Scientific Journal of Informatics
Vol. 10, No. 4, Nov 2023

p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Mapping of Social Vulnerability to Natural Hazards in Geodemographic Analysis Using Fuzzy Geographically Weighted Clustering

Deden Istiawan[1*], Ratri Wulandari[2], Sulastri[3]
[1]Software Engineering Department, Faculty of Science and Technology,
Institut Teknologi Statistika dan Bisnis Muhammadiyah Semarang, Indonesia
[2]Actuarial Science Department, Faculty of Science and Technology,
Institut Teknologi Statistika dan Bisnis Muhammadiyah Semarang, Indonesia
[3]Information Systems Department, faculty of information technology and industry,
Universitas Stikubank, Indonesia

Abstract.
Purpose: Assessing social vulnerability is essential for addressing environmental risks by developing suitable adaptation strategies and fostering a resilience mindset. However, relying solely on an index-based approach to measure social vulnerability has limitations as it only provides a broad overview. It is essential to recognize that various regions are influenced by distinct factors contributing to social vulnerability. This study aims to pinpoint specific community factors that impact vulnerability to natural disasters in various districts across Indonesia.
Methods: In this research, we determine the optimal number of clusters with the Cluster Validity Index (CVI). Furthermore, this research applies clustering analysis of social vulnerability to natural disasters at the district level using the Fuzzy Geographically Weighted Clustering (FGWC) algorithm.
Results: This research highlights varying social vulnerability profiles across Indonesia's diverse districts. Specifically, districts on the western side of Sumatra Island, such as Nias and Mentawai, and those in the eastern regions of Indonesia, including Nusa Tenggara, West Sulawesi, Central Sulawesi, North Sulawesi, the Southern Maluku Islands, and Papua, exhibit the most noticeable vulnerability. This vulnerability is particularly evident in socioeconomic indicators, family composition, housing conditions, and educational access.
Novelty: The results of this study provide valuable support for the government as a policymaker. By identifying priority areas and tailoring policies to address critical social vulnerability issues in each district, especially in the most vulnerable areas, the research offers a practical framework for targeted and effective disaster risk reduction and mitigation efforts.

Keywords: Disaster, Disaster mitigation, Social vulnerability, Clustering
Received September 2023 / **Revised** October 2023 / **Accepted** October 2023

INTRODUCTION

Indonesia is widely recognized as one of the countries most susceptible to natural hazards. Over recent years, the nation has witnessed remarkable events, including earthquakes, floods, and tsunamis. According to the World Risk Report 2022, Indonesia holds the third position out of 172 countries, and it is one of seven Asian nations among the top ten countries with the highest disaster risks [1]. Natural disasters present a substantial threat to the country's economy and human life. What is particularly noteworthy is that the same type of natural disaster can have a disproportionate impact on different groups of people, largely contingent on the resilience and preparedness of each group [2]. The capacity of community groups to effectively respond to disasters is intimately tied to their vulnerability to the disaster. Addressing these disparities in vulnerability is a crucial aspect of disaster risk reduction and management efforts [3], [4].

The initial concept of vulnerability pertains to potential losses during a natural disaster [5]. Social vulnerability, on the other hand, characterizes the extent, scale, or level of exposure and the incapacity to

*Corresponding author.
Email addresses: deden.istiawan@itesa.ac.id (Istiawan), ratri.wulandari@itesa.ac.id (Wulandari), sulastri@edu.unisbank.ac.id (Sulastri)
DOI: 10.15294/sji.v10i4.47418

cope with and rebound from the adverse consequences of a hazard or disaster [6], [7]. The impact of disasters on different community groups continually varies based on various socioeconomic and social-environmental factors that come into play [8]–[10]. Socially vulnerable communities are more likely to face higher casualties and property damage rates during disasters and tend to experience slower recovery processes afterward [11]. Furthermore, social vulnerability attributable to natural disasters is not uniformly distributed across regions [12]. It is crucial to clearly understand the spatial distribution of factors contributing to social vulnerability in each region to plan appropriate actions. This knowledge informs targeted and effective disaster preparedness and response efforts [13].

Social vulnerability assessment plays a pivotal role in disaster mitigation efforts, the development of suitable adaptation strategies, and cultivating a culture of resilience in the face of environmental hazards [14], [15]. Moreover, it aims to discern which specific groups of people are most susceptible to the impacts of natural disasters and pinpoint the primary factors contributing to social vulnerability [16]. Identifying these community groups in each region and understanding the mechanisms and reasons behind their vulnerability is fundamental in devising effective mitigation measures, allocating resources for preparedness, and formulating appropriate strategies [17], [18]. By addressing these critical questions, this research offers valuable policy insights for reducing social vulnerability in every district across Indonesia. It achieves the objective by mapping out priority areas of social vulnerability to natural disasters, enabling policymakers to direct resources and efforts toward the most vulnerable regions and populations, ultimately enhancing disaster resilience and preparedness on a local level.

The technology of computer science may be an instrumental strategic tool for efficient management [19]. Research on social vulnerability has a long history, with Susan L. Cutter being the first to model disaster vulnerability. She created the Social Vulnerability Index (SoVI) for the United States using the Principal Component Analysis (PCA) method [5], [20]. This method has also developed widely and has been used by numerous studies in Indonesia. Measuring social vulnerability to natural disasters in Indonesia was first initiated by Birkmann et al. [21], which calculated the social vulnerability of the community in Padang City. Then, in the broader area coverage, SoVI calculations were carried out by Siagian et al. [22], which measured social vulnerability at the district/city level using factor analysis methods. Kurniawan et al. [23] employed a distinct methodological approach, Structural Equation Modeling (SEM), for their Social Vulnerability Index (SoVI) calculations. Meanwhile, Nugraha et al. [24] conducted research that integrated economic and physical vulnerability to evaluate the extent of social vulnerability arising from natural disasters in Jepara Regency. Wijaya and Halim's [25] study identified social vulnerability and the primary factors affecting vulnerability at the district level in Indonesia using Principal Component Analysis (PCA).

Assessing social vulnerability through an index approach has limitations, primarily in providing a broad overview of social vulnerability conditions. It is essential to recognize that various regions are influenced by distinct factors contributing to social vulnerability [26], [27]. Additionally, using the Social Vulnerability Index (SoVI) may overlook the geographical context, even though social vulnerability can vary significantly based on geographical factors unique to each region [28]. Vulnerability assessment should involve measuring the potential damage and delving into the specific reasons why an area is exposed to various hazards. Rufat has pointed out that relying solely on an index to measure vulnerability can lead to regional homogenization and overlook the region-specific factors that impact social vulnerability [29].

To develop a timely and effective disaster risk management strategy, it is crucial to have a comprehensive understanding of the nature of the disaster and to assess the level of social vulnerability in each region [30]. Comparing social vulnerability as a metric becomes more meaningful when accompanied by visualization and spatial distribution methods. Analyzing spatial information is of utmost importance in disaster risk mitigation and reduction because the impact of hazards on communities can vary significantly from one region to another [31]. One approach that can be employed involves grouping regions based on socioeconomic statistical data through Geodemographic Analysis [32].

Geodemographic analysis (GDA) is an approach that can be used to manage, identify, and visualize the social vulnerability of an area [33]. The GDA approach extracts unique and hidden information from data and has proven to be widely applied in supporting effective policymaking [34]. The main goal of GDA is to produce clusters based on the socioeconomic status of residents in an area [35]. The Fuzzy Geographically Weighted Clustering (FGWC) algorithm is a clustering algorithm suitable for GDA by considering the influence of spatial effects in the form of population size and distance between regions [36].

Given Indonesia's vast geography and diverse socioeconomic conditions, each district or city may exhibit different social vulnerabilities. Given the frequency of natural disasters and the considerable potential for such events in Indonesia, a more in-depth analysis of social vulnerability is essential. Such analysis can be a valuable reference for local and central government efforts to prevent and respond to natural disasters. This research focuses on clustering analysis of social vulnerability to natural disasters at the district/city level using the Fuzzy Geographically Weighted Clustering (FGWC) algorithm. The aim is to provide a comprehensive picture of what factors influence social vulnerability in each region across Indonesia to provide critical input for disaster preparedness and emergency response efforts.

METHODS

Dataset Description

This study built upon the concepts of social vulnerability, drawing from the work of Cutter et al. [20] as its theoretical framework on social vulnerability as its foundation. In this research, 19 variables were employed and categorized into eight indicators, which were informed by prior studies [2], [10], [43], [12], [18], [37]–[42]. The dataset utilized in this study was sourced from secondary data provided by the Central Statistics Agency (BPS) and the National Team for the Acceleration of Poverty Reduction (TNP2K). The specific indicators and variables used in this study are stated in Table 1.

Table 1. Indicators and variables of social vulnerability

Indicators	Variables	References
Socio-economic status	Percentage of population that is not working (X_1) Percentage of population who have health social security (X_2) Percentage of poor people (X_3)	[12], [18], [40]
Age	Percentage of population aged 5 years and under (X_4) Percentage of population aged 65 years and over (X_5)	[2], [18], [22], [37], [39], [40], [42]
Family Structure	Average number of family members (X_6) percentage of households headed by females (X_7)	[10], [12], [18], [22], [37], [39], [40]
Gender	Percentage of female population (X_8)	[10], [12], [18], [22], [40], [42]
Population growth	Percentage population growth (X_9)	[3], [29], [39]
Housing Quality	The percentage of households that do not utilize electricity (X_{10}) The percentage of households lacking a drainage system (X_{11}) The percentage of households that utilize piped water (X_{12}) The percentage of households residing in regions or areas that are prone to disasters (X_{13})	[10], [18], [40], [41]
Homeownership	Percentage of households that rent a house (X_{14})	[10], [41], [42]
Education	The percentage of the population with low education (X_{15}) Percentage of the population who cannot read and write (X_{16}) Percentage of Households that did not receive disaster preparedness training (X_{17})	[10], [22], [40]
Special needs population	Percentage of individuals with disabilities (X_{18}) Percentage of individuals who have a chronic disease (X_{19})	[2], [18], [39]–[41]

Data Preprocessing

Low-quality data will lead to low-quality data analysis results [44]. Data preprocessing is essential so that data can be processed according to the tools used. Data preprocessing in this study was divided into several steps: data cleaning and data transformation.

Geodemographic Analysis (GDA)

The primary purpose of Geodemographic Analysis (GDA) is to create clusters based on the social and economic status of an area's population, making it easy to predict people's behavior if we know where they live and their habits [45]. GDA combines Geographical Information systems (GIS) and data mining algorithms. GDA uses clustering techniques to classify geodemographic data to facilitate analysis [35].

Fuzzy Geographically Weighted Clustering (FGWC)

Fuzzy Geographically Weighted Clustering (FGWC) is an improvement of the Fuzzy C-Means algorithm. The fuzzyfication process is a calculation of the crisp value or the input value into the degree of membership [46]. Fuzzyfication, which is the process of converting system inputs that have firm values into linguistic variables using membership functions stored in the fuzzy knowledge base [47]. It is more geographically aware because it involves the effects of population and distance between regions in calculating membership weights for each observation [48]. FGWC considers the influence of one region on another as the product of population and distance between the regions. The determination of group membership in FGWC calculated at each iteration is shown in the following equation:

$$\mu_i' = \alpha\mu_i + \beta \frac{1}{A} \Sigma_j^n \; w_{ij}\mu_j \tag{1}$$

Where μ_i' is the new membership value of object i. μ_i is the old membership value of object i. w_{ij} is a measure of weighting the number of interactions between regions, and A is a value to ensure the weighting value is not more than 1.
α and β are multipliers for the old membership value and the weighted value of the average membership of other observation units. The α dan β values are defined as follows:

$$\alpha + \beta = 1 \tag{2}$$

The membership weight (w_{ij}) is defined as follows:

$$w_{ij} = \frac{(m_i\, m_j)^b}{d_{ij}^a} \tag{3}$$

Where m_i is the population size of region i, m_j is the population size of region j, and d_{ij} is the distance between region i and region j. a and b are user-defined parameters. If the population effect is considered as important as the distance effect, then $a = b = 1$. FGWC incorporates geographical elements in Geodemographic Analysis (GDA) so that clusters are sensitive to environmental effects and will affect the values of cluster centers to create "geographically aware" clusters.

Cluster Validity Index (CVI)
Cluster performance is not only compared based on the objective function but also using validity indices. Validity indices used to evaluate the optimal number of clusters are the Partition Coefficient (PC), Classification Entropy (CE), Partition Index (SC), Separation Index (S), IFV Index, and Xie and Beni Index (XBI). A brief explanation of the validity indices is as follows.

Partition Coefficient (PC)
The Partition Coefficient (PC) reflects the overlap of fuzzy subsets and depends on the membership coefficients. The maximum PC value expresses the optimal number of groups. The partition coefficient (PC) is calculated using the following equation:

$$PC = \frac{1}{N}(\Sigma_{i=1}^{c}\Sigma_{j=1}^{N}\, \mu_{ij}^2) \tag{4}$$

Classification Entropy (CE)
Classification Entropy (CE) represents the fuzziness between clusters. Classification Entropy measures the degree of fuzziness of the cluster partition. The CE value is between [0, log c] from the equation. The minimum CE value expresses the optimal number of clusters.

$$CE = \frac{1}{N}\Sigma_{i=1}^{c}\Sigma_{j=1}^{N}\, \mu_{ij} \log\, \mu_{ij} \tag{5}$$

Partition Index (SC)
The Partition Index (SC) compares the degree of cluster compactness and separation. A minimum SC value indicates excellent clustering quality.

$$\mathrm{SC} = \sum_{j=1}^{c} \frac{\sum_{i=1}^{n} \mu_{ij}^{m} \|x_j - v_i\|^2}{n_j \sum_{k=1}^{c} \|v_k - v_j\|^2} \tag{6}$$

Separation Index (S)
The Separation Index (S) calculates the compactness and separation of each cluster. The minimum S index value expresses the optimal number of groups.

$$S = \frac{\Sigma_{i=1}^{c}\Sigma_{j=1}^{N}(\mu_{ij})^2\ ||x_j - v_i||^2}{N min_{i,k}||x_j - v_i||^2} \tag{7}$$

IFV Index
The IFV index has biased robustness and stability when validating spatial clustering. The maximum IFV index value reflects good spatial cluster separation.

$$IFV = -\frac{1}{c}\sum_{j=1}^{c}\left\{\frac{1}{n}\sum_{i=1}^{n}\mu_{ij}^{2}\left[log_2\, c - \frac{1}{n}log_2\mu_{ij}\right]\right\}\frac{max_{k,j}\|v_k - v_j\|^2}{\overline{\sigma_d}} \quad (8)$$

Xie and Beni index (XBI)

Together with the SC index, the XBI index indicates the variation between clusters and the clarity of separation. The minimum XBI index value expresses the optimal number of clusters.

$$XBI = \frac{\Sigma_{i=1}^{C}\Sigma_{j=1}^{N}\left(\mu_{ij}\right)^{m}||x_j - v_i||^2}{Nmin_{i,j}||x_k - v_i||^2} \quad (9)$$

Good cluster separation can be seen from the maximum PC and IFV values and the minimum CE, SC, S, and XBI. In this study, we implemented the FGWC algorithm and cluster validity index with the R programming language using the number of clusters from 2 to 4 to see the results and stability of the data clustering results.

RESULTS AND DISCUSSIONS

This section presents the results of clustering social vulnerability using the Fuzzy Geographically Weighted Clustering (FGWC) technique. The most optimal clustering results achieved with FGWC are used to characterize social vulnerability at the district level in Indonesia. Experiments were conducted on a computing platform with a 1.6 GHz Intel Core i5 CPU, 8 GB RAM, and Microsoft Windows 11 Home 64-bit operating system. Social vulnerability analysis involves identifying critical issues within each group. This study begins this analysis by forming groups using the FGWC algorithm, with groups ranging from 2 to 4. The initial parameter settings for FGWC in this study are as follows $m = 1.5, \alpha = 0.7, \beta = 0.3, a = 1, b = 1, \varepsilon = 1e - 5$.

Table 2. Evaluation of optimal number of clusters

Cluster	PC	CE	SC	S	IFV	XB
2	0.57022	0.61991	6.46258	2.17321	0.54181	12.0618
3	0.45244	0.92108	2.48717	2.02221	3.48908	9.89732
4	0.34373	1.19240	1.65164	3.85422	7.11897	8.17455

This research begins by finding the optimal number of clusters using the cluster validity index consisting of Partition Coefficient (PC), Classification Entropy (CE), Partition Index (SC), Separation Index (S), IFV Index, Xie and Beni Index (XBI) with the number of clusters from 2 to 5. Based on Table 2 of the six cluster validity indexes used in this study, three cluster validity indexes, namely SC, IFV, and XBI, show that the optimal number of clusters is 3. Thus, social vulnerability analysis uses 4 clusters as the most optimal number in this study.

By referring to Table 3, we can identify each cluster's social vulnerability characteristics. Clusters with the highest average in a particular variable indicate that the most dominating problems related to that variable are in the districts included in that cluster. Bold values indicate that the cluster is more vulnerable to a particular feature. The results from Table 3 imply that we cannot quickly determine which clusters are more vulnerable, as each cluster has different vulnerability characteristics.

Table 3. Social Vulnerability Characteristics

Variables	Cluster			
	1	2	3	4
Population not working (X_1)	6.42189	7.30181	6.37091	6.42179
Population with social health insurance (X_2)	20.72325	32.28580	14.11441	20.72451
Population in poverty (X_3)	12.59386	8.99397	17.59802	12.59340
Population aged 5 years and under (X_4)	9.36421	9.12211	10.05198	9.36419
Population aged 65 years and above (X_5)	4.87592	4.19856	3.83476	4.87589
Average family size (X_6)	3.91965	3.90599	4.15382	3.91964
Female head of household (X_7)	15.36038	14.20110	13.38710	14.20111
Female population (X_8)	49.56621	49.53100	49.36673	49.53097
Population growth (X_9)	1.31971	1.57940	1.48569	1.31973
Households that do not utilize electricity (X_{10})	4.65263	2.71275	13.49251	4.65228
Households without a drainage system (X_{11})	16.20067	9.19356	31.36892	16.19947
Households using piped water (X_{12})	15.03810	40.16055	11.81761	15.03825
households residing in regions prone to disasters (X_{13})	91.91438	91.87238	89.26508	91.91470
Households that rent a house (X_{14})	6.04764	15.07043	4.50298	6.04806
Population with low education (X_{15})	33.85385	25.42358	34.41166	33.85320
People who cannot read and write (X_{16})	8.04448	5.92778	11.17058	8.04420
Households that did not receive disaster preparedness training (X_{17})	98.68123	97.59715	98.66998	98.68121
Disabled individuals (X_{18})	6.53512	6.90407	6.49453	6.53502
Individuals with chronic diseases (X_{19})	6.40733	7.32666	6.40115	6.40723

As shown in Figure 1, among the 514 districts analyzed, Cluster 1 comprises 150 districts, Cluster 2 includes 109 districts, Cluster 3 encompasses 121 districts and Cluster 4 consists of 134 districts. Based on the average variables within each cluster, districts in Cluster 1 exhibit social vulnerability issues related to the population aged 65 years and over, female heads of households, and the female population. These districts are distributed across various islands, including Sumatra, Java, South Kalimantan, and Sulawesi. The elderly population faces heightened vulnerability during natural disasters due to their limitations, especially in terms of health, compared to other age groups. Typically, this group is more susceptible to health-related challenges than younger age groups. Additionally, they often rely on assistance and care from family members or healthcare workers. When a natural disaster occurs, the individuals responsible for their care may be distracted or unable to provide adequate assistance, elevating the health and safety risks for the elderly population.

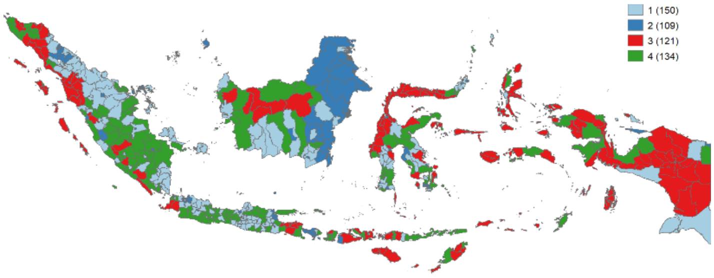

Figure 1. Cluster-based social vulnerability map using FGWC

In addition, within cluster 1, social vulnerability issues also relate to the female population and female-led households. The National Disaster Management Agency (BNPB) notes that the level of vulnerability of women, girls, and adolescents tends to increase in natural disaster situations. One of the contributing factors is the lack of access to information and participation of women in socialization activities on disaster management at the village level. This contributes to the high number of victims due to disasters. This aligns with the findings of research conducted by Sohrabizadeh et al. [49], which indicates that the female population typically has reduced access to resources and information that significantly impact their physical and mental well-being during and after a disaster. Such insights are crucial in the context of disaster management. This reflects the gender inequality that remains a severe problem in many societies, especially in disaster-prone areas. In addition, female-led households tend to be at greater risk of poverty than male-led households. Women are, therefore, a vulnerable group as they can experience high-stress levels when natural disasters occur. This is especially true if they act as the head of the household, who must be responsible for meeting the needs of all family members. Research by Lixin et al. [50] has demonstrated that the female population and the proportion of female-headed households have a notable and statistically significant impact on the extent and severity of social vulnerability to natural disasters.

In Cluster 2, as indicated by the average values of the variables within this cluster, the districts encompassed by Cluster 2 exhibit social vulnerability issues related to several factors. These factors include unemployment rates, population growth, households relying on piped water, households in rental housing, disabled individuals, and individuals with chronic diseases. Cluster 2 districts are spread across a limited portion of Sumatra, Java, and Sulawesi islands, with the majority situated on the island of Kalimantan. Notably, many of these districts are concentrated in North Kalimantan Province, East Kalimantan Province, and South Kalimantan Province. One significant factor contributing to the high population growth in this region is the Indonesian government's transmigration program initiated in 1954. This program was designed to alleviate population density on other islands, particularly Java, by encouraging population settlement in less densely populated areas, notably on the island of Kalimantan. The transmigration program promoted by the government as one of the population policies programs only sometimes brings good impacts. Behind the potential for a more secure life, this program also causes many people to seek rental housing. Renters are a vulnerable group that can experience more significant difficulties when natural disasters occur. In such situations, they may need help paying rent or finding a new place to live [41], and renters are harder to track, and there may be little data available [42]. The lack of inclusive disaster information, planning, and education for people with disabilities and chronic illnesses often results in people with disabilities and chronic illnesses being isolated and trapped within their homes and away from sources of assistance during a disaster [18]. Hence, it is crucial to prioritize the well-being of migrants by promoting their integration, enhancing land productivity, and offering support. The government can play a pivotal role by creating employment opportunities and providing training to improve their skills. Moreover, developing water infrastructure is also essential to reduce the costs associated with disaster responses and recovery efforts.

Cluster 3 emerges as the cluster with the most pronounced social vulnerability issues. Based on the average values of variables within this cluster, the districts belonging to Cluster 3 exhibit a range of social vulnerability problems. These problems are related to populations lacking social health insurance, impoverished populations, individuals aged five years and under, the average number of family members in households, households without access to electricity, households lacking proper drainage systems, populations with limited education, and individuals who cannot read or write. The districts in Cluster 3 are distributed across a wide area encompassing much of North Sumatra, West Sumatra, Sumatra as a whole, and Eastern Indonesia, which includes regions like Nusa Tenggara, Sulawesi, Maluku, and Papua. Addressing the specific social vulnerability issues identified in this cluster is crucial for enhancing disaster resilience and community well-being in these areas. These findings are consistent with the research of Wijaya et al. [25], stating that areas that show high levels of social vulnerability are mostly scattered in several regions, such as Nias and Mentawai Islands, Nusa Tenggara, West Sulawesi, Central Sulawesi and North Sulawesi, Southern Maluku Islands, and Papua. The elevated social vulnerability in these regions can be attributed to the low socioeconomic status of the local population. Severe economic inequality and limited access to resources are significant challenges in these areas, further amplifying their vulnerability to natural disasters. Additionally, geographical constraints are crucial in heightening social vulnerability in specific regions. Challenging access and communication to areas like the Nias Islands and the Central Mountains in Papua can impede disaster management and relief efforts, increasing local populations' risk and social vulnerability when facing natural disasters. Therefore, allocating special attention and resources to mitigate social vulnerability and enhance preparedness for natural disasters in these areas is imperative.

The observation that the districts in this cluster generally have larger average household sizes underscores the potential for increased social vulnerability. A higher number of family members, especially children, can heighten social vulnerability, particularly in disasters. Children under the age of 5, in particular, constitute a highly vulnerable group when it comes to disasters [51]. They heavily rely on adults for care and protection and may be unable to save themselves or make sound decisions during emergencies. Therefore, disaster response planners and operational teams must give special consideration to the unique needs of children and provide age-appropriate protection, medical care, psychosocial support, and education during disaster responses. Furthermore, the issues of access to electricity and clean water are prominent in this cluster, closely intertwined with the problem of poverty. Addressing these infrastructure and poverty-related challenges is vital to mitigating social vulnerability in these regions.

People with higher education tend to have more access to resources, information, and skills to help them cope with problems arising from natural disasters [51]. The high illiteracy rate in the region can be a severe problem during natural disasters. Illiteracy, or the inability to read and write, can hinder an individual's ability to access and understand necessary information during crises. This can have a dangerous impact, especially when critical information on evacuation, rescue, or relief needs to be understood and followed quickly. Emergency literacy training programs can help illiterate adults and children understand critical information during disasters. This could include understanding symbols, graphics, or signs used in disaster warnings. The combination of socialization on disaster risk reduction and the development of supportive infrastructure is a practical approach to increasing the region's resilience to disasters.

In Cluster 4, as evidenced by the average values of variables within this cluster, the districts in Cluster 4 exhibit social vulnerability issues primarily related to households residing in disaster-prone areas and households lacking access to disaster training. These districts are dispersed across much of Sumatra and Java, with a significant presence in Kalimantan's western and northern islands and a smaller representation in Sulawesi and Papua. Cluster 4 highlights social vulnerability concerns closely linked to informal education, particularly the lack of disaster training. The fact that approximately a quarter of Indonesia's districts fall into this cluster underscores the pivotal role of education in enhancing preparedness and reducing disaster risk. The absence of disaster education and training can lead to a limited understanding of disasters within the community [7]. This knowledge gap can result in a greater need for awareness about disaster prevention and recovery efforts, especially in natural disasters. Consequently, these areas become more vulnerable to disasters due to the population's need for adequate knowledge about disaster preparedness.

Disaster education can be integrated into the school curriculum as a compulsory subject. This helps ensure that the younger generation grows up with a better understanding of disasters. Disaster training should be addressed to students and adults, such as teachers, health workers, and community leaders. Socialization and education to the community on disaster risk reduction is a crucial first step. It helps raise people's awareness of existing disaster threats and the actions they can take to protect themselves and their communities. With a better understanding, communities will be better prepared to deal with disasters. A comprehensive approach that includes socialization, education, and infrastructure development is critical to reducing disaster risk and improving community preparedness for possible disaster threats. To address this issue, the government should increase its socialization efforts about disasters and how to reduce their risks, especially in areas living in disaster-prone areas. With better education, communities will be better prepared for disasters and can contribute to more effective prevention and recovery efforts.

CONCLUSION

This research underscores the significance of social vulnerability mapping as a vital tool for regional planning and emergency management. By applying the Fuzzy Geographically Weighted Clustering (FGWC) algorithm on social vulnerability indicators, the study identified an optimal number of four clusters, revealing the spatial distribution of social vulnerability in Indonesia. It became evident that each district possesses unique social vulnerability characteristics, highlighting the diversity of vulnerability aspects across regions. The study's findings pave the way for developing more tailored mitigation policies for each district or city based on its specific social vulnerability profile. Cluster 3, spanning most of North Sumatra, West Sumatra, and Eastern Indonesia, emerges as the region with the most pronounced social vulnerability issues, particularly concerning socioeconomic status, family structure, housing quality, and education. In conclusion, this research offers valuable insights into the indicators of social vulnerability to disasters in each district, aiming to support government efforts in designing appropriate programs to mitigate the adverse impacts of natural disasters. Given Indonesia's susceptibility to various natural disasters like earthquakes, tsunamis, volcanic eruptions, and floods, continued research and assessment of social vulnerability and disaster risk at the district level, or even finer scales, remain essential for informed disaster preparedness and response efforts throughout the country.

REFERENCES

[1] F. Atwii *et al.*, “World Risk Report 2022,” Stuttgart, Germany, 2022.

[2] S. A. Zarghami and J. Dumrak, “A system dynamics model for social vulnerability to natural disasters: Disaster risk assessment of an Australian city,” *Int. J. Disaster Risk Reduct.*, vol. 60, no. January, p. 102258, Jun. 2021, doi: 10.1016/j.ijdrr.2021.102258.

[3] K. K. Zander, R. Sibarani, J. Lassa, D. Nguyen, and A. Dimmock, “How do Australians use social media during natural hazards? A survey,” *Int. J. Disaster Risk Reduct.*, vol. 81, no. July, p. 103207, Oct. 2022, doi: 10.1016/j.ijdrr.2022.103207.

[4] P. A. Kaban, R. Kurniawan, R. E. Caraka, B. Pardamean, B. Yuniarto, and Sukim, “Biclustering Method to Capture the Spatial Pattern and to Identify the Causes of Social Vulnerability in Indonesia: A New Recommendation for Disaster Mitigation Policy,” *Procedia Comput. Sci.*, vol. 157, pp. 31–37, 2019, doi: 10.1016/j.procs.2019.08.138.

[5] S. L. Cutter, “Vulnerability to hazards,” *Prog. Hum. Geogr.*, vol. 20, no. 4, pp. 529–539, 1996.

[6] E. Polcarová and J. Pupíková, “Analysis of Socially Vulnerable Communities and Factors Affecting Their Safety and Resilience in Disaster Risk Reduction,” *Sustain.*, vol. 14, no. 18, 2022, doi: 10.3390/su141811380.

[7] V. Cerchiello, P. Ceresa, R. Monteiro, and N. Komendantova, “Assessment of social vulnerability to seismic hazard in Nablus, Palestine,” *Int. J. Disaster Risk Reduct.*, vol. 28, pp. 491–506, 2018, doi: 10.1016/j.ijdrr.2017.12.012.

[8] S. K. Aksha, L. Juran, L. M. Resler, and Y. Zhang, “An Analysis of Social Vulnerability to Natural Hazards in Nepal Using a Modified Social Vulnerability Index,” *Int. J. Disaster Risk Sci.*, vol. 10, no. 1, pp. 103–116, 2019, doi: 10.1007/s13753-018-0192-7.

[9] C. Armenakis, E. X. Du, S. Natesan, R. A. Persad, and Y. Zhang, “Flood risk assessment in urban areas based on spatial analytics and social factors,” *Geosci.*, vol. 7, no. 4, pp. 1–15, 2017, doi: 10.3390/geosciences7040123.

[10] B. M. de Loyola Hummell, S. L. Cutter, and C. T. Emrich, “Social Vulnerability to Natural Hazards in Brazil,” *Int. J. Disaster Risk Sci.*, vol. 7, no. 2, pp. 111–122, 2016, doi: 10.1007/s13753-016-0090-9.

[11] B. E. Flanagan, E. J. Hallisey, E. Adams, and A. Lavery, “Prevention ’ s Social Vulnerability Index,” *J. J Env. Heal.*, vol. 80, no. 10, pp. 34–36, 2020.
[12] D. K. Yoon, “Assessment of social vulnerability to natural disasters: A comparative study,” *Nat. Hazards*, vol. 63, no. 2, pp. 823–843, 2012, doi: 10.1007/s11069-012-0189-2.
[13] P. Krishnan *et al.*, “Framework for mapping the drivers of coastal vulnerability and spatial decision making for climate-change adaptation: A case study from Maharashtra, India,” *Ambio*, vol. 48, no. 2, pp. 192–212, 2019, doi: 10.1007/s13280-018-1061-8.
[14] W. Zhang, X. Xu, and X. Chen, “Social vulnerability assessment of earthquake disaster based on the catastrophe progression method: A Sichuan Province case study,” *Int. J. Disaster Risk Reduct.*, vol. 24, pp. 361–372, 2017, doi: 10.1016/j.ijdrr.2017.06.022.
[15] K. Krellenberg, J. Welz, F. Link, and K. Barth, “Urban vulnerability and the contribution of socio-environmental fragmentation: Theoretical and methodological pathways,” *Prog. Hum. Geogr.*, vol. 41, no. 4, pp. 408–431, 2017, doi: 10.1177/0309132516645959.
[16] D. Hao, D. Shei-Fei, and H. Li-Hua, “Research Progress of Attribute Reduction Based on Rough Sets,” *Comput. Eng. Sci.*, vol. 32, no. 6, 2010.
[17] T. B. Paveglio, C. M. Edgeley, and A. M. Stasiewicz, “Assessing influences on social vulnerability to wildfire using surveys, spatial data and wildfire simulations,” *J. Environ. Manage.*, vol. 213, pp. 425–439, 2018, doi: 10.1016/j.jenvman.2018.02.068.
[18] E. Mavhura and T. Manyangadze, “A comprehensive spatial analysis of social vulnerability to natural hazards in Zimbabwe: Driving factors and policy implications,” *Int. J. Disaster Risk Reduct.*, vol. 56, no. February, p. 102139, 2021, doi: 10.1016/j.ijdrr.2021.102139.
[19] P. Pampouktsi *et al.*, “Techniques of Applied Machine Learning Being Utilized for the Purpose of Selecting and Placing Human Resources within the Public Sector,” *J. Inf. Syst. Explor. Res.*, vol. 1, no. 1, pp. 1–16, 2023.
[20] S. L. Cutter, B. J. Boruff, and W. L. Shirley, “Social vulnerability to environmental hazards,” *Soc. Sci. Q.*, vol. 84, no. 2, pp. 242–261, 2003, doi: 10.1111/1540-6237.8402002.
[21] J. Birkmann, N. J. Setiadi, and N. Baumert, “Socio-economic Vulnerability Assessment at the Local Level in Context of Tsunami Early Warning and Evacuation Planning in the City of Padang, West Sumatra,” in *International Conference on Tsunami Warning (ICTW*, 2008, no. January, pp. 1–8.
[22] T. H. Siagian, P. Purhadi, S. Suhartono, and H. Ritonga, “Social vulnerability to natural hazards in Indonesia: Driving factors and policy implications,” *Nat. Hazards*, vol. 70, no. 2, pp. 1603–1617, 2014, doi: 10.1007/s11069-013-0888-3.
[23] R. Kurniawan *et al.*, “Construction of social vulnerability index in Indonesia using partial least squares structural equation modeling,” *Int. J. Eng. &Technology*, vol. 7, no. 4, pp. 6131–6136, 2018, doi: 10.14419/ijet.v7i4.
[24] A. L. Nugraha, M. Awaluddin, A. Sukmono, and N. Wakhidatus, “Pemetaan Dan Penilaian Kerentanan Bencana Alam Di Kabupaten Jepara Berbasis Sistem Informasi Geografis,” *Geoid*, vol. 17, no. 2, p. 185, 2022, doi: 10.12962/j24423998.v17i2.9370.
[25] Y. T. Wijaya and I. T. Halim, “Measuring and Profiling Social Vulnerability to Natural Disaster in Indonesia in 2019,” *J. Mat. Stat. dan Komputasi*, vol. 19, no. 1, pp. 183–194, Sep. 2022, doi: 10.20956/j.v19i1.21686.
[26] J. Birkmann *et al.*, “Framing vulnerability, risk and societal responses: The MOVE framework,” *Nat. Hazards*, vol. 67, no. 2, pp. 193–211, 2013, doi: 10.1007/s11069-013-0558-5.
[27] B. E. Flanagan, E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis, “A Social Vulnerability Index for Disaster Management,” *J. Homel. Secur. Emerg. Manag.*, vol. 8, no. 1, Jan. 2011, doi: 10.2202/1547-7355.1792.
[28] R. C. Nethery, D. P. Sandler, S. Zhao, L. S. Engel, and R. K. Kwok, “A joint spatial factor analysis model to accommodate data from misaligned areal units with application to Louisiana social vulnerability,” *Biostatistics*, vol. 20, no. 3, pp. 468–484, 2019, doi: 10.1093/biostatistics/kxy016.
[29] B. I. Nasution, R. Kurniawan, T. H. Siagian, and A. Fudholi, “Revisiting social vulnerability analysis in Indonesia: An optimized spatial fuzzy clustering approach,” *Int. J. Disaster Risk Reduct.*, vol. 51, no. May, p. 101801, 2020, doi: 10.1016/j.ijdrr.2020.101801.
[30] X. Guo and N. Kapucu, “Social Vulnerability Evaluation for Ankang City, China, using Fuzzy Analytic Hierarchy Process Method,” *J. Homel. Secur. Emerg. Manag.*, vol. 15, no. 3, Sep. 2018, doi: 10.1515/jhsem-2016-0037.
[31] S. W. M. Weis *et al.*, “Assessing vulnerability: an integrated approach for mapping adaptive capacity, sensitivity, and exposure,” *Clim. Change*, vol. 136, no. 3–4, pp. 615–629, 2016, doi: 10.1007/s10584-016-1642-0.

[32] A. Fekete, "Social Vulnerability (Re-)Assessment in Context to Natural Hazards: Review of the Usefulness of the Spatial Indicator Approach and Investigations of Validation Demands," *Int. J. Disaster Risk Sci.*, vol. 10, no. 2, pp. 220–232, 2019, doi: 10.1007/s13753-019-0213-1.

[33] F. Fatemi, A. Ardalan, B. Aguirre, N. Mansouri, and I. Mohammadfam, "Social vulnerability indicators in disasters: Findings from a systematic review," *Int. J. Disaster Risk Reduct.*, vol. 22, pp. 219–227, Jun. 2017, doi: 10.1016/j.ijdrr.2016.09.006.

[34] L. H. Son, B. C. Cuong, P. L. Lanzi, and N. T. Thong, "A novel intuitionistic fuzzy clustering method for geo-demographic analysis," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9848–9859, 2012, doi: 10.1016/j.eswa.2012.02.167.

[35] G. Grekousis and H. Thomas, "Comparison of two fuzzy algorithms in geodemographic segmentation analysis: The Fuzzy C-Means and Gustafson–Kessel methods," *Appl. Geogr.*, vol. 34, pp. 125–136, May 2012, doi: 10.1016/j.apgeog.2011.11.004.

[36] L. H. Son, "Enhancing clustering quality of geo-demographic analysis using context fuzzy clustering type-2 and particle swarm optimization," *Appl. Soft Comput.*, vol. 22, pp. 566–584, Sep. 2014, doi: 10.1016/j.asoc.2014.04.025.

[37] I. S. Holand, P. Lujala, and J. K. Rod, "Social vulnerability assessment for Norway: A quantitative approach," *Nor. Geogr. Tidsskr.*, vol. 65, no. 1, pp. 1–17, 2011, doi: 10.1080/00291951.2010.550167.

[38] D. Liu and Y. Li, "Social vulnerability of rural households to flood hazards in western mountainous regions of Henan province, China," *Nat. Hazards Earth Syst. Sci.*, vol. 16, no. 5, pp. 1123–1134, May 2016, doi: 10.5194/nhess-16-1123-2016.

[39] X. Guo and N. Kapucu, "Assessing social vulnerability to earthquake disaster using rough analytic hierarchy process method: A case study of Hanzhong City, China," *Saf. Sci.*, vol. 125, no. December 2019, p. 104625, 2020, doi: 10.1016/j.ssci.2020.104625.

[40] C. Guillard-Gonçąlves, S. L. Cutter, C. T. Emrich, and J. L. Zêzere, "Application of Social Vulnerability Index (SoVI) and delineation of natural risk zones in Greater Lisbon, Portugal," *J. Risk Res.*, vol. 18, no. 5, pp. 651–674, 2015, doi: 10.1080/13669877.2014.910689.

[41] W. Chen, S. L. Cutter, C. T. Emrich, and P. Shi, "Measuring social vulnerability to natural hazards in the Yangtze River Delta region, China," *Int. J. Disaster Risk Sci.*, vol. 4, no. 4, pp. 169–181, 2013, doi: 10.1007/s13753-013-0018-6.

[42] O. Drakes, E. Tate, J. Rainey, and S. Brody, "Social vulnerability and short-term disaster assistance in the United States," *Int. J. Disaster Risk Reduct.*, vol. 53, p. 102010, Feb. 2021, doi: 10.1016/j.ijdrr.2020.102010.

[43] A. L. Griego, A. B. Flores, T. W. Collins, and S. E. Grineski, "Social vulnerability, disaster assistance, and recovery: A population-based study of Hurricane Harvey in Greater Houston, Texas," *Int. J. Disaster Risk Reduct.*, vol. 51, no. July, p. 101766, 2020, doi: 10.1016/j.ijdrr.2020.101766.

[44] F. A. Husna, D. Purwitasari, B. A. Sidharta, D. A. Sihombing, A. Fahmi, and M. H. Purnomo, "A Clustering Approach for Mapping Dengue Contingency Plan," *Sci. J. Informatics*, vol. 9, no. 2, pp. 149–160, Nov. 2022, doi: 10.15294/sji.v9i2.36885.

[45] P. J. B. Brown and P. W. J. Batey, "Applications of geodemographic methods in the analysis of health condition incidence data," *Reg. Sci.*, vol. 70, no. 3, pp. 329–344, 1991.

[46] "Implementation of fuzzy tsukamoto in employee performance assessment," *J. Soft Comput. Explor.*, vol. 2, no. 2, Sep. 2021, doi: 10.52465/joscex.v2i2.52.

[47] K. Tyas, A. Ms Ubaidillah, and D. Rahmawati, "The application of the tsukamoto fuzzy method in controlling the dryer for shrimp cracker hygienization," *J. Student Res. Explor.*, vol. 1, no. 2, 2023, doi: https://doi.org/10.52465/josre.v1i2.143.

[48] G. A. Mason and R. D. Jacobson, "Fuzzy Geographically Weighted Clustering," in *Proceedings of the 9th International Conference on Geocomputation*, 2007, no. 1998, pp. 1–7.

[49] S. Sohrabizadeh, S. Tourani, and H. R. Khankeh, "The Gender Analysis Tools Applied in Natural Disasters Management: A Systematic Literature Review," *PLoS Curr.*, 2014, doi: 10.1371/currents.dis.5e98b6ce04a3f5f314a8462f60970aef.

[50] Y. Lixin, Z. Xi, G. Lingling, and Z. Dong, "Analysis of social vulnerability to hazards in China," *Environ. Earth Sci.*, vol. 71, no. 7, pp. 3109–3117, Apr. 2014, doi: 10.1007/s12665-013-2689-0.

[51] I. Armas and A. Gavris, "Social vulnerability assessment using spatial multi-criteria analysis (SEVI model) and the Social Vulnerability Index (SoVI model) - A case study for Bucharest, Romania," *Nat. Hazards Earth Syst. Sci.*, vol. 13, no. 6, pp. 1481–1499, 2013, doi: 10.5194/nhess-13-1481-2013.