# Spatial Pattern Analysis of Dengue Fever using Conditional Autoregressive in Makassar City, Indonesia

Sri Astuti Thamrin[1*], Siswanto[2,]Ansariadi[3], and Andi Kresna Jaya[4]
{tuti@unhas.ac.id[1], siswantosumargo@gmail.com[2],
ansariadi@gmail.com[3], andikresna@gmail.com[4]}

Department of Statistics,Hasanuddin University, Makassar, 90245[1], Department of Statistics,Hasanuddin University, Makassar, 90245[2], Department of Epidemiology, Hasanuddin University, Makassar, 90245[3], Department of Statistics, Hasanuddin University, Makassar, 90245[4],

**Abstract.**Makassar is one of the big cities in South Sulawesi Province which has a high incidence of dengue fever (DF). The number of DF incidence cannot be reduced effectively, even though various efforts have been made. The aim of this study is to analyse the factors that influence the spread of DF incidence using the Conditional Autoregressive (CAR) method. The spatial analysis of DF incidenceis also done using Moran'sIndex. In addition, a map of the spread of DF incidence is made based on the relating factors. This study used the number of DF incidence in Makassar city in the period 2012-2017. The results show that the number of population and socio-economic conditions as malnutrition families describes substantially the spread of DF incidence in Makassar city during the period of 2012-2017. The DF distribution has positive spatial autocorrelations from 2012-2017. The pattern of the spread of DF incidenceis clustered. Some sub-districts included in the category prone to the spread of DF incidence are Manggala, Rappocini, Tallo and Biringkanaya. For this reason, it is recommended that the health office of Makassar city should implement the regional-based disease management to eradicate and prevent the occurrence of DF incidence.

**Keywords:** Conditional autoregressive, Dengue fever, Moran's Index, Socio-economic, Spatial pattern, Spatial autocorrelation

## 1  Introduction

Dengue fever (DF) incidence in Indonesia is still quite high, with the number of incidences in 2017was 68,047 [1]. A total of 2,141 of DF incidence in Indonesia is in the South Sulawesi province in 2018 [2]. In the efforts to achieve the 2019 SDGs target for controlling infectious diseases, the South Sulawesi government continues to make movements in controlling DF. The potential spread of DF is influenced by several factors, including demographic factors such as high population, people who do not understand the behaviour of a healthy environment, many are still storing used goods so that they become a den of mosquitoes, contributing to the high incidence of DF. Makassar city Health Department data showed that in 2017 recorded 135incidences stricken with DF, but in 2018 the number increased twofold to 256incidences [3].

Some supports to reduce the number of DF incidencehave been put in place. One of these is the availability of human resources who are able to spatially analyse surveillance data. Spatial analysis is expected to provide benefits to determine the pattern of spread of DF by area, so as to identify the high-risk areas. It can also see a disease from various contexts so that better planning can be done in eradicating and preventing DF.

Several researches on the spatial analysis of DF have been done so far. For examples, socio-economic factors and climate related to the spread of the dengue disease in New Caledonia [4] and global ecological study of DF viruses on islands with the Bayesian approach [5]. Socio-economic and environmental factors related to the transmission of DF in urban environments in New Caledonia has done by Zellweger et al [6]. The spatial distribution of dengue incidence in Brazil in 2010 using the model Bayesian was discussed by Honorato et al. [7]. Temporal spatial analysis of DFhas been done by Acharya et al. [8], Ling et al. [9] and Huang et al. [10]. Geographical inequality in DF considering Bayesian approach using spatial random effects and geographical mapping of DF were done by Thamrin and Taufik [11] and Thamrin and Alimun [12], respectively. Spatial patterns and its relation with the economic social transmission of dengue in Queensland, Australia was discussed by Hu et al. [13].

The spatial aspect is considered important to be studied because of the spreads of DF through mosquitoes were very influenced by fluctuating climate. Aspects that are the focus of attention in spatial epidemiology are the potential factors affecting the incidence of the disease based on location [14]. One model often used in analysing the area especially in the field of epidemiology is a ConditionalAutoregressive (CAR) model [15]. Therefore,this paper will analyse the factors that influence the spread of DF incidence using CAR. Additionally, the map of incidence of DFwill also be drawn up based on the factors that influence it by considering the location (sub-districts) in the Makassar city in 2013-2017.

In this paper several things will be presented. The background of the problem will be discussed in the first section. Then, the next section, the material and methods related to spatial analysis and the CAR method will be described. The results and discussion and conclusions will be presented in the last section.


## 2  Material and Methods

### 2.1 Sources of Data dan Location

The data used in this paper is the secondary data which is available by the Health Department of Makassar city and the Central Statistics Agency of Makassar (BPS)in 2013-2017. Response variable used are the total population of patients who have been positively identified immunologically dengue (IgM and IgG) in 2013- 2017 in the Makassar city, South Sulawesi province, Indonesia. Then, the explanatory variables are population density ($X_1$), pre-prosperous family ($X_2$) and malnutrition ($X_3$).

Makassar city consists of 14 sub-districts as shown in **Figure 1**. The number of incidences of DF in Makassar city in 2013-2017 is 787. Details of the average number of occurrences of DF per year are presented in **Table 1**. In this paper, a map maker application for data processing was used. Shapefile (shp) which have been produced then opened in map maker application to be processed.
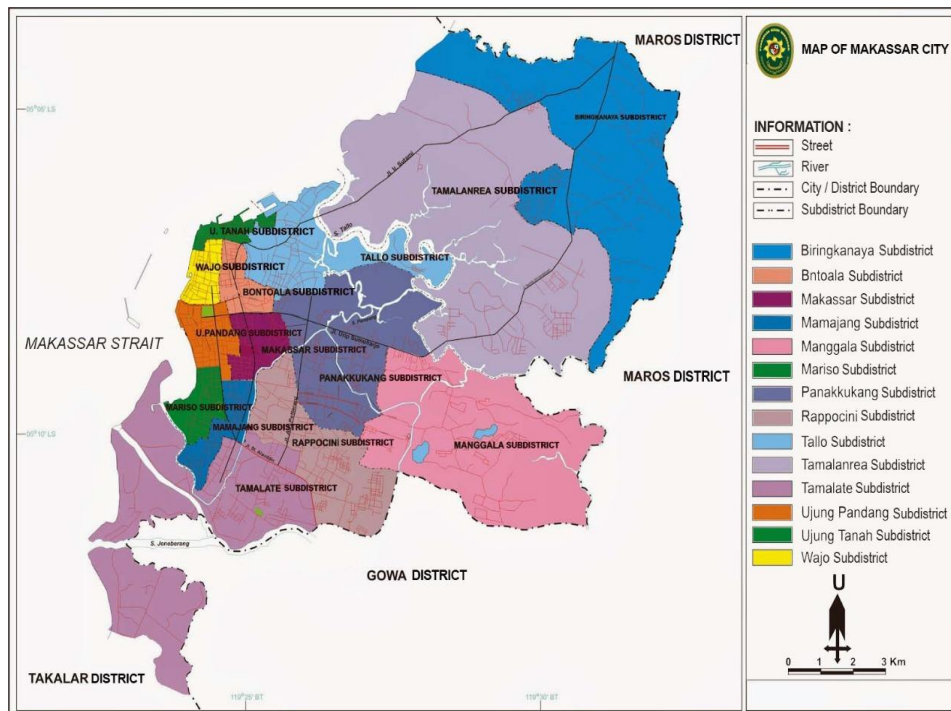
Fig 1. Map of Makassar city, South Sulawesi, Indonesia

Table 1. The mean and variance of the number of DFincidence inMakassar city in 2013-2017.

| Year | Average | Variance |
|------|---------|----------|
| 2013 | 18.929 | 216.379 |
| 2014 | 9.926 | 129.609 |
| 2016 | 17.714 | 153.604 |
| 2017 | 9.642 | 23.786 |

## 2.2 Spatial Analysis

Spatial data contains the information or the geographical location of a region. Locations on the spatial data must be measured to determine the spatial of effects. The information of location can be determined from two sources, namely an adjacency relationship and a distance [16]. The adjacency relationships reflect the relative locations of the spatial unit or location to another in a given space. The adjacency relationships of spatial units are usually formed based on the map.

Weighting or often referred to as a spatial weighting matrix is important in spatial analysis. Spatial weighting matrix is basically a matrix that describes the relationship within the region and it is obtained by distance or adjacency information. There are several possibilities for making a weights matrix items, namely contiguity rook, bishop and queen contiguity [16]. Basically, there is no theory that explains specifically in terms of the selection of weighting matrix for the dependent spatial models. However, in this paper, the queen

contiguity spatial weights matrix will be used. Matrix elements are defined as 1 for regions that are side by side or their vertices meet the area of concern, while other regions are defined as weighting matrix elements of zero.

The pattern of spatial relationships in geographical data consists of two methods, namely global and local. Global method is the analysis of patterns of spatial relationships on a broad scale to see the distribution of data, whether in cluster, spread or random. The method used are Moran'sIndex and Geary's C. Local method is the quantification of the spatial autocorrelation in a smaller area than the global. The method used is Local Moran's Index Indicator of Spatial Autocorrelation (LISA).

In this paper, method is limited only to the Moran's Index. Moran's Index is the most widely method is used to calculate the global spatial autocorrelation. Moran's Index is also used to measure the global spatial autocorrelation spatially interrelated within locations. This index is a development of the Pearson correlation. Spatial autocorrelation calculation with Moran's Index can be done in two ways, namely Moran'SIndex with no standardized spatial weighting matrix $W^*$ and Moran's Index with standardized spatial weighting matrix $W$ [16]. The range of values of the Moran's Index is $-1 \leq I \leq 1$. Moran's Index values of $-1 \leq I \leq 0$ indicates the negative spatial autocorrelation while the values of $0 < I \leq 1$ indicates the positive spatial autocorrelation. Moran's Index value of zero indicates no autocorrelation.

Testing the hypothesis on parameters $I$ can be done by determining the hypotheses, namely $H_0$:I = 0 (there is no spatial autocorrelation) and $H_1$: I $\neq$ 0 (there is a spatial autocorrelation). Then, the test statistic of Moran's Index $I$ statistic is derived in the form of standard normal random variable. It is based on the theory whereby the central limit theorem for large $n$ and variance known then $Z(I)$ will be a normal standard distribution [17]. $Z(I)$ can be written as:

$$Z_{count} = \frac{I - E(I)}{\sqrt{Var(I)}} \approx N(0,1) \tag{1}$$

The test in equation (1) will reject the initial hypothesis $H_0$ if the value of $Z_{count} > Z_\alpha$ or $Z_{count} < Z_\alpha$. Positive spatial autocorrelation between locations indicates that the observation has a close relationship.

LISA was introduced by Anselin [18] and provides a way to assess the level of local spatial patterns. In contrast to the previously described, Moran's Index indicates the global autocorrelation, the Moran's Index in LISA indicates the local autocorrelation. LISA identifies how the relationship within an observation location to the other observations' location. The index [15] can be obtained as $I_i = z_i \sum_j w_{ij} z_j$ with $z_i$ and $z_j$ is the deviation from the average. It can be written as $z_i = \frac{(x_i - \bar{x})}{\delta}$ and $z_j = \frac{(x_j - \bar{x})}{\delta}$ with $\delta$ a standard deviation of $x_i$ and $x_j$.

The CAR model is a conditional model of the observed random variable at one location when another location has known [18]. If $Z$ distributed normally then conditional probability function is:

$$f(Z(s_i)|\{Z(s_j): j \neq i\}) = (2\pi\sigma_i^2)^{-\frac{1}{2}}\exp{-\frac{1}{2\sigma_i^2}\left[\{Z(s_i) - \theta_i(\{Z(s_j): j \neq i\})\}^2\right]} \tag{2}$$

For $i = 1, ..., n$ with $f$ in equation (2) is the conditional probability function of $Z(s_i)|\{Z(s_j) = z(s_j); j = 1, ..., n; j \neq i\}$. $\theta_i$ and $\sigma_i^2$ is the mean and variance of conditional, respectively [19]. Based on the equation (2), CAR models can be written in the form:

$$(Z(s_i)|\{Z(s_j): j \neq i\}) = \mu_i + \rho \sum_{j=1}^{n} W_{ij}(Z(s_j) - \mu_j), \quad i = 1, ..., n. \tag{3}$$

with $\mu_i$ is the expected value to $i$, and $\rho$ is the spatial autocorrelation parameter that determines the size and property (positive or negative) of the spatial effect. The sum of the equation (3) is the average amount corresponding weighting of all others locations [20, 21].

# 3 Results and Discussion

Testing of Moran's Index is done by determining the weighting matrix first. In this paper, the weighting matrix used is the queen contiguity. It is weighting based on the proximity of certain observations with other observation sites. Comparison Moran's Index value, the expected value $E(I)$ and the variance value Var(I) are presented in **Table 2**. The linkage spatially for the spread of DF is measured through spatial autocorrelation by using the Moran's Index.

Table 2. Comparison of value of Moran's Index, E(I), Var(I) and p-value per year for DF in Makassar city

| Year | Moran's Index value | E(I) | Var(I) | p-value |
|------|--------------------|------|--------|---------|
| 2013 | 0.390 | -0.077 | 0.022 | 0.000 |
| 2014 | 0.082 | -0.077 | 0.022 | 0.144 |
| 2016 | 0.315 | -0.077 | 0.022 | 0.004 |
| 2017 | 0.087 | -0.077 | 0.024 | 0.146 |

Based on **Table 2**, Moran's Index value in 2013 and 2016 are $I= 0.390$ and $I= 0.315$, respectively and with p-values of 0.000 and 0.004. This suggests that there is likely spatial autocorrelation in the spread of DF in Makassar city. Then, in 2014 and 2017, the value of Moran's Index is $I= 0.082$ and $I= 0.087$, respectively with the p-values is 0.144 and 0.146. This shows that in 2014 1nd 2017 there were no spatial correlation. The patterns formed in 2014 and 2017 is random related to the spread of DF in the Makassar city.

Table 3 Comparative analysis of CAR model for incidence of DF per year in Makassar city

| Year | | Intercept | $X_1$ | $X_2$ | $X_3$ |
|------|------|-----------|-------|-------|-------|
| 2013 | Coefficient | 1.939 | $2.369\times10^{-4}$ | $-1.237\times10^{-3}$ | $-2.129\times10^{-2}$ |
| | Standard error | 8.014 | $9.318\times10^{-5}$ | $2.253\times10^{-3}$ | $3.988\times10^{-2}$ |
| | p-value | 0.809 | 0.011 | 0.583 | 0.594 |
| 2014 | Coefficient | -5.115 | $1.715\times10^{-4}$ | $-1.741\times10^{-5}$ | $-1.179\times10^{-2}$ |
| | Standard error | 3.621 | $4.318\times10^{-5}$ | $1.383\times10^{-3}$ | $2.811\times10^{-2}$ |
| | p-value | 0.158 | $7.140\times10^{-5}$ | 0.990 | 0.675 |
| 2016 | Coefficient | 2.325 | $1.818\times10^{-4}$ | $-1.142\times10^{-3}$ | $3.972\times10^{-3}$ |
| | Standard error | 4.408 | $3.514\times10^{-5}$ | $8.967\times10^{-4}$ | $1.583\times10^{-2}$ |
| | p-value | 0.598 | $2.302\times10^{-7}$ | 0.203 | 0.802 |
| 2017 | Coefficient | 3.777 | $3.252\times10^{-5}$ | $-2.479\times10^{-4}$ | $2.286\times10^{-2}$ |
| | Standard error | 1.870 | $2.838\times10^{-5}$ | $1.038\times10^{-3}$ | $2.099\times10^{-2}$ |
| | p-value | 0.043 | 0.025 | 0.811 | 0.028 |

Furthermore, **Table 3** presented CAR comparative analysis per year for DF in Makassar city which have been obtained by using equation (3). From **Table 3**, the results show that the spread of DF in one area affect the other areas. Based on the significance test, in 2013, the total population $(X_1)$ affect the number of DF incidence significantly. This means that any increase in the value of the average number of population $(X_1)$ cause the number of patients with DF rose as $2.369\times10^{-4}$. It also occurred in 2014 and 2016, only population $(X_1)$ affects significantly the number of DF incidence. Meanwhile in 2017, the number of population $(X_1)$ and the number of malnutrition $(X_3)$ describes the number of DF incidence significantly. This means that any increase in the value of the average number of population $(X_1)$ cause the

number of patients with DF rose as $3.252 \times 10^{-5}$ and any increase in the value of the average number of malnutrition ($X_3$) cause also the number of patients with DF rose as $2.286 \times 10^{-2}$. The results that are obtained is associated with the one that Teurlai et al. [4], Zellweger [6] dan Hu et al. [12] have done, that is socio-economic factors can be influenced the transmission of DF. Moreover, the result of this paper is similarly with the work of Ling et al. [9], that the number of populations can affect spatial patterns of the spread of DF.
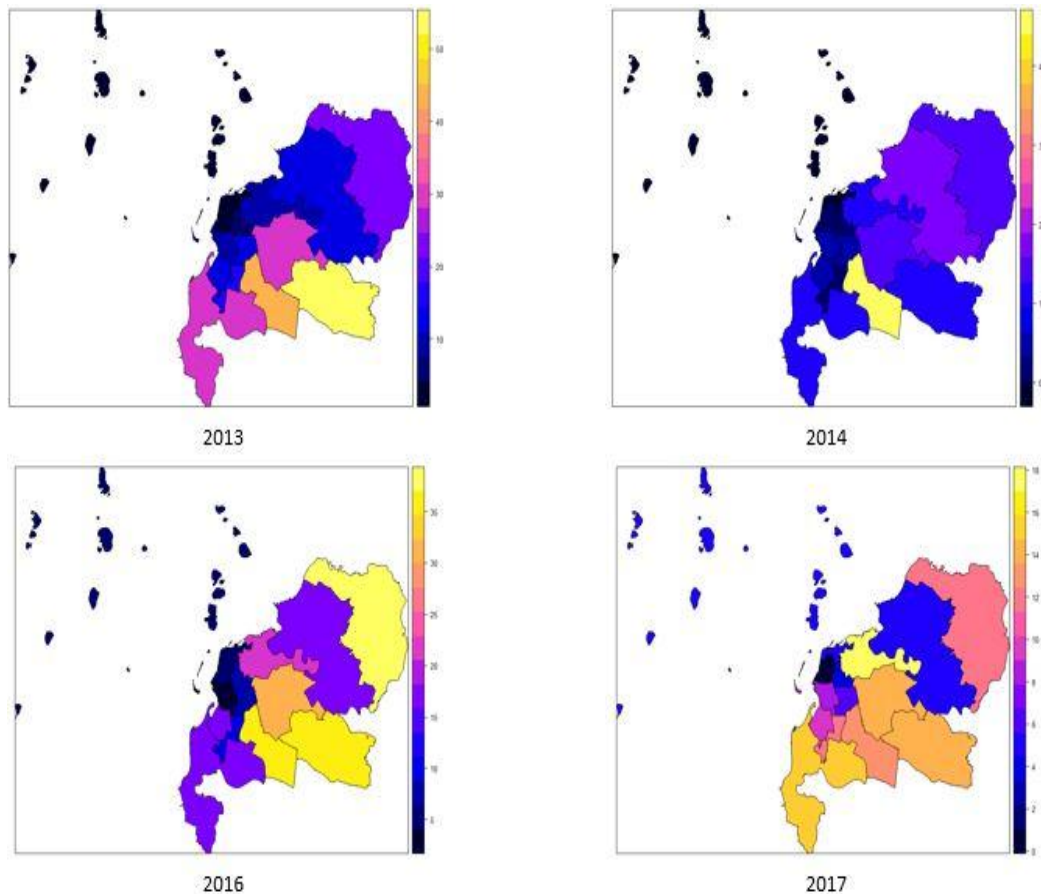


Fig 2. Map of the distribution of DF in Makassar city in 2013-2017

Map of the spread of the number of DF incidence in 2013-2017 in the Makassar city based on factors that influence can be seen in **Figure 2**. The factors of geography and socio-economics conditions in Indonesia influence the spread of DF in Makassar city. **Figure 2** shows the sub-districts with a high number of DF incidence occurred in the sub-districts of Manggala, Rappocini, Biringkanaya and Tallo, which are marked in yellow on the map. Sub-districts with the number of DF incidence are low occurred in the sub-districts of Wajo, Bontoala, Mamajang and Ujung Pandang. It was marked in black on the map. The clustering of incidence of DF in Makassar city has a tendency to follow by high population densities, the economic conditions of the community, especially nutrition, and the geographical conditions such as urban and rural areas. These results were confirmed by studies carried out by

Honorato et al. [7], Acharya et al. [8], Thamrin and Taufik [12] and Hu et al. [13]. The number of larvae is also likely to affect the spread of DF but has not been included as a contributing factor in this paper [5].

## 4 Conclusion

The spatial pattern analysis of DF from 2013 to 2017 in Makassar city, Indonesia has done using CAR method. The pattern of distribution of DFincidence occurring in Makassar city appears clustered. This shows that the environment is heterogeneous. The clustered pattern of distribution of DFincidence is an indicator that there is a concentration of mosquito vector habitat, so there is a potential for greater local transmission. In general, the clustering of incidence of DF in Makassar city has a tendency to follow high population densities, the economic conditions of the community, especially nutrition, and the geographical conditions such as urban and rural areas. As the impact, the surrounding area becomes susceptible to DF. This can be the basis of information in an effort to control strategies and prevention of DF in Makassar city. The large number of population factors and malnutrition status in several sub-districts in Makassar city influenced the process of transmission of DF from one person to another. The number of larvae is also a little likely to affect the spread of DF and it is important to include it as a contributing factor in future studies.

## References

[1]    Data and Information Center of the Ministry of Health. Dengue Situation in Indonesia in 2017. InfoDatin. Jakarta. ISSN: 2442-7659. (2018)

[2]    Antaranews. DinkesSulsel: 10 penderita DBD meninggal dunia. https://makassar.antaranews.com/berita/110282/dinkes-sulsel-10-penderita-dbd-meninggal-dunia, Online access January 31, 2019.(2019)

[3]    Antaranews. Dinkes Makassar mintawargaantisipasi DBD. https://makassar.antaranews.com/berita/111807/dinkes-makassar-minta-warga-antisipasi-dbd. Online accsess 12 February 2019. (2019)

[4]    Teurlai M, Menkès CE, Cavarero V, Degallier N, Descloux E, Grangeon JP., Guillaumot L., Libourel T., Lucio PS., Mathieu-Daude F., and Mangeas M. Socio-economic and Climate Factors Associated with Dengue Fever Spatial Heterogeneity: A Worked Example in New Caledonia. *PLOS Neglected Tropical Diseases* 9(12): e0004211 (2015)

[5]    Feldstein, L. R., J. S. Brownstein, O. J. Brady, S. I., Hay, and M. A. Johansson. Dengue on islands: a Bayesian approach to understanding the global ecology of dengue viruses. *Transaction Royal Society Tropical Medicine and Hygiene* 109: pp. 303–312 (2015).

[6]    Zellweger, R.M., Jorge C., Morgan M., Francois T, Alize M, Marc D, Christophe E, Myrielle DR, Birgit N and Magali T. Socio-economic and environmental determinants of dengue transmission in an urban setting: An ecological paper in Nouméa, New Caledonia. *PLOS Neglected Tropical Disease* 11(4): e0005471 (2017)

[7] Honorato, T., P. P. D. A. Lapa, C. M. M. Sales, B. R. Santos, R. Tristão-Sá, A. I. Bertolde, and E. L. N. Maciel. Spatial analysis of distribution of dengue incidence in Espírito Santo, Brazil, in 2010: Use of Bayesian model. RevistaBrasileira de Epidemiologia 17: pp. 150-159 (2014)

[8] Acharya, B.K., C. X. Cao., T. Lakes., W. Chen., and S. Naeem. Spatio temporal analysis of dengue fever in Nepal from 2010 to 2014. *BMC Public Health* 16: pp. 849 (2016)

[9] Ling CY, Gruebner O, Krämer A, Lakes T. Spatio-temporal patterns of dengue in Malaysia: combining address and sub-district level. *Geospatial Health* 9 (1): pp. 131–40 (2014)

[10] Huang, X., L. Yakob, G. Devine, F. D. Frentiu, S. Y. Fu and W. Hu. Dynamic Spatiotemporal Trends of Imported Dengue Fever in Australia. *Scientific Reports* 6: pp. 30360-30368 (2016)

[11] Thamrin, S. A., and I. Taufik. Spatial Random Effects Survival Models to Assess Geographical Inequalities in Dengue Fever Using Bayesian Approach: A Case Paper. *Journal of Physics: Conference Series* 979: pp. 1-9 (2018)

[12] Thamrin, S. A., and Alimun. Geographical Mapping of Dengue Fever Incidence 2012-2016 in Makassar, Indonesia. *Earth and Environmental Science* 279 (012013): pp. 1-8 (2019)

[13] Hu, W., A. Clements, G. Williams, and K. Mengersen. Spatial Patterns and Socio-ecological Drivers of Dengue Fever Transmission in Queensland, Australia. *Environmental Health Perspectives* 120: pp. 260-266 (2012)

[14] Waller LA, Gotway CA. *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons, Inc (2004)

[15] Oliveira VD. Bayesian Analysis of Autoregressive Models. *Annals Institute of Statistics and Mathematics* 64: pp. 107-133 (2012)

[16] Kosfeld R. *Spatial Econometric*. URL: *http://www.scribd.com*(2016)

[17] Lee J, Wong SD. *Statistical analysis with Arcview GIS*. New York: John Willey & Sons. Inc (2001)

[18] Anselin L. Local indicators of spatial association—LISA. *Geographical Analysis*. 27(2): pp. 93–116 (1995)

[19] Wall MM. A Close Look At The Spatial Structure Implied By The CAR and SAR Models. *Journal of Statistical Planning and Inference.* 121: pp. 311-324 (2004)

[20] Smith MJD. Statistical Analysis Handbook A Comprehensive Handbook of Statistical Concepts, Techniques and Software Toolhttp://www.statsref.com/HTML/index.html?car_models.html. (2015)

[21] Cressie NAC. *Statistics for Spatial Data*. John Wiley & Sons, Inc (1990)