Spatial patterns of hotel’s occupation rate

A Ghozali* and E D Syafitri
Department of Urban and Regional Planning, Institut Teknologi Kalimantan,
Balikpapan, 76127, Indonesia

*ghozali@itk.ac.id

Abstract. In Balikpapan the growth in the number of hotels increased significantly over the past 10 years but the income earned from this sector has tended to decline. Currently, it is found that hotel development growth becomes imbalance due to overcapacity and only exist in several regions. This study aims to identify hotel distribution patterns and hotel location clustering analysis based on the occupancy rate conducted in two stages using the Average Nearest Neighbor (ANN) and hot spot analysis method. In overview, hotels in Balikpapan, both star hotel and non-star hotel, significantly clustered spatially around the centers of urban activities. This grouping is not linear with the occupancy rate distribution pattern. The occupancy rate pattern of star hotels shows a spatially clustered structure with z-score to 2.69 (most significant). Conversely, the occupancy rate of non-star hotels has a random pattern with a z-score is closest to 0 (not significant). However, in general, the constructed hotels in an area which has high accessibility and closes to urban centers is the hot spot point for significant high occupancy rates.

1. Introduction
The growth of Balikpapan City as a destination city by immigrants, as well as being a center of business, tourism, and industry, primarily related to oil and gas activities, in East Kalimantan, is linearly related to the rapid growth of the accommodation sector, especially in the number of hotels. However, the main challenge arises when growth in the number of hotels is not controlled and concentrated in several regions [1]. The growth of hotels over the past decade shows a significant increase in both star and non-star hotels. In 2010, the number of hotels in Balikpapan accounted for more than 50 units and continued to increase to almost 90 units in 2018.

As a destination city of tourism and business, the hospitality industry in Balikpapan city should be able to contribute actively to local revenue [2]. However, in the last five years, the total contribution of the hotel sector to local revenue has continued to decline from 10% in 2013 to less than 6.8% in 2017. The lack of contribution to local revenue shows that this sector has low competitiveness. Many scholarly studies have been conducted to investigate the competitiveness of the manufacturing industry, especially in examining the competitiveness of the service sector, tourism destinations, and hotel sector. When the tourism and hospitality industry continue to develop in the global economy, competition among industry members will occur increasingly aggressive and unhealthy [3]. Lower competitiveness brings impact to the level of hotel occupancy rate in Balikpapan city in the last seven years. The hotel occupancy rate in Balikpapan city slightly decreased, resulting in around 47.8 % to 26.7% per month.

The excessive growth of the hotel development can result in an imbalance between demand and supply of these services where supply becomes higher than demand. Such conditions result in intense
competition among accommodation service providers, which in turn can lead to unfair competition. The hospitality industry is one of the leading sectors in the city driving economic growth from accommodation, output in the hotel sector will encourage increased output in other sectors such as restaurants, city transports, and trades. Oversupply may occur due to several factors such as easy procurement of funds and property appreciation, regardless of the intention of the hotelier to build excessively, but oversupply may not always be a consequence of over-development [4].

Previous studies have only focused on evaluating hotel services to increase consumer interest, improving internal industry management, and assessing the impact of hotel growth. For this reason, we try to conduct a spatial analysis of hotel distribution and occupancy rates to identify the degree of saturation in the area of the number of hotels. The results indicate that information on the statistical significance of hotel distribution can be obtained via hot spot analysis. Overall, the hotels in Balikpapan is constructed in an area the most which has high accessibility and closes to urban centers.

2. Method
Hotel classifications in the Balikpapan City, the same as international classifications, are generally based on the quality of services and facilities, such as star hotels, ranging from one-star hotels to five-star hotels, and non-star hotels. In this study, there will be many reviews on how to compare between star and non-star hotels. We carried out two stages in the data analysis to achieve the research objectives, namely (1) spatial pattern analysis of hotel distribution and (2) hotel location cluster analysis based on occupancy rate (hot spot and cold spot).

2.1. Data and scope of study
The study focuses on the entire administration area of Balikpapan City, as a part of the East Kalimantan province, directly adjacent to the Makassar Straits and the Balikpapan bay. The location of the city is at 0 - 1.5 S Longitude and 116.5 - 117.5 E Latitude, with an area of approximately 503.3 km². The city comprises of six subdistricts, namely Balikpapan Kota, -Tengah, -Barat, Selatan, -Utara, and -Timur. As the largest city in East Kalimantan, the number of tourist arrivals in the city reached 2.7 million in 2017, increased from 2.2 million in the previous year [5] visiting tourist attractions, going to transit, or mostly organizing business activities such as meetings, incentives, conventions and exhibitions.

This study uses spatial data such as hotel location coordinates, city boundary, sub-district boundary, and road network maps. In spatial data, we inserted necessary non-spatial data, including hotel classification, number of rooms (supply), and percentage of monthly occupancy (demand) at the attribute data loaded by the Geographic Information Systems (GIS). Spatial data for hotel locations were obtained from the field observation while the attribute data obtained from the Balikpapan City Tourism agency and Balikpapan City Development Planning Agency.

2.2. Spatial pattern analysis
The method used by the researcher to analyze hotel distribution patterns is the Average Nearest Neighbor (ANN) method, which is available in GIS Software. The basic principle of this method is to explain the distribution patterns of the location points by using statistical calculations based on distance, number of locations, and area, measured in Nearest Neighbor Index Values, z-scores and p-values [6]. The p-value and z-score are the degree of significance of the data so that the research hypothesis, generally taking into assumption that incidents are random (Ho), will be accepted or rejected. The Rejection's previous hypothesis results in the opposite conclusion by accepting an alternative hypothesis that the inclination is towards clustered patterns [7].

In this analysis, the hotel distribution pattern is divided into two groups according to the classification of stars and non-stars. Evaluation of object groups, clustered or random, is conducted by interpreting the value of the average Nearest Neighbor ranging from 0 to 2.15. If the index is less than 1, the inclination is towards a clustered pattern, whereas if the index is greater than 1, the tendency is more towards spread [7].
2.3. Hot spot analysis

Hot spots are areas which have sufficient concentrations of certain events or incidents then labeled as high concentration areas [8]. There is no boundary around the incident, but the gradient in the form of an imaginary line indicates the location where the hot spot was started [9]. In general, every hot spot variable measured will appear to be higher in an area, and conversely, in other parts will appear lower.

Getis-Ord Gi* is used to identify comparisons of feature values, high or low, including the inclination to make a cluster in the study area [10]. This tool works by observing at each feature in the context of adjacent features (spatial neighborhood). Therefore, a significant hot spot is shaped when the value of a feature is high; instead, the value for all neighboring features is also high. The Getis-Ord Gi* is depicted in the Gi* Z-score, a critical value of data and Gi* p-value, a level of significance of data. These values compare the local number for each feature and its neighbors to the sum of all features proportionally. A feature data should have a high value and be surrounded by other features with high values as well to be a statistically significant hot spot [9,11].

We use 5% of a significant level in this study. Consequently, significant level (p-value) will be evaluated to should be less than 0.05. If the local sum is very different from the local sum predicted and that difference is too large to be the result of random chance, a statistically significant z-score results [8,9,11]. A high z-score and a small p-value for the feature indicate a significant hot spot. A low negative z-score by with a small p-value indicates a significant cold spot — the higher (or lower) z-scores, the more intense the cluster [8]. Contrastingly, a z-score near zero means no spatial clustered.

In this study, the variable measured based on the Getis-Ord Gi* statistic is the occupancy rate at each hotel location using a fixed distance band in the GIS Software. To determine the fixed distance band, we use the Incremental Spatial Autocorrelation analysis, available in the GIS software. This analysis operates to choose the optimal analysis distance for cluster formation by evaluating a concatenation distance increases and measuring the intensity of spatial grouping for each distance, also shown in the z-score, which illustrates the intensification of the grouping. The optimal distance is indicated by the highest z-score, although sometimes in certain conditions will show several peaks.

3. Results and discussions

3.1. Analysis of hotel location distribution

Currently, Balikpapan city has 84 hotels consisting of non-star and star hotels accounted for 38 and 46 units, respectively. Table 1 below shows that the highest number of hotels, more than a quarter of existing hotels, is located in Balikpapan Kota and Balikpapan Selatan sub-districts. Although both sub-districts have narrower area than others have, they become the most favorite location of hotel development in the city due to closer from the center of government, offices, and business, as well as proximity to regional infrastructures such as international airports, harbors, and main roads. On the other hand, Balikpapan Timur sub-district only has one hotel even though not a star hotel.

| Sub-district       | 5-Star | 4-Star | 3-Star | 2-Star | 1-Star | Non-Star | Aggregate |
|--------------------|--------|--------|--------|--------|--------|----------|-----------|
| Balikpapan Barat   | 1      |        |        |        |        | 5        | 6         |
| Balikpapan Kota    | 2      | 6      | 8      | 3      | 2      | 10       | 28        |
| Balikpapan Selatan | 1      | 3      | 6      | 5      | 3      | 10       | 29        |
| Balikpapan Tengah  | 3      |        | 1      | 11     |       |          | 15        |
| Balikpapan Utara   | 1      | 1      |        |        | 3      |          | 5         |
| Balikpapan Timur   |        |        |        |        | 1      |          | 1         |
| **TOTAL**          | **3**  | **11** | **18** | **8**  | **6**  | **38**   | **84**    |

In contrast to the conditions of the previous sub-districts, the North, East, and West Balikpapan sub-districts are regions with diverse natural tourist destinations. There are various attractions such as Protected Forests—namely Sungai Wain, and Manggar protected forest, Balikpapan Botanical Gardens, Endemic species preservation - such as Borneo sun bears and Bekantan, Mangrove Ecotourism, Dragon
Fruit Agritourism and Coastal Tourism. Nevertheless, the regions are not enough to be called a suitable location for hospitality activities. Only a few hotels exist, especially non-star class in the regions.

Based on the distribution of hotel locations in the city, the analysis shows that they tend to make a cluster, even though at a glance, as in Figure 1 shows the hotels’ urban sprawl from West to East occurred. Spatial data analysis for the hotel location types, both star and non-star classes using the ANN method shows that the spatial patterns are clustered. The output shows that the nearest neighbor average ratio is less than 1, which is 0.62 with an average incident distance or the location gap of each hotel, in the study area is 398 meters. Therefore, the spatial pattern of the hotel distribution is considered to be a group with a z-score and p-value obtained was -6.72 and less than 0.001, respectively. Thus, the null hypothesis (Ho), which states that the spatial pattern of hotel distribution in Balikpapan occurred randomly, is rejected, while the alternative hypothesis (Hi) is accepted. Hereafter, the hotels in Balikpapan City show a clustered spatial pattern indeed.

Subsequently, if the location of the hotel is divided into star and non-star hotel, the previous conclusion about hotels’ cluster propensity remains indisputable. As shown in Table 3 below, both star and non-star hotels tend to make a group with consecutive obtained z-scores by 5.93 and -3.23. Even so, the level of significance of star hotel grouping is remarkably higher than non-star hotels have, indicated by the lower level of p-value. Star hotels agglomerate along the coastline of Balikpapan Kota and Balikpapan Selatan sub-district while non-star hotels gather around the intersection of the main roads, which are part of the sub-districts’ activity center.

| Locations | Nearest Neighbor Ratio | Mean Distance Observation (m) | Z-Score | P-Value | Pattern   |
|-----------|-----------------------|------------------------------|---------|---------|-----------|
| Non-Star Hotels | 0.73                  | 948                          | -3.23   | 0.00126 | Clustered |
| Star Hotels   | 0.54                  | 475                          | -5.93   | 0.00000 | Clustered |
| All types of Hotel | 0.62                  | 398                          | -6.72   | 0.00000 | Clustered |

The mean distance observation in Table 2 shows the pattern of grouping of each type of hotel. Due to a mean distance observation was almost one kilometer, the cluster of non-star hotels generates more than one neighborhood unit spatially. In contrast, the star hotels’ pattern was hardly distinguished how many groups might form because they have closer distances to each other by less than half a kilometer along the coastline. However, the analysis also identifies the outliers, known as unusual star hotel location, whose most far away from the others. The outliers may serve as an alternative hotel with different market option and offer a closer distance to tourist attractions such as hotel in north side dan east side. On the other hand, as a more affordable option, the distribution of non-star hotel groups inclines to maintain proximity to the center of environmental activities rather than being in the city center, which already occupied by the most star hotels.

3.2. Analysis of occupancy rate pattern

In the previous session, we were able to conclude that there was a clustering of hotels’ location distribution in the City. The following analysis was used to identify incidences of hotels' occupancy rates thus the intensity of the incident in particular locations classified as either high/low concentration or dispersed spatially. A spatial autocorrelation (Moran’s I) test method was applied to obtain the spatial patterns, which is used to test global and local spatial autocorrelation among continuous data based on the calculation of the cross products of the deviations from the mean. Regarding to calculation result, variable value at any one location is compared to the values at all other locations statistically [12].

Figure 1 and 2 compare and contrast the hotels’ grouping pattern regarding their occupancy rate between star hotel and non-star hotel. The occupancy rate expresses a ratio between the number of rooms booked and available room in each hotel measured in percentage in specific periods, monthly or annually [2]. Several factors stimulate hotel room occupancy consisting of several advantages externally, for instance, accessibility and strategic location since close to urban business center or attractive landscape and internal attraction such as performance, amenities, price, facilities, and convenience [12,13].
By analyzing the spatial distribution of the star hotels’ occupancy rates (z=2.69), it is evident that the critical value is high, and the null hypothesis can be rejected because the incidents account for the lower significant level (p=0.05). Therefore, as represented in figure 2, the overall data are clustered. With this result, the occupancy rate of star hotel in Balikpapan can be identified as at least 1 group of areas with high or low occupancy concentration levels. Besides, the concentration areas indicate the influence of locational character for personal concerning instead of other considerations such as economic, social, and individual knowledge.

Figure 1. Moran’s I value for occupancy rate of star hotel.  
Figure 2. Moran’s I value for occupancy rate of non-star hotel.

The application of Moran’s I gives significant results from the analysis of star hotel but contrast ones for non-star hotel. In fact, for star hotel type, the spatial autocorrelation establishes that there is clustering of both values, high and low values. Besides, for non-star hotels’ occupancy rate, the Moran’s I statistics says that the distribution of data is random, accounting for negative critical value (z=−0.035) and insignificant level (p=0.97). Therefore, the null hypothesis should be accepted, and there is no clustered pattern of the data. Thus, each non-star hotel has a vary occupancy rate and there is no favorite location being preferred the most by visitors.

3.3. Hot spot analysis based on occupancy rate

Both types of star and non-star hotels are considered local variables because they are different in several levels, such as facility, quality of service and market segmentation, affecting visitors' preferences to make decisions and hotel occupancy levels. Hot spot analysis is arranged to represent the formation of location patterns in each type of hotel. Principally, the analysis of hot spot areas in this study is focused more on locational aspect in intensity incident of hotels' occupancy rate each.

Hot spots for each type of hotels’ occupancy rate were calculated using the Getis-Ord Gi* function. The statistical Getis-Ord Gi* identifies spatial groups of high values (hot spots) and low values (cold spots). The output of the calculation is Gi* Z-score and Gi* P-value for each feature representing the statistical significance of the grouping of spatial values. A high z-score and small p-value for a hotel point indicate a spatial grouping of hot spot with significant high value whereas a low z-score and small p-value indicates a spatial cold spot with significant low values [14]. The higher value of Gi* Z-score obtained, the stronger the grouping is formed [14]. A z-score close to zero indicates that there are unclear spatial groupings or random patterns [12]. The results of this analysis can be presented in table 3.

Table 3. Getis-Ord Gi* Test Summary

| Group Observation | Hotel’s Location Observed | Gi* Z-Score | Gi* P-Value |
|-------------------|--------------------------|------------|------------|
|                   |                          | Min        | Max        | St. Dev    | Min        | Max        | St. Dev    |
| Non-Star Hotel    | 38                       | -0.057     | 4.460      | 0.999      | 0.000      | 0.964      | 0.216      |
| Star Hotel        | 46                       | -2.971     | 2.582      | 1.335      | 0.003      | 0.980      | 0.337      |
| All types of Hotel| 84                       | -2.180     | 4.400      | 1.950      | 0.000      | 0.947      | 0.221      |
Critical value for the overall occupancy rate of hotels in Balikpapan City varies between $z=-2.18$ to $p=4.40$ and significant level ranging from $p=0$ to $p=0.95$. The positive $Gi^* z$-score, being statistically significant with a higher value, indicates a hot spot or higher room sales achieved by hotel. Otherwise, the statistically significant negative $Gi^* z$-score indicates a cold spot. Even though all three groups of observations have a similar $Gi^* z$-score pattern, the establishment of hot spots cannot be generalized partially because each group consists of a different number of locations observed. From table 3, the non-star hotels’ critical value accounts for a shorter range of values from $z=-0.0057$ to $z=4.46$ and smaller standard deviation; hence, the $Gi^* z$-score in each member has insignificantly different. Contrastingly, by constituting the higher value of standard deviation, room occupied percentage in each star hotels have a diverse value defining a higher potential of significant spatial clustering, which can be formed.

![Figure 3. Hot spot distribution of hotels' occupancy rate.](image)

The results of the earlier analysis only measured autocorrelation score statistically and could not show how hot spots and cold spots formed. Therefore, the $Gi^* Z$-score is embedded in the geographical location of each hotel; thus, the hot spot pattern is visible by using GIS Software. Figure 2 illustrates spatial variability in hot spots and cold spots distribution patterns based on star hotels' occupancy rates. According to Moran’s I test results and $Gi^* Z$-score range of the star hotel, the hot spots of occupancy rate can be identified to 5 hotels location adjacent (red star point). They located around the T-junction between Sudirman road and MT. Haryono road (shown as “A”), has a significant level of room sales. This corridor is one of the business and trade centers in the Balikpapan City. Besides, according to cold spots formed, the star hotel situated along the north side of MT. Haryono road (shown as “B”) accounts for the lowest occupancy rate annually, even though the area lies as city culinary and small offices.

In contrast, the occupancy rate of non-star hotel represents non-clusters form since it has insignificant Moran’s I value, neither hot spot nor cold spot. Figure 2 also depicts the data on the distribution of remote spots of non-star hotels in the ratio of rooms occupied, showing that the occupancy rate of non-star hotel in Balikpapan can be recognized as random distribution. However, the analysis results also find two outliers (redpoint), namely Gajah-Mada Hotel and J-Icon Pop Hotel located near the city government, constitute the highest occupancy rate than other non-star hotels have. These hotels attract more visitors to stay measured significantly higher than other non-star hotels do even compared to non-star hotels in a similar location. The average occupancy rate of non-star hotels constitutes less than 30%, while the outliers can rent out amounts to more than 50% of available rooms annually in each hotel. By having those trends, it can be concluded that the consumers of non-star hotel may focus less on the location consideration unless the outliers.
The hot spot analysis able to estimate density distribution of events and identify statistically significant hot spots in the data. Figure 2 shows the established hot spots, both star and non-star hotels which are well distributed around the center of Balikpapan City. Thus, the area close to the city center remains the main factor that strongly influences consumer’s preferences the most for hotel selection.

4. Conclusion
With a useful database that can integrate spatial and non-spatial data, the use of GIS can help us to see the spatial patterns of hotel distribution more quickly and informatively so that it provides convenience in the decision-making process. This study shows that the use of GIS can deliver information that is easy to understand for capturing an urban phenomenon. Besides, the results of the study show that the distribution patterns of hotels in Balikpapan tend to be clustered, both in aggregate and partial, according to two hotel types such as star hotels and non-star hotels. In general, the hotel is developed in an area which has high accessibility and closes to urban centers. This factor also affects hotel selection trends. In downtown areas, as accumulated hot spots, existing hotels have higher occupancy rate, statistically significant, than in other areas.

Acknowledgments
This research was supported by the Ministry of Research, Technology and Higher Education of the Republic of Indonesia through a national competitive research grant scheme in 2019. We also express our appreciation to the City Government of Balikpapan, particularly Balikpapan City Tourism agency and Balikpapan City Development Planning Agency, for sharing their data during this research.

References
[1] Mutmainnah A N, Kolopaking L M and Wahyuni E S 2014 Urbanisasi di Kota Balikpapan: Formasi Sosial Keluarga Pendatang Miskin Jurnal Ilmu Sosial dan Ilmu Politik 18(1) 51-65
[2] Aliandi V D, Anggitasari and Retno H H 2013 Effect of Number of Tourists, Hotels, and Hotel Occupancy Rate on Hotel Tax Revenue (Case Study in Yogyakarta City) Diponegoro Journal of Economics 2(4) 2013 p 1-14
[3] Maya E D 2014 The Effect of Hotel Taxes and Restaurant Taxes on Local Revenues in Batu City Jurnal Pendidikan Ekonomi Universitas Negeri Surabaya 2(3) 2014
[4] Lee K S and Cheong I S 2011 Overcapacity of the US Lodging Industry: The Effect of Uncertainty and Incentive to Overbuild
[5] East Kalimantan’s Central Bureau of Statistics 2018 East Kalimantan in Figures Year 2018
[6] Yusrina N F 2018 Settlement Pattern Analysis Using Nearest Neighbors Approach for Study of Benefits of Tourism Objects in Prambanan District, Klaten Regency Jurnal Geografi, Edukasi, dan Lingkungan 2(1) p 46-55
[7] Masron T, Hussin W M T W, Nordin M N, Yaakub N F and Jamian M A H 2018 Applying GIS in Analysing Black Spot Areas in Penang, Malaysia The Indonesian Journal of Geography 50(2) 133-144
[8] Mrityunjay J and Nityananda S 2016 Modeling of Hotspot Detection Using Cluster Outlier Analysis and Getis-Ord GI* Statistics of Educational Development in Upper-Primary Level, India Modeling Earth System and Environment 2 60
[9] Prasannakumar V, Vijith H, Charutha R and Geetha N 2011 Spatio-temporal clustering of road accidents: GIS based analysis and assessment Procedia-Social and Behavioral Sciences 21 317-325
[10] Harris N L, Goldman E, Gabris C, Nordling J, Minnemeyer S, Ansari S and Potapov P 2017 Using spatial statistics to identify emerging hot spots of forest loss Environmental Research Letters 12(2) 024012
[11] Kalinic M and Krisp J M 2018 Kernel Density Estimation (KDE) vs. Hot-Spot Analysis – Detecting Criminal Hot Spots in the City of San Francisco Proceeding of the 21st Conference on Geo-Information Science
[12] Abdulhafedh A 2017 A Novel Hybrid Method for Measuring the Spatial Autocorrelation of Vehicular Crashes: Combining Moran’s Index and Getis-Gi* Statistic Open Journal of Civil Engineering 7 208-221

[13] Bernini C and Guizzardi A 2010 Internal and Locational Factors Affecting Hotel Industry Efficiency: Evidence from Italian Business Corporations Tourism Economics 16(4) p 883–913

[14] Mazzulla G and Forciniti C 2012 Spatial association techniques for analyzing trip distribution in an urban area European Transport Research Review 4(4) 217–233