

Entropy-Based Dynamic Graph Embedding for Climate Change Detection

Gen Li

Chung-Ang University

Jason Jung (✉ j2jung@gmail.com)

Chung-Ang University

Research Article

Keywords: EDynGE, climate, graph

Posted Date: December 29th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-133833/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on July 5th, 2021. See the published version at <https://doi.org/10.1038/s41598-021-92973-8>.

Entropy-based dynamic graph embedding for climate change detection

Gen Li¹ and Jason J. Jung^{1,*}

¹Department of Computer Engineering, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, Republic of Korea 06974

*j3ung@cau.ac.kr, j2jung@gmail.com

ABSTRACT

Climate change is a severe problem caused by abnormal climate events. The existing methods for detecting climate changes utilize statistical models to analyze the atmospheric temperature, but a climate event commonly comprises multiple meteorological data. To detect climate changes using meteorological data, we propose a novel dynamic graph embedding model based on graph entropy called EDynGE. A climate event is denoted as a graph, in which the nodes indicate meteorological data and edges indicate the correlation between nodes. Graph entropy measures the information of the climate event, and the EDynGE model clusters graphs based on graph entropy. We conducted experiments on real meteorological data. The results showed that the number of days of abnormal climate events has increased by 304.5 days in the past 30 years.

Introduction

Climate change is a severe problem that leads to the redistribution of global precipitation, melting of glaciers, and rise in sea levels^{1,2}. Furthermore, it endangers the balance of the natural ecosystem and threatens the survival of humans. The main reason for climate change is that the terrestrial greenhouse gas emissions cause the atmospheric temperature to rise on the mainland³. Existing research shows that in the 20th century, the world's average temperature showed an upward trend⁴. Therefore, climate change detection can help people identify the causes of ecosystem damage and suggest corresponding countermeasures⁵.

Existing data analysis methods for detecting climate changes are based on statistical models that identify temporal and spatial information from meteorological data⁶. As a simple example, most of the statistical analysis usually measures the average temperature and compares it with past climate conditions to detect whether the climate is changing. These methods focus on detecting climate change using single meteorological data. Climate events comprise multiple meteorological data, and the correlation of these data plays an essential role in climate change detection.

To solve this problem, we identified a spurious relationship between meteorological data in which two correlation data are not causally related⁷. A graph was constructed to model the climate event using these spurious relationships. The vertices of the graphs indicate meteorological data, whereas the edges indicate the spurious relationship between two vertices. To obtain information from the climate graph, we calculated the graph entropy using the spurious correlation coefficient. We constructed a dynamic graph embedding model based on graph entropy to cluster the climate graphs. Climate change was detected as an abnormal climate event in a time interval where the spurious correlation coefficients are different from most other time intervals, which is defined as follows.

Definition 1 (Abnormal climate event) *An abnormal climate event is defined as a graph G_i in the i^{th} time interval in which the weights of the edges are significantly different from those in other time intervals. It is formulated as $|e(G_i) - e(G_{i-1})| > \theta$ or $|e(G_i) - e(G_{i+1})| > \theta$, where $e(G_i)$ is the entropy of the climate event G_i and θ is the threshold for detecting climate change.*

Figure 1 shows the climate events in three time intervals where T indicates temperature, P indicates pressure, S indicates wind speed, G_i indicates the climate event in the i^{th} time interval, and the weight of an edge indicates the spurious correlation coefficient. The temperature at G_2 increases by 2 °C, which leads to a difference in the weights of edges and those of the other two time intervals. This indicates that the climate has changed at G_2 .

The main contributions of this work are as follows.

- We present the graph entropy to measure the information of climate events, which is calculated using the spurious correlation coefficient.
- We propose an entropy-based dynamic graph embedding model (EDynGE) that clusters climate events. Abnormal climate events are far from normal events and can be detected as climate change.

- We conducted experiments on real meteorological datasets using the EDynGE model. The results showed that the days of the abnormal climate events exhibit an upward trend from 1990–2020.

The remainder of this paper is organized as follows. In the next section, the methodology of entropy-based dynamic graph embedding for detecting climate change is detailed. Then, the experimental results are described. Finally, in the discussion section, the conclusions, limitations, and future works are provided.

Methods

This section describes the dynamic climate graph, graph entropy, and EDynGE model. A dynamic graph is used to model climate events in each time interval to detect climate change. The graph entropy measures information regarding the climate event. The meteorological datasets are allowed to collect from the China meteorological data service center (<http://data.cma.cn/en>) by registering an account.

Graph Construction

The graph is denoted as $G(V, E)$, where V and E denote the vertices and edges, respectively. For a climate event, the vertex indicates the meteorological data, the edge denotes the spurious relationship, and the weight w indicates the spurious correlation coefficient. The coefficient is calculated based on the causality and correlation between two time series, x and y . The time series causality is defined as follows.

Definition 2 (Time series causality) *The causality of two series is defined as that if one of the series improves the prediction of the other, which is formulated as*

$$C(x, y) = \begin{cases} 1 & \text{if } p \geq 0.05 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where x and y are two time series, $C(x, y)$ indicates the causality between them, and p is the probability that the two series are not causally related.

The Granger causality test is utilized to calculate the short-run causality between two meteorological time series, x and y ⁸. The test makes a null hypothesis that the two series are not causally related and includes two predictions. Firstly, it uses the past values of the series y as variables to predict the current y . Then, it uses the past values of series x and y as variables to predict the current y . If the prediction result obtained using the temporal information of two series x and y is better than the prediction only using the series y , then x helps predict y . The t-test was utilized to compare the difference between two prediction results⁹. The p value was used to denote the probability of the null hypothesis. If the p value was more than 0.05, then the two series x and y were said to not be causally related¹⁰.

The weight value of an edge in the graph indicates the spurious correlation coefficient that is calculated using the causality and Pearson correlation coefficient (PCC)¹¹, which is formulated as follows.

$$R(x, y) = \begin{cases} 1 & \text{if } C(x, y) = 0 \\ C(x, y) - |PCC| & \text{otherwise} \end{cases} \quad (2)$$

where $C(x, y)$ indicates the causality between the two series x and y . $R(x, y)$ indicates the spurious correlation coefficient between the two series. The spurious correlation coefficient is inversely proportional to the causality between the time series x and y . If the causality $C(x, y)$ between two series x and y is 0, then there is no spurious correlation between the two series. In this case, the corresponding spurious correlation coefficient $R(x, y)$ is 1.

Graph Entropy

The graph entropy is calculated based on information entropy¹². We assume that there are two independent events, x and y . The information of these events should be satisfied as $h(x, y) = h(x) + h(y)$, where $h(x)$ indicates the information of the event x , and $h(x, y)$ indicates the information of these two events occurring at the same time. The probability of these events should be satisfied as $p(x, y) = p(x) \times p(y)$, where p indicates the probability of the event. The information of the event x can be measured as $h(x) = -\log_2 p(x)$. Information entropy can be represented as the information of the event x times the probability of x , which is formulated as $e(x) = -p(x) \log_2 p(x)$. For a set of events X , the information entropy is formulated as $e(X) = -\sum_{i=1}^N p(x_i) \log_2 p(x_i)$, where N indicates the number of events in the set and x_i indicates the i^{th} event.

To calculate the graph entropy, we calculated the entropy for each vertex in the graph. The definition of the vertex entropy is defined as follows.

Definition 3 (Vertex entropy) Given a graph $G = (V, E)$, the entropy of the vertex v_i is defined based on the weight between the vertices v_i and v_j , which is formulated as $e(v_i) = \sum_{j=0, j \neq i}^N -w_{i,j} \log_2 w_{i,j}$, where N indicates the number of vertices. The weight value $w_{i,j}$ equals $R(v_i, v_j)$, which denotes the spurious correlation coefficient between two vertices v_i and v_j .

The graph entropy is calculated by summing the entropy of all vertices, which is formulated as $e(G) = \sum_{i=0}^N e(v_i)$. The dynamic graph entropy is composed of the graph at time interval $t \in [0, T]$, which is formulated as $\mathcal{E} = \{e(G_t) | t \in [0, T]\}$. The information of the climate event can be quantified using graph entropy. When one of the meteorological data points changes, the spurious relationship coefficients change, and the graph entropy also changes at the corresponding time interval. The abnormal climate event can be detected by obtaining graph entropy.

Entropy-Based Graph Embedding

The dynamic graph consists of graphs G_t in the time interval $t \in [0, T]$, which is formulated as $\mathcal{G} = \{G_t | t \in [0, T]\}$. Dynamic graph embedding is used to capture the temporal information of the dynamic graph \mathcal{G} for learning a mapping function $f: G_t \rightarrow g_t$, where g_t is an embedding vector of the graph G_t . The similarity of the entropy between the two graphs is formulated as $d(e(G_i), e(G_j)) = \sqrt{\|e(G_i) - e(G_j)\|_2^2}$. The object of the entropy-based graph embedding reduces the distance between two graphs with similar entropy. To address this problem, we construct a dynamic supervised graph, which is defined as follows.

Definition 4 (Dynamic supervised graph) Let $\mathcal{G} = \{G_t | t \in [0, T]\}$ denote the dynamic graph at time interval $t \in [0, T]$. For the graph G_t , the corresponding graph G_i can be found from the dynamic graph \mathcal{G} , where the similarity $d(e(G_i), e(G_t))$ of the entropy between the two graphs is the smallest. The dynamic supervised matrix is a set composed of the graph G_i , which is formulated as $\mathcal{S} = \{G_i | i \in [0, T]\}$.

As shown in Figure 2, the dynamic graph is formulated as $\mathcal{G} = \langle G_0, G_1, G_2 \rangle$. The entropy of vertex v_0 in the climate event G_0 is calculated as $e(v_0) = \sum_{j=0, j \neq i}^2 -w_{0,j} \log_2 w_{0,j} = -0.6 \times \log_2 0.6 - 0.4 \times \log_2 0.4 = 0.292$. The entropies of the vertices v_1 and v_2 in G_0 can be calculated as $e(v_1) = 0.241$ and $e(v_2) = 0.267$, respectively. The entropy of graph G_0 is $e(G_0) = 0.292 + 0.241 + 0.267 = 0.800$. The entropy of graphs G_1 and G_2 are $e(G_1) = 0.796$ and $e(G_2) = 0.848$, respectively. The similarity of the entropy between the graphs G_0 and G_1 is $d(e(G_0), e(G_1)) = 0.004$. The similarity between the graphs G_0 and G_2 is $d(e(G_0), e(G_2)) = 0.048$. Therefore, the nearest graph from G_0 is G_1 , the nearest graph from G_1 is G_0 , and the nearest graph from G_2 is G_0 . The dynamic supervised matrix is thus denoted by $\mathcal{S} = \langle G_1, G_0, G_0 \rangle$.

We utilize two autoencoders to reconstruct the dynamic graph and supervised graph. The two autoencoders share parameters with each other. Figure 3 shows the architecture of the EDynGE model, where G_t and S_t indicate the climate graph and the supervised graph, respectively. The embedding vectors of G_t and S_t are denoted as g_t and s_t , respectively. The autoencoder includes an encoder and decoder. We use y^i to indicate the i^{th} layer of the encoder, and \hat{y}^i is used to denote the i^{th} layer of the decoder. The autoencoder reconstructs the input data using the encoder and decoder to calculate the graph's embedding vector. The encoder uses non-linear functions to extract the features for mapping the graphs into the embedding space, which are formulated as

$$y^1 = \delta(W^1 G_t + b^1) \quad (3)$$

$$y^i = \delta(W^i y^{i-1} + b^i) \quad (4)$$

where δ is an activation function. W^i and b^i indicate the weight and basis in the i^{th} layer, respectively. The ReLU function is utilized as the activation for making the neural network non-linear, which is formulated as $f(y^i) = \max(0, y^i)$ ¹³. The decoder reconstructs the graph from the embedding vector, which is calculated by reversing the encoder's computation.

The purpose of dynamic embedding is to reduce the distance between two graphs that have a similar entropy in an embedding space. Therefore, we establish a loss function based on the similarity of the graph entropy in the embedding layer, which is formulated as $\mathcal{L}_s = \frac{1}{T} \sum_{t=1}^T \|g_t - s_t\|_2^2$, where T indicates the number of time intervals. The graph G_t and supervised graph S_t have the smallest similarity on graph entropy. Thus, the function \mathcal{L}_s reduces the loss between a_t and s_t to reduce the distance between two graphs in the embedding space. An autoencoder is used to reconstruct the input so that we have to establish a loss function for reducing the loss between the input and output, which are formulated as $\mathcal{L}_1 = \frac{1}{T} \sum_{t=1}^T \|G_t - \widehat{G}_t\|_2^2$. To avoid overfitting, we establish a regularization term that is formulated as $\mathcal{L}_{reg} = \frac{1}{2} \sum_{i=0}^I (\|W^i\|_2^2 + \|\widehat{W}^i\|_2^2)$, where W^i and \widehat{W}^i indicates the weight of the encoder and decoder in the i^{th} layer. The joint loss function is established using the functions \mathcal{L}_s , \mathcal{L}_1 , and \mathcal{L}_{reg} , which is formulated as

$$\mathcal{L} = \frac{1}{T} \sum_{t=1}^T \|g_t - s_t\|_2^2 + \frac{1}{T} \sum_{t=1}^T \|G_t - \widehat{G}_t\|_2^2 + \frac{1}{I} \sum_{i=0}^I (\|W^i\|_2^2 + \|\widehat{W}^i\|_2^2) \quad (5)$$

We utilize the gradient descent algorithm and backward propagation algorithm^{14,15} to train the model. Gradient descent is used to calculate the weight and basis in the output layer, which are formulated as $W^I = W^I - \eta \frac{\partial \mathcal{L}}{\partial W^I}$ and $b^I = b^I - \eta \frac{\partial \mathcal{L}}{\partial b^I}$, where I indicates the output layer. Each layer's weight and basis are calculated using the backward propagation algorithm, which calculates the partial derivation of the loss function based on the chain rule for updating each layer's weight and basis.

Results

In this section, we apply the EDynGE model to real meteorological data and use local outlier factor (LOF)¹⁶, isolation forest (IF)¹⁷, and box-plot (BP)¹⁸ methods to detect the abnormal climate events in an embedding space. IF shows that the distribution of outliers is sparse, and these outliers are far away from the normal observations with high density. Thus, the outliers can be easily separated. LOF detects outliers based on the density of the data points. The BP method is based on statistical indices for detecting outliers and requires the dataset to have a normal distribution.

Dataset

Daily climate data from the Chinese surface stations of 10 provinces were used to conduct experiments. According to the nationwide surface climate statistical method¹⁹, these datasets were derived from various provincial meteorological bureaus through statistical compilations. The datasets were collected from 194 basic and reference surface meteorological observation stations and automatic weather stations in China from 1951. Each dataset included 18 elements, including mean pressure, mean temperature, and precipitation. In this study, we collected meteorological data from 1990–2020 to evaluate the EDynGE model.

Evaluation Metrics

Because the datasets are unlabeled, we propose using two different ways to evaluate the EDynGE model. The first way is to label a certain number of data points as outliers. These data points are embedded vectors of climate events. The selection rules for these outliers are as follows. We assume that 10% of data points are selected as outliers in each dataset. The embedding vector of the t^{th} graph is denoted as g_t . The center of the embedding vectors can be formulated as $c = \frac{1}{T} \sum_{t=1}^T g_t$, where T is the number of time intervals. If the entropies of the graphs are similar, the embedding vectors of these graphs are close to each other, and the outliers are far from the normal observations. The 10% data points farthest from the center is selected as outliers. The EDynGE model can be evaluated using accuracy and F-score.

In the second way, we propose a hypothesis based on global warming that with increasing temperature, the number of days of abnormal climate events also increases. We counted the number of days with abnormal climate events every year, every five years, and every decade. If the number of days of abnormal climate events exhibited an upward trend, then climate was shown to have changed in the past 30 years. EDynGE can be used to detect this climate change.

Analysis of Results

Figure 4 shows the days of abnormal climate events in the four provinces obtained using the IF method. The results of every year show that the frequency of abnormal climate events exhibits an increasing trend. We calculated the days of abnormal climate events every five years, the results showed that four provinces exhibited a non-linear increasing trend. However, a local minimum value in Guangzhou and Shanghai was observed from 2005–2010, and Beijing had a local minimum value from 2000–2005. The results of every decade indicated that three provinces Beijing, Shandong, and Shanghai showed an upward trend. Guangzhou showed a falling trend first followed by a rising trend. According to the experimental results, the detected climate change results conform with the hypothesis in most cases.

To conduct a comparison, we utilized a graph convolutional neural network (GCN) and a dynamic graph to a vector-based model (dyngraph2vec) as baselines^{20,21}. The GCN uses convolutional kernels to capture the spatial information of vertices in the graph. Because GCN is applied to the static graph, it does not consider the temporal information of the dynamic graph. Dyngraph2vec is an unsupervised learning model for embedding dynamic graphs. It provides the three dyngraph2vec-based models, which are autoencoders (dyngraph2vecAE), recurrent neural network (dyngraph2vecRNN), and the autoencoder-based recurrent neural network model (dyngraph2vecAERNN). Dyngraph2vecAE cannot extract the temporal information from the dynamic graph since the model computes the embedding vectors by reconstructing the graphs. Dyngraph2vecRNN utilizes the idea of the skip-gram to consider the temporal information of the graphs²². It computes the embedding vector of the current graph by using the graphs around the current graph. Table 1 shows the accuracy of the models using the IF method under 10% outliers. According to the experimental results, the EDynGE model exhibits the best performance in all provinces. GCN achieves better accuracy than dyngraph2vecAE in 6 provinces. The dyngraph2vecAE method achieves better accuracy than the other two dyngraph2vec models. Because GCN, AE, and EDynGE do not capture the temporal information of the dynamic graph and they perform better than the other two models that extract the temporal features, the effect of temporal information is negligible.

Table 2 shows the performance of the EDynGE model with 10% outliers. According to the results, the IF method achieved the best accuracy and F1-score for the eight provinces. On the other hand, it achieved a lower accuracy and F1-score than the LOF method for Beijing and Shanxi. The BP method achieved a better F1-score than LOF for seven provinces, but the scores were worse than those obtained for IF. Overall, the experimental results show that IF exhibited the best performance for most provinces.

To obtain performances with different embedding sizes, we evaluated the stability of the EDynGE model using $mean \pm std$, where std indicates the standard deviation. Figure 5 shows the stability of the EDynGE model in Beijing. According to the experimental results, the IF method achieved the best performance based on the average F1-score and had the lowest std among the three outlier detection methods. This indicates that IF is the most stable method compared with the other two methods.

We conducted experiments to validate the EDynGE model with different ratios of outliers. Figure 6 shows the performance of the model on Beijing province by choosing three different ratios of outliers. According to the experimental results, the IF method achieved the best accuracy for 10% outliers. The LOF and BP methods achieved the best performance for 5% outliers. If the ratio of outliers was low, the imbalance of labels reduced the performance of the EDynGE model. Therefore, the three methods showed the lowest accuracy under 3% outliers.

Discussion

In this study, we proposed an EDynGE model to detect climate changes. The model uses the spurious correlation coefficient to calculate the graph entropy and reduces the distance between two climate events with similar graph entropy. We conducted experiments to validate the performance and stability of the EDynGE model for climate change detection. The results showed that the IF method exhibited better results than the LOF and BP methods by 32.5% and 12.7% in terms of F1-score, respectively. The EDynGE model performed better than the other dynamic graph embedding models by 37.9% in terms of accuracy. Based on global warming, we hypothesized that with an increase in temperature, the number of days of abnormal climate event exhibits an upward trend. The experimental results showed that the number of days of climate change increased by 304.5 days from 1990–2020, which agreed with the hypothesis. This indicates that the EDynGE model can detect climate change.

This study has some limitations. The EDynGE model can cluster graphs based on graph entropy. However, graphs with different neighbor structures have the same graph entropy in some cases. The EDynGE model cannot detect outliers with an abnormal neighbor structure. This indicates that the EDynGE model ignores the spatial information of the dynamic graph. To overcome this issue, we plan to construct a hybrid model that consists of the neighbor structure similarity and graph entropy similarity for detecting outliers in multiple time series. The second limitation is that the EDynGE model is based on an autoencoder that does not capture the temporal information from the dynamic graph. Although the temporal information is negligible in the research problem, it also needs to be considered in dynamic graph embedding. To overcome this problem, we plan to construct an autoencoder model using the long short-term memory architecture to discover temporal features from the dynamic graph.

We propose a novel idea to detect outliers from multiple time series. It utilizes the correlation of the time series to construct a dynamic graph and detects the outlier from the dynamic graph. The outlier detection problem is transformed from the multiple time series domain to the dynamic graph domain. It can help people find the causes of the outliers by obtaining the evolution of graphs. For example, in the financial time series, abnormal trends in the stock market can be detected and analyzed using the EDynGE model. Furthermore, the digital twin technology is developing rapidly. It utilizes the sensors to record the digital information for simulating the condition of object in the physical space. The proposed is able to detect the anomalies from the recorded digital signals to diagnose faults from the physical. For example, the transmission failure in the machines and the structural damage in the buildings can be detected by using the proposed idea.

References

1. Nerem, R. S. *et al.* Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proc. Natl. Acad. Sci.* **115**, 2022–2025 (2018).
2. Parker, V. T. & Boyer, K. E. Sea-level rise and climate change impacts on an urbanized pacific coast estuary. *Wetlands* **39**, 1219–1232 (2019).
3. Shen, M. *et al.* (micro) plastic crisis: Un-ignorable contribution to global greenhouse gas emissions and climate change. *J. Clean. Prod.* **254**, 120138 (2020).
4. Diffenbaugh, N. S. & Burke, M. Global warming has increased global economic inequality. *Proc. Natl. Acad. Sci.* **116**, 9808–9813 (2019).
5. Creutzig, F. *et al.* Towards demand-side solutions for mitigating climate change. *Nat. Clim. Chang.* **8**, 260–263 (2018).

6. Lobell, D. B. & Asseng, S. Comparing estimates of climate change impacts from process-based and statistical crop models. *Environ. Res. Lett.* **12**, 015001 (2017).
7. Runkler, T. A. Data analytics. *Wiesbaden: Springer. doi* **10**, 61–68 (2012).
8. Akadiri, S. S., Lasisi, T. T., Uzuner, G. & Akadiri, A. C. Examining the causal impacts of tourism, globalization, economic growth and carbon emissions in tourism island territories: bootstrap panel granger causality analysis. *Curr. Issues Tour.* **23**, 470–484 (2020).
9. Pearson, K. Contributions to the mathematical theory of evolution. *Philos. Transactions Royal Soc. London. A* **185**, 71–110 (1894).
10. Nuzzo, R. Scientific method: statistical errors. *Nat. News* **506**, 150 (2014).
11. Benesty, J., Chen, J., Huang, Y. & Cohen, I. *Noise Reduction in Speech Processing*, (Springer Science & Business Media, 2009).
12. Pele, D. T., Lazar, E. & Dufour, A. Information entropy and measures of market risk. *Entropy* **19**, 226 (2017).
13. Schmidt-Hieber, J. *et al.* Nonparametric regression using deep neural networks with relu activation function. *Annals Stat.* **48**, 1875–1897 (2020).
14. Ruder, S. An overview of gradient descent optimization algorithms. Preprint at: <https://arxiv.org/abs/1609.04747> (2016).
15. Cao, Q. & Parry, M. E. Neural network earnings per share forecasting models: A comparison of backward propagation and the genetic algorithm. *Decis. Support. Syst.* **47**, 32–41 (2009).
16. You, L. *et al.* Integrating aspect analysis and local outlier factor for intelligent review spam detection. *Futur. Gener. Comput. Syst.* **102**, 163–172 (2020).
17. Karczmarek, P., Kiersztyn, A., Pedrycz, W. & Al, E. K-means-based isolation forest. *Knowledge-Based Syst.* 105659 (2020).
18. Williamson, D. F., Parker, R. A. & Kendrick, J. S. The box plot: a simple visual method to interpret data. *Annals internal medicine* **110**, 916–921 (1989).
19. National meteorological information center. China meteorological data service center dataset. <http://data.cma.cn/en>.
20. Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M. & Monfardini, G. The graph neural network model. *IEEE Transactions on Neural Networks* **20**, 61–80 (2008).
21. Pareja, A. *et al.* Evolvegen: Evolving graph convolutional networks for dynamic graphs. In *AAAI*, 5363–5370 (2020).
22. Guthrie, D., Allison, B., Liu, W., Guthrie, L. & Wilks, Y. A closer look at skip-gram modelling. In *LREC*, vol. 6, 1222–1225 (2006).

Acknowledgements

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIP) (NRF-2020R1A2B5B01002207).

Author contributions statement

Conceptualization, G. L.; methodology, G.L.; software, G.L.; validation, G.L. and J.J.J.; formal analysis, G.L.; investigation, G.L.; resources, G.L.; data curation, G.L.; writing—original draft preparation, G.L.; writing—review G.L. and J.J.J.; supervision, J.J.J.; project administration, J.J.J.; funding acquisition, J.J.J. All authors have read and agreed to the published version of the manuscript.

Additional information

The authors declare no conflicts of interest.

Data availability

The real-meteorological datasets used in this study are available to be collected from the China Meteorological Data Service Center and allow the authors to download the data from the website <http://data.cma.cn/en> by registering an account.

Figure legends and Tables

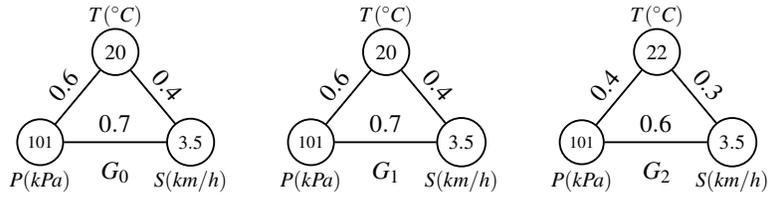


Figure 1. Example of an abnormal climate event.

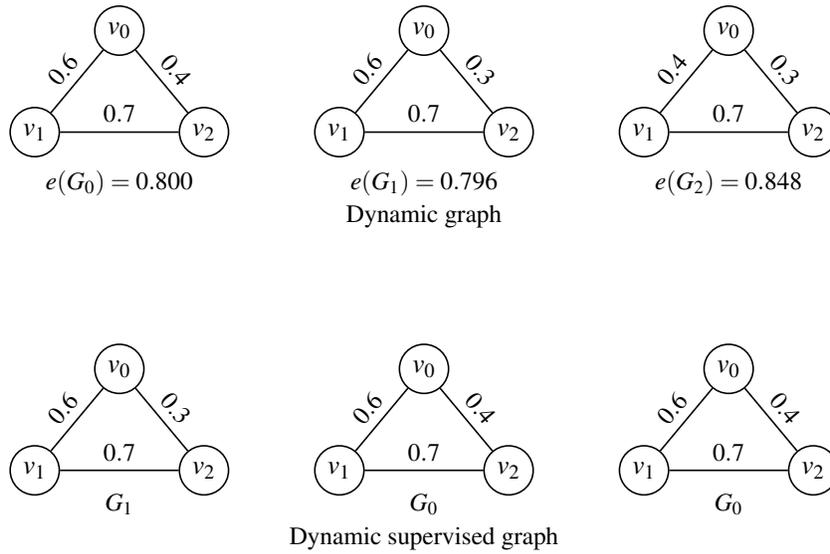


Figure 2. Example of the dynamic supervised graph.

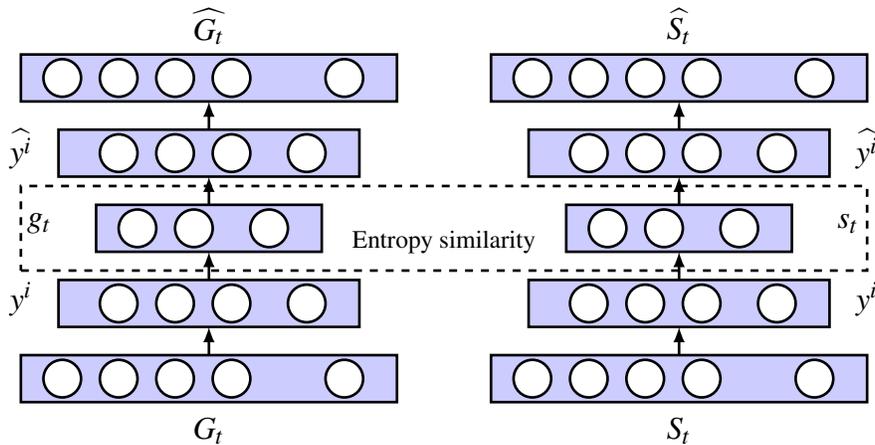


Figure 3. Architecture of EDynGE model.

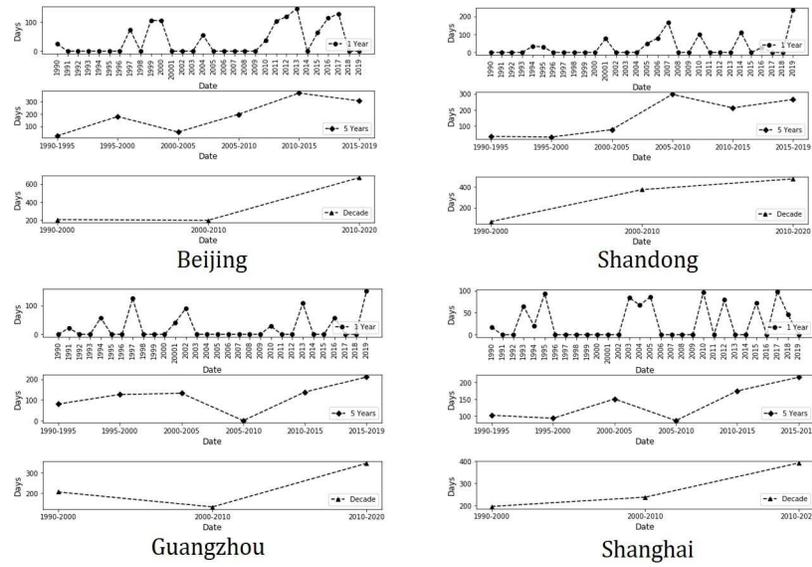


Figure 4. Days of abnormal climate events in the four provinces.

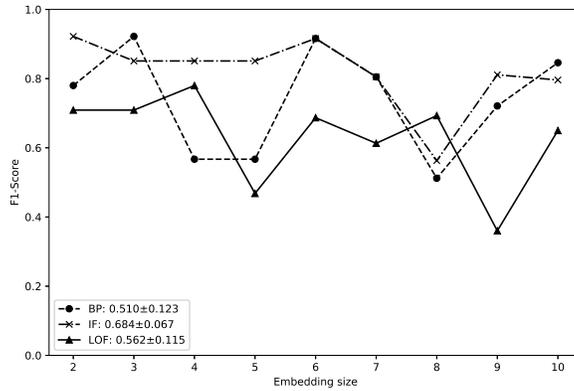


Figure 5. Stability of EDynGE model under the 10% outliers.

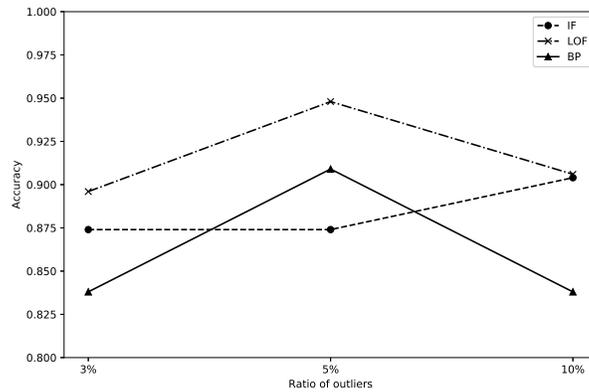


Figure 6. Performance for the city of Beijing.

Table 1. Comparison experiments under 10% outliers.

	GCN	AE	RNN	AERNN	EDynGE
Beijing	0.904	0.849	0.590	0.512	0.942
Shanghai	0.865	0.912	0.554	0.566	0.953
Guangzhou	0.741	0.832	0.423	0.570	0.953
Shandong	0.892	0.919	0.581	0.453	0.959
Hebei	0.922	0.860	0.680	0.539	0.959
Jiangsu	0.836	0.793	0.679	0.621	0.950
NeiMongol	0.894	0.785	0.646	0.613	0.960
Shanxi	0.908	0.824	0.632	0.644	0.941
Zhejiang	0.772	0.813	0.585	0.528	0.967
Tianjin	0.911	0.883	0.509	0.645	0.967

Table 2. Performance of the EDynGE model with 10% outliers.

	Accuracy			F1-score		
	IF	LOF	BP	IF	LOF	BP
Beijing	0.942	0.948	0.909	0.709	0.780	0.709
Guangzhou	0.953	0.940	0.860	0.922	0.709	0.780
Shanghai	0.953	0.939	0.878	0.851	0.709	0.922
Shandong	0.959	0.953	0.953	0.851	0.780	0.567
Hebei	0.959	0.899	0.953	0.851	0.468	0.567
Jiangsu	0.950	0.943	0.871	0.916	0.687	0.916
NeiMongol	0.960	0.934	0.825	0.805	0.613	0.805
Shanxi	0.941	0.950	0.887	0.563	0.693	0.512
Zhejiang	0.967	0.943	0.943	0.811	0.360	0.721
Tianjin	0.967	0.944	0.855	0.796	0.650	0.846

Figures

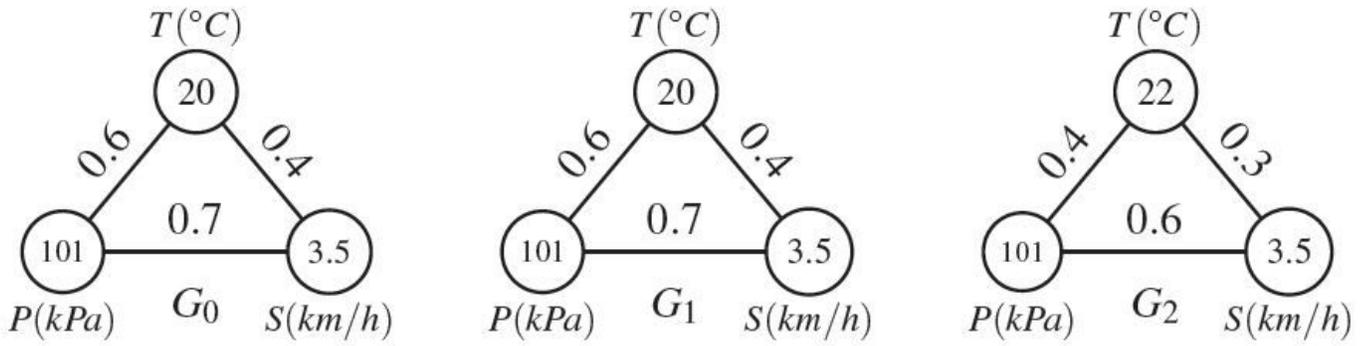


Figure 1

Example of an abnormal climate event.

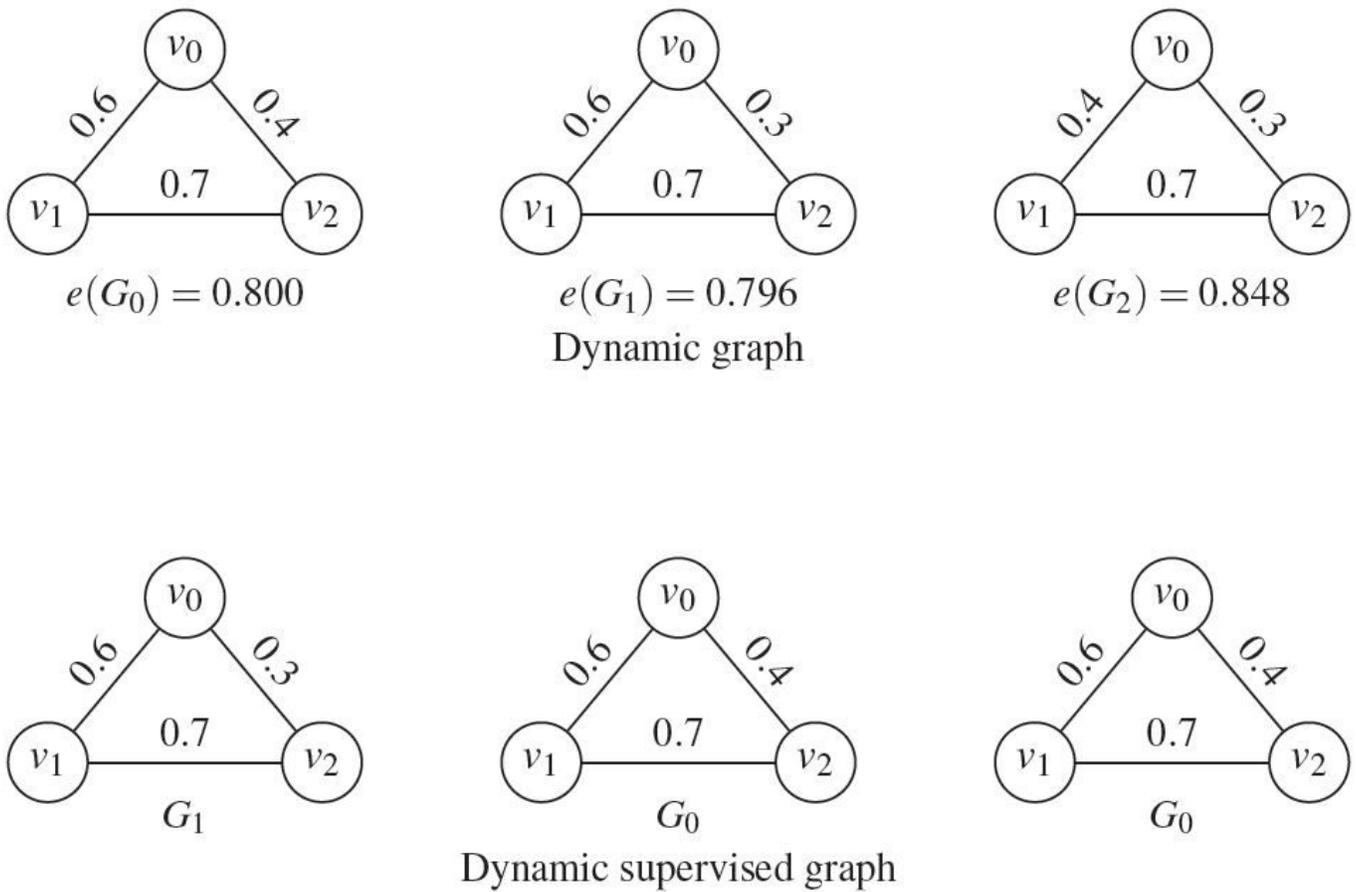


Figure 2

Example of the dynamic supervised graph.

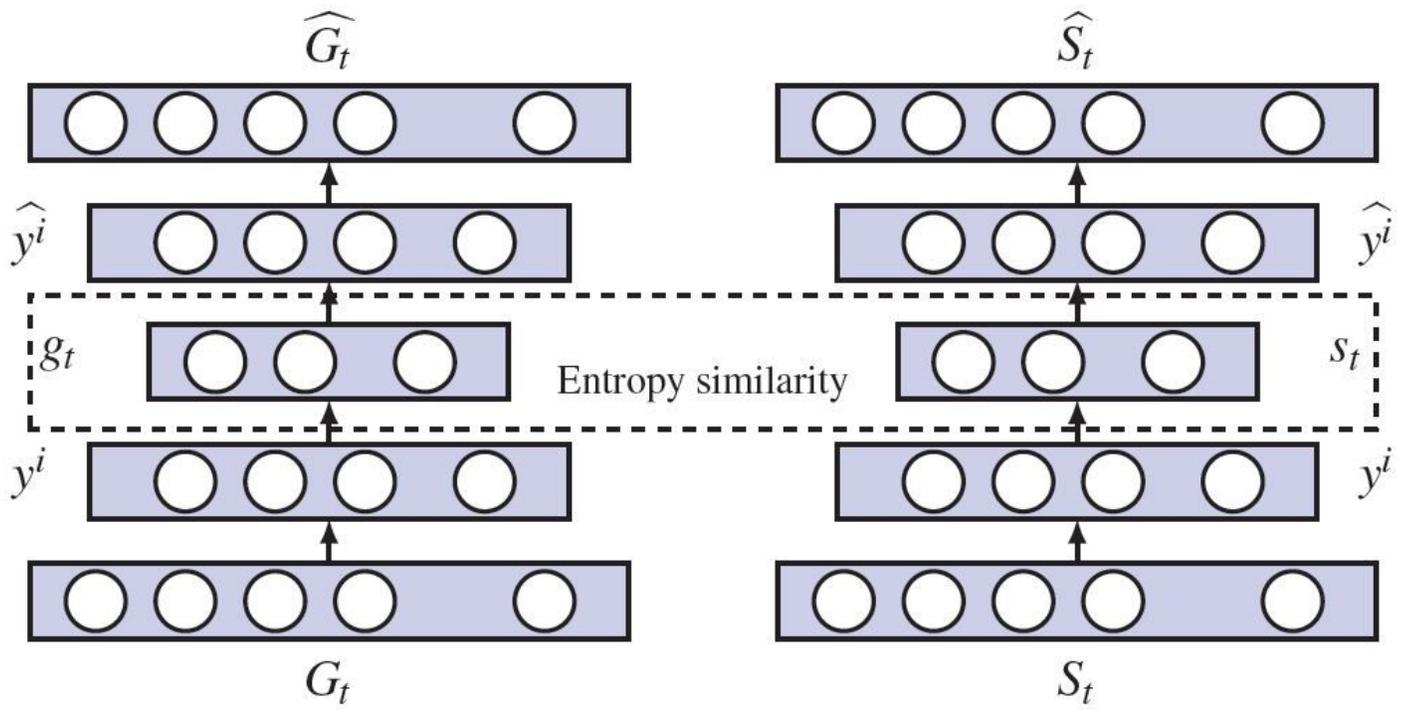
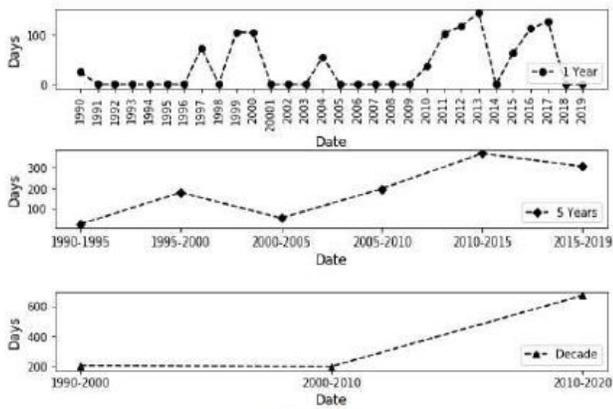
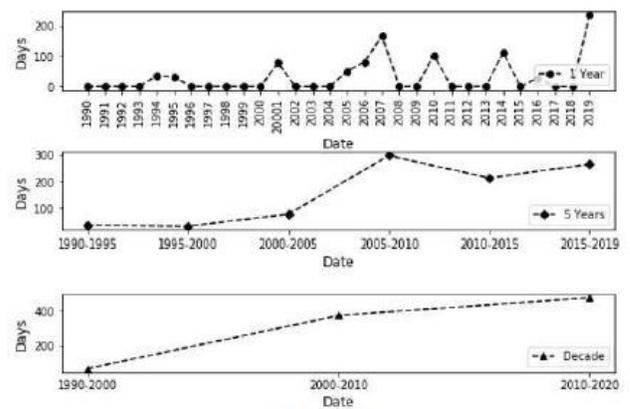


Figure 3

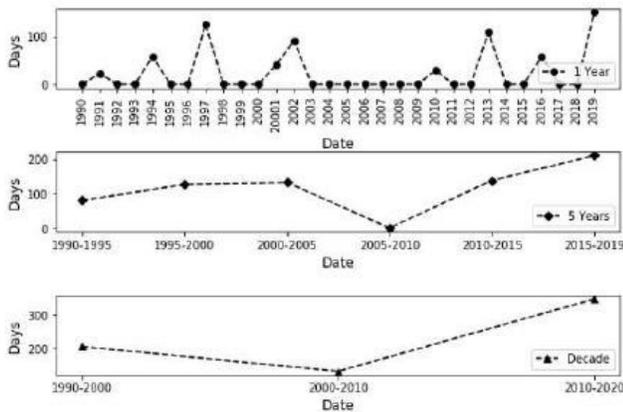
Architecture of EDynGE model.



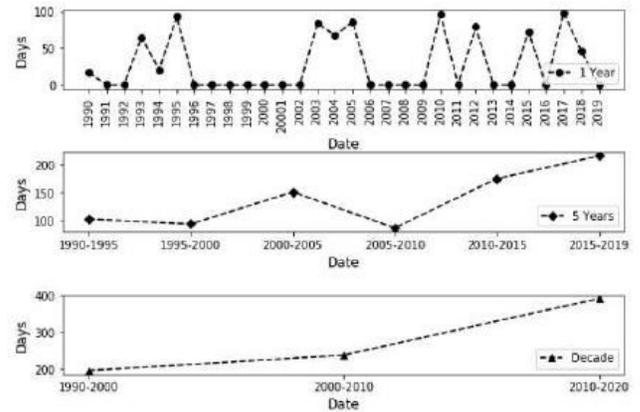
Beijing



Shandong



Guangzhou



Shanghai

Figure 4

Days of abnormal climate events in the four provinces.

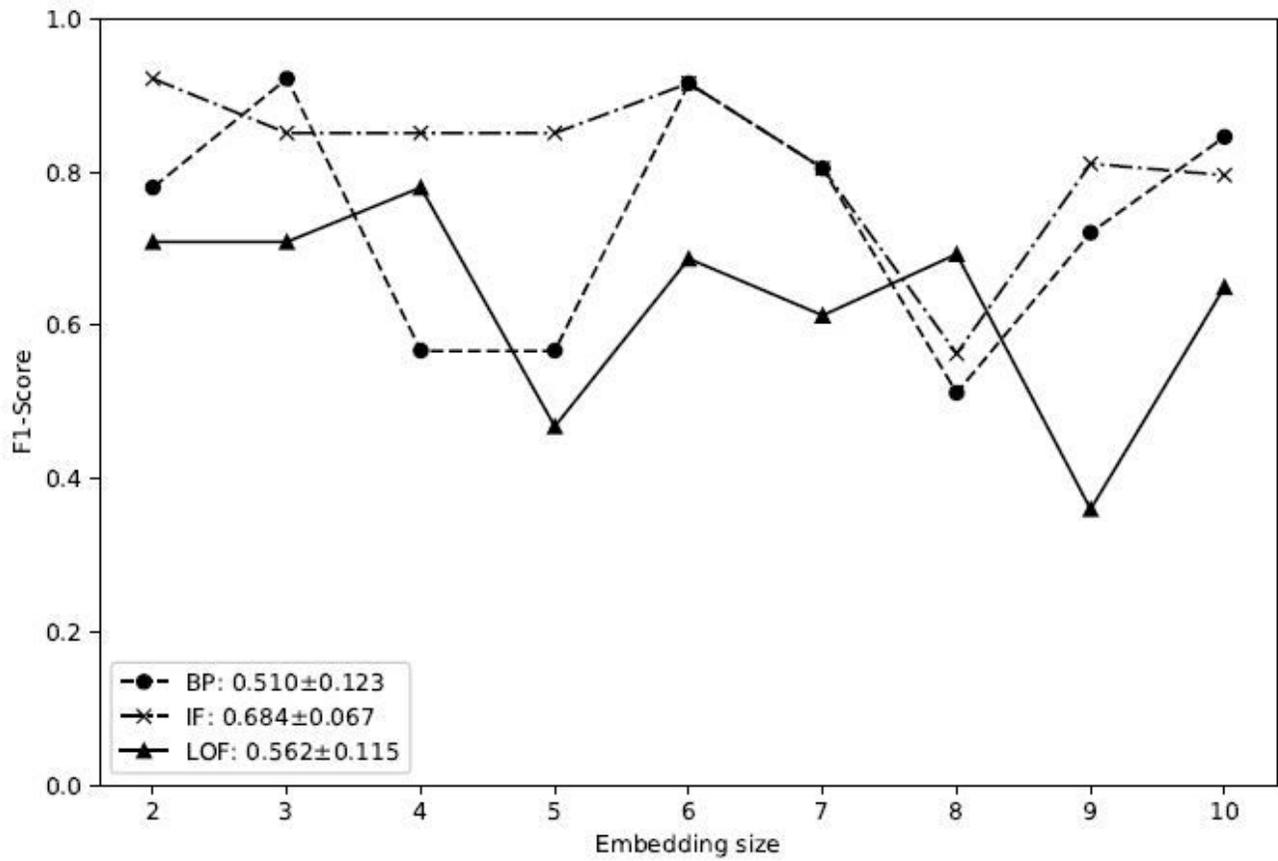


Figure 5

Stability of EDynGE model under the 10% outliers.

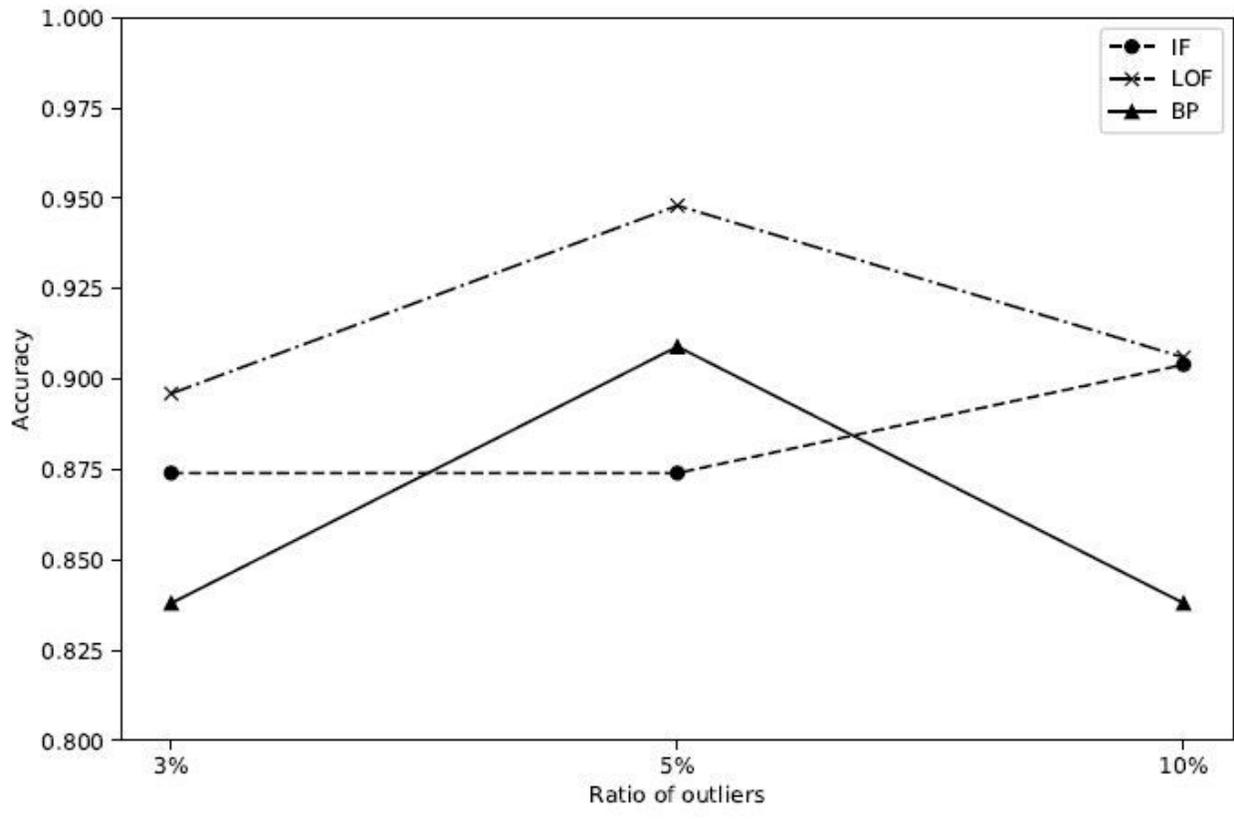


Figure 6

Performance for the city of Beijing.