Reconstructing time series into a complex network to assess the evolution dynamics of the correlations among energy prices

1 Introduction

Reconstructing complex networks to detect dynamics in time series is an important issue in nonlinear science. Many researchers have proposed arithmetic models to reconstruct time series into complex networks based on pseudo-periodic time series [1], the visibility graph algorithm [2], horizontal visibility graph [3], parametric natural visibility graph [4], limited penetrable visibility graph [5], and multiscale limited penetrable horizontal visibility graph [6]. Phase space reconstruction, chaos, motifs and coarse graining methods have also been applied to establish time series-based complex networks to identify the time series types [7–12]. For bivariate and multivariate time series, researchers have constructed complex network models based on the correlations among variables for characterizing nonlinear time series [13–17].

These algorithms and methods have been applied in the fields of engineering, finance, economics, social science and biology [18–21]. Notably, in the economic system, energy plays an important role in economic development. Crude oil, natural gas and coal are the most important types of energy and account for more than 85% of global primary energy consumption according to the British Petroleum (BP) Statistical Review of World Energy 2017. Thus, the fluctuations in crude oil, natural gas and coal prices will impact production costs, transportation costs, price levels and other elements in economic systems. Moreover, because of substitution and game effects among crude oil, natural gas and coal, the relations among energy prices will result in fluctuations. Then, the impact...
of such fluctuations can be transmitted to the economic system and create economic risks. Hence, the evolution of the correlations among energy prices must be investigated to guide decision making.

Many studies have assessed the long-term equilibrium among crude oil, natural gas and coal prices based on econometrics method [22–24]. These studies also found that the relationships among energy prices are not always stable, i.e. the associated variables change over time and exhibit structural breaks, change points and outliers [25–27]. However, traditional econometric methods cannot detect specific nonlinear and dynamic variations. As shown in Figure 1, the correlation coefficient between crude oil price and coal price is 0.6865 (positive correlation) for the entire period (2008 to 2017). However, the correlation coefficient varies patterns in different periods, e.g., the correlation coefficient is −0.8378 from May 31, 2016, to August 10, 2016 (50 days), and 0.9699 from April 3, 2012, to June 20, 2012 (50 days). In this case, the correlation coefficients are extremely different, and the relationship reflects a strong negative correlation. The traditional econometric research paradigm can be applied to reveal the overall relationship among energy prices but cannot analyze the evolution of this correlation. Thus, decisions should not be made based only on the relationship over an entire period. Instead, the evolution of the correlations among different energy prices must be understood.

Figure 1: The daily closing prices of coal, crude oil and natural gas (coal price: USD/ton, crude oil price: USD/barrel, natural gas price: USD/million British thermal units)

Studies that reconstruct time series into complex networks also provide new findings for traditional econometrics. Current econometric methods, such as time series modeling, regression modeling and co-integration testing, are mainly utilized for long-term trend analysis and forecasting and may ignore some important short-term details. Most economic time series are nonlinear and unsteady. To obtain stable statistical results, most econometric methods transform unsteady time series into stable time series. However, stability processing methods often lose the nonlinear fluctuations in the time series. Thus, the processing methods used may not encompass the various correlations among time series variables.

Our previous studies focused on the autocorrelation and autoregression of energy prices [28, 29] and the correlation and regression between two energy prices [30, 31] by reconstructing networks based on energy price time series. We found that a temporal sliding window and coarse graining methods were able to reflect the evolution process of fluctuations, relationships and patterns. Moreover, the current literature mainly focuses on the relationship between two types of energy prices [32], and the relationship among three types of energy prices in one model as rarely been studied. Additionally, although studies of the evolution of the correlation among crude oil, natural gas and coal prices are new, they can help us understand the evolution of the energy price system, fluctuations in energy prices and the associated effects on economic system.

Investigating the evolution dynamics and the relationships among variables is a significant way to understand the variations and evolution mechanisms of the relationships among variables. Hence, in this paper, we divide an energy prices time series into a series of segments based on temporal sliding windows. The periods selected for evaluation include 50, 100, 150 and 200 days. The temporal sliding windows divides the three types of energy prices into a series of segments. Each segment contains a corresponding number of values (50, 100, 150 and 200). Then, we define the correlations among the crude oil, natural gas and coal price time series as modes in each segment. Because the division process is based on temporal sliding windows, a series of modes will form an evolution process over time. Thus, we reconstruct the complex network according to the evolution process. Next, we study the evolution dynamics of the energy price time series by analyzing the structures of the complex networks in different periods. The empirical study results provide a statistical-based guide for decision making in different periods (e.g. short-term decision making, medium-term decision making and long-term decision making).

2 Data and methods

In this study, empirical future price data for coal, natural gas and crude oil are used to investigate the evolution dynamics of the time-varying correlations among energy price time series. The sample data in this paper include the daily closing prices of the London International Petroleum Exchange (IPE) for Richard Bay coal and Brent oil and the
in the jth temporal sliding window. Similarly, let \( r_{13j} \) represents the Pearson correlation coefficient between \( x_{1j}, x_{1_{j+1}}, \ldots, x_{1_{\omega+j}} \) and \( x_{3j}, x_{3_{j+1}}, \ldots, x_{3_{\omega+j}} \) in the jth temporal sliding window, and let \( r_{23j} \) represent the Pearson correlation coefficient between \( x_{2j}, x_{2_{j+1}}, \ldots, x_{2_{\omega+j}} \) and \( x_{3j}, x_{3_{j+1}}, \ldots, x_{3_{\omega+j}} \) in the jth temporal sliding window. Hence, we can generate the following matrix.

\[
R = \begin{bmatrix}
    cr_{121} & r_{131} & r_{231} \\
    cr_{122} & r_{132} & r_{232} \\
    \vdots & \vdots & \vdots \\
    cr_{\omega_n-\omega+1} & r_{\omega_n-\omega+1} & r_{23\omega_n-\omega+1}
\end{bmatrix}
\] (1)

In matrix \( R \), the correlation coefficients in each row describe the state of the relationship between the three energy prices on the \( \omega \)th day (Figure 2).

(3) Defining the correlation mode based on a coarse graining process

Because each element in the matrix \( R \) can take a variety of values, it is not possible to directly build the network based on the matrix. Thus, a coarse graining process is applied in the phase space. The aim of this process is to map a group of dots in the phase space to a mode, so that the essential features of the data can be analyzed. Therefore, the correlation coefficients in matrix \( R \) are divided into five levels as follows.

\[
p = f(r) = \begin{cases}
    A, & r \in (0.6, 1] \\
    B, & r \in (0.2, 0.6] \\
    C, & r \in (-0.2, 0.2] \\
    D, & r \in (-0.6, -0.2] \\
    E, & r \in [-1, -0.6]
\end{cases}
\] (2)

Let \( r \) represents an element in the matrix \( R \), e.g., \( r_{12j}, r_{13j}, \) or \( r_{23j} \), where \( j = (1, 2, \ldots, n - \omega + 1) \).

The coarse graining method processes all the elements in matrix \( R \) according to Eq. (2) to obtain the following modes.

\[
mode_j = \{ p(r_{12j}), p(r_{13j}), p(r_{23j}) \}
\] (3)

The mode forms include “AAA”, “AAB”, “AAC”, “BCE”, “CBD”, etc. Each mode denotes a state that reflects the interactions among the three energy prices. For example, in mode “AAA”, strong positive correlations exist among the three types of energy prices. In addition, in mode “ACE”, a strong positive correlation exists between the natural gas and coal prices (A), a weak correlation exists between the natural gas and crude oil prices (C), and a strong negative correlation exists between the coal and crude oil prices (E).
Constructing the evolution networks based on the correlation modes

Different types of correlation modes are used as nodes and the subsequent relationships among the correlation modes are the associated edges. The weight of each edge is determined by the frequency of transmission between the two correlation modes. Then, a complex network based on the interactions among and evolutions of the three energy prices is developed.

Figure 3: The sensitivity analysis of the length of the temporal sliding windows

To study the evolution of the correlations among the three energy prices in different periods, we set different temporal sliding windows with various lengths. First, a sensitivity analysis is conducted for temporal sliding windows of different lengths. As shown in Figure 3, as the size of the temporal sliding window \( \omega \) increases, the number of nodes, the number of edges and average shortest path length in the corresponding networks decrease, and the density of networks increases (details in the supplementary Table S1). The nodes represent correlation modes and the edges represent transmission patterns between two modes. Obviously, the variations in the correlations among energy prices are hidden as the size of the temporal sliding window increases. Traditional econometric methods cannot detect the variations in correlation modes. Thus, we choose different window sizes \( \omega \) to provide details of the evolution mechanisms of correlations among energy prices. In this paper, we select 50, 100, 150 and 200 days as the different periods in the empirical study. Finally, we obtain 4 directed, weighted network models, as shown in Figure 4.

3 Results

3.1 Identification of major correlation modes

We consider the correlation modes that frequently appear as major modes, which can be used to represent the major trends of the relationships among coal, crude oil and natural gas prices in a variety of periods associated with different sliding window sizes. Major modes can be identified via the weighted outdegree of nodes, and we defined \( w_{out}^{\text{ij}} \) as follows:

\[
w_{out}^{\text{ij}} = \sum_{j \in N_i} w_{ij}
\]

where \( N_i \) represents the total number of neighboring nodes for node \( i \) and \( w_{ij} \) is the weight of the edge from \( i \) to \( j \). The greater the value of the weighted degree of one node is, the more important the node in the evolution network. In Figure 4, we use the sizes of nodes to represent the weighted degrees of the nodes. Large nodes have higher weighted outdegrees than small nodes. Moreover, the number of major modes varies in different periods.

Theoretically, for different periods associated with different temporal sliding window sizes, there should be \( 5 \times 5 \times 5 = 125 \) types of modes. Practically, there are 80, 70, 56 and 51 modes when the sliding window sizes are 50, 100, 150 and 200, respectively. According to the cumulative percentage distribution displayed in Figure 5, only 40% of the total modes represents 70% of the market information. In fact, from a complete time series perspective, the correlation mode of the three energy prices is “BBA” (the correlation coefficient between the natural gas and coal prices is 0.35, the correlation coefficient between the natural gas and crude oil prices is 0.32, and the correlation coefficient between the coal and crude oil prices is 0.69). However, when \( \omega = 50 \), the occurrence probability of mode “BBA” is only 6.18%; when \( \omega = 100 \) and \( \omega = 150 \), the modes with
the two highest occurrence probabilities are “AAA” and “ABB”; and when $\omega = 200$, the modes with the two highest occurrence probabilities are “AAA” and “ABA”. This result suggests that the correlations among energy prices do not always reflect the “BBA” correlation mode and that different periods have different major correlation modes.

We list the top five correlation modes ranked by the weighted outdegree for four different periods associated with four temporal sliding window sizes (details in the supplementary Table S2). Note that the top five modes of the relationship among the three variables are generally highly related, e.g., “AAA”, “ABB” and “BBA”. As the sliding window size increases, the correlation among energy prices in the top five modes increases. Thus, the relationship among energy prices becomes stronger from the short to long term. For example, for $\omega = 50$, the fourth mode is “DDB”, which means the three energy prices are negatively correlated. Therefore, different statistical results should be considered when making decisions in different periods (short, medium or long term).

### 3.2 Analysis of the evolution dynamics of correlation modes

(1) The distribution characteristics of the transmission pattern

We identify the transmission pattern based on the edge weights of the two modes. From the traditional econometric perspective, econometric models focus on the long-term correlation among energy prices, which may ignore the short-term evolution of the correlation modes. The transmission pattern explains the relationships among modes. The transmission between mode $i$ and mode $j$ can be expressed as $mode_i \rightarrow mode_j$. Some modes may self-transmit, such as “AAB → AAB”, and most transmissions occur between two different modes, such as “AAB → BCD” or “BCD → AAB”. Additionally, the weight of an edge represents the number of times a transmission occurs between the two modes.

Figure 6 shows the cumulative distribution of the weights of modes. Notably, 20% of mode transmissions account for approximately 80% of the total weight. Thus, the key transmission patterns are not abundant, although hundreds of modes and transmission patterns exist. In this case, only a few types of transmission patterns play important roles in the evolution.

(2) The cluster effect in the evolution of correlation modes

The evolution process of correlation modes among energy prices exhibits a cluster effect. There are some cor-

---

Figure 5: The cumulative percentage distribution over four periods

Figure 6: The cumulative distribution of modes transmission at different time periods

Figure 7: Volatility cluster network diagrams at different time periods
relation modes in the network that frequently transmit to each other during the evolution process, and they are hardly transmit to other correlation modes. These correlation modes form a cluster in the network. Typically, there are several clusters in a network. From a time series perspective, the evolution process of correlation of three energy prices usually last for a given time in a cluster and then transmit into another cluster. A cluster reflects one of status of the energy market, and the relationships associated with the prices fluctuation in the three types of important primary energy last for different periods. Determining the duration of the existence of each cluster and the evolution trends among clusters can benefit the monitoring of energy price fluctuations and avoidance of energy market risk. In this paper, we divide the network into a series of clusters using the cluster division method proposed by Blondel [33]. In Figure 7, we show the different results of network division when the period is $\omega = 50, 100, 150$ or 200. Under certain conditions, the networks are divided into 6 clusters, and we replace each cluster with a node in Figure 7.

As shown in Figure 7, each node represents a cluster. The size of a node reflects the number of correlation modes contained in the cluster. In addition, the direction of the edge reflects the path of transmission between clusters, and the thickness of the edge indicates the probability of transmission between clusters. Figure 7 shows that there is no relation between some clusters; therefore, it is less probable for these clusters to directly interact (although indirect interactions through other clusters may occur).

This is an important and interesting finding, and the evolution exhibits rapid jumps due to fluctuations in the energy market. For example, in Figure 7, nodes with different colors represent different energy market statuses. When $\omega = 100$, if the energy market status is represented by a purple cluster, it can easily be transformed to a yellow cluster, but it is difficult to transform this cluster to a green cluster. Thus, the cluster effect results provide important information regarding market fluctuations that may occur.

(3) The duration of the cluster effect

The transmission relationship among the clusters is characterized by jumps. Figure 8 clearly shows the evolution characteristics of a cluster over time. In Figure 8, the different colors represent different clusters, and each cluster lasts for a specific period before a transformation occurs and new clusters form. The colors do not gradually vary but instead suddenly change. The evolution of the correlation modes has a more significant effect on clusters than was noted in previous studies [28, 31]. Additionally, this phenomenon becomes more distinct as the size of the time window increases. In some cases, a small color block is inserted into a larger block. In this scenario, a short-period fluctuation in energy prices occurs after a long period of relative stability in the energy market, and this fluctuation is followed by another relatively stable long period. Additionally, a strong relationship between the energy prices is clearly observed, and this relationship results in similar trends for the three energy prices.

Figure 8: Analysis of the fluctuating cluster effect over time at different time periods

Figure 9: Transmission time distribution of mode values in clusters at different time periods

Figure 9 shows the amount of time required for correlation modes to transform from inner clusters to alternative clusters. Based on the statistical results (the cumulative percentage plots in Figure 9), we find that approximately 30 days is a turning point at which one type of energy market status transforms into another status. The transmission among all modes in the cluster will not last
more than 110 days at different time scales. For example, when $\omega = 50$, the energy market status will not last more than 70 days. When $\omega = 100$, the temporal distribution is relatively uniform. Moreover, when $\omega = 150$ and $\omega = 200$, the duration of the transmission is only 30 days or less. These results can be used to understand changes in market patterns.

### 3.3 Media capability during market evolution

As noted above, one cluster will transform into other clusters. Therefore, the status of the energy market will change from one pattern to another. The understanding the transmission process is important for understanding the energy market and can provide important information for decision making. In this situation, the changes in energy market status are also important. During the evolution of correlation modes, different correlation modes have various roles. Some modes may act as a hub or bridge, and another may be a key point in the overall evolution network. Without those correlation modes working as hubs, the network may separate into different clusters. Thus, the modes on the shortest paths between modes or clusters tend to be the hubs in the network and have high media capabilities.

We measure the media capability associated with the evolution of correlation modes by calculating the betweenness centrality (details in the supplementary Table S2). Figure 10 shows that some nodes exhibit high media capability; therefore, when the status of the energy market changes, the transmission may pass through these nodes (correlation modes). Currently, the betweenness centrality calculation is based on the shortest path in the entire network. However, we should consider the weighted outdegree of the nodes in other clusters to measure the media capability between clusters. Thus, the betweenness centrality algorithm is extended to identify the media nodes in evolution networks. The improved betweenness centrality $BC_i$ algorithm is as follows:

\[
BC_i = \sum_{j}^{n} \sum_{k}^{n} \frac{s_{jk}(i)}{s_{jk}} \sum_{l}^{w_{l}} w_{l}^{\text{out}},
\]

\[
i \neq j \neq k \neq l, j < k, i \in A, l \in B, A \cap B = \emptyset
\]

where $s_{jk}(i)$ is the number of shortest paths between node $j$ and node $k$th at pass through node $i$; $s_{jk}$ is the total number of shortest paths between node $j$ and node $k$; $A$ and $B$ are different clusters; and $w_{l}^{\text{out}}$ is the weight of the edge from node $i$ to node $l$. Nodes $i$ and $l$ do not belong to the same cluster. Thus, a node with higher betweenness centrality has better media capability.

![Figure 10: Transmission media modes in different periods](image1)

![Figure 11: Transmission possibilities between clusters at different time periods](image2)
a very small probability that Cluster 6 transforms to Cluster 4 or Cluster 5.

Hence, in actual decision making, we should first determine what type of cluster the current energy market status is associated with. If the cluster is relatively stable, the duration of the new status will be long; otherwise, it will be short. When the duration is over, the clusters that are most and least likely to be transmitted can be determined based on the proposed approach. These results can provide guidance for the establishment of relevant energy price policies.

### 4 Conclusion

In this paper, we used the closing prices of natural gas, coal and crude oil in the futures market as time series of empirical data. Then, we divided the time series of the three variables into a series of segments by applying temporal sliding windows. After implementing a coarse graining method, the correlation coefficients of energy prices in each segment were defined as modes. Finally, the energy price time series were reconstructed into a directed and weighted network in which nodes are correlation modes and edges are the time-ordered relationships between correlation modes.

We set four different periods in the temporal sliding windows to study the samples in this paper. The temporal sliding windows could be set to any length (short, medium or longterm) and were used to study the evolution dynamics of correlations at different time scales. When $\omega = 50$, the modes of the three energy prices exhibited a positive correlation in general, but negative correlations were also observed. When $\omega$ increased, i.e. $\omega = 100$, 150 or 200, the ratio between the high positive correlation modes gradually increased. The interrelationships among the closing prices of the three types of energy are significant in the longterm and delayed in the short term.

Compared with traditional econometric methods, the reconstruction of time series into a complex network in this study provides some new findings. Notably, based on this method, the short-term evolution of the relationships among variables can be obtained. We show that variations in the correlation among energy prices are hidden as the size of the temporal sliding window increases based on sensitivity analysis. Thus, economic and management policies cannot be solely based on the results of an entire long-term period, as considered in traditional econometric models.

A few major correlation modes and transmission patterns play important roles in the evolution of correlations among energy prices. These major correlation modes and transmission patterns represent the main fluctuations in the energy market. In addition, we find that the major correlation modes in different periods differ from the correlation modes over the entire period of analysis. Additionally, different statistical results should be considered when making decisions for different periods (e.g., short, medium or longterm).

Moreover, the evolution of the correlation modes among energy prices exhibits a significant cluster effect that reflects different patterns in the energy price market. The energy price market rapidly evolves and is characterized by jumps, and each cluster will last for a period then transform into another cluster. Based on the distribution of the duration of the cluster effect at different time scales, we find that approximately 30 days is a turning point at which one type of energy market status transforms into another status.

Then, we improve the betweenness centrality algorithm to measure the media capability of correlation modes in the evolution process for different clusters. The correlation modes with high media capability can be used to determine whether the current energy market is in a transitional period. Moreover, based on the transmission probability between clusters, we can assess the evolution direction of the energy market. These results are useful for analyzing the fluctuations in the energy market and decision making.

Furthermore, we only consider three energy price time series in this paper. However, there are many variables in

![](https://example.com/table1.png)
real-world systems, such as economic, social, and financial factors. If more variables are present, the correlation mode will be long and complex. We only consider the correlation among energy prices, and there is not direction of causal relationship for energy prices. In a future study, we will try to consider more time series variables, the impact direction and the impact extent into the model to better understand the real-world systems.

Acknowledgement: This research is supported by the Fundamental Research Funds for the Central Universities and the Project of Humanities and Social Sciences of the Ministry of Education in China (No. 12YJC630043). The authors would like to express their gratitude to Prof. Haizhong An, who provided valuable suggestions.

References