Jurnal Teknologi Sistem Informasi dan Aplikasi ISSN: 2654-3788
Penerbit: Program Studi Teknik Informatika Universitas Pamulang e-ISSN: 2654-4229
Vol. 8, No. 2, April 2025 (138-149) DOI: https://doi.org/10.32493/jtsi.v8i2.54092

A Predictive Model for Urban Agglomeration Development in East Kalimantan Province Using Cluster Analysis

Anita Claudia Novedy[1*], Ibnu Sasongko[2], Endratno Budi Santosa[3], Firman Afrianto[4]

[123]Program Studi Perencanaan Wilayah dan Kota, Institut Teknologi Nasional Malang, Malang, Indonesia, 65145
[4]Ikatan Ahli Perencanaan Indonesia Provinsi Jawa Timur,Malang, Indonesia, 65142
e-mail: [1*]2224007@scholar.itn.ac.id, [2]ibnukoko59@gmail.com, [3]endratnobudi@gmail.com, [4]firmanafrianto@mail.ugm.ac.id
[*]Corresponding Author

Submitted Date: January 18, 2025 Reviewed: Febuary 09, 2025
Revised Date: March 10, 2025 Accepted Date: March 20, 2025

Abstract

The development of the new capital in East Kalimantan Province is guided by agglomeration as a strategic framework, aiming to reduce regional disparities and maintain a balance between the central area and its surrounding regions. This strategy is reinforced by the development of the transportation sector, which provides the foundation for both geographic and administrative connectivity across regions. The concentration of activities, including industry, trade, government, and infrastructure, is a key driver of economic growth and regional development. This study employs a quantitative approach supported by Geographic Information Systems (GIS), specifically the DBSCAN algorithm, using a road network shapefile (.shp) as the primary data. The results of the analysis show that Bontang, Balikpapan, and Samarinda City function as the trigger nodes driving regional activity concentration and growth, with development trends expanding towards the north and west. Predictive analysis using DBSCAN at distances of 1,600–2,400 meters indicates that 9 out of 10 regions in East Kalimantan Province have formed new growth centers, except that Mahakam Ulu Regency has not yet shown signs of urban agglomeration.

Keywords: Urban Agglomeration; DBSCAN Algorithm; Geographic Information Systems; Urban Growth

1 Introduction

As time passes, urban development will continue to evolve in step with the dynamic nature of human activities. The development of urban areas is shaped by three fundamental factors that determine the quality and quantity of cities: appearance, mass, and space. These elements function as products; time serves as the process; and cities serve as platforms for human activities (Larasati et al., 2022). Currently, urban development has attracted a significant portion of the rural population, who are moving to cities in search of greater opportunities to improve their welfare. The United Nations Department of Economic and Social Affairs (Kurniati et al., 2022) further supported this phenomenon by reporting that Indonesia had one of the highest urbanization rates in the world, reaching 56.64% in 2020. While urbanization leads to population growth exceeding the availability of vacant land, it also stimulates household consumption, investment growth, government spending, and financial benefits for the community (Aini, 2022). Rijal & Tahir (2022) suggested that urbanization acts as a catalyst for economic transformation by influencing human and natural resources, technology, income, and social dynamics. However, various challenges arise that cannot be fully addressed by urban policies alone. In response, the government continues its efforts to bridge the infrastructure gap between regions, because this disparity hinders regional development grounded in economic growth. The efforts to develop the region are consistent with the consolidation of activities, population, and connectivity among regions, resulting in the formation of urban areas characterized by integrated activity systems. This phenomenon is referred to as urban agglomeration.

The formation of the urban agglomeration brought several spatial challenges to the affected areas, including water and air pollution, environmental degradation, increased traffic congestion, and changes in land use to accommodate residential areas (Surya et al., 2021). Despite these issues, urban agglomeration can also bring multiple benefits. These benefits include driving economic growth, increasing human productivity, facilitating infrastructure development, and encouraging development to spread from central growth areas to surrounding regions. This process is known as the "spread effect".

Designating East Kalimantan Province as the location for Indonesia's new national capital is a significant milestone that provides a foundation for equitable, multidimensional growth—economic, social, and environmental—in a sustainable manner. The development of the capital city, located in the North Penajam Paser and Kutai Kartanegara Regencies, is expected to have far-reaching effects, not only in the core area but also in other parts of East Kalimantan. In the short term, infrastructure investments are expected to boost the economy by facilitating increased trade between Indonesian regions and supporting the growth of various sectors. This will create new employment opportunities. In the medium to long term, the capital's development is expected to drive sustained economic growth, encourage the growth of nontraditional sectors, and promote economic diversification across Kalimantan. This growth will likely enhance regional trade and help to reduce income inequality. The development will be supported by the strategic location of the well-developed cities of Balikpapan and Samarinda, which offer high accessibility and essential infrastructure, such as airports and ports, due to their location along the ALKI II route (Makassar Strait). This immense spatial potential is expected to trigger a rapid wave of urbanization, ultimately transforming living patterns, land use, urban management, and residential concentration throughout the region (Badan Perencanaan Pembangunan Nasional, 2021).

Alexander (2024) emphasized that agglomeration is a critical factor in the development of the new state capital in East Kalimantan Province. The goal is to reduce disparities and strike a balance between the central area and its surrounding regions. This balance will be achieved by developing the transportation sector to establish connectivity across regions, both geographically and administratively. Agglomeration plays a critical role in urban growth, especially in East Kalimantan Province, where significant transformation is underway. The concentration of activities, including industry, trade, government, and infrastructure, is a key driver of economic growth and regional development. Fostering such centralization is expected to the region and its surrounding areas by promoting a broader distribution of facilities and utilities.

For these reasons, it is crucial to conduct research examining how agglomeration develops across various regions of East Kalimantan, accounting for their respective sizes. This research could serve as a valuable reference for policymakers to guide the planning, implementation, and evaluation of development efforts and ensure fair and equitable spatial development throughout the province. A similar study was conducted by Kalpana et al. (2021), who discussed using road network data as a new approach to predicting urban agglomeration patterns in the Sri Lanka region. Similarly, Afrianto (2023) researched using road network data to indicate urban agglomeration expansion in the Greater Malang area. However, the existing research on predicting urban agglomeration development is considered limited because it relies on existing road network data, which does not capture future road networks and is thus not suitable for long-term projections. Therefore, the prediction results may be slightly inaccurate if future city development depends on changes to the road network.

Therefore, it is essential to conduct this research to predict the direction of urban agglomeration development in East Kalimantan using planned road network data as a new approach. Based on this research objective, the following research question will be answered through a series of analyses: Does the extent of road network development contribute to regional growth?

2 Methodology

The methodology of this research is explained in the sub-sections below.

Jurnal Teknologi Sistem Informasi dan Aplikasi **ISSN: 2654-3788**
Penerbit: Program Studi Teknik Informatika Universitas Pamulang **e-ISSN: 2654-4229**
Vol. 8, No. 2, April 2025 (138-149) **DOI: https://doi.org/10.32493/jtsi.v8i2.54092**

A. Research Location

This research is located in East Kalimantan Province, which is divided administratively into seven districts and three cities, with a total area of 127.346,92 km^2

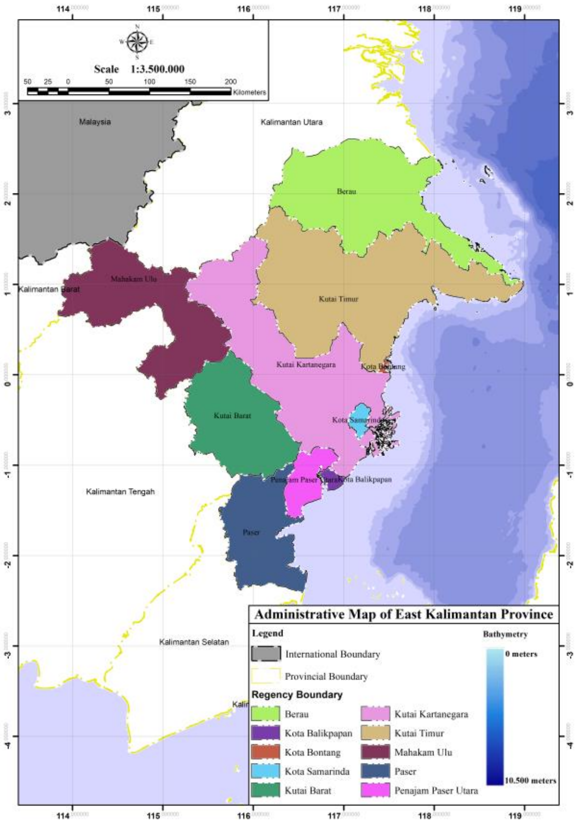


Figure 1. Map of Administrative Boundaries of Research Locations
Source: Processed by Author, 2024

Based on data from the Population and Civil Registration Service, East Kalimantan ranks third among Indonesia's provinces, after Central Kalimantan and West Kalimantan. As of 2024, the province is recorded as having a population of 3.850.679 people. The territorial boundaries of East Kalimantan are as follows:

North	:	North Kalimantan Province
South	:	South Kalimantan Province
East	:	Makassar Strait and Celebes Sea
West	:	West Kalimantan Province, Central Kalimantan province, and and state of Sarawak, Malaysia

B. Data Collection Method

This research uses secondary data, including administrative boundary data of the study area, existing road network, and planned road network data. All of this data was obtained from various institutions or spatial data providers in shapefile (.shp) format. The shapefile (.shp) road network is the primary dataset used to analyze point density and detect activity concentrations, and to delineate urban agglomeration clusters. The data used can be seen in table 1 below.

Table 1. Main Data

Data Source	Data Type
Indonesian Geospatial Information Agency downloaded via https://tanahair.indonesia.go.id/portal-web/	Administrative boundaries of East Kalimantan Province as a research location
	Symbology of Provincial and Regency Boundaries in the East Kalimantan Region
Open Street Map downloaded via https://download.geofabrik.de/	Existing Road Network
Spatial Planning data downloaded via https://gistaru.atrbpn.go.id/	Planned Road Network

Source: Data Collected, 2024

C. Data Processing Method

This research uses cluster analysis, which groups similar data into homogeneous clusters. The goal is to identify data sets with similar characteristics and group them into a single cluster (Birant & Kut, 2007). When processing the data, researchers use Geographic Information System (GIS) software, such as Quantum GIS, as tools. GIS software has long been used to collect, create, and distribute spatial data, and it is designed to make map creation easier. Technically, processing .shp road network data involves converting lines to intersection points and splitting multipart features into single-part features. Other preprocessing tasks are performed in ArcGIS because it efficiently manages spatial data. ArcGIS is preferred because it offers features in the toolbox. However, QGIS is selected for the clustering stage using the DBSCAN algorithm because its plugin features are more comprehensive and easier to access. Users can install the plugin for free. It is flexible, supports Python integration for advanced analysis, and allows for easy visualization of clustering results.

Spatial analysis in this study was conducted in stages to obtain a comprehensive understanding of point distribution patterns. First, Nearest Neighbor Analysis (NNA) was performed to evaluate the global pattern of point distribution, whether it was clustered, random, or dispersed. The NNA results provided an initial framework for the concentration and distribution of points, as well as guidance on determining appropriate parameters for further cluster analysis. Next, DBSCAN (Density-Based Spatial Clustering of Applications with Noise) was performed to identify the locations of specific clusters and points that did not belong to clusters (noise), thereby enabling a more detailed interpretation of spatial dynamics. This research was carried out through several stages of analysis, as follows:

a. Nearest Neighbor Analysis

Nearest neighbor analysis is a method used to determine the distances to the closest neighbors in a random pattern (Riadhi et al., 2020). This analysis then explains the distribution pattern of location points by considering distance, the number of location points, and area. The equations used in this analysis are as follows:

$$ANN = \frac{\bar{D}_O}{\bar{D}_E} \quad (1)$$

Where:

- $\bar{D}_O$ is the observed mean distance between each feature and its nearest neighbor.

$$\bar{D}_O = \frac{\sum_{i=1}^{n} d_i}{n} \quad (2)$$

- $\bar{D}_E$ is the expected mean distance for the features given in a random pattern.

$$\bar{D}_E = \frac{1}{2\sqrt{p}} \quad (3)$$

- Point density value or point in square kilometers (p) can then be obtained using the formula:

$$p = \frac{n}{A}$$

Where:

- n = number of points
- A = size of study area

Based on these equation, Clark & Evans (1954) put forward three classification on the Nearest Neighbor Index (NNI). If NNI <1 is significantly, the distribution is clustered, if NNI = 1, the results distribution is random, and if NNI > 1 significantly, the distribution is uniform (Kharitonov, 2015).

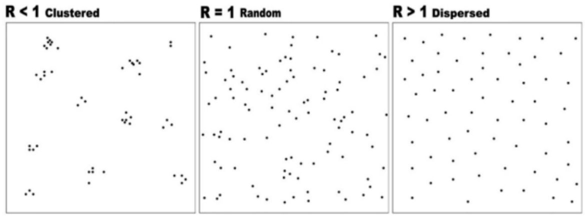


Figure 2. Classification of Nearest Neighbor Analysis
Source: Barret et al., 2017

This research uses primary data in the form of a combined road network, including existing and planned conditions, in shapefile (.shp) format. Vector data, represented as lines of road network, is then converted into points using the GIS software's data management tools. The new point-based data serves as the foundation for an average nearest-neighbor analysis using spatial statistics tools. In ArcGIS, this analysis yields key information, including the type of point distribution, the observed and expected average distance, the z-score, and the p-value. Figure 2 presents the road network map of East Kalimantan Province, which served as the material for this analysis.

b. Cluster Analysis

Cluster analysis can be utilized to group objects and cases into smaller, more manageable groups, where each group contains similar objects. This method uses a distance measure to explain the similarity or closeness between complex data points, simplifying the data into distinct clusters. One commonly used distance measure for this purpose is the Euclidean distance. Several methods are employed in spatial clustering algorithms, including K-means, K-medoids, and DBSCAN.

The DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm was selected for this research due to its superior ability to identify clusters of arbitrary shapes and its performance in large databases. In fact, it outperforms the K-medoids algorithm by more than 100 times in terms of efficiency (Ester et al., 1996). In this context, density is defined as the number of minimal data points (minPts) that exist within a defined epsilon (ε) radius around each data point (Budiman et al., 2016). The parameter minPts specifies the minimum number of points required to form a cluster. At the same time, epsilon signifies the maximum distance between two points for them to be considered part of the same cluster.

Subsequently, the collected data is processed in QGIS using various features. One such feature, the ability to transform a multipart to a single part on vector geometry, facilitates the segmentation of geometric features, such as road intersections, into discrete components. Furthermore, the net. The centrality feature is employed to measure the influence of a node within the road network. Finally, the clustering process is executed using the DBSCAN algorithm to identify patterns based on data density. Figure 3 illustrates the flowchart of the research methodology.

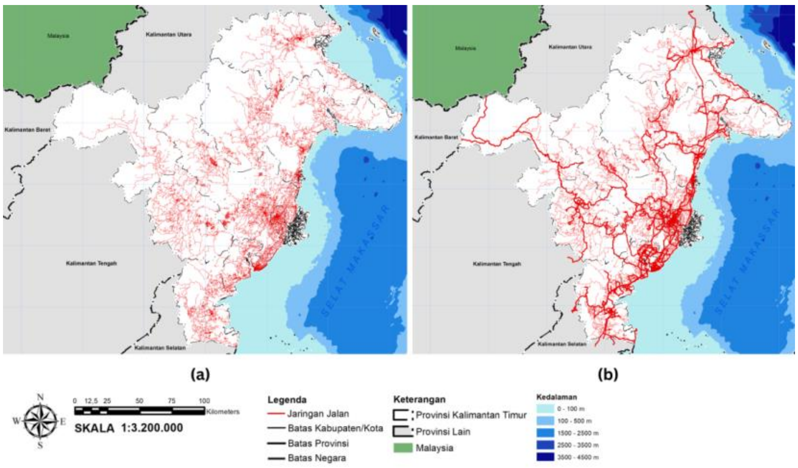


Figure 3. Existing Roads Network (a) Planned Roads Network (b)
Source: Processed by Author, 2024

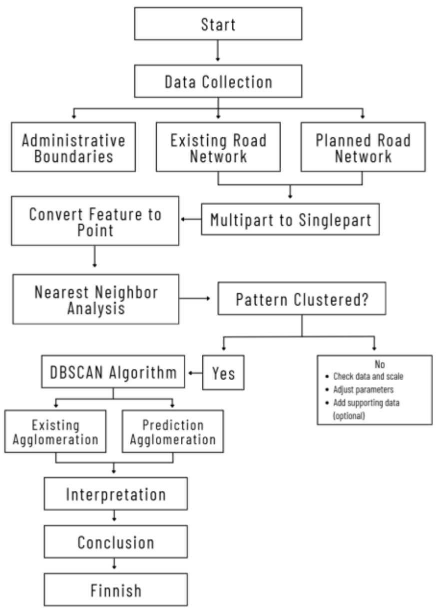


Figure 4. Research Flow Diagram

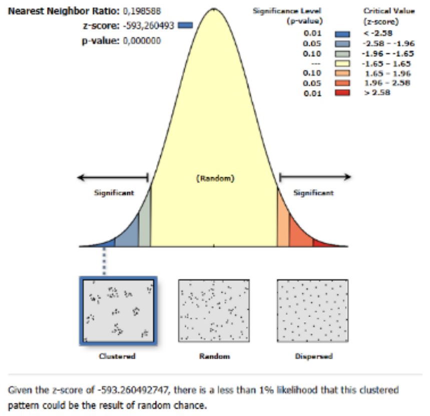
Jurnal Teknologi Sistem Informasi dan Aplikasi **ISSN: 2654-3788**
Penerbit: Program Studi Teknik Informatika Universitas Pamulang **e-ISSN: 2654-4229**
Vol. 8, No. 2, April 2025 (138-149) **DOI: https://doi.org/10.32493/jtsi.v8i2.54092**

3 Results and Discussions

A. Average Nearest Neighbor Analysis

The calculations using nearest-neighbor analysis yield an index indicating a clear clustering pattern, with an average distance between points of approximately 126 meters. This clustering is confirmed by the Nearest Neighbor Ratio, which is 0.198588, a value less than 1. This finding indicates that the distribution of data points is predominantly concentrated in specific regions rather than dispersed randomly across the area. The observed clustering suggests a high concentration of activities or features within relatively proximity, a phenomenon characteristic of agglomeration in urban areas. In addition, the p-value derived from the analysis is less than 0,05, with a p-value of 0,00 and a z-score of -593,260. These findings indicate that the ANN results demonstrate statistically significant clustering and strongly reject the random hypothesis (CSR) at the 99% confidence level. The study provides statistical evidence that the observed clustering is not attributable to random chance (ESRI, n.d.). Consequently, the null hypothesis (Ho), which posits a random distribution, can be rejected, and the alternative hypothesis (Ha), which states that the pattern of data points is statistically significant and indicative of real clustering behavior within the study area, can be accepted.

This analysis serves as the preliminary phase in understanding the region's spatial dynamics, thereby laying the foundation for subsequent cluster analysis. The observed mean distance of 126 meters serves as a critical foundation for subsequent analysis. This analysis will be conducted at progressively larger distances to map out how agglomerations form and expand between different regions. The approach is staged, meaning that the cluster analysis will be repeated at increasing distance thresholds to track the evolution of agglomerations. This approach facilitates a more nuanced understanding of the growth, interaction, and merging of clusters of activity over time across the study area. By implementing this systematic approach, the research endeavors will identify regions where agglomeration is more pronounced and where the emergence of novel clusters may be anticipated. This progressive analysis is essential for understanding the dynamic nature of urban expansion. Urban planners and policymakers must use this analysis to inform decisions on infrastructure development and resource allocation. Figure 4 presents the results of the nearest-neighbor analysis, providing a visual representation of the initial clustering pattern. This pattern will be further explored in the subsequent stages of the research.



Average Nearest Neighbor Summary	
Observed Mean Distance:	126,1066 Meters
Expected Mean Distance:	635,0161 Meters
Nearest Neighbor Ratio:	0,198588
z-score:	-593,260493
p-value:	0,000000

Figure 5. NNA Result
Source: Processed by Author, 2024

B. DBSCAN Analysis

The application of the DBSCAN algorithm in predicting the direction of urban development has also been carried out by Afrianto (2023) to identify patterns and directions of development in Greater Malang, which includes Malang City, Batu City, and Malang Regency, through road network intersections. This research served as a foundational reference point for subsequent research aimed at elucidating the patterns and expansion of development across a broader geographical scope, specifically in East Kalimantan Province.

In this case, the DBSCAN algorithm requires two main parameters: epsilon (ε) and minimum points (minPts). Based on the previous nearest-neighbor analysis, the epsilon value was determined to be 100 meters (rounded from the observed mean distance) and 500 meters (rounded from the expected mean distance) for the next range. Furthermore, the minPts value was determined based on the assumption that a settlement could be

Jurnal Teknologi Sistem Informasi dan Aplikasi **ISSN: 2654-3788**
Penerbit: Program Studi Teknik Informatika Universitas Pamulang **e-ISSN: 2654-4229**
Vol. 8, No. 2, April 2025 (138-149) **DOI: https://doi.org/10.32493/jtsi.v8i2.54092**

agglomerated at the research location if it was formed from at least five road intersections.

a. Existing Agglomerations

The analysis results visualization in Figure 5 indicates that agglomeration is concentrated within approximately 100 meters of the Samarinda City area, with the highest cluster intensity compared to other districts and cities in East Kalimantan Province. These findings indicate that Samarinda serves as the region's preeminent urban hub. Furthermore, Balikpapan City and Bontang City follow closely behind Samarinda, forming agglomerations with similarly high intensity. The results of the analysis show that Bontang, Balikpapan, and Samarinda City function as the trigger nodes driving regional activity concentration and growth, with development trends expanding towards the north and west.

As development progresses, other regions are beginning to show early signs of agglomeration, though at lower intensity. The southern part of East Kutai Regency, Berau Regency, and the eastern side of Kutai Kartanegara Regency (near Samarinda City), along with Paser Regency, are beginning to exhibit signs of emerging agglomeration, as indicated by areas marked in pink, representing low to medium cluster intensity. These regions are in the nascent stages of urban cluster formation, suggesting the likelihood of imminent growth and activity concentration.

Conversely, West Kutai Regency and Mahakam Ulu Regency have not shown any signs of agglomeration within 100 meters. These regions are largely undeveloped in terms of urban concentration, highlighting the uneven distribution of agglomeration across the province. The lack of cluster formation in these regions suggests that there are significant barriers to development, possibly due to inadequate infrastructure or other limiting factors, which will require further investigation.

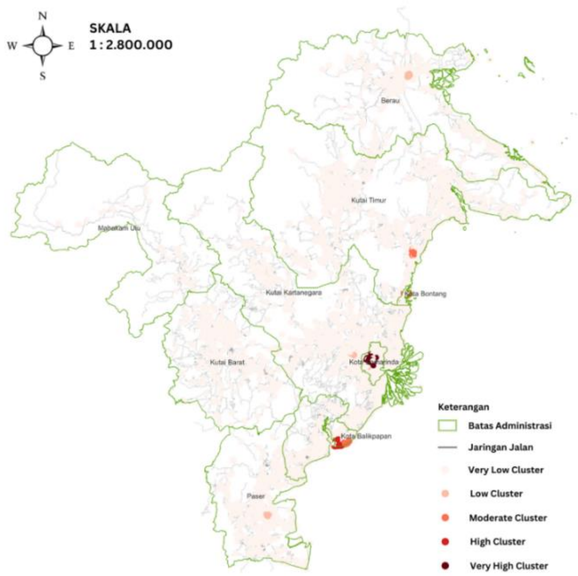


Figure 6. Agglomeration Formation with a 100 meters Distance
Source: Processed by Author, 2024

b. Agglomerations Development Process

The analysis of agglomeration development across various distances reveals distinct patterns and visualizations of growth. As shown in Figure 6, it is evident that the initial agglomeration exhibits a strong tendency to expand further north and west. The point of agglomeration, which was initially centered in Samarinda City with very high intensity, has gradually spread northward and westward into Kutai Kartanegara Regency. A similar pattern of agglomeration expansion has been observed in Balikpapan City, with signs of extending toward the southern side of Kutai Kartanegara Regency and Samarinda City. As the distance increases from 200 meters to 800 meters, a large cluster is predicted to form, linking these three key areas.

The agglomeration in Balikpapan City shows significant potential for further development, as indicated by the vigorous intensity of activity shown in Figure 6. This suggests that Balikpapan City is poised for substantial growth, closely following the trends seen in Samarinda City and Kutai Kartanegara Regency, particularly in Samboja District, which has been designated as one of the strategic locations for Indonesia's new capital, Nusantara.

As the distance between cluster points increases, there is a direct correlation with further

ISSN: 2654-3788
e-ISSN: 2654-4229
DOI: https://doi.org/10.32493/jtsi.v8i2.54092

development in other areas of East Kalimantan. New growth centers are emerging in regions such as Berau Regency, the western and southern parts of East Kutai Regency, West Kutai Regency, Paser Regency, and North Penajam Paser Regency, each with varying levels of agglomeration. This finding underscores a broader pattern of regional development, with significant progress observed as the maximum distance between cluster points increases.

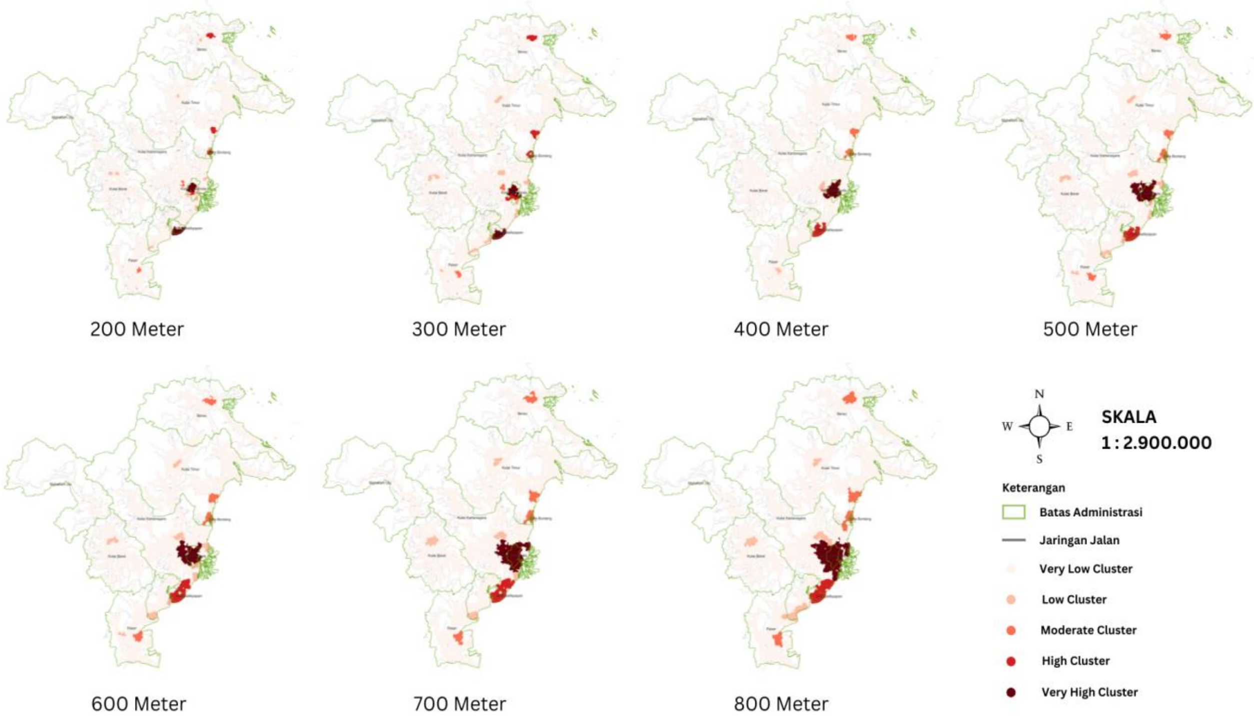


Figure 7. Visualization of the Development of Agglomeration until Large Clusters are Formed
Source: Processed by Author, 2024

However, in contrast to these expanding regions, Mahakam Ulu Regency stands out as the only area that has not yet shown any signs of growth or the emergence of new centers of activity. Despite the expansion observed in most districts and cities across East Kalimantan, the majority of Mahakam Ulu Regency remains underdeveloped. This indicates a lack of agglomeration in this region, highlighting the need for further exploration into the factors that may be inhibiting development there.

c. Agglomerations Prediction

Based on observations of agglomerations that have already extended up to 800 meters, the next logical step is to predict the pattern and direction of future development of activity concentrations across various districts and cities within East Kalimantan Province. The analysis focuses on potential developments at distances ranging from 1,600 to 2,400 meters. As shown in Figure 7, if agglomeration continues to expand within this range, rapid growth is anticipated, primarily towards the north and south. Conversely, the western region is expected to undergo a more gradual development. Although steadily, reflecting a slower but sure expansion in that direction. This extended growth is likely to result in the formation of clusters that focus and interact across regions, which in turn will stimulate the creation of new growth centers. These new centers have the potential to evolve into additional clusters, thereby reinforcing the overall agglomeration process. According to the visual analysis depicted in Figure 7, when the agglomeration distance is extended to 1,600 meters, development spreads significantly northward, particularly towards East Kutai Regency, where the cluster intensity is very high. This concentration of activity suggests that East Kutai Regency is likely to become a key area of growth as the agglomeration continues to expand in the coming years.

Concurrently, in the southern expansion, nearly all regions in North Penajam Paser Regency and Paser Regency exhibit signs of forming a substantial cluster that exhibits a propensity to also evolve towards the north, traversing the Sepaku District as a prospective location for the National Capital, so that it becomes one cluster with Balikpapan City, Samarinda City, Kutai Kartanegara Regency, Bontang City, as well as several areas in East Kutai district. Subsequently, the distance was once again expanded to 2,400 meters as shown in Figure 7b. It is predicted that these clusters will continue to grow, accompanied by rapid growth in new centers, such as Berau Regency, the central southern part of East Kutai Regency, Kutai Kartanegara Regency, and West Kutai Regency. Similarly, West Kutai Regency has experienced rapid development in North Penajam Paser Regency and Paser Regency compared to the previous distance, resulting in a very high cluster intensity.

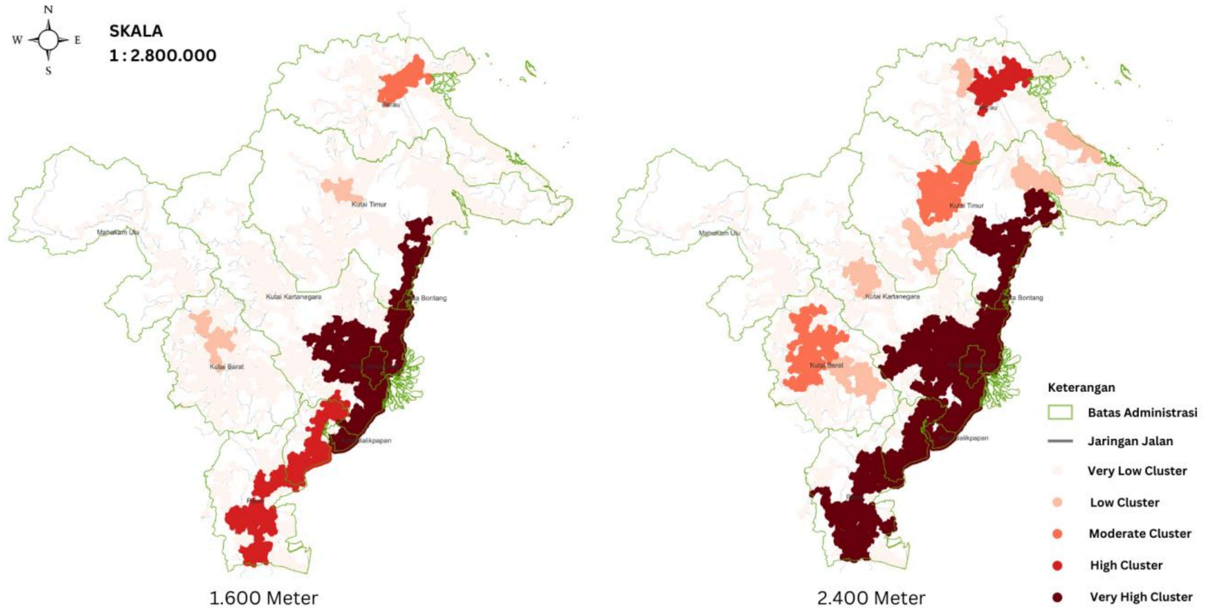


Figure 8. Prediction of Urban Agglomeration at a Distance of 1.600 Meters (a) and Prediction of Urban Agglomeration at a Distance of 2.400 Meters (b)
Source: Processed by Author, 2024

Concurrently, in the southern expansion, nearly all regions in North Penajam Paser Regency and Paser Regency exhibit signs of forming a substantial cluster that exhibits a propensity to also evolve towards the north, traversing the Sepaku District as a prospective location for the National Capital, so that it becomes one cluster with Balikpapan City, Samarinda City, Kutai Kartanegara Regency, Bontang City, as well as several areas in East Kutai district. Then, the distance was once again expanded to 2,400 meters as shown in Figure 7b above. It can be seen that these clusters are predicted to develop further, accompanied by the rapid growth of new centers, such as Berau Regency, the central side to the south of East Kutai Regency, and Kutai Kartanegara Regency. Likewise, West Kutai Regency is experiencing rapid development compared to the previous distance, to North Penajam Paser Regency and Paser Regency, which have very high cluster intensity. However, the predicament lies in the observation that Mahakam Ulu Regency, which has reached this analysis, has not shown the gradual development observed in the other regions. It has been observed that, before the expansion, all areas within the Mahakam Ulu Regency did not experience the effects of the spread, which extended up to 2,400 meters. This observation underscores the significance of conducting a thorough discussion and investigation to gain a comprehensive understanding of the phenomenon.

C. Discussion

The prediction results regarding the agglomeration areas in Balikpapan City, Bontang City, and Samarinda City, as growth centers, are expected to have a beneficial impact on the surrounding areas. The spatial spread observed in other regions has significant potential to contribute to the region's development. The research results show that the emergence of growth centers in different areas is increasingly expanding, in line with the spread effect theory, which states that regional development in growth centers will stimulate economic growth, which will, in turn, be followed by the development of the surrounding regions (Qibti & Hendarto, 2020). This phenomenon certainly provides an understanding that is in line with the theory of Cumulative Causation proposed by Myrdal (1957) that not all regions experience uniform growth, but rather tend to be concentrated in certain areas that have services and advantages, so that more developed regions will attract goods, people, and infrastructure development, while less developed areas tend to stagnate or even regress (Muta'ali, 2025). As shown in Figure 6, the development process of new growth centers is evident. The expansion of these growth centers, which form a cluster in nearly all regencies in East Kalimantan Province, has been proven.

However, a notable contrast to this research's results is that Mahakam Ulu Regency has not yet demonstrated any development in the predicted distance area. In contrast, other regions have shown potential growth and development. The investigation identified several fundamental factors contributing to the comparatively slower urban agglomeration growth in Mahakam Ulu Regency compared to surrounding areas. Despite its geostrategic significance as a gateway between Indonesia and Malaysia via Sarawak, the region has experienced a slowdown in its urban agglomeration growth. Infrastructure is the paramount factor in Mahakam Ulu Regency, an area widely regarded as lagging behind the rest of East Kalimantan Province. This disparity is evident in the absence of adequate electricity networks, telecommunications infrastructure, access to clean water, and efficient road networks, collectively indicating a state of considerable isolation in this region. The Mahakam Ulu Regency is accessible by land; however, the substandard condition of the roads poses significant challenges for the community, including increased susceptibility to damage and road flooding during periods of heavy rainfall. In addition to terrestrial access, the Mahakam Ulu Regency can be reached by waterway; however, this method is associated with considerable expense and the possibility of substantial riverine flows during precipitation events (Lavenia, 2024). This research is of critical importance for further exploration from various scientific perspectives, as it can provide a strong foundation for decision-making, planning, implementation, and evaluation of development efforts across the East Kalimantan Province.

4 Conclusion

The study demonstrates that urban agglomeration, initially concentrated around Balikpapan, Bontang, and Samarinda, has been a major driver of growth in East Kalimantan Province, encouraging the development of potential new activity centers. This is particularly relevant in the context of Indonesia's decision to establish its new capital, Nusantara, in East Kalimantan. The development of Nusantara, located in North Penajam Paser and Kutai Kartanegara Regencies, marks a significant milestone in Indonesia's goal of achieving equitable, multifaceted growth across the region. As the study highlights, the existing urban agglomeration patterns provide spatial momentum that aligns with the government's broader vision for Nusantara. The expansion of these agglomerations indicates a pronounced pattern, with notable growth in the northern and southern regions, while the western areas show a comparatively slower rate of development. It has been demonstrated that nine out of ten regions in the province have exhibited elevated levels of cluster intensity. Conversely, the Mahakam Ulu Regency has remained largely unaffected, a phenomenon primarily attributable to the region's underdeveloped infrastructure.

However, the study has certain limitations. It relies heavily on road network data, overlooking other essential infrastructure such as other transportation networks, energy, telecommunications, and water access, which are crucial too for comprehensive regional development. Furthermore, the study does not account for socio-economic factors such as population migration or local mobility patterns, both of which are essential to the planning of a new capital city like Nusantara. Using only the

DBSCAN clustering method may also limit the analysis, as integrating other clustering techniques could yield a more robust understanding of urban development.

Future research should address these gaps by incorporating a broader range of infrastructure data and socio-economic variables better to reflect the complex dynamics of Nusantara's development.

5 Future Work

The application of the DBSCAN algorithm as a clustering method to predict the development of urban spatial agglomeration in East Kalimantan Province provides a preliminary insight into the distribution patterns and growth directions of metropolitan areas. Accordingly, future research is encouraged to focus on validating these predictive results using the Cellular Automata–Markov Chain approach, a spatial modeling technique capable of simulating temporal land-use change dynamics. This approach is expected to assess the consistency and accuracy of the predictions, while also strengthening the foundation for spatial planning and policymaking in the region.

References

Afrianto, F. (2023). Fractal Dimensions Analysis of Urban Agglomeration at Road Intersections in Metropolitan Malang Raya. *IOP Conference Series: Earth and Environmental Science*, *1186*(1). https://doi.org/10.1088/1755-1315/1186/1/012010

Aini, A. F. (2022). Analisis Analisis Dampak Urbanisasi Terhadap Pertumbuhan Ekonomi Kota Surabaya. *Journal Economics and Strategy*, *3*(2), 60–67. https://doi.org/10.36490/jes.v3i2.425

Alexander, H. B. (2024). *Transportasi Berkelanjutan IKN Diwacanakan Tiru Aglomerasi Jadebotabek*. Www.Kompas.Com.

Badan Perencanaan Pembangunan Nasional. (2021). Handbook of the Relocation of Indonesia’s Capital City. *Kementrian PPN/Bappenas*, 1–29.

Birant, D., & Kut, A. (2007). ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data and Knowledge Engineering*, *60*(1), 208–221. https://doi.org/10.1016/j.datak.2006.01.013

Budiman, S. A. D., Safitri, D., & Ispriyanti, D. (2016). Perbandingan Metode K-Means Dan Metode Dbscan Pada Pengelompokan Rumah Kost Mahasiswa Di Kelurahan Tembalang Semarang. *Jurnal Gaussian*, *5*, 757–762. http://ejournal-s1.undip.ac.id/index.php/gaussian

Clark, P. J., & Evans, F. C. (1954). Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology*, *35*(4), 445–453. https://doi.org/10.2307/1931034

ESRI. (n.d.). *What is a z-score? What is a p-value?* Retrieved November 25, 2025, from https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/what-is-a-z-score-what-is-a-p-value.htm

Ester, M., Kriegel, H., Xu, X., & Miinchen, D.-. (1996). *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*.

Kalpana, L. D. C. H. N., Abenayake, C., Jayasinghe, A., Mahanama, P. K. S., & Sanjaya, N. (2021). A novel approach to measure the pattern of urban agglomeration based on the road network. *International Journal of Sustainable Development and Planning*, *16*(2), 251–262. https://doi.org/10.18280/IJSDP.160205

Kharitonov, S. (2015). *Nearest-neighbor distance method for the evaluation of distribution of different biological objects in plane or in line*. https://doi.org/10.13140/RG.2.1.1733.0643

Kurniati, S. A., Rahayu, P., & Istanabi, T. (2022). Peri-Urbanisasi Dan Dinamika Perkembangan Kawasan Perkotaan Sekunder (Studi Kasus: Bosukawonosraten). *Desa-Kota*, *4*(2), 167. https://doi.org/10.20961/desa-kota.v4i2.55247.167-180

Larasati, A. P., Rahman, B., & Kautsary, J. (2022). Pengaruh Perkembangan Perkotaan Terhadap Fenomena Pulau Panas (Urban Heat Island). *Jurnal Kajian Ruang*, *2*(1), 35. https://doi.org/10.30659/jkr.v2i1.20469

Lavenia, R. (2024). *10 Tahun Jadi Daerah Otonom Baru, Pengamat Unmul Menilai Mahakam Ulu Kaltim Masih Dianaktirikan*. Tribun Kaltim. https://kaltim.tribunnews.com/2024/07/17/10-tahun-jadi-daerah-otonom-baru-pengamat-unmul-menilai-mahakam-ulu-kaltim-masih-dianaktirikan

Muta’ali, L. (2025). *Teori dan Konsep Economic Geography*.

Qibti, M. H. M., & Hendarto, R. M. (2020). Analisis Spillover Effect Pertumbuhan

Ekonomi antar Kabupaten / Kota Di Kawasan. *Diponegoro Journal Of Economics*, *9*(4).

Riadhi, A. R., Aidid, M. K., & Ahmar, A. S. (2020). Analisis Penyebaran Hunian dengan Menggunakan Metode Nearest Neighbor Analysis. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *2*(1), 46. https://doi.org/10.35580/variansiunm12901

Rijal, S., & Tahir, T. (2022). Analisis Faktor Pendorong Terjadinya Urbanisasi di Wilayah Perkotaan (Studi Kasus Wilayah Kota Makassar). *Journal of Economic Education and Entrepreneurship Studies*, *3*(1), 262–276.

Surya, B., Salim, A., Hernita, H., Suriani, S., Menne, F., & Rasyidi, E. S. (2021). Land Use Change, Urban Agglomeration, and Urban Sprawl: A Sustainable Development Perspective of Makassar City, Indonesia. In *Land* (Vol. 10, Issue 6, p. 556). https://doi.org/10.3390/land10060556