J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 3655--3668

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Research on algorithm construction of time series data processing from the perspective of multi-resource fusion of virtual power plantk

Yan Shi[1,✉], Siteng Wang[1], Rui Zhang[1], Luxi Zhang[1], Yi Zhang[1]

[1] *State Grid Mengdong Power Supply Service Supervision Center, Tongliao, Inner Mongolia, 028000, China*

ABSTRACT

With the access of multiple renewable resources to virtual power plants, hundreds of millions of power time series data are generated every day. A sparse learning-based power data compression and reconstruction processing method is designed in the study, which effectively solves the problems of low computational efficiency in the data processing centre of the virtual power plant and the waste of storage resources. According to the vector principal component analysis method, the power data are compressed. Then the data reconstruction network model is constructed based on sparse learning to achieve the reconstruction of power data. The experimental test results show that the median absolute errors of reconstruction of active and reactive power data are 4.05 MW and 0.885 Mvar, respectively, and the percentages of absolute errors are not more than 5%, which makes the reconstruction performance highly stable. The method achieves high-quality power data compression and high-precision reconstruction processing, which is of great significance for improving the computational efficiency of the virtual power plant data centre and accelerating the digital transformation of the power grid.

Keywords: virtual power plant, power time series data, sparse learning, data compression, high-precision reconstruction

1. Introduction

Generally speaking, a virtual power plant is a virtual power plant, which does not have the physical power plant itself, but a management model or a set of systems, through supporting technologies to integrate small solar energy, wind energy and other new energy power generation devices, energy storage batteries and various controllable (regulated) electrical equipment (loads) scattered in different spaces [1-4], coordinate control, and form a controllable power supply equivalent to the external equivalent, assist the operation of the power system, and can participate in power market

✉ Corresponding author.
E-mail address: 13998216391@163.com (Y. Shi).

Received 20 March 2024; Revised 05 June 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-209

transactions, while optimizing resource utilization. Maintain the stability and safety of electricity consumption within and even across regions [5-7]. Virtual power plant can not only plan to consume power system power, but also to the power system reverse power output, more flexible and efficient 'peak shaving and valley filling' and other operations, and to obtain considerable economic returns [8-10], under the background of the new era, multi-virtual power plant multi-resource fusion of virtual power plant has an important role in the development of the virtual power plant.

Temporal data is data that is arranged in chronological order, usually including timestamps and values. Temporal data processing refers to the process of analysing, modelling, predicting and visualising these data. In today's big data era, temporal data processing has become an indispensable part of many industries, such as finance, logistics, and healthcare [11-14]. In the power system, time series data usually include multi-dimensional variables such as power load, voltage, current, power, and so on. The commonly used algorithms for time series data processing include ARIMA model, SARIMA model, Holt-Winters model, support vector regression model, etc [15-17]. These algorithms can help us to fit historical data and predict future trends, thus assisting decision-making and planning [18-19].

This paper proposes a series of innovative strategies for power regulation in virtual power plant data centres, including the use of deep learning models for power demand prediction, the introduction of temporal embedded word coding to improve prediction accuracy, and the construction of a new virtual power plant data processing centre system architecture. A data compression and reconstruction scheme for power time series data is proposed, which eliminates redundancy by calculating the similarity between data, and uses vector principal component analysis to compress power data. At the same time, a sparse learning network model consisting of n layers is constructed, and the reconstruction model for power time series data is solved by the ADMM algorithm to achieve the initialisation of the original data, and the model is trained by using the inverse gradient solving method to realise the reconstruction of power data. The 2023 operation data in a virtual power plant in a certain place is selected as the experimental dataset, and the experimental dataset is compressed and reconstructed by this method to test the efficiency of the data processing algorithm in this paper.

2. Construction of the virtual power plant data-processing centre system

As one of the electricity consumers on the load side of the virtual power plant, the data processing centre becomes a regulable resource on the load side, providing a means of regulation for the virtual power plant [20]. In the face of power fluctuations, in order to maintain the continuity and stability of service, the data processing centre needs to have the ability of early warning and lossless regulation. The traditional virtual power plant data processing centre system architecture is shown in Figure 1, which contains four parts: power consumption, power generation, power detection and container management. In the traditional data processing centre architecture, power supply and management is a crucial part.

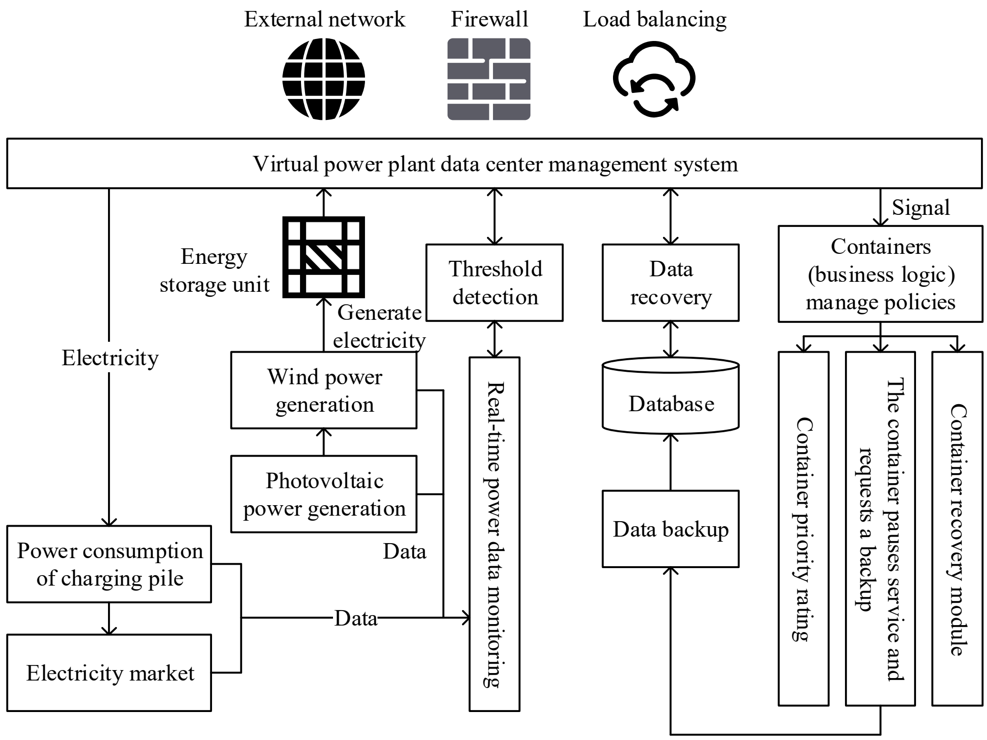


Fig. 1. Traditional virtual power plant data center system architecture

Conventional data processing centres are usually required to be equipped with an energy storage unit large enough to support their high loads. These energy storage units typically consist of a series of high-cost batteries or other energy storage devices designed to provide temporary power during power shortages or other emergencies. However, there are several obvious disadvantages to this setup:
1) High cost. A large energy storage unit is required to meet the buffer power needs of the data processing centre, and its high cost, coupled with the decay of the battery storage capacity during use, further increases the operating cost of the data processing centre.
2) Limited response time. Traditional processing data centres often rely on preset rules and thresholds for power shortage alerts and responses. This approach often only triggers alerts after a problem is about to occur or has already occurred, meaning that the data processing centre's containers only have a very limited amount of time to respond.
3) Passive Response Mode. Since alerts are usually triggered after a problem has occurred, this model is extremely reactive. This not only limits the flexibility of the data processing centre to respond to emergencies, but may also increase operational problems and costs.

The above problems need to be addressed by more advanced and flexible solutions to improve the efficiency and reliability of data processing centres. Therefore, the new virtual power plant data processing centre system architecture proposed in this paper is shown in Fig. 2, which contains four parts: power consumption, power generation, early warning, and container management [21]. The architecture employs a deep learning model for power demand prediction to improve the accuracy of early warning. A time window of 15 min provides enough time for the containers in the data processing centre to suspend running services and perform data backup, which greatly reduces the risk of data loss and service interruption caused by power shortage. In addition, the system architecture has good dynamic response capability and adaptability, and is able to automatically adjust resource allocation and backup strategy according to the prediction results.

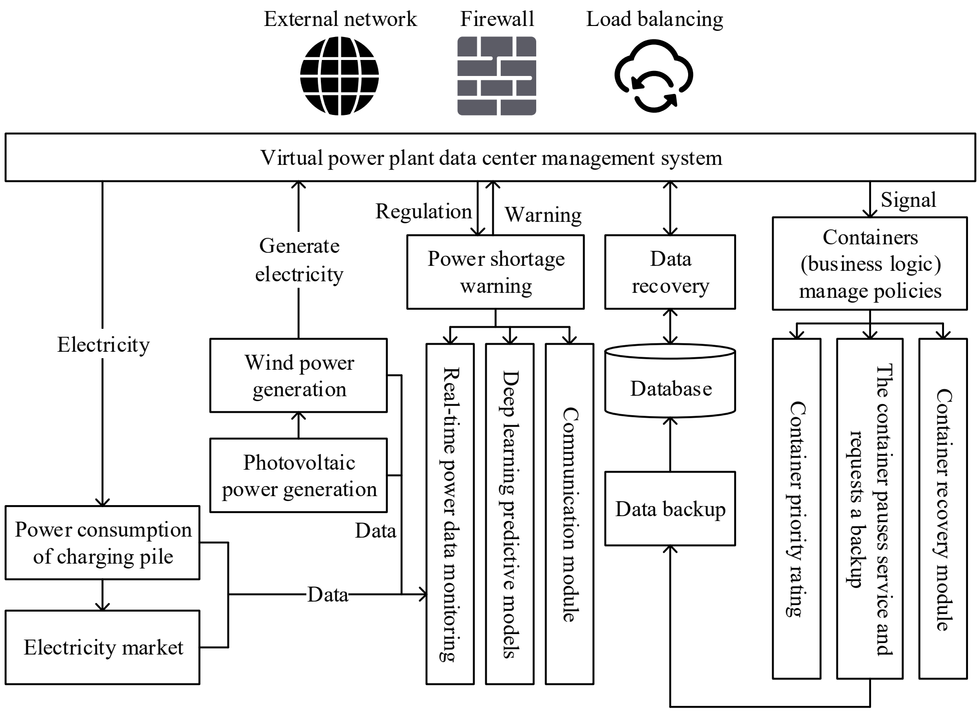


Fig. 2. Architecture of the new virtual power plant data processing center system

Overall, by introducing a data-driven intelligent forecasting and dynamic response mechanism, the system architecture not only improves the integration and management efficiency of distributed energy resources and the risk-resistant capability of the virtual power plant, but also fully prepares for the changing demands and challenges in the future, and is able to ensure a sustainable power supply.

3. Time-series data compression and reconstruction processing

Data acquisition devices in the power system collect a huge amount of power data, of which power timing data is a typical power data. The analysis and prediction of power timing data are crucial for the safe and stable operation of power grids. Power timing data is an important basis for condition monitoring and fault diagnosis of power grid equipment, which requires high real-time performance. Therefore, the compression and reconstruction processing of power time series data is a prerequisite for power time series data quality management.

3.1. Timing data structure characteristics

Temporal data is a sequence of observed values at fixed time intervals. In general, data L generated by a particular object or objects (e.g., a smart meter) at n point in time can be defined as:

$$L = \{l_0, l_1, \cdots, l_{n-1}\}, l_i = (v_i, t_i) \tag{1}$$

Where v_i is the value at point t_i. When calculating the maximum or minimum value or the variance or mean of a time series for a particular time period, the features can be used to quickly find and calculate the value, reducing the time overhead. Typical datasets are two-dimensional tables with time and attributes as rows and columns. Time series data generally contain the following major features:

1) Temporality. Temporality is the most basic feature of temporal data and has a sequential order and this order is time. The characteristics possessed by the data generated by the same data node have similarity, and at a certain point in time, other attribute values have uniqueness.

2) Large data volume. Time series data is usually a process of data accumulation, and new data are

generated every moment. Therefore, data growth is another reason for data redundancy.

3) Diversity. The data stored in the storage system come from various industries and are transmitted in various ways. For example, the current and voltage of a household user's electricity consumption, a current data ranges from 0 to 63A, while the voltage fluctuates up and down around 220V, and other attribute values are also distributed in different ranges, which is, to some extent, the diversity of data. On the other hand, data diversity can also be interpreted as data generated by different types of equipment.

4) Correlation. There may be a certain correlation between the attributes of time series data, for example, in the case of an electric meter where the voltage value is generally a fixed value, the power is proportional to the current value. From a mathematical point of view, the Pearson correlation coefficient is commonly used as a judgement criterion, and the correlation coefficient is the covariance under the standard scale.

5) Randomness. Whether it is time series data or other types of data, in the generation of the time may be affected by temperature, humidity, wind and other force majeure factors, which will form data noise. A common white noise is a series of independently distributed normal sequences, which are characterised by irrelevance and randomness.

3.2. Power data pre-processing

This paper takes the new virtual power plant data processing centre system constructed in Chapter 2 as an example of electric power time series data, which is generally floating point type data. Before encoding it, the input data is first normalised to integers, and the formula for normalisation is $f' = f' \times 10^d$. Where f is the original input data, f' is an integer, and d indicates the maximum number of decimal places in the data. Assuming that the input power data is the data shown in Fig. 3, the detailed steps of power data preprocessing are as follows:

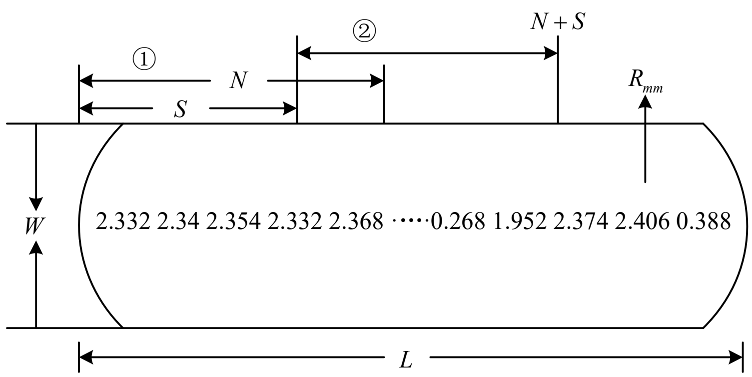


Fig. 3. Schematic diagram of power data processing

1) After applying normalisation, the output power data is represented as $[98, 96, 12, 46, 98, 40]^T$.

2) The difference between consecutive values is computed for a particular column of data representing an attribute value d_i, where $d_0 = n_0$, $d_{t>0} = n_i - n_{i-1}$, is computed for adjacent consecutive data.

3) Encoding the difference as a zero-order signed exponential Columbian code word with separators at every number of code words for variable length encoding.

4) Convert the non-negative integer 1 to binary by removing the lowest bit k and adding M.

5) Calculate the number of bits remaining and subtract that number by 1, which is the number of zeros to be added to the prefix.

6) Add the lowest k bits removed in step (1) back to the end of the bit string.

7) Using zero-order exponential Columbian coding all the code words are combined together and formed into a single string. The data types in the process are divided into signed and unsigned, where the unsigned code elements are structured as $[M, 0, \text{pre}]*[offset]$, $[M, 0]$ represents M zeros, * can

be viewed as a separator, and $[offset]$ represents the offset within the group, i.e., the desired coded non-negative integer, calculated as:

$$\begin{cases} M = [\log_2(CodeNum+1)] \\ offset = CodeNum + 1 - 2^M \end{cases} \tag{2}$$

However, for the difference between the power data used in this paper may be negative, so it is necessary to convert the signed form to the unsigned form, the conversion process is shown below:

$$S = (-1)^{k+1} Ceil(k/2) \tag{3}$$

Where S is a signed number; the $Ceil()$ function represents upward rounding. According to the formula calculated after the unsigned number, and then use the unsigned to encode.

8) Will be encoded into a binary string for segmentation, the subsequent compression process, each segment of the even bit encoding compression is a range of 0 ~ 1 between the number. The number is the compression accuracy of each segment of binary data, thus ensuring the accuracy of power data recovery and the advantages of compression time.

According to the above steps, the preprocessing of virtual power plant power noise data is completed.

3.3. Power Data Compression Processing

Vector Principal Component Analysis (VPCA) is used to compress the power data of the virtual power plant, and the specific processing steps are as follows:

1) It is assumed that the timing data of the virtual power plant consists of L_U set of dynamic synchronous voltage phase data and L_I sets of dynamic synchronous current phase data.

2) Taking L_U as an example, t represents the acquisition time of the dynamic synchronous phase data and h_s represents its acquisition frequency. Then the number of dynamic synchronous phase data acquisition is:

$$Q = t \times h_s \tag{4}$$

Denote the i st voltage phase quantity by $W_i \angle \beta_i$, where W_i refers to the amplitude sequence corresponding to the voltage phase quantity. β_i refers to the sequence of phase angles corresponding to the voltage phase quantity. The expressions for W_i and β_i are as follows:

$$\begin{cases} W_i = \left[W_{(1,i)}, W_{(2,i)}, \cdots, W_{(p,i)}, \cdots, W_{(Q,i)} \right] \\ \beta_i = \left[\beta_{(1,i)}, \beta_{(2,i)}, \cdots, \beta_{(p,i)}, \cdots, \beta_{(Q,i)} \right]^t \\ 1 \le p \le Q \\ 1 \le i \le L_U \end{cases} \tag{5}$$

where p is a positive integer, $\beta_{(p,i)}$ refers to the phase angle of the i rd voltage phase at moment p, and $W_{(p,i)}$ is the amplitude of the i th voltage phase at moment p.

3) Process L_I through the same steps.

4) Construct the original phase data matrix of L_U, denoted by Y, which is a complex matrix of $Q \times L_U$, with the following expression:

$$\begin{cases} Y = \left[W_1 \angle \beta_1, W_2 \angle \beta_2, \cdots, W_i \angle \beta_i, \cdots, W_{L_v} \angle \beta_{L_v} \right] \\ Y \in B_{Q \times L_v} \end{cases} \tag{6}$$

where $B_{Q\times L_v}$ is the complex matrix of $Q\times L_U$.

Normalise matrix Y to obtain matrix S, which is the complex matrix of $Q\times L_U$. Compute the covariance matrix L of matrix S, which is the complex matrix of $L_U\times L_U$.

Compute all the eigenvalues and the corresponding vector matrices for L. Where the eigenvalues are denoted as:

$$\delta=\{\delta_1,\delta_2,\cdots,\delta_{L_v}\} \tag{7}$$

where δ_{L_v} refers to the L_U nd eigenvalue.

The set of vector matrices corresponding to δ is denoted as:

$$V=\left[v_1,v_2,\cdots,v_{L_v}\right] \tag{8}$$

The q principal components are selected on the basis of the cumulative contribution ratio. In the set of vector matrices V, the first q eigenmatrices are selected to form a new eigenvector matrix, denoted by V_q, and V_q is the complex matrix of $L_U\times q$.

Calculate the principal component matrix F:

$$\begin{cases}F=SV_q\\F'\in Q\times q\end{cases} \tag{9}$$

where F' is the complex matrix of $Q\times q$.

Only the principal component matrix F and the eigenvector matrix V_q of L_U are stored to complete the compression process of the virtual power plant power time series data.

3.4. Power Data Reconstruction Processing

A data reconstruction network model is constructed based on sparse learning to achieve the reconstruction of power data [22]. The sparse learning network model constructed consists of n layer, and each layer consists of three sub-layers respectively. The first sub-layer is the reconstruction layer, which is able to achieve the reconstruction of one-dimensional components:

$$Z^{(u)}=\left[\varpi^o\varpi+\sum_{g=1}^{J}\theta_g^{(u)}\right]^{-1}\left[\varpi^oM+\sum_{g=1}^{J}\theta_g^{(u)}\left(\kappa_g^{(u-1)}-\mu_g^{(u-1)}\right)\right] \tag{10}$$

Where ϖ refers to the reconstruction factor, O refers to the reconstruction threshold, J is a positive integer, $\theta_g^{(u)}$ refers to the expansion function, and M refers to the number of reconstructions.

By initialising $\kappa_g^{(u-1)}$ and $\mu_g^{(u-1)}$ directly to 0, the above equation can be written as:

$$z^{(1)}=\left[\varpi^o\varpi+\sum_{g=1}^{J}\theta_g^1\right]^1\left(\varpi^oM\right) \tag{11}$$

The second sublayer is the nonlinear transformation layer, in which more general functions are learnt by means of linear functions. Given two inputs $\mu_g^{(u-1)}$ and $r_g^{(u)}$, the output of this layer can be defined as the following equation:

$$b_g^{(u)}=F_{PLF}(\mu_g^{(u-1)}+r_g^{(u)};\{\eta_j,\rho_{g,j}^{(u)}\}_{j=1}^{C_v}) \tag{12}$$

where F_{PLF} refers to a segmented linear function and $\{\eta_j,\rho_{g,j}^{(u)}\}_{i=1}^{C_V}$ refers to a series of control points.

The third sublayer is the multiplicative update layer in which, given three inputs $\mu_g^{(u-1)},r_g^{(u)},\kappa_g^{(u-1)}$, the output of the layer can be defined as:

$$\mu_g^{(u)}=\mu_g^{(u-1)}+\vartheta_g^{(u)}\times r_g^{(u)}\times\kappa_g^{(u-1)} \tag{13}$$

where $\vartheta_g^{(u)}$ refers to the parameter to be learnt.

The initialisation process of parameter Ψ is achieved by solving the reconstruction model of power timing data by ADMM algorithm [23]:

$$\omega = \arg\min_{\tau}\left\{\frac{1}{2}\|\ \Im\aleph - \wp\|_2^2 + \gamma\|\ \ \tau\|_1\right\} \tag{14}$$

Where, τ refers to the soft threshold, $\Im$ refers to the nonlinear threshold, $\aleph$ refers to the transformation threshold, $\wp$ refers to the distribution threshold, and γ refers to the loss function. Finally, the model is trained using inverse gradient solving, and the reconstruction of power data is completed using the trained model.

4. Analysis of examples

For the designed sparse learning-based virtual power plant power data compression and high-precision reconstruction processing method, its power data compression performance and reconstruction performance are tested through experiments. The operation data of 2023 in a virtual power plant is selected as the experimental dataset, and the experimental dataset is compressed and reconstructed by the designed method.

4.1. Processing of experimental data sets

4.1.1. Missing value processing:

1) Missing value detection. Due to the existence of a variety of circumstances may lead to missing power timing data, such as active power power timing data, there is a value of 0 data, and the existence of missing points makes it difficult to use the data. Therefore, in order to ensure the normal use of the data, it is necessary to firstly count these missing values, and then fill in the missing points on this basis. In this paper, we have implemented the ability to upload power time series data and visualise the data, missing value statistics and missing value details, in order to quickly and intuitively find out the location of the missing points in the time series data.

Figure 4 shows the visualisation of the raw power timing data after uploading the data on a certain day in 2023, where the sampling interval is 5 s. It can be found that there are 9 consecutive missing values (active power = 0 Kw) in the raw power timing data on that day.

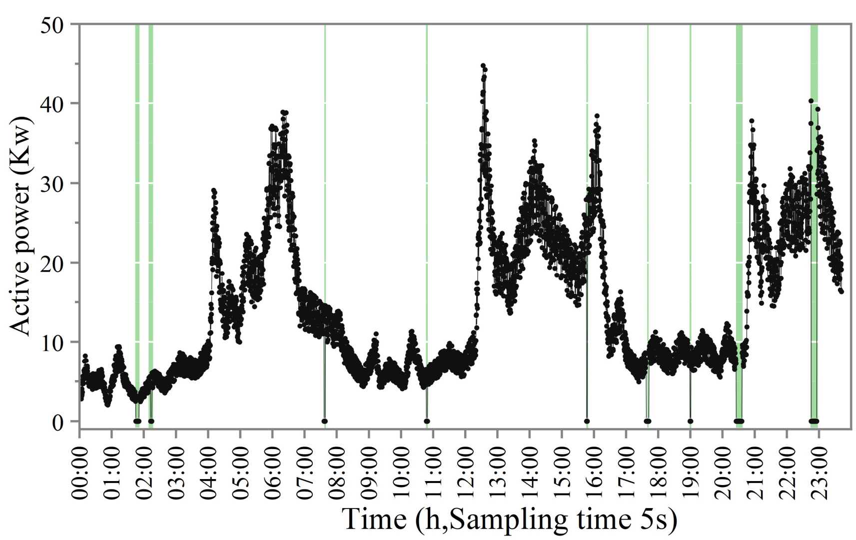


Fig. 4. Visualization of raw power timing data

2) Missing value filling. The visual display of the original power time series data can be more intuitive

to determine the degree of missing original data, if the number of missing points meets the data filling conditions, the missing value filling method can be used to fill the missing points, so that the data is more complete, providing more effective data for the back-order task. In this paper, the missing values are filled by the power time series data reconstruction processing method proposed in the previous section, and the missing value filling method can be selected to fill the data according to the actual needs. The result of power timing data filling is shown in Figure 5, where the intersection of the red vertical line and the folded line indicates the filling point (that is, the missing point before the original power timing data is not processed).9 places) After the missing values are filled in the processing, the subsequent analyses can be carried out.

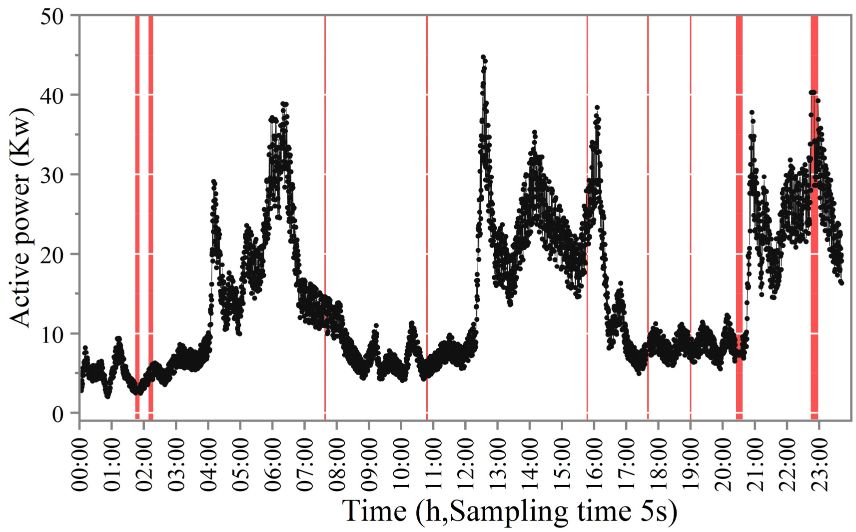


Fig. 5. Power timing data filling line diagram

4.1.2. Raw data analysis:

1) Seasonal component analysis. The seasonal characteristics of the power time series data are very obvious, for example, in the case of the data on the proportion of induction motors in the virtual power plant in the region, in summer, because of the high temperature, the air conditioning usage rate increases steeply leading to an increase in the residential electricity load. However, it may also be due to the danger of high-temperature operations, industrial production is reduced leading to a decrease in industrial electricity load. And both residential and industrial electricity loads increase in autumn and winter. In this paper, the histogram of seasonal component analysis of electricity time series data is implemented as shown in Fig. 6, which is able to show the frequency of occurrence of electricity time series data in each value interval in each season. When pointing to a specific numerical interval, the seasonal histogram shows the number of times each season appears in that numerical interval, and it can be clearly found that the proportion of electric induction machines for that year appeared the highest number of times in summer 320 times within the interval of 62.28 to 62.43%. Meanwhile, the winter, autumn, and spring were 271, 168, and 63 times in that interval, in that order.

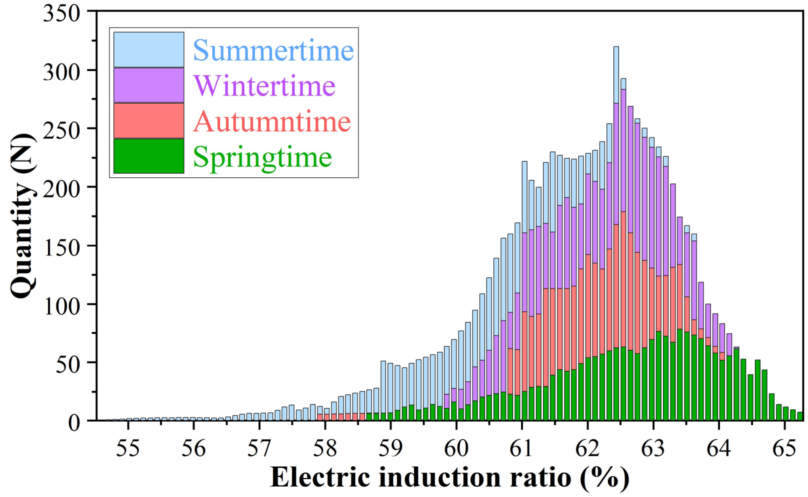


Fig. 6. Power time-series data seasonal component analysis histogram

Looking at the overall picture, the distribution of the proportion of induction motors in the spring, autumn and winter seasons ranges from about 58 per cent to a maximum of 65.3 per cent. The highest induction motor proportion is particularly noticeable in summer, indicating the highest peak stage of electricity load in the region in summer.

2) Intraday Monthly Analysis. Box-and-line diagrams are used to carry out intra-day and inter-month analysis of power time series data. Box-and-line diagrams are mainly used to reflect the characteristics of the distribution of the original data, and they can also be used to compare the characteristics of the distribution of multiple groups of data. In this paper, inter-monthly and intra-day boxplot visualisation and analysis of time-series electricity data is implemented. Figure 7 shows the inter-monthly boxplot of the induction motor ratio data for the 12 months of 2023 for the virtual power plant in the region, and it can be seen that there is no significant sudden change or inflection point in the change of the ratio of the induction motors during this period. During the month of August, the high frequency and duration of refrigeration equipment use due to hot weather causes the induction motor ratio to increase. Clicking on a particular box plot module displays the upper limit, upper quartile, median, lower quartile, and lower limit values of the induction motor ratio for that month, and the upper limit value of the electric induction machine ratio for the month of August is 64.71 per cent, the median is 62.99 per cent, and the lower limit value is 60.04 per cent.

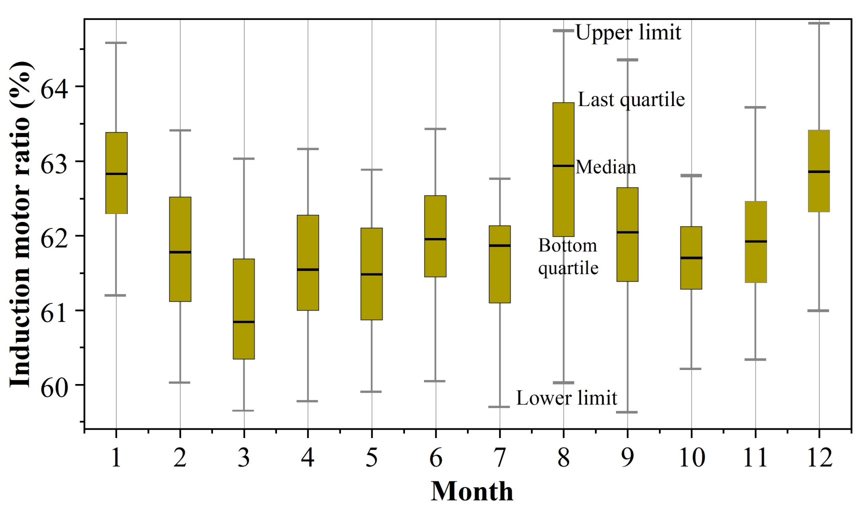
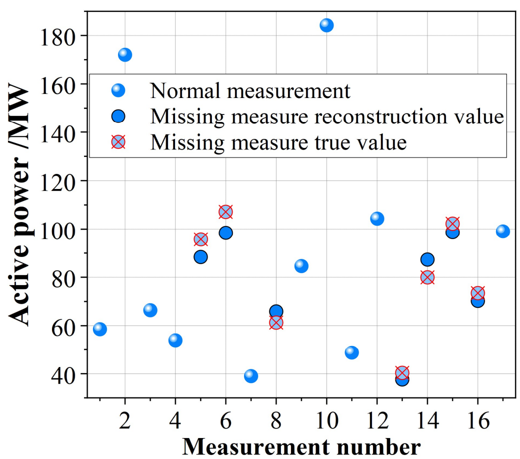


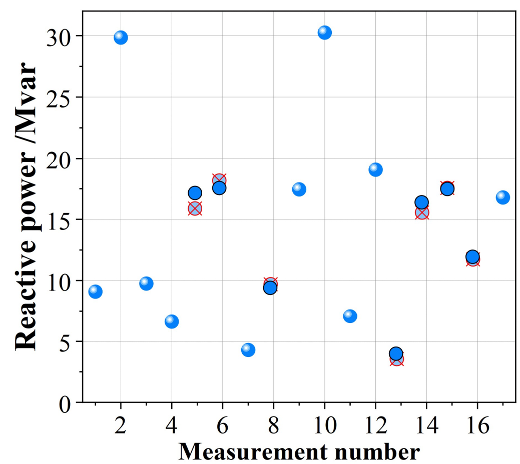
Fig. 7. Monthly boxplot of sequential power data

4.2. Power data reconstruction effect

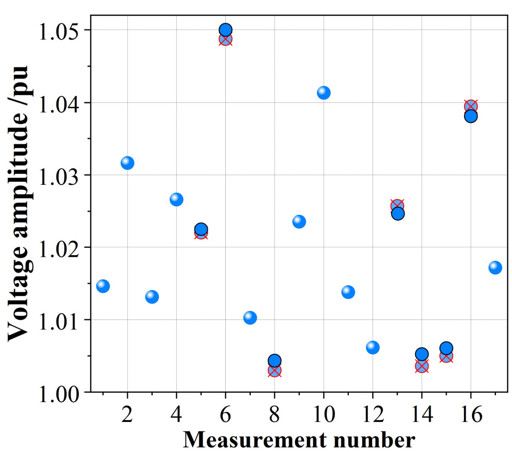
On the 2023 operation data in the virtual power plant as the experimental dataset, eight values of its binary complex matrix 1 are randomly set to 0 and the rest to $M_s(17\times1)$, indicating that the corresponding eight sets of measurement values in the system are randomly missing. The measured data containing missing values and M_s are used as inputs to the model, and taking one of the sets of data as an example, the reconstruction effect of sparse learning on the missing data of electric power is shown in Fig. 8, and (a) to (d) are the reconstruction effects of the data of active, reactive, voltage, and phase angle, respectively. It can be seen that the measurements numbered 5, 6, 8, 13, 14, 15, and 16 in the system are all lost due to communication failures and other problems, and the remaining 10 groups of measurements can still work normally. Based on the contextual relationship and the historical operation law, the reconstruction data given by sparse learning is shown as the blue dots in the figure. It can be seen that even in the case of nearly half of the measurement failures in the system, sparse learning can still give reconstruction results that are very close to the real value by combining the contextual relationship of the other measurements and the historical data change pattern.



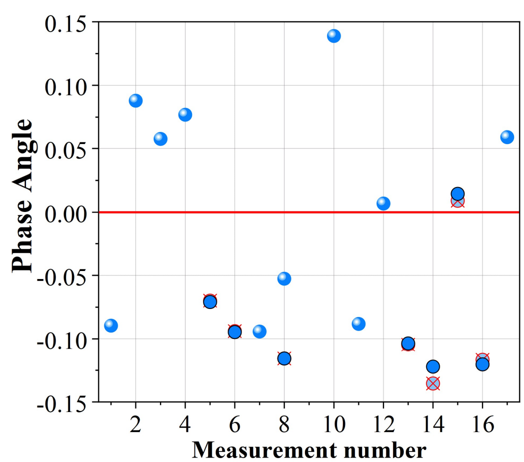
(a) Reconstruction of active power



(b) Rebuild reactive power



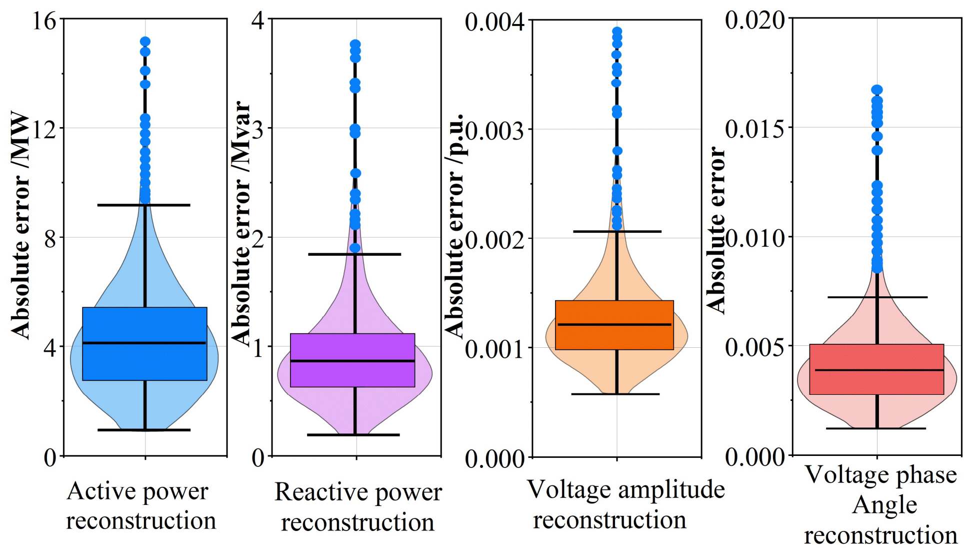
(c) Reconstruct voltage amplitude



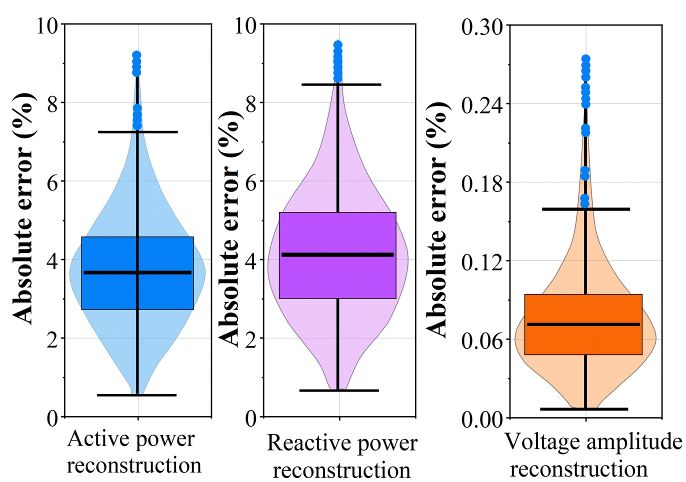
(d) Reconstructed voltage phase Angle

Fig. 8. Sparse learning model reconstructs missing data effects

Similarly, the other 1348 sets of data in the test set are randomly simulated to measure faults and reconstructed, and the distribution of absolute error (AE) and absolute percentage error (MAE) of the reconstructed active, reactive, voltage, and phase angle data in the test set are shown in Fig. 9. Among them, Fig. 9 (a) and (b) show the distribution of absolute error (AE), and absolute error percentage (MAE), respectively; the box-and-line plot shows the median, upper and lower quartiles, and outliers of the reconstruction error, and the violin plot shows the data distribution. The phase angle values do not have a quantitative size relationship, so the MAE metric is not applicable. It can be seen that even in the case of nearly half of the quantitative measurements of the system appearing randomly missing, the reconstruction results of 1348 groups of data still maintain a high reconstruction accuracy, and the distribution of the error is more concentrated. Among them, the median absolute error of active power reconstruction is 4.05 MW, and the median absolute error percentage is 3.67%. The median absolute error of reactive power reconstruction is 0.885Mvar, and the median percentage of absolute error is 4.13%. From the distribution of the error data, the reconstruction results are concentrated near the median, and the number of samples with large prediction errors is very small, and the outliers of the four measurements are less than 20, and the maximal reconstruction errors of active and reactive power are still not more than 5%, so the reconstruction performance is very stable.



(a) Data reconstruction absolute error distribution



(b) Data reconstruction absolute percentage error distribution

Fig. 9. Reconstruction error analysis

5. Conclusion

Power time series data, as the massive data generated during the operation of virtual power plants, is mainly applied to fault diagnosis, load forecasting and renewable energy prediction. However, at the same time, power time series data is also characterised by high redundancy, which needs to be compressed and reconstructed. In this paper, a sparse learning-based method is designed to achieve high-quality power timing data compression and high-precision reconstruction. The experimental results show that the designed method has the highest data compression ratio, and none of the integrated vector errors exceeds 5%, indicating that the method has certain applicability.

The paper verifies the effectiveness of sparse learning in data compression and reconstruction, but there are deficiencies in terms of time considerations. Further consideration of more features of time-series data to enrich the diversity of evaluation methods is a need for improvement in future research.

References

[1] Zamani, A. G., Zakariazadeh, A., & Jadid, S. (2016). Day-ahead resource scheduling of a renewable energy based virtual power plant. Applied Energy, 169, 324-340.

[2] Nafkha-Tayari, W., Ben Elghali, S., Heydarian-Forushani, E., & Benbouzid, M. (2022). Virtual power plants optimization issue: A comprehensive review on methods, solutions, and prospects. Energies, 15(10), 3607.

[3] Popławski, T., Dudzik, S., Szeląg, P., & Baran, J. (2021). A case study of a virtual power plant (VPP) as a data acquisition tool for PV energy forecasting. Energies, 14(19), 6200.

[4] Dulău, L. I., Abrudean, M., & Bică, D. (2014, September). Distributed generation and virtual power plants. In 2014 49th International Universities Power Engineering Conference (UPEC) (pp. 1-5). IEEE.

[5] Elgamal, A. H., Kocher-Oberlehner, G., Robu, V., & Andoni, M. (2019). Optimization of a multiple-scale renewable energy-based virtual power plant in the UK. Applied Energy, 256, 113973.

[6] Yang, X., & Zhang, Y. (2021). A comprehensive review on electric vehicles integrated in virtual power plants. Sustainable Energy Technologies and Assessments, 48, 101678.

[7] Zajc, M., Kolenc, M., & Suljanović, N. (2019). Virtual power plant communication system architecture. In Smart power distribution systems (pp. 231-250). Academic Press.

[8] Ropuszyńska-Surma, E., & Węglarz, M. (2019). The virtual power plant–A review of business models. In E3S web of conferences (Vol. 108, p. 01006). EDP Sciences.

[9] Wang, S., & Wu, W. (2021). Aggregate flexibility of virtual power plants with temporal coupling constraints. IEEE Transactions on Smart Grid, 12(6), 5043-5051.

[10] Tan, Z., Zhong, H., Wang, X., & Tang, H. (2020). An efficient method for estimating capability curve of virtual power plant. CSEE Journal of Power and Energy Systems, 8(3), 780-788.

[11] Pal, A., & Prakash, P. K. S. (2017). Practical time series analysis: master time series data processing, visualization, and modeling using python. Packt Publishing Ltd.

[12] Yao, S., Hu, S., Zhao, Y., Zhang, A., & Abdelzaher, T. (2017, April). Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In Proceedings of the 26th international conference on world wide web (pp. 351-360).

[13] Wilson, S. J. (2017). Data representation for time series data mining: time domain approaches. Wiley Interdisciplinary Reviews: Computational Statistics, 9(1), e1392.

[14] Sakurai, Y., Matsubara, Y., & Faloutsos, C. (2015, May). Mining and forecasting of big time-series data. In

Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (pp. 919-922).

[15] Susto, G. A., Cenedese, A., & Terzi, M. (2018). Time-series classification methods: Review and applications to power systems data. Big data application in power systems, 179-220.

[16] Koivisto, M., Das, K., Guo, F., Sørensen, P., Nuño, E., Cutululis, N., & Maule, P. (2019). Using time series simulation tools for assessing the effects of variable renewable energy generation on power and energy systems. Wiley Interdisciplinary Reviews: Energy and Environment, 8(3), e329.

[17] Amini, M. H., Kargarian, A., & Karabasoglu, O. (2016). ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation. Electric Power Systems Research, 140, 378-390.

[18] Ye, X., Lu, Z., Qiao, Y., Min, Y., & O'Malley, M. (2016). Identification and correction of outliers in wind farm time series power data. IEEE Transactions on power systems, 31(6), 4197-4205.

[19] Li, H., Yeo, J. H., Bornsheuer, A. L., & Overbye, T. J. (2020). The creation and validation of load time series for synthetic electric power systems. IEEE Transactions on Power Systems, 36(2), 961-969.

[20] Michael Neethu Elizabeth,Hasan Shazia,Biswal Birendra & Mishra Manohar. (2022). Reactive Power Compensation Using Electric Vehicle and Data Center by Integrating Virtual Power Plant. Electric Power Components and Systems(4-5),245-255.

[21] Popławski Tomasz,Dudzik Sebastian,Szeląg Piotr & Baran Janusz. (2021). A Case Study of a Virtual Power Plant (VPP) as a Data Acquisition Tool for PV Energy Forecasting. Energies(19),6200-6200.

[22] Zhang Jinsong,Liu Jian & Huang Zhenwei. (2023). Improved deep learning method for accurate flow field reconstruction from sparse data. Ocean Engineering.

[23] Keke Huang,Wenfeng Deng,Yichi Zhang & Hongqiu Zhu. (2020). Sparse Bayesian learning for network structure reconstruction based on evolutionary game data. Physica A: Statistical Mechanics and its Applications(C),123605-123605.