figure 3. The discriminator can compare the generated sequence $\hat{x}$ and the original sequence x and when it cannot distinguish the difference between $\hat{x}$ and x, the adversarial network is converged. In this case, the generator can simulate the distribution of the original sequence.

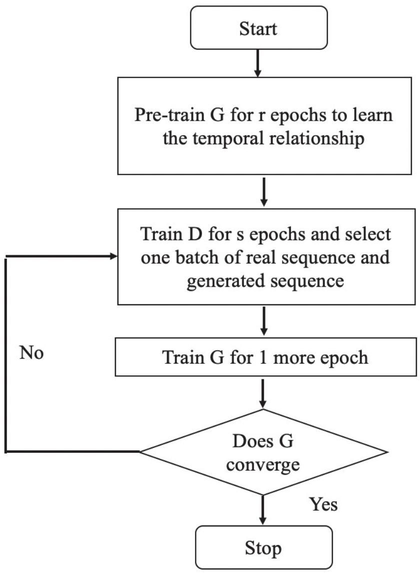

Fig. 3. The Training Process of Generative Adversarial Network

When the training process of GAN is finished, the parameter θ_G and θ_D is worked out and we can use the GAN for the time series imputation by adding random noise Z. However, there is one problem that how we can choose different random noise Z for different incomplete sequence X to impute the missing values. For this problem, we design the following objective function.

$$\min_{Z}\left\|X\circ M-G\left(Z+X\right)\circ M\right\|_2-\lambda D\left(G\left(Z+X\right)\right) \quad (13)$$

Where M is mask matrix to represent the missing values. There are two parts in this equation, the first part means that the non-missing values of the generated sequence should be as similar as that of the original incomplete sequence. The second part means the discriminator should cannot distinguish the generated sequence and the original incomplete sequence to guarantee the equilibrium of the generative adversarial network. λ is the hyper-parameter to balance the two parts. Since equation (13) is also an non-convex problem, the gradient descent method is used to solve it. By working out Z which can minimize equation (13), we can input it to the generator to get the missing values and it is shown in equation (14).

$$\widehat{X}=X\circ\left(1-M\right)+G\left(\widetilde{Z}+X\right)\circ M \quad (14)$$

IV. Case Study

A. The Training Dataset and Experiment Environment

The field wind farm cluster data is used for analysis. There are 20 wind farms in the dataset. There are 13041 samples in the dataset and the time interval of the samples is 15min. To illustrate the performance under different circumstances, the missing samples are randomly generated on the original data according to a certain miss rate p, and then the generative adversarial network is designed to learn the distribution of the data set and impute the missing values. The observation interval is 4 hour which means $X\in R^{20\times16}$ in this sample dataset. The test environment is Python 2.7 and TensorFlow 1.7.0. The root mean square error (RMSE) of the imputed samples and the real samples is selected to evaluate the performance of the imputation and prediction. The calculation method of RMSE is as follows.

$$RMSE=\sqrt{\frac{1}{K}\sum_{i=1}^{K}\left(x_{ti}-\widehat{x}_{ti}\right)^2} \quad (15)$$

In the practice, there are following common and classic methods for the time series imputation.

1) Mean Value Filled: Mean value filled method is a basic imputation process which replaces the missing value with the two values before and after the missing value.

2) K-Nearest Neighbor(KNN) [15]: The missing values can be represented by the mean value of k nearest neighbors. Therefore, the multivariate time series imputation can be realized by choosing k nearest neighbors in the high dimension space. In this case, k is fixed as 5.

3) Multivariate Imputation by Chained Equation(MICE) [16]: MICE is a kind of method widely used in imputation of solar data and traffic data. It is based on the chain of regression equation for the imputation of data. The iteration number is 1 and the neareast feature is 15.

4) Matrix Factorization (MF) [17,18]: Matrix Factorization method is also widely used in anomaly imputation of traffic data. It factorize the incomplete matrix into 2 low-rank matrix U and V by gradient descent. In the process of gradient descent, the $L1$ regulation is used on U and $L2$ regulation on V is considered. By the gradient process and the optimization of loss function, the missing values can be worked out at the optimal value of objective function. The learning rate is 0.01, rank is 8 and the L_2 penalty is 1e-5.

5) Generative Adversarial Network (GAN): It is the method which is used in this paper. When the training of the network is finished, the parameter which minimizes equation (11) will lead to the final imputation results. Research [14] used single direction GRU for the imputation and the effectiveness of this method has been tested in solar dataset. However, for the basic cell of the discriminator and generator, we can also use different units such as fully connected layer, and bidirectional GRUI to replace the single direction GRUI in refence [14]. Therefore in our case study, we also design model which is only based on the fully connected layer and model based on bidirectional GRUI. The hidden state number of the GRUI is 32. The hyperparameter setting method is the same with research [14].

B. The Imputation Results Under Different Missing Rate

In the dataset, the first 10121 samples are used for GAN training, 1000 samples for verification, and the last 1920 samples for testing. The results on the test set are as follows.

TABLE I. THE IMPUTATION RESULTS UNDER DIFFERENT MISSING RATE(%)

Missing Rate	Mean	KNN	MICE	MF	FC	GRUI	BGRUI
0.05	26.29	9.61	9.17	8.69	8.25	8.11	7.38
0.08	26.29	9.81	9.21	8.79	8.31	8.17	7.48
0.1	26.42	9.99	9.6	8.47	8.3	8.26	7.56
0.15	26.43	10.16	9.66	8.73	8.42	8.27	7.61
0.18	26.45	10.25	9.96	9.22	8.46	8.46	7.81
0.2	26.47	10.28	10.05	8.78	8.6	8.58	7.91
0.25	26.45	10.41	10.18	9.27	8.74	8.67	8.27
0.3	26.45	10.52	10.54	9.18	8.84	8.82	8.48
0.4	26.49	10.75	10.93	10.13	9.03	8.94	8.42

KNN, MICE and MF are the imputation methods described in part A. FC means the GAN method whose discriminator and generator are constructed only by the fully-connected later. GRUI means the GAN which only consider the single direction character of time series. BGRUI means the GAN which takes account of the bidirectional mechanism. According to the case study results, the RMSE of BGRUI is smaller than the other imputation methods. Compare to the simple mean value filled method, the RMSE of the imputed missing value decreased about 18%.

C. The Visualization of Different Imputation Methods

In the case of missing rate of 30%, the original data of the No. 1 wind farm in the dataset and the data imputed by different methods are displayed. The results are as follows.

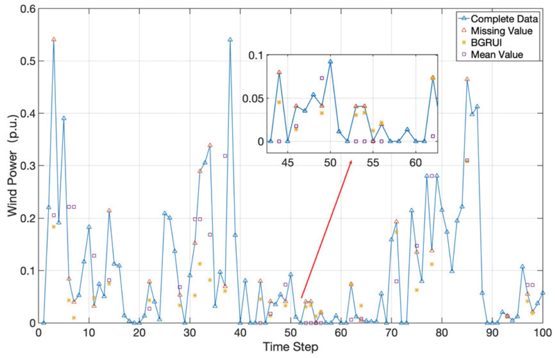

Fig. 4. The Imputation Results Comparison of Mean Value and GAN

The missing value in the dataset is marked by the red rectangle. The imputed results by BGRUI and mean value filled method is also listed in the figure. We can notice that the mean value fitted method cannot describe the fast fluctuation of wind power compare to the BGRUI method. Especially when there are some continual missing values, the mean value fitted method cannot consider the change of wind power in this interval. We also compare the BGRUI with other machine learning methods in figure 6.

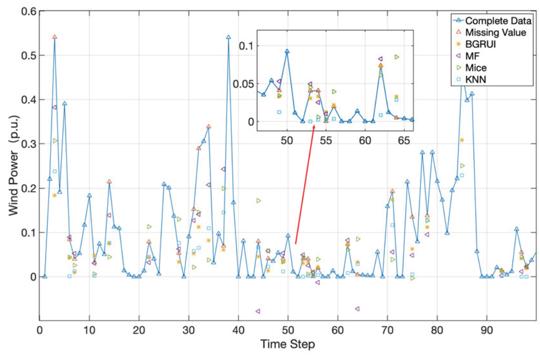

Fig. 5. The Imputation Results Comparison of Machine Learning and GAN

According to the results in figure 6, the imputed results of BGRUI is better than the other machine learning methods in most missing value cases. The ablation study of different units in the GAN is also conducted.

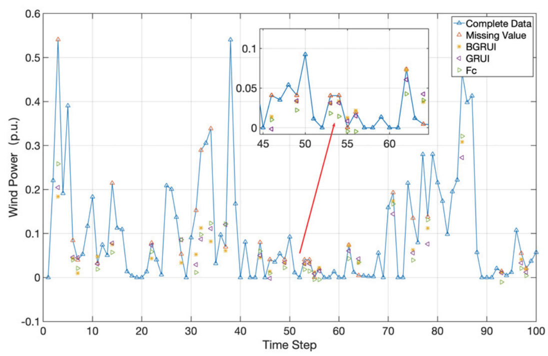

Fig. 6. The Imputation Results Comparison of Different GAN

It can be found that when the bidirectional mechanism is taken into account, the imputed results is closer to the real value in most cases.

D. The Sensitivity Analysis of Hyperparameter

In this part, the sensitivity of the hyper-parameter is analyzed. For the missing value reconstruction problem, the most influential parameter is λ in (11). Therefore, we compare different RMSE under different hyper-parameter when the missing rate is 0.2. The results is as follows.

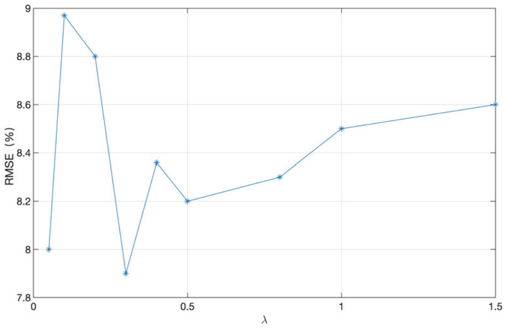

Fig. 7. The Sensitivity Analysis of Hyper-Parameter λ

According to the sensitivity analysis, the discriminative loss can help to reduce the reconstruction loss. In our test case, when the λ is set to 0.3, the smallest reconstruction loss can be reached.

E. The Comparison of Prediction Results

The same dataset are used to train a prediction model. the first 10121 samples are used for GAN training, 1000 samples for verification, and the last 1920 samples for testing. The deep learning model which consists of two layer LSTM are used to predict the forth hour wind power. The learning rate is 0.01. The results are as follows.

TABLE II. The Prediction Errors Under Different Missing Rate(%)

Missing Rate	Mean	MF	GRUI	BGRUI	Complete
0.05	16.82	16.05	14.21	14.89	12.86
0.1	16.93	16.18	15.64	15.26	12.86
0.2	17.08	16.75	15.99	15.69	12.86
0.4	17.63	16.92	16.87	16.21	12.86

In this case, the complete means the prediction results when there is no missing values. According to the results in table 2, the imputation model consists by the BGRUI can lead to more accurate prediction results. It indicates the more superior imputation ability of this method.

V. Conclusion

The imputation of missing data is crucial to enhance the quality of training data in practice. In this paper, through the related experiments, it is verified that the generative adversarial network based on bidirectional GRUI has better performance than other abnormal data imputation methods in the random data missing pattern. In the follow-up research, the results will be applied to the downstream tasks to verify the effectiveness of imputation results on the prediction results. Besides, when the duration of missing data is short, the continual missing pattern will also be discussed in future work.

References

[1] L. Li, Y. Li, B. Zhou, Q. Wu, X. Shen, H. Liu, et al. An adaptive time-resolution method for ultra-short-term wind power prediction[J]. International Journal of Electrical Power & Energy Systems, vol. 118, pp. 1-11, January 2020.

[2] H. Fan, X. Zhang, S. Mei, K. Chen, X. Chen. M2GSNet: Multi-Modal Multi-Task Graph Spatiotemporal Network for Ultra-Short-Term Wind Farm Cluster Power Prediction[J]. Applied Sciences, vol. 10, pp. 1-20, November, 2020.

[3] Y. Zhou, W. Hu, Y. Min, Le Z, B. Liu, R Yu, et al. A semi-supervised anomaly detection method for wind farm power data preprocessing[C] 2017 IEEE Power & Energy Society General Meeting. pp. 1-5, July, 2017.

[4] M. Xu, Z. Lu, Y. Qiao, N. Wang, S. Zhou. Application of change-point analysis to abnormal wind power data detection[C] 2014 IEEE PES General Meeting| Conference & Exposition. pp. 1-5, July, 2014.

[5] L. Zheng, W. Hu, Y. Min. Raw wind data preprocessing: a data-mining approach [J]. IEEE Transactions on Sustainable Energy, vol. 6, pp. 11-19, Jannuary, 2015.

[6] D. Zhang, W. Li, Y. Liu, et al. Reconstruction Method of Active Power Historical Operating Data for Wind Farm [J]. Automation of Electric Power Systems, vol. 38, pp. 14-24, March, 2014.

[7] X. Ye, Z. Lu, Y. Qiao, C. Liu. Identification and correction of outliers in wind farm time series power data[J]. IEEE Transactions on power systems, vol. 31, pp. 4197-4205, November, 2016.

[8] Y. Zhao, L. Ye, W. Wang, H. Sun, Y. Ju and Y. Tang. Data-driven correction approach to refine power curve of wind farm under wind curtailment[J]. IEEE Transactions on Sustainable Energy, vol. 9, pp. 95-105, Jannuary, 2018.

[9] M. Yang, J. Wang, J. Du. The complement of the missing data based on the extreme learning machine and granger test in wind power[J]. Journal of Northeast Electric Power University, vol. 39, pp. 9-16, October, 2019..

[10] Y. Zhao, L. Ye, Z. Li, X. Song, Y. Lang and J. Su. A novel bidirectional mechanism based on time series model for wind power forecasting[J]. Applied energy, vol. 177, pp. 793-803, June, 2016.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair and et al. Generative adversarial networks[J]. Communications of the ACM, vol. 63, pp. 139-144, October, 2020.

[12] Wei X, Gong B, Liu Z, et al. Improving the improved training of wasserstein gans: A consistency term and its dual effect[J]. arXiv preprint arXiv:1803.01541, 2018.

[13] Y. Luo, X. Cai, Y. Zhang, J. Xu and X. Yuan. Multivariate Time Series Imputation with Generative Adversarial Networks [C] Advances in Nerual Information Processing Systems, pp. 1596-1607, December, 2018.

[14] W. Zhang, Y. Luo, Y. Zhang, D. Srinivasan. Solargan: Multivariate solar data imputation using generative adversarial network[J]. IEEE Transactions on Sustainable Energy, vol. 12, pp. 743-746, Jannuary, 2020.

[15] T. Hastie, R. Tibshirani, G. Sherlock, M. Eisen, P. Brown and D. Botstein. Imputing missing data for gene expression arrays[J]. 1999.

[16] S. Buuren, K. Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in R[J]. Journal of statistical software, pp. 1-68, 2010.

[17] Y. Koren, R. Bell, C. Volinsky. Matrix factorization techniques for recommender systems[J]. Computer, vol. 42, pp. 30-37, August, 2009.

[18] Z. Liu, M. Hauskrecht. Learning linear dynamical systems from multivariate time series: A matrix factorization based framework[C] Proceedings of the 2016 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, pp. 810-818, 2016.