Ph.D. Dissertation of Education

Exploring Spatial Variations in the Influences of Health-threatening Anthropogenic Factors on Human Health: An ESDA Approach to Remote Sensing Data

인류발생적 건강위협 인자가 인체건강에 미치는 영향력의 공간적 변동 탐색: 원격탐사 데이터에 대한 ESDA 적 접근

February 2019

Graduate School of Education
Seoul National University
Social Studies Education Department
Geography Major

Lei Zhu

Exploring Spatial Variations in the Influences of Health-threatening Anthropogenic Factors on Human Health: An ESDA Approach to Remote Sensing Data

Advised by Prof. Sang-Il Lee

Submitting a Ph.D. Dissertation of Education

November 2018

Graduate School of Education
Seoul National University
Social Studies Education Department
Geography Major

Lei Zhu

Confirming the Ph.D. Dissertation written by
Lei Zhu
December 2018

Chair _____ (Seal)
Vice Chair _____ (Seal)
Examiner _____ (Seal)
Examiner _____ (Seal)
Examiner _____ (Seal)

Abstract

Exploring Spatial Variations in the Influences of Health-threatening Anthropogenic Factors on Human Health: An ESDA Approach to Remote Sensing Data

Lei Zhu

Dept. of Social Studies Education

Geography Major

The Graduate School

Seoul National University

The distribution pattern of disease and health care delivery are the two major strands of health geography. A great deal of resources and geographic analytical techniques have been utilized to conduct research on these two principal concerns. Due to the abundant information and the wide spatiotemporal scale, remote sensing has been widely applied in the studies of health geography since the 1970s. By using remotely sensed data, the main objective of this study is to explore spatial variations in the influences of health-threatening anthropogenic factors on human health under the exploratory spatial data analysis (ESDA) framework.

Diverse remote sensing data from Landsat 8 and the Visible Infrared Imaging Radiometer Suite (VIIRS) have been utilized in this study to extract potential health-threatening anthropogenic. In these two examples, the man-made impervious surface (IMS) and the exposure to artificial nighttime light (NTL) have been extracted as the most representative

potential health-threatening anthropogenic factors. Meanwhile, the discomfort index (DI) and non-accident mortality as well as the prevalence of breast cancer were considered as potential public health outcomes influenced by those factors.

Based on this, the distributive characteristics of potential health-threatening anthropogenic factors as well as their possible health outcomes and the spatial co-pattern detection between the two variables have been firstly analyzed. Results show that the potential health-threatening anthropogenic factors have a tendency to concentrate in urban area, while their possible health outcomes have various spatial distributive features. These two variables tend to have a positive spatial association. Next, spatial variations in the influences of health-threatening anthropogenic factors on human health has been estimated by spatial regression analysis. The results indicate that man-made IMS and NTL are the main influencing factors for the increase of the DI value and the increase of the breast cancer prevalence rate.

Keywords: health-threatening anthropogenic factors, human health, remote sensing, spatial variation, spatial association measure, health geography

Student Number: 2015-30770

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor professor Sang-Il Lee. He is an admirable scholar, and after four years study with him, I not only get the latest knowledge, also I acquire many "know-how" to do researches. I can see the light in his eyes when he talks about researches and I hope I can be a scholar like him in the future. He is a kind friend, he is always encouraging us to think big and think difference and to be better us, and he is also a good listener when we need him. Secondly, I would like to thank the rest of my dissertation committee members, Professor Jungyeop Shin, Professor Daeheon Cho, Professor Gunhak Lee, and Professor Jongmin Byun, for their comments on my dissertation and the guidance they provided. This work would not have been accomplished without their invaluable advice and counsel.

Also, I would like to express my thanks to everyone in GIS labs, thanks to whom, I received much happiness during my Ph.D. study. They are like families to me and always give a hand to me when I need help, encourage me when I feel down, and also take care of me when I feel sick. The thankful feelings are beyond expression, thanks to my Korean families.

I also would like to express my thankful heart to the friends in Geography education. I feel so lucky to meet all of you and I am very grateful for the help and accompany of yours. I can't imagine what my life will be without out your kindness and assistance.

Many thanks to my lifelong friends, they give me encouragements both on the academic and mental level. Thanks for accompanying and the power they give me when I need them.

Last but not the least, I would like to express my deepest gratitude to the most important people in my life, my parents, Yongxun Zhu and Yanli Zhang, for the pure love and support, and my little brother, Yunlong Zhu, for the happiness and support he gave me.

Contents

Abstract ..... **i**

Chapter 1. Introduction **1**

1.1 General backgrounds 1

1.2 The main objectives 3

1.3 The outline of this study 5

Chapter 2. Theoretical Underpinnings and Research Design **9**

2.1 Health geography, remote sensing, and spatial statistics 9

2.1.1 Health geography 9

2.1.2 Remote sensing techniques in health geography 11

2.1.3 Spatial statistical approach to health geography 14

2.2 Main concepts in this study 18

2.2.1 Anthropogenic factors 18

2.2.2 Exploratory spatial data analysis 20

2.3 Research Design 28

2.3.1 Health-threatening anthropogenic factors 28

2.3.2 Possible health outcomes 30

Chapter 3. Man-made Impervious Surface and Its Potential Heat-related Threats to Public Health 32

3.1 Introduction 32

3.2 Materials and Methods 35

3.2.1 Data 35

3.2.2 Methods 40

3.3 Analytical results 46

3.3.1 Spatiotemporal distribution of man-made IMS 46

3.3.2 Spatial distribution of health-threatening risk 51

3.3.3 Detection of the co-pattern between man-made IMS with DI value and the non-accident mortality 57

3.4 Discussions and conclusions 64

Chapter 4. Exposure to Nighttime Light and Its Potential Hormone-related Threats to Public Health 68

4.1 Introduction 68

4.2 Materials and Methods 71

4.2.1 Data 71

4.2.2 Methods 76

4.3 Analytical results 77

4.3.1 Human exposure to NTL 77

4.3.2 Spatiotemporal distribution of the breast cancer prevalence rate 80
4.3.3 Detection of the co-pattern between human exposure to NTL and the prevalence of breast cancer 84
4.4 Discussions and conclusions 92

Chapter 5. The Influences of Health-threatening Factors on Human Health 96
5.1 Introduction 96
5.2 The influence of man-made IMS on the DI value 96
5.2.1 Model description 97
5.2.2 Analytical results 98
5.3 The influence of man-made IMS on the non-accident mortality .101
5.3.1 Model description 101
5.3.2 Analytical results 102
5.4 The influence of NTL on the breast cancer prevalence 105
5.4.1 Model description 105
5.4.2 Analytical results 106
5.5 Discussions and conclusions 109

Chapter 6. General Conclusions and Discussions 112

Bibliography **117**

Appendix **136**

Appendix 1: The distribution of regression coefficients in the model on the DI value 136

Appendix 2: The distribution of regression coefficients in the model on the non-accident mortality 137

Appendix 3: The distribution of regression coefficients in the model on the breast cancer prevalence rate (significant coefficients) 138

국문초록 **139**

中文摘要 **141**

List of Figures

Figure 3-1 The description of object-based feature extraction 41

Figure 3-2 The process of object-based segmentation 43

Figure 3-3 The spatial distribution of the man-made IMS 49

Figure 3-4 The spatial distribution of the man-made IMS in *Dong* unit 50

Figure 3-5 The spatial distribution of the DI value in 1Km by 1Km grid unit 52

Figure 3-6 The spatial distribution of the DI value in *Dong* unit 52

Figure 3-7 The spatial distribution of the non-accident mortality 56

Figure 3-8 The scatterplot of local L_i between the man-made IMS and the DI value 59

Figure 3-9 The spatial distribution of local Li between the man-made IMS and the non-accident mortality concentration degree 62

Figure 4- 1 Average NTL, 2012~2015 73

Figure 4-2 Residential buildings in NTL 75

Figure 4-3 The spatial distribution of human exposure to NTL in *Si-Gun-Gu* unit 79

Figure 4-4 The spatial distribution of (a) the breast cancer prevalence rate in 2015; (b) the clusters of the breast cancer prevalence rate in 2015; and (c) the clusters of the breast cancer prevalence rate change from 2012 to 2015.... 81

Figure 4-5 The spatial distribution and the cluster distribution of local Pearson's r_i between human exposure to NTL and the breast cancer prevalence 87

Figure 4-6 The scatterplot of local Pearson's r_i between human exposure to NTL and the breast cancer prevalence 87

Figure 4-7 The spatial distribution and the cluster detection of local Pearson's $r_i$between human exposure to NTL and the breast cancer prevalence change 90

Figure 4-8 The scatterplot of Pearson's $r_i$between human exposure to NTL and the breast cancer prevalence change 90

Figure 5- 1 The spatial distribution of regression coefficients of the IMS ratio 100

Figure 5- 2 The spatial distribution of strongest influencing factors on the DI value 100

Figure 5- 3 The spatial distribution of regression coefficients of the IMS ratio 104

Figure 5- 4 The spatial distribution of the strongest influencing factors on the non-accident mortality 104

Figure 5- 5 The spatial distribution of regression coefficients of the mean_NTL 108

Figure 5- 6 Distribution of the strongest influencing factors on the breast cancer prevalence rate.....108

List of Tables

Table 2-1 Spatial analytical techniques 17

Table 3-1 Data 36

Table 3-2 Regions with the top and bottom 5 DI values 53

Table 3-3 Rank of the non-accident mortality 55

Table 3-4 Rank of the local L_i in different quadrants (X: ratio of man-made IMS; Y: DI) 60

Table 3-5 Rank of the local L_i in different quadrants 63

Table 4-1 Rank of human exposure to NTL (unit: nanoWatts/cm^2/sr) 78

Table 4-2 Rankings of the breast cancer prevalence rate 82

Table 4-3 Rankings of the breast cancer prevalence rate increase 82

Table 4-4 Rankings of the local ri in different quadrants (X: human exposure to NTL; Y: BCPR) 86

Table 4-5 Rankings of the local ri in different quadrants (X: human exposure to NTL; Y: BCPRI) 89

Table 5- 1 Data and sources (Influencing factors on the DI value) 98

Table 5- 2 Results of OLS and SEM (Influencing factors on the DI value) 99

Table 5- 3 Data and sources (Influencing factors on the non-accident mortality) 102

Table 5- 4 Results of OLS and SEM (Influencing factors on the non-accident mortality).....103

Table 5- 5 Data and sources (Influencing factors on the breast cancer prevalence rate) 106

Table 5- 6 Results of OLS and SEM (Influencing factors on the breast cancer prevalence rate) 107

Chapter 1. Introduction

1.1 General backgrounds

Health is one of the basic prerequisites of a productive entity or community, and the health of a nation is strongly associated with the nation's economy and living standard (Coombes, 1993). Since John Snow's pioneering research on the communication of cholera in the 19th century, placing diseases and health conditions within the geographical framework has long been an integrated practice of public health (Lipton *et al.*, 2009). Initially, studies on geography of health and disease remained a sub-field of medical and health service research, with being named geographic pathology, geographical epidemiology, geomedicine, to name a few. However, since 1940s, human geographers started to attend directly in disease and health studies by applying geographical disciplinary perspectives and developed *medical geography* as a recognized sub-discipline of human geography (Andrews and Moon, 2005).

Till now, there are two main tasks for medical geographers, one is mapping and modelling the spatial pattern of determinants and the diffusion of diseases, the other one is related to the location, distribution, accessibility and utilization of the health services (Kearns and Moon, 2002; Andrews and Moon, 2005; Gatrell and Elliott, 2014). Among which, researches on what are the distributive features of the diseases and what

are the relationships with other features can not only inform us the health inequalities among regions, present us the possible risk factors, and more essentially, assist us to make more accurate decisions and policy for human well-being.

In the studies on health geography, cartographic and statistical methods are usually adopted in the studies on disease incidence and spatial diffusion, while qualitative methods are often applied to study how people cope with their environment (Gatrell and Elliott, 2014). In 1970, Cline published an overview article "New Eyes for Epidemiologists: Aerial Photography and Other Remote Sensing Techniques" as the first extensive review of the epidemiological application of the remote sensed data (Cline, 1970). Whereafter, due to the launch of Landsat-1 about 50 years ago, an increasing number of health studies have utilized remotely sensed data to monitor, surveil, or map health risk factors (Cline, 1970; Beck *et al.*, 2000). Various remote sensors like Landsat's Multispectral Scanner (MSS) and Thematic Mapper (TM), the National Oceanic and Atmospheric Administration (NOAA)'s Advanced Very High Resolution Radiometer (AVHRR) and so on are used in the health geography studies (Thompson *et al.*, 1996; Clarke *et al.*, 1991; Beck *et al.*, 2000;). However, the main factors derived from remote sensed data are natural factors (Rogers and Randolph, 1991; Rodriguez *et al.*, 1996; Dister *et al.*, 1997), like vegetation cover, water bodies, soil wetness, to name a few, with little mention with the anthropogenic factors.

What's more, in the previous studies on the health geographical application of remotely sensed data adopted statistical methods such as correlation, regression analysis to study the risk factors derived from the remotely sensed data and the health outcomes (Linthicum *et al.*, 1987; Thompson *et al.*, 1996; Thomson *et al.*, 1997). But as Tobler (1970) proposed, everything is related to everything else, but near things are more related than distant things. The lack of the utilization of spatial statistical evaluation between the risk factors derived from remotely sensed data and the health outcomes is also a research gap to be solved.

Based on the previous studies, the primary objective of this study is to detect the spatial statistical association between the possible health-threatening anthropogenic factors derived from remote sensing and their possible health outcomes.

1.2 The main objectives

Based on the research background and the research limitations in the existing literatures proposed above, the objectives of this study are as follows:

1. Recently, similar with the terminology of "remote sensing", there are many similar succeeding researches termed as "social sensing (Liu *et al.*, 2015)", which means the big geospatial data on individual-level and the

related methods; and "human sensing (Lopez *et al.*, 2017)", which means sensing human beings with various methods like communication data and so on. Based on the previous studies, due to the difficult availability of health-related data and the merits of remote sensing on both large spatial and temporal scales, I strive to build the bridge between the health and the remote sensing on the aspect of anthropogenic factors, which I prefer to name as "health sensing". The expected results of "health sensing" are to quantify the health conditions through the remote sensing data, and thus take measures to lessen the possible health-threatening anthropogenic factors from both macro and micro scales.

2. For the reason that the studies on detecting anthropogenic factors, especially the health-threatening anthropogenic factors in remote sensing is still on its emerging stage, and also there is little related researches in Korea, the second objective of this study is to detect the possible health-threatening anthropogenic factors in Korea using remote sensing data. In this study, based on the previous researches, I chose "man-made IMS" and "exposure to NTL" as the possible health-threaten activities to detect in remote sensing data.

3. The third objective of this study is to detect the spatial and temporal distribution pattern of the possible health outcomes which are induced by the possible health-threatening anthropogenic factors. In accordance with the previous studies, I postulate that the possible threats to human health

is the heat-related threats for man-made IMS and the hormone disease for exposure to NTL (here, we use breast cancer).

4. The fourth objective of this study is to figure out the relation between the possible health-threatening anthropogenic factors in remote sensing and the possible health outcomes to human health. Besides the aspatial statistical methods, spatial statistical methods have been adopted to evaluate the spatial association between the variables.

5. The final objective, which is also the fundamental objective is to contribute to the improvement of the human health conditions by coming up with some suggestions to the policy makers and the government. With the spatial association results between possible health-threatening anthropogenic factors and their possible health, besides general suggestions on the global scale, some detailed local suggestions can also be proposed, which could be more practical and with great assistance.

1.3 The outline of this study

The study contains mainly five parts, and the flow map of this study is shown as Figure 1-1.

The first part is the introduction. In this part, I did a general introduction of this study, as well as the emerging questions and the objective of this study.

In the second part, I did the literature review and represented the theoretical background of this study.

The third and fourth chapters are two cases of the possible health-threatening anthropogenic factors derived in remote sensing and their possible threats to human health. For different cases, there are mainly four sections, including introduction, which introduce the research context, materials and methods used in the study, results, and conclusions and discussions.

In the third part, man-made IMS is seen as one of the possible health-threatening anthropogenic factors in remote sensing, and its possible threats to human health are represented by its spatial association with the discomfort index (DI) and non-accident mortality.

In the fourth part, NTL is seen as one of the possible health-threatening anthropogenic factors in remote sensing, and the breast cancer has been identifies as the possible threats to human health when excessive exposure to NTL.

In the fifth part, I design models to identify the influences of health-threatening anthropogenic factors on human health using both global and local regression models.

In each case chapter, I first explain the spatial distribution pattern of the possible health-threatening anthropogenic factors derived from

remote sensing, then characterize the spatial distribution of the possible health outcomes, and finally, I demonstrate the spatial association between two variables by quantifying the spatial statistical association between them.

The last chapter is the general conclusions and discussions. I gave a general conclusions, and some advantages as well as limitations in this study. What's more, I also propose the follow-up studies of this study which should be continued.

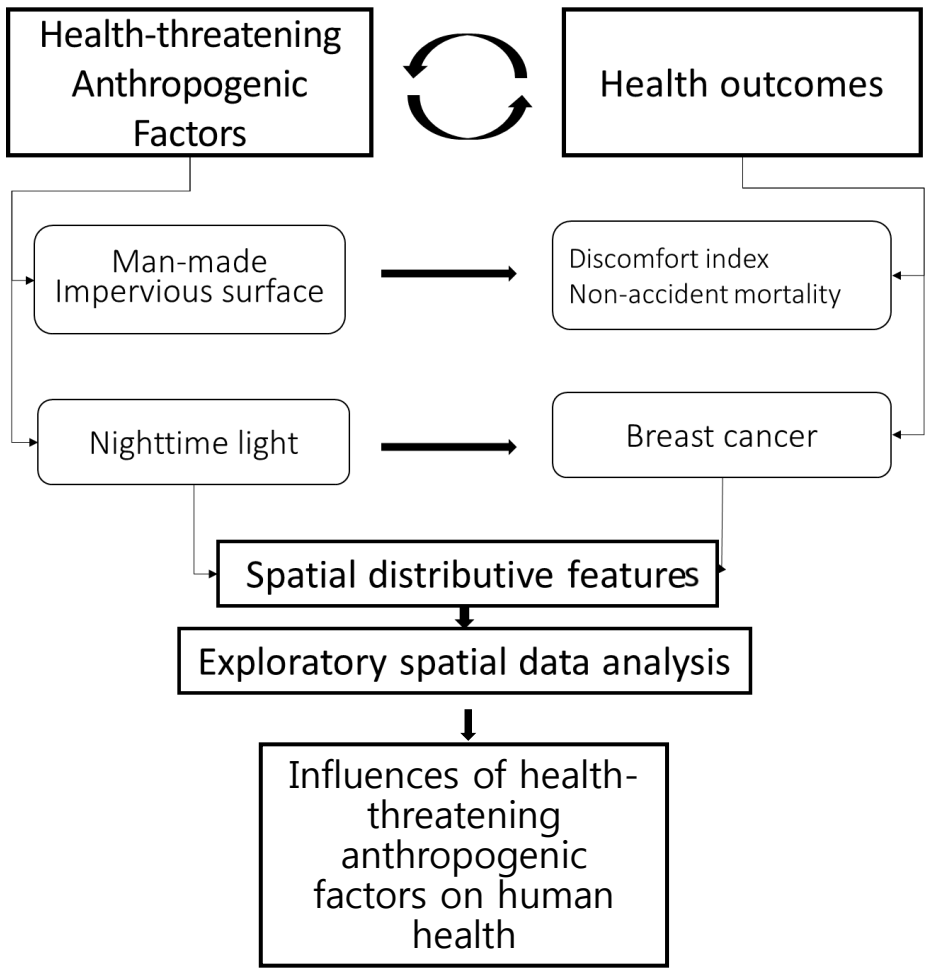


Figure 1-1 Research flow of this study

Chapter 2. Theoretical Underpinnings and Research Design

2.1 Health geography, remote sensing, and spatial statistics

2.1.1 Health geography

The conceptual origins of geographies of health can be traced to the ancient Greece as Hippocrates' "On airs, waters and places" (Hippocrates, 2000). According to Porta (2014), health geography is "a branch of science concerned with the spatial variations in environmental conditions related to health and disease".

In the 20th century, researches on geography of health and disease developed rapidly. At first, it remained a sub-field of medical and health research, variously named as "geographic pathology", "geomedicine", and "geographical epidemiology", to name a few. Whereas, by the 1940s, human geographers started to participate directly in disease and health researches, applying the geographical perspectives and developing "medical geography" as a recognizable sub-discipline of human geography (Andrews and Moon, 2005).

There are two traditions of medical geography, one is the health system planning, or so-called health care geography, another one is

geographical epidemiology, or the so-called disease geography (Mayer, 1982). Before 1960s, the medical geography has been dominated by the mapping of the disease (Gilbert, 1958; Howe, 1989; Brown *et al.*, 2009) . For example, Valentine Seaman's dot map of the yellow fever in New York in 1796 (Seaman, 1798), the well-known John Snow's map of the deaths distribution from cholera in the Broad Street district of London in 1854 (Snow, 1855), and Peterman's map of deaths from cholera in London in 1832 (Petermann, 1852). Beginning in the 1960s, quantitative analysis and newly developed theories began to be applied in the studies of medical geographies. For example, dislike the existing cholera studies , Pyle (1969) linked the society hierarchy with the cholera in the United States using the diffusion theory.

For some researchers, the utilization of "medical" and "health" geography remains considerable slippage, while the terms could be interchangeable in other cases (Brown *et al.*, 2009). However, for the staunch health geographers, health geography emerges since 1980s due to the shift from attention to medical world towards an increasing concern with wellbeing and "broader social models" of health and health care services.

Health geography is a sub-discipline of diversity and distinction. Both of the two traditional streams of medical geography, that is to say, disease geography and health care geography are continued in contemporary

health geography. While its attentiveness to new theories, its wide-ranging empirical concerns, and its diverse and flexible methodological approaches, establish a substantively "new" area of geographical researches.

2.1.2 Remote sensing techniques in health geography

Remote sensing is the process of collecting data of objects or landscape features without coming into directly contact with them (Albert *et al.*, 2003). The purpose of remote sensing is not only acquire information, but also the application of the acquired information. Human have used remotely sensed data to interpret the landscape even from the earliest hunter-gathers. Image data allows us to receive data that are difficult to receive by other means, such as measuring the sizes, areas, depths and heights, and seeing differences over time.

Since 1960s, the remote sensing images have been used in diverse research fields. The most significant reason for the remote sensing application on the geography of health is that the medical geography is "rooted in the idea that disease-causing microbes, or the infected insects and other creatures that transmit these microorganisms to people or animals, normally reside in identifiable environments" (Travis, 1997)., and remote sensing offers the capability to measure many characteristics of these environments, and record them (Cline, 1970).

Cline (1970) published an overview article "New Eyes for

Epidemiologists: Aerial Photography and Other Remote Sensing Techniques" as the first extensive review of the epidemiological application of the remote sensed data. In the following studies, the main factors they derived with remote sensing data are vegetation cover, landscape structure, and water bodies (Beck *et al.*, 2000).

The main application of remote sensing on health geography can be classified into two types in general: (1) remote sensing for disease vector research; (2) remote sensing for disease prevention.

The remote sensing for disease vector research may be the most widely application of remotely sensed data on the health geography. Mosquitos, trematodes, tsetse flies and ticks are the main disease vector studied in these researches. For instance, Rodriguez *et al.* (1996) analyzed the relationship between aerial photography derived land cover type, which may be the breeding and resting places of malaria mosquito, and some village factors like density of cattle and horses, with the mosquito abundance. Results showed that breeding sites located at low elevation sites in flooded unmanaged pastures were the most important determinants of mosquito abundance. Dister *et al.* (1997) analyzed the relation of Ixodes scapularis nymphs (the vector of Lyme disease) abundance with vegetation structure, moisture (wetness) ad vegetation density (greenness) derived from remote sensed data and found that high-risk regions were significantly greener and wetter. Rogers and Randolph (1991) found mortality of tsetse flies, the size of male and female tsetse

and the abundance were significantly correlated with the factors derived from meteorological remote sensing data and suggested that the information can be used to predict the mortality and abundance of tsetse and to produce maps of high risk areas of disease transmission.

For the disease prevention, Barinaga (1993) published an article named "Satellite Data Rocket Disease Control Efforts Into Orbit", arguing the possibility of application of remote sensing into the disease controlling. In this article, the author gave some examples like malaria distribution in the remote sensing images to evaluate the disease control capability of remote sensed data. Also, the author said "despite the growing enthusiasm, disease prevention by satellite still has to prove itself for actual disease control".

There are also some researches utilizing remotely sensed data on medical geography in South Korea (hereafter, Korea). For example, Sithiprasasna *et al.* (2005) used both QuickBird and Landsat remote sensed imagery to identify the larval and adult anopheline mosquito habits in Korea, and showed that the immature collections of Anopheles senensis were significantly associated with land-used. The author also highly evaluated the classified remotely sensed data in the use of estimate the distribution of immature and adult mosquito populations. In order to compare the costs of two malaria-controlling methods, Claborn *et al.* (2002) applied remote sensing and geographic information systems to

estimate the size of vector larval habits around two U.S. Army camps in Gyeonggi province, Korea, and demonstrated this method allowed extrapolation of larval surveillance data to a regional scale.

All in all, as Cline (1970) suggested, "there may be a whole spectrum of applications on health geography awaiting the imaginative researchers".

2.1.3 Spatial statistical approach to health geography

Unlike the general data, data observed or measured in specific locations or within specific regions are spatial data. Analyzing spatial data with general statistical techniques may be flawed because the general statistical techniques are based on the aspatial assumptions (Lee, 2002). Instead, spatial statistics is rational during spatial data analysis. As Lee (2002) suggested, "spatial statistics is not simply a bundle of methods or statistical techniques, but is a new perspective on geographically referenced data with a theoretical integrity centered on space."

Various spatial statistical approaches have been applied in the study of disease pattern and the health care delivery, which are the two basic themes of health geography. The central role of spatial statistics in health geography is to (1) evaluate the differences of disease data in different geographical regions; (2) identify disease "clusters"; (3) observe the patterns of disease; (4) assess the significance of the determinant factors (Waller and Gotway, 2004). According to Hungerford (1991), there are

three issues that are critical: whether a disease is clusterd; is there any similar distribution of two diseases or a disease and its potential risk factor; and whether there are any specific relationship between the values of the same variable in different regions.

For the evaluating the differences of disease data among different regions, the spatial patterns of the disease are always the interests of both the epidemiologists and health geographers. For example, Walter (1993) used the spatial autocorrelation indices, which are Moran's I, Geary's c, and adjacency statistics to analyzed the different cancer incidence of different regions. Bhunia and Shit (2019) also adopted GIS and spatial statistic tools to analysis the spatial-temporal patterns and distribution of diseases. What's more, disease mapping as the visualization of disease distribution, also attract the attention of researches since decades before. For example, Walter and Birnie (1991) create the disease atlases of 49 different regions in 1991. Pickle *et al*. (1996) made an all-causes mortality atlas and Devesa *et al*. (1999) made a separate cancer mortality atlas of the United States.

The disease cluster identifying is under studying for many decades. For example, Alexander *et* al. (1989) found an evidence of spatial clustering of Hodgkin disease, which was along with other epidemiologic evidence. Kulldorff and Nagarvalla (1995) develops a new methods to detect and inference foe the spatial clusters of a disease. Brooker *et al*. (2004) analyze

the spatial clusters of malaria and the potential risk factors in a highland area of western Kenya using spatial scan statistics approach. And Sankoh *et al*. (2001) also adopts this method to study the clustering of childhood mortality. Dearden *et al*. (2019) also evaluate the health status change in Britain from 1991 to 2011 by analyzing the spatial autocorrelation among the neighbor regions.

Literatures on health geography in the mid-1980s revealed that quite a few literatures were interested in the spatial analytic approaches as aids in understanding both the disease distributive pattern and the health care delivery systems (Gesler, 1986). According to Albert *et al.* (2003), the spatial analysis approaches applied in health geography can be sorted as spatial analytic techniques for points, lines, areas and surfaces, as shown in Table 2-1.

Spatial statistics have both global and local scales. The global spatial statistical indicators indicate the "average" and the "integral" spatial association, while the local ones are the decomposition of the global ones, and can reflect the local instabilities in the overall association (Anselin, 1995; Lee, 2001).

Table 2-1 Spatial analytical techniques

Points	**Surfaces**
Mean center/ standard distance	Isolines
Standard deviational ellipse	Trend surface analysis
Gradient analysis	Power series polynomials
Nearest neighbor	Fierier series
Variance/ mean ratio test	
Quadrat analysis	**Map comparisons**
Space clustering	Lorenz curves
Space-time clustering	Coefficient of area correspondence
	Correlation coefficient
Lines	Difference maps
Random walk	
Vectors	**Relative spaces**
Graph theory:	Case-control matching
Nodality	Acquaintance networks
Connectivity	Multidimensional scaling
Nodal hierarchies	Cluster analysis
Flow analysis	
Areas	
Location quotients	
Standardized mortality ratios	
Poisson probability	
Space clustering	
Space-time clustering	
Autocorrelation measure	

Source: Albert *et al.*, 2003: 12.

2.2 Main concepts in this study

2.2.1 Anthropogenic factors

"Anthropogenic" refers to things have been generated by human (Wikipedia, https://en.wikipedia.org/wiki/Anthropogenic). Anthropogenic factor similarly means human activity. Health-threatening anthropogenic factors represent for those human activities that could be harmful to human health. Actually, the harmful influence of anthropogenic factor has been studied in various fields, such as the impact on vegetation (Smiet, 1992; Skole *et al.*, 1994; Wackernagel and Rees, 1998), animals (Gaston *et al.*, 2013; Naguib, 2013; Eriksen *et al.*, 2014), soil (Brookes, 1995; Bridges and Oldeman, 1999; Oldeman *et al.*, 2017), waters (Novotny, 1994; Vörösmarty *et al.*, 2010; Chenoweth *et al.*, 2014), geomorphology (Walling, 2006; Goudie, 2018), climate and the atmosphere (Arnfield, 2003; Berglund, 2003).

With the intensifying of the impact of health-threatening anthropogenic factors, there are more and more human morbidity and mortality due to the affected environment by human activity. In the ten largest contributors to global disability-adjusted life-years (DALYs), there are at least four of them which are affected by the consequence of human impact on the natural environmental (Forouzanfar *et al.*, 2016). For example, Rosenthal *et al.* (2014)proved that the IMS cover had a significant positive correlation on the heat-related mortality ratio. Ibald-Mulli *et al.*

(2001) studied the effects of air pollution on blood pressure, and found that the exposure to ambient air pollution could be related to the change in cardiovascular autonomic control, and thus may have association to the high blood pressure. Ponticiello *et al.* (2015) also concluded that workers exposed to urban pollution have an additional risk for body mass index (BMI) increasing. In recent years, NTL pollution have attracted the attention of researchers, and studies have proven that the exposure to NTL can induce hormone-based cancers like breast cancer, prostate cancer and so on (Kloog *et al.*, 2008; Kloog *et al.*, 2009; Kloog *et al.*, 2010).

Therefore, the capabilities for frequent and rapid observation of those human activities on both large spatial and temporal scales could improve the understanding of human spatial distribution and its impact on environment. With its diverse spatial and temporal scales, remote sensing techniques can be adopted.

Actually, many health-threatening anthropogenic factors can be detected using remote sensing techniques, such as the IMS (Bauer *et al.*, 2007; Lu *et al.*, 2011; Zhang and Huang, 2018), NTL pollution (Kloog *et al.*, 2009; Kloog *et al.*, 2010; Bennie *et al.*, 2014), air pollution (Gupta *et al.*, 2006; Garland *et al.*, 2008; Martin, 2008), water pollution (Ritchie *et al.*, 2003), to name a few.

2.2.2 Exploratory spatial data analysis

As Tukey (1977) demonstrates, exploratory data analysis (EDA) is a detective work, which can be a numerical detective work, counting detective work or graphical detective work, and EDA usually can discover the indications, mostly quantitative ones. Anselin (1999) says that the EDA for statistical analysis is to let the data speak for themselves without imposing much on them. For the criticism of spatial data as "data rich but theory poor" (Openshaw, 1991; Anselin, 1999), the EDA seems to be a good solution. However, the characters of spatial data, such as the existence of the spatial autocorrelation, may violate the independent assumption in EDA studies. Therefore, it is necessary to develop the specialized approaches of exploratory spatial data analysis (ESDA) taking the nature of spatial data into account.

One step endeavor to the real ESDA is the spatialized EDA for that the location information is combined, such as the spatial distribution of Chernoff faces on maps (Anselin, 1999). It becomes more close to ESDA with the description of the spatial heterogeneity or spatial trends, as suggests by Anselin (1999).

The definition of ESDA is " a collection of techniques to describe and visualize spatial distribution, identify stypical locations or spatial outliers, discover patterns of spatial association, clusters or hot spots, and suggest

spatial regimes or other forms of spatial heterogeneity" (Anselin, 1994; Anselin, 1998; Lee, 2002).

According to Unwin (1996) and Lee (2002), the spatial data analysis (SDA) can be divided into exploratory spatial data analysis (ESDA), confirmatory spatial data analysis (CSDA), and prescriptive spatial data analysis (PSDA). ESDA is data-driven and inductive, while CSDA is model-driven and deductive (Lee, 2002). Although the boundary between ESDA and CSDA is blurry, the ESDA is suggested to be more appropriate with the spatial data analysis. As Haining *et al.* (2000) and Lee (2002) stated, ESDA aims at seeking spatial patterns, identifying the clusters and outliers, formulating hypotheses and assessing spatial models based on spatial data. Together with the cooptation with GIS and other current research platform, ESDA can give full play to its advantages on the field of spatial analysis, especially on health geography (Owusu-Edusei and Owens, 2009; Yao *et al.*, 2012).

This study adopted ESDA to evaluate the spatial distribution of the possible health-threatening anthropogenic factors extracted in remote sensing and their possible health outcomes, as well as their spatial association. In the following part, I introduce the spatial analytical techniques I utilized in this study.

2.2.2.1 The standardized score of dissimilarity

In order to give insight into the spatial distribution characteristics of health outcomes, it is necessary to quantify the proportion of disease cases in different regions. Several methods quantifying the proportion have been widely utilized. One of the methods is simply calculating the ratio of the observed population to the total population, which is also called the column-proportions (Lee, 2007). The standardization of the column-proportions and the Location Quotient (LQ) are also extensively employed. However, there is one common important drawback of these methods: it ignores the reality that the total population in different regions are different. That is to say, we cannot differentiate whether the ratios are to a region with a large population or a little one. So it is necessary to propose a method considering the different total populations of regions.

The standardized score of dissimilarity (SSD) is a spatial separation statistical measure proposed by Lee (2007; 2008) by combining the spatial association measures and spatial chi-square statistics. Unlike previous segregation indexes like index of dissimilarity (Duncan and Duncan, 1955), SSD can detect the spatial dissimilarity patterns between two groups as well as the spatial dependence of the patterns. By standardizing the row-proportions of some group into the population proportion of the region in the total population, SSD cannot only identity the representative regions

with high heterogeneity, but take the different population in different regions into account. The formula of SSD is as follows:

$$SSD = \frac{\frac{X_i}{X} - \frac{Y_i}{Y}}{\sqrt{\frac{\sum_i (\frac{X_i}{X} - \frac{Y_i}{Y})^2}{n}}}$$

Where X is the total population of observed group; X_i is the population of group X in region i; Y is the total population of the criteria group; Y_i is the population of group Y in region i. The positive values of SSD indicate the absolute concentration of the observed group in this region, and the larger the value is, the higher the concentration degree is, and vice versa.

2.2.2.2 Cluster detection

G_i and G_i^* help us to detect local "pockets" of spatial dependence that may not be evident when using global statistics. The Moran's I statistics has its own weakness that it isn't able to discriminate between patterns that have high values dominant or low values dominant, but G_i and G_i^* statistics can (Getis and Ord, 1992; Ord and Getis, 1995).

$$G_i^* = \frac{\sum_{j=1}^{n} \mathrm{w}_{ij}\mathrm{x}_j - \bar{x}\sum_{j=1}^{n} \mathrm{w}_{ij}}{S\sqrt{\frac{n\sum_{j=1}^{n} \mathrm{w}_{ij}{}^2 - (\sum_{j=1}^{n} \mathrm{w}_{ij})^2}{n-1}}}$$

Where $\mathrm{S} = \sqrt{\frac{\sum_{j=1}^{n} \mathrm{x}_j{}^2}{n} - \bar{x}^2}$

Where n is the number of spatial units indexed by i and j; x is the variable interest; $\bar{x}$ is the mean of x; and w_{ij} is an element of a spatial weight matrix.

2.2.2.3 Bivariate spatial association measure

Correlation coefficients measure the linear association between two variables, but when they are spatial data, these correlation coefficients are deficient in that they are aspatial and disregard the spatial characteristics (Haining, 1991). But the presence of spatial autocorrelation indicates that a certain amount information is shared and not independent, which violates the independent prediction of many standard statistics (Lee, 2017). Bivariate spatial association measure estimates the relationship between two variables, with taking the spatial topological relationship among observations into account. A bivariate spatial association measure should capture the numerical co-varying ("point-to-point" association (Hubert *et al.*, 1985)) as well as the spatial clustering ("spatial association")(Lee, 2001). Although the bivariate spatial association problem has long been recognized, Wartenberg (1985) firstly did the comprehensive trial to devise a parametric bivariate spatial association measure, with proposing a matrix algebraic form for the bivariate Moran's I, but the measure has some drawbacks. In order to refine the shortage, Lee (2001) developed a bivariate spatial association index L for both global scale (for detecting

the bivariate spatial dependence) and local scale (for exploring the bivariate spatial heterogeneity).

In this study, the bivariate spatial association measure Lee's L was used to detect the co-pattern between possible health-threatening anthropogenic factors extracted in remote sensing and the possible health threats. Both global L and local L_i were used, of which the formula are as follows:

$$L = \frac{n}{\sum_i (\sum_j v_{ij})^2} \times \frac{\sum_i \left[\left(\sum_j v_{ij} (x_j - \bar{x}) \right) \times \left(\sum_j v_{ij} (y_j - \bar{y}) \right) \right]}{\sqrt{\sum_i (x_i - \bar{x})^2} \times \sqrt{\sum_i (y_i - \bar{y})^2}}$$

$$L_i = \frac{n \times \left[\left(\sum_j v_{ij} (x_j - \bar{x}) \right) \times \left(\sum_j v_{ij} (y_j - \bar{y}) \right) \right]}{\sqrt{\sum_i (x_i - \bar{x})^2} \times \sqrt{\sum_i (y_i - \bar{y})^2}}$$

Where n is the number of spatial units indexed by i and j; x, y are the variables of interest; $\bar{x}$, $\bar{y}$ are the mean of x, y; and v_{ij} is an element of a general spatial weight matrix V .

2.2.2.4 Local Pearson's r_i

A local Pearson's r_i indicates the degree of numerical correspondence between two values at a location, which is also called "point- to-point" association (Hubert *et al.*, 1985; Lee, 2001; Lee, 2004; Lee *et al.*, 2013). The formula of local Pearson's r_i is as follows (Lee, 2004):

$$r_i = z_{xi} \times z_{yi}$$

Where x and y are two variables, and i represents for the i^{th} location. And the z scores are standardized by the mean and standard deviation of each variable. The mean value of local Pearson's r_i is a global Pearson's r between two variables.

2.2.2.5 Spatial error model

I adopt regression analysis in order to identify the influences of the health-threatening anthropogenic factors on human health. In the traditional regression, the error term have a mean of zero ($\mathrm{E}[\varepsilon_i] = 0, \forall_i$), and they are identically and independently distributed (i.i.d.). Therefore, the variance of the error term is constant ($\mathrm{Var}[\varepsilon_i] = \sigma^2$), and they are uncorrelated, $\mathrm{E}[\varepsilon_i \varepsilon_j] = 0, \text{for all } i, j$ (Anselin, 2009). However, aggregated spatial data are characterized by dependence (or spatial autocorrelation) and heterogeneity (or spatial structure), which may invalidate certain classic methodologies (Anselin, 2013; Lee, 2017).

Spatial dependence is applied in the regression model in two major ways, one is the spatial lag model (SLM), and the other is the spatial error model (SEM). SLM is appropriate when the research focus is on the spatial interaction, while SEM is appropriate when the research is concerned with correcting for the biasing influence of the spatial dependence due to the use of spatial data (Anselin, 2013). For this study, spatial interaction is not the focus, instead, the biasing influences of the spatial data are to be

considered. Therefore, SEM is adopted in the influences analysis of health-threatening anthropogenic factors on human health.

The SEM is as follows:

$$y = X\beta + \lambda W\varepsilon + \eta$$

Where λ is the spatial autoregressive coefficient, W is the spatial weight matrix, and η is the random error term.

2.2.2.6 Geographically weighted regression

In order to estimate the influences of every factor in the specific locations, geographically weighted regression (GWR) model is utilized. In the classic linear regression a stationary process were assumed, that is, we purpose the same relationship throughout the entire study region and presume that the same stimulus will have same response in all units of the study region (Fotheringham *et al.*, 2009). Consequently, if there is spatial dependence, the single resulting parameter represents for the average of the heterogeneous process operating over space. Instead, parameters estimated to vary over space is essential, and that is the core of GWR. The GWR model can be described in the following equation:

$$y_i = \beta_{i0} + \sum_{k=1}^{p-1} \beta_{ik} x_{ik} + \varepsilon_i$$

$$\beta'(i) = (\mathbf{X}^{\mathrm{T}}\mathbf{W}(\mathrm{i})\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}(\mathrm{i})\mathbf{Y}$$

Where y_i is the dependent variable value at location i, x_{ik} is the value of the k th covariate at location i, β_{i0} is the intercept, β_{ik} is the regression coefficient for the k th covariate, p is the number of regression terms, and ε_i is the random error at location i. $\mathbf{W}(i)$ is a spatial matrix of weights specific to location i such that the nearer neighbors of i are giver greater weight than the further ones.

2.3 Research Design

2.3.1 Health-threatening anthropogenic factors

In this study, two of the most representative health-threatening anthropogenic factors in remote sensing are selected. One is man-made IMS, and the other is artificial NTL. There are two main reasons: the representativeness of two factors and the necessity.

The representativeness of two factors

IMS represents the anthropogenic features in which water cannot infiltrate through, such as roads and rooftops (Weng *et al.*, 2008). According to previous studies, anthropogenic land cover takes up about 40 percent of the Earth's surface and large amount of natural-dominated landscapes have been transformed into IMS (Sterling and Ducharne, 2008; Xu, 2010). Therefore, IMS is identified as one of the most representative anthropogenic factors of the remote sensing. Also, many studies have

proved the relationship between IMS and heat. Consequently, IMS can be seen as one of the most representative health-threatening anthropogenic factors in remote sensing.

NTL data is one of the few nocturnal remote sensing sources. NTL data have been frequently used as a proxy of human activities, such as population estimation (Sutton *et al.*, 2001; Amaral *et al.*, 2005; Yoo *et al.*, 2011), economic or energetic activities quantification (Amaral *et al.*, 2005; Doll *et al.*, 2006; de Miguel *et al.*, 2014), and urban extraction or urbanization estimation (Liu *et al.*, 2012; Kim, 2014; Sharma *et al.*, 2016). Also, studies proved that the exposure to NTL can induce some hormone diseases. Therefore, NTL could also be seen as one of the most representative health-threatening anthropogenic factors in remote sensing.

The necessity

For IMS and possible heat-related threats, Meehl and Tebaldi (2004) predict that there will be more intense, more frequent and longer lasting heat waves in the near future. Therefore, studies on heat prevention and heat-related threats are necessary. Also, studies show that 83% of the world population and more than 99% of the American and European are under light pollution (Falchi *et al.*, 2011). Thus, identifying the adverse effect of NTL on human health, and implementing measures to reduce the influences are of necessity.

2.3.2 Possible health outcomes

According to the previous studies, this study selects the discomfort index (DI) and the non-accident mortality as possible heat-related health outcomes of man-made IMS, and the breast cancer prevalence as the possible health outcome due to the exposure to NTL.

Discomfort Index was first proposed by Thom (1959) and was modifies by Sohar *et al.* (1963) and used to express the impact of climate factors on human. DI values are mainly influenced by temperature and humidity. Therefore, the possible increase in temperature due to man-made IMS would influence the DI value, and influence human health subsequently.

As one of the most common and serious health outcomes of heat, non-accident mortality is also receiving attention (Choi *et al.*, 2005; Robine *et al.*, 2008; Barriopedro *et al.*, 2011; Rosenthal *et al.*, 2014; Cao *et al.*, 2018). Thus, non-accident mortality is selected as a possible heat-related health outcome of the man-made IMS.

Studies show that excessive exposure to NTL can increase the risk of breast cancer (Kloog *et al.*, 2009; Stevens, 2009; Kloog *et al.*, 2010; Falchi *et al.*, 2011). The primary hypothesis of the carcinogenic characteristics of NTL is due to the suppression of melatonin, which can induce hormone-

related cancers (Kloog *et al.*, 2009; Kloog *et al.*, 2010). Consequently, this study selects breast cancer as a possible health outcome of NTL.

Therefore, the main research of this study consists of three parts in general. In Chapter 3 and 4, the detection of co-pattern between the health-threatening anthropogenic factors in remote sensing and their health outcomes are discussed. In chapter 5, models designed to identify the influences of the health-threatening anthropogenic factors on human health are discussed.

Chapter 3. Man-made Impervious Surface and Its Potential Heat-related Threats to Public Health

3.1 Introduction

Temperature distribution is determined by latitude, vegetation distribution, height above sea level, topography, city size and atmosphere stability (Stewart and Oke, 2009). The rapid urbanization leads to the alteration of natural landscapes to impervious land, with changing surface radiation, thermal characteristics and humidity over the urban areas (Meng *et al.*, 2018). Actually, there have been many studies showing urban areas have high temperature, among which urban heat island effect is seen as one of the most important manifestations of urban-induced climate change (Cao *et al.*, 2018). There are also studies indicating the development of urbanization have impacts on the heat-island effects and the global-warming caused effect in Korea (Gim *et al.*, 2018; Nam *et al.*, 2018).

Whereas, the global climate has been proved to be warmer in the past century, and this trend is predicted to be continued in the next 100 years (Pachauri *et al.*, 2014). Heat could be seen as the deadliest of all atmospheric phenomenon (Sheridan and Kalkstein, 2004). Many studies

have proved the adverse effects of heat on human health, regional economies and ecosystems, and even the lasting and intensifying of heat waves in the future due to the global temperature increases (Meehl and Tebaldi, 2004; Song and Wu, 2018). Especially for health, studies show that exposure to the heat can be related to morbidity and mortality like heat syncope, cramps, exhaustion as well as heat stroke (Sheridan and Kalkstein, 2004). Robine *et al.* (2008) studied that the additional deaths exceeded 70,000 due to the heat wave in August 2003 in Europe. Barriopedro *et al.* (2011) also estimated that the extreme high temperature in summer 2010 have induced about 55,000 deaths. In Korea, there are also studies going into the exceed deaths due to the high temperature. Choi *et al.* (2005) have studied the summertime disease-related mortality due to the high apparent temperature and showed that blood circulation-related and cancer-related mortality increased due to the increase of the apparent temperature, and the elderly persons over 65 years old are more vulnerable to mortality due to abnormal heat waves in Seoul.

As mentioned before, DI was used to express the impact of climate factors (temperature, humidity) on human, and have been widely used in diverse studies. For example, Angouridakis and Makrogiannis (1982) studied the DI value from 1950 to 1975 during warm seasons (May to September) in Thessaloniki, Greece, and found a quite increase of DI values during June and August. Tselepidaki *et al.* (1992) also proved that the

discomfort degree was higher during June to September in Athens, Greece using DI values.

Seoul Metropolitan Region (SMR) has long been the center of Korea for policies, economy and culture, and the population and the employment of SMR are shown to increase constantly (Kim *et al.*, 2018). During the process of population intensifying, whether the urban-related heat have more serious influence on human health, and the degree of the impact are to be evaluated.

Man-made impervious land (IMS) have long been recognized as the proxy of urbanization, and knowledge of impervious surface changes is important for understanding the urban environment and human activity(Zhang and Huang, 2018). For the importance of the IMS, many studies have focused on the monitoring of IMS, among which, remote sensing methods have been preferred for its long spatial and temporal resolution (Slonecker *et al.*, 2001; Bauer *et al.*, 2007; Lu *et al.*, 2011; Zhang and Huang, 2018).

This study aims to analyze the spatial distribution of man-made IMS and its possible threats to human health, such as DI values and mortality. Ultimately, we evaluate the spatial association between the man-made IMS and its possible threats to human health in order to find the co-pattern between the two variables.

There are four sections in this chapter. In Section 2, I describe the data and methods used in this study. In Section 3, I provide an analysis of

the spatial distribution of man-made IMS as well as the spatial distribution of the possible threats to human health (DI values and non-accident mortality), and the results of the spatial bivariate correlation between the two variables. Section 4 is the discussion and conclusions, in which I suggest academic and practical implications and limitations of this study.

3.2 Materials and Methods

3.2.1 Data

This study takes Seoul Metropolitan Region (SMR) as the study area, which includes Seoul, Inchon, and Gyeonggi-do. Seoul is the capital and the most developed city in Korea, and as the adjacent regions of Seoul, Inchon and Gyeonggi-do are being influenced highly by Seoul. Both 1Km by 1Km grid and *Dong* administrative unit have been used in this study.

The data used in this study include remote sensing images used to extract IMS, heat-related mortality data to analysis the spatiotemporal distribution pattern of heat-related mortality and meteorological data as ancillary data. For the heat-related threats are the most representative in summer, data in summer season (from June to August) were used in this study.

Table 3-1 Data

Category	Data	Source
Remote sensing data	Landsat 8 OLI/TIRS	United States Geological Survey (USGS) https://glovis.usgs.gov/app
Meteorological data	Temperature(℃) Relative humidity (%)	Korea Meteorological Administration http://www.kma.go.kr/
Mortality data	Non-accident mortality	Korea National Statistical Office http://kostat.go.kr/

3.2.1.1 Remote sensing data: man-made impervious surface

In this study, man-made IMS is seen as one of the human activities in remote sensing. Due to the large spatiotemporal coverage and high temporal frequency, remote sensing data have been widely used to extract and monitor the impervious surfaces (Zhang and Huang, 2018). Remote sensing data with medium spatial resolution (10-100m) have been widely used in impervious surface extraction. For example, Powell *et al.* (2008) adopted Landsat Multispectral Scanner (MSS), Thematic Mapper (TM), and Enhanced Thematic Mapper Plus (ETM+) to estimate the amount of land that were converted into impervious surfaces due to the urbanization and residential development. Bauer *et al.* (2007) depicts the methods and

results for estimating and mapping the impervious surfaces using Landsat TM and ETM+ for the state of Minnesota in 1990 and 2000.

Medium remote sensing images are usually used at a macro level of impervious surface extraction, i.e. the urban expansion, with ignoring subtle changes within urban areas (Zhang and Huang, 2018). In order to acquire the higher extraction accuracy and more detailed change of impervious surfaces, high resolution (<10m) remote sensing images have also been widely adopted. Cablk and Minor (2003) used high-resolution IKONOS imagery to directly detect impervious cover in South Lake Tahoe, California, USA and produced accurate identification for both commercial and residential impervious cover, and showed both high overall accuracy and user accuracy. Lu *et al.* (2011) selected two study areas with diverse urban developments, sizes and spatial patterns to explore suitable methods for extracting impervious surface using Quick bird imagery, and showed that spectral confusion of impervious surface with other land covers like water/wetland, bare lands, and the impacts of shadows are ineradicable and manual editing is necessary. Zhang and Huang (2018) applied high-resolution images with object-based classification integrating multiple features to monitor the impervious coverage in Shenzhen, China, and showed that there were some alteration of impervious surface to pervious surfaces, and there was also a decreasing trend of impervious surface after 2012.

Though subtle spatial distribution and change pattern of impervious surface is significant, this study only takes account of the general distribution and coverage of it, so medium-resolution images Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) data are used. Landsat 8 OLI/TIRS images consist of nine spectral bands with a spatial resolution of 30m and two thermal bands with a downscaled spatial resolution of 30m (100m originally).

Two scenes of images (path: 116, row: 34; path: 115, row: 34) cover the total study area, the images acquired on 26th August 2017 and 20th September 2017 respectively were downloaded from the the United States Geological System (USGS) Visualization Viewer (GloVis, https://glovis.usgs.gov/app?fullscreen=0) and conducted in the following analysis. Two scenes of images were pre-processed in ENVI 5.3 with mosaicking and subseting, and then converted into Universal Transverse Mercator projection.

3.2.1.2 Health data: non-accident mortality

According to the previous studies, non-accident mortality is tend to be utilized frequently as a proxy of the heat mortality (Choi *et al.*, 2005; Son *et al.*, 2016; Cao *et al.*, 2018).

Mortality data were downloaded from Korean National Statistical Office (KNSO, http://kostat.go.kr/). KNSO provides the Cause of Death

Statistics every year, which includes the cause of death (the International Classification of Disease Revision 10, ICD-10: A00-Y89), the sex, the case and rate of mortality and the age standardized mortality rate of each *Si-Gun-Gu* administrative boundary. Here, external causes of mortality (ICD-10: V01-Y89) have been excluded, and all the other causes of mortality (non-accident mortality) (ICD-10: A00-R99) were taken into analysis.

Annual mortality data of each *Si-Gun-Gu* administrative unit from January 1st 2015 to December 31st 2016 were downloaded and the average value of two years was used to do the following analysis.

3.2.1.3 Meteorological data

In this study we used temperature (℃) , relative humidity (%) to analyze the relationship between man-made impervious surface and the meteorological environment. The meteorological data were derived from the Korean Meteorological Administration (KMA, http://www.kma.go.kr/index.jsp). Automatic Weather System (AWS) is the equipment for observing real time meteorological condition by recording local temperature, wind, and precipitation etc. every minute. And there are about 480 AWS sites in total throughout Korea.

There are totally 97 AWS sites in Seoul Metropolitan Region, and hourly meteorological information of these sites from Januray 1st 2015 to December 31st 2016 were downloaded and converted into daily average

values, which was used to calculated DI values between annual average of 2015 and 2016 respectively.

3.2.2 Methods

3.2.2.1 Man-made IMS extraction

For the estimating and mapping of impervious surface, pixel-based methods and object-based methods are the main methods utilized. Object-based methods aims at delineating reality usable objects from remote sensing imagery, and studies show that it outperforms the pixel-based methods (Blaschke, 2010). Therefore, object-based classification and the combination of traditional pixel-based methods and object-based methods become more and more pervasive. Shackelford and Davis (2003) have built an object-based method upon a pixel-based fuzzy classification to extract the roads, buildings, and non-road, non-building impervious surface in dense urban areas using IKONOS imagery, and got high classification accuracies. Wang *et al.* (2004) mapped mangroves with IKONOS imagery using pixel-based method, object-based method and a hybrid method that integrates pixel-based and object-based method, and results showed that the combination of pixel-based and object-based method achieved the best accuracy. Zhang and Huang (2018) monitored

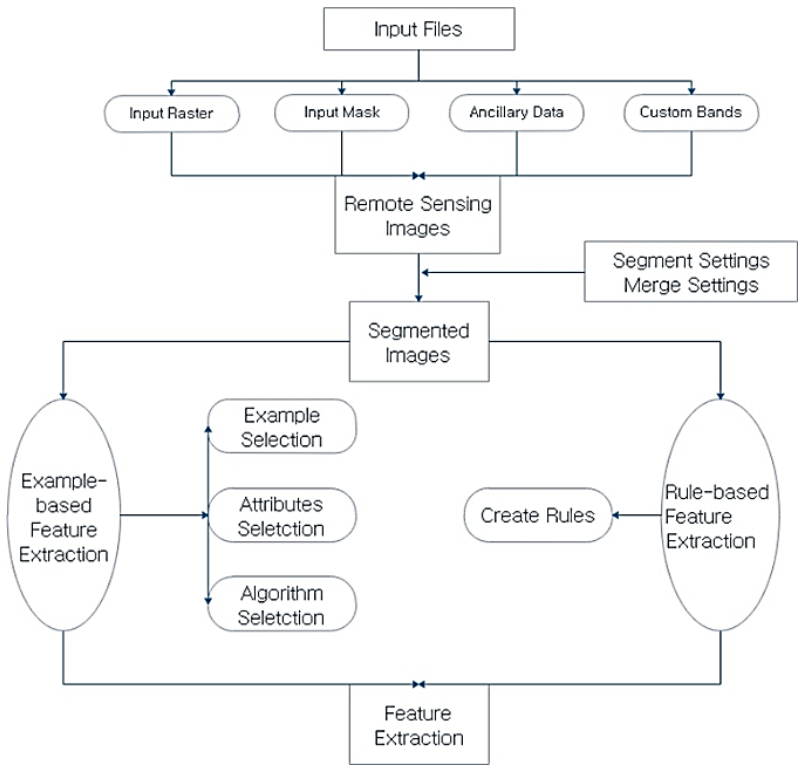


Figure 3-1 The description of object-based feature extraction

the impervious surface using object-based classification method integrating multiple features, including spectral, textural, shape and class-related features within high-resolution remote sensing images, and showed an accurate and reliable impervious surface information.

Object-based classification is based on image segmentation (Blaschke, 2010). In the data processing, segmentation is the first step. Segmentation is the process of partitioning the image into objects corresponding to real-world features by grouping neighboring pixels which have common characteristics. Then example-based feature extraction and rule-based feature extraction have been used. Example-based feature extraction is the process of using supervised training data to assign unknown objects to one or more known features. Rule-based feature extraction is the process of taking a segmented image in to groups using one or more rules that users can build based on the users knowledge or certain features.

By testing, Edge segment algorithm with a scale level of 50, and Full Lambda Schedule merge algorithm with a merge level of 60 is the most proper settings for our remote sensing images to be segmented. For the example-based feature extraction, all attributes of spectral, texture and spatial of all bands were selected, and Support Vector Machine (SVM) algorithm with radial basis kernel type was utilized.

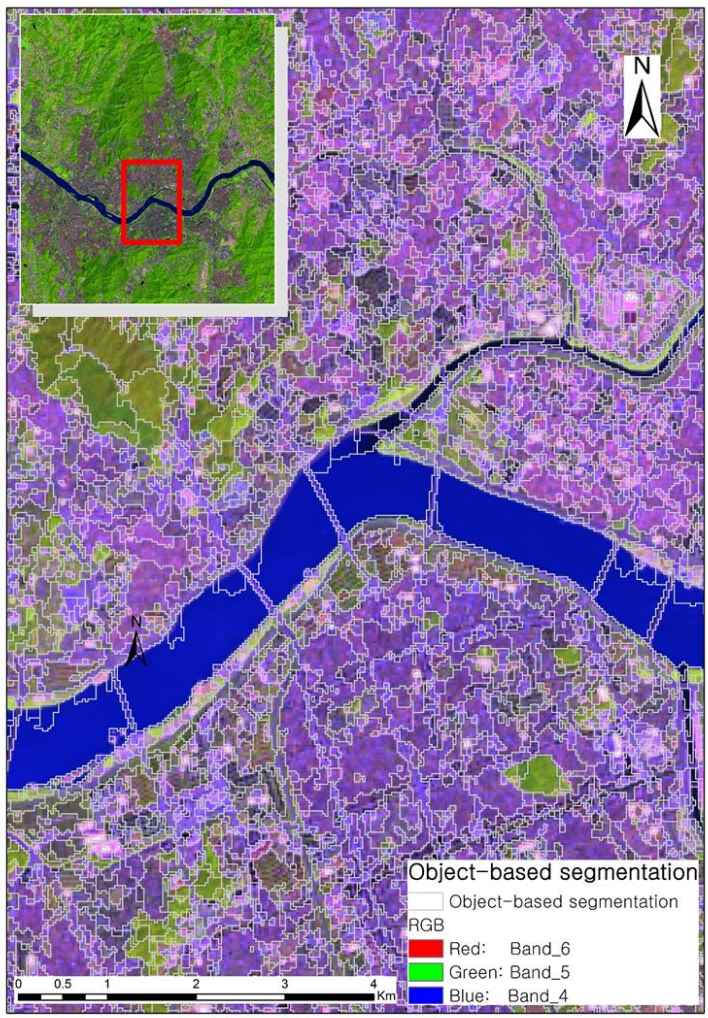


Figure 3-2 The process of object-based segmentation

3.2.2.2 Land surface temperature extraction

In order to improve the accuracy to extract the man-made IMS, land surface temperature (LST) was used as ancillary data. The LST were derived from Landsat 8 TIRS 1and TIRS 2 bands. The main difference between the new TIRS and previous TM/ETM sensors (apart from the differences of the sensor design) is the presence of two TIR bands in the atmosphere window between 10 and 12 μm, which shows an advancement over the previous single thermal band (Jiménez-Muñoz *et al.*, 2014). In order to acquire the LST, at-satellite brightness temperature (T_B) and the land emissivity are needed. The following equation is used to get the spectral radiance from the digital number (DN) of the TIRS bands(USGS, 2015):

$$L_\lambda = 0.0003342 * DN + 0.1$$

where L_λ is the spectral radiance, and DN is the digital number.

The next step is to convert the spectral radiance to T_B(USGS, 2015):

$$T_B = \frac{K_2}{\ln(\frac{K_1}{L_\lambda} + 1)}$$

where T_B is effective at-satellite temperature in K, L_λ is the spectral radiance in W/m^2 ster μm); and K_1 and K_2 are the pre-launch calibration constants. For Landsat 8, the K_1 for band 10 is 774.8853 Mw cm^{-2}sr^{-1}μm^{-1}, K_2 for band 10 is 1321.0789 K; and the K_1 for band 11 is 480.8883, K_2 for band 11 is 1201.1442 K.

The above temperature values are referenced to a black body, therefore, corrections for spectral emissivity (ε) became necessary according to the nature of land cover (Weng *et al.*, 2004). The emissivity corrected land surface temperature (LST) is calculated as follows (Artis and Carnahan, 1982):

$$\mathrm{LST} = \frac{\mathrm{T}_B}{1 + \left(\lambda \times \frac{\mathrm{T}_B}{\rho}\right) ln\varepsilon}$$

where λ is the wavelength of emitted radiance ($\lambda = 11.5\mu\mathrm{m}$) ((Markham and Barker, 1985) will be used), $\rho = \mathrm{h} \times \mathrm{c}/\sigma$($1.438\times10^{-2}$m K), σ=Boltzmann constant (1.38×10^{-23}J/K), h =Planck's constant (6.626×10^{-34}Js), and c=velocity of light (2.998×10^{8}m/s)(Weng *et al.*, 2004).

LST for TIRS 1 and TIRS 2 were calculated respectively, and then the average of them was used as the final estimated LST.

3.2.2.3 Discomfort Index

DI was first proposed by Thom (1959) and was adjusted by Sohar *et al.* (1963) as the following:

$$\mathrm{DI} = 0.5T_a + 0.5T_w$$

where T_a (℃) is dry-bulb temperature, and T_w (℃) is wet-bulb temperature.

Dry-bulb temperature refers to the air temperature measured by a thermometer exposed to the air, but shielded from moisture and radiation.

In this study, air temperature is used as the dry-bulb temperature (Song and Wu, 2018). The wet-bulb temperature is the temperature a parcel of air will have when it were cooled to saturation with a 100% relative humidity by the evaporation of water into it, with the latent heat being supplied by the parcel (Song and Wu, 2018).

In this study, wet-bulb temperature was calculated according to Stull (2011) 's method which can be expressed as follows:

$$T_w = T_a \times \mathrm{atan}[0.15977(RH + 8.313659)^{0.5}] + \mathrm{atan}(T_a + RH) - \mathrm{atan}(RH - 1.676331) + 0.00391838 \times RH^{1.5} \times \mathrm{atan}(0.023101RH) - 4.686035$$

Where T_a (℃) is dry-bulb temperature, RH is relative humidity (%); and T_w(℃) is wet-bulb temperature.

3.3 Analytical results

3.3.1 Spatiotemporal distribution of man-made IMS

Man-made IMS was extracted to 1Km by 1Km grids and each *Dong* unit. The area of man-made IMS and the area proportion to the *Dong* unit area were calculated, meanwhile the hot spots and cold spots of man-made IMS were evaluated, as shown in Figure3-3.

For the 1Km by 1Km grid unit, we can see the detail distribution of man-made IMS in the whole SMR. The total area of the study region is

11960 km^2, within which there are 2,675.3 km^2 (about 22.4% of the total area) man-made IMS. In Seoul, the total area is 611 km^2 (with 609 grids in total), while the area of man-made IMS area is 386.7 km^2 (about 63.3% of the area of Seoul, and about 14.5% of the total man-made IMS in the study area), and there are 582 grids have man-made IMS (about 95.7% of the total grids of Seoul), which shows a high density of man-made IMS in Seoul. On the other hand, the total area of Incheon is about 950 km^2 (with 950 grids), while the total area of man-made IMS in Incheon is 329.4 km^2 (about 34.7% of the total area of Incheon). However, 894 grids (about 94.1% of the total grids of Incheon) have man-made IMS. This indicates the relative decentralization of man-made IMS in Incheon. This situation seems more severe in Gyeonggi-do. The total area of Gyeonggi-do is about 10,400 km^2, with a total man-made IMS of 1,959.2 km^2 (about 18.8% of the total area). However, there are 8343 grids (about 80% of the total grids) distributed with man-made IMS, showing the high decentralization of man-made IMS.

Among the 11,960 girds, 484 girds with a high man-made IMS above 0.9 km^2 (90% of 1Km^2), of which 197 grids (40.7% of the total grids above 0.9 km^2, about 183.9km^2) are in Seoul, which takes up to 47.6% of the total man-made IMS area in Seoul. In Incheon, 94 grids (19.4% of the total grids above 0.9 km^2, about 90.2km^2) also have an area above 0.9 km^2, and 193 grids (39.9% of the total grids above 0.9 km^2, about 184km^2) with an area above 0.9 km^2 distribute in Gyeonggi-do, which only covers 9.4% of

the total man-made IMS area in Gyeonggi-do. Results show that the grid with man-made IMS above 0.9 km^2 are mainly distributed in Seoul and Gyeonggi-do, and there is a high centralization of man-made IMS in Seoul.

A cluster detection of man-made IMS for the 1Km by 1Km grid was conducted, as shown in Figure. 3-3 (b). Almost all regions in Seoul (except for the northern part of Gwanak-gu, and the southern part of Eunpyeong-gu, Jongno-gu, Gangbuk-gu and Dobong-gu respectively), all regions of Incheon except for Ganghwa-gun, and Goyang-si, Bucheon-si, Gwangmyeong-si, Siheung-si, Ansan-si, Suwon-si, Guri-si and other regions around Seoul in Gyeonggi-do are mainly distributed with hot spots. Meanwhile, the colds spots are mainly distributed in Ganghwa-gun of Incheon, and the peripheries of Gyeonggi-do, such as Yeoncheon-gun, Pocheon-si, Yangpyeong-gun and Gapyeong-gun.

In regard to the *Dong* unit, there are 1,128 units in total, of which 746 units (66.1%) have a man-made IMS area ratio higher than 50%, and 320 units (28.3%, 174 units in Seoul, 57 units in Incheon and 89 units in Gyeonggi-do), have an area ratio above 90%. The units with a high man-made IMS area ratio are mainly distributed in Nam-gu, *Dong*-gu of Incheon, Yeongdeungpo-gu, Guro-gu, and Yangcheon-gu of Seoul, and Paldal-gu of Suwon-si and so on. The *Dong* units with a man-made IMS area ratio above 90% are distributed discretely, as shown in Figure 3-4.

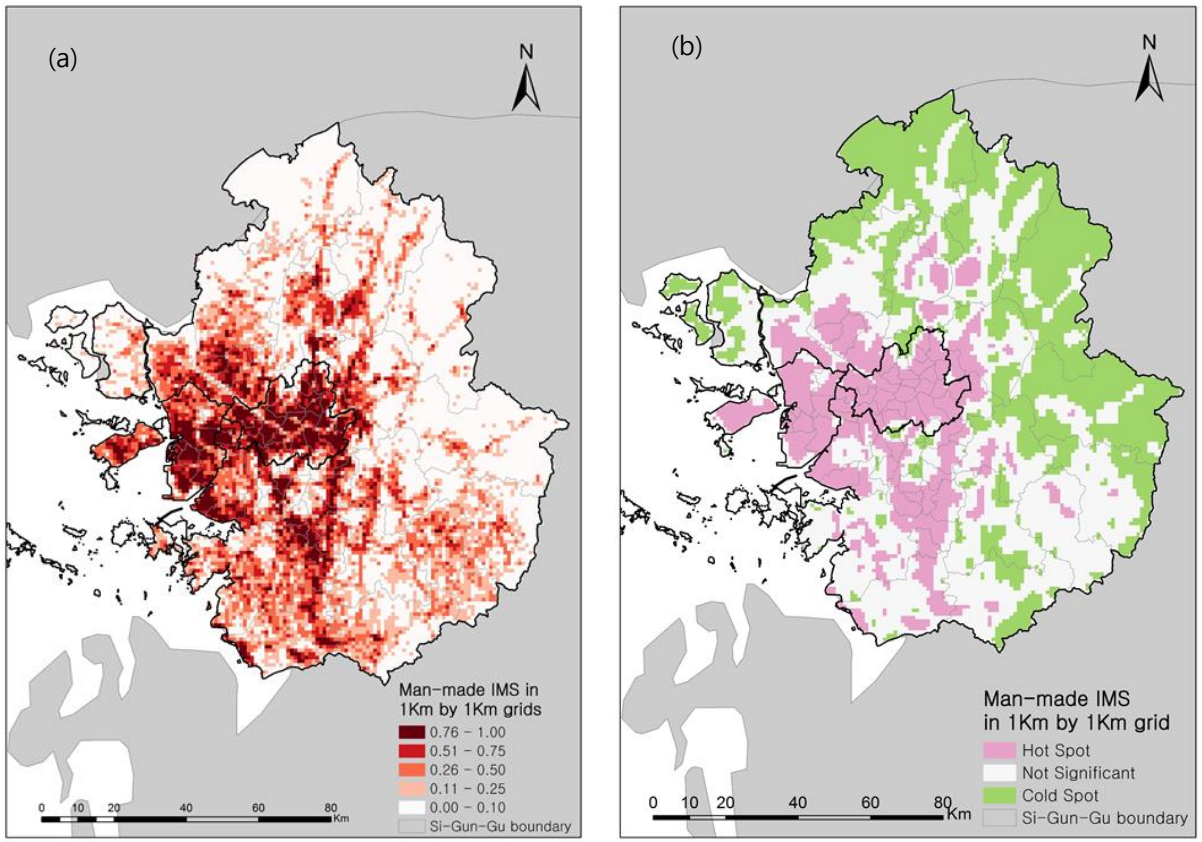


Figure 3-3 The spatial distribution of the man-made IMS: (a) the man-made IMS area in 1Km by 1Km grid; (b) cluster detection of the man-made IMS 1Km by 1Km in grid unit

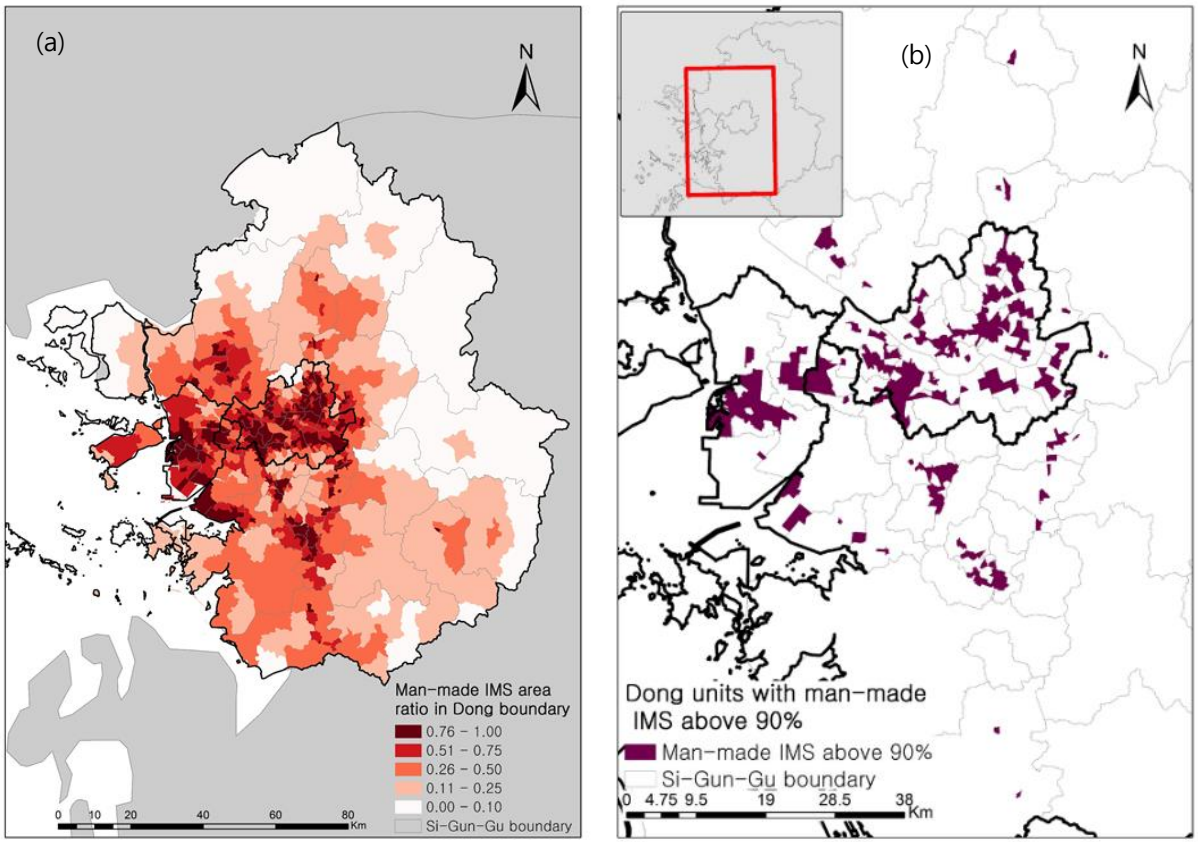


Figure 3-4 The spatial distribution of the man-made IMS in *Dong* unit: (a) man-made IMS ratio in *Dong* unit; (b) *Dong* units with man-made IMS above 90%

3.3.2 Spatial distribution of health-threatening risk

3.3.2.1 Spatial distribution of DI

The 1Km by 1 Km grid and the *Dong* unit were analyzed. The spatial distribution of the DI value, and the spatial cluster detection of DI in both the grid unit and the *Dong* unit have been conducted.

The results in the grid unit can show us the particular spatial distribution of DI, as shown in Figure 3-6. Natural break was used in the classification of the DI value. We can see that in general, the high DI values concentrate in Seoul. Most regions of Seoul, regions in Incheon except for Ganghwa-gun and Jung-gu, and some southern regions of Gyeonggi-do, like Hwaseong-si, Pyeongtaek-si, Anseong-si and Icheon-si have high DI values. On the other hand, Yeoncheon-gun, Pocheon-si, Gapyeong-gun and Yangpyeong-si have relatively low DI values.

By conducting spatial cluster analysis, we can see that the hot spots are concentrated mainly in Seoul, and some regions in southern Gyeonggi-do, such as Pyeongtaek-si, Ansan-si and Icheon-si. Cold spots are mainly distributed in the northeastern Gyeonggi-do, such as Yeoncheon-gun, Pocheon-si, Gapyeong-gun.

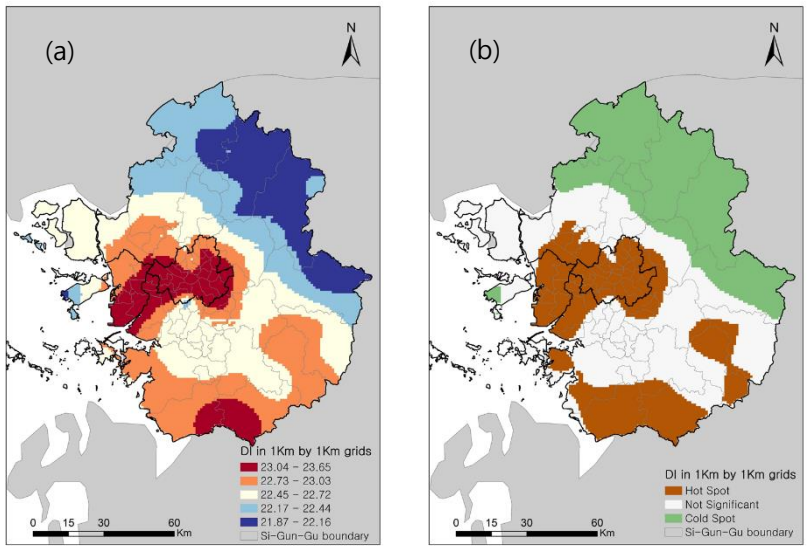


Figure 3-5 The spatial distribution of the DI value in 1Km by 1Km grid unit: (a) spatial distribution of DI; (b) cluster detection of DI

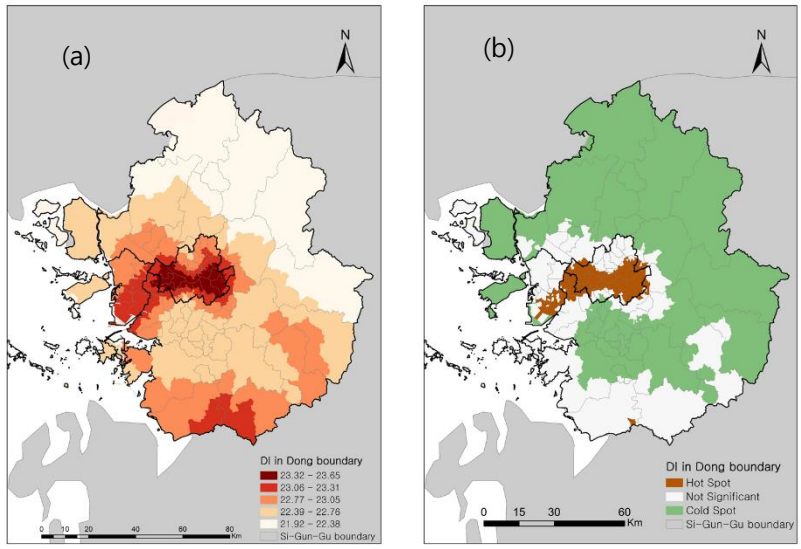


Figure 3-6 The spatial distribution of the DI value in *Dong* unit: (a) spatial distribution of DI; (b) cluster detection of DI

In order to have a clearer understanding of the DI distribution, DI has also been analyzed in *Dong* units, as shown in Figure 3-6 and Table 3-2. By classifying the average DI value in each *Dong* unit using Natural Break method, we can see that the highest values are concentrated within Seoul. Ten regions with highest DI values are all distributed in Yeongdeungpo-gu, Yangcheon-gu and Gangnam-gu. On the other hand, ten regions with lowest DI values are mainly concentrated in the peripheries of Gyeonggi-do, which are Gapyeong-gun, Pocheon-si and Yangpyeong-gun. Similarly, the result of the spatial cluster analysis of DI values in the *Dong* unit shows that hot spots of DI values are centralized in majority regions in Seoul and a few regions surrounding Seoul, while cold spots are mainly in the peripheries of Gyeonggi-do.

Table 3-2 Regions with the top and bottom 5 DI values

Rank (Top)	*Si-Do*	*Si-Gun-Gu*	*Eup-Myeon-Dong*	DI
1	Seoul	Yeongdeungpo-gu	Yangpyeong 1-dong	23.648
2	Seoul	Yangcheon-gu	Muk 1-dong	23.646
3	Seoul	Yeongdeungpo-gu	Dangsan 1-dong	23.640
4	Seoul	Yangcheon-gu	Muk 5-dong	23.62
5	Seoul	Gangnam-gu	Cheongdam-dong	23.61
Rank (Bottom)	*Si-Do*	*Si-Gun-Gu*	*Eup-Myeon-Dong*	DI
1	Gyeonggi-do	Pocheon-si	Ildong-myeon	21.92
2	Gyeonggi-do	Pocheon-si	Idong-myeon	21.93
3	Gyeonggi-do	Pocheon-si	Yeongbuk-myeon	21.99
4	Gyeonggi-do	Pocheon-si	Yeongjung-myeon	22.00
5	Gyeonggi-do	Yangpyeong-gun	Danwol-myeon	22.00

3.3.2.2 Spatial distribution of mortality

Non-accident mortality data have been analyzed in the *Si-Gun-Gu* spatial unit. The total number of mortality, the mortality rate and the SSD which represent the concentration of the mortality, have been studied, as shown in Figure 3-7. The results of the spatial distribution of total non-accident mortality numbers indicate a dense concentration in Seoul, and its surrounding regions such as Suwon-si, Seongnam-si, Bucheon-si and so on, suggesting a high mortality density in these regions. Opposite with the distribution of mortality population, the spatial distribution of the non-accident mortality shows a relatively low proportion of mortalities to the regional population in most regions of Seoul and Incheon, and some regions of Gyeonggi-do such as Goyang-si, Uiljeongbu-si, Suwon-si and so on. However, the peripheral Gyeonggi-do shows a relatively high mortality rate. The mortality rate can demonstrate the proportion of mortalities to the total population of each region, but it does not consider the difference of the total population in each region. That is to say, even with a constant mortality rate between two regions, the region that has a smaller population show a higher risk of mortality. Thus, the SSD of mortality has also been analyzed, which represents the mortality concentration of each region. The result shows that most peripheral regions of Gyeonggi-do have a high concentration degree of non-accident mortality, of which Nam-gu and Ganghwa-gu in Incheon, Yangpyeong-gu and Gangbuk-gu in Seoul, and Pocheon-si have the highest concentration

degree of non-accident moratlity, while Gangnam-gu, Songpa-gu in Seoul, Yeongtong-gu in Suwon-si, Hwaseong-si, and Seocho-gu in Seoul have the lowest concentration degree of non-accident mortality.

Table 3-3 Rankings of the non-accident mortality

Rank (Top)	*Si-Gun-Gu*	Case	Rank (Bottom)	*Si-Gun-Gu*	Case
1	Bucheon-si	3158	1	Gwacheon-si	238
2	Nowon-gu	2526	2	Yeoncheon-gu	394
3	Namyangju-si	2461	3	Dong-gu of Incheon	472
4	Bupyeong-gu	2390	4	Gapyeong-gun	530
5	Gangseo-gu	2317	5	Uiwang-si	578
		Rate (%)			Rate (%)
1	Ganghwa-gun of Incheon	1.04	1	Yeongtong-gu of Suwon-si	0.22
2	Yeoncheon-gun	0.86	2	Suji-gu of Yongin-si	0.27
3	Gapyeong-gun	0.85	3	Gangnam-gu	0.28
4	Yangpyeong-gun	0.79	4	Osan-si	0.29
5	Yeoju-si	0.69	5	Giheung-gu of Yongin-si	0.30
		SSD			SSD
1	Nam-gu of Incheon	1.79	-1	Gangnam-gu	-2.51
2	Yangpyeong-gun	1.71	-2	Sopa-gu	-2.41
3	Ganghwa-gun of Incheon	1.70	-3	Yeongtong-gu of Suwon-si	-2.23
4	Pocheon-si	1.56	-4	Hwaseong-si	-2.12
5	Gangbuk-gu	1.40	-5	Seocho-gu	-1.66

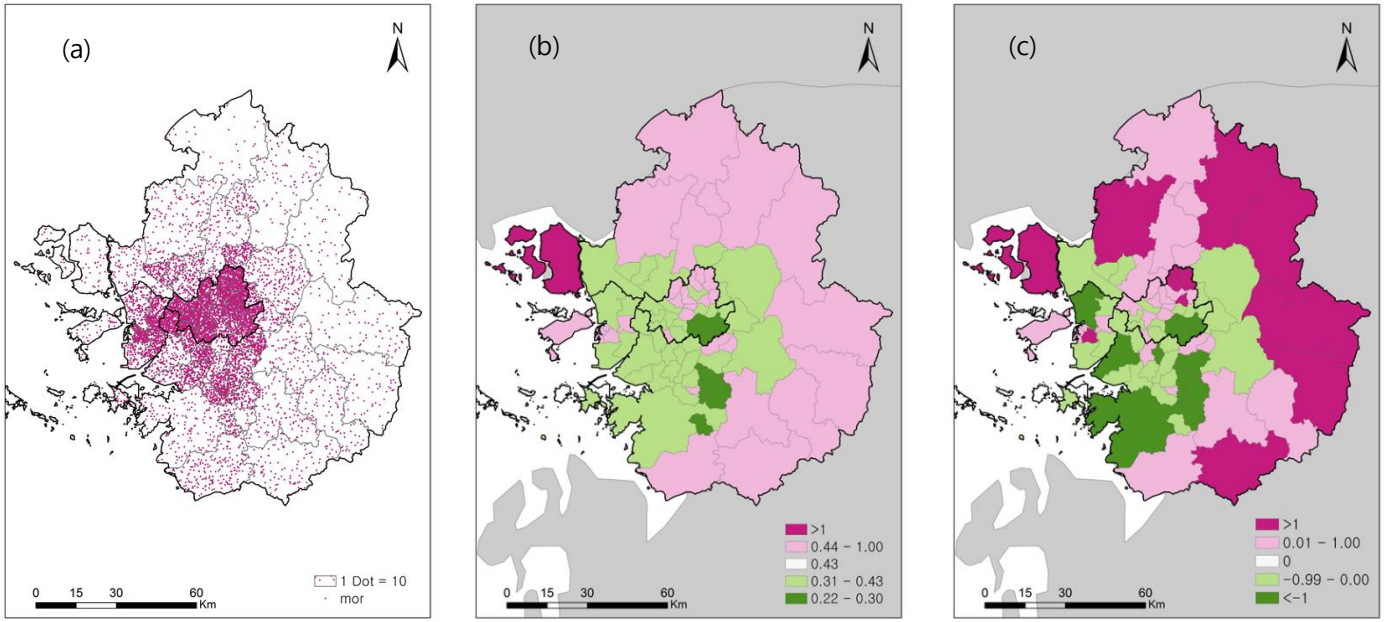


Figure 3-7 The spatial distribution of the non-accident mortality: (a) the mortality cases; (b) the mortality rate; (c) the mortality concentration degree

3.3.3 Detection of the co-pattern between man-made IMS with DI value and the non-accident mortality

Bivariate spatial association measure Lee's L was used to seek the relationship between man-made IMS and health-threatening risk. During the calculation process, man-made IMS was set as X and the DI value and the SSD of mortality were Y.

3.3.3.1 The detection of co-pattern between IMS ratio and DI values

When analyzing the bivariate spatial autocorrelation between IMS ratio and the DI values, two spatial units, the 1Km by 1Km grid and *Dong* unit were both analyzed. The global Lee's L shows the global spatial dependence between two variables, while the local Lee's L_i indicates the bivariate heterogeneity. High L values mean high spatial dependence between two variables, and high L_i values mean high positive bivariate spatial associations (high X with high Y, and vice versa).

The value of global Lee's L of man-made IMS and DI in the grid unit is 0.52 (with $p<0.000$), showing there is a relatively high positive spatial dependence between the two variables, and indicating the co-distribution of two variables. The scatterplot of L_i between two variables is shown as Figure3-8. High-high means high value of man-made IMS and the high value of DI, shown in deep pink, is mostly distributed in high-developed regions, such as most regions of Seoul, and some regions of Incheon and

Gyeonggi-do around Seoul. High-low means there is a high property of man-made IMS, but there is a low DI, indicating the environment is more comfortable for human. These regions are shown in pink, which are distributed in some regions of Jung-gu of Incheon, Dongducheon-si, Dongan-gu of Anyang-si, and so on. The low-high indicates although there is little man-made IMS, there is high DI, which is colored in light green, and Anseong-si, Icheon-si, Gwangju-si and Yeoju-si have some regions with this characteristics. And low-low means low man-made IMS area with low DI, and these regions are mainly distributed in the peripheries of Gyeonggi-do, such as Yeoncheon-gun, Pocheon-si, Gapyeong-gun, Yangpyeong-gun and so on.

We can see the bivariate spatial association pattern between man-made IMS and DI in the *Dong* unit from the Figure 3-8 (b). In the *Dong* unit, the global Lee's L is 0.48 (with $p<0.000$), and there are less hot spots and more cold spots. Similarly, regions that have high man-made IMS with high DI value, are distributed mainly in Seoul, like Seongdong-gu, Gangnam-gu, Songpa-gu, Yeongdeungpo-gu, Yangcheon-gu and so on. Others like Bupyeong-gu and Nam-gu of Incheon, Wonmi-gu of Bucheon-si also display a high density of man-made IMS with high DI values. The majority of peripheries of Gyeonggi-do shows low man-made IMS with low DI values, indicating a livable environment.

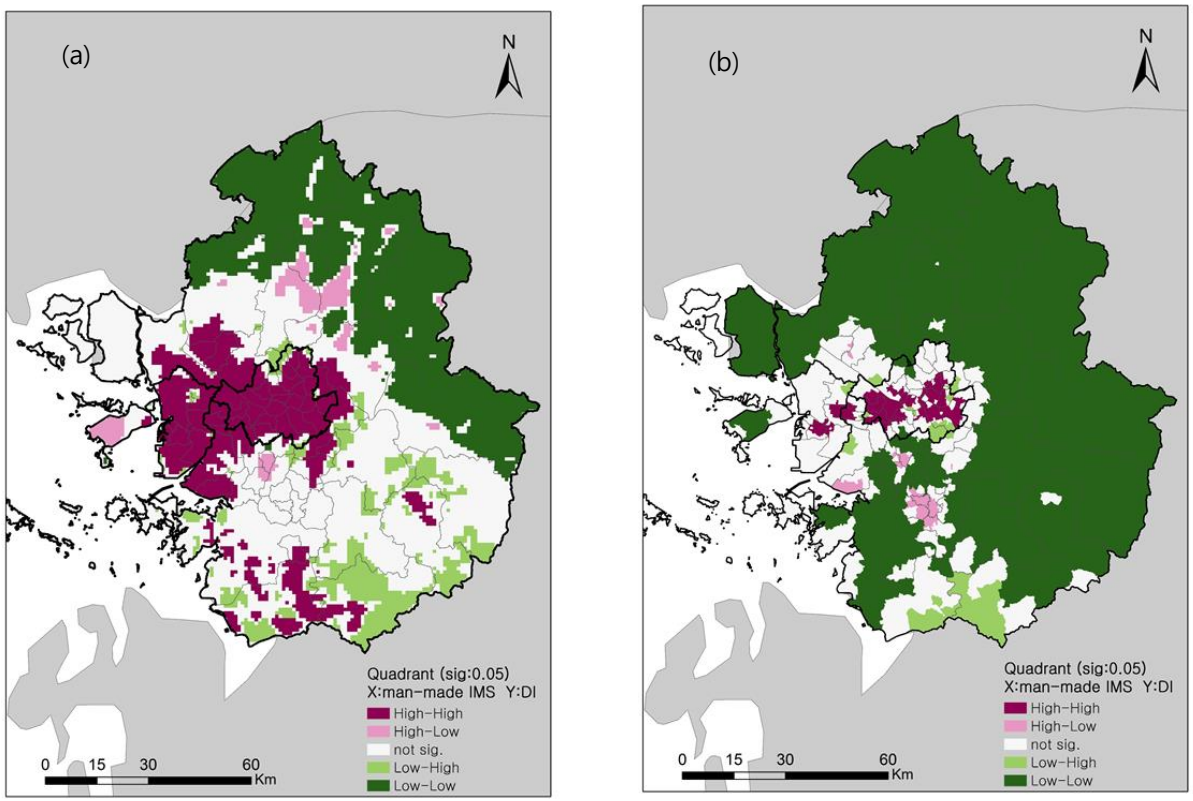


Figure 3-8 The scatterplot of local L_i between the man-made IMS and the DI value: (a) 1Km by 1Km grid unit; (b) *Dong* unit

Table 3-4 Rankings of the local L_i in different quadrants (X: ratio of man-made IMS; Y: DI)

High-High				High-Low			
Rank	*Si-Gun-Gu*	*Eup-Myeon-Dong*	Local L_i	Rank	*Si-Gun-Gu*	*Eup-Myeon-Dong*	Local L_i
1	Yeongdeungpo-gu	Mullae-dong	1.78	1	Anyang-si	Hogye 2-dong	-1.45
2	Yeongdeungpo-gu	Yeongdeungpo-dong	1.77	2	Anyang-si	Pyeongan-dong	-1.39
3	Yeongdeungpo-gu	Dorim-dong	1.66	3	Anyang-si	Bukgye-dong	-1.28
4	Gangseo-gu	Hwagok 8-dong	1.63	4	Anyang-si	Sinchon-dong	-1.27
5	Yeongdeungpo-gu	Singil 3-dong	1.60	5	Anyang-si	Hogye 2-dong	-1.24
Low-High				Low-Low			
Rank	*Si-Gun-Gu*	*Eup-Myeon-Dong*	Local L_i	Rank	*Si-Gun-Gu*	*Eup-Myeon-Dong*	Local L_i
1	Anseong-si	Miyang-myeon	-0.82	1	Pocheon-si	Idong-myeon	5.32
2	Anseong-si	Seoun-myeon	-0.70	2	Gapyeong-gun	Buk-myeon	5.24
3	Seochok-gu	Naegok-dong	-0.66	3	Pocheon-si	Ildong-yeon	2.03
4	Anseong-si	Gongdo-eup	-0.61	4	Pocheon-si	Yeongbuk-myeon	5.02
5	Anseong-si	Anseong 2-dong	-0.58	5	Gapyeong-gun	Jojong-myeon	4.94

3.3.3.2 The detection of co-pattern between IMS ratio and the non-accident mortality

When analyzing the bivariate spatial autocorrelation between IMS ratio and the non-accident mortality, *Si-Gun-Gu* administrative unit has been used.

The global Lee's L between the man-made IMS ratio and the concentration degree of non-accident mortality is -0.08, with a p value less than 0.001. The result shows that there is little correlation between the man-made IMS ratio and the non-accident mortality.

From the spatial distribution of the Local Lee's L_i shown as Figure3-9, we can see that some regions in Seoul show a high man-made IMS area ratio with a high concentration degree of non-accident mortality. Many regions also show a high man-made IMS area ratio with low non-accident mortality concentration degree, so as Suwon-si. Hwaseong- si, Yongin-si, Ansan-si are mainly distributed with cold spots with a low man-made IMS area ratio with a low concentration of non-accident mortality. The peripheral Gyeonggi-do mainly shows a low proportion of man-made IMS with a high concentration of non-accident mortality.

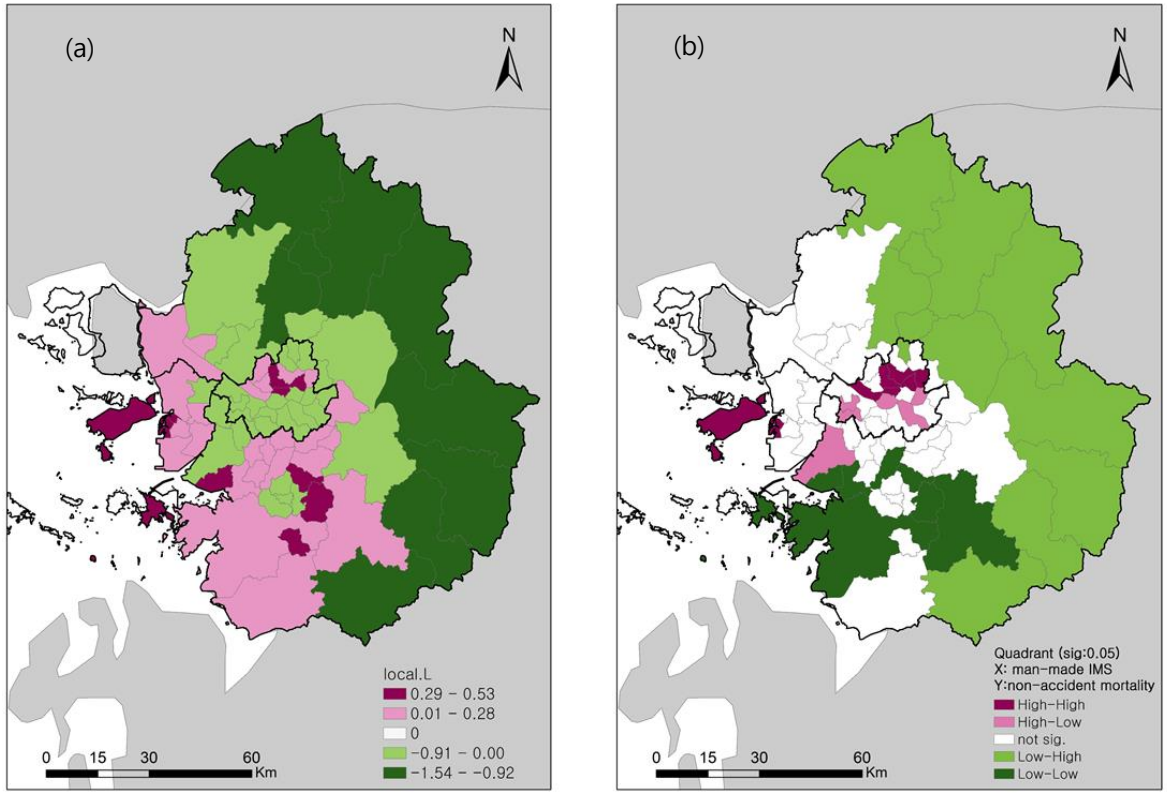


Figure 3-9 The spatial distribution of local L_i between the man-made IMS and the non-accident mortality concentration degree: (a) distribution of local L_i; (b) scatterplot map

Table 3-5 Rankings of the local L_i in different quadrants

(X: ratio of man-made IMS; Y: concentration degree of non-accident mortality)

High-High			High-Low		
Rank	*Si-Gun-Gu*	Local L_i	Rank	*Si-Gun-Gu*	Local L_i
1	Jung-gu of Incheon	0.40	1	Yangcheon-gu	-0.48
2	DonGdaemun-gu	0.37	2	Gangnam-gu	-1.47
3	Jung-gu	0.37	3	Gwangjin-gu	-0.39
4	Jongno-gu	0.34	4	Yongsan-gu	-0.34
5	Jungnang-gu	0.28	5	Guro-gu	-0.30
Low-High			Low-Low		
Rank	*Si-Gun-Gu*	Local L_i	Rank	*Si-Gun-Gu*	Local L_i
1	Gapyeong-gun	-1.54	1	Giheung-gu of Yongin-si	0.52
2	Yangcheon-gun	-1.45	2	Suji-gu of Yongin-si	0.43
3	Dongducheon-si	-1.41	3	Danwon-gu of Ansan-si	0.40
4	Yeoju-si	-4.10	4	Sangnok-gu of Ansan-si	0.27
5	Pocheoin-so	-1.02	5	Cheoin-gu of Yongin-si	0.24

3.4 Discussions and conclusions

Analyzing man-made IMS helps us to understand urban development and human activity (Zhang and Huang, 2018). However, IMS can also give rise to a more serious heat problem in urban regions including the heat island effect, which can threaten human health in various aspects. In this study, man-made IMS has been seen as possible health-threatening anthropogenic factors and have been extracted using remote sensing data. On the other hand, DI and non-accident mortality have been used as a proxy of possible threats to human health. Local correlations between the factors have also been analyzed. The main results are as follows:

(1) The total man-made IMS in the SMR is 2,675.3 km^2, which covers about 22.4% of the total area, while the proportion of man-made IMS is 63.3% for Seoul, 34.7% for Incheon, and 18.8% for Gyeonggi-do. For the grid unit, the spatial distribution pattern of man-made IMS in Seoul is more centralized, while those in Incheon and Gyeonggi-do are decentralized. The hot spots of man-made IMS mainly are distributed in the majority regions of Seoul, Incheon, and some surrounding regions of Seoul in Gyeonggi-do such as Suwon-si and so on. In the *Dong* unit, 66.1% of the regions have a man-made IMS area ratio above 50% and 28.3% above 90%. The units with a high man-made IMS area ratio are mainly distributed in Nam-gu, Dong-gu of Incheon, Yeongdeungpo-gu, Guro-gu, and Yangcheon-gu of Seoul, and Paldal-gu of Suwon-si and so on. The

regions with a man-made IMS area ratio above 90% have a discrete distribution pattern.

(2) The spatial distribution of DI has also been analyzed in the 1Km by 1Km grid and *Dong* administrative unit. For both spatial units, the general trend of DI distribution has a high DI value concentrations in Seoul. For the grid unit, most regions of Seoul and Incheon (except for Ganghwa-gun and Jung-gu), and some regions in Gyeonggi-do, such as Hwaseong-si, Pyeongtaek-si and so on have high DI values, while the north-east regions of Gyeonggi-do, such as Yeoncheon-gun, Pocheon-si etc. have low DI values. The hot spots of DI values are mainly concentrated in Seoul and some regions in Gyeonggi-do like Pyeongtaek-si, Ansan-si and Icheon-si, while regions in northeastern Gyeonggi-do like Yeoncheon-gun, Pocheon-si, Gapyeong-gun etc. are distributed with cold spots. For the *Dong* unit, the hot spots of man-made IMS are concentrated in Seoul, while majority of Gyeonggi-do are distributed with cold spots.

(3) The value of global Lee's L of man-made IMS and DI in the grid unit is 0.52, indicating a high positive spatial dependence between the two variables. The hot spots are mainly distributed in Seoul and some surrounding regions of Seoul, while the cold spots are mainly distributed in the peripheral Gyeonggi-do. For the *Dong* unit, the global Lee's L is 0.48, which shows a high positive bivariate spatial autocorrelation between the two variables. However, the majority of Gyeonggi-do shows a low

methods in the previous studies, this study was proposed. The purpose of this study is to detect the spatial co-pattern between NTL and the prevalence of breast cancer, different with the preexist studies, the human exposure to NTL, which was used to modified with the residential buildings have been used in this study, which is expected to respect for the exposure to NTL more accurately.

There are mainly four parts in this study. Besides the background introduction in the first section, I demonstrated the data and methods used in this study in the second section. In the third section, I analyzed the spatial distribution of the human exposure to NTL, and the prevalence rate of breast cancer, and demonstrated the results of the co-pattern between two variables. And in the fourth section, I concluded this study and addressed some discussions.

4.2 Materials and Methods

4.2.1 Data

The study takes Seoul Metropolitan Region as the study area, which includes Seoul, Incheon and Gyeonggi-do. Here, *Si-Gun-Gu* administrative unit is used as the basic unit. Except for the distant islands, there are 78 spatial units in total.

Data used in this study are the NTL remote sensing data to qualify people's exposure to NTL, the building data to identify the residential

buildings, and the morbidity of breast cancer in each spatial unit to evaluate its relationship with NTL.

4.2.1.1 NTL data

In 2011, NASA and NOAA launched the first Visible Infrared Imaging Radiometer Suite (VIIRS) instrument carried on the Suomi National Polar Partnership (SNPP). The VIIRS collects low light imaging data and study shows that VIIRS data has improvements over the U.S. Air Force Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS)(Elvidge *et al.*, 2013). Version 1 suite of average radiance composite images of NTL are being produced from VIIRS Day/Night Band (DNB) by the Earth Observation Group. Before averaging, the DNB data is conducted to exclude the impact of stray light, lightning, lunar illumination and cloud cover. The version 1 monthly series data has two different configurations: *vcm* and *vcmsl*, between which *vcm* excludes any data impacted by stray light and has a higher quality (Mills *et al.*, 2013). Monthly average composite of VIIRS DNB data from Apirl 2012 to December 2015 with a spatial resolution of 15 arc seconds (about 500m) was downloaded from National Center for Environmental Information of National Oceanic Atmospheric Administration (NOAA) (https://ngdc.noaa.gov/eog/viirs/download_dnb_composites.html). Due to the bad quality of monthly average composites of May, June and July in 2012, 2013, 2014, and 2015 respectively, we excluded these data, and got

33 scenes of NTL imagery data in total. All the images were converted to the WGS 1984 UTM zone 52N and resampled into 500m. In order to evaluate the long-term exposure to NTL, we take the average NTL value from Apirl 2012 to December 2015 to respect for the long-term average NTL, as Figure.4-1.

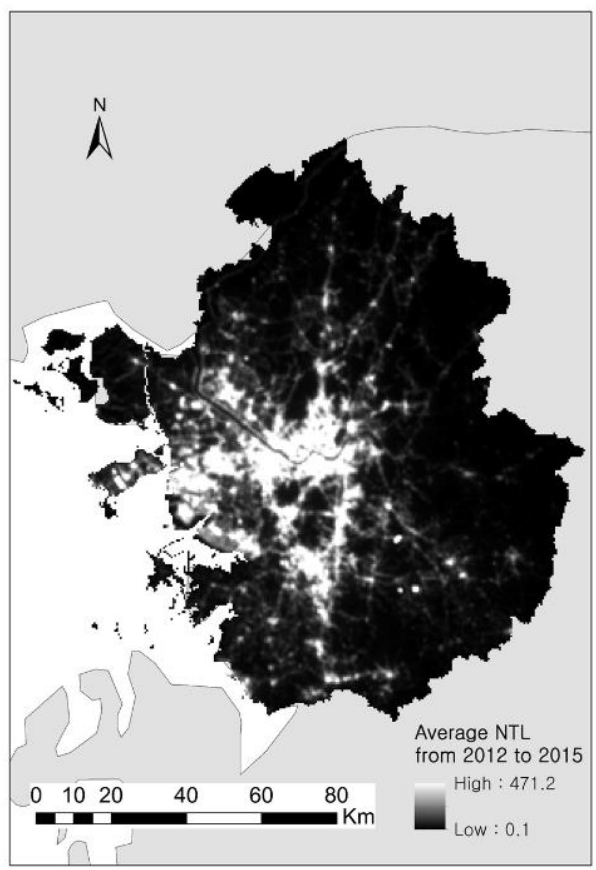


Figure 4- 1 Average NTL, 2012~2015

4.2.1.2 Building use data

In order to evaluate the people's exposure to NTL, building use data were used to identify the residential buildings. Ministry of Public Administration and Security of Korea is producing the Information Map (digital map) of the Road Name Address in order to promote the cognition and application of the Road Name Address. In the digital map of the Road Name Address, there are several layers such as building, building cluster, road interval, and actual width of road and so on. Shape files with UTM-K coordinate and Bessel ellipsoid in each Si-Do administrative unit are available. The layer of building includes information of each building like building use, number of floors, and date of construction and so on. By utilizing the building layer data derived from the Road Name Address system, we can identify the residential buildings in each study unit. In this study, building use data of 2015 were obtained from the Road Name Address System(https://www.juso.go.kr/addrlink/devLayerRequestWrite.do) and residential buildings are extracted in each *Si-Gun-Gu* unit, as shown in Figure 4-2.

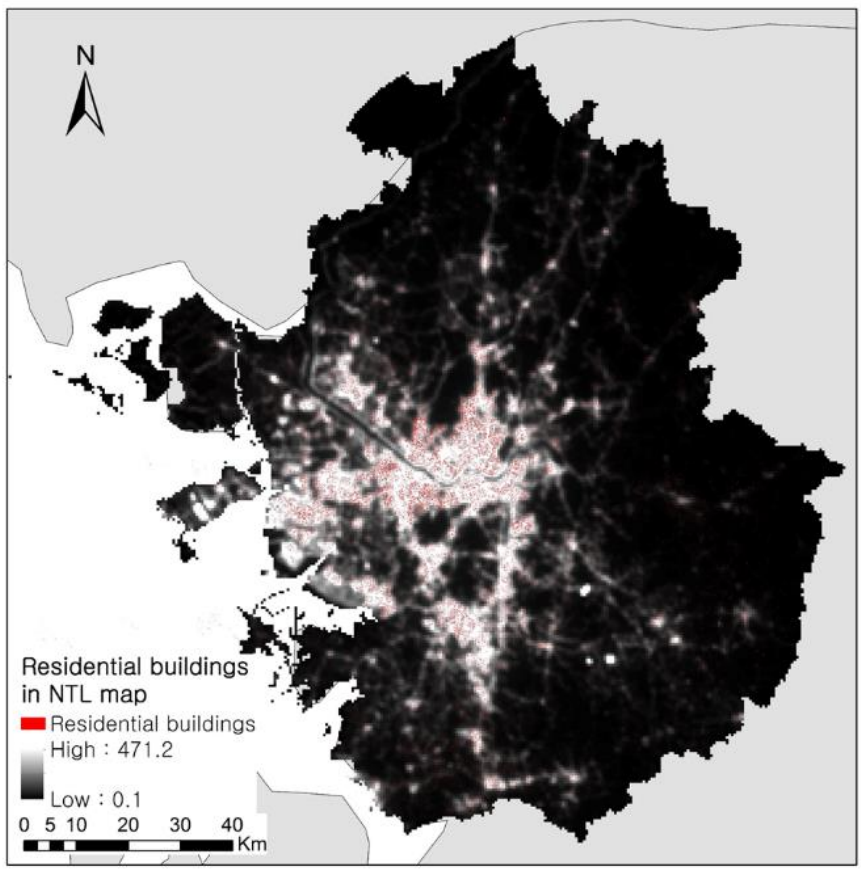


Figure 4-2 Residential buildings in NTL

4.2.1.3 Breast cancer data

Judgement results of breast cancer checkup in each *Si-Gun-Gu* administrative unit are provided from 2010 to 2016 in Korea Statistical Information Service (KOSIS, http://kosis.kr/). Yearly value of normal case, benign disease case, breast cancer suspicion case, defer judgement case, and existing breast cancer case are included. Here, we choose preexistent breast cancer cases of each year to represent for the breast cancer cases. Data from 2012 to 2015 were used, and prevalence rate of each year, and the increase of prevalence rate of breast cancer were calculated.

4.2.2 Methods

4.2.2.1 Prevalence rate

For a given disease, prevalence means the total number of existing cases within a specific time period, which can provides a summary of the current burden of the interested disease within the population (Waller and Gotway, 2004).

The formula of the prevalence rate is as follows:

$$p_i = \frac{pop_i^{case}}{pop_i^{total}}$$

Where p_i means the prevalence rate of the disease in region i; pop_i^{case} means the existing case number of the disease in region i, and pop_i^{total} means the total population of the observed group in region i. Here, we use the number of the existing breast cancer cases in each region as pop_i^{case}, and the total women

population of each region as pop_i^{total}, to calculate the breast cancer prevalence rate in each study region.

4.3 Analytical results

4.3.1 Human exposure to NTL

The average image of NTL from 2012 to 2015 were generated to signify long-term NTL. The exposure of residential buildings to NTL has been analyzed by extracting the residential buildings in the building use data and overlaying them with NTL images, which was used to represent the human exposure to NTL. The residential buildings' exposure to NTL in each *Si-Gun-Gu* unit have been extracted, and the average value of the residential buildings' exposure to NTL in each unit was used for the human exposure to NTL in this unit.

As shown in Table 4-1, the region with the highest human exposure to NTL is Jung-gu in Seoul, which is 107.27 nanoWatts/cm^2/sr averagely, while the lowest human exposure is in Gapyeong-gun in Gyeonggi-do, the value of which is only 3.49 nanoWatts/cm^2/sr. We can see that Seoul, Incheon (except for Ganghwa-gun), and some surrounding cities of Seoul in Gyeonggi-do such as Uiijeongbu-si, Goyang-si, Bucheon-si, Gwangmyeong-si, Anyang-si, Gunpo-si, Seongnam-si, Suwon-si and so on are showing high human exposure to NTL. The result shows a high

concentration of human exposure to NTL with Seoul as the center. Similarly, the result of the cluster detection of human exposure to NTL also shows that the hot spots, which means the high human exposure to NTL clusters, are mainly distributed in regions of Seoul, and Jung-gu and Nam-gu of Incheon. The low human exposure to NTL values are mainly concentrated in the peripheries of Geyeonggi-do.

Table 4-1 Rank of human exposure to NTL (unit: nanoWatts/cm^2/sr)

Rank (Top)	*Si-Gun-Gu*	Human exposure to NTL	Rank (Bottom)	*Si-Gun-Gu*	Human exposure to NTL
1	Jung-gu	107.27	1	Gapyeong-gun	3.49
2	Paldal-gu of Suwon-si	73.22	2	Yeoju-si	3.59
3	Yeongdeungpo-gu	72.92	3	Yangpyeong-gun	3.62
4	Gwangjin-gu	65.89	4	Yeoncheon-gun	4.61
5	Wonmi-gu of Becheon-si	64.42	5	Ganghwa-gun	5.38

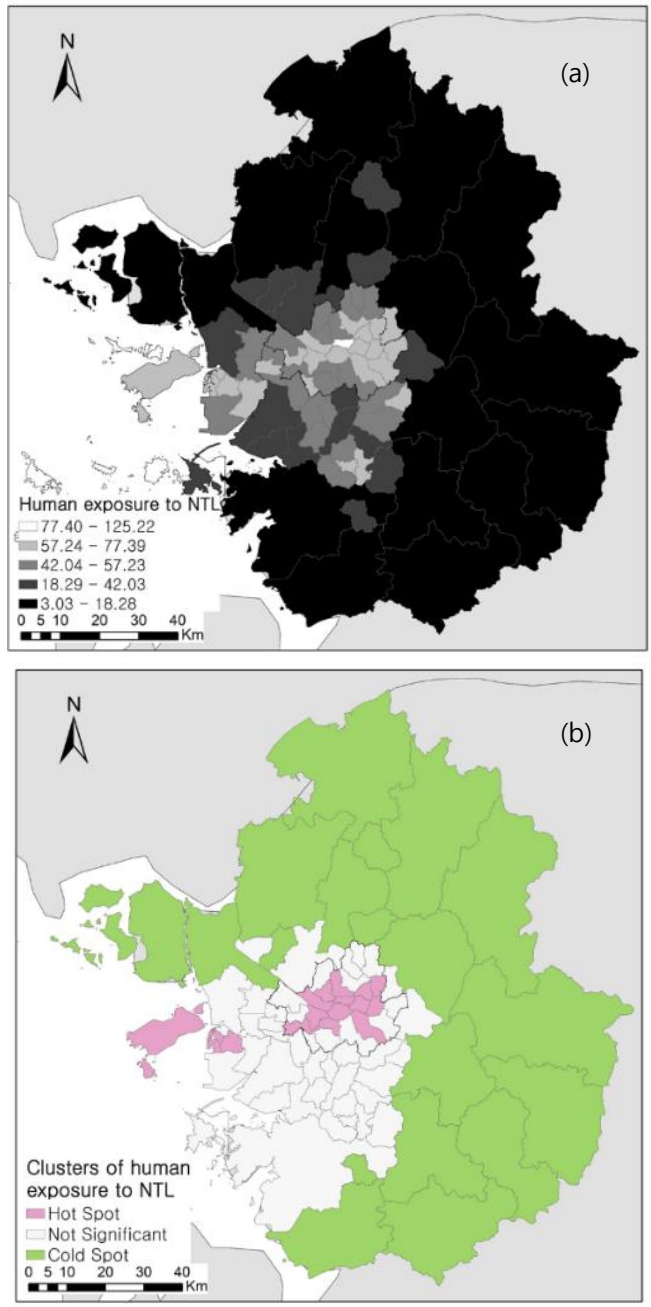


Figure 4-3 The spatial distribution of human exposure to NTL in *Si-Gun-Gu* unit: (a) human exposure to NTL; (b) cluster detection of human exposure to NTL

4.3.2 Spatiotemporal distribution of the breast cancer prevalence rate

Prevalence refers to the total existing cases over a specific time frame and indicates the current burden of the disease within the population (Waller and Gotway, 2004). Here, we use the breast cancer prevalence rate in women in each unit to represent for the current burden of breast cancer. The proportion of existing breast cancer to the total population of women in each unit was calculated from 2012 to 2015, which is the prevalence rate in each year. The prevalence change rate is also calculated by dividing the prevalence rate of 2012 by using that of 2015 to show the breast cancer prevalence rate trend.

By analyzing the breast cancer prevalence in 2015, we can see that regions with the high prevalence rate are mainly concentrated in Seoul, Goyang-si, Gimpo-si, Bucheon-si and Suwon-si, among which Jangan-gu of Suwon-si, Jongno-gu, Yongsan-gu of Seoul, Paldal-gu of Suwon-si and Nam-gu of Incheon show highest breast cancer prevalence rates in 2015, while Gyeyang-gu, Seo-gu and Yeonsu-gu of Incheon, Siheung-si, Ansan-si show the lowest values, as shown in Table 4-2.

The clusters of breast cancer prevalence in 2015 were detected, with the hot spot showing the clusters of high breast cancer prevalence rate and cold spot showing the clusters of low ones. Result shows that hot spots are maily distributed in northwestern regions of Seoul, such as

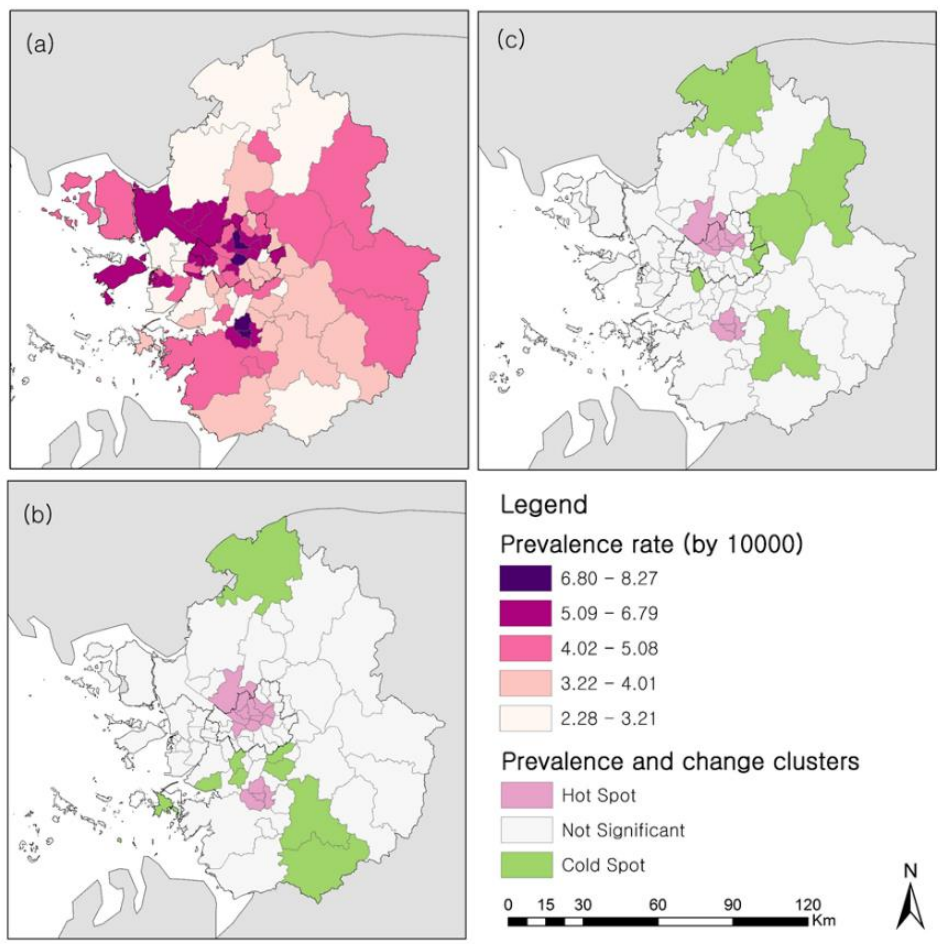


Figure 4-4 The spatial distribution of (a) the breast cancer prevalence rate in 2015; (b) the clusters of the breast cancer prevalence rate in 2015; and (c) the clusters of the breast cancer prevalence rate change from 2012 to 2015

Table 4-2 Rankings of the breast cancer prevalence rate

(BCPR, per 10000 women)

Rank (Top)	*Si-Gun-Gu*	BCPR	Rank (Bottom)	*Si-Gun-Gu*	BCPR
1	Jangan-gu of Suwon-si	8.27	1	Gyeyang-gu of Inchoen	2.28
2	Jongno-gu	8.03	2	Siheung-si	2.32
3	Yongsan-gu	7.74	3	Ansan-si	2.38
4	Paldal-gu of Suwon-si	7.47	4	Seo-gu of Incheon	2.48
5	Nam-gu of Incheon	6.79	5	Yeonsu-gu of Incheon	2.57

Table 4-3 Rankings of the breast cancer prevalence rate increase

(BCPRI, per 10000 women)

Rank (Top)	*Si-Gun-Gu*	BCPRI	Rank (Bottom)	*Si-Gun-Gu*	BCPRI
1	Jongno-gu	4.53	1	Hanam-si	-0.5
2	Nam-gu of Incheon	3.96	2	Manan-gu of Anyang-si	-0.08
3	Jangan-gu of Suwon-si	3.92	3	Anseong-si	0.02
4	Paldal-gu of Suwon-si	3.69	4	Yeoncheon-gun	0.02
5	Yongsan-gu	3.57	5	Gwangju-si	0.13

Yeongdeungpo-gu, Yongsan-gu, Mapo-gu, Seongdong-gu, Seodaemun-gu, Dongdaemun-gu, Jung-gu, Jongno-gu, Eunpyeong-gu and Seongbuk-gu. Also, Deogyang-gu of Goyang-si and Suwon-si also show the clusters of the high breast cancer prevalence in 2015. On the other side, regions like Yeoncheon-gun, Danwon-gu of Ansan-si, Manan-gu of Anyang-si, Gunpo-si, Sujeong-gu and Bundang-gu of Seongnam-si, Cheoin-gu of Yongin-si, and Anseong-si show the clusters of low breast cancer prevalence rate in 2015.

From 2012 to 2015, Jongno-gu of Seoul, Nam-gu of Incheon, Jangan-gu and Paldal-gu of Suwon-si, Yongsan-gu of Seoul show the highest increase of breast cancer prevalence rate throughout the whole regions, while Hanam-si and Manan-gu show a decreasing trend, together with Anseong-si, Yeoncheon-gun and Gwangju-si, they rank the bottom in the breast cancer prevalence increase, as described in table 4-3.

The change clusters of breast cancer prevalence rate from 2012 to 2015 were also analyzed. Here, hot spot means the clusters of high increase of breast cancer prevalence rate, which mainly distribute in Deogyang-gu in Goyang-si, Jongno-gu, Jung-gu, Seodaemun-gu, Dongdaemun-gu etc. in Seoul and Suwon-si, while cold spot indicates the concentration of the low change of breast cancer prevalence rate, which mainy distribute in Yeoncheon-gun, Gapyeong-gun, Namyangju-si and so on, as shown in Figure 4-4 and Table 4-3.

4.3.3 Detection of the co-pattern between human exposure to NTL and the prevalence of breast cancer

4.3.3.1 Local Pearson's r_i between human exposure to NTL and the breast cancer prevalence in 2015

Local Pearson's r_i between human exposure to NTL and the breast cancer prevalence rate of 2015, as well as that between human exposure to NTL and the breast cancer prevalence change between 2012 to 2015 were analyzed to figure out the local correlation between human exposure to NTL and the prevalence rate of breast cancer.

The global Pearson's r between human exposure to NTL and the breast cancer prevalence rate is 0.39, showing a moderate positive correlation between the two variables, which indicates that regions with high human exposure to NTL have the high possibility of a high breast cancer prevalence rate, and vice versa. Spatial distribution of local Pearson's r_i was shown in Figure 4-5. 46 regions of the total 78 regions (59%) show a local positive correlation between human exposure to NTL and breast cancer prevalence, such as most regions of Gyeonggi-do and Incheon, and some regions in Seoul. However, 32 regions (41%) display a negative correlation between the two variables, and these regions mainly distributed in Goyang-ri, Uijeongbu-si, Namyangju-si, Seongnam-si and

so on. The results of the cluster detection of local Pearson's r_i between the two variables show high positive values of r_i cluster in Yeoncheon-gun, Dongducheon-si, Anseong-si, Jangan-gu and Paldal-gu in Suwon-si, and some regions in Seoul like Yongsan-gu, Mapo-gu and so on.

Among those regions with positive local Pearson's r_i there are 23 regions show a high human exposure with high breast cancer prevalence in 2015, shown in Figure 4-6, such as regions in Seoul like Jung-gu, Jongno-gu and so on, and Suwon-si, and Jung-gu and Nam-gu in Incheon. While other 23 regions show low human exposure to NTL and low breast cancer rate in 2015. At the same time, 32 regions show a negative correlation, 22 regions show a high human exposure to NTL and a relatively low breast cancer prevalence rate in 2015. Those regions are mainly concentrated in Seoul and its surrounding regions. The regions with low human exposure to NTL but high breast cancer prevalence in 2015 were also detected, which mainly distributed in Gyeonggi-do, such as Gimpo-si, Goyang-si, Uijeongbu-si, Namyangju-si, Yeoju-si, Osan-si and Gwacheon-si etc.

Table 4-4 Rankings of the local r_i in different quadrants (X: human exposure to NTL; Y: BCPR)

	High-High			High-Low	
Rank	*Si-Gun-Gu*	Local r_i	Rank	*Si-Gun-Gu*	Local r_i
1	Jung-gu	5.33	1	Gwangjin-gu	-1.11
2	Yongsan-gu	2.75	2	Jungwon-gu of Seongnam-si	-0.96
3	Paldal-gu of Suwon-si	2.54	3	Guri-si	-0.70
4	Nam-gu of Incheon	1.48	4	Gangnam-gu	-0.63
5	Jongno-gu	1.47	5	Bupyeong-gu of Incheon	-0.54
	Low-High			Low-Low	
Rank	*Si-Gun-Gu*	Local r_i	Rank	*Si-Gun-Gu*	Local r_i
1	Gimpo-si	-0.67	1	Anseong-si	2.59
2	Deogyang-gu of Goyang-si	-0.57	2	Pocheon-si	1.68
3	Ilsandong-gu of Goyang-si	-0.28	3	Yeoncheon-gun	1.63
4	Gwacheon-si	-0.22	4	Paju-si	1.54
5	Ilsanseo-gu of Goyang-si	-0.1	5	Yangju-si	1.10

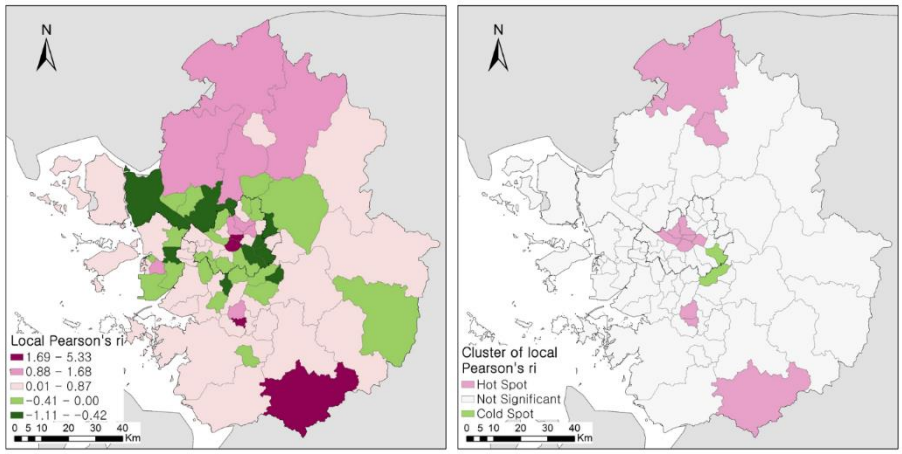


Figure 4-5 The spatial distribution and the cluster distribution of local Pearson's r_i between human exposure to NTL and the breast cancer prevalence

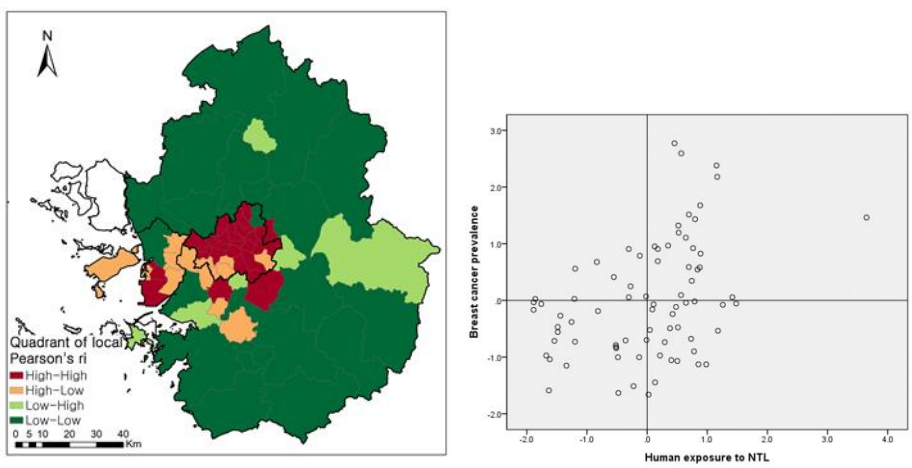


Figure 4-6 The scatterplot of local Pearson's r_i between human exposure to NTL and the breast cancer prevalence

4.3.3.2 Local Pearson's r_i between human exposure to NTL and the breast cancer prevalence change from 2012 to 2015

In order to detect the co-relation between human exposure to NTL and the breast cancer prevalence further, the bivariate correlation between human exposure to NTL and the breast cancer prevalence change from 2012 to 2015 have been studied.

The global Pearson's r between the two variables is 0.31, showing a moderate positive correlation, which indicates that in the regions where human are highly exposed to NTL, there are possibilities that the breast cancer prevalence rate will increase, and vice versa. The spatial distribution of local Pearson's r_i is shown as Figure 4-7. There are 44 regions out of 78 regions (56%) have a positive correlative value between human exposure to NTL and the breast cancer prevalence increase. Regions in Seoul (Jung-gu, Jongno-gu, Yongsan-gu etc.), and those in Incheon (Seo-gu, Jung-gu etc.) as well as regions in Gyeonggi-do (Paju-si, Pocheon-si, Hanam-si etc.) are included. While there are 34 regions (44%), such as Gangnam-gu, Seocho-gu etc. in Seoul, Ganghwa-gun etc. in Incheon and regions like Yeoju-si, Yangju-si and Guri-si in Gyeonggi-do, indicate a negative correlation between human exposure to NTL and the breast cancer prevalence rate increase. By detecting the cluster of local Pearson's r_i between two variables, hot spots, which mean the cluster of high positive correlation, concentrate in Seodaemun-gu, Jung-gu, Jongno- gu

Table 4-5 Rankings of the local ri in different quadrants (X: human exposure to NTL; Y: BCPRI)

	High-High			High-Low	
Rank	*Si-Gun-Gu*	Local r_i	Rank	*Si-Gun-Gu*	Local r_i
1	Jung-gu	4.88	1	Gwangjin-gu	-1.34
2	Paldal-gu of Suwon-si	2.11	2	Gangnam-gu	-1.19
3	Yongsan-gu	1.96	3	Guri-si	-1.12
4	Namgu of Incheon	1.82	4	Dong-gu of Incheon	-0.98
5	Jongno-gu	1.47	5	Jungwon-gu of Seongnam-si	-0.61
	Low-High			Low-Low	
Rank	*Si-Gun-Gu*	Local r_i	Rank	*Si-Gun-Gu*	Local r_i
1	Yeoju-si	-2.00	1	Yeoncheon-gun	2.69
2	Ganghwa-gun of Incheon	-1.47	2	Anseong-si	2.62
3	Deogyang-gu of Goyang-si	-0.88	3	Pocheon-si	2.34
4	Gwacheon-si	-0.73	4	Yangpyeong-gun	2.13
5	Yangju-si	-0.71	5	Gwangju-si	1.89

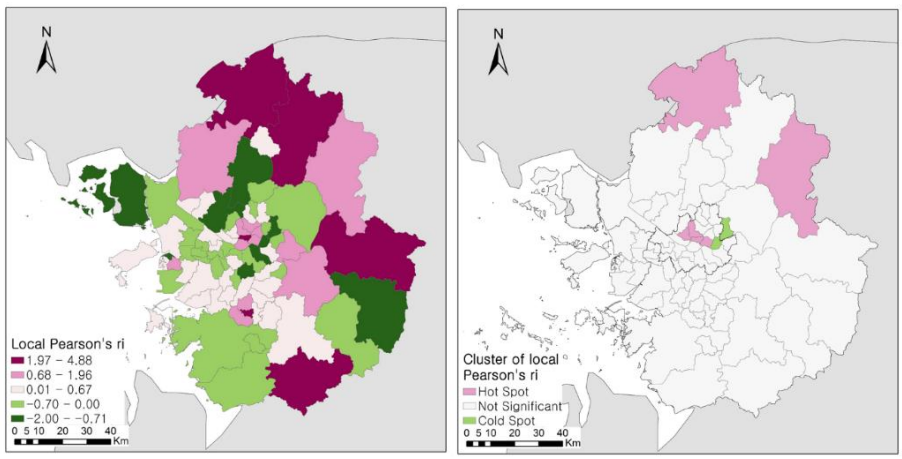


Figure 4-7 The spatial distribution and the cluster detection of local Pearson's r_i between human exposure to NTL and the breast cancer prevalence change

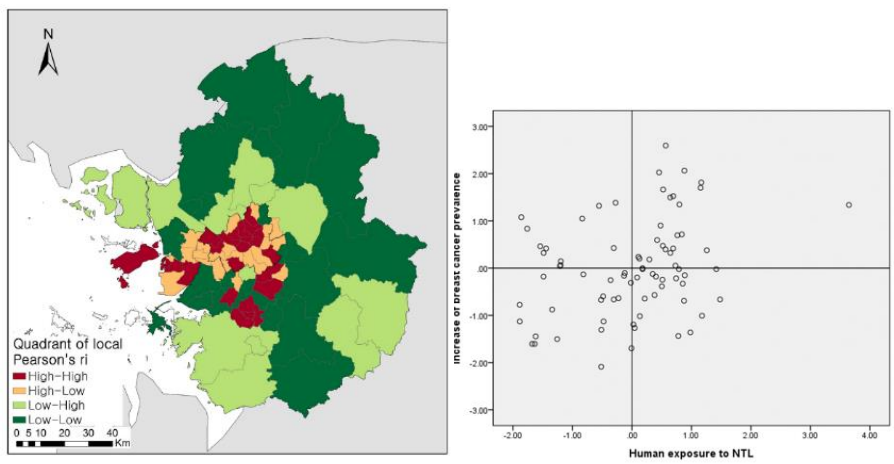


Figure 4-8 The scatterplot of Pearson's r_i between human exposure to NTL and the breast cancer prevalence change

and Seongdong-gu in Seoul, as well as Yeoncheon-gun and Gapyeong-gun in Gyeonggi-do. While the cold spots, indicating the cluster of high negative correlation, distribute in Guri-si and Gwangjin-gu in Seoul.

Among those 44 regions with positive correlation between human exposure to NTL and the increase of breast cancer prevalence, there are 23 regions showing the high human exposure to NTL with high increase of breast cancer prevalence, which are mainly concentrated in Seoul, like Gwanak-gu, Gangseo-gu, Yangcheon-gu etc., as well as Suwon-si, and a few regions of Incheon and Seongnam-si. At the same time, 21 regions show low human exposure to NTL with low increase of breast cancer rate, like Seo-gu of Incheon, Paju-si, Anseong-si etc. On the other hand, 22 regions have a high human exposure to NTL but low breast cancer prevalence increase, and these regions mainly distributed in Seoul and surrounding regions, as shown in Figure 4-8. And low human exposure to NTL but with high breast cancer prevalence increase occur in 12 regions such as Gwanghua-gu of Incheon, Gimpo-si, Uijeongbu-si, Namyangju-si and so on.

4.4 Discussions and conclusions

This study used NTL as a proxy of human activities, studied the spatial distribution of residential building modified human exposure to NTL during 2012 to 2015, and the prevalence rate of breast cancer in 2015 and the prevalence change during the study period. In addition, I analyzed the spatial co-pattern between residential building modified human exposure to NTL and the prevalence and the prevalence change of breast cancer. The main results are as follows:

(1) Local Pearson's r_i can detect the "point-to-point" correlation between two variables well.

(2) High human exposure to NTL are mainly concentrated in Seoul, Incheon (except for Ganghwa-gun), and the surrounding cities of Seoul in Gyeonggi-do. The region with the highest human exposure to NTL is Jung-gu in Seoul, which is 125.22 nanoWatts/cm^2/sr averagely, while the lowest human exposure is in Ganghwa-gun of Incheon, the value of which is only 3.03 nanoWatts/cm^2/sr. Hot spots of human exposure to NTL distribute in regions of Seoul like Jung-gu, Jongno-gu etc, and Jung-gu and Nam-gu of Incheon, while the cold spots are mainly distributed in the peripheries of Gyeonggi-do.

(3) The results of spatiotemporal distribution of breast cancer prevalence show that regions with the high prevalence rate are mainly concentrated

in Seoul, Goyang-si, Gimpo-si, Bucheon-si and Suwon-si, among which Jongno-gu and Yeongdeungpo-gu of Seoul, as well as Jangan-gu and Paldal-gu of Suwon show highest breast cancer prevalence rates in 2015. Hot spots of breast cancer prevalence rate in 2015 are mainly distributed in northwestern regions of Seoul, while some regions in Ansan-si, Anyang-si, Gunpo-si, -si, Yongin-si, and Anseong-si show the clusters of low breast cancer prevalence rate in 2015. On the other hand, hot spots of breast cancer prevalence increase are detected in some regions of Seoul, Goyang-si and Suwon-si, while cold spots are detected in regions in Gyeonggi-do like Gwangmyeong-si, Yeoncheon-gun etc.

(4) The global Pearson's r between human exposure to NTL and the breast cancer prevalence rate in 2015 is 0.39, showing there is a moderate positive correlation between two variables. For the local Pearson's r_i analysis, 46 regions shows positive correlation. There are 23 regions such as Jung-gu, Jongno-gu in Seoul, Suwon-si etc. show a high human exposure with high breast cancer prevalence in 2015. While other 23 regions show low human exposure to NTL and low breast cancer rate in 2015. At the same time, there are 32 regions (mainly concentrated in Seoul and its surrounding regions) show a negative correlation, among which 22 regions show a high human exposure to NTL and relatively low breast cancer prevalence rate in 2015. While some regions of Gyeonggi-do like Gimpo-si, Goyang-si etc. show low human exposure to NTL but high breast cancer prevalence in 2015.

(5) The global Pearson's r between human exposure to NTL and the increase of breast cancer prevalence is 0.31, showing a moderate positive correlation. For the local Pearson's r_i analysis, there are 44 regions have a positive correlative value between two variables, within which there are 23 regions like Gwanak-gu, Gangseo-gu in Seoul and Suwon-si etc. showing the high human exposure to NTL with high increase of breast cancer prevalence. 21 regions like Seo-gu of Incheon, Paju-si, Anseong-si etc. show low human exposure to NTL with low increase of breast cancer rate. On the other hand, there are 34 regions showing negative correlation. Among those 22 regions (Seoul and surrounding regions) have a high human exposure to NTL but low breast cancer prevalence increase while low human exposure to NTL but with high breast cancer prevalence increase occur in 12 regions such as Ganghwa-gu of Incheon, Gimpo-si, Uijeongbu-si, Namyangju-si and so on.

The results implied the concentration of human exposure to NTL in Seoul and its surrounding regions, including some cities in Gyeonggi-do such as Suwon-si, Seongnam-si and so on. Inevitably, the human exposures to NTL are also different and even have the same degree of NTL. For example, the frequency outdoor activity, and the setting and quality of the curtain and its quality all have influences on the exposure to NTL. However, this study only consider the average NTL degree of the residential buildings. Therefore, the results of this study may be biased due to the different behaviors.

The relatively high spatial association between exposure to NTL and the prevalence and prevalence change of breast cancer indicate that we should consider some actions to reduce the unnecessary light at night and night-shift work, and rouse the awareness of the disadvantages of excessive exposure to NTL.

Chapter 5. The Influences of Health-threatening Factors on Human Health

5.1 Introduction

After the detection of the co-pattern between health-threatening anthropogenic factors and health outcomes, models are designed to detect the influences of health-threatening anthropogenic factors on human health. The influences have been examined on the IMS ratio on the DI value, the IMS ratio on the non-accident mortality and NTL on the breast cancer prevalence rate, respectively. The health outcomes have been considered as dependent variables, while the possible influencing factors according to previous studies have been selected as the independent variables. Global and local regression models are all designed to figure out the influences both in the whole region and in each local unit level. The influences of health-threatening anthropogenic factors are analyzed in detail according to the objectives of this study.

5.2 The influence of man-made IMS on the DI value

In order to figure out the influence of the man-made IMS on the DI value, the DI value is define by the temperature and relative humidity. Therefore, factors impacting temperature or relative humidity could be seen as an influencing factor of the DI value. Studies show that the

vegetation has a relationship with the reduction in the land surface temperature (Kaufmann *et al.*, 2003; Yue *et al.*, 2007), and also vegetation has a relationship with precipitation (Yang *et al.*, 1997). Elevation and population are also an obvious influencing factor of temperature. Therefore, we take the DI value as the dependent value, while vegetation, elevation and population are selected as the independent variables.

5.2.1 Model description

Ordinary least square (OLS) and SEM are used to evaluate the influences of the vegetation, elevation, population and man-made IMS to the DI value. While the data and sources are as shown as table 5-1. In addition, GWR is used to calculate the influence of the factors in every unit, which can explain the spatial heterogeneity of factors and processes in detail.

Here, the IMS ratio is calculated by dividing the total region of every unit by the man-made IMS area in this region, and NDVI and the DEM are the average value of the unit.

Table 5- 1 Data and sources

	Category	Data	Source	Date
Dependent variable	DI value	○	○	○
Independent variable	IMS ratio	○	○	○
	Vegetation	NDVI	MOD13A1 https://lpdaac.usgs.gov/	2015.01.01-2016.12.31
	Elevation	DEM	SRTM DEM http://srtm.csi.cgiar.org/	2016
	Population	Population	KOSIS http://kosis.kr/	2015.01.01-2016.12.31

5.2.2 Analytical results

By calculation, the IMS ratio, NDVI, and population all show positive correlation with the DI value, while DEM shows a negative correlation with DI value. Due to the merits of SEM on processing spatial data, a lower Akaike information criterion (AIC) value is demonstrated. For the OLS model and SEM, the IMS ratio and population show significant positive correlation, while DEM show significant negative correlation, indicating the increase of IMS ratio and population will increase the DI value, and increase of DEM will reduce the DI value, and *vice versa*. Therefore, the IMS ratio has the strongest influence on the DI value, with the highest standardized coefficient.

Table 5- 2 Results of OLS and SEM

	OLS			SEM		
	Coefficient	Standardized Coefficient	*p*-value	Coefficient	Standardized Coefficient	*p*-value
Intercept	22.503		0.000	22.851		0.000
IMS ratio	0.714***	0.461***	0.009	0.361**	0.233**	0.042
NDVI	0.046	0.044	0.963	0.010	0.002	0.987
DEM	-0.002*	-0.292*	0.076	-0.001*	-0.194*	0.097
Population	0.000***	0.193***	0.023	0.000*	0.085*	0.064
R^2	0.62			0.88		
AIC	150.218			86.463		

* p value≤0.1; ** p value≤0.05; ***p value≤0.01

The distribution of the coefficient of the IMS ratio in Figure 5-1 shows most regions have a positive correlation between the IMS ratio and DI values, expect for some regions in Seoul, like Gangnam-gu, Seocho-gu *etc.* and some regions in Gyeonggi-do like Seongnam-si, Gwangju-si and so on.

For the distribution of the highest influencing factors on the DI value shown in Figure 5-2, we can see that the IMS ratio is the strongest influencing factor on the DI value in some regions of Seoul, Incheon, and the majority regions of north Gyeonggi-do. However, the DI values in the other regions are mostly influenced by the average DEM.

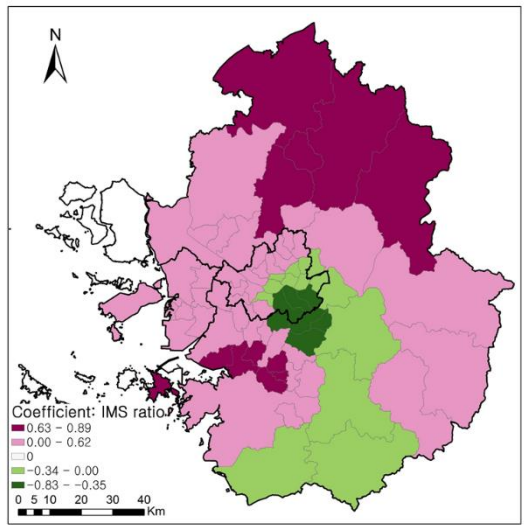


Figure 5- 1 The spatial distribution of regression coefficients of the IMS ratio

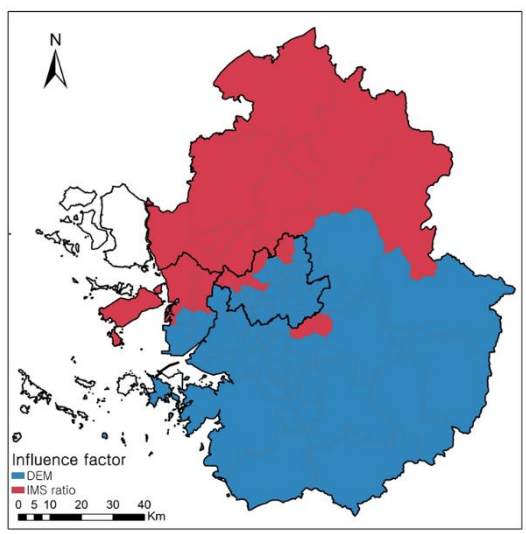


Figure 5- 2 The spatial distribution of strongest influencing factors on the DI value

5.3 The influence of man-made IMS on the non-accident mortality

Man-made IMS is assumed to influence non-accident mortality indirectly due to the positive influence of man-made IMS to temperature (Gabriel and Endlicher, 2011). In addition, other factors influence on the mortality, such as the age structure (Jeong and Jun, 2013), and the accessibility to the medical organizations. Therefore, this study designates non-accident mortality as the dependent variable, and the impervious land ratio, the elderly and the accessibility to the medical organizations as the independent variables.

5.3.1 Model description

OLS and SEM are used to evaluate the influences of the man-made impervious surface, elderly ratio and the accessibility to the hospital. While the data and sources are as shown as Table 5-2. In addition, GWR is used to calculate the influence of the factors in every unit, which can explain the spatial heterogeneity of factors and processes in detail.

Here, population above 65 years old are seen as elderly population (Jeon, 2017), and the elder SSD is calculated to represent for the concentration of elders. For the accessibility of the medical organization, hospitals (병원) and the general hospitals (종합병원) in the SMR are selected as the destination, while the residential buildings in each unit are

taken as the origins, the average value of the OD distance in every unit was taken as the distance to hospital.

Table 5- 3 Data and sources

	Category	Data	Source	Date
Dependent variable	Non-accident mortality	Mortality data	Korea National Statistical Office http://kostat.go.kr/	2015.01.01-2016.12.31
Independent variable	IMS ratio	○	○	○
	Elder SSD	Elder	Korea National Statistical Office http://kostat.go.kr/	2015
	Hospital accessibility	Hospital	Local data https://www.localdata.kr/	2015.01.01-present

5.3.2 Analytical results

SEM also performs better than OLS according to the results. By calculation, the IMS ratio, elderly SSD and the distance to the hospital all a have positive correlation with non-accident mortality. However, the IMS ratio and the distance to the hospital are only significant in the OLS model, and among those factors, elderly SSD has the strongest influence on non-mortality.

We can figure out the model fitness of every unit, and the strongest influence factors of every unit for the results of GWR, as seen in Figure 5-3 and Figure 5-4.

Table 5- 4 Results of OLS and SEM

	OLS			SEM		
	Coefficient	Standardized Coefficient	*p*-value	Coefficient	Standardized Coefficient	*p*-value
Intercept	-0.2160		0.044	0.1500		0.3656
IMS ratio	0.0030*	0.1220*	0.100	0.0006	0.0980	0.3189
Elder SSD	0.7720***	0.7750***	0.000	0.8670***	0.8630***	0.0000
Hosdis	0.0000*	0.1400*	0.092	0.0000	-0.0250	0.7800
R^2	0.621			0.769		
AIC	147.313			121.270		

* p value≤0.1; ** p value≤0.05; ***p value≤0.01

For the GWR results of coefficient distribution of IMS ratio, we can see that for the most regions of Seoul and Gyeonggi-do, there is a negative correlation between the IMS ratio and the non-accident mortality, while Incheon, Gimpo-si, Hwaseong-si *etc.* show positive correlation between the two factors. Regions like Incheon, Siheung-si, Bucheon-si and so on show stronger correlation between them.

For the strongest influencing factors, we can see that for the majority of Gyeonggi-do, the elder concentration is the strongest influencing center of non-accident mortality, while most regions in Seoul, Seo-gu, Gaeyang-gu of Incheon, and Gimpo-si, Bucheon-si, Guangmyeong-si *etc.* have the distance to the hospital as the strongest influencing factor of non-accident mortality, which is the most concentrate region of hospitals. However, Seocho-gu, Gangnam-gu of Seoul and Gwacheon-si, Seongnam-si *etc.* show the IMS ratio as the highest influencing factors of non-accident mortality.

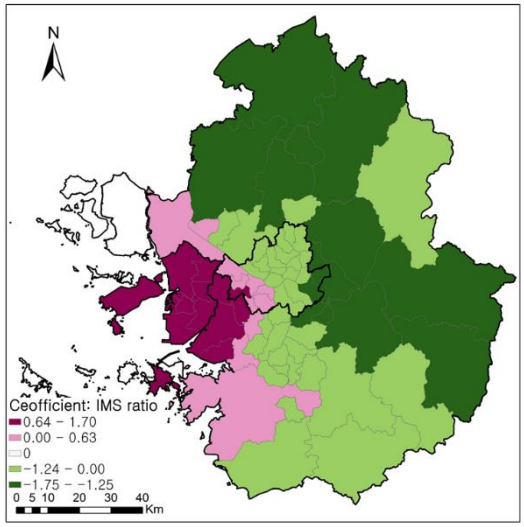


Figure 5- 3 The spatial distribution of regression coefficients of the IMS ratio

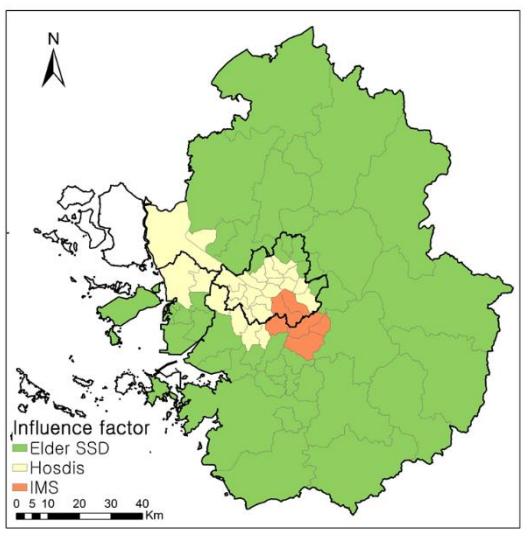


Figure 5- 4 The spatial distribution of the strongest influencing factors on the non-accident mortality

5.4 The influence of NTL on the breast cancer prevalence

One study shows that there are one million new breast cancer cases, which occupies about 18% of all female-type cancers (McPherson *et al.*, 2000). Breast cancer is the primary cause of cancer death among women globally (Bray *et al.*, 2004). Studies show that excessive exposure to NTL can increase breast cancer risk (Davis *et al.*, 2001; Kloog *et al.*, 2008; Pauley, 2004; Stevens, 1987; Stevens, 2005). In order to figure out the influence of NTL exposure to breast cancer, I apply the prevalence rate of breast cancer (BCPR) as the dependent variable, and possible influencing factors such as fertility, per capita GRDP (Kloog *et al.*, 2008), obesity (Kim, 2008), female education lever, and alcohol intake as independent variables.

5.4.1 Model description

This study also adopted OLS and SEM as the global regression model, while GWR as the local regression model to estimate the influence of NTL to the breast cancer prevalence rate. The data and sources are as shown as Table 5-5. The mean_NTL is the mean nighttime light in each unit from 2012 to 2015, in that there may be a time lag between the exposures to NTL and the inducing of breast cancer, and also there will be an incubation period of breast cancer. The fertility is the birth rate of every unit by dividing the fertile women by the number of births. While the obesity,

education, and workrate of female are considered. Thereinto, education is the rate of females with education degree higher than bachelor.

Table 5- 5 Data and sources

	Category	Data	Source	Date
Dependent variable	BCPR	Breast cancer	Korea National Statistical Office http://kostat.go.kr/	2015
Independent variable	Mean_NTL	NTL	○	○
	Fertility	Birth rate		2015
	per cap GRDP	GRDP		2015
	Alcohol	High Alcohol intake rate	Korea National Statistical Office	2015
	Obesity	Obesity rate	http://kostat.go.kr/	2015
	Education	Education degree		2015
	Workrate	Worker rate		2015

5.4.2 Analytical results

We can see that for the mean_NTL, per capita GRDP, and alcohol have a positive correlation with breast cancer prevalence, while fertility, obesity, education, and workrate all have a negative correlation with the breast cancer prevalence rate. Among which, mean_NTL, fertility and alcohol have significant correlation with breast cancer prevalence rate for both OLS and

SEM, and mean_NTL and fertility have stronger influences. SEM also shows a better performance than OLS.

Table 5- 6 Results of OLS and SEM

	OLS			SEM		
	Coefficient	Standardized Coefficient	*p*-value	Coefficient	Standardized Coefficient	*p*-value
Intercept	58.910**		0.037	58.97**		0.011
Mean_NTL	0.129**	0.356**	0.016	0.142***	0.460***	0.006
Fertility	-22.555***	-0.511***	0.002	-21.793***	-0.434***	0.001
Per capita GRDP	0.106	0.651	0.259	0.088	0.506	0.302
Alcohol	1.285**	0.334**	0.043	1.005*	0.206*	0.067
Obesity	-0.688	-0.155	0.455	-14.377	-0.161	0.394
Education	-7.525	-0.082	0.688	-13.419	-0.273	0.405
Workrate	-15.853	-0.650	0.273	0.055	-0.589	0.311
R^2	0.508			0.523		
AIC	151.569			148.812		

For GWR, we can see that the coefficient of mean_NTL is higher than 0.31, with an increasing trend from the southern to the northern region, with the highest coefficient distributed in the Pocheon-si, Yeonchen-gun Paju-si and so on. For the strongest influencing factors in each unit, we find that most regions of Seoul, Incheon, and norther Gyeonggi-do have

NTL as the strongest influencing factor for the breast cancer prevalence rate, while other regions have fertility as the strongest influencing factors.

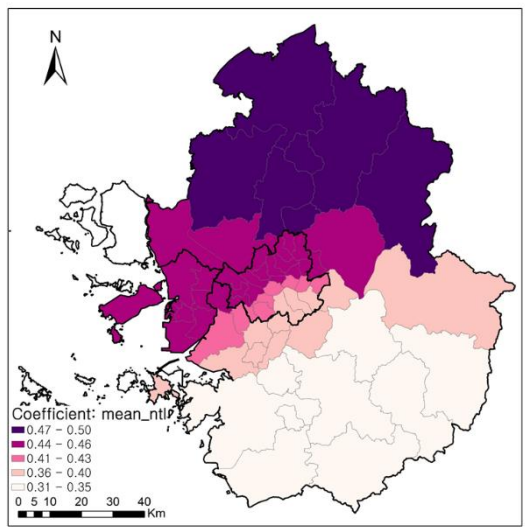


Figure 5- 5 The spatial distribution of regression coefficients of the mean NTL

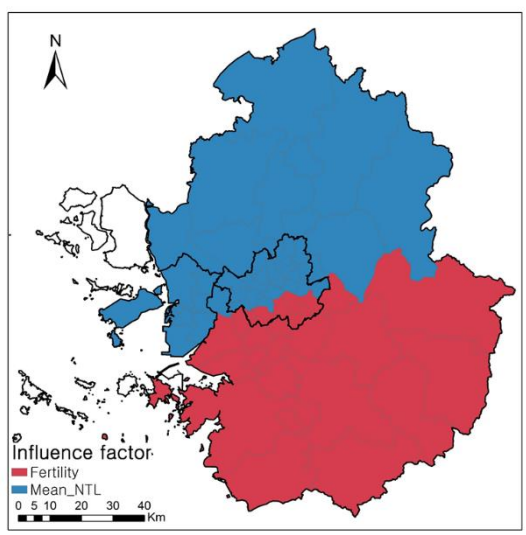


Figure 5- 6 Distribution of the strongest influencing factors on the breast cancer prevalence rate

5.5 Discussions and conclusions

Regression modeling is conducted to detect the influences of health-threatening anthropogenic factors on human health, after the detection of co-pattern between the two factors. For the global regression model, OLS and SEM and the local regression model, GWR are performed to identify the influencing factors, in order to emphasize the influences of health-threatening anthropogenic factors on human health. The results for anthropogenic factors in GWR are discussed. And the results are as follows.

First, the IMS ratio, DEM and population have significant influences on the modelling of influencing factors of DI value. However, the IMS ratio and population have a positive correlation, and DEM has a negative correlation. Among the influencing factors, the IMS ratio has the strongest influence on the DI value. Through the results of GWR, we can see that a majority of the regions have a positive correlation between the two factors, and the regions in Gyeonggi-do like Yeoncheon-gu, Pocheon-si and so on have the highest coefficient. Some regions of northern Seoul, Incheon and most regions of northern Gyeonggi-do show that the strongest influencing factors of DI values are the IMS ratio.

Second, the IMS ratio, elder concentration and the distance to the hospital have a positive correlation for the influencing factors of non-accident mortality. However, the strongest influencing factors of non-accident mortality is the concentration of the elderly. For the results of

GWR, we can see that only some regions in Incheon, and Gyeonggi-do such as Gimpo-si, Bucheon-si and so on show a positive correlation between the IMS ratio and non-accident mortality, while other regions are almost negative. The strongest influencing factors of non-accident mortality are the elder concentration in most Gyeonggi-do regions, the distance to the hospital in Ginpo-si and some regions in Incheon and Seoul and so on, and the IMS ratio in Gangnam-gu, Seocho-gu and some regions in Gyeonggi-do like Seongnam-si and Gwacheon-si.

Third, mean_NTL, and alcohol show a significant positive relationship with for the modeling of the influencing factors of the breast cancer prevalence rate, with fertility showing a negative correlation. The mean-NTL also shows a strong influence. For GWR, we can see an increasing trend of influences of mean_NTL from the southern to the northern regions. For most regions of Seoul, Incheon, and northern Gyeonggi-do, the highest influencing factors of the breast cancer prevalence rate is the NTL.

The results show a positive correlation of health-threatening anthropogenic factors and the potential health outcomes, and a strong influence on health outcomes. Therefore, some measures should be considered to reduce the health-threatening anthropogenic factors for reducing their influences on human health. Vegetation and other measures can be adopted to reduce the heat effect for man-made IMS.

Publicity in the harm of NTL and the forced outage of some types of unnecessary artificial lightings like some commercial lighting can be considered for reducing NTL.

Chapter 6. General Conclusions and Discussions

The consistent increase of the population in urban regions can scale up the adverse impact of humans to the natural environment, including bad air quality, ocean pollution, water security, to name a few. Those environmental problems can and already have counteracted our health, by increasing the risks of morbidity and mortality. Two major challenges are how to quantify and figure out the distribution of humans and human activities. Remote sensing has the capability to give insight into biotic and abiotic components on the earth for large temporal and spatial scales. Since the availability of remote sensing in scientific research conducted on the application of remote sensing on human health. However, there are limitations of the existence studies, including the lack of anthropogenic factors extraction and analysis on remote sensing and the infrequent use of spatial statistical measures when quantifying the relationship between the factors derived from remote sensing and possible health outcomes.

Based on previous research and limitations, I extracted the possible health-threatening anthropogenic factors in remote sensing and their possible health outcomes, and I analyzed the spatial distribution of the two variables and the spatial association between the two variables. And the main conclusions are as follows:

(1) When application remote sensing on the health-related analysis, remote sensing techniques can be utilized to extract the possible health-threatening anthropogenic factors, like the man-made IMS, the artificial light at night, to name a few, instead of extracting natural factors like vegetation, wetness *etc.* only.

(2) Remote sensing can give insight into the spatial and temporal distribution of the possible health-threatening anthropogenic factors in large spatial and temporal scales of less cost and time consumption.

(3) Instead of the traditional proportion and statistical methods used in the previous studies, row-standardized proportion method can give a more accurate result by taking the different total populations in different regions. And the utilization of spatial statistical methods, and the local statistical methods can acquire more rigorous and detailed relation between the variables.

(4) Man-made IMS have shown a high concentration in majority regions of Seoul, Incheon and some surrounding regions of Seoul in Gyeonggi-do. Similarly, the high DI values are also have a trend of concentration with Seoul as the center, while the spatial distribution pattern of non-accident mortality show low value cluster in Seoul and high value clusters in the peripheral Gyeonggi-do. The results proved a high spatial bivariate association between the man-made IMS ratio and the DI values, showing high ratio of man-made IMS accompanies with high DI

values, while it showed little relation between the man-made IMS ratio and the non-accident mortality, , which assumed to be associated with the high income and the better accessibility to the health care.

(5) For the exposure to NTL, the residential building modified exposure to NTL are mainly concentrated in majority regions in Seoul and Incheon, and some major cities in Gyeonggi-do. The distribution pattern of breast cancer prevalence rate also tend to concentrated in and around Seoul. The results have shown high spatial bivariate correlation between the exposure to NTL and the breast cancer prevalence rate, and a high local correlation between the NTL and the prevalence rate change.

(6) Models of the influencing factors on DI value, non-accident mortality and the breast cancer prevalence rate are designed to figure out the influences of the health-threatening anthropogenic factors. Results show that the man-made IMS and NTL have strong influence on the increase of DI value and the breast cancer prevalence, indicating the influences of the health-threatening anthropogenic factors on human health. For the non-accident mortality, although man-made IMS shows positive correlation, the strongest influencing factors is the elder concentration degree, instead of the man-made IMS.

There are some implications and limitations of this study. The main contribution of this study is the innovation of application of remote sensing to health-related research by extracting the anthropogenic factors

instead of the natural factors in remotely sensed data. With the merits of remotely sensed data, we can derive the possible health-threatening anthropogenic factors in large temporal and spatial scale, which can be of great help for decision-making. The utilization of spatial association measures and the new data standardization method are also advantages of this study. What's more, the human activities extracted in remote sensing are very practical and closely relate to our practical life, which can give us more practical implications in our life. The results are also expected to improve the environment impacted on human activities and give the basis for decision-making.

There are also limitations in this study. This study only researched on the correlation between the two variables, with little considering on the cause and effect. The pathogenic factors of the diseases are greatly diverse, therefore, analyzing the pathogenic factors of the diseases is beyond the scope of this study. However, the analysis on the correlation between the two variables also proved to be valuable. Another limitation is the limitation of the research data, including the lack of data in smaller spatial unit and more accurate information of the patients.

This study has analyzed the spatial and temporal distribution of the possible health-threatening anthropogenic factors in remote sensing and their possible threats, with the spatial association measure between the two variables. Therefore, there are some follow-up studies that can be

conducted. For example, based on the distribution of health-threatening anthropogenic factors in remote sensing, the risk map of human activities to human health can be generated. Other factors can also be introduced for more detailed analysis of the pathogenic factors of diseases. These studies are all expected to contribute in building a better human-environment relationship.

Bibliography

Albert, D. P., Gesler, W. M., and Levergood, B., 2003, *Spatial Analysis, GIS and Remote Sensing: Applications in the Health Sciences*: Chelsea: Ann Arbor Press.

Amaral, S., Câmara, G., Monteiro, A. M. V., Quintanilha, J. A., and Elvidge, C. D., 2005, Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data, *Computers, Environment and Urban Systems,* 29(2), 179-195.

Amiri, R., Weng, Q., Alimohammadi, A., and Alavipanah, S. K., 2009, Spatial–temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz Urban Area, Iran, *Remote sensing of environment,* 113(12), 2606-2617.

Andrews, G. J. and Moon, G., 2005, Space, place, and the evidence base: Part I—an introduction to health geography, *Worldviews on Evidence-Based Nursing,* 2(2), 55-62.

Angiuli, E. and Trianni, G., 2014, Urban mapping in Landsat images based on normalized difference spectral vector, *IEEE Geoscience and Remote Sensing Letters,* 11(3), 661-665.

Angouridakis, V. and Makrogiannis, T., 1982, The discomfort-index in Thessaloniki, Greece, *International Journal of biometeorology,* 26(1), 53-59.

Anselin, L., 1994, Exploratory spatial data analysis and geographic information systems, *New Tools for Spatial Analysis,* 17, 45-54.

Anselin, L., 1995, Local indicators of spatial association—LISA, *Geographical Analysis,* 27(2), 93-115.

Anselin, L., 1998, Exploratory spatial data analysis in a geocomputational environment, Longley, P. A., Brooks, S. M., McDonnell, R., and MacMillan B., eds., *Geocomputation, A Primer,* New York: Wiley, 77-94.

Anselin, L., 2009, Spatial regression, Fotheringham, A. S. and Pogerson P. A., eds., *The SAGE Handbook of Spatial Analysis,* Thousand Oaks: SAGE Publications, 255-276.

Anselin, L., 2013, *Spatial Econometrics: Methods and Models*, Boston: Kluwer Academic Publishers.

Arnfield, A. J., 2003, Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island, *International Journal of Climatology,* 23(1), 1-26.

Artis, D. A. and Carnahan, W. H., 1982, Survey of emissivity variability in thermography of urban areas, *Remote Sensing of Environment,* 12(4), 313-329.

Ban, Y., Jacob, A., and Gamba, P., 2015, Spaceborne sar data for global urban mapping at 30m resolution using a Robust Urban Extractor, *ISPRS Journal of Photogrammetry and Remote Sensing,* 103, 28-37.

Barinaga, M., 1993, Satellite data rocket disease control efforts into orbit, *Science,* 261(5117), 31-33.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M., and García-Herrera, R., 2011, The hot summer of 2010: redrawing the temperature record map of europe, *Science,* 332(6026), 220-224.

Bauer, M. E., Loffelholz, B. C., and Wilson, B. 2007. Estimating and mapping impervious surface area by regression analysis of landsat imagery, Weng Q. H., ed., *Remote Sensing of Impervious Surfaces*, Boca Raton: CRC Press, 31-48.

Beck, L. R., Lobitz, B. M., and Wood, B. L., 2000, Remote sensing and human health: new sensors and new opportunities, *Emerging Infectious Diseases,* 6(3), 217-227.

Bennie, J., Davies, T. W., Duffy, J. P., Inger, R., and Gaston, K. J., 2014, Contrasting trends in light pollution across europe based on satellite observed night time lights, *Scientific Reports,* 4, 3789.

Bennie, J., Duffy, J. P., Davies, T. W., Correa-Cano, M. E., and Gaston, K. J., 2015, Global trends in exposure to light pollution in natural terrestrial ecosystems, *Remote Sensing,* 7(3), 2715-2730.

Berglund, B. E., 2003, Human impact and climate changes—synchronous events and a causal link? , *Quaternary International,* 105(1), 7-12.

Blaschke, T., 2010, Object based image analysis for remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing,* 65(1), 2-16.

Bray, F., McCarron, P., and Parkin, D. M., 2004, The Changing Global Patterns of Female Breast Cancer Incidence and Mortality, *Breast cancer research,* 6(6), 229-239.

Bridges, E. and Oldeman, L., 1999, Global assessment of human-induced soil degradation, *Arid Soil Research and Rehabilitation,* 13(4), 319-325.

Brookes, P., 1995, The use of microbial parameters in monitoring soil pollution by heavy metals, *Biology and Fertility of Soils,* 19(4), 269-279.

Brown, T., McLafferty, S., and Moon, G., 2009, *A Companion to Health and Medical Geography*, Oxford, UK: Wiley-Blackwell.

Cablk, M. and Minor, T., 2003, Detecting and discriminating impervious cover with high-resolution ikonos data using principal component analysis and

morphological operators, *International Journal of Remote Sensing,* 24(23), 4627-4645.

Cao, Q., Yu, D., Georgescu, M., Wu, J., and Wang, W., 2018, Impacts of future urban expansion on summer climate and heat-related human health in eastern china, *Environment International,* 112, 134-146.

Chenoweth, J., Hadjikakou, M., and Zoumides, C., 2014, Quantifying the human impact on water resources: a critical review of the water footprint concept, *Hydrology and Earth System Sciences,* 18(6), 2325-2342.

Choi, G., Choi, J., and Kwon, H., 2005, The Impact of high apparent temerature on the increase of summertime disease-relatred mortality in seoul: 1991-2000, *Journal of Pretive Medical and Public Health,* 38(8), 283-290.

Claborn, D. M., Masuoka, P. M., Klein, T. A., Hooper, T., Lee, A., and Andre, R. G., 2002, A cost comparison of two malaria control methods in kyunggi province, republic of korea, using remote sensing and geographic information systems, *The American Journal of Tropical Medicine and Hygiene,* 66(6), 680-685.

Clarke, J. I., Rhind, D. W., Becket, C., Wilkes, A., Sadler, G., and Short, J., 1992, Population data and global environmental change, International Social Science Council, Programme on Human Dimensions of Global Environmental Change, UNESCO, Paris, France.

Clarke, K. C., Osleeb, J. P., Sherry, J. M., Meert, J. P., and Larsson, R. W., 1991, The use of remote sensing and geographic information systems in unicef's dracunculiasis (guinea worm) eradication effort, *Preventive Veterinary Medicine,* 11(3-4), 229-235.

CLINE, B. L., 1970, New eyes for epidemiologists: aerial photography and other remote sensing techniques, *American Journal of Epidemiology,* 92(2), 85-89.

Coombes, Y. J., 1993, A Geography of the New Public Health, Ph.D. dissertation, Queen Mary University of London.

Davis, S., Mirick, D. K., and Stevens, R. G., 2001, Night shift work, light at night, and risk of breast cancer, *Journal of the National Cancer Institute,* 93(20), 1557-1562.

de Miguel, A. S., Zamorano, J., Castaño, J. G., and Pascual, S., 2014, Evolution of the energy consumed by street lighting in spain estimated with dmsp-ols data, *Journal of Quantitative Spectroscopy and Radiative Transfer,* 139, 109-117.

Dister, S. W., Fish, D., Bros, S. M., Frank, D. H., and Wood, B. L., 1997, Landscape characterization of peridomestic risk for lyme disease using satellite imagery, *The American Journal of Tropical Medicine and Hygiene,* 57(6), 687-692.

Doll, C. N., Muller, J.-P., and Morley, J. G., 2006, Mapping regional economic activity from night-time light satellite imagery, *Ecological Economics,* 57(1), 75-92.

Duncan, O. D. and Duncan, B., 1955, A methodological analysis of segregation indexes, *American Sociological Review,* 20(2), 210-217.

Ehrlich, P. R., Ehrlich, A. H., and Daily, G. C., 1993, Food security, population and environment, *Population and Development Review,* 1-32.

Elvidge, C., Sutton, P., Tuttle, B., Ghosh, T., and Baugh, K. E., 2009, Global urban mapping based on nighttime lights, *Global Mapping of Human Settlements: Experiences, datasets and prospects,* Boca Raton: CRC Press, 129-144.

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., Davis, E. R., and Davis, C. W., 1997, Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption, *International Journal of Remote Sensing,* 18(6), 1373-1379.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., and Hsu, F.-C., 2013, Why VIIRS data are superior to DMSP for mapping nighttime lights. In *Proceedings of the Asia-Pacific Advanced Network*.

Elvidge, C. D., Imhoff, M. L., Baugh, K. E., Hobson, V. R., Nelson, I., Safran, J., Dietz, J. B., and Tuttle, B. T., 2001, Night-time lights of the world: 1994–1995, *ISPRS Journal of Photogrammetry and Remote Sensing,* 56(2), 81-99.

Eriksen, M., Lebreton, L. C., Carson, H. S., Thiel, M., Moore, C. J., Borerro, J. C., Galgani, F., Ryan, P. G., and Reisser, J., 2014, Plastic pollution in the world's oceans: more than 5 trillion plastic pieces weighing over 250,000 tons afloat at sea, *PloS one,* 9(12), e111913.

Falchi, F., Cinzano, P., Duriscoe, D., Kyba, C. C., Elvidge, C. D., Baugh, K., Portnov, B. A., Rybnikova, N. A., and Furgoni, R., 2016, The new world atlas of artificial night sky brightness, *Science Advances,* 2(6), e1600377.

Falchi, F., Cinzano, P., Elvidge, C. D., Keith, D. M., and Haim, A., 2011, Limiting the impact of light pollution on human health, environment and stellar visibility, *Journal of Environmental Management,* 92(10), 2714-2722.

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., Brauer, M., Burnett, R., Cercy, K., and Charlson, F. J., 2016, Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the global burden of disease study 2015, *The Lancet,* 388(10053), 1659-1724.

Fotheringham, A. S., Brunsdon, C., and Charlton, M., 2009, Geographically weighted regression, Fotheringham A. S., Brunsdon C., Charlton M., eds, *The SAGE Handbook of Spatial Analysis,* Chechester: Wiley, 243-254.

Gabriel, K. M. and Endlicher, W. R., 2011, Urban and rural mortality rates during heat waves in berlin and brandenburg, Germany, *Environmental Pollution,* 159(8-9), 2044-2050.

Garland, R., Yang, H., Schmid, O., Rose, D., Nowak, A., Achtert, P., Wiedensohler, A., Takegawa, N., Kita, K., and Miyazaki, Y., 2008, Aerosol optical properties in a rural environment near the mega-city guangzhou, china: implications for regional air pollution, radiative forcing and remote sensing, *Atmospheric Chemistry and Physics,* 8(17), 5161-5186.

Gaston, K. J., Bennie, J., Davies, T. W., and Hopkins, J., 2013, The ecological impacts of nighttime light pollution: a mechanistic appraisal, *Biological Reviews,* 88(4), 912-927.

Gatrell, A. C. and Elliott, S. J., 2014, *Geographies of Health: An Introduction*, Chichester: Wiley Blackwell.

Gesler, W., 1986, The uses of spatial analysis in medical geography: a review, *Social Science & Medicine,* 23(10), 963-973.

Getis, A. and Ord, J. K., 1992, The analysis of spatial association by use of distance statistics, *Geographical Analysis,* 24(3), 189-206.

Gilbert, E. W., 1958, Pioneer maps of health and disease in england, *The Geographical Journal,* 124(2), 172-183.

Gim, H.-J., Ho, C.-H., Kim, J., and Lee, E. J., 2018, Urbanization may reduce the risk of frost damage to spring flowers: a case study of two shrub species in south korea, *PloS one,* 13(2), e0191428.

Goudie, A. S., 2018, *Human Impact on the Natural Environment*, Cambridge: Blackwell Pub.

Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Lee, Y., and Kumar, N., 2006, Satellite remote sensing of particulate matter and air quality assessment over global cities, *Atmospheric Environment,* 40(30), 5880-5892.

Haining, R., 1991, Bivariate correlation with spatial data, *Geographical Analysis,* 23(3), 210-227.

Haining, R., Wise, S. and Ma, J., 2000, Designing and implementing software for spatial statistical analysis in a gis environment, *Journal of Geographical Systems,* 2(3), 257-286.

Hippocrates, A. F., 2000, *On Airs, Waters, and Places*, DC Stevenson: Web Atomics.

Howe, G. 1989. Historical evolution of disease mapping in general and specifically of cancer mapping, Boyle, P., Muir, C. S., Grundamann, E., eds, *Cancer Mapping*, Berlin: Springer, 1-21.

Hubert, L. J., Golledge, R. G., Costanzo, C. M., and Gale, N., 1985, Measuring association between spatially defined variables: an alternative procedure, *Geographical Analysis,* 17(1), 36-46.

Ibald-Mulli, A., Stieber, J., Wichmann, H.-E., Koenig, W., and Peters, A., 2001, Effects of air pollution on blood pressure: a population-based approach, *American Journal of Public Health,* 91(4), 571-577.

Jeon, C., 2017, A Study on the Spatial Distribution Characteristics and Type of the Low-Income Single Elderly in Seoul, Ph.D. Dissertation, Seoul National University.

Jeong, J. E. and Jun, M. i., 2013, Spatial concentrations of the elderly and its characteristics in the Seoul Metropolitan Area, *Journal of the Korean Regional Science Association,* 29(1), 3-18.

Jiménez-Muñoz, J. C., Sobrino, J. A., Skoković, D., Mattar, C., and Cristóbal, J., 2014, Land surface temperature retrieval methods from landsat-8 thermal infrared sensor data, *IEEE Geoscience and Remote Sensing Letters,* 11(10), 1840-1843.

Kaufmann, R., Zhou, L., Myneni, R., Tucker, C., Slayback, D., Shabanov, N., and Pinzon, J., 2003, The effect of vegetation on surface temperature: a statistical analysis of ndvi and climate data, *Geophysical Research Letters,* 30(22), 2147.

Kearns, R. and Moon, G., 2002, From medical to health geography: novelty, place and theory after a decade of change, *Progress in Human Geography,* 26(5), 605-625.

Kim, H., Lee, N. and Kim, S.-N., 2018, Suburbia in Evolution: exploring polycentricity and suburban typologies in the Seoul Metropolitan Area, South Korea, *Land Use Policy,* 75, 92-101.

Kim, J., 2008, Risk Factors Affecting the Onset of Breast Cancer, Ph.D. Dissertation, Chung-ang University.

Kim, O., 2014, Identifying urban areas using time-series remote sensing techniques, *Journal of the korean Cartographic Asspociation,* 14(2), 119-126.

Kloog, I., Haim, A., Stevens, R. G., Barchana, M., and Portnov, B. A., 2008, Light at night co-distributes with incident breast but not lung cancer in the female population of israel, *Chronobiology International,* 25(1), 65-81.

Kloog, I., Haim, A., Stevens, R. G., and Portnov, B. A., 2009, Global co-distribution of light at night (lan) and cancers of prostate, colon, and lung in men, *Chronobiology International,* 26(1), 108-125.

Kloog, I., Stevens, R. G., Haim, A. and Portnov, B. A., 2010, Nighttime light level co-distributes with breast cancer incidence worldwide, *Cancer Causes & Control,* 21(12), 2059-2068.

Lee, H.-E., Lee, J., Jang, T.-W., Kim, I.-A., Park, J. and Song, J., 2018, The relationship between night work and breast cancer, *Annals of Occupational and Environmental Medicine,* 30(1), 11-17.

Lee, S.-I., 2001, Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's I, *Journal of Geographical Systems,* 3(4), 369-385.

Lee, S.-I., 2002, Spatial Association Measures for An Esda-Gis Framework: Developments, Significance Tests, and Applications to Spatio-Temporal Income Dynamics of US Labor Market Areas, 1969-1999, Ph.D. Dissertation, The Ohio State University.

Lee, S.-I., 2004, Spatial data analysis for the US regional income convergence, 1969-1999: A critical appraisal of β-convergence, *Journal of the Korean Geographical Society,* 39(2), 212-228.

Lee, S.-I., 2007, A spatial statistical approach to residential differentiation (i): developing a spatial separation measure, *Journal of the Korean Geographical Society,* 42(4), 616-631.

Lee, S.-I., 2008, A spatial statistical approach to residential differentiation (ii): exploratory spatial data analysis using a local spatial separation measure, *Journal of the Korean Geographical Society,* 43(1), 134-153.

Lee, H., Lee, S.-I. and Cho, D., 2013, Exploring the spatiality of school choice through residential mobility: a preliminary case study of elementary school students in Seoul, *Journal of the Korean Geographical Society*, 48(6), 897-913.

Lee, S.-I., 2017, Correlation and spatial autocorrelation, Shekhar, S., Xiong, H., and Zhou, X., eds, *Encyclopedia of GIS*, Cham: Springer, 360-368.

Lee, S.-I., 2018, Spatializing the Pearson's correlation coefficient: an experimental comparision of three relevant techniques, *Journal of the Korean Geographical Society,* 53(5), 761-776.

Linthicum, K. J., Bailey, C. L., Davies, F. G., and Tucker, C. J., 1987, Detection of rift valley fever viral activity in kenya by satellite remote sensing imagery, *Science,* 235(4796), 1656-1659.

Lipton, R., Gorman, D., Wieczorek, W. F., Banerjee, A., and Gruenewald, P, 2009, Geography and public health, Khosrow-pour, M., ed., *Encyclopedia of Information Science and Technology, Second Edition*, Hershey: Idea Group Reference, 1634-1645.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., and Shi, L., 2015, Social sensing: a new approach to understanding our socioeconomic environments, *Annals of the Association of American Geographers,* 105(3), 512-530.

Liu, Z., He, C., Zhang, Q., Huang, Q., and Yang, Y., 2012, Extracting the dynamics of urban expansion in China Using DMSP-OLS nighttime light data from 1992 to 2008, *Landscape and Urban Planning,* 106(1), 62-72.

Lopez, J. M. R., Heider, K., and Scheffran, J., 2017, Frontiers of urbanization: identifying and explaining urbanization hot spots in the south of Mexico City using human and remote sensing, *Applied Geography,* 79, 1-10.

Lu, D., Hetrick, S. and Moran, E., 2011, Impervious surface mapping with quickbird imagery, *International Journal of Remote Sensing,* 32(9), 2519-2533.

Markham, B. L. and Barker, J. L., 1985, Spectral characterization of the landsat thematic mapper sensors, *International Journal of Remote Sensing,* 6(5), 697-716.

Martin, R. V., 2008, Satellite remote sensing of surface air quality, *Atmospheric Environment,* 42(34), 7823-7843.

Mayer, J. D., 1982, Relations between two traditions of medical geography: health systems planning and geographical epidemiology, *Progress in Geography,* 6(2), 216-230.

McPherson, K., Steel, C., and Dixon, J., 2000, ABC of breast diseases: breast cancer—epidemiology, risk factors, and genetics, *BMJ: British Medical Journal,* 321(7261), 624-628.

Meehl, G. A. and Tebaldi, C., 2004, More intense, more frequent, and longer lasting heat waves in the 21st century, *Science,* 305(5686), 994-997.

Meng, Q., Zhang, L., Sun, Z., Meng, F., Wang, L., and Sun, Y., 2018, Characterizing spatial and temporal trends of surface urban heat island effect in an urban main built-up area: a 12-year case study in Beijing, China, *Remote Sensing of Environment,* 204, 826-837.

Mills, S., Weiss, S., and Liang, C., 2013, VIIRS Day/Night Band (DNB) Stray Light Characterization and Correction, Earth Observing Systems XVIII, International Society for Optics and Photonics, August 26~29, San Diego, USA.

Naguib, M., 2013, Living in a noisy world: indirect effects of noise on animal communication, *Behaviour,* 150(9-10), 1069-1084.

Nam, W.-H., Baigorria, G., Hong, E.-M., Kim, T., Choi, Y.-S., and Feng, S., 2018, The fingerprint of climate change and urbanization in South Korea, *Atmosphere,* 9(7), 273-289.

Novotny, V., 1994, *Water Quality: Prevention, Identification and Management of Diffuse Pollution*, New York: Van Nostrand-Reinhold Publishers.

Oldeman, L., Hakkeling, R., and Sombroek, W. G., 2017, World map of the status of human-induced soil degradation: an explanatory note, International Soil Reference and Information Centre.

Ord, J. K. and Getis, A., 1995, Local spatial autocorrelation statistics: distributional issues and an application, *Geographical Analysis,* 27(4), 286-306.

Owusu-Edusei, K. and Owens, C. J., 2009, Monitoring county-level chlamydia incidence in texas, 2004–2005: application of empirical bayesian smoothing and exploratory spatial data analysis (ESDA) methods, *International Journal of Health Geographics,* 8(1), 12.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., Church, J. A., Clarke, L., Dahe, Q., and Dasgupta, P., 2014, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, New York: Cambridge University Press.

Pauley, S. M., 2004, Lighting for the human circadian clock: recent research indicates that lighting has become a public health issue, *Medical hypotheses,* 63(4), 588-596.

Petermann, A., 1852, *Statistical Notes to the Cholera Map of the British Isles: Shewing the Districts Attached in 1831, 1832, and 1833*, London: John Betts.

Ponticiello, B. G., Capozzella, A., Di Giorgio, V., Casale, T., Giubilati, R., Tomei, G., Tomei, F., Rosati, M. V., and Sancini, A., 2015, Overweight and urban pollution: preliminary results, *Science of The Total Environment,* 518, 61-64.

Porta, M., 2014, *A Dictionary of Epidemiology*, Oxford: Oxford University Press.

Powell, S. L., Cohen, W. B., Yang, Z., Pierce, J. D. and Alberti, M., 2008, Quantification of impervious surface in the snohomish water resources inventory area of western Washington from 1972–2006, *Remote Sensing of Environment,* 112(4), 1895-1908.

Pyle, G. F., 1969, The Diffusion of cholera in the united states in the nineteenth century, *Geographical Analysis,* 1(1), 59-75.

Ritchie, J. C., Zimba, P. V. and Everitt, J. H., 2003, Remote sensing techniques to assess water quality, *Photogrammetric Engineering & Remote Sensing,* 69(6), 695-704.

Robine, J.-M., Cheung, S. L. K., Le Roy, S., Van Oyen, H., Griffiths, C., Michel, J.-P., and Herrmann, F. R., 2008, Death toll exceeded 70,000 in europe during the summer of 2003, *Comptes Rendus Biologies,* 331(2), 171-178.

Rodriguez, A. D., Rodriguez, M. H., Hernandez, J. E., Dister, S. W., Beck, L. R., Rejmankova, E., and Roberts, D. R., 1996, Landscape surrounding human settlements and anopheles albimanus (diptera: culicidae) abundance in Southern Chiapas, Mexico, *Journal of Medical Entomology,* 33(1), 39-48.

Rogers, D. and Randolph, S., 1991, Mortality rates and population density of tsetse flies correlated with satellite imagery, *Nature,* 351(6329), 739-741.

Rosenthal, J. K., Kinney, P. L. and Metzger, K. B., 2014, Intra-urban vulnerability to heat-related mortality in New York City, 1997–2006, *Health & Place,* 30, 45-60.

Seaman, V., 1798, Article Ⅱ, *The Medical Repository of Original Essays and Intelligence, Relative to Physic, Surgery, Chemistry, and Natural History (1797-1800),* 1(3), 315-336.

Shackelford, A. K. and Davis, C. H., 2003, A Combined Fuzzy Pixel-Based and Object-Based Approach for Classification of High-Resolution Multispectral Data over Urban Areas, *IEEE Transactions on GeoScience and Remote sensing,* 41(10), 2354-2363.

Sharma, R. C., Tateishi, R., Hara, K., Gharechelou, S., and Iizuka, K., 2016, Global mapping of urban built-up areas of year 2014 by combining modis multispectral data with viirs nighttime light data, *International Journal of Digital Earth,* 1-17.

Sheridan, S. C. and Kalkstein, L. S., 2004, Progress in heat watch–warning system technology, *Bulletin of the American Meteorological Society,* 85(12), 1931-1941.

Shi, K., Huang, C., Yu, B., Yin, B., Huang, Y., and Wu, J., 2014, Evaluation of npp-viirs night-time light composite data for extracting built-up urban areas, *Remote Sensing Letters,* 5(4), 358-366.

Sithiprasasna, R., Lee, W. J., Ugsang, D. M., and Linthicum, K. J., 2005, Identification and characterization of larval and adult anopheline mosquito habitats in the Republic of Korea: potential use of remotely sensed data to estimate mosquito distributions, *International Journal of Health Geographics,* 4(1), 17.

Skole, D. L., Chomentowski, W. H., Salas, W. A., and Nobre, A. D., 1994, Physical and human dimensions of deforestation in Amazonia, *BioScience,* 44(5), 314-322.

Slonecker, E. T., Jennings, D. B., and Garofalo, D., 2001, Remote sensing of impervious surfaces: a review, *Remote Sensing Reviews,* 20(3), 227-255.

Smiet, A. C., 1992, Forest ecology on java: human impact and vegetation of montane forest, *Journal of Tropical Ecology,* 8(2), 129-152.

Snow, J., 1855, *On the Mode of Communication of Cholera*, London: John Churchill.

Sohar, E., Adar, R. and Kaly, J., 1963, Comparison of the environmental heat load in various parts of Israel, *Bull Res Counc Israel E,* 10, 111-116.

Son, J.-Y., Lane, K. J., Lee, J.-T., and Bell, M. L., 2016, Urban vegetation and heat-related mortality in Seoul, Korea, *Environmental Research,* 151, 728-733.

Song, Y. and Wu, C., 2018, Examining human heat stress with remote sensing technology, *GIScience & Remote Sensing,* 55(1), 19-37.

Sterling, S. and Ducharne, A., 2008, Comprehensive data set of global land cover change for land surface model applications, *Global Biogeochemical Cycles,* 22(3).

Stevens, R. G., 1987, Electric Power Use and Breast Cancer: A Hyptothesis, *American Journal of Epidemiology,* 125(4), 556-561.

Stevens, R. G., 2005, Circadian disruption and breast cancer: from melatonin to clock genes, *Epidemiology,* 16(2), 254-258.

Stevens, R. G., 2009, Light-at-night, circadian disruption and breast cancer: assessment of existing evidence, *International Journal of Epidemiology,* 38(4), 963-970.

Stewart, I. D. and Oke, T., 2009, A new classification system for urban climate sites, *Bulletin of the American Meteorological Society,* 90(7), 922-923.

Stull, R., 2011, Wet-bulb temperature from relative humidity and air temperature, *Journal of Applied Meteorology and Climatology,* 50(11), 2267-2269.

Sutton, P., Roberts, D., Elvidge, C., and Baugh, K., 2001, Census from heaven: an estimate of the global human population using night-time satellite imagery, *International Journal of Remote Sensing,* 22(16), 3061-3076.

Tan, M., 2016, An intensity gradient/vegetation fractional coverage approach to mapping urban areas from DMSP/OLS nighttime light data (in press), *IEEE Journal of Selected Topics In Applied Earth Observation and Remote Sensing*.

Thom, E. C., 1959, The Discomfort Index, *Weatherwise,* 12(2), 57-61.

Thompson, D. F., Malone, J. B., Harb, M., Faris, R., Huh, O. K., Buck, A. A., and Cline, B. L., 1996, Bancroftian filariasis distribution and diurnal temperature differences in the southern nile delta, *Emerging infectious diseases,* 2(3), 234-235.

Thomson, M., Connor, S., Milligan, P., and Flasse, S., 1997, Mapping malaria risk in africa: what can satellite data contribute? , *Parasitology Today,* 13(8), 313-318.

Tobler, W. R., 1969, Satellite confirmation of settlement size coefficients, *Area,* 1(3), 30-34.

Tobler, W. R., 1970, A computer movie simulating urban growth in the detroit region, *Economic Geography,* 46(sup1), 234-240.

Travis, J., 1997, Spying diseases from the sky: satellite data may predict where infectious microbes will strike, *Science News,* 152(5), 72-73.

Tselepidaki, I., Santamouris, M., Moustris, C., and Poulopoulou, G., 1992, Analysis of the summer discomfort index in Athens, Greece, for cooling purposes, *Energy and buildings,* 18(1), 51-56.

Unwin, A. R., 1996, Exploratory spatial analysis and local statistics, *Computational statisctics*, 11, 387-400.

USGS, 2015, Landsat 8 (L8) Data Users Handbook, *USGS*.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., and Liermann, C. R., 2010, Global threats to human water security and river biodiversity, *Nature,* 467(7315), 555-582.

Wackernagel, M. and Rees, W., 1998, *Our Ecological Footprint: Reducing Human Impact on the Earth*, Gabriola Island: New Society Publishers.

Waller, L. A. and Gotway, C. A., 2004, *Applied Spatial Statistics for Public Health Data*, Hoboken: John Wiley & Sons.

Walling, D., 2006, Human impact on land–ocean sediment transfer by the world's rivers, *Geomorphology,* 79(3-4), 192-216.

Wang, L., Sousa, W., and Gong, P., 2004, Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery, *International Journal of Remote Sensing,* 25(24), 5655-5668.

Wartenberg, D., 1985, Multivariate spatial correlation: a method for exploratory geographical analysis, *Geographical Analysis,* 17(4), 263-283.

Weng, Q., Hu, X., and Lu, D., 2008, Extracting impervious surfaces from medium spatial resolution multispectral and hyperspectral imagery: a comparison, *International Journal of Remote Sensing,* 29(11), 3209-3232.

Weng, Q., Lu, D., and Schubring, J., 2004, Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies, *Remote Sensing of Environment,* 89(4), 467-483.

Wilson, E., 2001, *The Diversity of Life*, New York: W. W. Norton.

Xu, H., 2010, Analysis of impervious surface and its impact on urban heat environment using the normalized difference impervious surface index (NDISI), *Photogrammetric Engineering & Remote Sensing,* 76(5), 557-565.

Yan, W. Y., Shaker, A., and El-Ashmawy, N., 2015, Urban land cover classification using airborne lidar data: a review, *Remote Sensing of Environment,* 158, 295-310.

Yang, W., Yang, L., and Merchant, J., 1997, An assessment of AVHRR/NDVI-ecoclimatological relations in Nebraska, USA, *International Journal of Remote Sensing,* 18(10), 2161-2180.

Yao, J., Murray, A. T., Agadjanian, V., and Hayford, S. R., 2012, Geographic influences on sexual and reproductive health service utilization in rural Mozambique, *Applied Geography,* 32(2), 601-607.

Yoo, S.-H., Han, S.-H., Heo, J., and Sohn, H.-G., 2011, Variable selection for estimating population using DMSP-OLS night-time image, *Korean Journal of Remote Sensing,* 27(1), 69-74.

Yuan, X., Zhu, C., Wang, M., Mo, F., Du, W., and Ma, X., 2018, Night shift work increases the risks of multiple primary cancers in women: a systematic review and meta-analysis of 61 articles, *Cancer Epidemiology and Prevention Biomarkers,* 27(1), 25-40.

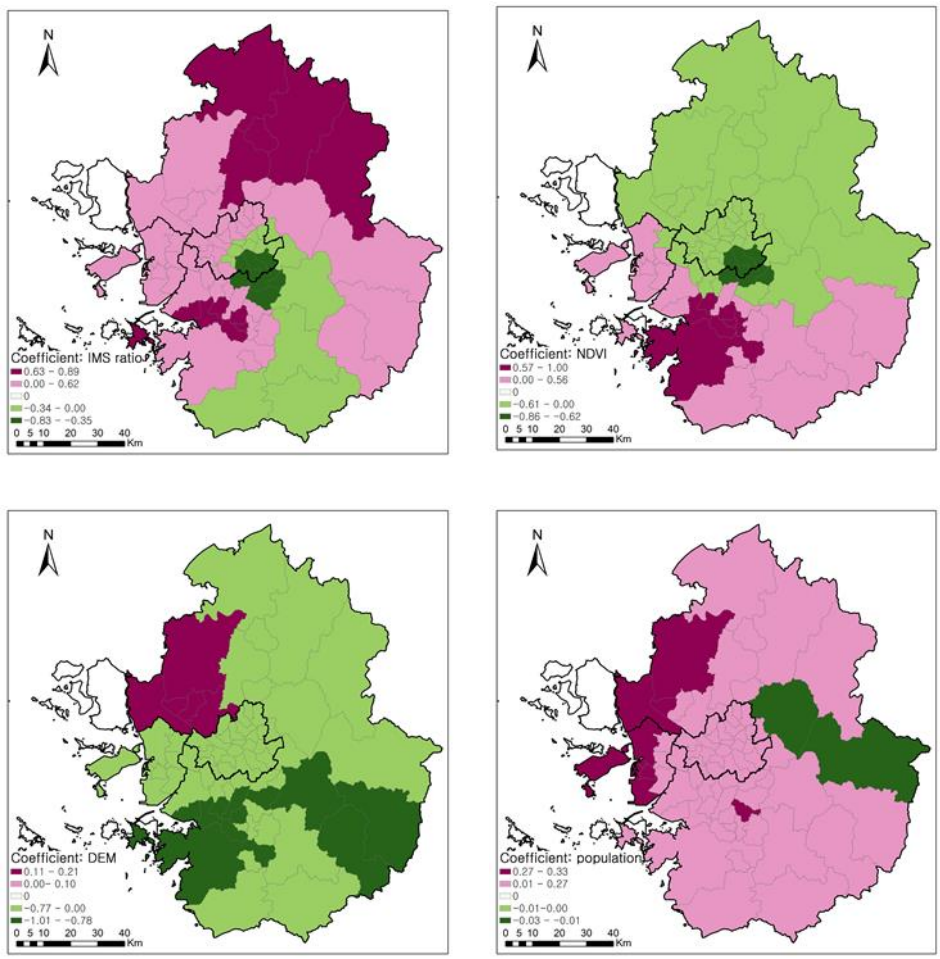
Yue, W., Xu, J., Tan, W., and Xu, L., 2007, The relationship between land surface temperature and ndvi with remote sensing: application to Shanghai Landsat 7 Etm+ data, *International Journal of Remote Sensing,* 28(15), 3205-3226.

Zhang, J., Zhou, Z., Shuai, G., and Liu, H., 2016, Support vector data description model to map urban extent from National Polar-Orbiting Partnership Satellite–Visible Infrared Imaging Radiometer Suite Nightlights and normalized difference vegetation index, *Journal of Applied Remote Sensing,* 10(2), 026012.

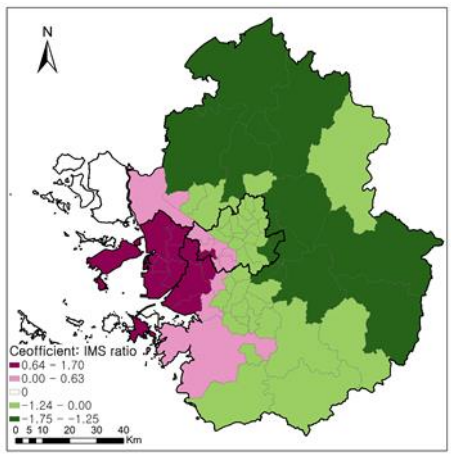
Zhang, T. and Huang, X., 2018, Monitoring of urban impervious surfaces using time series of high-resolution remote sensing images in rapidly urbanized areas: a case study of Shenzhen, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 11(8), 1-17.

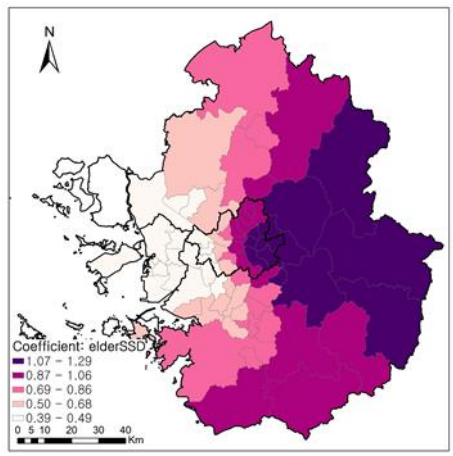
Appendix

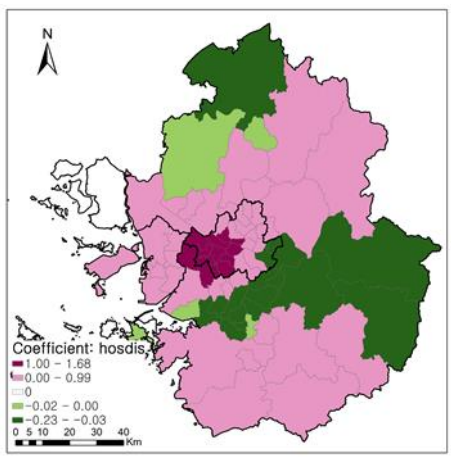
Appendix 1: The distribution of regression coefficients in the model on the DI value



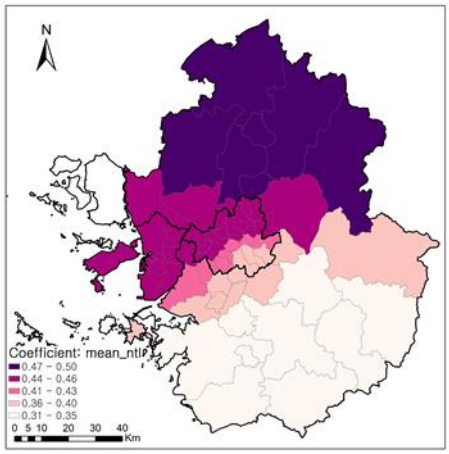
Appendix 2: The distribution of regression coefficients in the model on the non-accident mortality

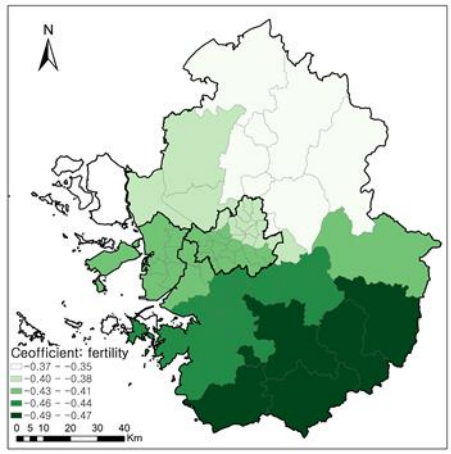


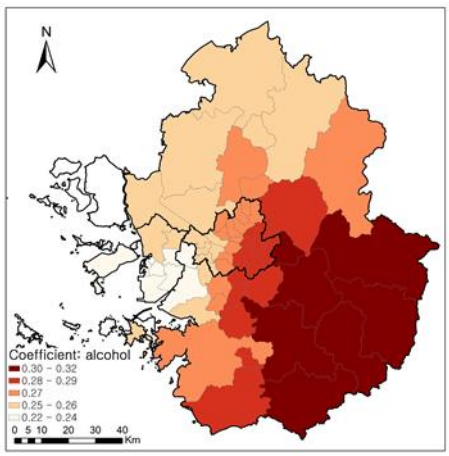




Appendix 3: The distribution of regression coefficients in the model on the breast cancer prevalence rate (significant coefficients)







국문초록

인류발생적 건강위협 인자가 인체건강에 미치는 영향력의 공간적 변동 탐색:

원격탐사 데이터에 대한 ESDA 적 접근

주뢰

사범대학 사회교육과 지리전공

서울대학교 대학원

질병과 보건서비스 제공의 공간적 분포는 보건지리학의 두 가지 주요 주제이다. 두 가지 주요 관심사에 관한 연구를 수행하기 위해 많은 자료와 공간적 분석 기법들이 사용되어 왔다. 그 중에 원격탐사 기술은 다양한 정보와 광범위한 시·공간적 스케일의 장점으로 인해 1970년대 이후 보건지리학에서 널리 사용되어 왔다. 본 연구는 탐색적 공간 데이터 분석(exploratory spatial data analysis, ESDA)의 관점에서 원격탐사 데이터를 사용하여 잠재적인 인류발생적 건강위협 인자가 인체 건강에 미치는 영향에 있어서의 공간 변동을 파악하고자 한다.

이 연구에서는 잠재적인 인류발생적 건강위협 인자를 추출하기 위해 Landsat 8과 Visible Infrared Imaging Radiometer Suite(VIIRS)로부터 얻어진 다양한 원격탐사 데이터를 사용하였다. 두 가지 위성영상 자료에서 잠재적인 인류발생적 건강위협 인자를 추출한 결과, 인공 불투수면(impervious surface, IMS)과 인공 야간 조명(nighttime light, NTL)에 대한 노출이 가장 대표적인 인자로 추출되었다. 한편,

불쾌지수(discomfort index, DI), 비사고 사망률, 그리고 유방암 유병률은 이들 요인들에 의해 영향을 받는 잠재적인 건강 결과로 여겨진다.

이를 바탕으로 본 연구는 우선 잠재적인 인류발생적 건강위협 인자와 건강결과의 공간적 분포 특성 및 공간적 연관성을 파악하였다. 그 결과, 잠재적인 인류발생적 건강위협 인자들은 도시 지역에 집중되는 경향이 있는 반면, 건강결과는 다양한 공간적 분포 특징들을 가진 것으로 나타났다. 그리고 공간적 연관성에 있어 이들 두 변수들은 양(+)의 공간적 상관관계를 갖는 경향을 보였다. 다음으로 본 연구는 공간회귀분석을 이용하여 인류발생적 건강위협 인자가 건강에 미치는 영향력에 대한 공간적 변동을 분석하였다. 분석 결과, 인공 불투수면과 야간 불빛은 불쾌지수 및 유방암 유병률 증가의 주된 요인으로 파악되었다.

주요어: 인류발생적 건강위협 인자, 인체 건강, 원격탐사, 공간적 변동, 공간적 연관성, 보건지리학

학번: 2015-30770

中文摘要

人为健康威胁要素对健康影响的空间分布特征研究：“探索性空间数据分析（ESDA）”方法对遥感数据的应用

朱蕾

师范学院 社会教育学科 地理专业

首尔国立大学

疾病和医疗保健服务的空间分布是健康地理研究的两大主要部分。各种数据资源和地理分析技术已经在这两个问题的研究中广泛应用。其中，因丰富的信息量以及大尺度的时空范围，遥感技术自 20 世纪 70 年代以来已经广泛应用于健康地理学研究中。通过使用遥感数据，本研究的主要目的是利用探索性空间数据分析（Exploratory Spatial Data Analysis, ESDA）检测潜在的人为健康威胁要素对人体健康影响的空间异质性。

本研究利用 Landsat 8 影像和 Visible Infrared Imaging Radiometer Suite (VIIRS)影像提取遥感影像中潜在的人为健康威胁要素。在这两个影像应用中，人工不透水面 (impervious surface, IMS)和夜间灯光 (nighttime light, NTL)被提取为最具代表性的潜在人为健康威胁要素。同时，不适指数(discomfort index, DI)，非事故死亡率以及乳腺癌的发病率被认为是这些因素对健康可能造成的影响。

首先，本研究分析了上述人为健康威胁要素的分布特征，以及不透水面与不适指数，非事故死亡的空间相关性，夜间风光与乳腺癌发病率的空间相关性。其次，本研究用空间回归模型分别分析了不透水面对不适指数，非事

故死亡的影响，以及夜间灯光对乳腺癌发病率的影响。结果表明，潜在的人为健康威胁要素倾向于集中在城市地区，而其对健康的影响则有明显的地理分布的异质性。两个变量之间倾向于具有正向的空间关联性。同时，人工不透水面和夜间灯光被证明是分别影响不适指数及乳腺癌发病率的主要因素。

关键词：人为健康威胁要素，健康，遥感技术，空间异质性，空间相关性，健康地理学

学号：2015-30770