

# Exploratory Analysis in Mapping of Asthma Risk in Western Australia

Hendra Kurnia Febriawan and Carla Maria da Silva Sodre

Received: Nov 2017 / Accepted: June 2018

©2018 Faculty of Geography UGM and The Indonesian Geographers Association

**Abstract** Exploratory Data Analysis (EDA) is one of the spatial analysis methods to obtain the spatial pattern and get a relationship between the dependent and independent (explanatory) variables. The purpose of this study is to investigate the prevalence of asthma in Western Australia using EDA and confirmatory data analysis (CDA). Four provided variables (humidity, annual rainfall, EVI (Enhanced Vegetation Index) and SEIFA (Socio-Economic Indexes)) have been tested to get the best relationship that underlying the asthma prevalence in Western Australia. The EDA result indicates that high level of humidity, high level of annual rain, the positive value of EVI and low level of SEIFA index are the main factors that contribute to the asthma prevalence. However, the result of CDA reveals that only low level of SEIFA that contributes to the asthma prevalence. The negative value of EVI and low level of annual rain become dominant factors in the disease.

**Keywords:** Exploratory Analysis, Spatial Analysis, Asthma, Western Australia.

**Abstrak** Analisis Data Eksploratori (ADE) merupakan salah satu metode analisis spasial untuk mendapatkan pola spasial dan hubungan antara variabel dependen dan independen. Tujuan penelitian ini adalah untuk meneliti persebaran asma di Australia Barat menggunakan metode ADE dan analisis data konfirmatori (ADK). Empat variabel (kelembaban, curah hujan tahunan, EVI (Enhanced Vegetation Index) dan SEIFA (Socio-Economic Indexes)) di uji untuk mendapatkan hubungan yang mendasari persebaran asma. Hasil dari ADE mengindikasikan bahwa tingkat kelembaban yang tinggi, curah hujan yang tinggi, nilai EVI positif dan nilai SEIFA negative merupakan faktor utama pada persebaran asma. Namun, hasil dari ADK menunjukkan bahwa hanya nilai SEIFA yang rendah yang berkontribusi pada persebaran asma. Nilai EVI negative dan curah hujan yang rendah merupakan faktor dominan pada persebaran asma.

**Kata kunci:** Analisis Eksploratori, Analisis Spasial, Asma, Australia Barat

## 1. Introduction

Asthma is one of the chronic diseases in societies and the main factors responsible outside the environmental factors are still not confirmed [Tattersfield, Knox, Britton, & Hall, 2002]. Asthma is one of the most common chronic diseases in Australia and becomes a National Health Priority Area in Australia. Although the mortality of this disease has fallen by 70% from its peak in the 1980s, this number is still high compared with other countries, especially in young people between 5 – 34 years old [Reddel, Sawyer, & Everett, 201]. Mayo Clinic [2016] assumes that the triggers of asthma are broad and different from each person. There are several main causes such as air substances (pollen, dust mites, mould spores, pet dander, particles of cockroach waste), respiratory infections such as the common cold, physical activity, cold air, as well as air pollutants and irritants [Mayo Clinic, 2016].

There are many studies regarding Asthma prevalence worldwide. Those studies usually used common statistical analysis such as multinomial logistic regression using SAS 9.1 [Black et al., 2011],

univariate statistics using STATA statistical software [To et al., 2012], statistical analysis using SPSS and R [Mainardi et al., 2013], and questionnaire statistical analysis [Hekking et al., 2015]. The spatial approaches in the Asthma prevalence are also used widely in many studies. Those studies used various approaches from the quite simple methods such as kriging interpolation in spatial patterns analysis [Gorai, Tchounwou, & Tuluri, 2016], multivariate linear regression and correlation analysis [Skarkova et al., 2015; Krstić, 2011] to the more complex methods such as structured additive regression (STAR) model [Chien & Alamgir, 2014], Spearman correlation coefficient [Anderson et al., 2012], Logistic-binomial models based on a paired-poisson algorithm and global autocorrelation [Shankardass, Jerrett, Dell, Foty, & Stieb, 2015] and combination global (Moran's I) and Local Indicators of Spatial Association (LISA) [Crighton, Feng, Gershon, Guan, & To, 2012].

There are studies of Asthma prevalence in Western Australia that were usually concentrated in the Perth metropolitan region. However, these studies usually utilized the common statistical methods such as ANOVA or Kruskal-Wallis test [Collison et al., 2011], with minimum use of spatial approaches and only for visualization maps purposes, for instance Buffer analysis in ArcGIS and Conditional logistic regression [Cook, Devos, Pereira, Jardine, & Weinstein, 2011],

**Hendra Kurnia Febriawan and Carla Maria da Silva Sodre**

Post-Graduate of Department of Spatial Sciences, Curtin University-  
Building 207, Kent Street, Bentley, Western Australia 6845.

Correspondent Email: [henz.febriawan@gmail.com](mailto:henz.febriawan@gmail.com), [ba.carlasodre@gmail.com](mailto:ba.carlasodre@gmail.com)

Bayesian spatial methods [Speldewinde, Cook, Davies, & Weinstein, 2011], Conditional logistic regression using SAS [Pereira, De Vos, Cook, & D'Arcy, 2010] and Conditional logistic regression [Pereira, De Vos, & Cook, 2009a].

For that reasons, this study is aimed to investigate the prevalence of asthma in Western Australia and identify the primary causal factors that related to that prevalence using exploratory spatial data analysis approach. The explanatory data analysis is not a complicated method to investigate the patterns and factors that are underlying this disease. Thus, it can complement the existing studies regarding Asthma prevalence in Western Australia in particular.

## 2. The Methods

The data that was used in this study was provided by Department of Spatial Sciences, Curtin University, Western Australia. The data in shape files contains the feature of suburbs in Western Australia. The data contains several factors that are related with asthma such as the number and percentage of people infected by asthma, the cars, buses and heavy vehicles information, some particles in atmosphere (CO, NO<sub>2</sub>, O<sub>3</sub>, SO<sub>2</sub>, PM<sub>10</sub>, PM 2.5), Enhanced Vegetation Index (EVI), annual rain, humidity and Socio-Economic Indexes for Areas (SEIFA). However, this study only uses four main factors (humidity, annual rain, enhanced vegetation index (EVI) and Socio-Economic Indexes for Areas (SEIFA)).

Many studies reveal that humidity over 50 per cent can lead to a higher incidence of asthma since the humid air is denser and harder to breathe and humid air harbors fungus, moulds and dust mites that trigger asthma [Bottrell, 2011]. The humidity levels above 80% are also dangerous for asthmatic patients since the humid air can reduce the number of negative ions in the air that generally ease the process of breathing [Perry, 2017].

The heavy rain can burst the pollen particles become which much smaller and can be dispersed by wind over quite a distance. These small particles can quickly enter the airways of small children and can cause an acute asthma attack in those who react to pollen [Australia Wide First Aid, 2018].

The enhanced vegetation index (EVI) is an index that represents the level of vegetation density as the product of Terra and Aqua Moderate Resolution Imaging Spectroradiometers (MODIS) [Jiang & Huete, 2008]. EVI will return with positive values that represent the health vegetation (the more higher numbers indicate the healthier vegetation and negative values indicate a lack of vegetation such as water, rock, and soil [Davinci.asu.edu, 2013]. Vegetation is vital regarding producing pollen since the time of pollinations depend on the flowering and growth of plants. This pollen can expose the susceptible people with allergic diseases such as asthma [Kececi, 2017].

Socio-Economic Indexes for Areas (SEIFA) is a number described as the level of the socio-economic index in an area. The baseline index is designed in 1,000, which the score above 1,000 indicates an area of socio-economic advantage and the score below 1,000 indicates an area of socio-economic disadvantage. In addition, the lower SEIFA also correlates with lower health status and high-risk factors of mortality [Health Service, 2016]. Cunningham [2010] found that Asthma diagnoses were reported more in Indigenous people than non-Indigenous people in Australia (in the age group 18 – 64 years) since these Indigenous people tend to experience more in racism, and discrimination, marginalization and dispossession, chronic stress and exposure to violence. The study in Perth city area also showed that the geographic variability in risk of Asthma distribution is not fully represented by socio-economic disadvantage [Pereira, Vos, & Cook, n.d.].

Exploratory data analysis is aimed to detect the pattern in spatial data by using graphical views of the data, highlight meaningful clusters of observations, unusual observations, or relationship between variables and it can be used to develop hypotheses [Pisati, 2012]. The tools can use the description of each variable, choropleth maps, correlation graphs and matrix graphs. The flow chart of the study can be seen in figure 1.

Maps are an important tool for displaying the spatial data for exploratory analysis purpose since maps can help to detect the patterns of the data and spatial relationship between phenomena [Pisati, 2012]. In this study, the asthma percentage variable was chosen as the primary variable and find the correlation with other variables. There are several regions that contain zero value of asthma and these regions were not included in the analysis. The asthma percentage was displayed in the map using ArcGIS 10.5 by classified it into three classes, low asthma percentage (< 8%), medium asthma percentage (8 - 10 %) and high asthma percentage (> 10%).

The other variables were also visualized in the map using ArcGIS into three classes. The humidity was classified using Equal Interval into three categories: low humidity (<51%), medium humidity (51 – 63 %) and high humidity (>63%). The annual rain was classified using Equal Interval into three categories: low rain (<536 mm), medium rain (536 – 826 mm) and high rain (>826 mm). The SEIFA was classified using Equal Interval into two categories: low SEIFA (<1000) and high SEIFA (>1000). The EVI was classified into two categories: positive EVI and negative EVI. The summary of the classification process can be seen in table 1.

The proportion of each asthma percentage from the entire population then was calculated and the autocorrelation was performed to get the correlation between the asthma percentage and other variables. The spatial pattern shows the distribution of the data using choropleth maps. In this regard, the high asthma percentage was chosen as the main variable and then

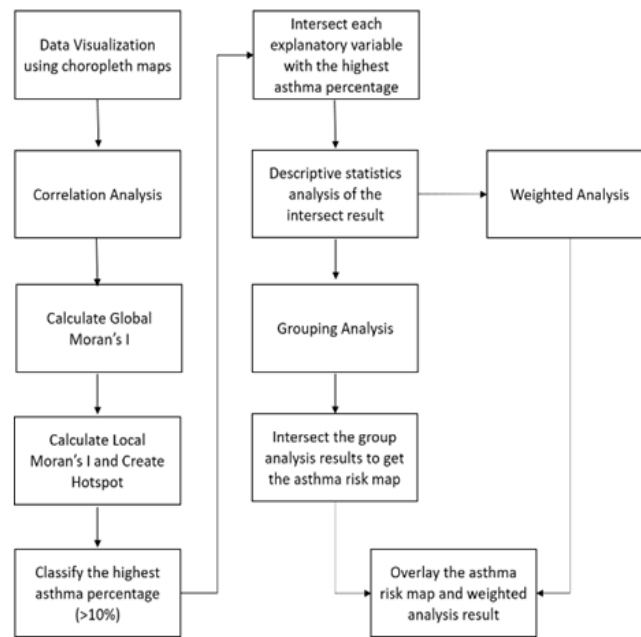


Figure 1. Flow-chart of Exploratory Data Analysis

Table 1. Asthma Percentage Variable

Layer	Class	Values	Description
Asthma percentage	1	< 8%	low asthma
	2	8 - 10 %	medium asthma
	3	> 10%	high asthma
Humidity	1	<51%	low humidity
	2	51 – 63 %	medium humidity
	3	>63%	high humidity
Annual rain	1	<536 mm	low rain
	2	536 – 826 mm	medium rain
	3	>826 mm	high rain
SEIFA	1	<1000	low SEIFA
	2	>1000	high SEIFA
EVI	1	<0	positive EVI
	2	>0	negative EVI

was compared to the other variables. The correlation of each distribution variable that has been chosen then was evaluated and overlaid with the main variable to get the most influential factor causing asthma. The correlation was examined using Locally Weighted Scatterplot (LOESS) smoothing tool in the PAST software. The descriptive statistical analysis was conducted to get the more precise result. The result of this correlation then was used to build the hypothesis.

In order to examine the spatial pattern and spatial autocorrelation, the global and local autocorrelation were performed. There are two types of spatial autocorrelation: global spatial autocorrelation (Global Moran's I statistics) and local spatial autocorrelation (Anselin's local Moran's I (LISA)). The Moran's I will

have a value between -1 (dispersed) to 1 (clustered) and also the larger the Z-score than 1.96 or less than -1.96 will show a significant value at the 95% confidence level [Górniak, 2016; Song et al., 2014]. The area with positive spatial autocorrelation will show a cluster of similar phenomena and negative spatial autocorrelation will depict a cluster of different phenomena [Rousta, Doostkamian, Haghghi, Ghafarian Malamiri, & Yarahmadi, 2017]. In addition, the hot spot location with significant statistic will have high values with its surroundings (High-high or Low-low values) [Zhou et al., 2017].

Confirmatory Spatial Data Analysis is the quantitative processes of modelling, estimation and validation necessary for the analysis of spatial

components and it is aimed to find a good correlation between predicted (hypothesis) and observed values [Lopes, São-carlense, Carlos, & São-carlense, 2007]. The weighted mean maps can also be used to confirm the proximity of each influence variables since it shows the mean value of each variable that is already weighted by the high majority values. In the confirmatory analysis, the high asthma percentage region (>10%) was used to get the most significant influence of other major factors in this region. Thus, that region subset was classified and separated from other regions. The analysis was conducted by intersecting the high asthma percentage with other variables and calculates the area of intersection (overlay) to get the proportion of each variable to the high asthma percentage. In addition, the weighted analysis was performed to get the concentration location of high asthma with other variables. The weighted analysis is usually used to measure the spatial characteristic of features when they are weighted by a variable [ArcMap, 2017]. In this study, the high asthma region was weighted with each of the variables and displays the result in a map.

The Grouping Analysis tool in ArcGIS was used to identify distinct groups of areas based on the relation between asthma percentage and each of the explanatory variables. This tool has a purpose to classify and to cluster the areas based on specified attributes and similar statistics using their spatial distribution and the combined indexes [Delclòs-alió, Miralles-guasch, & Jacobs, 2018; Ronchi, Salata, Arcidiacono, & Ronchi, 2018]. This tool uses K-Means algorithm to partition areas into groups based on the chosen variable, reclassify the values in iterations and when the groups are stable, then it will create group similarity and group differences [Delclòs-alió, Miralles-guasch, & Jacobs, 2018; Lomba et al., 2017]. In addition, Venkatramanan, Chung, Rajesh, & Lee [2015] explained that this tool examines the divisor of the groups using the Calinski-Harabasz pseudo F-statistic that shows group similarity and between-group differences. In this study, the non-spatial constraints were used for the calculation since the purpose is to get the statistical differences of the variable without considering the spatial contiguous lag. The variables were assigned into three groups of regions that show the relationship between high, medium and low asthma percentages with each of the variable.

### 3.Result and Discusion

The distribution of asthma percentage in Western Australia can be seen in figure 2(a) and can be noticed that the distribution of high asthma percentage spread in the south part of Western Australia. There are several huge regions that contain zero number percentage of asthma (green colour) where then were excluded in the analysis. The high percentage asthma regions occupied 64.8% of the total region (exclude the zero number), only 28% of regions contain medium percentage of asthma and the low percentage asthma regions existed

in 7.1% from the entire region. Those calculations were conducted by neglecting the 0% asthma regions. Perth city itself only contains medium level of asthma percentage. Although it was believed that the proportion of cars, buses and heavy vehicles are high in this city, it seems that those variables are not the main causal factors in the asthma prevalence.

It can be seen from the figure 2(c) that over 70% of Western Australia regions have low annual rain and the high annual rain only occupied 10% of the total region and located in the northern and southern part of Western Australia. Almost of 85% from the total area in Western Australia have negative EVI and the positive EVI areas are only located in the northwest and south part of Western Australia with only 15% of the total area (figure 2(e)). In addition, figure 2(f) also depicts that over 80% of the Western Australia regions have low SEIFA index (less than 1000). Figure 2(c) and (d) show an interesting similarity. The high annual rain occurred in the north and south part of Western Australia, where these locations have medium to high humidity. Although the areas with high humidity level are located in the southwest of Western Australia (figure 2(d)), these areas contain negative and positive EVI (figure 2(e)).

It can be seen from the table 2 that the humidity level` in the region is medium-high with the average is 62.2%. The annual rain considerably varied from 246.77 mm to 1116.06 mm per year, but since the median is 444.47 mm and less than the average value (540.61 mm), it means that the majority of the regions have low annual rain. In general, the regions contain less of vegetation density since the EVI index shows that the mean and median are below 0 (-0.003139 mm and -0.003131 mm respectively). In addition, generally, the regions in Western Australia can be categorized in disadvantage region since the mean and median of the SEIFA index are 981.54 and 991.15 respectively and these values are below the minimum criteria of advantage region (should be >1000).

The relationships between the dependent variable (Asthma Percentage) and all of the explanatory variables also can be obtained using Locally Weighted Smoothing (LOESS). The result of LOESS shows the relationship of explanatory variables with the asthma percentage in a scatterplot graph (figure 3). The correlation between asthma percentage and other explanatory variables revealed that in general all of the variables have a positive relationship with the asthma percentage (figure 3). While the annual rain and EVI show a robust positive relation until a particular point and then decrease, the humidity and SEIFA indicate a gradual increase, followed by a stable trend, and finally increases with a strong positive trend. The global Moran's I autocorrelation of all variables were calculated using ArcGIS and the result can be seen in the table 3.

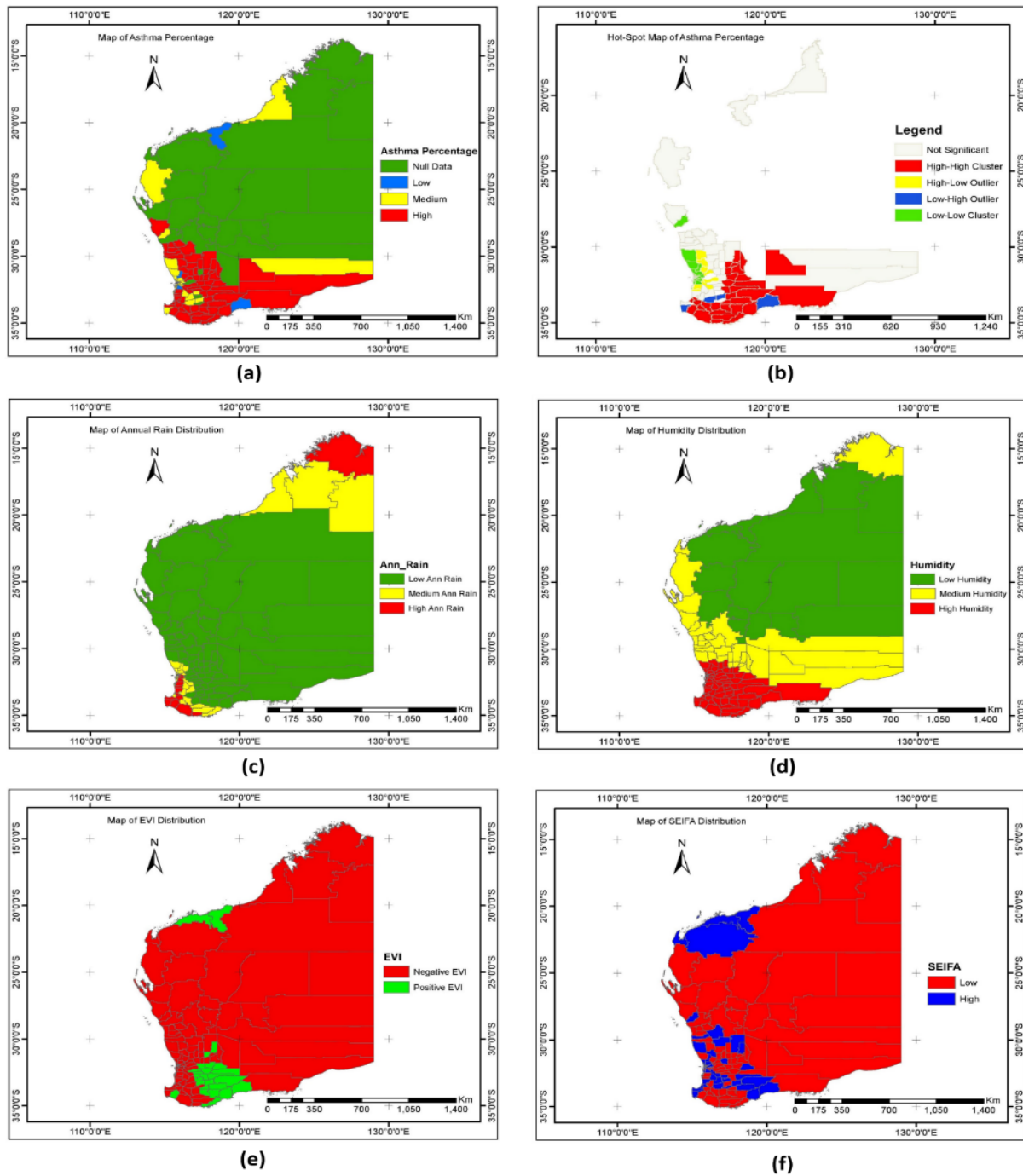


Figure 2. (a) Map of Asthma Percentage, (b) Hot-Spot Map of Asthma Percentage, (c) Map of Annual Rain Distribution, (d) Map of Humidity Distribution, (e) Map of EVI Distribution, (f) Map of SEIFA Distribution.

Table 2. The Summary of Descriptive Analysis

Measures	Variables				
	Asthma Percentage (%)	Humidity (%)	Annual Rain (mm)	EVI	SEIFA
Mean	9.973	62.803	540.609	-0.00314	981.543
Median	10.356	65.020	444.470	-0.00313	991.150
Max	12.477	75.572	1116.060	0.01025	1126.490
Min	5.744	38.730	246.777	-0.02109	597.616
Sum	1107.019	8666.858	74604.071	-0.43322	135452.977
Std Deviation	1.454	8.404	227.652	0.00474	85.802



Table 3. Moran's I autocorrelation result

Variable	Moran's I	Z-value	P-value	Interpretation
Asthma Percentage	0.395	15.681	0	Clustered
Annual Rain	0.350	13.881	0	Clustered
Humidity	0.552	22.207	0	Clustered
EVI	0.230	9.372	0	Clustered
SEIFA	0.314	12.589	0	Clustered

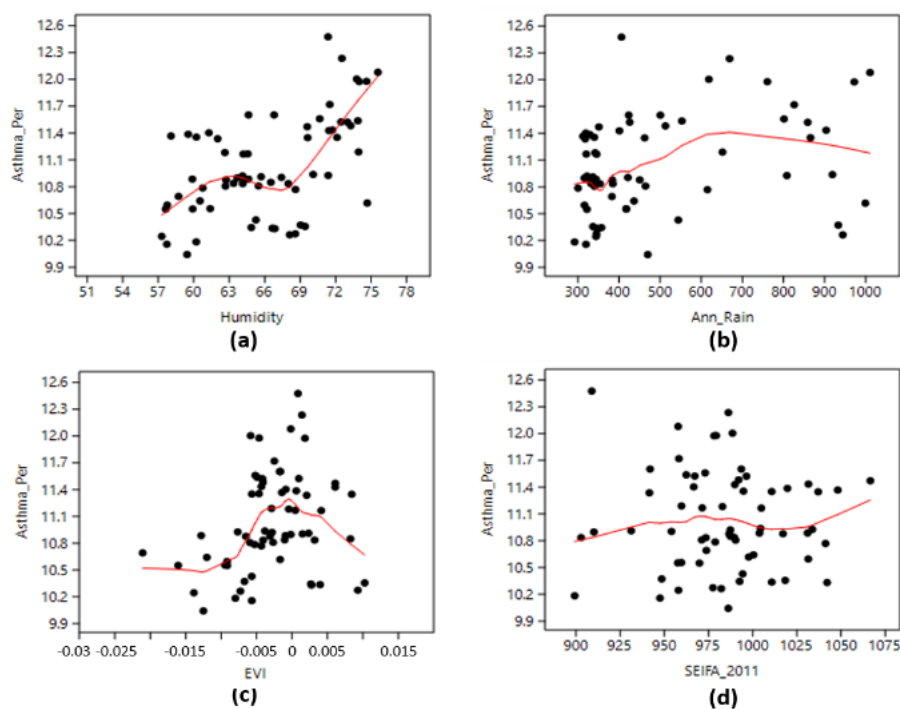


Figure 3. Scatter plot of Asthma percentage in LOESS with the explanatory variables (a) Humidity, (b) Annual rain, (c) EVI, (d) SEIFA 2011

The result of spatial autocorrelation Moran's I shows that all variables indicate a robust global clustering. This result also confirms the choropleth maps in figure 2 that all of the variables have a cluster pattern. The hotspot map of the asthma percentage (figure 2(b)) depicts that the high-high (hotspot) values of the asthma is located in the south part of the region and has high-low and low-high values in the surrounding areas. The hotspot area also has a similar value of the choropleth map of asthma percentage (figure 2(a)). It means that those areas need to be investigated further related to other explanatory variables in the confirmatory analysis. From those results, it can be concluded as a hypothesis that the main factors that influence of asthma prevalence are the high level of humidity, high level of annual rain, the positive value of EVI and low level of SEIFA index.

The map of the high asthma percentage region before and after classification can be seen in figure 5 (a) and (b). The figure 5 show maps as a result from the intersection of each variable and depict the distribution of variables in particular area, which is the high percentage asthma regions. The detail of each portion variable can be seen in table 4.

Table 4. Proportion of Each Variable in High Asthma Percentage Region

Variables	Percentage (%)
SEIFA low	79.73
SEIFA high	20.27
EVI plus	20.03
EVI minus	79.97
Rain low	88.14
Rain medium	6.25
Rain high	5.61
Humidity medium	53.44
Humidity high	46.56

The correlations between the dependent variable (Asthma Percentage) and each of explanatory variable are essential and have been discussed previously. The correlation among each explanatory variable also can reveal the relation among these variables, particularly between humidity, annual rain and EVI since these variables are related with natural phenomenon and

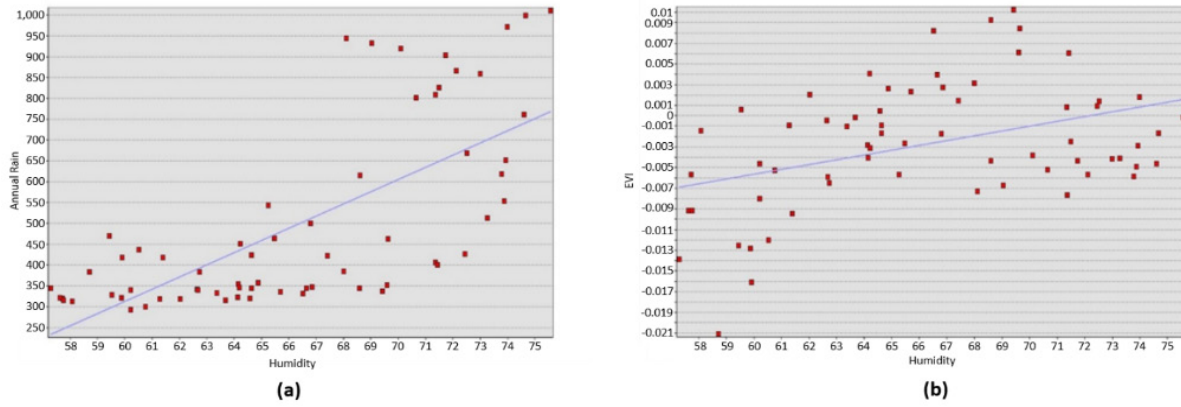


Figure 4. Scatterplot Matrix Graph of a) Humidity and Annual Rain, b) Humidity and EVI

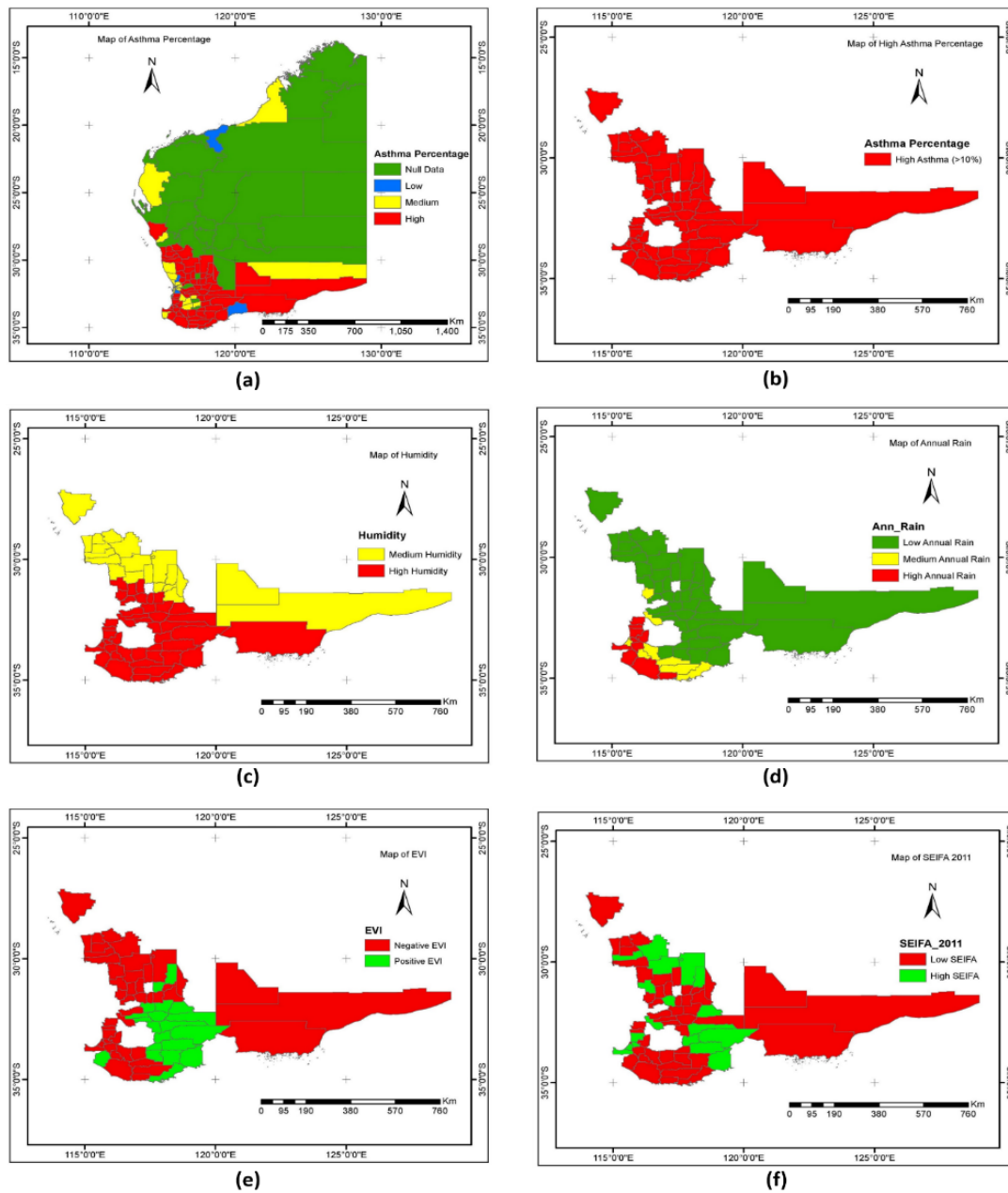


Figure 5. (a) Map of Asthma Percentage before classification, (b) Map of High Asthma Percentage after classification, (c) Map of Humidity in High Asthma Region, (d) Map of Annual Rain in High Asthma Region, (e) Map of EVI in High Asthma Region, (f) Map of SEIFA 2011 in High Asthma Region

possibly are related to each other. The correlation between each variable in the region can be seen using the scatterplot matrix graph and can be seen in figure 4.

Figure 4 depicts the correlation between humidity and annual rain also EVI. It can be seen that those three variables have positive correlation, although positive EVI and high annual rain only contribute to less effect to the asthma prevalence.

The result of grouping and weighting analysis can be seen in figure 6. It can be noticed that location of weighted mean values from each of the variables are closed to each other (with only EVI location that is separated 124 km from the other locations). This result shows that all of the variables are connected and give the strong correlation to the high percentage of asthma.

The maps of grouping analysis (figure 6) depict that the areas with high asthma percentage and high variables (humidity, annual rain, EVI) and low SEIFA variable and show the more obvious impacts of the areas where the explanatory variables contribute significantly more than in the previous map of intersection variables within the high asthma areas. Those maps have a similar pattern to each other and the most risky areas are located in the west south and central south of the region (figure 6). Those areas also experienced the most definite influence of the contributing factors. In addition, the overlay locations of the weighted mean of each variable (figure 6(f)) show that although those locations are located close to each other, those locations are located outside the most risky areas of asthma.

This study investigates the underlying factors of the prevalence of asthma in Western Australia using exploratory data analysis. The visualization and correlation analysis showed that the high level of humidity, high level of annual rain, the positive value of EVI and low level of SEIFA index contributes to the asthma prevalence in Western Australia. The 2011 SEIFA survey showed that the Goldfields region in WA where has an average SEIFA score 826.4 (low SEIFA) reported 10% of Asthma diagnosis [Health Service, 2016], the Pilbara region in WA where has an average SEIFA score 1,028 (high SEIFA) reported had only 8.8% of asthma diagnosis [Health Service, 2015a] and the Wheatbelt region in WA where has an average score of 925.8 (low SEIFA) reported had 10% of asthma diagnosis [Health Service, 2015b]. Those facts show that the low SEIFA has a correlation with the higher percentage of asthma diagnosis.

The comparison of each variable in the high asthma percentage and low asthma percentage can be used to find the key variable to reduce the prevalence of asthma. Both regions (low asthma percentage and high asthma percentage) have a low level of annual rain, which means that rain is not the primary factor in asthma prevalence. However, the significant differences between these two regions are the low asthma regions have a more significant proportion of positive EVI and low humidity percentage than the high asthma regions.

It is noticed that in the low asthma regions there is no region that has low SEIFA index, compared with the high asthma regions that have 79.73% of low SEIFA index. It means that the low SEIFA index has a high correlation with the asthma prevalence.

Based on the data analysis, the three main factors that can lead to a low percentage of asthma in the regions are the positive index of EVI, low humidity and high SEIFA index. It can be seen that the humidity (both high and medium) has almost equal proportion in the region (46.56% and 53.44% respectively). It means that the high percentage of asthma regions are influenced by more 57% humidity. The negative regions of EVI give more influence than the positive regions with 79.97% and 20.03% respectively. This fact has an inverse relationship with the hypothesis, which states that positive regions of EVI will give the higher influence to asthma. In contrast, the only small portion of high annual rain regions (5.61%) where are located in the high asthma percentage regions and the low annual rain dominated the high asthma regions (88.14%). This result also does not confirm the hypothesis (the high level of annual rain will give more effect to asthma). In addition, 79.73% of the high asthma percentage was dominated by the low index of SEIFA and only 20.27% that contains the high index of SEIFA. This result confirmed the hypothesis that the lower SEIFA correlates with lower health status and high-risk factors of mortality. Erdman et al. [2015] suggested that the evergreen vegetation plantation in an urban area could reduce the number of asthma cases in the older adults. Andrusaityte et al. [2016] also explained that the insignificantly increased risk of asthma in young people was caused by the green vegetation in the environment. In addition, another key in managing and relieving asthma symptoms is by controlling the humidity levels [Compact Appliance, 2015]. Koster [2016] explained that humidity that can trigger dust mites and mould, which is one of the leading causes of asthma, could be controlled by using equipment such as refrigerated air conditioners, dehumidifiers and humidifiers. Regarding SEIFA index, the high percentage of asthma regions have a low SEIFA index, means that those regions are more disadvantage (poverty) than the regions that have high a SEIFA index. Jones, Lawson, Robson, Buchanan, & Aldrich [2004] argue that poverty as one of the factors that can increase asthma prevalence and worse results since it can lead to the lack of health-care access and management. For that reasons, the policies regarding green plantation, controlling humidity and increasing the SEIFA index need to be initiated by the government.

Although the grouping analysis shows a different pattern with the weighted analysis, however, it seems that the result of grouping analysis is more reliable than the latter method. The result of grouping analysis reveals that the areas where have the high asthma percentage and were influenced significantly by the explanatory variables are located in the west south and central south



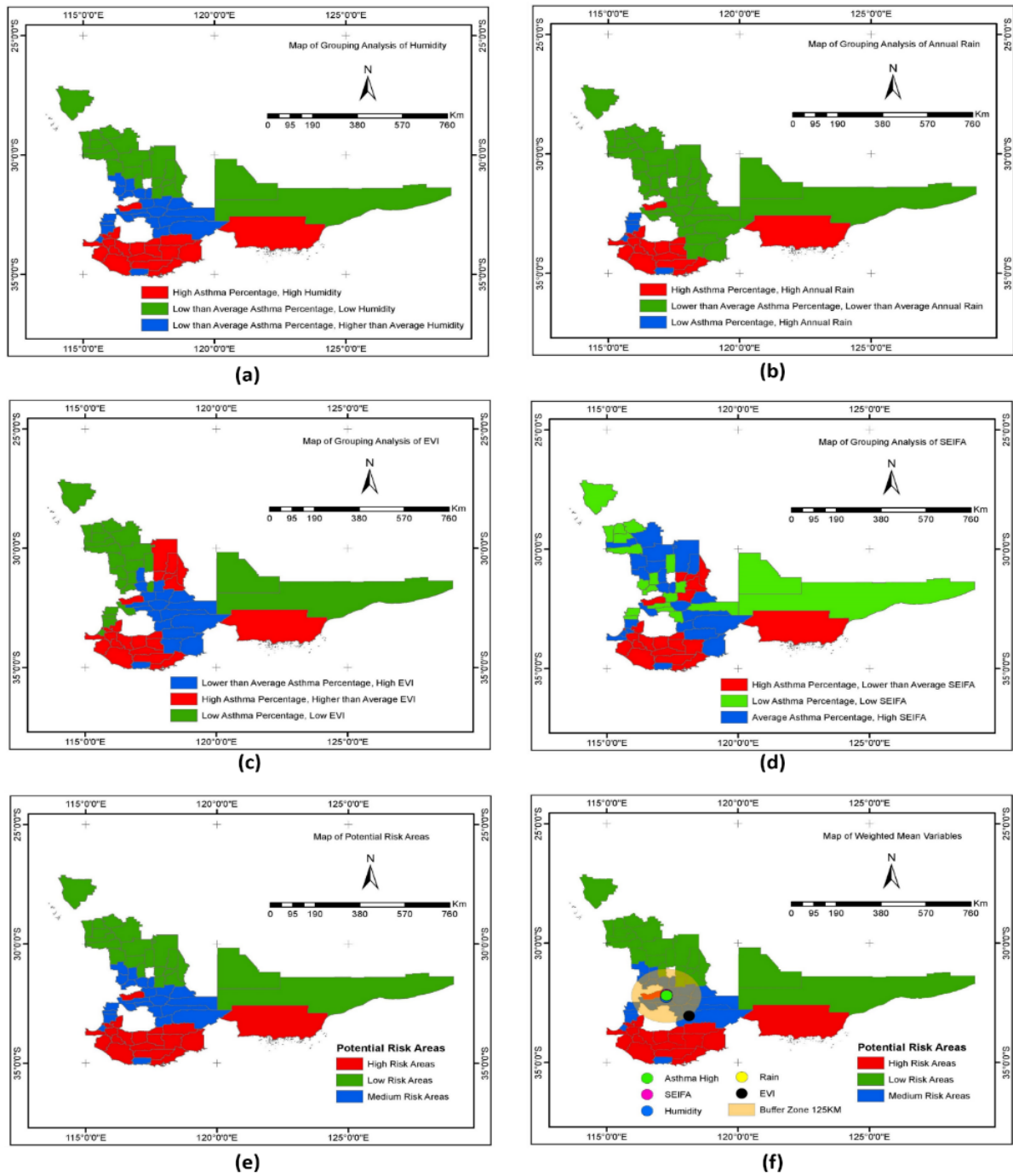


Figure 6. (a) Map of Grouping Analysis of Humidity, (b) Map of Grouping Analysis of Annual Rain, (c) Map of Grouping Analysis of EVI, (d) Map of Grouping Analysis of SEIFA, (e) Map of Potential Risk Areas, (f) Map of Weighted Mean Variables

of Western Australia.

In this study, the Perth City as a metropolitan and major urban centre of WA has a medium of asthma percentage (8 - 10 %). However, considering that the city only has a small proportion of the entire WA region, it can be assumed that the Perth City also experienced a significant prevalence of asthma. Pereira, De Vos, & Cook [2009b] noticed that the SEIFA and air pollution were two significant factors in the increasing amount of asthma in the Perth region. Pereira, De Vos, & Cook [2009b] found that the higher Socio-Economic

Status (SES) experienced more asthma cases and they argued that the low SEIFA regions be associated with the less risk of asthma. These results contradict with the result of this study and need a further investigation. In addition, Pereira et al. [2010] found that air pollution such as O<sub>3</sub>, PM<sub>10</sub>, NO<sub>2</sub> and CO were associated with the asthma prevalence in the Perth region during the 2002 to 2009 period. Finally, although the explanatory spatial data analysis can be a useful method to reveal the pattern and underlying factors of asthma, the regression analysis needs to be conducted in the further

study to get the significance of those factors in the asthma prevalence.

#### 4. Conclusion

The study tries to investigate the significant causal factors of asthma prevalence in Western Australia using spatial analysis approach. Many provided variables have been tested to get the best correlation with the asthma percentage variable and four variables (humidity, annual rainfall, EVI and SEIFA) were chosen and tested with high asthma percentage variable. The exploratory data analysis including data visualization using choropleth maps, correlation analysis, Moran's I spatial pattern analysis and hotspot analysis have been conducted to get the hypothesis of the factors that contribute to the asthma prevalence. The hypothesis as the first guidance of the analysis has been concluded with the high level of humidity, high level of annual rain, the positive value of EVI and low level of SEIFA index are the main factors that influence of asthma prevalence. The confirmatory data analysis uses maps intersect, grouping analysis and weighted analysis. The result of grouping analysis reveals that the most risky areas of asthma are located in the west south and central south of Western Australia. The low SEIFA has a strong correlation with the higher percentage of asthma diagnosis. Although humidity did not contribute significantly to the asthma prevalence, the negative regions of EVI and low annual rain give more influence to the disease. These results reject the hypothesis and indicate the need of further study to explain that.

#### 5. Acknowledgement

The authors would like to acknowledge the Department of Spatial Science, Curtin University, Perth, Western Australia for the provision of the spatial data. The authors would also like to acknowledge Associate Professor Cecilia Xia and Roula Zougheibe for their useful ideas and comments.

#### References

- Anderson, H. R., Butland, B. K., Donkelaar, A. Van, Brauer, M., Strachan, D. P., & Clayton, T. (2012). Research | Children's Health Satellite-based Estimates of Ambient Air Pollution and Global Variations in Childhood Asthma Prevalence. *Environ Health Perspect*, 120(9), 1333–1339. <https://doi.org/10.1289/ehp.1104724>
- Andrusaityte, S., Grazuleviciene, R., Kudzyte, J., Bernotiene, A., Dedele, A., & Nieuwenhuijsen, M. J. (2016). Associations between neighbourhood greenness and asthma in preschool children in Kaunas, Lithuania: a case-control study. *BMJ Open*, 6(4), e010341. <https://doi.org/10.1136/bmjopen-2015-010341>.
- ArcMap, D. (2017). Using weights. Retrieved from <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/using-weights.htm>.
- Australia Wide First Aid. (2018). Thunderstorm Asthma. Retrieved October 27, 2017, from <https://www.australiawidefirstaid.com.au/thunderstorm-asthma/>.
- Black, M. H., Anderson, A., Bell, R. A., Dabelea, D., Pihoker, C., Saydah, S., ... Lawrence, J. M. (2011). Prevalence of Asthma and Its Association With Glycemic Control Among Youth With Diabetes. *Pediatrics*, 128(4), e839–e847. <https://doi.org/10.1542/peds.2010-3636>.
- Bottrell, J. (2011). Here's Why Humidity and Cold Air Trigger Asthma. Retrieved from <https://www.healthcentral.com/article/heres-why-humidity-and-cold-air-trigger-asthma>.
- Chien, L.-C., & Alamgir, H. (2014). Geographic disparities of asthma prevalence in south-western United States of America. *Geospatial Health*, 9(1), 97. <https://doi.org/10.4081/gh.2014.8>.
- Collison, A., Herbert, C., Siegle, J. S., Mattes, J., Foster, P. S., & Kumar, R. K. (2011). Altered expression of microRNA in the airway wall in chronic asthma: miR-126 as a potential therapeutic target. *BMC Pulmonary Medicine*, 11. <https://doi.org/10.1186/1471-2466-11-29>.
- Compact Appliance. (2015). Breathe Easy: Why Controlling Humidity Can Help Reduce Asthma Symptoms. Retrieved October 25, 2017, from <https://learn.compactappliance.com/humidity-and-asthma/>.
- Cook, A. G., Devos, A. J. B. M., Pereira, G., Jardine, A., & Weinstein, P. (2011). Use of a total traffic count metric to investigate the impact of roadways on asthma severity: A case-control study. *Environmental Health: A Global Access Science Source*, 10(1), 1–8. <https://doi.org/10.1186/1476-069X-10-52>.
- Crighton, E. J., Feng, J., Gershon, A., Guan, J., & To, T. (2012). A spatial analysis of asthma prevalence in Ontario. *Canadian Journal of Public Health*, 103(5), 384–389.
- Cunningham, J. (2010). Socioeconomic status and self-reported asthma in Indigenous and non-Indigenous Australian adults aged 18-64 years: analysis of national survey data, 1–11.
- Davinci.asu.edu. (2013). evi. Retrieved from <http://davinci.asu.edu/index.php?title=evi>
- Delclòs-alió, X., Miralles-guasch, C., & Jacobs, J. (2018). Land Use Policy Looking at Barcelona through Jane Jacobs's eyes: Mapping the basic conditions for urban vitality in a Mediterranean conurbation. *Land Use Policy*, 75(April), 505–517. <https://doi.org/10.1016/j.landusepol.2018.04.026>.
- Erdman, E., Liss, A., Gute, D. M., Rioux, C., Koch, M., & Naumova, E. N. (2015). Does the presence of vegetation affect asthma hospitalizations among the elderly? A comparison between rural, suburban ..., (May). <https://doi.org/10.24102/ijes.v4i1.526>.
- Gorai, A. K., Tchounwou, P. B., & Tuluri, F. (2016).

- Association between ambient air pollution and asthma prevalence in different population groups residing in Eastern Texas, USA. *International Journal of Environmental Research and Public Health*, 13(4). <https://doi.org/10.3390/ijerph13040378>
- Górniak, J. (2016). The Spatial Autocorrelation Analysis for Transport Accessibility in Selected Regions of the European Union. *Comparative Economic Research*, 19(5), 25–42. <https://doi.org/10.1515/cer-2016-0036>.
- Health Service, G. W. A. (2015a). Pilbara – population and health snapshot, 1–6. Retrieved from <http://www.wapha.org.au/wp-content/uploads/2015/12/Regional-Profile-2016-Pilbara-population-and-health-snapshot-FINAL.pdf>
- Health Service, G. W. A. (2015b). Wheatbelt – population and health snapshot astern Primary Health Service. Retrieved from <http://www.wapha.org.au/wp-content/uploads/2016/01/Wheatbelt-snapshot-2016.pdf>
- Health Service, G. W. A. (2016). Goldfields – population and health snapshot. Retrieved from <http://www.wapha.org.au/wp-content/uploads/2015/12/Regional-Profile-2016-Goldfields-population-and-health-snapshot-FINAL-002.pdf>
- Hekking, P. P. W., Wener, R. R., Amelink, M., Zwinderman, A. H., Bouvy, M. L., & Bel, E. H. (2015). The prevalence of severe refractory asthma. *Journal of Allergy and Clinical Immunology*, 135(4), 896–902. <https://doi.org/10.1016/j.jaci.2014.08.042>.
- Jiang, Z., & Huete, A. R. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10), 3833–3845.
- Jones, V. F., Lawson, P., Robson, G., Buchanan, B., & Aldrich, T. (2004). The Use of Spatial Statistics to Identify Asthma Risk Factors in an Urban Community. *Pediatric Asthma, Allergy and Immunology*, 17(1), 3–13. <https://doi.org/http://dx.doi.org/10.1089/088318704322994895>
- Kececi, M. C. (2017). Monitoring Pollen Counts and Pollen Allergy Index Using Satellite Observations in East Coast of the United States. South Dakota State University.
- Koster, L. (2016). Indoor humidity and your family's health. Retrieved from <https://www.nationalasthma.org.au/news/2016/indoor-humidity>
- Krstić, G. (2011). Asthma prevalence associated with geographical latitude and regional insolation in the United States of America and Australia. *PLoS ONE*, 6(4). <https://doi.org/10.1371/journal.pone.0018492>.
- Lomba, A., Strohbach, M., Jerrentrup, J. S., Dauber, J., Klimek, S., & Mccracken, D. I. (2017). Making the best of both worlds : Can high-resolution agricultural administrative data support the assessment of High Nature Value farmlands across Europe? *Ecological Indicators*, 72, 118–130. <https://doi.org/10.1016/j.ecolind.2016.08.008>.
- Lopes, S. B., São-carlense, A. T., Carlos, S., & São-carlense, A. T. (2007). Exploratory and Confirmatory Spatial Data Analysis Tools in Transport Demand Modeling. *Computers in Urban Planning and Urban Management*.
- Mainardi, T. R., Mellins, R. B., Miller, R. L., Acosta, L. M., Cornell, A., Hoepner, L., ... Perzanowski, M. S. (2013). Exercise-Induced Wheeze, Urgent Medical Visits, and Neighborhood Asthma Prevalence. *Pediatrics*, 131(1), e127–e135. <https://doi.org/10.1542/peds.2012-1072>.
- Mayo Clinic, S. (2016). Asthma Causes. Retrieved from <http://www.mayoclinic.org/diseases-conditions/asthma/basics/causes/con-20026992>
- Pereira, G., De Vos, A. J. B. M., & Cook, A. (2009a). Residential traffic exposure and children's emergency department presentation for asthma: A spatial study. *International Journal of Health Geographics*, 8, 1–10. <https://doi.org/10.1186/1476-072X-8-63>.
- Pereira, G., De Vos, A. J. B. M., & Cook, A. (2009b). Residential traffic exposure and children's emergency department presentation for asthma: A spatial study. *International Journal of Health Geographics*, 8, 1–11. <https://doi.org/10.1186/1476-072X-8-63>.
- Pereira, G., De Vos, A. J. B. M., Cook, A., & D'Arcy, C. (2010). Vector fields of risk: A new approach to the geographical representation of childhood asthma. *Health and Place*, 16(1), 140–146. <https://doi.org/10.1016/j.healthplace.2009.09.006>.
- Pereira, G., Vos, A. De, & Cook, A. (2009). Removing the Effects of Socioeconomic Status on the Geographic Distribution of Asthma Events.
- Perry, M. (2017). Asthma and Humidity. Retrieved from <http://www.healthguidance.org/entry/10912/1/Asthma-and-Humidity.html>
- Pisati, M. (2012). Exploratory spatial data analysis using Stata 2012 German Stata Users Group meeting Spatial data Overview Dot maps Diagram maps Choropleth maps Multivariate maps, 1–91.
- Reddel, H. K., Sawyer, S. M., & Everett, P. W. (2015). Asthma control in Australia: a cross-sectional web-based survey in a nationally representative population, 202(May), 492–498. <https://doi.org/10.5694/mja14.01564>.
- Ronchi, S., Salata, S., Arcidiacono, A., & Ronchi, S. (2018). An indicator of urban morphology for landscape planning in Lombardy ( Italy ). *Management of Environmental Quality: An International Journal*. <https://doi.org/10.1108/MEQ-05-2017-0048>.
- Rousta, I., Doostkamian, M., Haghighi, E., Ghafarian Malamiri, H. R., & Yarahmadi, P. (2017). Analysis

- of spatial autocorrelation patterns of heavy and super-heavy rainfall in Iran. *Advances in Atmospheric Sciences*, 34(9), 1069–1081. <https://doi.org/10.1007/s00376-017-6227-y>.
- Shankardass, K., Jerrett, M., Dell, S. D., Foty, R., & Stieb, D. (2015). Spatial analysis of exposure to traffic-related air pollution at birth and childhood atopic asthma in Toronto, Ontario. *Health and Place*, 34(2), 287–295. <https://doi.org/10.1016/j.healthplace.2015.06.001>.
- Skarkova, P., Kadlubiec, R., Fischer, M., Kratenova, J., Zapletal, M., & Vrabel, J. (2015). Refining of Asthma Prevalence Spatial Distribution and Visualization of Outdoor Environment Factors Using Gis and Its Application for Identification of Mutual Associations. *Central European Journal of Public Health*, 23(3), 258–266. <https://doi.org/10.21101/cejph.a4193>.
- Song, J., Du, S., Feng, X., & Guo, L. (2014). Landscape and Urban Planning The relationships between landscape compositions and land surface temperature: Quantifying their resolution sensitivity with spatial regression models. *Landscape and Urban Planning*, 123, 145–157. <https://doi.org/10.1016/j.landurbplan.2013.11.014>.
- Speldewinde, P. C., Cook, A., Davies, P., & Weinstein, P. (2011). The hidden health burden of environmental degradation: Disease comorbidities and dryland salinity. *EcoHealth*, 8(1), 82–92. <https://doi.org/10.1007/s10393-011-0686-x>.
- Tattersfield, A. E., Knox, A. J., Britton, J. R., & Hall, I. P. (2002). *Asthma*, 360, 1313–1322.
- To, T., Stanojevic, S., Moores, G., Gershon, A. S., Bateman, E. D., Cruz, A. A., & Boulet, L. P. (2012). Global asthma prevalence in adults: Findings from the cross-sectional world health survey. *BMC Public Health*, 12(1), 204. <https://doi.org/10.1186/1471-2458-12-204>.
- Venkatramanan, S., Chung, S. Y., Rajesh, R., & Lee, S. Y. (2015). Comprehensive studies of hydrogeochemical processes and quality status of groundwater with tools of cluster , grouping analysis , and fuzzy set method using GIS platform : a case study of Dalcheon in Ulsan City , Korea. *Environmental Science and Pollution Research*, 22(15), 11209–11223. <https://doi.org/10.1007/s11356-015-4290-4>.
- Zhou, J., Tu, Y., Chen, Y., & Wang, H. (2017). Estimating Spatial Autocorrelation With Sampled Network Data. *Journal of Business & Economic Statistics*, 35(1), 130–138. <https://doi.org/10.1080/07350015.2015.1061437>.