International Journal of Disaster Risk Reduction 110 (2024) 104664

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

Post-hazard supply chain disruption: Predicting firm-level sales using graph neural network

Shaofeng Yang [a, *], Yoshiki Ogawa [a], Koji Ikeuchi [b, c], Ryosuke Shibasaki [d], Yuuki Okuma [e]

[a] *Centre for Spatial Information Science, The University of Tokyo, Chiba, Japan*
[b] *The University of Tokyo, Tokyo, Japan*
[c] *Foundation of River & Basin Integrated Communications (FRICS), Tokyo, Japan*
[d] *Department of Engineering, Reitaku University, Chiba, Japan*
[e] *Mitsubishi Research Institute, Inc., Tokyo, Japan*

ARTICLE INFO

Keywords:
Supply chain risk management
Hazard management
Supply chain disruption
Graph neural network
Explainable artificial intelligence

ABSTRACT

Evaluating the damage caused by supply chain disruptions triggered by natural hazards has an important role in supply chain risk management. This study proposes a novel approach to predict annual sales change rates for firms post-hazard. The approach employs graph neural networks, a machine learning technique that considers both internal factors such as sales and number of employees and external factors such as inter-firm relations and damages incurred from hazards. The model was trained on firm and flood data from previous floods in Japan and was found to outperform baseline models that did not consider inter-firm relations. To address the model's explainability problem, we employed Explainable Artificial Intelligence techniques to identify the factors influencing sales post-disruption and analysed the importance of firms and trades in the supply chain from a spatial perspective. The findings highlight the crucial role of business partners in a firm's supply chain and the impact of distance between suppliers on a firm's sales. The study contributes to improved supply chain risk management practices in industry as well as government, ultimately enhancing supply chain resilience.

1. Introduction

The frequency of supply chain disruptions (SCDs) triggered by natural hazards has been increasing [1]. Hurricane Maria, Hurricane Harvey, and Typhoon Hagibis are some examples of natural hazards that caused losses amounting to billions of dollars. The characteristics and frequency of natural hazards differ by region [2]. For example, the frequency of floods is significantly higher in South-East Asia. In Japan, the estimated economic damage from floods in 2019 exceeded USD 15 billion, a new record [3]. SCDs can slow down production activity in affected firms, and the impacts spread across the region through supply chains (SCs) to firms across the country or globally [4]. Therefore, how to mitigate damage from SCDs is a significant issue.

To ensure that companies can continue stable operations in the event of an SCD, it is necessary to accurately understand how supply chains are changed by SCDs and to make appropriate decisions to mitigate damage, such as restructuring the supply chain or reviewing internal controls. In this context, it is important to quantitatively predict how each company's performance, that is, sales, will change as a result of SCD. Many researchers have conducted studies on the impact of SCDs triggered by natural hazards on supply

* Corresponding author. 414, Kashiwa Research Complex, 5Kashiwanoha, Kashiwa-shi, Chiba, 277-8561, Japan.
E-mail address: sfyang@csis.u-tokyo.ac.jp (S. Yang).

https://doi.org/10.1016/j.ijdrr.2024.104664
Received 31 March 2023; Received in revised form 29 June 2024; Accepted 7 July 2024
Available online 8 July 2024

chains and firm performance. Several approaches have been proposed, mainly using input–output (I–O) models and computable general equilibrium (CGE) models to estimate how much the direct damage suffered by the impacted area affects the macroeconomy indirectly through SCs [5]. However, these studies have failed to consider the complex inter-firm relationships in SCs. Researchers have claimed that the structure of links between firms in the network, referred to as inter-firm topology [6], largely influences the damage caused by SCD and the subsequent recovery process. For example, capital relations between firms affect their stock prices [7], and geographical proximity, which indicates the relationship of physical distance between firms, may influence firm performance [8]. Thus, there is a complex relationship between firms, with demand and supply for each being volatile and unpredictable due to the uncertainty of SCs.

Machine learning (ML) is considered a better estimator of supply chain uncertainty because it can handle complex interdependencies between a very large number of factors with non-linear relational patterns affecting output results [9]. Therefore, ML can form the basis for developing methods to effectively optimise various supply chain objectives, such as controlling uncertainty in SCs [10]. However, the acceptance of ML in corporate decision-making [11] has faced challenges because ML models are 'black boxes' that only produce precise results but without detailed explanations, that is, the results are not easy to interpret. This is known as the explainability problem, which makes it difficult to implement ML in decisions that directly concern the interests of the company. To facilitate the deployment of artificial intelligence (AI) techniques such as ML in practice, many researchers have developed explainable artificial intelligence (XAI) techniques that make it easier to use ML models [12]. Despite the mature use of AI in the field of supply chain risk management (SCRM), the use of XAI in this field has not been adequately researched, such as how useful it is to use XAI in order to quantitatively identify the factors influencing firm sales after SCD.

While previous studies have provided valuable insights into the economic impacts of natural hazards on SCs, they often overlook the complex inter-firm relationships and the dynamic nature of supply chain networks. Traditional models (such as I–O and CGE) do not adequately capture the non-linear and interdependent relationships within SCs. Moreover, there is a lack of integration between economic models and more granular data that can reflect the immediate and cascading impacts of SCDs, presenting a significant challenge for effective SCRM and hazard preparedness. Given the identified gaps, this study aims to answer two research questions (RQs):

1. How can we develop a model that integrates complex inter-firm relationships and hazard data to predict the impact of SCDs on individual firm recovery performance?
2. What are the key factors influencing firm sales after an SCD, and how can these be identified and quantified using advanced ML techniques?

To address these RQs, this study has the following objectives:

1. To develop a post-SCD individual firm sales predicting model based on a graph neural network (GNN), leveraging data on Japanese firms. The GNN is designed to handle data with network structures, such as roads and SCs, and can be trained to consider inter-firm topology.
2. To evaluate the usefulness of the proposed model by training it with actual flood data and large-scale inter-firm topology data. This will involve verifying the model's performance in predicting firm sales post-hazard.
3. To apply XAI methods, such as GNNExplainer and SHapley Additive exPlanations (SHAP), to the proposed GNN model to identify the key factors affecting firm sales after a hazard and subsequently provide actionable insights for SCRM decision-making in industry and government.

The study makes two key contributions. First, it develops a GNN-based post-hazard sales forecasting model for individual firms that considers hazard damage and inter-firm topology of the firms. The development of this new method advances existing research on the use of AI to assess damage and predict demand changes in SCs after a hazard. Second, the XAI interpretation of the GNN model learned from real SC data reveals the key factors affecting individual firm sales after a hazard. The findings can encourage the adoption of AI techniques in actual SCRM decision-making.

The remaining paper is organised as follows. Section 2 summarises previous studies on conventional approaches to estimating SC damage triggered by hazards and ML approaches in SCRM. Section 3 describes the data used, and Section 4 outlines the procedure for preparing the training data, proposed and baseline models, and training model interpretation method. Section 5 summarises the experimental results and the interpretation results of the proposed model, Section 6 discusses the results and their implications for managers, and Section 7 presents the conclusion and future work ideas.

2. Literature review

2.1. Approaches to estimating damage to SCs caused by natural hazards

Many studies on estimating damage to supply chains during natural hazard-related SCDs have used methods based on the IO and CGE models. Both models are based on the Social Accounting Matrix method, which identifies the flow of transactions between all industrial sectors [13]. The IO model assumes a time-invariant production function with fixed levels of production technology, etc. in all sectors, and uses matrix data (e.g. input–output table) that capture the transaction flows of intermediate inputs and production in different sectors to estimate the impact of losses in one sector triggered by a natural hazard in other sectors [14]. CGE models assume stable behaviour of economic actors and equilibrium conditions in the demand and supply functions and then predict how demand, supply, and prices in various markets will change triggered by natural hazards [15].

Recent studies have further extended these models by incorporating more granular data and advanced simulation techniques. For instance, Nagurney et al. [16] developed a multicommodity international trade network equilibrium model under hazard scenarios, quantifying the impacts of disruptions to supply, transportation, and demand capacity, as well as exchange rates, on global trade. Ivanov [17] explored the ripple effect in SCs, proposing strategies for resilience through digitalisation and advanced modelling techniques. Additionally, Choi [18] examined the dynamic impacts of COVID-19 on global SCs, using real-time data analytics to highlight vulnerabilities and suggest adaptive strategies. Nagurney [19] applied supply chain game theory network modelling under labour constraints during the COVID-19 pandemic to illustrate the critical role of labour availability in maintaining supply chain functionality. These recent works highlight the evolving nature of SCRM and the increasing importance of integrating real-time and network-based analyses to understand and mitigate the impacts of natural hazards.

Moreover, recent advances in using firm-level data have furthered the understanding of hazard impacts on supply chains. For instance, Carvalho et al. [20] and Inoue and Todo [21] developed firm-level post-hazard simulation models that enable the estimation of impacts on the macroeconomy from a micro perspective, rather than relying on conventional aggregated statistical data. Gill et al. [22] developed a high-performance computing framework to simulate firm-level hazard impacts by integrating big data from multiple sources including firm-level data, emphasizing fine-grained simulations to understand and mitigate disruption ripple effects. Yang et al. [23] used inter-branch transaction data, which is more granular than firm-level data, to study urban flooding impacts. They found that many firms outside core hazard areas are affected through supply chain linkages, highlighting the importance of considering indirect effects on firms. These studies have highlighted the importance of considering the firm-level inter-firm relationships and the dynamic nature of supply chain networks in understanding the broader economic impacts of natural hazards.

However, previous studies have often overlooked the complex inter-firm relationships and geographical damage conditions in their models. Recent advances aim to address these gaps. Shao [24] found that supply chain geographic dispersion has a negative impact on a firm's disruption mitigation capabilities. Lawson et al. [25]found that geographic distance lengthens response time to SCDs, while industry-relatedness reduces it. Additionally, response time worsens as more firms in an industry are affected by the same recall.

Despite these advancements, the challenge with the previous studies is that, apart from inter-firm business relationships, other relationships such as investment and regional networks have not been considered. Nor have they considered geographical damage conditions. This study proposes a model for predicting the sales of individual firms that uses natural hazard damage data to identify the damage to each firm through the geographic information system (GIS) and considers the various relationships between firms (inter-firm topology).

2.2. ML approaches in SCRM and resilience

The ML-based studies in SCRM mainly deal with topics such as supplier selection, supplier segmentation, supply chain risk prediction, demand and sales estimation, production and inventory management, and transportation [11]. For example, sales problems are addressed by combining sales data for retail stores or the whole firm with ML models such as long-short-term memory and extreme gradient boosting (XGBoost), etc [26,27]. Kamble et al. [28] utilised the Technology Acceptance Model and Technology-Organisation-Environment framework to predict blockchain adoption in SCs, providing managers with a decision support system to enhance organisational competitiveness through blockchain technology. Rajesh [29]introduced grey causal modelling (GCM) to analyse complex causal relationships in SCs, focusing on resilience and sustainability. This methodology incorporates events, outcomes, objectives, and effects, offering a robust framework for analysis. Singh et al. [30] explored the role of the Digital-Twin technology in enhancing food supply chain resilience and sustainability using GCM, demonstrating its impact on production efficiency, profitability, and waste reduction. Additionally, Baryannis et al. [31] emphasised the trade-off between the performance and interpretability of ML models in predicting supply chain risks, advocating for interpretable models that enable decision-makers to understand and trust predictions. Ghazal and Alzoubi [32] highlighted fusion-based supply chain collaboration using ML techniques, showing how these approaches improve decision-making, reduce risks, and enhance overall resilience through better information sharing and teamwork. Yang et al. [33] used multi-agent deep reinforcement learning to optimise firm-level recovery strategies during floods, incorporating dynamic conditions and behavioural strategies.

Despite these advancements, many studies still rely on input data extracted separately for each firm. However, a firm's performance is affected by its network relationships with various stakeholders [34]. Researchers have also begun to emphasise the relationship between inter-firm topology and impact spillovers [35]. Therefore, an approach that considers inter-firm topology is necessary. In recent years, studies that used GNN for network data as this ML technology had the advantage of considering network topology [36,37]. Some studies have used GNN for firms, mostly to predict firm stock prices [7]. In these studies, past stock prices are used for firm features, while the connections between firms are trade and investment relationships. Among the studies using GNNs for SCRM topics, Wu et al. [38] addressed the problem of classifying industries of firms by mining hidden information obtained from SCs through a classification model based on the GNN and the supply chain network dataset of listed firms in China. The model shows higher performance than the ML model without GNNs. Furthermore, Kosasih and Brintrup [39] posed the supply chain visibility problem as a link prediction between firms and proposed a method to automatically detect potential trade links unknown to customers through a link prediction model using GNNs and actual automobile supply chain network data. The results show better performance than existing methods do.

Compared with previous studies, our proposed model offers numerous advantages: (1) the use of GIS data on flood inundation status allows the damage situation at individual enterprises to be considered; (2) the impact of SCD on supply chains can be analysed at the level of individual firms; and (3) the complex relationships between companies are considered, which enables the impact of external factors, such as other companies, as well as internal factors on post-SCD sales to be clarified. To the authors' knowledge, there are no studies on the development of sales prediction models based on GNN and real supply chain data for individual firms for post-hazard periods, and this study is the first attempt to do so.

3. Data

3.1. Firm-level SC data: corporate credit research dataset

This study utilises the corporate credit research dataset (2017–2018) held by the private credit survey firm Teikoku Databank Ltd. This dataset contains corporate profile information data for 1.46 million firms (1.46 million firm headquarters and 0.59 million branches) across Japan. The total sales value of all firms in this database covers about 90 % of the value of the Economic Census, a high coverage. The data contain the unique firm code for each firm, as well as profile information including office locations, industry, annual sales, and number of employees. The dataset is characterised by the high coverage of small and medium-sized firms (SMEs), for which obtaining profile information is usually difficult. The data pertain mostly to firms that engage in credit trades, that is, business-to-business (B2B) trades, and also contain information on business and investment relationships between firms (Fig. 1). The estimated volume of inter-firm trades has been developed by Tamura et al. [40].

3.2. Hazard area

The target hazard is Typhoon Hagibis, which caused record-breaking rainfall in July 2018 over a wide area, mainly in western Japan [41]. In some areas, the total rainfall on the 11 days from 28 June to 8 July was 2–4 times higher than the average rainfall in previous years. River overflows and landslides occurred simultaneously, and many firms suffered direct flood damage as well as human casualties. SCDs occurred as a result of factory shutdowns and blockage of major roads. In this study, we use the flood data for Ozu City, Ehime Prefecture, and Kurashiki City, Okayama Prefecture, two of the most affected locations. In Kurashiki City, the depth of inundation is estimated to have exceeded 5 m over a wide area, going up to 5.4 m, while some 4600 houses were inundated in Ozu City. Fig. 2 presents the flood damage in these cities estimated by the Geospatial Information Authority of Japan [55].

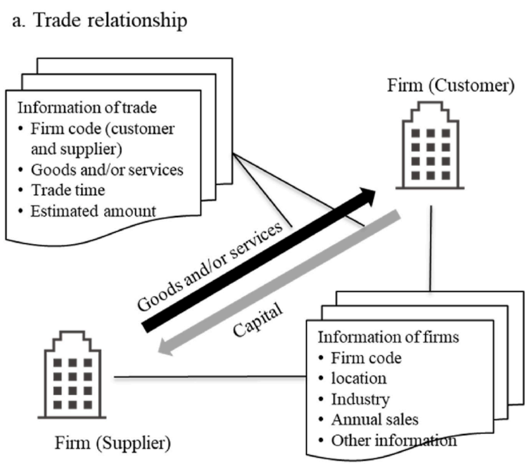

Fig. 1. Block diagram of corporate credit research data.

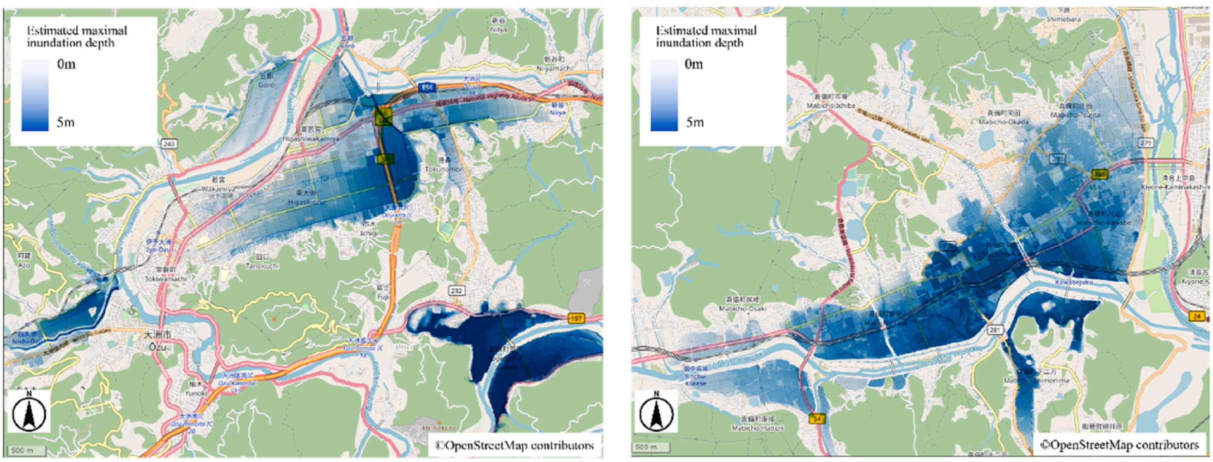

Fig. 2. Estimated flood damage due to Typhoon Hagibis (left: Kurashiki City, Okayama Prefecture; right: Ozu City, Ehime Prefecture).

4. Methodology

Fig. 3 presents the flowchart of the methodology. First, we extract the affected firms and their business partners from the supply chain and flood data and develop the data on inter-firm relationship networks and individual firm features. Next, we build and train a GNN model and baseline models. Finally, the XAI methods, GNN Explainer and SHAP, are used to identify the influencing factors for post-hazard changes in company sales.

An inter-firm network is denoted as $G = (V,E)$ where $V = \{v_1,v_2,\ldots,v_n\}$ is the set of n firm nodes and E is the set of links connecting these firm nodes. The links between the nodes are represented by the adjacency matrix $A \in \{0,1\}^{n*n}$.

4.1. Data processing

4.1.1. Features of firm nodes V

First, the affected firms and firms on their SCs were identified using the corporate credit research dataset of 2017, before the hazard. The impact on these firms was estimated by spatial join using the geospatial information system technique. We identified 770 affected firms (625 headquarters and 348 branches) affected by the flood, categorised as Tier 0 firms. Another 17,225 firms with direct business relations with Tier 0 firms were extracted as Tier 1 firms. In total, 17,995 firms were used as training data.

Next, a total of 67 features in four categories (information on the firms' trade, profile, network, and hazard) were calculated from the corporate credit research dataset, as shown in Table 1, and used as explanatory variables. The logic behind selecting these features was to comprehensively capture various dimensions that could influence a firm's performance post-hazard.

1. Profile: This category includes general information such as industry and firm growth rates, providing context on the firms' characteristics. Such information is crucial as it helps understand the inherent resilience or vulnerability of different types of firms (e.g. industrial classification and firm size) [42].
2. Trade: Features in this category include the number of trades by direction in the SC, type of business partner, and distance between firms, which are critical in understanding the operational interdependencies and the flow of goods and services. Previous studies have highlighted the importance of trade dynamics in supply chain resilience and disruption propagation [43].
3. Hazard: This includes the inundation depth of the firm's or business partner's headquarters or branches and the damage rate by industry sector calculated from the inundation depth (Kuroda et al., 2020). These features directly measure the physical impact of the hazard on the firms, which is crucial for assessing direct damage and its economic implications [44].
4. Network: This is an indicator of centrality in network theory; its value indicates the importance of the firm in the supply chain. Network centrality metrics have been shown to be significant predictors of a firm's ability to withstand and recover from disruptions [45].

The scale of the feature values for the explanatory variables is 17,995 firms * 67 = approximately, 1.2 million. The objective variable is the percentage change in annual sales before and after the hazard for each firm (ratio of 2018 to 2017). The basic statistics are presented in Table 2.

4.1.2. Matrix A of inter-firm network

A firm's performance is influenced by its various network relationships, including trade, capital, and other stakeholders [34]. Geographical proximity, which indicates interaction between firms in the same region, also affects firm performance [8]. Therefore, we consider the regional network and create seven directed adjacency matrixes from seven types of inter-firm network data from before the hazard, as shown in Table 3. In addition to the six types of networks of suppliers, customers, business partners, investees, in-

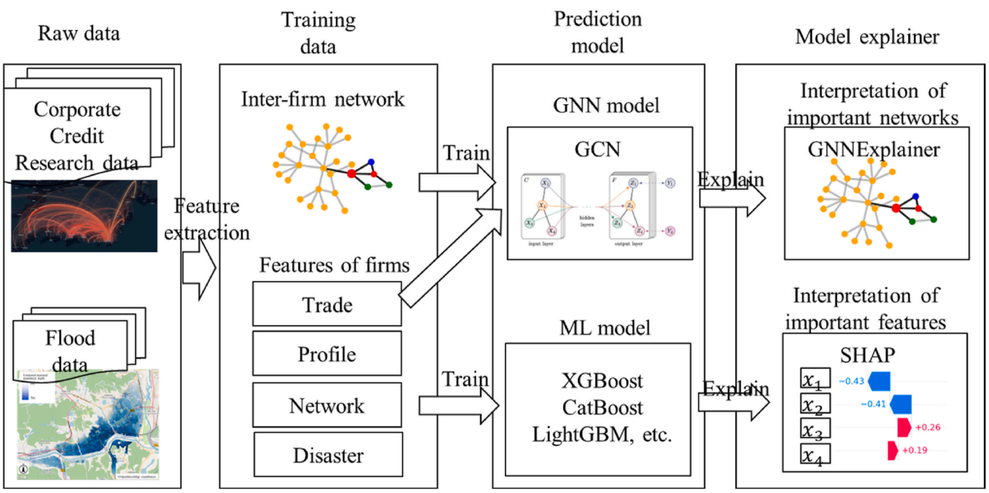

Fig. 3. A flow chart of the methodology.

Table 1
Firm features and objective variables.

Type	Description	Direction	Tier	notation	Direction	Tier	Notation
Trade	Average annual growth in trade value over the last three years	Customer	1	F1	Supplier	1	F2
	Total value of all trades		1	F6		1	F5
			0–2	F15		0–2	F9
	Average of inter-firm distance to business partners		1	F17		1	F11
	Maximum value of inter-firm distance to business partners		1	F18		1	F12
	Number of business partners		1	F8		1	F7
			0–2	F16		0–2	F10
	Unit distance average trade value (Average trade value per unit distance calculated from the total value of trades with all suppliers and the total value of inter-firm distances)		1	F19		1	F13
	Entropy of business partners' industries (the higher the entropy, the more diverse the business partners).		1	F20		1	F14
	Average annual growth rate of sales over the past three years		1	F4			
Profile	Annual average growth rate in the number of employees over the past three years					0	F3
	Agriculture						F55
	Forestry, hunting						F56
	Fishing						F57
	Mining industry						F58
	Construction						F59
	Manufacturing						F60
	Wholesale and retail trade, restaurants						F61
	Finance and insurance						F62
	Real estate						F63
	Transport and communications						F64
	Electricity, gas and water supply						F65
	Service						F66
	Public service						F67
	Number of branches (warehouses, factories, etc.)						F43
Network	Local SC network degree correlation				Both	0–2	F21
	Local SC network cluster coefficient					0–2	F22
	Betweenness centrality					All	F23
	Degree centrality					All	F24
	Eigenvector centrality					All	F25
	Closeness centrality					All	F26
	Sum of betweenness centrality of business partners	Customer	1	F35	Supplier	1	F27
	Average of betweenness centrality of business partners		1	F36		1	F28
	Sum of closeness centrality of business partners		1	F37		1	F29
	Average of closeness centrality of business partners		1	F38		1	F30
	Sum of eigenvector centrality of business partners		1	F39		1	F31
	Average of eigenvector centrality of trading partner		1	F40		1	F32
	Sum of degree centrality of business partners		1	F41		1	F33
	Average of degree centrality of trading partner		1	F42		1	F34
Hazard	Number of branches located in the flood area						F44
	Flag whether the headquarters are located in the flood area						F45
	Inundation depth for the headquarters						F46
	Capital damage rate for the headquarters						F47
	Total inundation depth for branches						F48
	Total capital damage ratio for branches						F49
	Average of inundation depth of branches						F50
	Average of damage rate for branches						F51
	Flag whether the headquarters or branches are in the affected area						F52
	Number of affected business partners	Customer	1	F54	Supplier	1	F53
Objective	Percentage change in annual sales before and after hazard (Ratio of post-hazard sales in 2018 to pre-hazard sales in 2017)						

Table 2
Basic statistics for the objective variable (percentage change in sales in the year before and after the hazard).

Count	Mean	Std	Min	25 %	50 %	75 %	Max
17,995	1.074245	0.35532	0	0.951526	1.018512	1.139158	7.200238

vestors, and the same city, a network called ALL, which considers all relationships, is created by the seven types of network matrixes. Note that ALL is a matrix of logical sums of the six types of adjacency matrix. The size of each network matrix is approximately 17,995*17,995 = 324 million. Also, 80 % of the data in the same network are used as training data and 20 % as test data, and the training results of the 5-fold cross-validation are the output.

Table 3
Seven types of inter-firm networks.

Notation	Number of links	Description
Suppliers	102,982	Relationships between firms connected by supplier relationships
Customers	102,982	Relationships between firms linked by customer relationships
Business partners	205,964	Relationships between firms linked by supplier or customer relationships
Investees	5365	Relationships between firms linked by investee relationships
Investors	5365	Relationships between firms connected by investors relationships
Regional	1,871,208	Relationships between firms located in the same city
ALL	2,293,866	Logical sum of adjacency matrixes of the above six types of networks

4.2. Graph convolutional network (GCN)

GCN is one of the GNNs and a model for adapting neural networks to network data, which gives the advantage of fast learning and the ability to handle large networks [46]. It is able to incorporate latent relationships in the connections between nodes in the network into the learning model. Not only the information of target nodes but also that of their neighbouring nodes are aggregated and used in the learning data. Taking the adjacency matrix A of the network G and the feature matrix X of the set of nodes V as input layers, the $l+1$ layer of the GCN can be expressed as follows.

$$H^{l+1} = \sigma\left(\widehat{D}^{-0.5}\widehat{A}\widehat{D}^{-0.5}H^{l}W^{l}\right) \tag{1}$$

where I_n is the unit matrix, $\widehat{D}$ is the order matrix of $\widehat{A}$, $\widehat{A} = A + I_n$ and the $\widehat{D}^{-0.5}\widehat{A}\widehat{D}^{-0.5}$ term is used to normalise A. H^{l+1} is the output of layer l, where $H^1 = X$. W^l is the weight matrix and σ the activation function ReLU. Fig. 4 illustrates the architecture of the GCN, where information about the target node and its neighbours is quantified using the adjacency matrix, unit matrix, and order matrix of the network between nodes, and then standardised and sent to the neural network, so that the predictions of each firm node become output, as in Equation (1).

In addition, the loss function is trained and updated to minimise the loss function using the Adam-based batch gradient descent optimisation algorithm [47]. The loss function is used as the mean absolute error (MAE), expressed by equation (2).

$$MAE = 1/n\sum_{i=1}^{n}\left|\widehat{y}_i - y_i\right|, \tag{2}$$

where y_i is the ground truth and $\widehat{y}_i$ is the predicted value. MAE and mean squared error (MSE) are used to evaluate the model. MSE is expressed as

$$MSE = 1/n\sum_{i=1}^{n}\left(\widehat{y}_i - y_i\right)^2 \tag{3}$$

The smaller the values of MAE and MSE, the better the prediction performance of the model, because the predictions then get closer to the measured values.

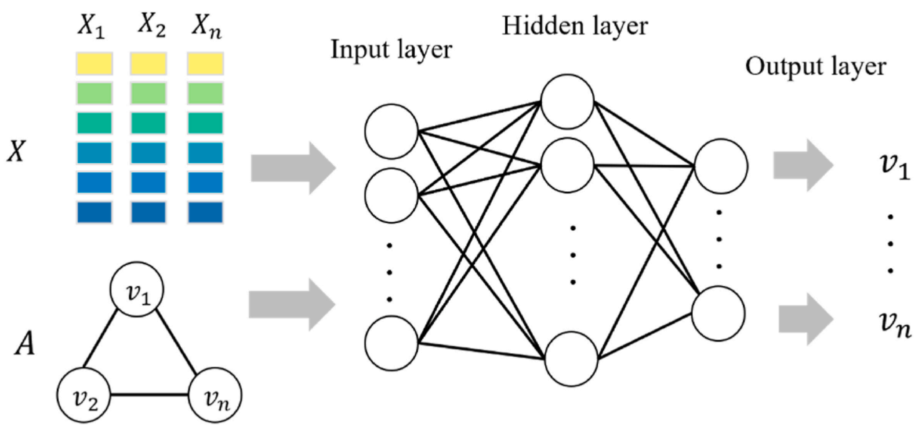

Fig. 4. The GCN architecture.

4.3. Baseline model

To demonstrate the usefulness of the proposed method, the following five models were prepared as baseline models. These models were selected because they are widely used in ML and do not consider the relationships between companies, which allows for a clear comparison with the GNN approach. For all approaches, the hyperparameters were tuned by validating the MAE of each model using the 5-fold cross-validation and grid-search method (Table 4).

1. Multiple-layer perceptron neural network (MLPNN): This network uses back-propagation to solve supervised learning problems with multiple layers of neurons. It can learn complex non-linear patterns in the data, making it suitable for capturing intricate relationships within the dataset.
2. Random forest (RF): This is an ensemble method for classification and regression that creates a large number of decision trees and then averages the results [48]. RF is robust to overfitting, handles various data types well, and can manage missing data effectively.
3. Extreme gradient boosting (XGBoost): This uses an optimised distributed gradient boosting algorithm that ranks among the most powerful ML models available [49]. XGBoost is highly efficient and effective, capable of capturing variable importance and interactions within the data.
4. Light gradient boosting machine (LightGBM): This model utilises a novel technique of gradient-based one-side sampling to reduce the feature space dimension and further improve the speed of training [50]. LightGBM is efficient with large datasets and high-dimensional data, offering fast training times.
5. Categorical boosting (CatBoost): These are boosting variants of linear models designed for categorical features that use prior information about the data to optimise splits [51]. CatBoost handles categorical data efficiently and effectively without extensive preprocessing.

These baseline models provide a diverse and robust set of benchmarks to evaluate the performance of the proposed method. By comparing with models that do not account for relationships between companies, we can effectively highlight the advantages and improvements offered by the GNN approach.

4.4. Explain model

4.4.1. SHAP

SHAP, developed by Lundberg and Lee [52], is used as an interpretation method for influencing factors for trained ML models. The output value, called the SHAP value, is a measure of the contribution of each feature to the output of the model. The SHAP value Φ_i is calculated based on equation (4):

$$\Phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} (|S|\,!\,(N - |S| - 1)\,!\,/N!\,) \times (v(S \cup \{i\}) - v(S)), \tag{4}$$

where N is the set of all variables, S a subset of N excluding i, and v the utility function; the $v(S \cup \{i\}) - v(S)$ term calculates the marginal contribution of feature i. This formula is used to evaluate the difference in predicted results caused by the inclusion of feature i in the output of the prediction model for all feature combinations other than a certain feature i. Based on this, SHAP values for individual variables can be determined, helping to identify the factors influencing prediction results.

Table 4
Hyperparameter tuning.

Model	Parameters	Significance	Search range	Optimal values
MLPNN	hidden_layer_sizes	Number of neurons in the hidden layer	10,50,100	100
	learning_rate_init	Initial learning rate for weight updates	0.0001,0.001,0.01	0.0001
RF	n_estimators	Number of trees	40,100,200	40
	min_samples_split	Minimum number of samples for nodes split	2,5,10	2
	max_depth	Maximum depth of a tree	4,8,12	4
CatBoost	n_estimators	Number of trees	40,100,200	200
	learning_rate	Shrinkage coefficient of each tree	0.01,0.1,0.2	0.1
	max_depth	Maximum depth of a tree	4,8,12	4
LightGBM	n_estimators	Number of trees	40,100,200	40
	learning_rate	Shrinkage coefficient of each tree	0.01,0.1,0.2	0.01
	max_depth	Maximum depth of a tree	4,8,12	8
	colsample_bytree	Subsample ratio of columns for tree construction	0.6,0.8,1.0	1.0
XGBoost	n_estimators	Number of trees	40,100,200	200
	learning_rate	Shrinkage coefficient of each tree	0.01,0.1,0.2	0.01
	max_depth	Maximum depth of a tree	4,8,12	4
	colsample_bytree	Subsample ratio of columns for tree construction	0.6,0.8,1.0	0.6
GCN	hidden_layer_sizes	Number of neurons in the hidden layer	16,32,64,128	128
	learning_rate_init	Initial learning rate	1e-5,1e-4,1e-3,1e-2	1e-3
	weight_decay	Parameter to manage over fitting	1e-4,5e-4	5e-4
	epochs	Number of epochs	50,100,200,300	300

4.4.2. GNNExplainer

GNNExplainer, a method for interpreting GNN models, takes a trained GNN model and training data as input and outputs the subgraphs that are important for prediction [53]. It identifies the subgraphs of the entire graph trained by the GNN model, which have the maximum amount of mutual information with the GNN prediction. By solving the optimisation problem for the objective function in equation (5), the (G_s, X_s) important for the prediction results of the GNN model can be obtained.

$$\max_{G_s} MI\left(Y, \left(G_s, X_s\right)\right) = H\left(Y\right) - H\left(Y \middle| G = G_s, X = X_s\right), \tag{5}$$

where G_s is the subgraph of the graph to be computed, X_s the feature mask of the node corresponding to G_s, Y the prediction result of the learned GNN model, H the entropy, and MI the mutual information. GNNExplainer is used to identify the firms and trades that influence prediction results.

4.5. Experimental environment

The experimental environment is conducted in Python 3.8.5, with PyTorch Geometric 1.6.0 for the GNN model and GNN Explainer, shap0.41.0 for SHAP, CatBoost for CatBoost1.1, LightGBM for Lightgbm3.3.3, and XGBoost used XGBoost1.6.1, RF and MLPNN used scikit-learn 1.1.2.

5. Results

5.1. Comparison of model performances

Table 5 shows the results of the comparison of the performance of the trained GCN models and the five baseline models, with MAEs ranging from 0.212 (XGBoost) to 0.128 (GCN, regional network). The seven GCN models outperform all baseline models and inter-firm topology. The forecasting method using GCN models considering inter-firm topology is beneficial for predicting the sales of a firm post-hazard. The best-performing baseline model without considering inter-firm topology is CatBoost's MAE = 0.188.

Among the seven models, the model with the lowest MAE and MSE was the one using the regional network. This suggests that firms located in the same region may be linked, influencing each other's performance. For example, in Japan, many SMEs form business cooperatives with neighbouring firms through business collaboration, creating a situation where firms are more likely to link their performance to each other's. The next lowest MAE is the model using the customer network, indicating a significant influence of the buyer firms. The worst among the GCN models is the investment one. Finally, ALL's performance does not give the best results, and its predicting performance is less than that of the other networks, perhaps because ALL includes investment relations with weak predictive performance.

5.2. Model interpretation results

5.2.1. Explanation of firm features using SHAP for the CatBoost model

Using SHAP, the magnitude of the influence of the 67 features on the prediction of the objective variable (average of the absolute SHAP values) was calculated for the CatBoost model with the best performance at baseline. Fig. 5 shows the results—the higher the SHAP value, the more important the feature to the prediction of the objective variable.

The top 10 influencing factors, in descending order of importance, are average annual growth rate of the number of employees (F3), average annual growth rate of sales (F4), eigenvector centrality (F25), sum of eigenvector centrality of business partners (F39), number of business partners (F16), degree centrality (F24), construction (F59), unit distance average trade value (F13), maximum physical distance to business partners (F12) and average of eigenvector centrality of customers (F40). Among these, three features, F3, F4 and F25, have a relatively high influence on the prediction of the objective variable, implying that they are important for the recovery of sales of post-hazard companies.

Next, to analyse how the top 10 most important features contribute to the objective variable, the correlation coefficient between the SHAP value and the feature value is calculated. This correlation coefficient indicates the contribution of each feature to the objec-

Table 5
Comparison of model performances.

Model		MAE	MSE
GCN	Supplier	0.160	0.073
	Customer	0.136	0.052
	Business partner	0.160	0.058
	Investee	0.172	0.068
	Investor	0.174	0.084
	Regional	**0.128**	**0.046**
	ALL	0.161	0.068
Baselines	XGBoost	0.212	0.216
	LightGBM	0.203	0.215
	CatBoost	0.188	0.191
	RF	0.204	0.285
	MLPNN	0.208	0.241

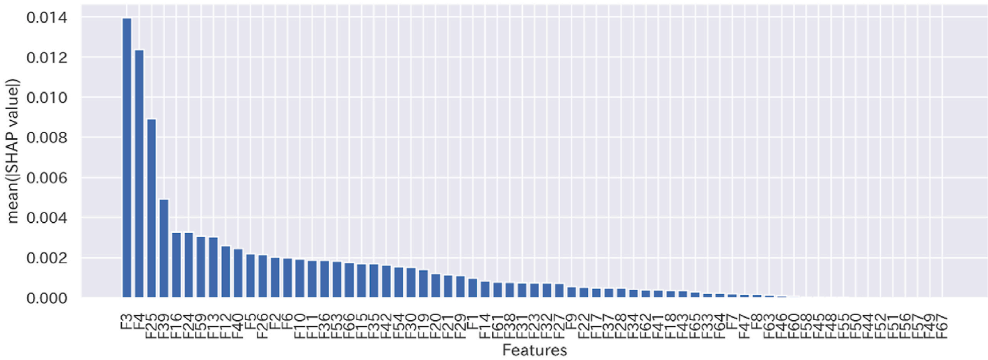

Fig. 5. The relative importance of each feature.

tive variable; a positive correlation coefficient contributes positively to the objective variable, and a negative correlation contributes in a negative direction. Fig. 6 shows the correlation diagram between the value of each feature and the SHAP value.

5.2.1.1. Features with positive correlation coefficients. The higher the value of the relevant feature, the more likely it is to contribute to an increase in sales of the post-hazard firms.

- Sum of eigenvector centrality of business partners (F39) = 0.71: Firms with sales destinations with high eigenvector centrality may be able to maintain stable trade even after a hazard, as these customers are important firms in the SCN. Having stable sales channels is a necessary condition for maintaining sales after a hazard. Therefore, companies would do better to prioritise companies with high eigenvector centrality in the SCN when selecting customers.
- Number of business partners (F16) = 0.68: Companies with a relatively large number of customers can avoid a significant decrease in sales with some of their customers, after a hazard, as long as they have stable business with other customers. Thus, having many customers can lessen the risk of indirect damage due to flooding.
- Unit distance average trade value (F13) = 0.68: A higher transaction value relative to the distance between companies and suppliers implies that sales are more likely to recover after a hazard. This suggests that, even when the distance to suppliers is

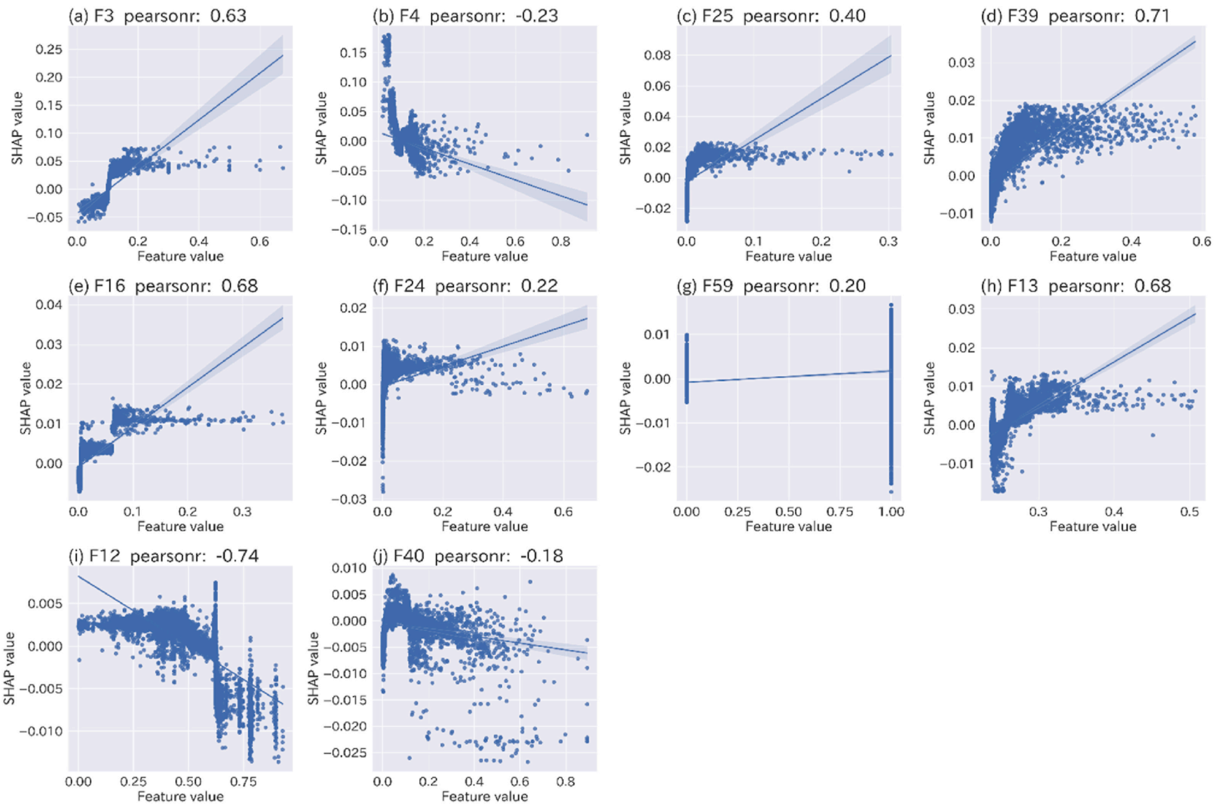

Fig. 6. Correlation diagram between SHAP values and feature values. (The x-axis is the value of the feature, and the y-axis is the SHAP value).

great, large amounts of transactions can be considered by businesses that are not dependent on logistics and do not incur transport costs (e.g. electronic files, services and other transactions delivered via the Internet). Such companies may be able to continue their business even after flooding, as it is relatively easy to change suppliers in the event of logistics stagnation or damage to suppliers.

- Average annual growth rate of employees (F3) = 0.63: Companies with a higher annual growth rate of employees are considered more able to restart or continue their business steadily after a hazard, as they have more human resources available to allocate to the development of internal SCRM, making it easier to put in place an SCRM system. In addition, companies that seek to add more staff are more likely to be proactive in expanding their business or launching new businesses. This means that even if core businesses are shut down due to SCD caused by flooding, sales can be maintained by shifting to other businesses and taking other measures.

5.2.1.2. Features with negative correlation coefficients. The higher the value in the negative direction, the more likely it is to contribute to a decrease in sales of post-hazard firms.

- Maximum value of inter-firm distance to business partners (F12) = −0.74: If the distance to suppliers is greater, then sales are likely to decrease. In particular, when flooding occurs, there may be landslides on roads, in addition to flood damage. This could result in longer road restoration times and logistical stagnation, which could conceivably bring firm sales and operations to a halt for a long time. When selecting suppliers, it is advisable to choose companies situated close to the firm. Ensuring stable purchases in the event of logistical disruption or stagnation following a hazard is one of the most important ways of maintaining sales.

5.2.2. Explanation of the SC network using GNNExplainer for the GNN model

GNNExplainer was used to interpret the model with a business partner network. For each firm's SC, the importance of each trade for the rate of change in sales was calculated. Fig. 7 shows a visualisation of the importance of each trade on the SC for a sample firm in the construction industry, whose sales increased after the hazard. Many of the trades judged to be important are related to firms whose sales increased after the hazard. We assume that the increase in sales of these firms led to higher sales by the sample firm. This indicates that any SCD could be fatal for the sample firm.

Next, we analyse the important trades in the SC by industry, direction of trade, and tier, and determine the proportion of important trades for each firm's business partners using eight indicators: number of affected firms, business partners, and trades; same subsector, major sector, prefecture, city and trade value, direction of trade in the supply chain tier, and the average value for all firms (Fig. 8). For example, the result of the indicator for the same prefecture is the percentage of important trades in the supply chain of the target firm in all trades between the firm and its partners in the same prefecture. The results of such indicators are used to analyse the trends and characteristics of important trades on the overall supply chain.

First, the results by tier show that the proportion of important trades with Tier 1 client firms is relatively high. This is because the amount of trades with sales counterparties is directly related to the firm's own sales, and therefore the firm's own sales are likely to be affected. Further, a comparison between the city and the prefecture shows that the city has a high proportion of important sales, which suggests that sales destinations that are closer together are more likely to be important in terms of trade. In terms of industry, the real estate sector has a relatively high proportion of important trades with firms that have been affected with respect to their clients. When a damaged firm that rents its own property withdraws at the point of sale, it is not easy to find a new tenant, and therefore the target firm's sales are affected. In addition, we find wholesale, retail, and construction industries relatively more important in trades with the same industry. This could have a significant impact on their own sales, as the main customers are in the same industry sector.

Next, to analyse trends geographically, important trades were extracted in the SCs of all firms, and the proportion of important trades in the total number of trades in each prefecture was calculated (Fig. 9). A high proportion of important ordering trades is found in the prefectures near large cities such as Tokyo and Osaka, and near hazard areas such as Ehime Prefecture. A high proportion of important ordered trades are in prefectures far from the hazard area, such as Kyushu and Tohoku, suggesting that the important SC net-

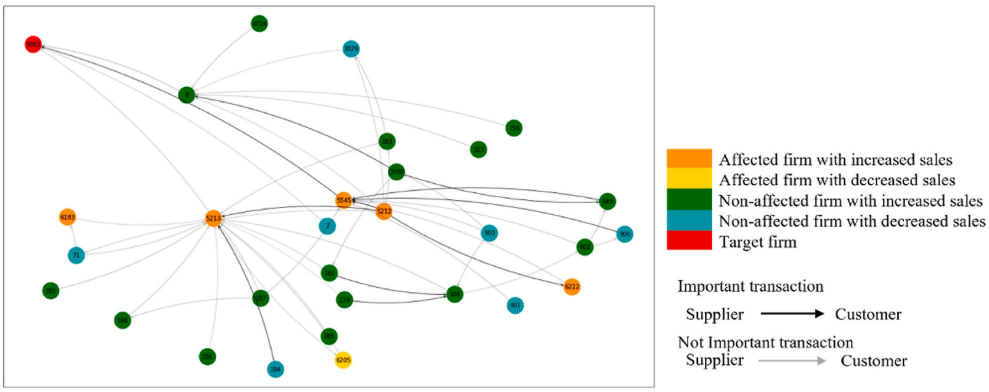

Fig. 7. Visualisation of important trades in the supply chain of a sample firm with increased sales.

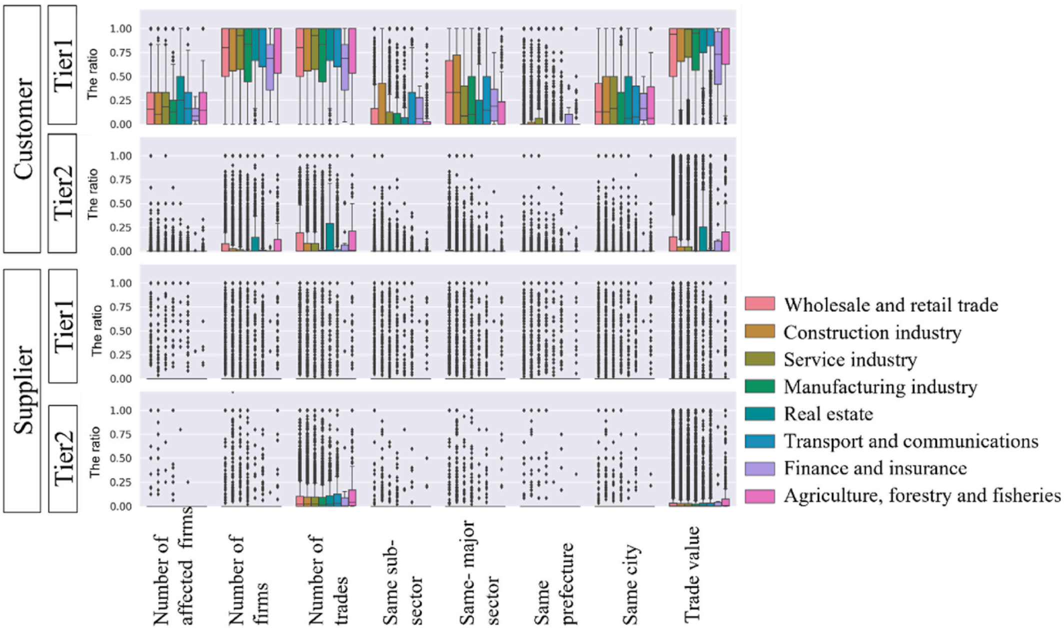

Fig. 8. Percentage of significant trades in the supply chain of affected firms (calculated using 8 indicators, by tier and direction of trade).

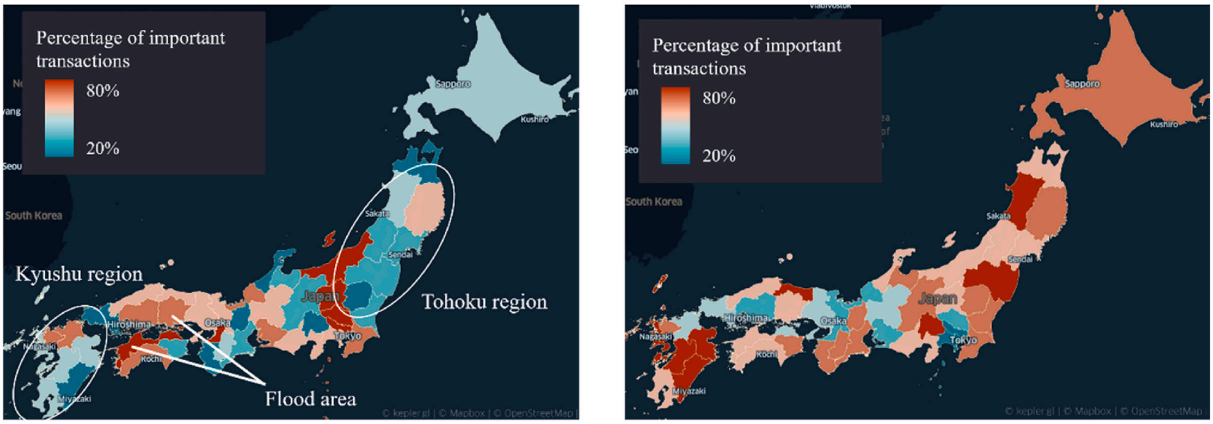

Fig. 9. Percentage of important trades in each prefecture in the SC of affected firms (left: ordering; right: ordered).

work for ordered trades is not dependent on distance. Identifying important trades geographically determines the trades and firms on a supply chain network that are potentially important for economic recovery after SCD.

6. Discussion and managerial implications

6.1. Main findings

The major results, including the identification of key influencing features and their impact on post-hazard sales recovery, are summarised in Table 6.

The main findings of this study are as follows. First, the model for predicting sales change for individual firms that considers inter-firm topology performs better than the baseline model. This implies that inter-firm topology affects the performance of post-hazard firms and is related to damage spillovers, which is consistent with the claims made by previous studies [35]. Further, the links between firms that affect performance are not only trade and investment relations, but geographical networks as well. This result is consistent with previous studies claiming that geographical proximity between firms affects firm performance [8]. Second, the influencing factors for improved post-hazard firm sales were identified and each factor was ranked. The most important factors include high eigenvector centrality in the supply chain network of the customer and low physical distance from the suppliers to the company. These results make a novel contribution to the literature. In conventional practice, risk management decisions are generally made by empirical rules, but our findings allow identifying the most important from a large number of factors so that a firm's limited resources can be used for them as effectively as possible.

Table 6
Summary of major results.

Section	Key Findings	Details	Comparison with Existing Literature
5.1	Model performance comparison	The trained GNN models outperformed the five baseline models. The best-performing GNN model is the one considering geographic relation between firms	Similarity: Consistent with Feng et al. [7] and Kosasih and Brintrup [39] in that the performance of the GNN model considering inter-firm topology y outperforms the baseline models. Difference: This study shows that the regional network model performs better than the business relation network, which is not specifically addressed in these studies.
5.2.2	The top 10 influencing factors(in descending order of importance)	F3: Average annual growth rate of employees F4: Average annual growth rate of sales F25: Eigenvector centrality F39: Sum of eigenvector centrality of business partners F16: Number of business partners F24: Degree centrality F59: Construction F13: Unit distance average trade value F12: Maximum physical distance to business partners F40: Average of eigenvector centrality of customers	Similarity: Aligns with existing studies on the importance of factors affecting firm performance, such as industry, physical distance, supply chain network centrality, etc. [24,25,45] Difference: Our study captures explanatory variables with comprehensive dimensions and quantifies the exact impact, including positive and negative, of each variable on post-hazard company sales, which previous studies have not done.
5.2.1.1	Features with Positive Correlation Coefficients	F39 (0.71): Stable trade from high-centrality business partners. F16 (0.68): Large number of customers reduces sales risk. F13 (0.68): High trade value over distance implies resilience. F3 (0.63): High employee growth indicates strong SCRM capabilities.	
5.2.1.2	Features with Negative Correlation Coefficients	F12 (−0.74): Greater distance to suppliers increases sales risk due to logistical issues.	
5.2.2	Explanation of the SC network	- Importance of Trades: Trades with firms whose sales increased post-hazard are crucial. - Tier Analysis: High importance of Tier 1 trades due to direct sales impact. - Industry Analysis: High importance in real estate, wholesale, retail, and construction sectors. - Geographical Trends: High importance of trades in and near large cities and hazard areas, and distant prefectures indicating SC network resilience.	Similarity: Similar to Kosasih and Brintrup [39], which employed a GNN model for supply chain analysis and utilised Integrated Gradient to highlight influential features. Difference: Detailed Analysis: Provides a granular, region-, industry-, and trade-specific trend analysis, adding depth to existing research on supply chain disruptions. Comprehensive Trade Indicators: Analyse the importance of trades using multiple indicators considering the direction of trade and tier in the supply chain, offering a detailed understanding of supply chain dynamics not extensively covered in the literature. Geographical Perspective: Conducts geographical analysis to identify key areas for recovery, a novel contribution to the literature that often overlooks the SCN's spatial dimension.

6.2. Implications

In the field of SCRM, a number of data-driven methods including ML have been proposed in recent years [54]. However, most of them use only the internal information of a single company, such as its products, factories, warehouses, etc. as input data. When obtaining data from other companies, the maximum limit would be the basic information from business partners or information from investor relation documents. The data on companies across Japan collected by the credit research firms that we used allow us to utilise information on not only individual companies but also other companies that have any relationship with these companies. This enables the prediction model to comprehensively collect and learn from the information on the supply chain, and output promising prediction results to provide useful insights for decision-making.

The interpretation of ML models is difficult and their effective use in practice has been challenging [11]. However, the explanatory variables used in our prediction model are all kinds of extracted information, including company profiles, hazards, network centrality of companies based on network theory, transaction-related, and geographical information. These not only enhance the performance of the prediction model but also improve its interpretability by identifying influencing factors using the XAI methodology, providing rich decision-making material for SCRM decision-making.

6.2.1. Theoretical implications

Advanced integration of models and diverse data sources. This study highlights the importance of integrating diverse data sources to enhance the comprehensiveness and accuracy of predictive models in SCRM. By utilising information from a variety of companies, including those indirectly related, the study provides a more holistic view of the supply chain. Furthermore, by incorporating both inter-firm relationships and geographical information, the model addresses the complex, non-linear interactions within SCs that traditional models often overlook. This integration enhances the SCRM theoretical framework by providing a more refined understanding of how inter-firm dynamics influence resilience and recovery.

Enhancing explainability in AI models. The application of XAI techniques, such as SHAP and GNNExplainer, adds a new dimension to the use of AI in SCRM. By making the model's decision-making process transparent, our approach allows for better interpretation and trust in ML predictions, addressing the black box problem commonly associated with ML models in corporate settings[11].

6.2.2. Practical implications

Improved decision-making for SCRM. The identification of key factors that influence sales recovery post-hazard, such as the eigenvector centrality of business partners and the physical distance to suppliers, provides actionable insights for SCRM managers. Companies can strategically allocate resources to strengthen relationships with high-centrality partners and consider geographic proximity in their supplier selection to enhance resilience.

Policy development for hazard recovery. For government policymakers, the ability to predict which firms are critical for economic recovery post-hazard can guide the formulation of targeted support measures. By identifying vulnerable SCs and crucial firms, policymakers can design more effective interventions, such as financial aid or tax relief, to support economic stability and recovery at the microeconomic level.

6.3. Use case scenarios

Considering real-world use case scenarios, we believe that our research results are useful to two types of decision-makers. First, corporate SCRM managers can use our research findings to develop SCRM strategies. Our methodology allows us to predict market demand after a hazard by predicting supply chain sales. This enables a preliminary understanding of the impact on the company's own affiliated companies. In addition, by identifying the influencing factors for sales recovery, damage can be mitigated through focused investment and improvement measures. It is considered that having a resilient SC is an essential condition for increasing firms' competitive advantage in the market and enhancing their industry position.

The second type of user can be the SCRM-related policymakers in government organisations. Our results provide a spatial picture of firms that are important for post-hazard economic recovery and supply chains that are vulnerable to SCD through predicting sales at the firm level and the extraction of important transactions in supply chains across Japan. These results can be used to guide companies in building resilient supply chains in normal times and to provide appropriate support measures at the firm level, such as post-hazard taxation measures and financial assistance. This would support the practice of ‘evidence-based policymaking’ promoted by the Japanese government in recent years and contribute to transparent and credible policies.

7. Conclusion

This study proposes a method for predicting sales of individual firms after SCD from natural hazards, based on a GNN that considers inter-firm topology. It trains a model with seven types of inter-firm networks and 67 firm features extracted from actual supply chain data. The proposed GNN model outperforms the baseline model without considering inter-firm topology. In addition, XAI methodology is used for prediction model interpretation to analyse the factors affecting the sales of firms in the post-hazard period in two aspects: firm features and the supply chain structure features. The results make several contributions to SCRM practice and decision-making in industry and government organisations.

However, the study has some limitations that need to be addressed in future research. First, updating the study with recent hazard data and expanding the data sources to include international firm data will enhance the relevance and accuracy of the findings, ensuring the ongoing applicability of the research. As the present analysis is based on data from 2017 to 2018 because the hazard occurred in 2017, it is important to incorporate more recent data. Furthermore, to better understand the long-term economic impacts, it is essential to analyse data spanning several years post-hazard. Second, **f**uture research should integrate fluctuations in consumer demand into the sales prediction model. This enhancement will provide a more accurate estimation of the impact of hazards on the macroeconomy, considering the dynamic nature of consumer behaviour in post-hazard scenarios. Third, **f**urther testing of the model is needed by applying it to different types of natural and man-made hazards. For instance, extending the model to analyse the impact of earthquakes, typhoons, and the COVID-19 pandemic will improve our understanding of supply chain resilience across diverse scenarios. Forth, **f**uture research could delve deeper into the policy implications of the findings, particularly how governments can leverage predictive models for hazard preparedness and response. This includes exploring the economic impacts of different policy interventions and their effectiveness in enhancing supply chain resilience. By addressing these areas, future research can build on our findings to provide a more comprehensive and robust framework for SCRM to deal with natural hazards.

Disclosure statement

The authors report there are no competing interests to declare.

CRediT authorship contribution statement

Shaofeng Yang: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yoshiki Ogawa:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Koji Ikeuchi:** Writing – review & editing, Methodology, Conceptualization. **Ryosuke Shibasaki:** Writing –

review & editing, Resources, Methodology, Conceptualization. **Yuuki Okuma:** Writing – review & editing, Methodology, Conceptualization.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used DeepL (https://www.deepl.com/) in order to translate. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This study was supported by JSPS KAKENHI Grant Number JP20K15001, and was carried out as part of joint research with Teikoku Databank Ltd. We would like to thank Editage (www.editage.com) for English language editing.

References

[1] M. Coronese, F. Lamperti, K. Keller, F. Chiaromonte, A. Roventini, Evidence for sharp increase in the economic damages of extreme natural disasters, Proc. Natl. Acad. Sci. USA 116 (43) (2019) 21450–21455, https://doi.org/10.1073/pnas.1907826116.
[2] H.C. Winsemius, Jeroen C.J. H. Aerts, Beek van, P.H. Ludovicus, Marc F.P. Bierkens, A. Bouwman, B. Jongman, Jaap C.J. Kwadijk, W. Ligtvoet, Paul L. Lucas, Detlef P. van Vuuren, Philip J. Ward, Global drivers of future river flood risk, Nat. Clim. Change 6 (4) (2016) 381–385, https://doi.org/10.1038/nclimate2893.
[3] AON, Weather, climate & catastrophe insight: 2019 annual report, https://www.aon.com/global-weather-catastrophe-natural-disasters-costs-climate-change-2020-annual-report/index.html, 2020.
[4] I. Manuj, J.T. Mentzer, Global supply chain risk management, J. Bus. Logist. 29 (1) (2008) 133–155, https://doi.org/10.1002/j.2158-1592.2008.tb00072.x.
[5] W.J.W. Botzen, O. Deschenes, M. Sanders, The economic impacts of natural disasters: a review of models and empirical studies, Rev. Environ. Econ. Pol. 13 (2) (2019) 167–188, https://doi.org/10.1093/reep/rez004.
[6] A. Nair, J.M. Vidal, Supply network topology and robustness against disruptions – an investigation using multi-agent model, Int. J. Prod. Res. 49 (5) (2011) 1391–1404, https://doi.org/10.1080/00207543.2010.518744.
[7] F. Feng, X. He, X. Wang, C. Luo, Y. Liu, T.-S. Chua, Temporal relational ranking for stock prediction, ACM Trans. Inf. Syst. 37 (2) (2019) https://doi.org/10.1145/3309547, Article 27.
[8] J. Brache, C. Felzensztein, Geographical co-location on chilean sme's export performance, J. Bus. Res. 105 (2019) 310–321, https://doi.org/10.1016/j.jbusres.2017.11.044.
[9] N. Ampazis, Forecasting demand in supply chain using machine learning algorithms, Int. J. Artif. Life Res. 5 (1) (2015) 56–73, https://doi.org/10.4018/IJALR.2015010104.
[10] I. Minis, Applications of neural networks in supply chain management, in: Handbook of Research on Nature-Inspired Computing for Economics and Management, IGI Global, 2007, pp. 589–607.
[11] D. Ni, Z. Xiao, M.K. Lim, A systematic review of the research trends of machine learning in supply chain management, International Journal of Machine Learning and Cybernetics 11 (7) (2020) 1463–1482, https://doi.org/10.1007/s13042-019-01050-0.
[12] G. Mugurusi, P.N. Oluka, Towards explainable artificial intelligence (xai) in supply chain management: a typology and research agenda. Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems, in: IFIP WG 5.7 International Conference, APMS 2021, Nantes, France, September 5–9, 2021, Proceedings, Part IV, 2021.
[13] S. Kelly, Estimating economic loss from cascading infrastructure failure: a perspective on modelling interdependency, Infrastructure Complexity 2 (1) (2015) 7, https://doi.org/10.1186/s40551-015-0010-y.
[14] S. Hallegatte, An adaptive regional input-output model and its application to the assessment of the economic cost of katrina, Risk Anal. 28 (3) (2008) 779–799, https://doi.org/10.1111/j.1539-6924.2008.01046.x.
[15] A. Rose, I. Sue Wing, D. Wei, A. Wein, Economic impacts of a California tsunami, Nat. Hazards Rev. 17 (2) (2016) 04016002, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000212.
[16] A. Nagurney, D. Hassani, O. Nivievskyi, P. Matyshev, Quantification of international trade network performance under disruptions to supply, transportation, and demand capacity, and exchange rates in disasters, Transportation, and Demand Capacity, and Exchange Rates in Disasters (2024) https://doi.org/10.2139/ssrn.4810971, March 29, 2024.
[17] D. Ivanov, Predicting the impacts of epidemic outbreaks on global supply chains: a simulation-based analysis on the coronavirus outbreak (covid-19/sars-cov-2) case, Transport. Res. E Logist. Transport. Rev. 136 (2020) 101922, https://doi.org/10.1016/j.tre.2020.101922.
[18] T.-M. Choi, Risk Analysis in Logistics Systems: A Research Agenda during and after the Covid-19 Pandemic, vol. 145, Elsevier, 2021 102190.
[19] A. Nagurney, Game theory and the covid-19 pandemic, in: Tutorials in Operations Research: Emerging Optimization Methods and Modeling Techniques with Applications, INFORMS, 2021, pp. 83–130, https://doi.org/10.1287/educ.2021.0226.
[20] V.M. Carvalho, M. Nirei, Y.U. Saito, A. Tahbaz-Salehi, Supply chain disruptions: evidence from the great east Japan earthquake, Q. J. Econ. 136 (2) (2021) 1255–1321, https://doi.org/10.1093/qje/qjaa044.
[21] H. Inoue, Y. Todo, Firm-level propagation of shocks through supply-chain networks, Nat. Sustain. 2 (9) (2019) 841–847, https://doi.org/10.1038/s41893-019-0351-x.
[22] A. Gill, M. Lalith, Y. Ogawa, H. O-tani, T. Ichimura, K. Fujita, F. Gatti, M. Hori, Seamless simulations of earthquake disasters and economy at 1: 1 scale utilizing big-data, in: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2021.
[23] S. Yang, Y. Ogawa, K. Ikeuchi, Y. Akiyama, R. Shibasaki, Estimation of the economic im-pact of urban flood through the use of big data on inter-branch office transactions, Journal of Japan Society for Natural Disaster Science 38 (S06) (2019) 185–199, https://doi.org/10.24762/jndsj.38.S06_185.
[24] X.-F. Shao, Supply chain characteristics and disruption mitigation capability: an empirical investigation in China, Int. J. Logist. Res. Appl. 16 (4) (2013) 277–295, https://doi.org/10.1080/13675567.2013.815695.
[25] B. Lawson, A. Potter, F.K. Pil, M. Holweg, Supply chain disruptions: the influence of industry and geography on firm reaction speed, Int. J. Oper. Prod. Manag. 39 (9/10) (2019) 1076–1098, https://doi.org/10.1108/IJOPM-04-2018-0225.
[26] S. Kohli, G.T. Godwin, S. Urolagin, Sales Prediction Using Linear and Knn Regression, Advances in Machine Learning and Computational Intelligence, Singapore, 2021.

[27] D. Rohaan, E. Topan, C.G.M. Groothuis-Oudshoorn, Using supervised machine learning for b2b sales forecasting: a case study of spare parts sales forecasting at an after-sales service provider, Expert Syst. Appl. 188 (2022) 115925, https://doi.org/10.1016/j.eswa.2021.115925.

[28] S.S. Kamble, A. Gunasekaran, V. Kumar, A. Belhadi, C. Foropon, A machine learning based approach for predicting blockchain adoption in supply chain, Technol. Forecast. Soc. Change 163 (2021) 120465, https://doi.org/10.1016/j.techfore.2020.120465.

[29] R. Rajesh, An introduction to grey causal modelling (gcm): applications to manufacturing, supply chains, resilience, and sustainability, Artif. Intell. Rev. 56 (7) (2023) 6267–6293, https://doi.org/10.1007/s10462-022-10314-1.

[30] G. Singh, R. Rajesh, Y. Daultani, S.C. Misra, Resilience and sustainability enhancements in food supply chains using digital twin technology: a grey causal modelling (gcm) approach, Comput. Ind. Eng. 179 (2023) 109172, https://doi.org/10.1016/j.cie.2023.109172.

[31] G. Baryannis, S. Dani, G. Antoniou, Predicting supply chain risks using machine learning: the trade-off between performance and interpretability, Future Generat. Comput. Syst. 101 (2019) 993–1004, https://doi.org/10.1016/j.future.2019.07.059.

[32] T. Ghazal, H. Alzoubi, Fusion-based supply chain collaboration using machine learning techniques, Intelligent Automation & Soft Computing 31 (3) (2022) 1671–1687, https://doi.org/10.32604/iasc.2022.019892.

[33] S. Yang, Y. Ogawa, K. Ikeuchi, R. Shibasaki, Y. Okuma, Modelling the behaviour of corporations during the flood damage recovery process using multi-agent deep reinforcement learning, Journal of Flood Risk Management 15 (4) (2022) e12845, https://doi.org/10.1111/jfr3.12845.

[34] S. Aureli, M. Del Baldo, The network contract and its influence on strategic decision of internationalization: evidence from Italy, Int. J. Appl. Behav. Econ. 2 (2) (2013) 23–40, https://doi.org/10.4018/ijabe.2013040102.

[35] V.K. Chauhan, S. Perera, A. Brintrup, The relationship between nested patterns and the ripple effect in complex supply networks, Int. J. Prod. Res. 59 (1) (2021) 325–341, https://doi.org/10.1080/00207543.2020.1831096.

[36] S. Hu, S. Gao, L. Wu, Y. Xu, Z. Zhang, H. Cui, X. Gong, Urban function classification at road segment level using taxi trajectory data: a graph convolutional neural network approach, Comput. Environ. Urban Syst. 87 (2021) 101619, https://doi.org/10.1016/j.compenvurbsys.2021.101619.

[37] F. Yuan, Y. Xu, Q. Li, A. Mostafavi, Spatio-temporal graph convolutional networks for road network inundation status prediction during urban flooding, Comput. Environ. Urban Syst. 97 (2022) 101870, https://doi.org/10.1016/j.compenvurbsys.2022.101870.

[38] D. Wu, Q. Wang, D.L. Olson, Industry classification based on supply chain network information using graph neural networks, Appl. Soft Comput. 132 (2022) 109849. https://doi.org/10.1016/j.asoc.2022.109849.

[39] E.E. Kosasih, A. Brintrup, A machine learning approach for predicting hidden links in supply chain with graph neural networks, Int. J. Prod. Res. (2021) 1–14, https://doi.org/10.1080/00207543.2021.1956697.

[40] K. Tamura, H. Takayasu, M. Takayasu, Diffusion-localization transition caused by nonlinear transport on complex networks, Sci. Rep. 8 (1) (2018) 5517, https://doi.org/10.1038/s41598-018-23675-x.

[41] Ministry of Land Infrastructure Transport and Tourism, Disaster and disaster prevention information: damage caused by the heavy rainfall in july 2018, Retrieved 7/26 from. https://www.mlit.go.jp/saigai/saigai_180703.html, 2018.

[42] H. Liu, H. Tatano, S. Samaddar, Analysis of post-disaster business recovery: differences in industrial sectors and impacts of production inputs, Int. J. Disaster Risk Reduc. 87 (2023) 103577.

[43] M.M.H. Chowdhury, M. Quaddus, R. Agarwal, Supply chain resilience for performance: role of relational practices and network complexities, Supply Chain Manag.: Int. J. 24 (5) (2019) 659–676, https://doi.org/10.1108/SCM-09-2018-0332.

[44] S.E. Chang, N. Nojima, Measuring post-disaster transportation system performance: the 1995 kobe earthquake in comparative perspective, Transport. Res. Pol. Pract. 35 (6) (2001) 475–494, https://doi.org/10.1016/S0965-8564(00)00003-3.

[45] D. Acemoglu, V.M. Carvalho, A. Ozdaglar, A. Tahbaz-Salehi, The network origins of aggregate fluctuations, Econometrica 80 (5) (2012) 1977–2016, https://doi.org/10.32604/iasc.2022.019892.

[46] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, arXiv preprint arXiv:1609.02907 (2016).

[47] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[48] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.

[49] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2016) https://doi.org/10.1145/2939672.2939785, San Francisco, California, USA.

[50] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm: a highly efficient gradient boosting decision tree, Adv. Neural Inf. Process. Syst. 30 (2017).

[51] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, A. Gulin, Catboost: unbiased boosting with categorical features, Adv. Neural Inf. Process. Syst. 31 (2018).

[52] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017).

[53] Z. Ying, D. Bourgeois, J. You, M. Zitnik, J. Leskovec, Gnnexplainer: generating explanations for graph neural networks, Adv. Neural Inf. Process. Syst. 32 (2019).

[54] P. Wichmann, A. Brintrup, S. Baker, P. Woodall, D. McFarlane, Extracting supply chain maps from news articles using deep neural networks, Int. J. Prod. Res. 58 (17) (2020) 5320–5336, https://doi.org/10.1080/00207543.2020.1720925.

[55] Geospatial Information Authority of Japan, Information on the heavy rainfall in july 2018, Retrieved 7/26 from. https://www.gsi.go.jp/BOUSAI/H30.taihuu7gou.html, 2018.