Author's response

We got second review from two reviewers during discussion interval. The reviewers talked about many factors in the paper. We tried to reflect comments on revised manuscript. We believe that the paper is substantially improved as the result of the revision. We wrote point-by-point response (in blue) to reviewer's comments (in red). Also, we put revised part of the manuscript in black.

<**Reviewer 1**>

1) In order to a more legible text, it would be better if the results and discussion were reported in the present tense.

   ➢ Following the above comment, we changed the Result and Discussion part as present tense. Because both parts are too long to put in the Author's response, it would be appreciate to check changes of the Result and Discussion part in the revised paper or revised paper with track change.

2) Line 62 is not clear.

   ➢ In Line 62, we tried to express that using a correlation coefficient as a weight makes a network reflect relationship between regions. Because a link weight is one of the most important input value for constructing a network and both analysis (Vital node identification and Multi-resolution community detection) use it as important input data. The current expression about Line 62 is not clear for understanding the meaning. Therefore we changed it like below.

   - (Line 62) Using the correlation coefficient makes that a network reflect relationship between regions. The link weight is one of the most important inputs for constructing and analyzing a network.

3) Some phrase (Line 105, Line 107, Line 119) need refences.

   ➢ We put references on lines 105, 107, and 119 on each line.

   - (Line 105) The most widely used methodology is the correlation coefficient (Donges et al., 2009).
   - (Line 107) While various previous studies used the Pearson correlation coefficient for links, they tend to derive inaccurate values if they are applied to nonlinear data (Zadian et al., 2018).

     ※ Zadian M.A., Haapasilta, V., Relan, R., Paasonen, P., Kerminen, V. M., Junnien, H., Kulumala, M. and Foster, A.S.: Exploring non-linear associations between atmospheric new-particle formation and ambient variables: a mutual information

approach, Atoms. Chem. Phys., 18(7), 12699-12714, https://doi.org/10.5194/acp-18-12699-2018, 2018.

- (Line 119) While some methods of identifying important nodes have been developed, these methods have limitations and are only applicable to certain types of networks (Mester et al., 2021).

  ※ Mester, A., Pop, A., Mursa, B. E. M., Grebla, H., Diosan, L. and Chira, C.: Network Analysis Based on Important Node Selection and Community Detection, Mathematics, 9(18), 2294, https://doi.org/10.3390/math9182294, 2021.

4) (Line 123-125) If you are computing the degree accounting for the weight of links, the actual name is strength. Degree defines just the number of neighbors, as you secified.

➢ About the comment, we checked exrpession about Degree and Strength in other papers (Newman, 2001, Barrate etal. 2004, Opsahl et al. 2010) and found that our expression needs to be fixed. For a weighted network, it is more suitable to express degree as strength. We appreciate for letting know about the expression.

※ Newman, M. E. J.: Scientific collaboration networks. II. Shortest path, weighted networks, and centrality, physical reveiw E, 64, 016132, doi: 10.1103/PhysRevE. 64.016132., 2001.

※ Barrat, A., Barthelemy, M., Pastor-Satorras, R. And Vespignani, A. : The architecture of complex weighted networks, Proceedings of the National Academy of Sciences, 101(11), 3747-3752, https://doi.org/10.1073/pnas.0400087101, 2004.

※ Opsahl, T., Agneessens, F. And Skvoretz, J. : Node centrality in weighted networks : Generalizing degree and shortest paths, Social Networks, 32(3), 245-551, doi:10.1016/j.socnet.2010.03.006, 2010.

- (Line 135) First Calculate strength(k_i) of each node in a weighted network

5) (Line 189) Why has Kuala Lampur belonging coefficient so low?

➢ Precipitation of Kuala Lumpur gets alot of influence from the Boreal winter season or Australia Summer monsoon (Sigh and Xiaosheng, 2020). However, the other cities except for Kuala Lumpur, the primary influencer of precipitation is the Indian monsson and East Asia monsoon. Therefore, Kuala Lumpur has different precipitation characteristics and it makes lowering the belonging coefficient of Kuala Lampur. We put this explaination in Line 189 for authors.

- Unlike other cities, Kuala Lumpur is significantly influenced by the Boreal winter season and Australia Summer monsoon (Sigh and Xiaosheng, 2020). Therefore, Kuala Lumpur has a different rainfall pattern, and it makes lower results in the belonging coefficient.

  ※ Singh, V. and Xiaosheng, Q.: Study of rainfall variabilities in Southeast Asia using long-term gridded rainfall and its substantiation through global climate indices, Journal of Hydrology, 585, 124320, https://doi.org/10.1016/j.jhydrol.2019.

124320, 2020.

6) (Line 205) ″The relationships in figure 6 were derived from the characteristics of East Asian Rainfall″. I am not sure about the meaning of this phrase : do you mean that the characteristics of rainfall cause the relationship between groups ? or do you mean that the relationship between the groups can be explained by the synoptic atmospheric circulation, as you explain later? Please, clarify.

➢ Our intention about Line 205 is more close to later meaning. Most of rainfall in East Asia is concentrated on summer season. And synoptic atmospheric circulation (Indian monsoon and East asia monsoon) is a main influence factor on summer rainfall. Therefore, many previous studies concluded that the synoptic atmospheric circulation makes characteristics of rainfall and rainfall relationship in East Asia. But our expression did not clearly express the intention. So, we changed Line 205 like below.

- The relationships in figure 6 are derived from synoptic atmospheric circulation in East Asia.

7) The phrases (Line 65, 71, 151, 178, 186, 230, 259) need to be rewritten.

➢ We appreciate comments on English expression. We checked not only the lines suggested by the comments but also whole paper.

- (Line 65) Past studies about complex network considered only spatial factors, while we add also temporal factors.
- (Line 71) In this research, the major cities in East Asia are analyzed (Fig. 1).
- (Line 151) In this research, we design a rainfall network as a weighted and undirected graph.
- (Line 178) The threshold value is the 95th quantile of the calculated belonging coefficient, 0.06, in order to form a group of nodes with strong relationship.
- (Line 186) Seoul has low belonging coefficients with the nearby nodes because of its location in the Korean peninsula. The area is influenced by maritime air mass in summer and continental air mass in winter. Therefore, the precipitation of Seoul is affected by both features and has different characteristics.
- (Line 230) The adjacency information entropy was calculated and compared to check the effects of nodes in the network.
- (Line 259) Once created the network, vital node identification and clustering analysis were conducted using adjacency information entropy and multi-community detection.

<Reveiwer 2>

1) Where is Centre for Research on the Epidemiology of Disasters (CRED) ?

- Centre for Research on the Epidemiology of Disasters has been active for over 40 years in the fields of international disaster and conflict health studies with activities linking relief, rehabilitation and development. The Centre promotes research, training and technical expertise on humanitarian emergencies, particularly public health and epidemiology. CRED's research scope includes natural disasters and crisis caused by civil strife, conflict or others. CRED offers EM-DAT (International Disaster database). This database was developed in 1988 to rationalize decision making for disaster preparedness, while also providing an objective base for vulnerability assessment and priority setting. EM-DAT contains essential core data on the occurrence and effects of disasters worldwide, running from 1900 to the present. If reviewer wants to find more detailed information about CRED, the below link is connected to main web page of CRED

  https://www.cred.be/

  We added information of CRED in Introduction part for authors who have any knowledge about the CRED.

  - (Introduction) According to disaster database of the Centre for Research on the Epidemiology of Disasters (CRED), which offer essential core data on the occurrence and effects of disasters all over the world, annual average of 165 flood disasters occurred worldwide during 2000 to 2020, resulting in 5,278 deaths and economic damage up to $29 million.

2) Check the Refs about Watts and Strogatz (1998) and Barabási and Albert (1999). Also check the rest

- We checked the Refs about the two papers and found they were omitted. Therefore, two papers were added in the Reference part. Also we checked other papers whether those were correctly written in Reference part.

  - (Reference) Watts, D. J. and Strogatz, S. H.: Collective dynamics of 'small-world' networks, Nature, 393, 440-442, https://doi.org/10.1038/30918, 1998.
    Barabási, A. L. and Albert, R.: Emergence of scaling in random networks, Science, 286, 509-512, https://doi.org/10.1126/science.286.5439.509, 1999.

3) Seperate all paragraphs by space and/or indent

- We checked all paragraphs and put space in the beginning of each paragraph

4) This sentence is confusing because there is no mention of rain nor precipitation in this document (This sentence = ‘we used selected cities from the World Bank Group report called “East Asia’s Changing Urban Landscape” (The World Bank, 2015)’)

➢ We agreed with the reveiwer’s comment. We changed the sentence like the below accordingly.

- (Study area and materials) Among East Asia cities, we used selected cities selected by Haraguchi et al. (2019). They chose the cities with more than 5 million people and a high degree of urbanization. Because rainfall often causes numerous floods in the region, and with growing urbanization, more and more cities suffer from small and medium-sized frequent floods, as well as large-scale low-probability floods (The World Bank, 2015).
- (Reference) Haraguchi, M., Kim, S. and Lall, U.: Correlated Risks for Heavy Precipitation in Mega-cities in East Asia, in: American Geophysical Union, Fall Meeting 2019, Meeting in San Francisco, U.S.A, 8-14 December, abstract #GC43G-1316, 2019.
  Disaster Risk Management Overview: https://www.worldbank.org/en/region/eap/brief /disaster-risk-management, last access: 29 May 2015.

5) How did you manage to gather the data from countries outside China ?

➢ In the research, we used the APHRODITE (Asian Precipitation-Highly-Resolved Observation Data Integration Towards Evaluation of Water Resources). The APHRODITE is a long-term daily grid precipitation dataset that combines rain gauge data, remotely sensed data, and geographic information over Asia. International Research Institute for Climate and Society (IRI) at Columbia University provides the data on a 0.25-degree grid for Asia. Therefore, rainfall data of countries outside China were collected from grid data.

6) Please expand your figure cpation tell us what the figure is trying to show the reader, before searching in the text. This comments applies to all figures and some of the more complex tables.

➢ We remembered this comment in the first revision stage. While we changed almost all captions in the table and figures, we did not change the caption in figure 1. In this time, we checked all captions again and revised some captions. We hope we reflect the comment correctly on the captions.

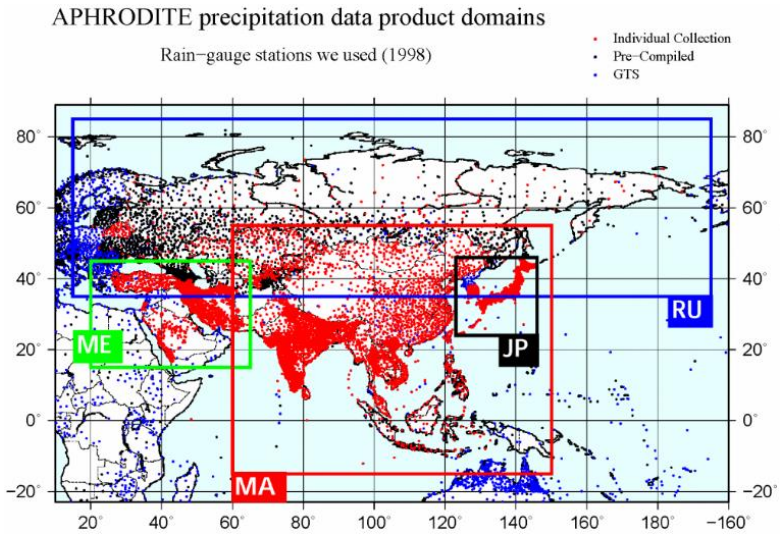


Figure 1. APHRODITE data (http://aphrodite.st.hirosaki-u.ac.jp/products.html) of cities; Dots are example of st ation data distribution that were used for making data(Red: individual collections, Black: pre-complied data, Blu e: global telecommunication networks (GTS) based data obtained from Global summary of day (NCEI/NOAA)). Rectangles show the domain of precipitation data. In this research, precipitation data on a 0.25 degree grid for Monsoon Asia (MA) for the period from 1981 to 2015.

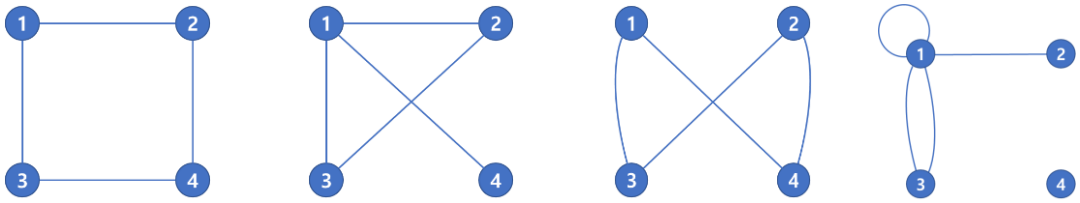


Figure 2. Various shapes of networks with the same number of nodes and links; Each network has 4 nodes an d 4 links. However, those shows different shapes and have different topological characteristics.

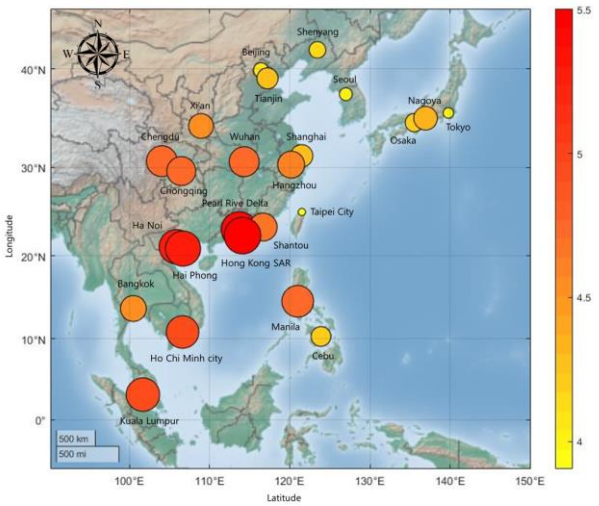


Figure 3. Adjacency information entropy value of cities; color and size of circle are respectively proportional to the entropy and rank. The right side of bar shows the adjacency information entropy values of nodes; except for Taipei city, nodes near the South China Sea, which had higher values;

7) Comments about Table 1 : Use 3 significant figure for all

- About Table 1, there were several problems. Therefore, we wrote all values in the table with three decimal places and changed the font to 'Times New Roman'.

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	2.160	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	5.900	2.105	4.619
Bangkok	1.101	4.067	3.694	7.308
Tianjin	3.903	9.932	2.544	4.910
Shantou	2.211	4.853	2.195	5.279
Chengdu	3.782	6.241	1.650	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	6.460	1.321	2.177
Shenyang	1.590	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	7.970	1.286	2.311
Xi'an	1.527	4.099	2.684	5.054
Ha Noi	4.053	8.870	2.189	4.417
Chongqing	2.790	5.751	2.061	4.890
Cebu	4.126	6.759	1.638	4.923
Hai Phong	3.801	9.530	2.507	5.023

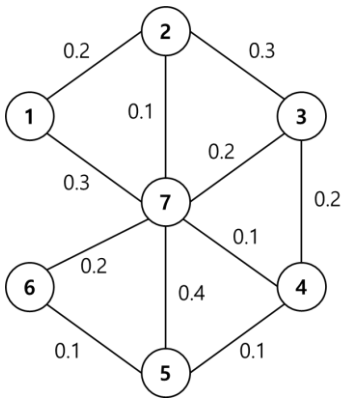
8) Why mutual information in equation (1) is expressed as I(A,B), not MI(A,B) ?

- The papers (for example, Donges et al. (2009), Ghorbani et al. (2021)) about mutual information usually express mutual information as I(A,B). So we used the same expression in this paper. But if it confuses the readers, we changed I(A,B) into MI(A,B).

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

9) Comments about Vital node identification method : is the adjacency degree is same to inverse-distance weight, like Kringing ?

- A_i is not an inverse distance weight. It is just total weight of neighbor nodes of node i. Let me explain it with orginal paper (Xu et al., 2020). If there is a sample network like below,



degree of node must be calculated first. This network has weight. Threrefore, the degree is equivalent to the sum of link weight.

Node	1	2	3	4	5	6	7
k_i	0.5	0.6	0.7	0.4	0.6	0.3	1.6

By using degree of node, Adjacency degree (A_i) can be calculated like below :

$$A_1 = k_2 + k_7 = 0.6 + 1.6 = 2.2$$
$$A_2 = k_1 + k_3 + k_7 = 0.5 + 0.7 + 1.6 = 2.8$$
$$A_3 = k_2 + k_4 + k_7 = 0.6 + 0.4 + 1.6 = 2.6$$
$$A_4 = k_3 + k_5 + k_7 = 0.7 + 0.6 + 1.6 = 2.9$$
$$A_5 = k_4 + k_6 + k_7 = 0.4 + 0.3 + 1.6 = 2.3$$
$$A_6 = k_5 + k_7 = 0.6 + 1.6 = 2.2$$
$$A_7 = k_1 + k_2 + k_3 + k_4 + k_5 + k_6 + k_7 = 0.5 + 0.6 + 0.7 + 0.4 + 0.6 + 0.3 = 3.1$$

As you can see, A_i is sum of degree of neighbors nodes to node i. If node i has many links or links with high weight, it has high value of A_i. It would be better put examples of Vital node identification for authors. But this can make paper longer and detailed explanation is already in the reference paper. So we recommend to authors for checking the reference paper like bleow

- The procedure of the method is as follows (more detailed explanation about the Vital node identification method is in Xu et al. (2020)):
  First. Calculate degree(k_i) of each node in the network

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

  Here, j is the neighbor of node i. Γ_i is the set of neighbors of node i. 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.

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

  A_i is total weight of neighbor nodes of node i.

10) Comments about Multiresolution detection method : does the path in Line 41 mean the shortest path ?

➢ The path in calculating link intensity is 'Distinct path'. k-distinct path can be denotead path between node i and j with k-edges if the path has no identical intermediate nodes or edges with any other distinct paths. We added this explain into 3.3.

- The First, Define Distinct path
  The simple(not repeating links between nodes) and elementary (not repeating nodes) path θ between node i and j with k-edges is denoted as a k-edge distinct path, if the path has no identical intermediate nodes or edges with any other distinct paths.
  Second, calculate the link intensity (I_P) of each link.

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

  Here, $\sigma(path_p(v_i, v_j))$ is the sum of link weights in p-edge distinct paths 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.
  Third, identify the links with link intensity greater than the selected threshold and create a group of nodes with the identified links.

$$v_j \in V, \quad I_P(e_{ij}) > t$$
$$v_j \in c_u, \quad I_P(e_{ij}) > t \quad - \quad (7)$$

  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).

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

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

11) Comments about Table 2 : Suggest the table should be able to fit the whole table on one page instead of overflow.

- We changed the table's width and put Talbe 2 on one page.

Table 2. Average, maximum, and minimum link weights of each node; Parentheses next to 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)

12) About mutual information results, certain cities have the smallest value with other cities. In the paper, author told that this result came from the location of cities. However, the reason for the result is judged to be due to climate, not location. Therefore, authoer need to elaborate about it.

- When looking at the cities in the Minimum column of Table 2, except for Taipei, cities are located in the outskirt of the study area. Regarding their location, we initially hypothesized that their locations would cause different rainfall patterns from cities in central, which makes low value in mutual information. However, after receiving the above comment, we agreed that another potential factor might be climatic. We will mention that this is one of the limitations of our study and that future studies can explore this. We propose that future studies analyze each city's rainfall characteristic with geographical data and compare them to find a mechanism behind these different rainfall patterns. In the revised version, we changed Line 162 related to the comment like below and put this problem in the discussion part.

- (Line 162) Two cities have a common feature that their location is in the outskirt of the study area.
- (Discussion) In the mutual information result, we could find that many cities had the lowest value with Tokyo and Beijing. We tried to find why this result came out, but we could not find any differences in the rainfall data. Therefore, future studies should collect and analyze other climate factors and geographical factors for finding unique rainfall characteristics in Tokyo and Beijing.

13) It would be helpful put the reference first line in 4.2 section.

- Thanks for your comment about putting reference into the first line of 4.2 section. However, we already put the reference, which the reviewer recommended to us, in 3.2 section (3.2 Vital node identification using adjacency information entropy). In section 3.2, readers can recognize which paper has background and detailed information about the vital node identification method.

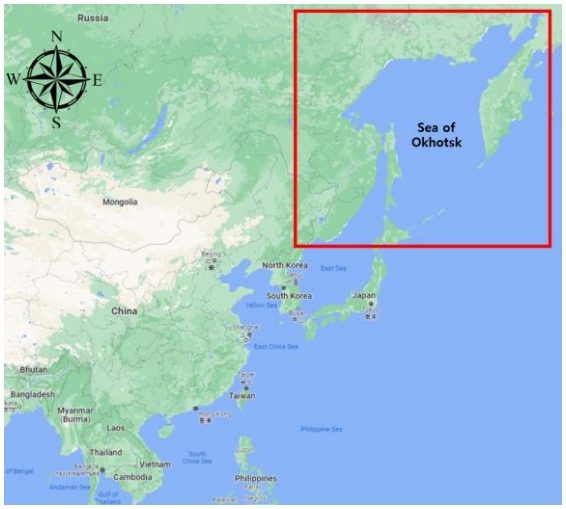
14) Is ″95th quartile″ right express? Quartiles are 25%, 50% and 75%.

- We did not correctly express the word ″95th quartile″. Quantiles are points in a distribution related to the rand order of values in that distribution. On the other hand, quartile refers to the point where the data is divided into quarters, as the reviewer commented. Therefore, we accordingly fixed this, and now we express ″95th quantile″.

  - (Line 179) The threshold value is the 95th quantile of the calculated belonging coefficient, 0.06, in order to form a group of nodes with strong relationship.

15) Where is location of sea of Okhotsk?

- The sea of Okhotsk is marginal sea of the western Pacific ocean. It is located between Russia's Kamchatka penisula on the east, the Kuril islands on the southeast, Japan's island of Hokkaido on the sotuh, the island of Sakhalin along the west, and a stretch of eastern Siberian coast along the west and north. We tried to express the location of the sea of Oktotsk in the figures but it is far from study area. Therefore we added the location in Line 211.



< Location of Sea of Okhotsk (from Google map) >

- (Line 211) The Indian monsoon moves northwest from the Bay of Bengal, passing mainland China into the Sea of Okhotsk, which is located between the Russia's Kamchatka peninsula and Japan's island of Sakhalin.

16) Comment on Reference: Please separate each article, and/or ident the lines after the lead one. This will make it easier for the reader to find one referred to in the text.

- Now we clearly understood and revised all list of references accordingly.

17) Comment on English expression

- We checked all the comments related to English expression. We appreciate comments on English expression. Once again, we checked the English and appropriately revised the whole manuscript. These changes can be seen with the track-change version of our revised paper.