Analysis of rainfall relationships in East Asia using a complex network

Kyunghun Kim[1], Jaewon Jung[2], Hung Soo Kim[1], Masahiko Haraguchi[3], and Soojun Kim[1]

[1]Department of Civil engineering, INHA University, Incheon, 22212, Rep. of Korea
[2]Center for Hydrology and Ecology, Incheon, 22212, Rep. of Korea
[3] Research Institute for Humanity and Nature, Kyoto, Japan

Correspondence to: Soojun Kim (sk325@inha.ac.kr)

Abstract. Concurrent floods in multiple locations pose systemic risks to the interconnected economy in East Asia through supply chains. Despite the significant economic impacts, the understanding of the interconnection between rainfall patterns in the region is yet limited. Here, we analyzed spatial dependence in rainfall patterns of the 24 mega-cities in the region using complex analysis theory and discussed the technique's applicability. Each city and correlation coefficient was represented by a node and a link, respectively. Vital node identification and clustering analysis were conducted using adjacency information entropy and multi-community detection. The groups were clustered to reflect the spatial characteristics of climate. In addition, the climate links between each group were identified through the cross-mutual information considering the delay time for each group. We conclude that complex network analysis could be a valuable method for analyzing the spatial relationship between climate factors.

Keywords: Concurrent floods, Complex network, Adjacency information entropy, Multi-community detection

1 Introduction

East Asia accounts for 54% of the global supply chain, providing a wide range of services and products across the world (Ann et al, 2020). However, East Asia is prone to major floods. According to the Centre for Research on the Epidemiology of Disasters (CRED) disaster database, between 2000 and 2020, an annual average of 165 flood disasters occurred worldwide, resulting in 5,278 deaths and economic damage up to $29 million. While more than 22% of these flood disasters have occurred in East Asia, more than 60% of global-flood deaths and economical damage worldwide are in the region. For instance, Thailand recorded 813 deaths and $40 million worth of damages from floods in 2011 (Haraguchi and Lall, 2015), while China recorded 300 fatalities and $4.5 million in damages from floods in 2019 (CRED). These flood damages occurred in several areas of East Asia simultaneously. Even though floods occurred simultaneously at distant places, the impacts of floods will propagate through supply chains, incurring economic losses in the entire region. In this sense, concurrent flooding causes severe life and economic losses in multiple countries at the same time, disrupting the global economy more severely. For example, in 2020, concurrent floods in East Asia inundated automobile factories in Thailand, disrupting automobile supply, adversely affecting China's rare-earth and fertilizer industries along the Yangtze river, and affecting the global rare-earth industry (AON. 2020). Changes in rainfall characteristics caused by climate change are some of the primary causes of concurrent floods in East Asia. These changes occur across all regions, and the changed characteristics affect each other, resulting in even more significant changes (Wang et al., 2020). Therefore, it is vital to investigate the relationships among rainfall patterns in each region. Many studies have been conducted to identify rainfall relations in East Asia. Most of them investigated the relationship between major East Asian countries using statistical techniques (Jee-Hoon et al., 2008, Kosaka et al., 2011, Deng et al., 2014), and some demonstrated connections between weather factors or indicators, sea-level temperature, and monsoons. (Renguan et al., 2003, Lau and Kim., 2006, Li et al., 2010, Jianqi and Huijun, 2012, Renguan, 2017). Researchers have also used teleconnection methods to discover relationships between precipitation in East Asia and other parts of the world (Kripalani and Kulkarni, 2001, Sahai et al., 2003, Riyu and Zhongda, 2009, Lin, 2014, Maity et al., 2020.). These studies have been used to anticipate

rainfall in East Asia and aid in the preparation of flood disasters. In this study, we investigated the usefulness of complex network concepts for relationship analysis.

Complex network theory, developed by Leonard Eüler in 1735, expresses and analyzes a subject or phenomenon as a graph. In the late 1990s, Watts et al. (1998) and Barabási and Albert (1999) extended the analytical technique, making the theory fundamental in network science. A complex network can display a complicated phenomenon as a simple graph. Information obtained from the methodology can be used to identify the characteristics of subjects, their physical behavior, and the roles and relationships of the phenomenon's components. Complex network analysis has also been used in various fields because of its high applicability. For example, researchers have applied it to social networks (Michael et al., 2010), world trade (Brett et al., 2007), air transportation nets (Cardillo et al., 2013), patterns in human migration (Davis et al., 2013), and others. The analytical method has also been used in the fields of hydrology and meteorology fields to discover new patterns and relationships (Donges et al., 2009, Scarsoglio et al., 2013, Boers et al., 2015, Hongjun et al., 2020, Wolf et al., 2020). About precipitation related researches, the method had been used to analyze extreme rainfall patterns around the world (Boers et al., 2019), track rainfall events caused by typhoons (Ozturk et al., 2018), and study the spatial connectivity of rainfall (Ihsan et al., 2018) to determine new information or characteristics.

With the encouraging results of previous rainfall-related studies, this study applied complex network theory to rainfall in East Asia to understand the relationships between the rainfall patterns in each region. A complex network defines connectivity using correlation methods. Therefore, characteristics can be analyzed from the relationships. In addition, for clustering analysis, complex network-based methods consider the entire network, rather than the regions independently, unlike many other traditional methods. This feature results in a more accurate clustering (Xiang et al., 2020). Despite this advantage, one of the challenges in complex network theory is to identify thresholds, which determines whether the links existed. While no perfect methodology exists to clearly address this challenge, new methodologies are constantly being proposed. In this study, we assumed that each region (node) is connected with all the other regions (nodes) in the network, and that each connection (link) has a correlation coefficient as a weight. Because the weights are used as input data in each analysis enabled the relationship between regions to reflect in the network and be analyzed. We assessed the effects of each region through centrality analysis and grouped the regions according to clustering analysis. Subsequently, mutual information (MI) was calculated with a time lag (i.e., cross-mutual information) to identify the relationships between each group. Before related complex network studies, they only considered spatial factors. But in this study, we added temporal factors in relationship.

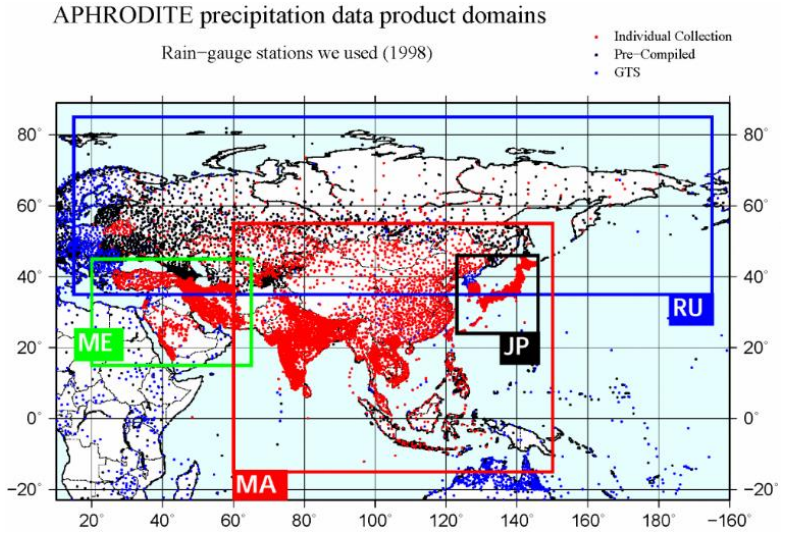
The remainders of this paper were organized as follows. Section 2 describes the study area and data used in this study. The complex network theory and related indicators are detailed in Section 3. Section 4 presents the results of the complex network analysis of East Asia and a discussion of these results. Section 5 presents the conclusions of this study.

2 Study area and materials

In this study, the major cities in East Asia were studied (Fig. 1). Among the cities in East Asia, we used selected cities from the World Bank Group report called "East Asia's Changing Urban Landscape" (The World Bank, 2015). However, we excluded Surabaya, Jakarta, and Badung (Indonesia) from the selected cities because of the changes in the location of rainfall observation since 2007. Instead, we included Ho Chi Minh City, Hai Phong (Vietnam), and Cebu (Philippines), which are economically active. Thus, a total of 24 cities were selected.

Figure 1. Selected 24 major cities in East Asia

This study used daily precipitation data from the Asian Precipitation – Highly Resolved Observational Data Integration Toward Evaluation (APHRODITE) girded precipitation dataset (Akiyo et al., 2012). The APHRODITE data contains long-term, high-
80 resolution daily rainfall data of the Asian continent obtained from the dense precipitation observation data network (Fig. 2). The data were obtained from the APHRODITE Water Resource project conducted by the Research Institute for Humanity and Nature (RHIN) and the Meteorological Research Institute of Japan Meteorological Agency (MRI/JMA) and have been used in many studies because of their high definition.



85 **Figure 2.** APHRODITE data **(http://aphrodite.st.hirosaki-u.ac.jp/products.html)**

Daily rainfall data for each city consisted of observations from January 1, 1981, to December 31, 2015. The basic statistics for each city's rainfall data are the same as those listed in Table 1.

Table 1. Basic statistics values for rainfall data of cities; Basic statistics contain average, standard deviation, coefficient of variation and skewness;

Station	Average (mm/day)	Standard deviation	Coefficient of Variation	Skewness
Pearl River Delta	4.277	9.416	2.202	4.063
Tokyo	3.637	10.736	2.952	7.044
Shanghai	2.985	6.446	2.16	3.993
Beijing	1.316	4.377	3.325	6.395
Manila	6.241	12.514	2.005	5.023
Seoul	3.457	9.427	2.727	5.543
Osaka	2.802	5.9	2.105	4.619
Bangkok	1.101	4.067	3.694	7.308
Tianjin	3.903	9.932	2.544	4.91
Shantou	2.211	4.853	2.195	5.279
Chengdu	3.782	6.241	1.65	3.246
Ho Chi Minh City	4.551	11.314	2.486	4.919
Nagoya	3.176	7.762	2.444	4.492
Wuhan	4.583	11.278	2.461	4.427
Hong Kong SAR	4.889	6.46	1.321	2.177
Shenyang	1.59	4.946	3.11	6.183
Taipei	6.955	16.128	2.319	6.192
Hangzhou	3.614	7.413	2.051	3.981
Kuala Lumpur	6.196	7.97	1.286	2.311
Xi’an	1.527	4.099	2.684	5.054
Ha Noi	4.053	8.87	2.189	4.417
Chongqing	2.79	5.751	2.061	4.89
Cebu	4.126	6.759	1.638	4.923
Hai Phong	3.801	9.53	2.507	5.023

The city with the largest rainfall on average is Taipei, which is approximately six times more than that of Bangkok, which has the lowest. Bangkok has the largest variation, and Kula Lumpur has the least.

3 Methodologies

3.1 Complex network analysis

Complex network analysis effectively visualizes a subject or phenomenon using a network and analyzes its characteristics, components, and relationships among nodes in the network. To apply complex network analysis, nodes and links must be defined. A node is a fixed element that serves as a point of intersection/junction within a network. For example, in the global airways network, airports become nodes. A link is an element that connects each node. In a global airways network, airways are the links. Defining these two elements is crucial in the analysis because even networks with the same number of nodes and links can assume various forms (Fig. 3).

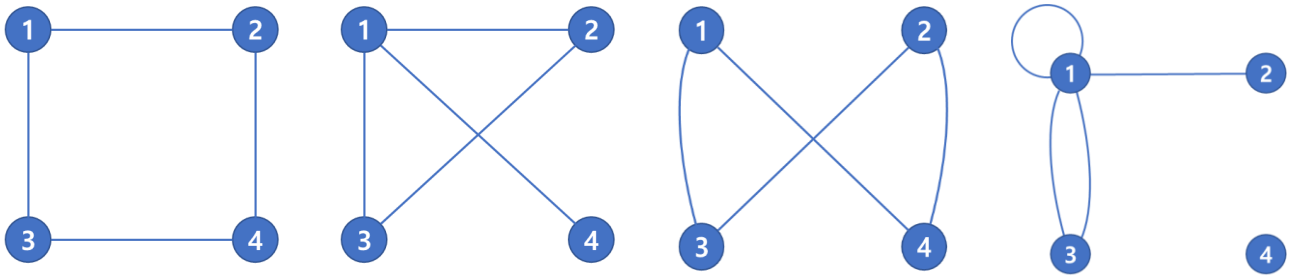


Figure 3. Various shapes of networks with the same number of nodes and links

In a complex network, the links are the most influential aspects of the network. This is because the type and characteristics of the graph vary depending on the type of link used and how it is defined. Based on the directionality and weight of the link, the network can be an undirected/directed network or an unweighted/weighted network. Generally, it is easy to define links in systems such as transportation systems or power grid system, which show clear connections between elements. However, if uncertainty occurs in the connections like social networks, researchers must define them. The most widely used methodology is the correlation coefficient. Depending on the value of the correlation calculated between two nodes, the researcher can define whether a link exists. While various previous studies used the Pearson correlation coefficient for links, they tend to derive inaccurate values if they were applied to nonlinear data. To address this problem, some researchers have utilized MI as an alternative (e.g., Donges et al., 2009, Kyunghun et al., 2019, Ghorbani et al., 2021). MI is based on the information and probability theory. For two variables (A and B), it quantifies and represents the amount of information of B contained by the variable A.

$$I(A,B) = \sum_{b\in B}\sum_{a\in A} p(a,b)log(\frac{p(a,b)}{p(a)p(b)}) \quad (1)$$

Here, $p(a)$ and $p(b)$ are probability distributions of variables, and $p(a,b)$ is the joint probability density function of the variables. MI values range from 0 to ∞, and an MI of 0 indicates that the two variables are independent of each other. MI can consider the nonlinearity of the data and has the advantage of calculating the correlation between different data sizes (Goyal, 2014).

3.2 Vital node identification using adjacency information entropy

Important nodes in a network have various effects on the structure or function of the network. Identification of these nodes is of practical and theoretical value (Xiang et al., 2020). For example, if a government identifies which places have a key role in power grids and traffic networks, it can effectively invest and create defense measures to prepare for blackouts and traffic jams. While some methods of identifying important nodes have been developed, these methods have limitations and are only applicable to certain types of networks. Xiang et al. (2020) developed a new methodology based on information entropy, making it applicable to all types of networks. This method has better results than the existing methods. The procedure of the method is as follows:

First. Calculate degree(k_j) of each node in the network

$$k_i = \sum_{j\in\Gamma_j} w_{ji} \quad (2)$$

Here, Γ_j is a group of nodes that form links with node j. w_{ji} is weight of link that connect node j and node i. If a network is unweighted, degree is the number of neighbor nodes.

Second, calculate the adjacency degree (A_i) of each node.

$$A_i = \sum_{j\in\Gamma_j} k_j \quad (3)$$

Third, calculate the selection probability (P_{i_j}).

$$P_{i_j} = k_i / A_j \quad (4)$$

Final, calculate the adjacency information entropy (E_i).

$$E_i = \sum_{j\in\Gamma_j} (P_{i_j} log_2 P_{i_j}) \quad (5)$$

After comparing the calculated adjacency information entropy of each node, the importance is determined according to the descending power.

3.3 Multiresolution community detection in weighted complex networks

A complex network consists of many nodes and links. Some of the nodes with strong relationships or similar characteristics can be clustered together. These clusters have several features and perform specific network functions. However, the cluster results depend on the level of analysis. Therefore, the multi-resolution community detection method can be a useful method for understanding complex networks (Newman, 2012). Several cluster analysis methods have been used for complex networks, but they require intense computations for complicated network shapes and focus only on graphical properties (Long and Liu, 2019). To address these problems, Long and Liu (2019) proposed a new clustering methodology using an intensity-based community detection algorithm (ICDA) in weighted networks. The procedure for the proposed method is as follows:

First, calculate the link intensity (I_P) of each link.

$$I_P(e_{ij}) = \begin{cases} \sum_{p=1}^{P} \alpha_p \times \frac{\sigma(path_p(v_i, v_j))}{\min(w_i, w_j)}, & e_{ij} \in E \\ 0, & otherwise \end{cases} \quad (6)$$

Here, $\sigma(path_p(v_i, v_j))$ is the sum of link weights in the path through p links from node i (v_i) to node j (v_j), P is the parameter of the path, and α_p is a polygonal effect parameter. For edge e_{ij} between node i and node j, w_i and w_j are their respective strengths.

Second, identify the links with link intensity greater than the selected threshold and create a group of nodes with the identified links.

$$\begin{aligned} &v_j \in V, \quad I_P(e_{ij}) > t \\ &v_j \in c_u, \quad I_P(e_{ij}) > t \end{aligned} \quad (7)$$

Here, t ($0 < t \leq 1$) is the selected threshold, and c_u is a group of nodes.

Third, calculate the belonging coefficient (I_P) of the nodes in node set u (c_u).

$$I_P(c_u, v_j) = \sum_{v_i\in c_u} I_P(e_{ij}) \quad (8)$$

This method has the advantages of forming groups more accurately and a faster computational speed than other methods.

4 Application and Results

4.1 Construction of East Asia rainfall network

In this study, we described the shape of a rainfall network as being weighted and undirected. In the network, each node was selected from 24 major cities, and link weights represented shared knowledge between nodes. Each node had different link weights. Table 2 compares the results of the link weights of the nodes.

Table 2. Average, maximum, and minimum link weights of each node; Parentheses under link weights are nodes that form a maximum or minimum value for target nodes;

Node	Average	Maximum (Node)	Minimum (Node)
Pearl River Delta	0.352	1.674 (Hong Kong SAR)	0.203 (Tokyo)
Tokyo	0.226	0.528 (Nagoya)	0.140 (Tianjin)
Shanghai	0.253	1.076 (Hangzhou)	0.153 (Tokyo)
Beijing	0.232	0.850 (Tianjin)	0.130 (Osaka)
Manila	0.294	0.467 (Ho Chi Minh City)	0.219 (Shanghai)
Seoul	0.227	0.276 (Ha Noi)	0.155 (Tokyo)
Osaka	0.244	0.870 (Nagoya)	0.130 (Beijing)
Bangkok	0.272	0.520 (Ho Chi Minh City)	0.194 (Shanghai)
Tianjin	0.252	0.850 (Beijing)	0.143 (Osaka)
Shantou	0.284	0.861 (Hong Kong	0.183 (Beijing)
Chengdu	0.293	0.621 (Chongqing)	0.200 (Shanghai)
Ho Chi Minh City	0.308	0.520 (Bangkok)	0.217 (Shanghai)
Nagoya	0.254	0.870 (Osaka)	0.139 (Beijing)
Wuhan	0.292	0.529 (Hangzhou)	0.195 (Taipei City)
Hong Kong SAR	0.364	1.674 (Pearl River Delta)	0.215 (Tokyo)
Shenyang	0.240	0.367 (Tianjin)	0.166 (Tokyo)
Taipei	0.220	0.333 (Shantou)	0.147 (Beijing)
Hangzhou	0.279	1.076 (Shanghai)	0.164 (Beijing)

Kuala Lumpur	0.308	0.377 (Wuhan)	0.207 (Beijing)
Xi'an	0.271	0.508 (Chengdu)	0.188 (Osaka)
Ha Noi	0.341	1.162 (Hai Phong)	0.238 (Tokyo)
Chongqing	0.289	0.621 (Chengdu)	0.207 (Beijing)
Cebu	0.246	0.354 (Manila)	0.179 (Beijing)
Hai Phong	0.342	1.162 (Ha Noi)	0.235 (Shanghai)

According to Table 2, the ranges of average, maximum, and minimum link weights were 0.22–0.37, 0.27–1.67, and 0.13–0.24, respectively. The average and minimum values had a narrow range, whereas the maximum values had a relatively wide range. We observed that the cities with the maximum values for each node were closely located. This is because the characteristics of rainfall in cities located in similar areas are similar; thus, the value of the MI is high. Each node had a maximum value with several different cities, while the minimum value was for certain cities such as Beijing and Tokyo. Beijing and Tokyo were selected as the cities with the lowest MI value eight and seven times, respectively. This was because the two cities are on the outskirts of East Asia.

4.2 Vital node identification by adjacency information entropy

For the network, we applied vital node identification (VNI) to determine the influence of nodes. VNI can be used to analyze all types of networks and more precisely determine the effects of nodes more accurately. ~~The results for the 24 nodes are shown in Table 3.~~

~~**Table 3. Rank of nodes according to adjacency information entropy**~~

Rank	Node	Adjacency information entropy
1	Hong Kong SAR	5.5022
2	Pearl River Delta	5.3833
3	Hai Phong	5.2789
4	Ha Noi	5.2701
5	Kuala Lumpur	4.9206
6	Ho Chi Minh	4.9182
7	Manila	4.7668
8	Chengdu	4.7560
9	Wuhan	4.7441
10	Chongqing	4.7092
11	Shantou	4.6543
12	Hangzhou	4.5983
13	Bangkok	4.5198
14	Xi'an	4.5117
15	Nagoya	4.3093
16	Shanghai	4.3023
17	Tianjin	4.2954
18	Cebu	4.2148
19	Osaka	4.1917

20	Shenyang	4.1507
21	Beijing	4.0442
22	Seoul	3.9890
23	Tokyo	3.9742
24	Taipei City	3.9049

The cities with high ranking nodes were located around the South China Sea (Fig. 4). In addition, they had a large average of MI. Cities with the low-rank nodes were in the northeast outskirts, except for Taipei, and had a low mean of MI. From these results, we can deduce that the location of a node affects its influence. High-rank nodes were in the center of the map and had many neighboring nodes. The characteristics of low-rank nodes were diametrically opposed to those of high-rank nodes. However, location is not the only factor affecting vital node identification. Despite its proximity to the South China Sea, Taipei had a low ranking because its link weights were the smallest on average (0.220).

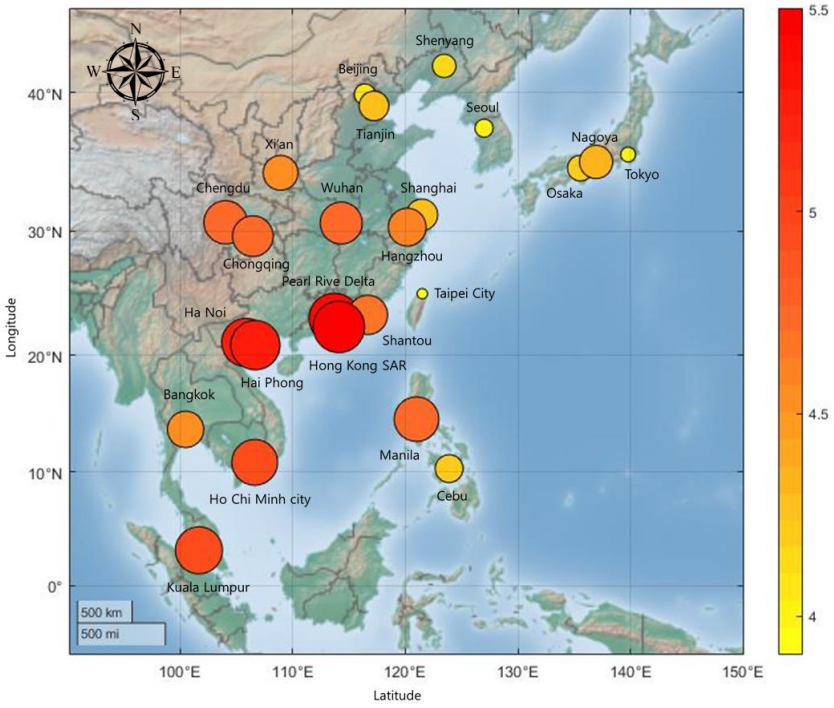
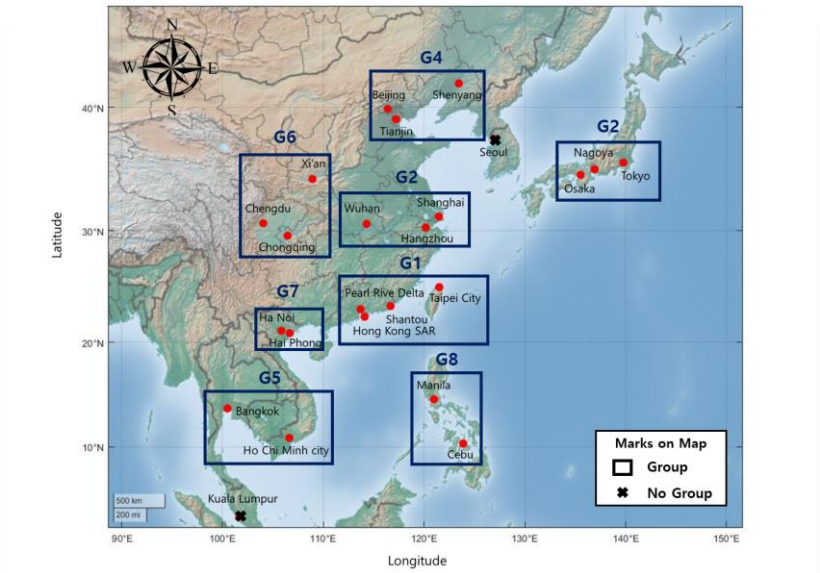


Figure 4. Adjacency information entropy value of cities; Color and size of circle are respectively proportion to the entropy and rank; Ride side of bar shows the adjacency information entropy values of nodes; Except for Taipei city, nodes near south china sea had higher values;

4.3 Clustering analysis using multiresolution community detection

In the clustering analysis, multiresolution community detection was applied to 24 nodes to create groups. After calculating the belonging coefficient, we determined the groups based on the threshold value. To form a group of nodes with strong relationships, the threshold value was the 95th quartile of the calculated belonging coefficient, 0.06.



185 **Figure 5. Group of nodes using multiresolution community detection; There are 8 groups in the East Asia; G1(Pearl River Delta, Hong Kong SAR, Shantou, Taipei City), G2(Osaka, Nagoya, Tokyo), G3(Wuhan, Hangzhou, Shanghai), G4(Tianjin, Shenyang, Beijing), G5(Bangkok, Ho Chi Minh City), G6(Xi'an, Chengdu, Chongqing), G7(Hanoi, Haiphong), G8(Manila, Cebu); Seoul and Kuala Lumpur did not make group with other nodes.**

~~**Table 4. Groups of nodes and their components**~~

Group	Components
G1	Pearl River Delta, Hong Kong SAR, Shantou, Taipei City
G2	Osaka, Tokyo, Nagoya
G3	Wuhan, Shanghai, Hangzhou
G4	Shenyang, Beijing, Tianjin
G5	Bangkok, Ho Chi Minh City
G6	Chengdu, Xi'an, Chongqing
G7	Ha Noi, Hai Phong
G8	Manila, Cebu
Other	Seoul, Kuala Lumpur

190

Nodes in close proximity formed a group (Fig. 5). The cities of Seoul (South Korea) and Kuala Lumpur (Malaysia) were not clustered with the others. Because of location of Seoul, it had low belonging coefficients with nearby nodes. Seoul is in the Korean peninsula. It is influenced by maritime air mass in summer and by continental air mass in winter. Therefore, precipitation of Seoul is affected by both features and has different characteristics. This feature made Seoul distinguish between

195 G2 and G4. For the Kuala Lumpur node, the belonging coefficient calculated with other nodes was between 0.03 and 0.05.

~~**Table 5. Representative cities in the groups**~~

Group	Representative city
G1	Hong Kong SAR
G2	Nagoya
G3	Shanghai
G4	Tianjin
G6	Chengdu

4.4 Relationship between node groups

Nodes were grouped based on their belonging coefficients (Section 4.3). The relationships between the groups were determined using cross-mutual information analysis. The cross-mutual information is a methodology for calculating MI by adding time lags between targets. It can estimate an appropriate correlation coefficient by considering the time intervals for geographically distant points. In this study, the time lag ranged from −10 to 10 days, and we checked the maximum cross-mutual information value and corresponding time lag of each group.

~~**Table 6. Maximum cross-mutual information and lag time**~~

	G1	G2	G3	G4	G5	G6	G7	G8
G1	-	0.312 (−3)	0.333 (2)	0.312 (2)	0.438 (−9)	0.472 (2)	0.430 (−1)	0.388 (3)
G2	0.312 (3)	-	0.343 (1)	0.273 (2)	0.374 (−8)	0.398 (2)	0.323 (−9)	0.333 (−10)
G3	0.333 (−2)	0.343 (−1)	-	0.271 (0)	0.369 (−9)	0.496 (1)	0.355 (−3)	0.332 (−6)
G4	0.312 (−2)	0.273 (−2)	0.271 (0)	-	0.409 (−5)	0.403 (1)	0.361 (−1)	0.344 (−10)
G5	0.438 (9)	0.374 (8)	0.369 (9)	0.409 (5)	-	0.607 (3)	0.551 (1)	0.564 (3)
G6	0.472 (−2)	0.398 (−2)	0.496 (−1)	0.403 (−1)	0.607 (−3)	-	0.535 (−2)	0.486 (4)
G7	0.430 (1)	0.323 (9)	0.355 (3)	0.361 (1)	0.551 (−1)	0.535 (−2)	-	0.432 (−5)
G8	0.338 (−3)	0.333 (10)	0.332 (6)	0.344 (10)	0.564 (−3)	0.486 (6)	0.432 (5)	-

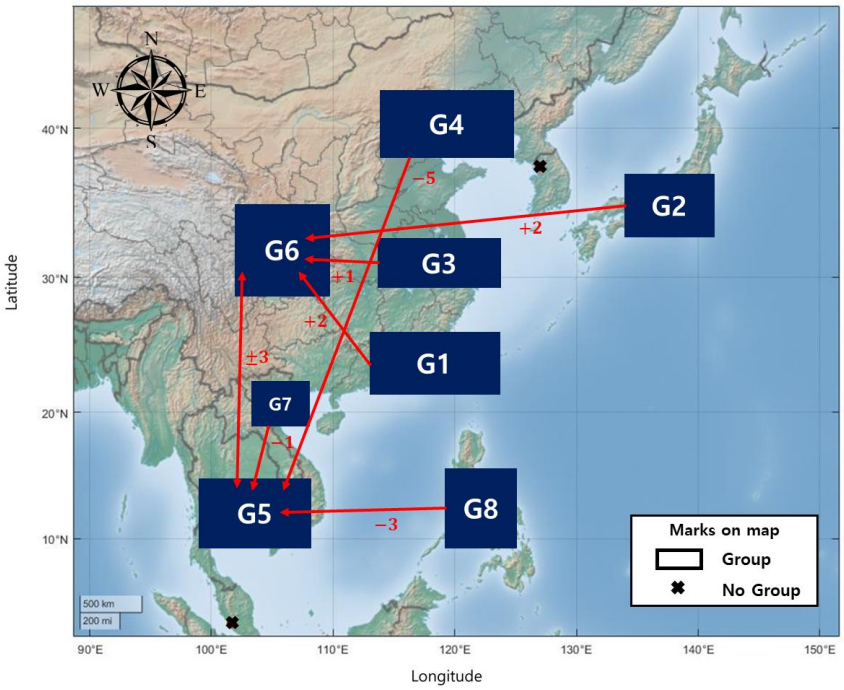


Figure 6. The maximum cross mutual information relationship and its time lag value; Each arrow points out the maximum relationship group and the numbers under the arrows express the lag time(days) of the maximum cross mutual information value; The figure shows relationship of groups and influence time intervals in East Asia;

As Figure 6 shows, most of the groups have strong relationships with G5 or G6 with maximum cross-mutual information values. G5 and G6 have a maximum cross-mutual information value with each other, and this value is larger than other cross-mutual information results. This result indicated that the two regions have a comparatively high relationship. A comparison of the lag times that forms the maximum cross-mutual information indicated that the maximum values were less than five days. Therefore, East Asian regions can affect each other within five days. The relationships in figure 6 were derived from the characteristics of East Asian rainfall. Indian and East Asian monsoons are major factors affecting rainfall in East Asia. The Indian monsoon brings very humid wind from the sea that conveys a large quantity of vapor across India and the Bay of Bengal to East Asia. This can reach the northern part of China (Renguang, 2017). If the Indian monsoon is strong, large amount of rainfall can occur in India and northern China. This characteristic was observed in the relationship between G5 and G4 (G5 is the first place affected by the Indian monsoon in East Asia, and G4 is the one in northern China). The Indian monsoon moves northwest from the Bay of Bengal, passing mainland China into the Sea of Okhotsk. G5, G6, and G7 in this pathway are related to each other by the Indian monsoon. The effect of the South China Sea, which supplies vapor to the mainland, was contained in the relationship between G1 and G6. In the summer, vapor from the South China Sea causes much rainfall in southern China and arrives in the mainland (Kanaly et al., 1996). Similar to the Indian Monsoon, the East Asian monsoon affects East Asian rainfall. The East Asian monsoon begins in the Western Pacific, moves eastward through Indonesia, and ends in Japan and South Korea. If it is strong, it affects southern Vietnam and Thailand (Rehe and Akimasa, 2002). This was observed in the relationship between G8 and G5. In the summer, there is an anomalous anticyclone between China and Korea. The anomalous anticyclone located in the western sea forms a clockwise wind cycle throughout China, Korea, and Japan (Rengquang, 2017). This created wind cycle transports vapor from Japan to the east and the center of China. This phenomenon formed relationships between G2 and G6 and between G3 and G6.

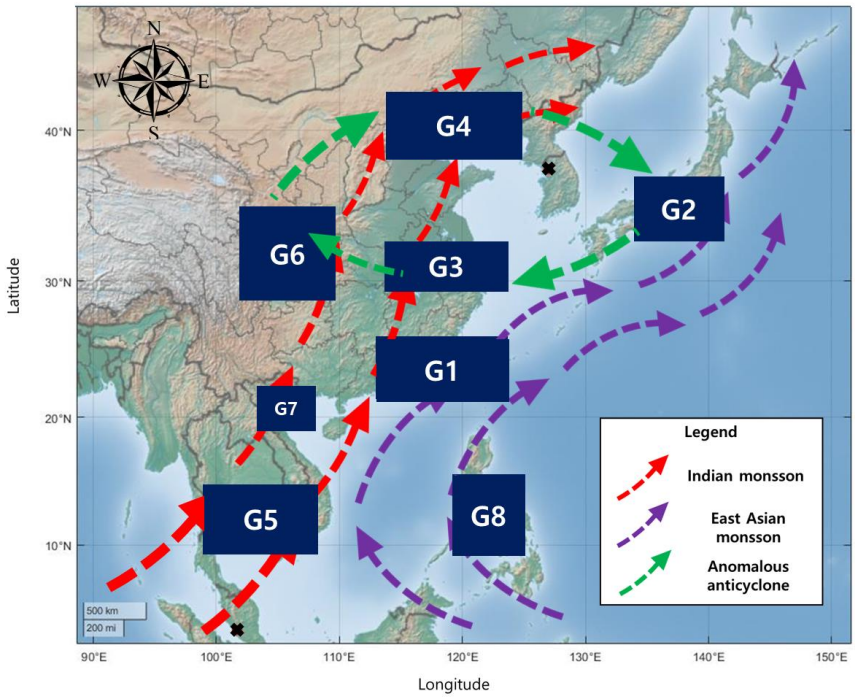


Figure 7. Major water vapor transport routes in East Asia; The routes could explain the reasons why the relationship of groups was made like Figure 6; Indian monsoon brings vapor from Indian Ocean, East Asian monsoon gets vapor from Pacific Ocean and east china sea; Anomalous anticyclone provide vapor in east china, Korea and Japan;

4.5 Discussion

Complex network analysis has the advantage of reducing complex phenomena or systems to a graph form, making it easier to determine characteristics. In addition, it can be used to analyze the effects of network components, perform clustering analysis, etc. In this study, we used these merits to examine the relationships between major cities in East Asia.

To create a rainfall network, we first calculated the MI between nodes and used it as the link's weight. Thus, the network could reflect the correlation of rainfall in each city and was used as the most important factor in subsequent analyses. To check the effects of nodes in the network, the adjacency information entropy was calculated and compared. The results indicated that nodes surrounding the South China Sea were highly ranked, and the node's location was an important factor in identifying vital nodes. The South China Sea is one of the primary vapor providers in East Asia, and the two significant monsoons pass through it. Vapor from the South China Sea first affects coastal cities and then moves to other cities in the continent. Thus, rainfall from some cities affects the neighboring cities. Based on this phenomenon, cities located in the South China Sea ranked high. As described in Section 4.3, the coefficient of each node was calculated by using the link weight. Each group consisted of nodes located nearby, and their coefficients were significantly higher than those of the other nodes. The correlation values of representatives in groups were higher than the nodes in the same group. We observed that a city with a high correlation with nearby cities had a significant influence. After clustering, we applied cross-mutual information analysis to determine the relationships between groups. During the analysis, the lag time was considered because the groups were geographically separated. The cross-mutual information results were interpreted using the rainfall characteristics of East Asia. Two monsoons (Indian and East Asian monsoons) and anomalous anticyclones affected group relationships. An interesting result was the strong relationship between G5 and G6. Even with G7 between them, they have a strong correlation. Previous research has primarily focused on the relationship between southern China (G1) and regions surrounding the East China Sea (G5, G7, and G8) (Tie and Xiushu, 2008, Hu et al., 2014, Zhao et al., 2017). These studies analyzed the effects of monsoons in the East

China Sea but did not expand the region to G6. Therefore, research into the physical interpretation of the link between the G5 and G6 regions is required.

The complex network facilitated a simple analysis of the relationship between East Asian cities. Unlike previous studies, we considered temporal factors in the relationship. Through this, we observed new relationships and characteristic of rainfall in East Asia. During the analysis, two methods (vital node identification and multiresolution community detection) were very useful for analyzing the network and made reliable results. From the results of research, we could find out that our research framework is helpful process for studying relationships in regions. The frame contains not only topological analysis, but also statistical analysis and considers temporal factor. Also, in the result, it reflects climate cycle factors and reveals its characteristics.

5 Conclusions

Concurrent floods in East Asia inundate the firm's production facilities at multiple locations simultaneously, causing supply chains disruptions at the global level. In this study, we analyzed the spatial relationships between major cities in East Asia using a complex network. The East Asia rainfall network was composed of major cities (nodes) and correlation coefficients (links). After the network was created, vital node identification and clustering analysis were conducted using adjacency information entropy and multi-community detection. Cross-mutual information defined relationships between cluster groups in East Asia. The results revealed that the network reflected the rainfall characteristics of East Asia and the relationships significantly affected vital nodes and clustering analysis. In addition, we observed that Southeast Asia and northwest China have a strong relationship. The study observed that although the computational burdens of implementing complex network analysis is not so high, the method accurately reflects the relationship between regional rainfall and can be used to analyze the relationships between various weather factors. In a subsequent study, we will evaluate the applicability of the complex network methodology to interpret key climate factors, such as ENSO, IOD, and NAO, which have complex interconnection characteristics.

Code/Data availability

The Asian Precipitation – Highly Resolved Observational Data Integration Toward Evaluation (APHRODITE) girded precipitation dataset is available online at APHRODITE's Water Resources (http://aphrodite.st.hirosaki-u.ac.jp/).

Author contributions

K. Kim and J. Jung designed and conducted the research. K. Kim analyzed the results, with feedback from H. S. Kim, M. Haraguchi, and S. Kim. K. Kim prepared the paper, with contributions from all co-authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by a grant(2021-MOIS36-002) of Technology Development Program on Disaster Restoration Capacity Building and Strengthening funded by Ministry of Interior and Safety (MOIS, Korea).

References

2015 International Bank for Reconstruction and Development, The World Bank publication, Washington DC, USA, 2015.

Adam, P., Steve, B., Michal, L., Brian, K., and Gaurav, S.: AON Empower Results: Global Catastrophe Recap report, AON publications, London, England, 2020.

Akiyo, Y., Kenji. K., Osamu. A., Atsushi. H., Natsuko. Y., and Akio, K.: APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges, Bulletin of the American Meteorological Society, 93(9), 1401-1415, https://doi.org/101175/BAMS-D-11-00122.1, 2012.

Ann, M. U., James, K., Vanessa, M., and Kate, W.: 2020 Global Supply Chain Disruption and Future Strategies Survey Report, Foley & Lardner LLP publications, Detroit, USA, 2020.

Boers, N., Donner, R.V., Bookhagen, B., and Kurths, J.: Complex network analysis helps to identify impacts of the EL Niño Southern Oscillation on moisture divergence in South America, Climate Dynamics, 45, 619-632, https://doi.org/10.1007/s00382-014-2265-7, 2015.

Boers, N., Goswami, B., Rheinwalt, A., Bookhagen, B, Hoskins, B. and Kurths, J.: Complex networks reveal global pattern of extreme-rainfall teleconnections, Nature, 566, 373-377, https://doi.org/10.1038/s41586-018-0872-x, 2019.

Brett, W. B., Richard, A. H., and Tamara, G. K.: Temporal analysis of semantic graphs using ASALSAN, Seventh IEEE International Conference on Data Mining (ICDM 2007), 33-42, https://doi.org/10.1109/ICDM.2007.54, 2007.

Cardillo, A., Gomez-Gardenes, J., Zanin, M., Romance, M., Papo, D., Pozo, F., and Boccaletti, S.: Emergence of network features from multiplexity, Scientific Reports, 3, 1344, https://doi.org/10.1038/srep01344, 2013.

Chenghao, W. and Zhi-Hua, W.: A network-based toolkit for evaluation and intercomparison of weather prediction and climate modelling, Journal of Environmental Management, 268(110709), 1-11, https://doi.org/10.1016/j.jenvman.2020.110709, 2020.

Davis, K., D'Odorico, P., Laio, F., and Ridolfi, L.: Global spatio-temporal patterns in human migration: A complex network perspective, PLOS ONE, 8(1), 1-8, https://doi.org/10.1371/journal.pone.0053723, 2013.

Deng, Y., Gao, T., Yao, X., and Xie, L.: Regional precipitation variability in East Asia related to climate and environmental factors during 1979-2012, Nature Scientific Reports, 4, 5693, https://doi.org/10.1039/srep05693, 2014.

Donges, J., Zou, Y., Marwan, N., and Kurths, J.: Complex networks in climate dynamics, The European Physical Journal Special Topics, 174, 157-179, https://doi.org/10.1140/epjst/e2009-01098-2, 2009.

Ghorbani, M. A., Karimi, V., Ruskeepaa, H., Sivakumar, B., Pham, Q. B., Fatemeth, M., and Nazly, Y.: Application of complex networks for monthly rainfall dynamics over central Vietnam, Stochastic Environmental Research and Risk Assessment, 35, 535-548, https://doi.org/10.1007/s00477-020-01962-2, 2021.

Goyal M. K..: Monthly rainfall prediction using wavelet regression and neural network: an analysis of 1901-2002 data, Assam, India, Theoretical and Applied Climatology, 118, 25-34, https://doi.org/10.1007/s00704-013-1029-3, 2014.

Han, X., Ouarda TBMJ, Rahman, A., Haddad, K., Mehrotra, R., and Sharma, A.: A Network Approach for Delineating Homogeneous Regions in Regional Flood Frequency Analysis, Water Resource Research, 56(3), 1-18, https://doi.org/10.1029/2019WR025910, 2020.

Hongjun, J., Mungjin, L., Jongsung, K., Jaewon, J., Jaewon, K., and Hung, S. K.: Stream gauge network grouping analysis using community detection, Stochastic Environmental Research and Risk Assessment, 35, 781-795, https://doi.org/10.1007/s00382-014-2265-7, 2015.

Hu, W., Wu, R., and Liu, Y.: Relation of the South China Sea Precipitation Variability to Tropical Indo-Pacific SST Anomalies during Spring to Summer Transition, Journal of Climate, 27(14), 5451-5467, https://doi.org/10.1175/JCLI-D-14-00089.1, 2014.

Ihsan, N., Sivakumar, B., Fitsum, M. W., Srivastan, V. R., Minh, T. V., and Shie-Yui, L.: Spatial connections in regional climate model rainfall outputs at different temporal scales: Application of network theory, journal of Hydrology, 556, 1232-1243, https://doi.org/10.1016/j.jhydrol.2017.05.029, 2018.

Jee-Hoon, J., Baek-Min, K., Chang-Hoi, H., and Yeon-Hee, N.: Systematic variation in wintertime precipitation in East Asia by MJO-induced extratropical vertical motion, Journal of Climate., 21(4), 788-801, https://doi.org/10.1174/2007JCLI1801.1, 2008.

Jianqi, S. and Huijun, W.: Changes of the connection between the summer North Atlantic Oscillation and the East Asian summer rainfall, Journal of Geophysical research Atmospheres Climate and Dynamics, 117(D8), 1-12, https://doi.org.10.1029/2012JD017482, 2012.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewsk, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne and D. Josepth: The NCEP/NCAR 40-Year Reanalysis Project. Bull. Amer. Meteor. Soc., 77, 437-471. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kosaka, Y., Xie, S., and Nakamura, H.: Dynamics of Interannual Variability in Summer Precipitation over East Asia, Journal of Climate, 24(20), 5435-5453, https://doi.org/10.1175/2011JCHLI4099.1, 2020.

Kripalani, R. H. and Kulkarni, A.: Monsoon rainfall variations and teleconnections over South and East Asia, International Journal of Climatology, 21(5), 603-616, https://doi.org/10.1002/joc.625, 2001.

Kyunghun, K., Hongjun, J., Daegun, H., Soojun, K., Taewoo, L., and Hung, S. K.: On complex network construction of rain gauge stations considering nonlinearity of observed daily rainfall data, Water, 11(8), https://doi.org/10.3390/w11081578, 2019.

Lau, K. M. and Kim K. M.: Observational relationships between aerosol and Asian monsoon rainfall, and circulation, Geophysical Research Letter Atmospheric Science, 33(21), 1-5, https://doi.org/10.1029/2006GL027546, 2006.

Li, H., Dai, A., Zhou, T., and Lu, J.: Responses of East Asian summer monsoon historical SST and atmospheric forcing during 1950-2000, Climate Dynamics, 34, 501-514, https://doi.org/10.1007s00382-008-0482-7, 2010.

Lin, Z.: Intercomparison of the impacts of four summer teleconnections over Eurasia on East Asian rainfall, Advances in Atmospheric Sciences, 31, 1366-1376, https://doi.org/10.1007/s00376-014-3171-y, 2014.

Long, H., and Liu, X.: Multiresolution community detection in weighted complex networks, International Journal of Modern Physics C, 30(2&3), 1-15, https://doi.org/10.1142/S0129183119500165, 2019.

Maity, R., Chanada, K., Dutta, R., Ratnam, J. V., Nonaka, M., and Behera, S.: Contrasting features of hydroclimatic teleconnections and the predictability of seasonal rainfall over east and west japan, Meteorological Applications, 27(1), 1-20, https://doi.org/10.1002/met.1881, 2020.

Masahiko, H., and Upmanu, L.: Flood risks and impacts: A case study of Thailand's flood in 2011 and research questions for supply chain decision making, International Journal of Disaster Risk Reduction, 14, 256-272, https://doi.org/10.1016/j.ijdrr.2014.09.005, 2015.

Michael, S., Renaud, L., and Stefan, T.: Multirelational organization of large-scale social networks in an online world, Proceedings of the National Academy of Sciences of the United States of America, 107(31), 13636-13641, https://doi.org/10.1073/pnas.1004008107, 2010.

Newman, M. E. J.: Analysis of weighted networks, Phys, Rev., E70, 1-9, https://doi.org/10.1103/PhysRevE.70.056131, 2004.

Ozturk, U., Marwan, N., Korup, O., Saito, H., Agarwal, A., Grossman, M. J., Zaiki, M., and Kurths, J..: Complex networks for tracking extreme rainfall during typhoons, Chaos, 28, 1-8, https://doi.org/10.1063/1.5004480, 2018.

Preethi, B., Mujumdar, M., Kripalani, R. H., Amita, P., and Krishnan, R: Recent trends and tele-connections among South and East Asian summer monsoons in a warming environment, Climate Dynamics, 48, 2489-2505, https://doi.org/10.1007/s00382-016-3218-0, 2016.

Renguan, W., Zeng-zhen, H., and Ben, P. K.: Evolution of ENSO-related rainfall anomalies in East Asia, Journal of Climate, 16(22), 3742-3758, https://doi.org/10.1175/1520-0442(2003)016<3742:EOERAI>2.0.CO;2, 2003.

Renguang, W.: Relationship between Indian and East Asian summer rainfall variations, Advances in Atmospheric Sciences, 34, 4-15, https://doi.org/10.1007/s00376-016-6216-6, 2017.

Renhe, Z., and Akimasa, S.: Moisture Circulation over East Asia during EL Nino Episode in Northern Winder, Spring and Autumn, Journal of the Meteorological Society of Japan, 80(2), 213-227, https://doi.org/10.2151/jmsj.80.213, 2002.

Riyu, L. and Zhongda, L.: Role of Subtropical Precipitation Anomalies in Maintaining the Summertime Meridional Teleconnection over the Western North Pacific and East Asia, Journal of Climate, 22(8), https://doi.org/10.1175/2008JCLI2444.1, 2009.

Sahai, A., Pattanaik, D., Satyan, V., and Alice, M.: Teleconnections in recent time and prediction of Indian summer monsoon rainfall, Meteorology and Atmospheric Physics, 84, 217-227, https://doi.org/10.1007/s00703-002-0595-1, 2003.

Scarsoglio S., Laio F., and Ridolfi, L.: Climate Dynamics: A Network-Based Approach for the Analysis of Global Precipitation, PLOS ONE, 8(8), 1-11, https://doi.org/10.1371/journal.pone.0071129, 2013.

Tie, Y., and Xiushu, Q.: Study on lightning activity and precipitation characteristics before and after the onset of the South China Sea summer monsoon, Climate and Dynamics, 113(D14), 1-10, https://doi.org/10.1029/2007JD009382, 2008.

Wang, Z., Mu, J., Yang, M., and Yu, X.: Reexaming the mechanisms of East Asian summer monsoon changes in response to non-East-Asian anthropogenic aerosol forcing, Journal of Climate, 33(8), 2929-2944, https://doi.org/10.1175/JCLI-D-19-0550.1, 2020.

Wolf, F., Bauer, J., Boers, N., and Donner R. V.: Event synchrony measures for functional climate network analysis: A case study on South America rainfall dynamics, Chaos, 30(3), 1-12, https://doi.org/10.1063/1.5134012, 2020.

Xiang, X., Cheng, Z., Qingyong, W., Xianqiang, Z., and Yun, Z.: Identifying vital nodes in complex networks by adjacency information entropy, Nature Scientific Reports, 10, 1-12, https://doi.org/10.1038/s41598-020-59616-w, 2020.

Zhao, S., Liu, Z., Chen, Q., Wang, X., Shi, J., Jin, H., Liu, J., and Jian, Z.: Spatiotemporal variations of deep-sea sediment components and their fluxes since the last glaciation in the northern South China Sea, Science China Earth Sciences, 60, 1368-1381, https://doi.org/10.1007/s11430-016-9058-6, 2017.