Research on non-intrusive monitoring of large power data industrial users

Wenfeng Li¹, Hongkun Bai¹, Bo Wu¹, Feifei Bu¹, Yuanpeng Hua¹, Hao Lu² and Yongchun Yang³

1 State Grid Henan Electric Power Company Institute of Economics and Technology, Zhengzhou, 450000, China
2 State Grid Jiaozuo Power Supply Company, Jiaozuo, 454150, China
3 State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Baoding, 071003, China
E-mail: yongchunyang@126.com

Abstract. This paper constructs a non-intrusive monitoring method based on wavelet packet decomposition and fast Fourier decomposition technology to monitor the key production processes of high-energy-consuming and high-pollution industrial enterprises by utilizing the large data of power consumption of key industrial users such as iron and steel. Taking a steel plant in Henan as an example, using the constructed non-intrusive monitoring method, the load curve of the main production process is decomposed to achieve effective monitoring of the main processes of iron making, steel making and steel rolling. At the same time, through the typical link power capacity measurement, the company's steel production capacity is obtained, which can provide technical support for the supervision of key industrial users such as steel.

1. Introduction
With China's economic development entering a new stage, the development environment of various industries has changed, especially in the traditional industries such as iron and steel, non-ferrous metals, such as "capacity removal", "environmental protection management", "energy conservation and emission reduction" and so on. However, at present, there is a lack of effective means to monitor the capacity and output of iron and steel, non-ferrous and other enterprises, and the policy of overcapacity is facing difficulties in supervision. Henan Province is rich in mineral resources, widely distributed and has obvious regional advantages [1]. However, in the face of severe environment such as continuous capacity removal of key industries, increasing efforts to prevent and control air pollution, it is necessary to innovate monitoring methods to accurately grasp the implementation effect of capacity removal, energy saving and emission reduction of key enterprises. Therefore, this paper studies the non-intrusive monitoring method for industrial users based on large power data, which provides key technical support for the effective monitoring of the production situation of key enterprises such as iron and steel and electrolytic aluminium.

At present, the widely used load monitoring system is divided into intrusive monitoring and non-intrusive monitoring. Traditional intrusive monitoring is to install sensors at each load to monitor the operation of each location. Non-intrusive load monitoring (NILM) was originally proposed by Hart in the 1980s, and its core is the efficient decomposition of loads [2]. Non-intrusive load monitoring is to install monitoring equipment at the power entrance. By monitoring the voltage, current and other signals at the entrance, the type and operation of a single load in the overall load can be obtained by using an effective decomposition algorithm.
The American Academy of Electric Power Sciences has developed a non-intrusive load monitoring system. This method is based on the steady-state power change value, which can better detect the input/exit of a single power equipment, but it will fail or misjudge the simultaneous input/exit of multiple power equipment [3-5]. In reference [6], a non-intrusive power load monitoring method based on transient power information is proposed by using abrupt signal detection method. It is proposed that the significant change part of transient power curve can be used as the identification mark of power equipment, and the type of power equipment can be identified to a certain extent. After that, non-intrusive load monitoring has been applied in various fields. Considering voltage disturbance, fuzzy logic pattern recognition and neural network, non-intrusive load monitoring has been further improved [7-9].

This paper studies the non-intrusive monitoring method for industrial users based on large power data, which provides key technical support for the effective monitoring of the production situation of key enterprises such as iron and steel, electrolytic aluminium and so on.

2. Non-intrusive monitoring method for industrial users with large power data

The load curve of industrial users is a non-stationary signal with long power consumption time and heavy load. The signal contains a variety of feature information, and the overall feature is more complex than that of ordinary residential users. However, industrial production generally has typical links, such as iron and steel enterprises, mainly including ironmaking, steelmaking and rolling processes, each process has different load characteristics and cycle characteristics. Based on the non-stationary characteristics of industrial user load and the periodic characteristics of each production process, this paper uses the combination of wavelet packet decomposition and fast Fourier transform to decompose the signal and transform the spectrum of industrial user load information.

2.1. Data preprocessing method

Before using the data for spectrum analysis, the original data collected need to be preprocessed. Data preprocessing mainly includes data cleaning, data integration and data transformation.

2.1.1. Missing value completion. In this study, missing value completion is divided into two steps: calculating the missing data rate in the window by sliding window, deciding whether to complete the missing data according to the missing rate; and completing missing data by mapping the neighboring data sets.

In the process of calculating the missing rate, it is necessary to ensure that the time series at both ends of the sliding window have corresponding measurement values. As shown in Figure 1, the length of sliding window is 13 time series, only window (4) satisfies the calculation conditions, and the data missing rate is 6/13=47%.

![Figure 1. Sketch of sliding window.](image)

2.1.2. Abnormal value processing. The abnormal value of the original data of industrial user load mainly refers to that the value at a certain time point is much larger than the maximum value in the data set nearby.
In this study, outlier processing is divided into two steps: outlier recognition in the selected data set of sliding window using outlier algorithm; outlier modification into the average value of the corresponding data set of sliding window.

The specific flow of outlier algorithm is as follows:

1. Determine the outlier distance
   \[ S = \frac{\text{Max}_t - \text{Min}_t}{a}, a = 2, 3, \ldots, 10 \]
   where \( a \) is the outlier constraint value, \( \text{Max}_t \) is the maximum value in the current sliding window, \( \text{Min}_t \) is the minimum value in the current sliding window, and \( S \) is the outlier distance value.

2. Calculate the outlier coefficient of each point
   \[ s_i = n - \sum_{j=1}^{n} \left\lfloor \frac{m_i - m_j}{S} \right\rfloor, j = 1, 2, \ldots, n \]
   where \( n \) is the number of values contained in the current sliding window, \( j \) is the traversal of each value, \( m_i \) and \( m_j \) represent the values corresponding to the \( i \)-th point and the \( j \)-th point, respectively, and \( s_i \) is the outlier coefficient of the \( i \)-th point in the current sliding window.

3. Find out one or more points whose outlier coefficient is less than a certain value.

4. Change the found point value to the average value of other points.

2.2. Wavelet packet decomposition module

Because of the long production cycle of industrial users, the corresponding signal frequency of each process is mainly concentrated in the low frequency band. By decomposing the signal with wavelet packet, the low-frequency information of the signal is obtained, and the spectrum of the low-frequency signal is analyzed.

Wavelet packet decomposition is an improvement and extension of wavelet decomposition\(^{[10]}\). The traditional wavelet decomposition first decomposes the signal into approximation part and detail part, then further decomposes the approximation part into approximation part and detail part, and repeats such decomposition until the decomposition requirements are met. WPD decomposes the details at the same time. The time scale of WPD is arbitrary, and there will be no time-frequency fixing problem. Therefore, it is mostly used in time-frequency analysis, which can better reflect the characteristic information of the signal.

In non-intrusive monitoring of large data industrial users, according to the size of scale factor \( j \), WPD decomposes Hilbert space \( L^2(R) \) into orthogonal sums of multiple wavelet subspaces \( W_j (j < \mathbb{Z}) \), as shown in Figure 2, and subdivides each wavelet subspace according to binary. Where \( U_j \) is the \( n \)-th (\( n = 0, 1, 2, \ldots, 2^j \)) wavelet subspace, \( j \) is called the scale of wavelet subspace, and its corresponding orthogonal basis is \( u_j(t) = 2^j u^n(2^j t - k) \), \( k \) is the translation factor.

\[
\begin{array}{c|c|c|c|c|c|c|c}
   & U_j(V_0) & U_j(V_1) & U_j(W_1) & U_j(W_2) & U_j(V_2) & U_j(W_3) & U_j(V_3) \\
   \hline
   U_j(V_1) & U_j(W_2) & U_j(V_2) & U_j(W_3) & U_j(V_3) & U_j(W_4) & U_j(V_4) & U_j(W_5) \\
   \hline
   \vdots & \hline
   \end{array}
\]

**Figure 2.** Wavelet packet space decomposition.

In the decomposition process, if the scale is small enough, the sampling sequence \( f(k\Delta t) \) of the function \( f(t) \) in \( L^2(R) \) space is directly used as the coefficient \( d_0(k) \) in \( U \) space. According to the algorithm principle of orthogonal wavelet transform, the wavelet decomposition coefficients of order \( j \) and \( k \) can be written as follows:
\[ d^n_j(k) = \sum_m h_0(m - 2k)d_{j-1}^{n/2}(m), \text{ (n is even)} \]
\[ d^n_j(k) = \sum_m h_1(m - 2k)d_{j-1}^{(n-1)/2}(m), \text{ (n is odd)} \]

where, \( h_0 \) and \( h_1 \) are two orthogonal mirror filters, and the decomposition coefficients of stage \( j \) are expressed by the coefficients of stage \( j-1 \). According to this recursive relation, the decomposition coefficients of each level corresponding to signal \( f(k) \) can be obtained. After \( j \)-level decomposition, the corresponding frequency band interval of each subspace is:

\[ \{ [0, \frac{f_s}{2^{(j-1)}}); [\frac{f_s}{2^{(j-1)}}, \frac{2f_s}{2^{(j-1)}}); [\frac{2f_s}{2^{(j-1)}}, \frac{3f_s}{2^{(j-1)}}); \ldots; (\frac{2^j-1}{2^{j-1}}f_s, \frac{f_s}{2^{j-1}}]; \} \]

where \( f_s \) is the sampling frequency of the signal \( f(t) \).

By calculating the energy spectrum of the load signal in each frequency range, the main information of the industrial user load signal can be determined in which frequency range according to the size of the energy spectrum, so that the corresponding signal of the energy spectrum can be used for spectrum analysis. The formula for calculating the energy spectrum is [11]:

\[ E(j, n) = \sum [d^n_j(k)]^2 \]

Among them, the recursive algorithm of formula (4) and formula (5) is used to calculate the wavelet packet decomposition coefficient \( d^n_j(k) \).

2.3. Fast Fourier transform module

The load signal decomposed by wavelet packet is decomposed into several signals of different frequencies by improved fast Fourier transform. The characteristic information of industrial user load signal is extracted effectively, which lays a foundation for determining the operation and cycle of various production processes.

The improved fast Fourier transform method makes use of the odd, even, virtual and real characteristics of discrete Fourier transform to improve the DFT [12]. In this study, the FFT extracted by time and the FFT extracted by frequency are included.

For finite discrete data \( x(n), n = 0, 1, \ldots, N-1 \) DFT is defined as:

\[ X(k) = \sum_{r=0}^{N-1} x(r)W_N^{-kr} \quad k = 0, 1, \ldots, N-1; W_N = e^{-\frac{2\pi}{N}} \]

In order to decompose a large point DFT into a small point DFT, the length of the sequence \( N \) must be a composite number. In this paper, \( N = 2m \) (\( m \) is a positive integer). The sequence \( x(n) \) is decomposed into two groups according to odd and even terms by time-extracted FFT:

\[
\begin{align*}
    x(2r) &= x_1(r) \\
    x(2r+1) &= x_2(r)
\end{align*}
\]

\( r = 0, 1, \ldots, \frac{N}{2} - 1 \)

Then the DFT of two \( N/2 \) points is:

\[ X(2r) = \sum_{n=0}^{N/2-1} x_1(n)W_N^{nr} \]
\[ X(2r+1) = \sum_{n=0}^{N/2-1} x_2(n)W_N^{nr} \]

3. Case Analysis of Non-intrusive Monitoring for Large Power Data Industry Users
Steel is an important traditional pillar industry in Henan Province. The production process of steel is complex, especially in the whole process steel industries, the production system is huge. By using the non-intrusive monitoring method provided in this study, the operation and cycle of the main production processes can be determined, and the industrial production process can be monitored in detail according to the electricity consumption situation of industries, thus providing technical means for monitoring the implementation effect of "de-productivity". It can also obtain the output of steel production for a period through the calculation of typical links and technological power capacity and provide technical support for supervision of key industrial users to reduce production capacity and energy consumption.

3.1. Preprocessing of Large Load Data in Steel Industries
The data preprocessing method is used to preprocess the load data, which can reduce the impact of some abnormal values and missing data, and more accurately reflect the real power consumption situation of users.

Taking the original load data of a steel industry in Henan Province (not multiplied by the comprehensive ratio) as an example, the data set collected by the meter on a certain day appears to be missing. As shown in Figure 3, the sampling time interval of the data set is 15 minutes, the missing time interval is 21:00-22:00, and a total of 5 values are missing.

According to the actual missing situation of data sets, the mapping method of adjacent data sets is selected to fill all missing data. The length of sliding window is set to 12, which means the missing data is filled by data sets within 3 hours, and the maximum missing rate is set to 50%. When the missing rate is greater than the set value, the missing data is not filled completely. In the above case, when the sliding window moves to 19:30, its window range is 79-90, and the current data missing rate is $\frac{5}{12}=42%<50\%$. It satisfies the condition of data completion. The data set after completion is shown in Figure 4.
3.2. Decomposition of process flow in steel industry by non-intrusive technology

By preprocessing the data of electricity load from March to September in the steel industry, the original load data of electricity consumption in the industry is improved. Combining with the comprehensive ratio 264000. The sampling period is 5 minutes, with a total of about 63000 power points.

Wavelet packet decomposition in the third layer is used to decompose the original load signal of the industry, and the energy spectrum is calculated. The three-layers decomposition has 2, 4 and 8 bands, totaling 14 bands. The signal decomposition diagram of each band is shown in Figure 5.

Figure 4. Completed ammeter data.

(a) Wavelet Packet Decomposition in the First and Second Layers
The sampling period is 5 minutes, and the corresponding sampling frequency is 1/300. The signal of the third low frequency band filters out the high frequency useless signal and contains most of the load signal information. The frequency band signal is decomposed by FFT to obtain the power spectrum information of steel industry, as shown in Figure 6. Through spectrum analysis, the frequency of rolling, steelmaking and ironmaking in this steel industry is about 1/1200Hz, 1/1800Hz and 1/3600Hz, and the period is about 20min, 30min and 60min, which is consistent with the actual production situation. After obtaining different process flow curves and calculating the average value of the waveform curve, the absolute value of the difference between each point and the average value is calculated, a reasonable threshold is set and the points whose absolute value is less than the threshold value and the slope is positive is screened out. If the point is greater than the average value, it is the starting point, and if it is less than the average value, it is the stop point. According to the decomposition of different process flow in the enterprise, the start-up and shutdown of each process flow in a one-day cycle are obtained as shown in Figure 7.
Figure 7. Electrical decomposition diagram for each process.
Through non-intrusive decomposition of the process flow in steel plants, the industrial production process can be monitored in detail, so that the production situation of industries can be effectively monitored "remotely", and the occurrence of "should stop but not" production situation can be prevented, thus providing technical means for monitoring the implementation effect of "no capacity". Through non-invasive decomposition to judge the start-up and stop of each process, the operation status of each link can be more carefully grasped. According to the number of start-up and shut-down times per day and the output of each link in a single start-up operation, the daily output of each link can be preliminarily estimated. The results are compared with the average power consumption of steel and iron to calculate the production capacity and output, to grasp the production situation of enterprises more reliably and further test the implementation of "de-production capacity".

3.3. Analysis of Production Capacity and Output of Steel Industries

Usually, the electricity consumption of steelmaking and rolling links in steel industries accounts for about 40% of the total electricity consumption, and that of ironmaking links accounts for about 20% of the total electricity consumption.

According to the monthly electricity consumption information, the energy consumption of steel and iron produced by steel industry is estimated, as detailed in Table 1.

| Table 1. Calculation of electricity consumption of steel produced by steel enterprises. |
|-----------------------------------------------|
|                                | March | April | May    | June   | July   | August  | September |
| Power consumption of steel production (MWh) | 30608.04 | 34840.22 | 37348.53 | 37342.49 | 37633.43 | 35510.88 | 32982.25 |
| Power consumption of iron production (MWh)  | 15304.02 | 17420.11 | 18674.26 | 18671.24 | 18816.71 | 17755.44 | 16491.12 |

Through visiting and investigating several steel industries and statistical data of relevant departments [13-15], it is shown that the power consumption per ton of steel and iron is 430-470 kWh and 180-220 kWh respectively. According to the monthly power consumption of steel industries and the power consumption per ton of steel and iron, the production range of steel and iron can be calculated, as shown in Table 2.

| Table 2. Estimation of output of steel enterprises. |
|-----------------------------------------------|
|                                | March | April | May     | June    | July     | August   | September |
| Steel production (Ton)          |       |       |         |         |          |          |           |
| Maximum                        | 71181.4 | 81023.7 | 86857.04 | 86843.00 | 87519.61 | 82583.4 | 76702.9 |
| Median                         | 68152.4 | 77575.9 | 83161   | 83147.56 | 83795.37 | 79069.2 | 73438.95 |
| Minimum                        | 65123.4 | 74128.1 | 79464.96 | 79452.11 | 80071.13 | 75555.0 | 70175.0 |
| Iron production (Ton)          |       |       |         |         |          |          |           |
| Maximum                        | 85022.3 | 96778.4 | 103745.9 | 103729.1 | 104537.3 | 98641.3 | 91617.36 |
| Median                         | 77293.0 | 87980.3 | 94314.47 | 94299.22 | 95033.92 | 89673.9 | 83288.51 |
| Minimum                        | 69563.7 | 79182.3 | 84883.02 | 84869.30 | 85530.52 | 80706.5 | 74959.66 |

4. Conclusion
This paper studies the non-intrusive monitoring method for industrial users based on large data of electric power, realizes the effective decomposition of the main production processes of steel industries, and takes a steel industry in Henan as an example, realizes the decomposition of the main production processes such as ironmaking, steelmaking and rolling, determines the operation cycle of each process, and estimates the production capacity of the industry. Comparing the calculated capacity with the capacity limitation stipulated by the national policy, it is helpful for the supervisory department to control the production situation of industrial users and better implement the national policies and regulations. Non-intrusive monitoring of industrial users can help relevant departments monitor the production situation of industries according to real-time electricity consumption, and reasonably analyze the accomplishment of the goals of "de-productivity" based on relevant data. Then, by analyzing the production capacity and power consumption of industrial users, it can grasp the production level and energy consumption level of industrial users and provide indicators support for further formulation of the "de-productivity" target, form a "closed-loop policy" to guide industrial users to transform and upgrade to high-efficiency productivity, and can also help users understand their electricity consumption habits and main influencing factors, provide industrial users with electricity optimization suggestions, and further guide industrial users to use electricity scientifically, economize on electricity and improve their electricity efficiency. Therefore, through the research results of this paper, we can better grasp the implementation effect of key enterprises' policies such as capacity removal, energy saving and emission reduction, monitor the operation of enterprises, and provide key support for the formulation and adjustment of policy initiatives.

5. References
[1] Wu Xianli. Problems and Countermeasure of Supply Side reform in Henan's Industrial Field [J]. Journal of Luoyang Institute of Science and Technology (Social Science), 2018, 33(05): 34-38.
[2] Hart G W. Nonintrusive appliance load monitoring [J]. Proceedings of the IEEE, 1992, 80(12): 1870-1891.
[3] W.C. Bush, M. Bons, Marc. Nonintrusive End-use Monitoring in Europe[C]. 1998 DA/DSM Proceedings, London 1998
[4] W.C. Bush. Single Point End-use Energy Disaggregation (SPEED) [C]. Proceedings of the International End-Use Monitoring Symposium, Amsterdam, The Netherlands, September 1998
[5] Steven Drenker, Ab Dader. Nonintrusive Monitoring of Electric Loads[J]. IEEE Computer Applications in Power, 1999, 12(4): 47-51
[6] S.B. Leeb, S.R. Shaw, J.L. Kirtley. Transient Event Detection in Spectral Envelope Estimates for Nonintrusive Load Monitoring[C]. IEEE PES Winter Meeting, New York, January, 1995: 95-105
[7] Robert W. Cox, LCDR Patrick L. Bennett, LCDR T. Duncan Mc Kay, et al. Using the Non-Intrusive Load Monitor for Shipboard Supervisory Control[C]. Electric Ship Technologies Symposium, 2007: 523-530
[8] Robert Cox, Steven B. Leeb, Steven R. Shaw, et al. Transient Event Detection for Nonintrusive Load Monitoring and Demand Side Management Using Voltage Distortion[C]. Applied Power Electronics Conference and Exposition, 2006: 1751-1757
[9] JG Roos, IE Lane, EC Botha, et al. Using Neural Networks for Nonintrusive Monitoring of Industrial Electrical Loads[C]. Instrumentation and Measurement Technology Conference, 1994, 3: 1115-1118
[10] Wu Ting, Yan Guozheng, Yang Banghua. EEG feature extraction in brain computer interface based on wavelet packet decomposition [J]. Chinese Journal of Scientific Instrument, 2007, 28(12):2230-2234.
[11] Li Shiling, Li Zhi, Li Hesheng. The Method of Roller Bearing Fault Monitoring Based on Wavelet Packet Energy Feature [J]. Journal of System Simulation, 2003(1).
[12] Qian Hao, Zhao Rongxiang. Interharmonic Analysis Based on Interpolated FFT Algorithm[J].Proceedings of CSEE,2005, 25(21):87-91.
[13] Tian Jinglong. Energy consumption and energy saving analyses of top ten iron & steel plant[J]. Energy of Metallurgical Industry, 2007(06):3-7.

[14] Wang Weixing. Analysis on Energy Consumption Status and Energy Saving Potential of China's Iron and Steel Industry[J]. China Steel Focus, 2017(8):50-58.

[15] Liu Xiangbin. Study on Reasonable Production and Optimization of Electricity in Iron and Steel Enterprises[D]. Northeastern University, 2012

Acknowledgments
This work supported by the The National Natural Science Foundation of China (51607068); The Fundamental Research Funds for the Central Universities(2017MS090).