Curve tracking and comparison during electricity spot trading based on judgment methods for curve similarity

Shuai Zhang\textsuperscript{1,2}, Shijun Chen\textsuperscript{1,2*}, Guangwen Ma\textsuperscript{1,2}, Yanmei Zhu\textsuperscript{1,2}, Chunhua Tao\textsuperscript{3}

\textsuperscript{1} College of Water Resource & Hydropower, Sichuan University, Chengdu 610065, China;
\textsuperscript{2} State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, China;
\textsuperscript{3} Dadu River Basin Hydropower Development Co., Ltd. of National Energy Group, Chengdu 610041, China

Abstract. Accurate tracking of load curves during transactions in the electricity generation process and timely identification and adjustment of aberrant electricity prices are key to ensuring profits from bidding in the production of enterprises. By selecting three judgement methods for curve similarity, \textit{i.e.} Euclidean metric, correlation coefficient and mean square error (COE-MSE), and discrete Fréchet distance, this study tracked and compared an electricity generation curve, a transaction curve, and a bidding and declaration curve of a hydropower station where reservoirs have strong regulating capacity in a river basin in the Sichuan Province power grid. The advantages and disadvantages of the three methods were analysed and compared. The results demonstrate that the COE-MSE and discrete Fréchet distance can meet the requirements of electricity generation according to the transaction results when electricity generation enterprises participate in electricity spot trading. Furthermore, the discrete Fréchet distance shows advantages in identifying eigenvalues and tracking trends of the curves.

1. Introduction

It has been five years since a new round of the reformation of the electricity system in China and the reform has been deepened to a significant extent. By the end of 2019, trial operation or settlement of spot trading has been successively conducted in eight pilot spot markets for trading electricity, represented by Guangdong and Sichuan Provinces, based on the stable and orderly development of mid and long-term contract transactions. This marks the full development of spot trading. The new transaction modes have imposed more stringent requirements on the electricity sale and dispatch operations on the generation side.

With the continuous expansion of the electricity system and constant liberalisation of energy marketization, the power grid and users have become more diversified in terms of electricity purchase options, which imposes more onerous requirements on the quality of electricity on the generation side \cite{1,2}. At present, hydropower is the most widely used renewable energy source in China, and is an indispensable part of the power source structure in the region with its rich water resources \cite{3}. For example, by the end of 2019, the installed capacity of hydropower stations accounted for nearly 80% of the total installed capacity of the electricity system in Sichuan Province. The uncertain demands on
water inflow and load faced by hydropower stations lead to fluctuations of load and problems affecting the quality of electricity, thus influencing the overall stability and reliability of the power grid \[^4\]. In the absence of markets with auxiliary services in the current market structure, large-scale regulatory hydropower stations need to undertake the task of peak load and frequency regulation to balance electricity in the system \[^5\], which increases the fluctuation of load for electricity generators to a certain extent.

During electricity spot trading, the real-time transaction results of load will be sent to a hydropower station in the form of a dispatch order through a trading centre. Faults in unit operation, adjustment of network constraints, and execution errors when handling the dispatching order are more likely to cause the deviation of the actual electricity generation from the dispatch order, leading to the generation of unqualified electricity. Unqualified electricity is an important factor influencing the quality of electricity. The accumulation of unqualified electricity in many hydropower stations in the electricity system results in a large deviation of intraday electric generation and dispatch curves from the actual electricity consumption curve, thus affecting the overall stability of the electricity system and increasing the complexity of the transaction and settlement system \[^6\]. To standardise the power dispatch behaviours and improve the operation and management levels of hydropower stations, government departments have issued implementation rules for dispatching and management to restrict unqualified electricity generation.

Cascade hydropower stations have complex hydraulic and electric connections \[^7\]. The generation of unqualified electricity in hydropower stations in the upper reaches of a river implies a change in water discharge processes at hydropower stations, thus affecting the fulfilment of the electricity generation target and increasing the difficulties in dispatching operations at hydropower stations in the lower reaches. Therefore, it is necessary to explore a scientific and effective method of tracking real-time electricity generation, compared with the transaction results of load in the real-time markets and monitor the deviation in electricity generation, thus minimising generation of unqualified electricity.

The load for electricity generation in hydropower stations and transactions in the real-time markets are sorted chronologically to form a time series curve \[^8\], which can be tracked and compared with the help of judgment methods for curve similarity. At present, the judgment methods for curve similarity are clearly defined and completely deduced with detailed proofs thereof: these form a complete theoretical system and have been widely used. For instance, in recent years, as a secure and efficient encryption technology, dynamic signature verification has attracted much attention and its core technology is based on judgment of curve similarity in human biology \[^9\]. By abstracting a DNA sequence into two-dimensional (2-d) curves, Guo-Sen Xie \textit{et al.} \[^10\] compared similarity of DNA sequences of different species based on the principle of judgment of curve similarity. Rongheng Lin \textit{et al.} \[^11\] clustered load curves of smart grid systems by virtue of enhanced Pearson similarity. The Fréchet distance can be used for calculation without translation and stretching of the original curves, so it has become a commonly used method for judging similarity \[^12\]. After some improvements, it has been developed into a discrete Fréchet distance measure \[^13\-14\], which can better meet the requirements for judgment of curve similarity.

Electricity spot trading imposes more onerous requirements on the refined management of dispatching operation of cascade hydropower operations. At present, existing research mainly focuses on the bidding and declaration methods in the day-ahead and real-time markets and design of algorithms for high-efficient clearing models and does not pay enough attention to the deviation in the production and operation of hydropower stations after a transaction. Moreover, electricity generation enterprises of the cascade hydropower stations lack special tracking and statistical methods for unqualified electricity generation during production after a transaction. The aim of this study is to match real-time transaction results of load with spot trading in cascade hydropower stations and improve operation and management levels of electricity-generating plants and the quality of electricity on the generation side. Based on the mainstream judgment principles and methods of curve similarity, the time sequences of intraday load produced in the processes of declaration and clearing in the spot market and production were tracked and compared. The following results are obtained:
(1) Based on the typical pilots of spot market trading of high-proportion renewable energy (electricity), namely Trading Rules in Spot Power Markets of Sichuan Province, the compositions of spot trading curves in detail and the causes and influences of unqualified electricity were analysed. Moreover, the necessity of tracking and comparing load curves during trading was illustrated. (2) From the definition of curve similarity, the calculation method for indices used for judging curve similarity selected in this research and the criteria for judging similarity were introduced. (3) Taking the real-time transaction and execution process of load on a certain day in a large frequency-modulation hydropower station in Sichuan Province as an example, the applicability of the algorithms was verified. The calculation results demonstrate that the discrete Fréchet distance can track and compare real-time transaction results of load in electricity generation, timeously adjust unqualified electricity, and minimise the influence on stability of the power grid.

The rest of the study is arranged as follows: Section 2 introduces the trading rules in electricity spot markets in Sichuan Province and explains tracking and comparison of spot trading curves. Section 3 illustrates the definition of curve similarity and introduces the principle and calculation process of three judgment methods for curve similarity. By taking declaration and clearing during spot trading and production load curves in the typical peak-load regulatory hydropower station in Sichuan Province on a certain day as examples, Section 4 compares the advantages and disadvantages as well as applicability of the three judgment methods: conclusions are drawn in Section 5.

2. Curve tracking and comparison during spot trading

When electricity generation enterprises are not involved in real-time market bidding, the load curve after market-clearing is composed of the following parts: an inter-provincial power-transmission curve, a preferential load curve, a curve of decomposition of mid and long-term contract power, and a clearing curve in a day-ahead market. Each part was settled by different methods. The first two were settled by inter-provincial transaction price and catalogue price, while the mid and long-term contract power and the cleared electricity in the day-ahead market were settled based on the contract for difference (Fig. 1). Electricity generation needs to be organised according to the transaction curve of load on the generation side, but the deviation of the actual electricity generation from the transaction curve readily occurs due to unit operation errors and human error in the dispatch process. If the deviation exceeds ± 2%, the generated electricity is considered as unqualified, and is assessed in accordance with relevant regulations.

There are complex non-linear hydraulic and electric connections among the hydropower stations in a group of cascade hydropower stations. The electricity generation process of hydropower stations is restricted by water influx: when the electricity generation process of hydropower stations in the upper reaches changes the reservoir discharge, it affects the storage process at hydropower stations in the lower reaches. In particular, the task of peak load and frequency regulation of the power grid undertaken by hydropower stations increases uncertainty of water influx to the lower reaches. According to the statistics pertaining to unqualified electricity in the actual production process, more than 90% of unqualified electricity is produced in the process of increasing or decreasing load. Frequent adjustment of load can increase human errors and the AGC fault rate, which is the main reason for the deviation in electricity generation.

The generation of unqualified electricity is often accompanied by the loss of profit from electricity generation among hydropower stations. If hydropower stations fail to timeously reduce load, the generated electricity in a certain period will exceed the transacted electricity. In this case, unqualified electricity is positive and this excessive electricity generated will affect the security and stability of the power grid. In addition, the electricity generation side should bear certain costs for assessing unqualified electricity. On the contrary, failure of hydropower stations in timeously increasing load, insufficient water influx, low water head during unit operation, and insufficient capacity of transmission lines on the power grid can result in the electricity generated in a certain period being less than the transacted electricity. In this case, unqualified electricity is negative. The deviation on the markets on the generation side can be settled according to the settlement method of the contract for
difference. Based on economic experience, the price in the real-time markets is usually higher than that in the day-ahead markets, and the profits of electricity generation enterprises from electricity generation will suffer accordingly.

![Compositions of curves in the day-ahead market and settlement methods](image)

**Fig. 1** Compositions of curves in the day-ahead market and settlement methods

For curve tracking during spot trading, through real-time interaction with the monitoring system of hydropower stations, the output process data pertaining to electricity generation at hydropower stations are collected at 15-minute intervals during spot trading. Furthermore, the transaction results of load in the real-time markets are obtained from the spot declaration system of the trading centre to form the load transaction curve and actual electricity generation curve of hydropower stations. For curve comparison, the judgment methods for similarity were used. If the similarity is strong, the electricity generation will be continued according to the transaction results of load; otherwise, the load needs to be adjusted according to the real-time markets, and electricity is then generated according to the adjusted load curve. The specific steps are presented in Fig. 2

### 3. Judgment methods for curve similarity

**Definition 1**: for two functions \( f_1(x) \) and \( f_2(x) \),

\[
\text{d}(f_1, f_2) = \int_{c_0}^{c_1} |f_1(x) - f_2(x)| \, dx
\]

represents the distance between two curves. \([c_0, c_1]\) is the domain of definition of the functions. For the given \( \varepsilon \), if \( \text{d}(f_1, f_2) \leq \varepsilon \), then \( f_1(x) \) and \( f_2(x) \) are similar; otherwise, they are dissimilar.

The curve can be expressed as a combination of multi-segment functions after translation and stretching. Therefore, in the calculation of curve similarity, it is necessary to segment the curves first and transform the curves segment-by-segment, then conduct matching calculation. The distance between the transformed curves can be defined as follows:

\[
d(\tilde{f_1}, \tilde{f_2}) = \int_{c_0}^{c_1} |k f_1(x) + h - f_2(x)| \, dx
\]

where, \( k \) and \( h \) indicate the scaling ratio and translation ratio in the vertical direction, respectively.

The curves in this study can be expressed as corresponding continuous functions in a 2-d coordinate system and the process of segmentation and transformation is very simple during comparison. According to the research status of judgment of curve similarity in the world, the following three methods were selected for judging similarity.

#### 3.1. Euclidean metric

The Euclidean metric is the most commonly used distance definition at present. In a 2-d coordinate range, the Euclidean metric is the real distance between two points within a specific range and is the length of the line between two points after coordinate representation.

The corresponding coordinates of each point on the two curves in the same coordinate system are \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\) and \((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\). The Euclidean metric between the points is given by:

\[
p = \sqrt{(x_i - X_i)^2 + (y_i - Y_i)^2}
\]
The average Euclidean metric of each point on the two curves is given by:

$$\overline{\rho_i} = \frac{1}{n} \sum_{i=1}^{n} \rho_i$$  \hspace{1cm} (3)

In the similarity judgment, whether the average Euclidean metric is close to 0 is first judged. If the deviation from 0 is large, the curves have no similarity; when the mean approaches 0, the discreteness of the average Euclidean metric of each point can be judged based on indices of mathematical statistics, such as range and variance. The more concentrated the mean distance, the stronger the similarity.

Fig. 2 Technology roadmap

3.2. COE-MSE

The method is a comprehensive judgment method combining the correlation coefficient (COE) with the mean square error (MSE). The points on the two curves are plotted in the same coordinate system with the same abscissa, that is, the coordinates of each point on the two curves are $(x_i, y_i)$ and $(x_i, y_i')$ ($i = 1, 2, \ldots, n$). The ordinates of each point on the two curves constitute two sequence sets, namely $\{y_1, y_2, \ldots, y_n\}$ and $\{y'_1, y'_2, \ldots, y'_n\}$. According to Formulae (4) and (5), the COE and MSE are separately calculated:

$$COE = \frac{\frac{1}{n} \sum_{i=1}^{n} y_i y_i'}{\sqrt{\left(\frac{1}{n} \sum_{i=1}^{n} y_i^2\right) \left(\frac{1}{n} \sum_{i=1}^{n} y_i'^2\right)}}$$  \hspace{1cm} (4)

$$MSE = \frac{\sum_{i=1}^{n} (y_i - y_i')^2}{n}$$  \hspace{1cm} (5)

The larger the absolute COE, the stronger the correlation. To be specific, the closer the COE is to 1 or - 1, the stronger the correlation; the closer the COE is to 0, the weaker the correlation. If the sequences constituted by ordinates of each point on the two curves are strongly correlated or extremely correlated (COE $> 0.8$), it can be considered that the trends of the two curves are basically consistent.

Based on the same trend in the curves, if the MSE of the sequences constituted by ordinates of each point on the two curves is close to 0, the similarity of the two curves is very strong. By successively...
observing errors of the ordinates of two points at the same position, if most of them are 0 and only some have large errors, this indicates that the curves have strong similarity. If the MSE is large and all of errors of the ordinates of two points at the same position are non-zero, it is considered that the two curves are dissimilar.

3.3. Discrete Fréchet distance

This is a comprehensive method used to judge curve similarity comprehensively by combining the identification of eigenvalue points on the curves with the numerical calculation of the similarity indices. In the judgment, the eigenvalue points on the curves are first identified, which are usually the highest (lowest) points. The definition and criterion are as follows: one point on the curve that is larger (smaller) than the two points in front and rear thereof belongs to the highest (lowest) point. If the two curves are similar, the difference in the number of the highest (lowest) points on the two curves is not larger than three, and the time difference of the highest (lowest) points is no more than two intervals; otherwise, the similarity of the two curves is weak.

Definition 2 [13]: (1) Given a polygon chain \( P = (P_1, P_2, P_3, \ldots, P_n) \) with \( n \) highest points, \( k \) disjoint non-empty subsets \( \{P_i\}_{i=1}^{n-k} \) are obtained by dividing the highest point of \( P \) using a \( k \)-walk along \( P \), so that \( P = (P_{n-k+1}, \ldots, P_n) \) and \( 0 = n_0 < n_1 < \cdots < n_k = n \).

(2) Given two polygon chains \( A = (a_1, \ldots, a_m) \) and \( B = (b_1, \ldots, b_n) \), there is one paired walk composed of a \( k \)-walk \( \{a_i\}_{i=1}^{m-k} \) along \( A \) and a \( k \)-walk \( \{b_i\}_{i=1}^{n-k} \) along \( B \) separately. Therefore, for \( 1 \leq i \leq k \), \( |A_i| = 1 \) or \( |B_i| = 1 \), that is, \( A_i \) or \( B_i \) exactly contains a highest point.

(3) The cost of the paired walk \( W = \{(A_i, B_i)\} \) along the chains \( A \) and \( B \) is expressed as follows:

\[
 d^w_F(A, B) = \max_{(a, b) \in A \times B} \text{dist}(a, b) \tag{6}
\]

The discrete Fréchet distance between chains \( A \) and \( B \) is expressed as follows:

\[
 d_F(A, B) = \min_W d^w_F(A, B) \tag{7}
\]

The discrete Fréchet distance can be calculated based on a man-dog model. Assuming that the two curves \( A \) and \( B \) are the walking routes of a man and a dog, respectively, they can be generalised into ordered point series \( (A_1, A_2, \ldots, A_m) \) and \( (B_1, B_2, \ldots, B_n) \). When walking, the man and the dog walk from the pair of starting points \((A_1, B_1)\) and walk forward along their respective curves to the pair of end points \((A_m, B_m)\). The discrete Fréchet distance \( d_F(A, B) \) is the minimum distance between the man and the dog.

When the man and dog are in any position \((A_i, B_i)\) on their respective curves, they move according to the following three situations:

(a) When the man steps forward and the dog stays still, their positions are \((A_{i+1}, B_i)\).

(b) When the man stays still and the dog steps forward, their positions are \((A_i, B_{i+1})\).

(c) As the man and dog step forward at the same time, their positions are \((A_{i+1}, B_{i+1})\).

In conclusion, \( d_{Fr} \) can be calculated according to the recursive Formula (8).

\[
 d_F(A, B) = \max \left\{ \min_{(a, b) \in A \times B} \text{dist}(a, b) \right\}
\]

\[
 d_F(A_n, B_m) = \max \left\{ \min_{(a, b) \in A_n \times B_m} \text{dist}(a, b) \right\}
\]

\[
 d_F((A_1, A_2, \ldots, A_{n-1}), (B_1, B_2, \ldots, B_{m-1})) = \max_{n \neq 1} \min_{m \neq 1} d_F(A_n, B_m) \tag{8}
\]

where, \( d_F(A_n, B_m) \) indicates the Euclidean metric between points \( A_n \) and \( B_m \), which can be regarded as strings. The discrete Fréchet distance between \((A_1, A_2, \ldots, A_{n-1})\) and \((B_1, B_2, \ldots, B_{m-1})\) can be calculated based on Formula (8). When the string is finally reduced to the starting point \((A_1, B_1)\), the calculation stops. In this case, \( d_F((A_1), (B_1)) = d_F(A_1, B_1) \).
4. Comparative analysis of examples

The hydropower station M is a hydropower station that has a reservoir with strong regulating capacity on the lower reaches of the river basin. The reservoir has a total storage capacity of 5.122 × 10^9 m^3 and its regulating capacity of 3.826 × 10^9 m^3 is nevertheless an incomplete annual regulatory capacity. The hydropower station is equipped with six generation units, with a total installed capacity of 3600 MW, supplying electricity to the Sichuan power grid; because the hydropower station is close to the load centre and the installed capacity is sufficient, the load fluctuation of the power grid can be matched by AGC in real time. Therefore, the hydropower station is the main hydropower station for peak load and frequency regulation in the Sichuan power grid. Besides participating in the market transaction of electricity, the hydropower station provides auxiliary services of peak load and frequency regulation for the power grid. In the combined electricity generation of cascade hydropower stations, the load adjustment of hydropower station M will cause significant fluctuations of water influx to cascaded hydropower stations located downstream. Therefore, it is necessary to study scientific and reasonable measures available to track the real-time electricity generation of the hydropower station and compare this with the transaction results of load to avoid the generation of unqualified electricity.

By utilising crawling technology, network data, declared load, and clearing results in real-time markets for hydropower station M can be obtained from the declaration and release platform of electricity spot markets. The intra-day real-time electricity generation process can be obtained by accessing the data interface of the monitoring system of the hydropower station. The hydropower station generates electricity in strict accordance with the clearing results in the real-time markets and does not apply for load adjustment to the power grid dispatching department. The transacted load of the hydropower station is arranged chronologically to form three load curves, namely an actual load curve (reality), a declared load curve (declaration), and a transaction curve of load (transaction).

By using the comparison methods for similarity described in Section 3, the similarity of Reality & Declaration (RD) curves was compared. To verify the applicability of the algorithm proposed in this study, with Reality & Transaction (RT) curves as the controls, the same method was used to judge the similarity and the results are summarised below.

The Euclidean metric between the two curves was calculated and the curve similarity was judged based on indices of mathematical statistics. Fig. 4 (a) demonstrates the distribution of Euclidean metrics of RD and RT curves in time. Table 1 and Fig. 4 (b) show calculation results of the indices of mathematical statistics for the Euclidean metric between the two curves and their similarity comparison.

| Index | Average | Max | Min | Range | S^2     |
|-------|---------|-----|-----|-------|---------|
| RD    | 200.89  | 574.45 | 5.13 | 569.32 | 18993.81 |
| RT    | 25.95   | 69.78 | 0.66 | 69.12  | 289.06  |

As shown in Fig. 4, intra-day Euclidean metrics between the RT curves are very small and in [0,100], with the maximum value being 69.78. The calculated mathematical statistics also show that the average Euclidean metric is small. The small variance indicates that the Euclidean metrics at each time point are near to the average, which is in line with the criterion for judging the similarity of load curves. In contrast to the RD curves, the Euclidean metrics in some time periods are small, but their variance is large. The Euclidean metrics between points are evenly dispersed between the maximum and minimum values. According to the definition of curve similarity based on the Euclidean metric, the two curves are dissimilar. According to Euclidean metric-based similarity judgment, the following conclusions can be drawn: the RT curves have similarity at all-time points in the day. In other words, the inter-day electricity generation of the hydropower station is conducted according to the transaction curve of load, while the single-point deviation of the RD curves is relatively large, so there is no similarity.
Based on the COE-MSE method, the conclusions of curve similarity judgement are obtained by combining curve trends with single-point errors. The calculation results of two indices of the RD and RT curves are listed in Table 2. COEs of the two pairs of curves are 1.0, that is, the curves are strongly correlated, proving that the reality curve is strongly consistent with declaration and transaction curves in terms of trend. Based on Fig. 5, the deviation of the two curves at each time point was analysed.

Fig. 5 (a) displays the size and distribution of relative deviations of the two pairs of curves at each time point. Fig. 5 (b) demonstrates several statistical indices of absolute and relative deviations thereof. As shown in Fig. 5 (a), the relative deviation of the RT curves is less than 2%, which meets the requirements of deviation assessment in electricity markets, while that of the RD curves is large. In terms of comparison of statistical indices, the maximum, minimum, and average absolute and relative deviations of the RD curves are larger than those of the RT curves. In summary, the calculated COEs of the two pairs of curves show that they (separately) have excellent similarity in terms of trend therein. The deviations of the RT curves are much smaller than those of the RD curves based on meeting the requirements for load deviations. Compared with the COE-MSE method, the RT curves have extremely high similarity, while the RD curves only show consistency in terms of trend.
Table 2 COE-MSE calculation results for the two pairs of curves

| Index | COE   | MSE   |
|-------|-------|-------|
| RD    | 1.00  | 59153 |
| RT    | 1.00  | 959   |

Fig. 5 Comparison of deviations of the RD and RT curves, (a) and (b)

The discrete Fréchet distance was used to judge the similarity separately between the RT and RD curves, respectively and the highest (lowest) points of the two pairs of curves were collected and compared. As illustrated in Fig. 6, red and blue separately represent the highest and lowest points on the curves. For the RT curves, the numbers of the highest points are 15 and 13, while those of the lowest points are 11 and 13, respectively. There are 12 highest points and 15 lowest points on the declaration curve. In terms of time distribution, there are 22 groups of the highest (lowest) points appearing in the same period on the RT curves. Although the other four groups do not appear in the same interval, their time difference is less than two intervals, however, only 10 groups the highest (lowest) points appear in the same period on the RD curves, and there are multiple groups of eigenvalue points with a time difference larger than two intervals. Therefore, the similarity of the RT curves is stronger than that of the RD curves from the perspective of eigenvalue points.

The numerical calculation results show that the discrete Fréchet distance between the RT curves is 63.83, and that between the RD curves is 441.28. According to the distribution of eigenvalue points and the definition of curve similarity based on the discrete Fréchet distance, the RT curves are determined to be similar, that is, the electricity generation process is conducted according to the transaction curve of load. The calculated discrete Fréchet distances between the RD curves are large, and do not meet the definition of curve similarity. The numbers of eigenvalue points on the two curves are similar, but there are differences in their time distribution. Therefore, it can be judged that the two curves have a certain similarity in terms of trend, but there is a significant difference between single points thereon.

The numerical calculation results show that the discrete Fréchet distance between the RT curves is 63.83, and that between the RD curves is 441.28. According to the distribution of eigenvalue points and the definition of curve similarity based on the discrete Fréchet distance, the RT curves are determined to be similar, that is, the electricity generation process is conducted according to the transaction curve of load. The calculated discrete Fréchet distances between the RD curves are large, and do not meet the definition of curve similarity. The numbers of eigenvalue points on the two curves are similar, but there are differences in their time distribution. Therefore, it can be judged that the two curves have a certain similarity in terms of trend, but there is a significant difference between single points thereon.
Fig. 6 The highest (lowest) points on the RT and RD curves

In conclusion, the COE-MSE and discrete Fréchet distance meet the demands for comparing similarity of transaction curves, but the latter can track an electricity generation process in real time, which is conducive to improving the level of refinement of the electricity generation process of a hydropower station. On this basis, through an ASP.NET platform, the judgment method, namely the discrete Fréchet distance was programmed in the C# language to form a function module for tracking and comparing transaction curves, which has been used as the main module of the quotation system in electricity spot trading.

5. Conclusion

The continuous deepening and development of electricity spot trading is the only way forward for electricity marketization. Traditional electricity generation according to dispatch orders is forced to follow declaration and clearing results in the markets, which imposes more onerous demands on the refined production and operation on the generation side. In addition to participating in the market transaction of electricity, the task of peak load and frequency regulation at large-scale hydropower stations lead to significant fluctuations of load for electricity generation enterprises. Furthermore, the uncertainty of climate, water influx, and load demands can easily result in the deviation of the actual electricity generation process from the transaction results in the markets, reducing profits accruing from electricity generation and affecting the stability of the power grid. Therefore, in the production process, the electricity generation side needs to track and compare the load in real time, adjust the load deviation in a timely manner, and strictly follow the market clearing results.

Based on the compositions of the load curves after a hydropower station participates in spot trading, this study analysed the causes and adverse effects of load deviations in electricity generation, and expounded the necessity of real-time tracking of electricity generation curves and similarity comparison with market clearing results of load. In this research, the Euclidean metric, COE-MSE, and discrete Fréchet distance were selected as the three main judgment methods for curve similarity to describe similarity definitions, calculation methods, and judgment principles. The actual electricity generation process at a hydropower station that has a reservoir with strong regulating capacity in the basin on a certain day and the transaction and declaration processes of load in real-time markets were taken as examples. Based on this, the similarity of TR and RD curves was judged by using the three methods based on interactive data acquisition from related systems and the reasons for judgment results were analysed. Moreover, the advantages and disadvantages of the three methods were summarised. The results demonstrate that the Euclidean metric can only preliminarily judge the similarity through numerical indices, however, the COE-MSE and discrete Fréchet distance can judge the similarity in terms of both trend and error, thus meeting the demand for tracking and comparing load curves during spot trading; the latter also offers more advantages in real-time curve tracking and calculation efficiency.
Acknowledgment
This work is supported by the National Key Research and Development Programme of China (Grant nos 2016YFC0402208 and 2018YFB0905204).

References
[1] Lee G J, Heydt G T. An interactive-dynamic mechanism conceptualizing the cost and benefit of electric power quality [J]. Electric Power Components and Systems, 2003, 31(9): 879-890.
[2] Xu Y-H, Xiao X-N. Power quality problems in deregulated power systems [J]. Power System Technology, 2004, 28(22): 48-52.
[3] C.-T. Cheng, J.-J. Shen, X.-Y. Wu et al., “Operation challenges for fast-growing China’s hydropower systems and respondent to energy saving and emission reduction,” Renewable and Sustainable Energy Reviews, vol. 16, no. 5, pp. 2386-2393, 2012.
[4] Liang X. Emerging Power Quality Challenges Due to Integration of Renewable Energy Sources [J]. IEEE Transactions on Industry Applications, 2017, 53(2): 855-866.
[5] S. Song, H. Guo, P. Yang et al., “Research on demand response strategy of the electricity deviation assessment,” The Journal of Engineering, vol. 2017, no. 13, pp. 2221-2225, 2017.
[6] S. C. Z. h. X. Qing, “Day-ahead electricity market design based on market interpretation of optimization theory,” Automatic of Electric Power System, vol. 2, pp. 56-22, 2016.
[7] Apostolopoulou D, McCulloch M. Optimal Short-Term Operation of a Cascaded Hydro-Solar Hybrid System: A Case Study in Kenya [J]. IEEE Transactions on Sustainable Energy, 2018, PP:1878-1889.
[8] YANG Hai-min, PAN Zhi-song, BAI Wei. Review of Time Series Prediction Methods [J]. Computer Science, 2019, 46(1): 21-28.
[9] A. Efrat, Q. Fan, and S. Venkatasubramanian, “Curve Matching, Time Warping, and Light Fields: New Algorithms for Computing Similarity between Curves,” Journal of Mathematical Imaging and Vision, vol. 27, no. 3, pp. 203-216, 2006.
[10] G. S. Xie, X. B. Jin, C. Yang et al., “Graphical Representation and Similarity Analysis of DNA Sequences Based on Trigonometric Functions,” Acta Biotheor, vol. 66, no. 2, pp. 113-133, Jun, 2018.
[11] R. Lin, B. Wu, and Y. Su, “An Adaptive Weighted Pearson Similarity Measurement Method for Load Curve Clustering,” Energies, vol. 11, no. 9, 2018.
[12] L. Holm, and C. Sander, “Mapping the protein universe,” Science, vol. 273, no. 5275, pp. 595-602, Aug, 1996.
[13] K. Bringmann, and W. Mulzer, “Approximability of the discrete Frechet distance,” Journal of Computational Geometry, vol. 7, no. 2, pp. 46-76, 2016.
[14] T. M. Chan, and Z. Rahmati, “An improved approximation algorithm for the discrete Fréchet distance,” Information Processing Letters, vol. 138, pp. 72-74, 2018.
[15] J. G. Alcázar, C. Hermoso, and G. Muntingh, “Similarity detection of rational space curves,” Journal of Symbolic Computation, vol. 85, pp. 4-24, 2018.