

# Northumbria Research Link

Citation: He, Xuan, Hussain Shah, Syed Bilal, Wei, Bo, Liu, Zheng and Wang, Wei (2021) Comparison and Analysis of Network Construction Methods for Seismicity Based on Complex Networks. Complexity, 2021. p. 6691880. ISSN 1076-2787

Published by: Hindawi

URL: <https://doi.org/10.1155/2021/6691880> <<https://doi.org/10.1155/2021/6691880>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/45407/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria**  
**University**  
NEWCASTLE

## Research Article

# Comparison and Analysis of Network Construction Methods for Seismicity Based on Complex Networks

Xuan He <sup>1</sup>, Syed Bilal Hussain Shah,<sup>2</sup> Bo Wei <sup>3</sup> and Zheng Liu<sup>4</sup>

<sup>1</sup>College of Medicine & Biological Information Engineering, Northeastern University, Shenyang 110169, China

<sup>2</sup>School of Software, Dalian University of Technology, Dalian 116024, China

<sup>3</sup>Department of Computer and Information Sciences, Northumbria University, Newcastle Upon Tyne NE1 8ST, UK

<sup>4</sup>School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China

Correspondence should be addressed to Xuan He; [hexuan@bmie.neu.edu.cn](mailto:hexuan@bmie.neu.edu.cn)

Received 22 December 2020; Revised 12 January 2021; Accepted 20 January 2021; Published 2 February 2021

Academic Editor: Wei Wang

Copyright © 2021 Xuan He et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The approach of the complex network has well described seismic complex systems. In this paper, this is the first time three classical network construction methods for seismicity are compared. By using the same dataset from the Southern California Seismic Network, three networks are constructed. They all present the scale-free, small-world properties, a strength-degree correlation, and an assortative mixing feature. However, they show some differences in the hierarchical clustering feature. On observing the evolution results, three measures show a similar correlation with seismicity dynamics, but one measure shows a different result. These results show that different network construction methods will present some similarities and differences in network properties. This situation needs to be considered, especially when discussing a predictive indicator of seismicity.

## 1. Introduction

Network science is widely used in many fields in the real world to describe complex systems' characteristics. In order to represent a complex system as a graph, nodes are usually used to represent research objects, while edges represent the relationships between research objects. Scientists represent complex systems as graphs from different perspectives in various fields, such as brain networks [1], protein-protein networks [2], social networks [3], Internet topology [4], and transportation networks [5–7]. Complex networks prove to be an effective method to study the complex system.

Due to some unknown dynamics of the earth's crust, seismic activity has been proven to be a complex system with temporal and spatial characteristics [8]. Recently, seismic complex systems have been described by the approach of the complex network [9]. The most significant advantage is that we no longer study seismic activity from some small local areas or study one big shock but consider the relationships between seismic events from a broader geographical scope. Most of the proposed methods [9–13] can construct a

complex earthquake network only from the main elements of magnitude, time, and location and have achieved precious results. They discovered that the earthquake network is scale-free and small-world. They also discovered that the networks' topological characteristics change over time, corresponding to the large earthquakes [14–16]. Resaei et al. [17] found that the PageRank value is an appropriate alarming clue before the event's occurrence, which is worthwhile in hazard probabilistic evaluation of earthquakes. However, the discovery of these laws is based on different network construction methods. Our recent study found that the conclusions drawn by different network construction methods will be different to some extent. As far as we know, no researchers have compared and analyzed these differences.

Abe and Suzuki proposed the earthquake network construction method for the first time in 2004 [9]. They first divided the geographical region into many small equal-sized cells. If any event occurred in the cell, the cell is represented by a node in the network. Two successive events defined an edge from the former to the latter between two nodes. In this

way, the complex event-event correlation was represented by the edge. Many studies were based on this method. We called it a time-series (TS) construction method in this paper. But researchers proposed many network construction methods for seismicity from a different point of views. Douglas [13] proposed a time window-based (TW) method to construct a global earthquake network. Xuan [11] proposed a space-time influence domain (STID) method to construct an earthquake network considering the time, location, and magnitude of the earthquakes. Other researchers proposed other construction methods, such as similar activity patterns [12], mutual information [10], and hybrid model [18], etc.

The time-series (TS) construction method, time window-based (TW) method, and space-time influence domain (STID) are all based on the temporal and spatial characteristics of seismicity and only consider the factors of time, location, and magnitude. The TS method constructs networks according to events' time points. The TW method constructs networks according to a time window, which includes many time points. The STID method constructs networks according to both a time window and a space window. Therefore, these three methods are similar to some extent. Therefore, we will compare and analyze the topological characteristics of these three methods in this paper.

The rest of this paper is organized as follows. In Section 2, the three construction methods are introduced in detail. The data used in our study and the results are presented in Section 3. Finally, in Sections 4 and 5, the discussions and conclusions of this paper are presented.

## 2. Construction Methods

As previously mentioned, we will introduce three different construction methods in detail in this section. For convenience, we will use abbreviated names TS, TW, and STID instead of a time-series method, time window-based method, and space-time influence domain. The first step before constructing the network is to make possible nodes in the network to be built. The same step of these three methods is to divide the geographical region into equal square cells of side  $SC \times SC$ , where a cell will become a node of the network every time it is located therein.

*2.1. Time-Series Method.* Through the nonextensive statistical mechanism, Abe and Suzuki et al. [9] studied the space and time interval of successive seismic events. They found that two successive seismic events are indivisibly related to each other, regardless of their distance. So, Abe and Suzuki proposed an earthquake network construction method based on time series. In their work, nodes representing the grid cells geographically were linked when successive earthquakes occurred in them.

As shown in Figure 1(a), we give a schematic diagram of the network construction method based on time series. We suppose that the events occurred in a particular sequence, which is shown at the top of the diagram. Then, we can see that each node in the graph has only one edge pointing to itself and one edge pointing to other nodes except for the

node, which has the event that occurs at last. So, there will not be isolated nodes in the networks constructed by this method.

*2.2. Time Window Method.* In the time window method, Douglas et al. [13] proposed a time window. In this time window, the node where the first event occurred will be connected to all nodes within this window by directed edges but respecting the time order of events. Then, the time window moves forward and then restarts the above steps until all events are traversed.

To illustrate this process, we show a diagram in Figure 1(b). We suppose that a time window of value equal to  $T$  has been adopted in a given dataset of seismic events. The first window is  $w_1$ , which has four events (A, B, C, and D), so A has a connection to B, C, and D, respectively. Then, we move forward to the time window. The second window is  $w_2$ , which also has four events (B, C, D, and E). B has a connection to C, D, and E, respectively. After recursively processing the raw data, the network constructed by the time window is generated.

Douglas et al. found from the results of variation of the number of communities in the earthquake network with the size of the time window that small windows will cause the network of earthquakes to be very fragmented and consequently with few communities, and large windows will produce massive clusters also causing a decrease in the number of communities. So, they conclude that the ideal window size for the global earthquake networks is  $T = 3000$  s and for California is  $T = 2000$  s.

*2.3. Spatial-Temporal Influence Domain Method.* He et al. [11] generally divide the cause-and-effect relationships between earthquakes into two categories: a direct effect correlation and an indirect effect correlation. For the method of time series, it is an indirect effect correlation. It considers some pairs of earthquakes, which are far away from each other, exhibit consistent correlations. We can also call this category "remote triggering." However, if we want to analyze the specific cause-and-effect relationship, it is usually inexplicable. So He et al. think that earthquakes tend to occur near the epicenter and shortly after the mainshock. They suppose that a larger earthquake will release more energy and trigger a greater number of earthquakes. They call this a direct influence correlation.

Moreover, Gardner et al. [19] found a reasonable approximation to formulate the relationship between the distance interval  $R$  and the magnitude  $M$  of an earthquake as

$$\log_{10} R = a_1 M + b_1, \quad (1)$$

where  $a_1$  and  $b_1$  are constants. It indicates that the influence radius grows exponentially with the magnitude. Similarly, a relationship between the time interval  $T$  and the magnitude  $M$  of an earthquake is expressed as

$$\log_{10} T = a_2 M + b_2, \quad (2)$$

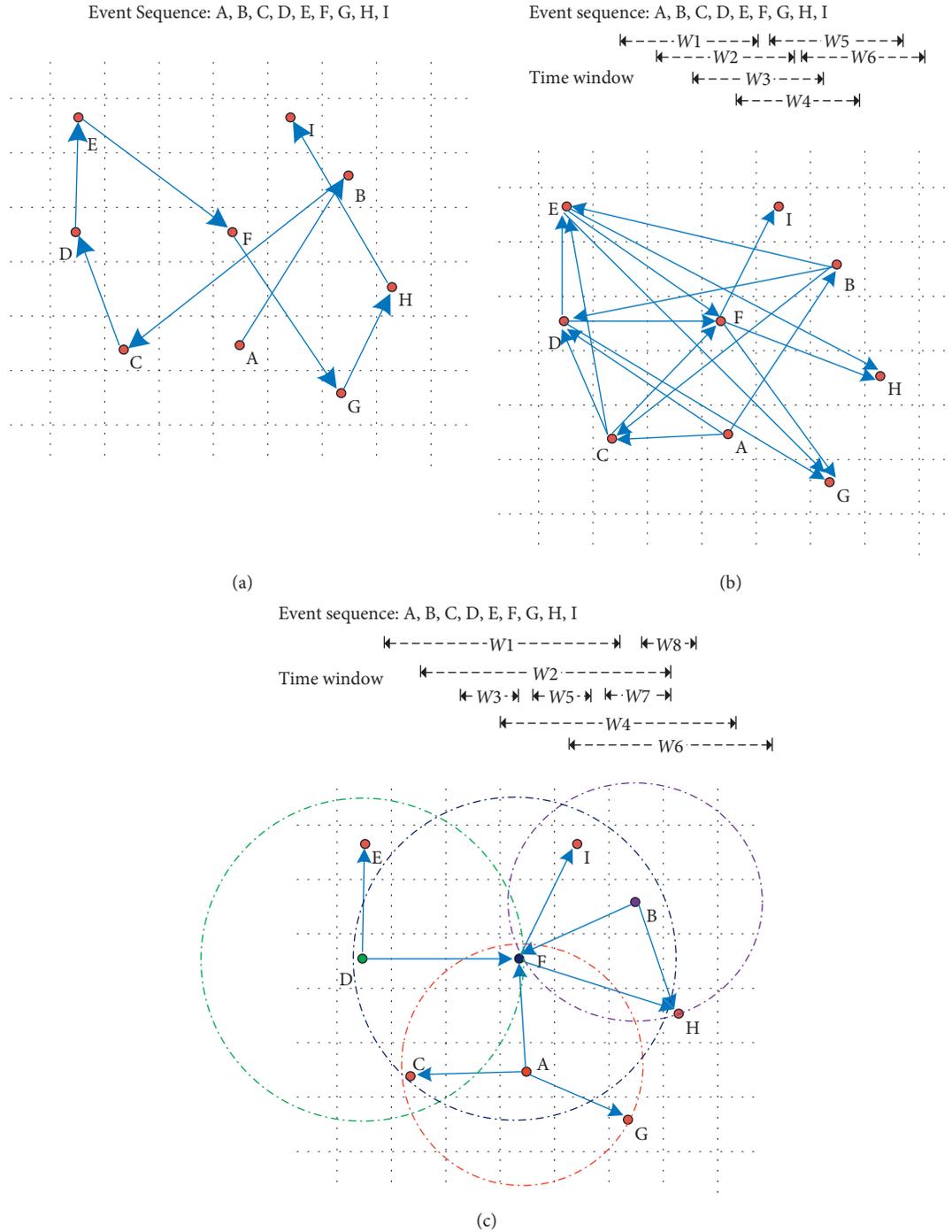


FIGURE 1: The schematic diagram of network construction methods. The time windows are represented by  $w_i$ , and  $i$  is the window number. Event sequence represents the sequence of the events occurring in the example. (a) Time series, (b) time window, and (c) space-time influence domain. The dashed colored circles in (c) represent the node's space influence domain with the same color.

where  $a_2$  and  $b_2$  are constants as above. For instance, if a magnitude 6.5 earthquake occurred, it will influence the region that less than 180 km away from it within 400 days. So, if another earthquake happened in this domain in both space and time, we define an edge from the magnitude 6.5 earthquake to this earthquake. In order to obtain the values of  $a_1$ ,  $b_1$ ,  $a_2$ , and  $b_2$  in equations (1) and (2), we fit the relation curve using the specific influence range from [19].

According to the fitting result, the values of  $a_1$ ,  $b_1$ ,  $a_2$ , and  $b_2$  are 0.3456, 0.0329, 0.0218, and 0.1471, respectively.

To illustrate this process, we show a diagram in Figure 1(c). We suppose that a time window and a space window have been adopted in a given dataset of seismic events. The size of space and time window is calculated by formulae (1) and (2) according to the magnitude of the event. The first window is  $w_1$ , which has seven events (A, B,

C, D, E, F, and G). At this time, we will not directly connect A with the other nodes. We should compare both space and time influence domains. The dashed colored circles in Figure 1(c) represent the space influence domain of the node with the same color. If the other nodes in this time window are also in the space influence domain of A, then A will connect to this node. For example, in the space influence domain of A, there are three nodes of events C, F, and G. So, A has three connections with these nodes. Although B is in the time influence domain or time window of A, it is not within the space influence domain of A. So, A will not connect with B at this time.

Through the brief illustration, we can see that these three methods are similar to some extent. TS method only considers successive time points, while the TW method considers a time window. Specifically, this method takes into account the influence of time. Then, the STID method considers both the impact of time and space. In the next section, we will compare and analyze the topological characteristics of these three methods.

### 3. Results and Discussion

Earthquakes in California, Japan, and New Zealand are the most frequent earthquakes in the world. The seismic data in these areas has much useful information and provides continuous support for seismic research. California is located in the Pacific Ring of Fire, with frequent earthquakes and volcanic activity. Many scientists have completed their study and achieve some valuable results by the seismic activity data in California.

Therefore, we choose the earthquake dataset from the Southern California Earthquake Data Center (SCEDC) [20] covering the region  $32^{\circ}\text{N} - 36^{\circ}\text{N}$  latitude and  $114^{\circ}\text{W} - 122^{\circ}\text{W}$  longitude in the period between January 1, 1990, and December 31, 2009. The total number of events with arbitrary values of magnitude in this region is 335633. The size of each cell is set to be 10 km.

**3.1. General Information.** Using the same dataset and setting the same parameters, three networks constructed by different methods are generated. First, we compare the general network information of the networks. Let  $G = (V, E)$  be a graph with  $|V| = n$  nodes and  $|E| = e$  edges. As shown in Table 1, we find that the numbers of nodes of TS and TW are approximate. However, the number of nodes of STID has only 1306 nodes in the network. For the number of edges, we observe similar results as above. The number of edges of TS and TW is more than ten times that of STID. For the average degree of the networks, TS and TW present a more clustering network while STID does not present the property as the above two methods. It is a very interesting phenomenon. Although these three methods all consider the space and time complexity of seismicity, they present different clustering degrees of networks. This may be because STID has two limitations at the same time. If an earthquake has a smaller magnitude, it may not influence many nodes. Therefore, there can be some isolated nodes due to the

construction process. This phenomenon also has a presence in the network of TW. However, because the time limitation still will produce many connections, it seems that TW does not present a pronounced sparse situation as STID.

**3.2. Scale-Free.** The scale-free characteristic points out that the degree distribution of nodes in the network satisfies the power-law distribution, namely,  $P(k) \sim k^{-\alpha}$ .  $P(k)$  is the degree distribution of network nodes. That is the probability that a randomly selected node has degree  $k$ . This feature is entirely different from random graphs. The degree distribution of random graphs is usually binomial distribution or Poisson's distribution, and the probability will decay rapidly when  $k$  is more significant than a certain threshold. However, the degree distribution of network nodes satisfying the power-law distribution has a long tail phenomenon. That is, there is no apparent attenuation threshold. If we draw  $P(k)$  on the double logarithmic axis, it will appear as an approximately straight line because  $\log P(k) \sim \alpha \log k$ . Figure 2 shows the degree distribution of the networks constructed by the above three methods. For comparison, we plot them in the same figure. We found that both the TW method and the STID method have scale-free characteristics after calculation. The TS method has some fluctuations when the degree value is less than 10, showing a scale-free characteristic with truncation. This result is consistent with the former studies [9, 11].

**3.3. Small-World.** Watts and Strogatz provided the term "small-world" for the complex network. The average path length  $L$  of the network is at the level of the characteristic path length  $L_r$  of the random network,  $\lambda = L/L_r \approx 1$ . But the clustering coefficient  $C$  is much higher than the random network  $C_r$ ,  $\gamma = C/C_r \gg 1$ . The small-world characteristic is quantified by the so-called small-world index  $S$  defined as

$$S = \frac{\gamma}{\lambda} = \frac{C/C_r}{L/L_r}. \quad (3)$$

When the value of  $S$  is much greater than 1, it indicates that there is a small-world characteristic. Through calculation, the network metrics  $C_r$  and  $L_r$  are, respectively, derived from a set of random networks with the same network scale.

The results are shown in Table 2. The values of  $S$  of the three networks are all much more significant than 1, indicating that the networks constructed by these three methods all have a small-world characteristic. The results are consistent with the former studies [11, 21]. Compared with TS and TW, STID presents a more sparse network as analyzed above, and it still presents a small-world characteristic.

**3.4.  $k$ -Core Decomposition.**  $k$ -core [22, 23] is one of the centrality indices of a node in complex networks. Based on a recursive pruning of the least connected nodes, it allows us to disentangle the hierarchical structure of networks by progressively focusing on their central cores. The  $k$ -core decomposition was recently applied to several real-world networks, such as the Internet [24] and software networks

[25]. It was turned out to be an essential tool for the visualization of complex networks and latent comprehension relations in the structure. In this paper, we use  $k$ -core decomposition to analyze the complex network structure and aim to find the core of the network.

Some basic definitions of  $k$ -core [26] are as follows.

*Definition 1.* A subgraph  $H = (C, E|C)$  induced by the set  $C \subseteq V$  is a  $k$ -core or a core of order  $k$  iff  $\forall v \in C$ : degree  $(v) \geq k$ , and  $H$  is the maximum subgraph with this property.

Therefore, a  $k$ -core of  $G$  can be obtained by recursively removing all the nodes of degree less than  $k$ , until all the nodes in the remaining graph have a degree at least  $k$ . Also, we will use the following definitions.

*Definition 2.* A node  $V_i$  has coreness  $k$  if it belongs to the  $k$ -core but not to  $(k+1)$ -core.

*Definition 3.* A  $k$ -shell  $S_k$  is composed of all the nodes whose coreness is  $k$ . The maximum value  $k$  such that  $S_k$  is not empty is denoted by  $k_{\max}$ . Thus, the  $k$ -core is the union of all shells  $S_c$  with  $c \geq k$ .

He et al. [27] first use  $k$ -core decomposition in the field of earthquake networks. They observed the evolution of the maximum coreness and found that it has some sudden changes with significant shocks. The highest core with high clustering feature tends to be located in the area that can directly or indirectly cause the major shocks.

In this work, we use  $k$ -core decomposition to analyze the clustering feature of the three networks. The coreness represents the hierarchical structure of the network. The greater the value of coreness is, the more hierarchies the network has. By comparing the result in Table 3, we found that TW has the greatest coreness while STID has the smallest coreness. To see the differences clearly, we visualize the topological graphs by Gephi [28]. The red dots in Figure 3 represent all the nodes in the highest core, while the green dots represent the other nodes in the network. There are noticeable differences. The highest core of the STID network has a more clustering feature. Moreover, the Landers earthquake (M 7.3) at 11:57:34.13 on June 28, 1992, has occurred in the region where all the nodes with the highest coreness located. However, the highest core of TS and TW shows a broader distributed feature.

**3.5. Evolution.** In previous studies, researchers usually study the change of network topology characteristics over time. This part will compare the similarities and differences of the topological characteristics calculated by the three network construction methods. Here, we need to use a correlation measure. This measure must reflect the dynamics of seismic activity. By calculating the correlation between this measure and the characteristics obtained by the three methods, we can see which method better describes seismic activity dynamics. In some papers, they use the total number of earthquake events. However, the network construction method is related to the number of earthquake events. These two metrics do not exist independently. We decided to use

TABLE 1: General network information of the three networks.

Networks	$n$	$e$	$d$
TS	2097	70412	67.155
TW	2048	90270	88.154
STID	1306	5633	8.626

$n$  represents the number of nodes;  $e$  represents the number of edges;  $d$  represents the average degree of the networks.

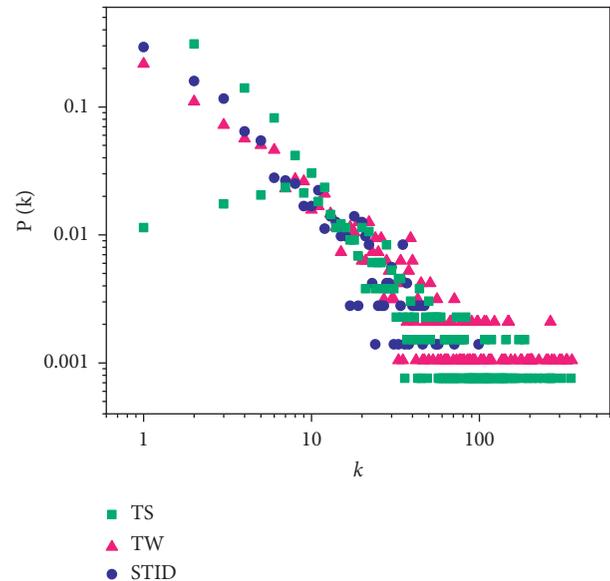


FIGURE 2: Degree distribution comparisons of TS (green dots), TW (pink dots), and STID (blue dots).

TABLE 2: The average path length, clustering coefficient, and small-worldness index of the networks.  $L$  represents the average path length while  $L_r$  in the brackets represents that of a random network.  $C$  represents the average clustering coefficient while  $C_r$  in the brackets represent that of a random network.  $S$  is the small-worldness index.

Networks	$L (L_r)$	$C (C_r)$	$S$
TS	2.459 (3.252)	0.657 (0.009)	96.54
TW	2.106 (2.454)	0.819 (0.026)	36.70
STID	4.834 (5.020)	0.580 (0.005)	120.46

the sum of earthquake magnitudes [29] as an essential measure to compare the similarities and differences of the network construction methods. We will compare the network characteristics number of nodes, average degree, betweenness, coreness, and entropy of the three networks. Since these measures' values have different orders of magnitudes, the log of these measures is used to distinguish better the differences [17]. In this regard, we first obtained the log of these measures and divided them to the most considerable log value for normalization, respectively.

First of all, we divide the above data between January 1, 1990, and December 31, 2009, into 20 data segments. The time scope of one segment is one year. Segment data is used to construct a network. Then, we obtain twenty values of each characteristic and observe the change over time. The

TABLE 3: The coreness of earthquake networks constructed by TS, TW, and STID methods.

Networks	coreness
TS	45
TW	81
STID	20

results are shown in Figures 4–6.  $\rho_{TS}$ ,  $\rho_{TW}$ , and  $\rho_{STID}$  represent the Pearson correlation between the characteristic and sum of earthquake magnitudes for TS, TW, and STID methods, respectively. The larger the value of  $\rho$  is, the better the methodology describes the dynamics of seismic activity.

First, we can see from Figures 4–6 that all the characteristics in the year 1992 present the most considerable value as a peak. On the one hand, the Landers earthquake (M 7.3) has occurred at 11:57:34.13 on June 28, 1992. It is the largest earthquake of these years. On the other hand, we obtained the log of these measures and divided them to the most considerable log value for normalization. Therefore, all the measures overlap at the point of the year 1992.

From the number of nodes (shown in Figure 4), the three networks have approximate values. They all have a reasonable correlation with the dynamics of seismic activity. STID has the highest value of  $\rho$ . In Figure 5, we can see the results of average degree evolution. TW and STID both present a higher correlation, implying that these methods can reflect seismic activities' dynamics. Moreover, TS also shows a positive correlation. It indicates that the average degree of all these three methods is a proper measure of seismicity.

Betweenness centrality is the number of shortest paths that pass through the nodes. A node with higher betweenness centrality would have more control over the network in an earthquake network because more energy will pass through the node. We can see from Figure 6 that the STID method is based on the influence of an earthquake. This may be the reason why it presents a better correlation with the sum of earthquake magnitudes. We believe this is an interesting result. We will further study the betweenness of a specific node.

**3.6. Weighted Properties.** On observation of the networks, events may occur between the same pair of nodes, which implies that multiple edges may exist between them. We believe that the main difference in the networks is derived from the multiple edges due to the construction process. So, in this section, we will analyze the weighted properties of the networks. Weighted network representation allows the consideration of the dynamics of traffic occurring on the network. Here, the weight  $w$  indicates the strength of the link between nodes. It is defined, associated with an edge, as the total number of multiple edges between its two end nodes. The greater the weight  $w$  is, the more active the interaction in the system is. A weighted earthquake network will add another dimension to the description of the seismicity. In this study, self-loops attached to a single node should be removed.

The nodal strength distributions for the three methods are plotted in Figure 7(a). We can see that a power-law function of the form  $P(s) \sim S^{-\eta}$  has been observed to fit well for these three methods. The nodal strength distribution of STID shows the heavy tail as the weight distribution. This is an interesting phenomenon. In the form of a simple network, the STID network presents a more sparse network compared with TS and TW. However, in the form of a weighted network, it is evident that the STID network has many nodes with greater nodal strength. It implies that the STID network has many multiple edges between the same pair of nodes, which will disappear when seen as a simple network.

In Figure 7(b), the average nodal strengths of all  $k$  degree nodes  $\langle s(k) \rangle$  are plotted against  $k$ . It has been observed that the average strength scales almost linearly for TS and TW networks exhibiting strength-degree correlation. In a weighted earthquake network, the nodal strength represents the total volume of interactions between the node and its neighbors. This indicates that the large degree nodes have a large number of edges and have a large number of weights. For the STID network, this result is not a strict linear correlation. When  $k$  is greater than 20, the average strength seems not to exhibit linearly to the degree due to its fluctuations, which implies that a node with a large degree is not necessarily with large weights for the STID network.

In Figure 7(c), we plot the weighted average neighbor degree  $\langle k_{nn}^w \rangle$  with degree  $k$ . Here, we address the question as to whether the preferential attachment exists in the three networks. If  $\langle k_{nn}^w \rangle$  decays with  $k$ , it indicates a disassortative mixing of the network. Otherwise, it means an assortative mixing of the network. We can see that the STID network presents a noticeable assortative mixing feature. It suggests that nodes with significant values of degree tend to be connected with each other. For TS and TW networks, they present an assortative mixing feature when  $k$  is smaller than 30. However, when  $k$  is greater than 30, it is not very obvious for the assortative mixing feature.

In Figure 7(d), the weighted average clustering coefficient  $\langle C^w(k) \rangle$  of all  $k$  degree nodes is plotted with  $k$ . In the early work, Abe and Suzuki [30] have proved the hierarchy structure of the unweighted earthquake network. Here, Figure 7(d) shows that the weighted network does not show the hierarchy structure. This may occur due to the fact that the cells used in their work are three dimensional in comparison to two-dimensional cells used by us. For TW and STID network, the weighted average clustering coefficient asymptotically decays with  $k$ , indicating a hierarchical topology of the networks.

As shown in Figure 7, we can see that the weighted average neighbor degree and the weighted average clustering coefficient of these networks show different results. It is an interesting phenomenon, and we believe that it is necessary to study the weighted network based on different construction methods.

For the evolution of the average nodal strength, we can see from Figure 8 that all the networks present a high correlation with the dynamics of seismic activities, which is  $\rho_{TS} = 0.97$ ,  $\rho_{TW} = 0.96$ , and  $\rho_{STID} = 0.94$ . This can explain

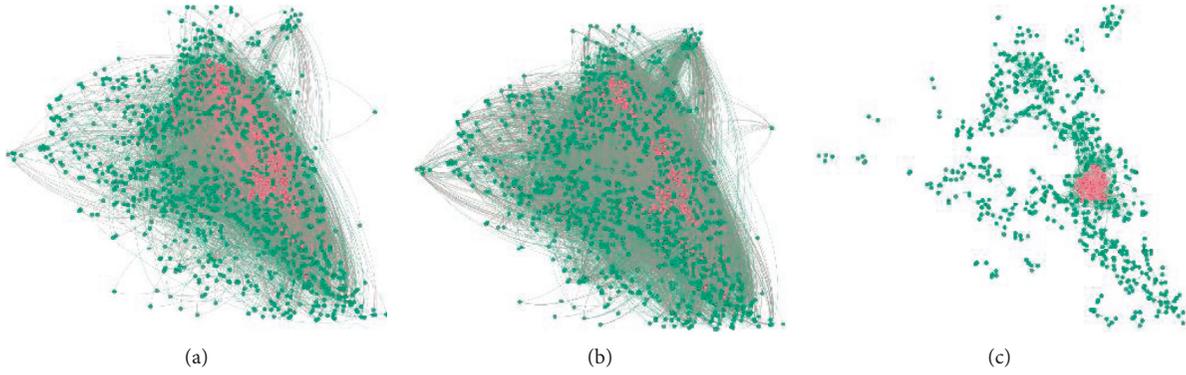


FIGURE 3: The visualization of the highest core nodes (red dots) and other nodes (green dots) of the networks constructed by the method of (a) TS, (b) TW, and (c) STID.

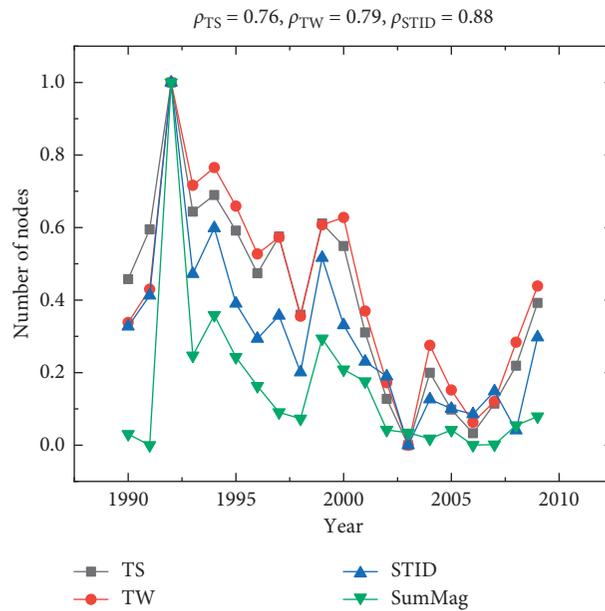


FIGURE 4: Evolution of the number of nodes of the networks constructed by TS, TW, and STID. SumMag represents the sum of earthquake magnitudes.

that the weighted form of networks is more helpful for earthquake network research than the unweighted one, which is also mentioned in their study [31, 32].

#### 4. Discussions

The advantage of the STID method is that it will show a better capture of dynamic characteristics than other methods when focusing on a specific geographical region. Using  $k$ -core decomposition, the highest core of the STID network has a more clustering feature. Moreover, the Landers earthquake ( $M$  7.3) at 11:57:34.13 on June 28, 1992,

has occurred in the region where all the nodes with the highest coreness located. It is because that STID focuses more on the direct influence, and its construction method is based on how important a node is, while the TS and TW methods are more similar, with better connectivities, and can capture global seismic activity characteristics. From the general information results, we can see that STID has fewer nodes and edges while TS and TW have many connections. Therefore, if the researcher focuses more on a specific region, we suggest using the STID method. If they focus more on global connectivity, we suggest the TS or TW method. Finally, it is still an important topic, which needs to be further studied.



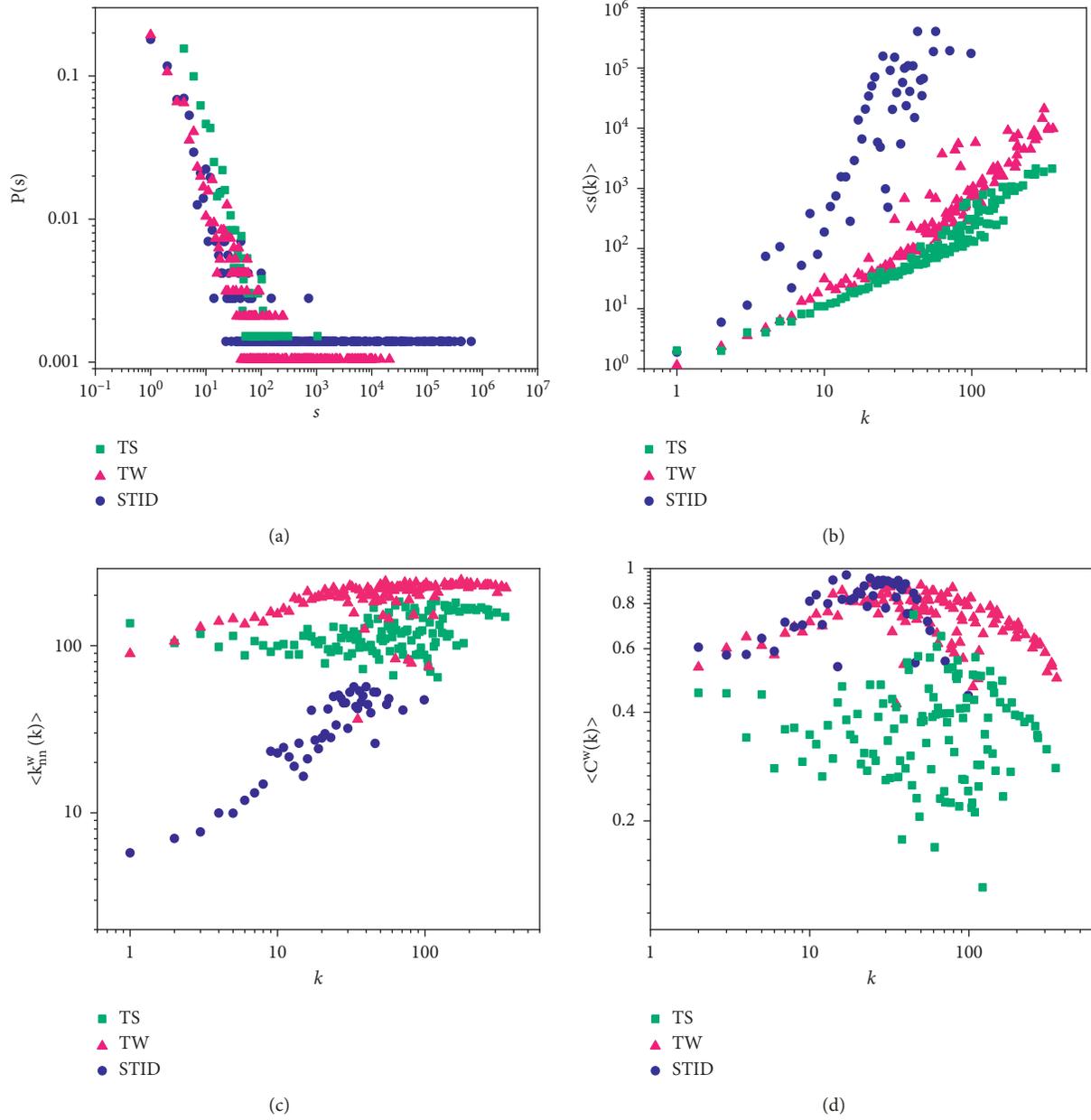


FIGURE 7: The weighted properties of the networks. (a) Strength distribution. (b) Plot of the average nodal strength  $\langle s(k) \rangle$  against the degree  $k$ . (c) Weighted average neighbor degree  $\langle k_{nn}^w \rangle$  against the degree  $k$ . (d) Weighted average clustering coefficient  $\langle C^w(k) \rangle$  against the degree  $k$ . The green squares, pink triangles, and blue dots represent TS, TW, and STID methods.

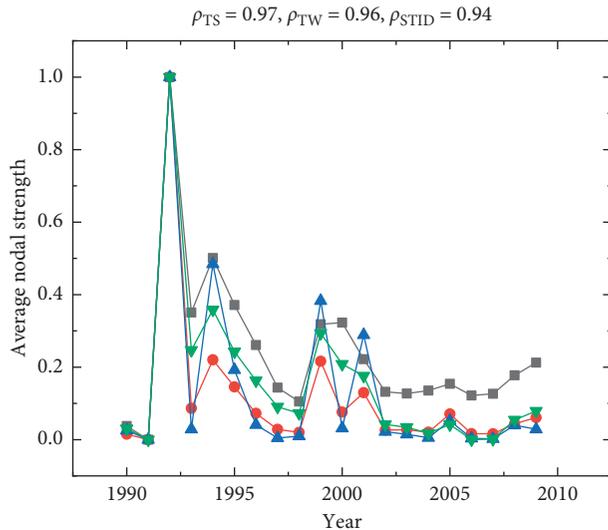


FIGURE 8: Evolution of the average nodal strength of the networks constructed by TS, TW, and STID. SumMag represents the sum of earthquake magnitudes.

## 5. Conclusions

This paper has compared and analyzed three classic earthquake network construction methods for seismicity, which are based on time series, time window, and space-time influence domain. We first give a detailed comparison of the construction process among these methods by visualization. Then, we found that these three networks all present scale-free and small-world properties. In the weighted form of earthquake networks, their nodal strength distributions all obey a power-law property and present a strength-degree correlation and an assortative mixing feature. However, the networks present some differences in some measures. They have different numbers of nodes, edges, and average degree. For TW and STID network, the weighted average clustering coefficient asymptotically decays with  $k$ , indicating a hierarchical topology of the networks. However, this feature is not observed in the weighted form of the TS network. We use decomposition to find the most central nodes. The highest core of the STID network has a more clustering feature, covering the great shock, while the highest core of TS and TW shows a broader distributed feature. On observing the evolution results of some measures for the three networks, we found that, except for betweenness, the number of nodes, average degree, and average nodal strength all show a strong correlation with seismic activities' dynamics. It implies that some of the network properties by different earthquake network construction methods will present some differences. In other words, when we discuss network characteristics as possible predictive indicators of seismicity in the future, the difference in network construction methods also needs to be taken as an essential consideration.

## Data Availability

The data used to support the findings of this study are available from the Southern California Earthquake Data Center (SCEDC) at [http://service.scedc.caltech.edu/eq-catalogs/date\\_mag\\_loc.php](http://service.scedc.caltech.edu/eq-catalogs/date_mag_loc.php).

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

The authors thank SCSN (Southern California Seismic Network) for providing the seismic data of the study area. This work was supported by the National Natural Science Foundation of China (NSFC) under Grant nos. 61806048, 61701096, and 61771121.

## References

- [1] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.
- [2] S. Damian, F. Andrea, W. Stefan et al., "STRING v10: protein-protein interaction networks, integrated over the tree of life," *Nucleic Acids Research*, vol. 43, no. D1, pp. 447–452, 2015.
- [3] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [4] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, no. 5916, pp. 892–895, 2009.
- [5] J. R. Banavar, A. Maritan, and A. Rinaldo, "Size and form in efficient transportation networks," *Nature*, vol. 399, no. 6732, pp. 130–132, 1999.
- [6] W. Wang, F. Xia, H. Nie et al., "Vehicle trajectory clustering based on dynamic representation learning of internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2020.
- [7] W. Wang, X. Zhao, Z. Gong et al., "An attention-based deep learning framework for trip destination prediction of sharing bike," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.
- [8] S. Stein and M. Wyssession, *An Introduction to Seismology, Earthquakes, and Earth Structure*, John Wiley & Sons, Hoboken, NJ, USA, 2009.
- [9] S. Abe and N. Suzuki, "Scale-free network of earthquakes," *EPL (Europhysics Letters)*, vol. 65, no. 4, pp. 581–586, 2004.
- [10] A. Jiménez, "A complex network model for seismicity based on mutual information," *Physica A*, vol. 392, no. 10, pp. 2498–2506, 2004.
- [11] X. He, H. Zhao, W. Cai, Z. Liu, and S.-Z. Si, "Earthquake networks based on space-time influence domain," *Physica A: Statistical Mechanics and its Applications*, vol. 407, pp. 175–184, 2014.

- [12] J. N. Tenenbaum, S. Havlin, and H. Eugene Stanley, "Earthquake networks based on similar activity patterns," *Physical Review E*, vol. 86, no. 4, Article ID 046107, 2012.
- [13] D. S. R. Ferreira, J. Ribeiro, A. R. R. Papa, and R. Menezes, "Towards evidences of long-range correlations in seismic activity," 2014, <https://arxiv.org/abs/1405.0307>.
- [14] S. Abe and N. Suzuki, "Dynamical evolution of clustering in complex network of earthquakes," *The European Physical Journal B*, vol. 59, no. 1, pp. 93–97, 2007.
- [15] S. Abe and N. Suzuki, "Dynamical evolution of the community structure of complex earthquake network," *EPL (Europhysics Letters)*, vol. 99, no. 1, pp. 313–316, 2012.
- [16] M. Lin, X. X. Fan, G. Wang, and G. Zhao, "Network structure entropy and its dynamical evolution for recurrence networks from earthquake magnitude time series," *The European Physical Journal B*, vol. 89, no. 5, p. 131, 2016.
- [17] S. Rezaei, H. Moghaddasi, and A. H. Darooneh, "PageRank: an alarming index of probable earthquake occurrence," *Chaos: an Interdisciplinary Journal of Nonlinear Science*, vol. 29, no. 6, Article ID 063114, 2019.
- [18] S. Rezaei, H. Moghaddasi, A. H. Darooneh, and M. Zare, "Forecasting earthquakes by hybrid model of pattern informatic and pagerank methods," *Bulletin of the Seismological Society of America*, vol. 109, no. 5, pp. 1605–1614, 2019.
- [19] J. K. Gardner and L. Knopoff, "Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?" *Bulletin of the Seismological Society of America*, vol. 64, no. 5, pp. 1363–1367, 1974.
- [20] California Institute of Technology, *SCEDC. Earthquake Catalogs SCSN Catalog Search (1932-Present)*, California Institute of Technology, Pasadena, CA, USA, 2020, [http://service.scedc.caltech.edu/eq-catalogs/date\\_mag\\_loc.php/](http://service.scedc.caltech.edu/eq-catalogs/date_mag_loc.php/).
- [21] S. Abe and N. Suzuki, "Small-world structure of earthquake network," *Physica A*, vol. 377, no. 1-2, pp. 357–362, 2004.
- [22] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, " $k$ -core organization of complex networks," *Physical Review Letters*, vol. 96, no. 4, Article ID 040601, 2006.
- [23] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.
- [24] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, "A model of Internet topology using  $k$ -shell decomposition," *Proceedings of the National Academy of Sciences*, vol. 104, no. 27, pp. 11150–11154, 2007.
- [25] H. Zhang, H. Zhao, W. Cai, J. Liu, and W. Zhou, "Using the  $k$ -core decomposition to analyze the static structure of large-scale software systems," *The Journal of Supercomputing*, vol. 53, no. 2, pp. 352–369, 2010.
- [26] M. Z. V. Batagelj, "Generalized cores," *Advances in Data Analysis and Classification*, vol. 5, no. 1, pp. 129–145, 2011.
- [27] X. He, H. Zhao, W. Cai, G.-G. Li, and F.-D. Li, "Analyzing the structure of earthquake network by  $k$ -core decomposition," *Physica A: Statistical Mechanics and its Applications*, vol. 421, pp. 34–43, 2015.
- [28] Gephi, 2020, <https://gephi.org/>.
- [29] N. Lotfi, A. H. Darooneh, and F. A. Rodrigues, "Centrality in earthquake multiplex networks," *Chaos: an Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 6, Article ID 063113, 2018.
- [30] S. Abe and N. Suzuki, "Complex earthquake networks: hierarchical organization and assortative mixing," *Physical Review E*, vol. 74, no. 2, Article ID 026113, 2007.
- [31] X. He, L. Wang, H. Zhu, and Z. Liu, "Statistical analysis of complex weighted network for seismicity," *Physica A: Statistical Mechanics and its Applications*, vol. 563, Article ID 125468, 2021.
- [32] A. Chakraborty, G. Mukherjee, and S. S. Manna, "Weighted network analysis of earthquake seismic data," *Physica A: Statistical Mechanics and its Applications*, vol. 433, pp. 336–343, 2015.