SPATIAL INFRASTRUCTURE SUSTAINABILITY ASSESSMENT FRAMEWORK USING MULTI-CRITERIA ANALYSIS

A Dissertation
Presented to
The Academic Faculty

by

Mengmeng (Helen) Liu

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in
Computational Science & Engineering in
the School Civil and Environmental Engineering

Georgia Institute of Technology
August 2019

Approved by:

Dr. J. David Frost, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. William J. Drummond
School of City and Regional Planning
Georgia Institute of Technology

Dr. Yi-Chang J. Tsai
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Duen Horng (Polo) Chau
School of Computational Science & Engineering
Georgia Institute of Technology

Dr. Catherine Ross
School of City & Regional Planning
Georgia Institute of Technology

Date Approved: [May 10th, 2019]

I dedicate this dissertation work to my parents, for unconditional love, endless support, and encouragement. Thank you for teaching me to believe in myself and to pursue my dream! Thanks for encouraging me to be myself!

To my husband, Qingsong for your patience, love, friendship, help, support, and encouragement. I could not finish this dissertation so quick without your help and support. Thanks to my picky appetite, your cooking skills improved a lot!

To my lovely and sweet daughter, Emma, who brings me so much happiness than I imagine. I love you more than anything! Your shining smile, your adorable giggle, your beautiful eyes, your endless energy, you are the perfect girl in the world! I am so lucky to be your mom and your friend. I hope you grow happily and healthily, be yourself!

I love you all dearly!

ACKNOWLEDGEMENTS

The past six years have been a long, challenging, and beautiful journey. The research work presented here would not have been accomplished without the support and encouragement of many people. Foremost, I would like to express my sincere gratitude to my advisor Dr. J. David Frost for the continuous support of my Ph.D. study and research, for his help and insightful comments and suggestions for my research, for his patience, generous, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. His wisdom and insights on many things, including research, career, personal development, and life, inspired me deeply and helped me overcome difficulties in life and research. I could not have imagined having a better advisor and mentor for my Ph.D. study. David is not only my advisor but also my best friend! I have a great time to work with him.

Besides my advisor, I would like to thank the rest of my thesis committee: Dr. Duen Horng (Polo) Chau, Dr. William J. Drummond, Dr. Catherine Ross, and Dr. Dr. Yi-Chang J. Tsai, for their encouragement, insightful comments, suggestions, assistance, and questions. I also want to thank Dr. Bistra Dilkina, Dr. Randall L. Guensler, Dr. Jenny Hirsch, and Dr. Haiying Huang for their help, support, and valuable suggestions for my research. I want to thank Dr. Haobing Liu, Xiaodan Xu, Yingping Zhao, for their kindly help of energy consumption estimation in the pilot study. I gratefully acknowledge the funding provided by the Center of Serve Learn and Sustain to support my pilot study in this thesis. Thanks to the Ray C. Anderson Fellowship for supporting part of my research.

My time at Georgia Tech was enjoyable due to the many friends and groups that became a part of my life. I want to acknowledge Fangzhou Liu, Jie Cao, Yixuan Sun, Yue Xu, Tianlong Xu, Yuzhi Guo, Mahdi Roozbahani, Jackson Su, Fikret Atalay, Alejandro Martinez, Shahrzad (Sherie) Roshankhah, Karie Yamamoto, Seth Mallet, Ariel Segel, Nimisha Roy, Sangy Hanumasagar, Andres F. Peralta, Rodrigo Borela, Longde Jin, Wencheng Jin, Cheng Zhu, Jiaojun Liu, Prashanth**,** Jintao Huang He, as well as other former and present students of Georgia Tech GeoSociety for their friendship, accompany and kind help. I really enjoyed the conversations and fun time during meetings, geo-parties, tailgates, trips, and so on. Thanks my teammates and friends, Arezoo Shirazi, Caleb Robinson, Aroua Gharbi, Xin Cao, Jingtian Zhang, Tong Jin, Qingsong Wen, Xiaoxue Zhang, Jian Yuan, Xin Xing, Ajitesh Jain, Wenlong Zhang, Xiao Zhang, and so on who works together with me on class projects or papers. You are great teammates and friends.

I also want to thank my friend Yige Gao, Ying Shao, Jing Ren, Ying Jing, Rong Li, Weihua Zhao, Suhua Liu, Yuhuan Ren, Qing Xu, Winston Yang, Shuli Liu, and so on (too many to list here but you know who you are!) for providing support and friendship that I needed. I want you all to know that you count so much to me. I would like to thank my wonderful and nice neighbors, Lina Yang, Jin Suo, Yang Yang, Mei Ouyang, Jiajia, Rui Xiong, Hualan Jiang, Ximei Zhang, Duanshu Song, for their generous help, even some of them have returned China. I also want to thank my doctors who keep me healthy in the past six years, especially Dr. Minghui Li who saves my life and gives me many encouragements and help! She is the best doctor in the world!

Lastly, I would like to thank my family for all their love and encouragement. For my parents who raised me with unconditional love and supported me in all my pursuits. I love them so much, and I would not have made it this far without them. For my parents-in-law who help me taking care of my daughter and support me in my study at Georgia Tech. Most of all for my loving, supportive, encouraging, and patient husband Qingsong whose faithful support during the final stages of this Ph.D. is so appreciated. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AHP	Analytic Hierarchy Process
ANN	Artificial Neural Networks
ANP	Analytic Network Process
API	Application Program Interface
ARC	Atlanta Regional Commissions
CBA	Cost-Benefit Analysis
CI	Consistency Index
CRS	Coordinate Reference System
CR	Consistency Ratio
CSI	Composite Sustainability Index
CSV	Comma Separated Variables
CTV	Contribution To Variance
CV	Coefficient of Variation
DI	Detour Index
DPSEEA	Driving force-Pressure-State-Exposure-Effect-Action
DPSIR	Driving force-Pressure-State-Impacts-Response
EEA	European Environment Agency
EPI	Environmental Pressure Indicators
EU	European Union
EUROSTAT	Statistical Office for the European Communities
GIS	Geographical Information Systems
GoldSET	Golder's "Sustainability Evaluation Tool"

GPI	Genuine Progress Indicator
GUI	Graphical User Interface
HDI	Human Development Index
IDW	Inverse Distance Weighting
IPCC	Intergovernmental Panel on Climate Change
IQM	Interquartile Mean
ISEW	Index for Sustainable Economic Welfare
IU	International Unit
LCA	Life Cycle Assessment
LCI	Life Cycle Inventory Analysis
LCIA	Life Cycle Impact Assessment
LEHD	Longitudinal-Employer Household Dynamics
LHS	Latin Hypercube Sampling
LODES	LEHD Origin-Destination Employment Statistics
MCA	Multi-Criteria Analysis
MCS	Monte Carlo Simulation
MFA	Material Flow Assessment
NRPA	National Recreation and Park Association
OD	Origin-Destination
PCA	Principle Component Analysis
PMC	Primitive Monte Carlo
PRD	Pedestrian Route Directness
PSR	Pressure-State-Response
QGIS	Quantum GIS
RCI	Random Consistency Index

SA	Sensitivity Analysis
SD	System Dynamics
SDI	Sustainable Development Index
SEA	Strategic Environmental Assessment
SILENT	Sustainable Infrastructure, Land-use, Environment and Transport Model
TAZ	Traffic Analysis Zone
TBL	Triple-Bottom-Line
TIN	Triangulated Irregular Networks
WHO	World Health Organization

SUMMARY

The world's natural resources are presently under increasing pressure and humanity is facing multiple sustainable challenges, such as environmental degradation, pollution, water shortage, climate change, poverty, social inequalities, carbon-intensive lifestyles, and so on. Sustainability is the only way to resolve the economic, environmental and social challenges simultaneously. It has become a prominent concept in societal and political discourses around the world and serves as a major guideline for political actions and future societal development. However, in order to help decision makers (policy makers) determine which actions should or should not be taken in an attempt to make our society sustainable or to assess the efficiency of actions engaged, a sustainability assessment framework is needed to evaluate the sustainability of proposed projects or plans in short-term and long-term perspectives.

Sustainability assessment frameworks clarify what to measure, what to expect from measurement and what kind of indicators to use when evaluating sustainability. There are many frameworks proposed by researchers, institutions or organizations. Among them, this thesis primarily considered and reviewed Driving force-Pressure-State-Impacts-Response (DPSIR), Driving force-Pressure-State-Exposure-Effect-Action (DPSEEA), Impact-based frameworks, Stakeholder-based frameworks, Golder's "Sustainability Evaluation Tool" (GoldSET), and Spatial Frameworks. Each of these frameworks has its advantages and disadvantages. Some consider the inter-linkages between components of sustainability (e.g., DPSIR, DPSEEA), some considers all aspects of sustainability (e.g., Triple-Bottom-Line (TBL), Multi-Criteria Analysis (MCA), GoldSET), some involving the stakeholder

into the process of sustainability assessment (i.e., Stakeholder-based frameworks, GoldSET), some considers the spatial difference of sustainability (e.g., SILENT). However, when evaluating the sustainability of one system or one project, a good sustainability assessment framework not only needs to consider all the aspects of sustainability, the inter-linkages, and dynamics changes happened in one system but also need to involving stakeholders into the process of sustainability assessment.

In this thesis, the author proposes a spatial sustainability assessment framework, which integrates DPSIR, TBL, Life Cycle Assessment (LCA), and stakeholder-based methods. The framework can make full use of the advantages of each method and provides a scientific and comprehensive method to spatially assess the sustainability of infrastructure at different spatial and temporal scales. The framework has components to define sustainability assessment objectives, to select indicators, to calculate the value of some indicators, to conduct multi-criteria decision analysis (pre-analysis, weighting, and aggregation of indicators), to evaluate the uncertainty and sensitivity of the assessment results, and to visualize sustainability evaluation results in multiple ways. The framework engages policymakers and other participants in the process of sustainability evaluation, by letting them set the weight of indicators with the help of the Analytic Hierarchy Process (AHP) or Principle Component Analysis (PCA) in the framework. Besides, there is a spatial database to manage the data, indicators and other information used during the process of sustainability assessment. The framework is fully data-driven, and is a general framework, which can be easily adapted and used in other disciplines and other application areas.

The framework is implemented as a plugin in Quantum Geographic Information System (QGIS), named "Spatial Sustain Assess". All the functions provided in QGIS can be used seamlessly in the framework. The framework provides various tools to help conduct sustainability assessment, including network analysis, spatial statistical analysis, comparison analysis, aggregation of attributes to larger scale, Pre-analysis of indicators (normalization, correlation, PCA), AHP, aggregating indicators, uncertainty and sensitivity analysis of sustainability assessment results, data explorations, and multiple visualizations of sustainability assessment results. It also provides a "wizard" to guide users conducting a sustainability assessment of their projects, step by step. Due to the limit of time, the author only implemented part of the functions in the proposed framework, and many works need to be done in the future. First, the current functions in the framework need to be optimized by decreasing their processing time. Second, more functions or tools need to be added to the framework to help evaluate sustainability, such as DPSIR to help select individual indicators, tools to evaluate commonly used individual sustainability indicators, and life cycle assessment. Finally, the framework needs to address the multi-scale problems to make the sustainability assessment at different spatial scales consistent. Fifth, a spatial database of sustainability evaluation indicators need to be well designed and constructed, so that it better supports the spatial sustainability assessment framework.

As an application of the "Spatial Sustain Assess" plugin, the author uses it to evaluate the sustainability performance of moving interstates within the Atlanta Perimeter (I-285) underground. First, twelve indicators are selected from economic, environmental, and social dimensions of sustainability for the sustainability assessment. After analyzing the correlations between indicators, ten relatively independent indicators are retained for

MCA. AHP is used to derive the weight of each indicator and a "linear additive model" is used to aggregate indicators to build Composite Sustainability Index (CSI) for economic, environmental and social dimension, as well as the overall CSI. Finally, uncertainty and sensitivity analysis are conducted for the sustainability assessment. Based on the sustainability assessment results, after moving interstates underground, both the overall sustainability and the sustainability at each dimension in the study area are improved, especially for the census tracts that are located within a 2-km buffer of interstates. However, further studies are needed to evaluate better the sustainability of moving interstates underground. For example, improving the measurement of some indicators used in this study with more data available, such as energy cost, equity, traffic flows, and detour index. Adjusting or updating the assumptions of the scenario after moving interstates underground, such as converting some of the newly released lands into commercial buildings or residential buildings, reconnection more roads, or redesigning some current roads (e.g., increase traffic capacity). Adding more indicators in the process of sustainability assessment, such as the ecological footprint, noise pollution level of residential places, traffic congestion levels, resilience or vulnerability of the transportation system to disasters, bad weather, or disruptions.

CHAPTER 1. INTRODUCTION

1.1 Background of the Research

The world's natural resources are presently under increasing pressure and humanity is facing multiple sustainable challenges, such as environmental degradation, soil erosion, desertification, pollution, famine, water shortage, extinction of species, global warming, climate change, poverty, population growth, rapid urbanization, social inequalities, and carbon-intensive lifestyles.

Sustainability is considered by many the best way to address the vast, complex and interrelated economic, environmental and societal problems and is deemed highly imperative for the sake of current and future generations (Waas et al., 2011). Sustainability can resolve the economic, environmental and social challenges simultaneously (Stumpf et al., 2015). In this sense, sustainable development not only represents a solution for these challenges, but offers a set of principles for a new view of value and a new way to understand human being's relationship with others in the world (Waas et al., 2011).

Nowadays, sustainability has become a prominent concept in societal and political discourses around the world and serves as a major guideline for political actions and future societal development (Stumpf et al., 2015). By the end of the 20th century, in response to a growing environmental crisis and inequalities in global development, the international community adopted sustainable development as a leading development model (Waas et al., 2011). Sustainable urban development has been on the top of the agenda in almost every city across the world (Yigitcanlar and Dur, 2010).

However, a framework or tool is needed to evaluate the global to local integrated nature-society systems in short- and long-term perspectives, in order to help decision makers or policy makers determine which actions should or should not be taken in an attempt to make our society sustainable or to assess the efficiency of actions engaged, (Bourdic et al., 2012; Singh et al., 2009; Yigitcanlar and Dur, 2010). There is a call for sustainability assessments from the local scale, such as sustainable development of cities and neighborhoods, to the global scale, for example, the United Nations sustainability goals; and from the product level, e.g., eco labels, to the sector level (Zijp et al., 2015). There have been various studies which have proposed different methods for sustainability assessment (Singh et al., 2009). However, only a few of them have an integral approach that takes into account all of the environmental, economic and social aspects during sustainability assessment. Besides sustainability is more than an aggregation of different aspects, it is also about their inter-linkages and the dynamics developed in a system (Graymore et al., 2009; Yigitcanlar and Dur, 2010). Thus a good sustainability assessment framework not only considers all aspects of sustainability, (environmental, economic, and social), as well as other factors such as resilience, but further it should also include inter-linkages between different aspects and their dynamic changes.

1.2 Problem Statement

In 2016, an estimated 54.5% of the world's population lived in urban settlements. By 2030, urban areas are projected to include 60% of people globally, and one in every three people will live in cities with at least half a million inhabitants. As mentioned before, there are many sustainable challenges to the world. Cities are at the heart of the problems or challenges facing this planet, thus developing a positive and sustainable mode of urban living is the only way to be able to sustain social life (James et al., 2015).

Infrastructures as the fundamental facilities and systems serving a country, city, or other areas (James et al., 2015), are critical to sustainable community development. They consist of basic facilities such as transportation, communications, power supplies, water supplies, buildings, and sanitation utilities. They should not only deliver their service efficiently but also in a way that helps cities towards addressing sustainability challenges. Evaluating their spatial sustainability is key to making infrastructures develop sustainably and contribute to the sustainability of the city. Therefore, it is urgently necessary to build a framework to assess the sustainability of infrastructures and to provide useful suggestions or information for the planning, design, investment and construction of new infrastructure systems and the rehabilitation, reuse or optimization of existing infrastructures.

As indicated by the nature of sustainability, it is necessary to incorporate its spatial and temporal scales or aspects into developing a sustainability assessment framework. The spatial assessment of sustainability should consider the spatial difference of distribution or allocation of natural and human made resources, population, economic development level, energy consumption structure, and culture, as well as spatial patterns and spatial locations of infrastructures. For example, more attention is paid to critical infrastructures such as transportation systems, electrical power systems, water supply systems and so on, when assessing the sustainability of the whole infrastructure system, because "the critical infrastructures are so vital to society that their incapacitation or destruction would have a debilitating effect on security, national economic security, national public health or safety, or any combination thereof" (Riedman, 2016). Meanwhile, temporal changes of infrastructure also need to be considered, since it can help in understanding the stages and development trends of infrastructure. Therefore, a spatial sustainability assessment framework is needed to evaluate the sustainability of infrastructures thoroughly and completely.

1.3 Purpose of Study

This study aims to develop a spatial infrastructure sustainability assessment framework. The framework will provide a scientific and comprehensive method to spatially assess the sustainability of infrastructure at different spatial and temporal scales. It engages stakeholders or people from different areas into the process of sustainability assessment. It will also provide multiple ways to present the sustainability assessment results clearly and effectively, so that it can assist decision-makers in determining which actions should or should not be taken in an attempt to make infrastructure sustainable. Besides, there is a spatial database to manage the data, indicators and other information used during the process of sustainability assessment. In the database, the users can also find robust, quantitative, science-based and cross-scale indicators for sustainability assessment.

1.4 Scope of the Study

In this study, spatial sustainability can be evaluated at multiple spatial scales, such as the national scale, state scale, local scale, infrastructure system component scale, or individual infrastructure element scale. In the spatial sustainability assessment framework, one can evaluate sustainability from different aspects, including environmental, economic, social and resilient (technique) perspectives. These four aspects are important dimensions of sustainability. During the evaluation, it is not only necessary to evaluate these four dimensions separately, but also to consider the inter-linkages between them. Even though the framework proposed in this study is for civil infrastructures, it is also a general framework, which can be easily used in other subjects or disciplines. Besides, the visualization and interpretation of sustainability evaluation results are also included in this framework. Visualization is very important and plays a very critical role when explaining the evaluation results to the public or policy makers. Multiple visualization methods will be included in this framework so that users can see the evaluations results from different perspectives. To illustrate the use of the proposed spatial sustainability framework, the sustainability of moving interstates within the Perimeter (I-285) Atlanta underground will be evaluated. Under this proposed framework, users can also compare different projects, planning propositions and policies to see which one is more sustainable or contributes more to the sustainability (or sustainable development objectives and agenda) of our city or community, state, national or even the world.

1.5 Key Terminology

1.5.1 Sustainability

There is an ongoing debate on the definition of “sustainability”. Ali-Toudert and Li (2017) give a schematic summary of 12 sustainability models, which represent different understanding or interpretation of sustainability. Sustainability is a multi-dimensional concept that takes into account different aspects of natural and human-made society, such as the environment, ecological system, economy, human society, and resources. Sustainability should encourage multi-function, diversity, and effective usage of resources.

Its main point is to seek dynamic balance within one system or between different systems and to ensure that everyone or everything has equal rights or opportunities in a finite period. First, sustainability is the equivalent allocation of resources and opportunities between generations or different contemporary people and is the equivalent development of different areas on earth. Second, sustainability is to seek a dynamic balance of all the things on the earth, such as the dynamic balance between economy and environment (Dong et al., 2015), and maintain their states in balance, which asks every system to develop and incorporate sustainably within its capacity. Sustainability is a necessary and sufficient condition for a population to be at or below carrying capacity (Daily and Ehrlich, 1996). Third, sustainability has a finite period and may change over time, because all systems or subsystems have limited longevity (Costanza and Patten, 1995). This is why it is necessary to consider temporal scales of sustainability in an assessment framework. Sustainability should stimulate markets for "green" products and services, promote environmentally friendly consumption, and contribute to urban economic development by creating a cleaner environment (Technion, 2017).

1.5.2 *Sustainable Development*

The concept of sustainable development emerged in the early to mid-1980s as an attempt to bridge the gap between environmental concerns about the increasingly evident ecological consequences of human activities and socio-political concerns about human development issues (Robinson, 2004). Brundtland (1987) defined sustainable development as "development which meets the needs of current generations without compromising the ability of future generation to meet their own needs". Sustainable development is not a fixed state of harmony, but rather a process of change in which the exploitation of resources, the direction of investments, the orientation of technological development, and institutional change are made consistent with future as well as present needs (Brundtland, 1987). This is the most popular definition of sustainable development.

The term "sustainability" is commonly used as a synonym of "sustainable development" (Waas et al., 2011). However, sometimes their meanings are different. Some scholars assert that "sustainable development" is primarily about development/economic growth, whereas "sustainability" focuses attention on the ability of humans to continue to live within environmental constraints (Robinson, 2004; Waas et al., 2011). Throughout this thesis, the author will use the terms "sustainability" and "sustainable development" interchangeably.

1.5.3 *Sustainable Assessment*

Sustainability can only be assessed when one look at systems and subsystems as hierarchically interconnected over a range of spatial and temporal scales (Costanza and Pattern, 1995). Sustainability assessment is a process integrating sustainability issues into decision-making by identifying and assessing the sustainability of future consequences of current and planned actions (Hugé et al., 2011; Waas et al., 2014). Its main purpose is to help decision makers to determine which actions should or should not be taken in an attempt to make society sustainable (Singh et al., 2009). It gives decision makers an interpretation of their sustainability, an influence of their actions, and structuring the inherent multi-dimension complexity of sustainability into operational information units (e.g., indicators) to feed their decision-making process (Waas et al., 2014).

1.5.4 *Resilience*

Resilience is “the measure of the persistence of systems and of their ability or capacity to prepare for threats, to absorb impacts, to recover from and to successfully adapt to changes, disturbance or adverse events.” (Holling, 1973; Marchese et al., 2018). It becomes evident that a resilient system must be redundant, diversified and efficient, autonomous and collaborative at the same time (stiff, flexible and adaptive). The system also needs to be innovative and be capable of learning from the past and of facing future uncertainties. However, to some extent these characteristics are contradictory. For example, cities are vulnerable to disasters for many reasons, such as complex infrastructures, a high concentration of population, and so on. However, cities are also resilient to disasters because they contain a greater amount of resources for recovery and reconstruction, as well as more specialized skill, expertise, and innovation (Gasparini et al., 2014).

Resilience focuses on the response of systems to both extreme disturbances and persistent stress. Resilience is seen as a necessary precondition for sustainability (Derissen et al., 2011; Perrings, 2006). Increasing the resilience of a system makes it more sustainable, without resilience a system can only possess fragile sustainability (Marchese et al., 2018). And “any consideration of sustainability without accounting for resilience would render the goal of sustainable development unrealized.” (Puppala and Bheemasetti, 2018). Figure 1-1 illustrates how the resilience of a system can impact that system's sustainability and addresses how a resilient system can become sustainable after recovering from a disruption through the adaptive component of resilience. More resilient systems can better achieve and maintain a sustainable operation, but they do not necessarily increase the sustainability of a system (Marchese et al., 2018). In this study, resilience is considered as one important dimension or aspect of sustainability. Studies which treat resilience as one component of sustainability can also be found in (Marchese et al., 2018), (Achour et al., 2015) and (Perrings, 2006).

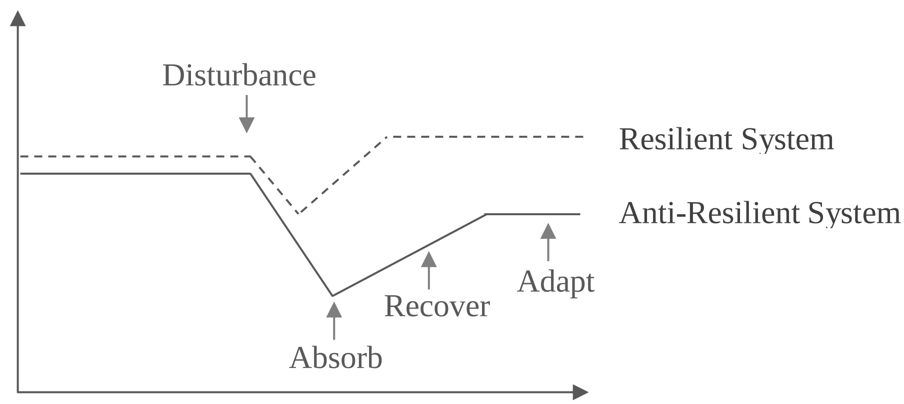


Figure 1-1 Resilience Component of Sustainability (Adapted from Marchese et al., 2018)

1.5.5 *Equity*

Equity is a measure of the relative similarity among individuals or groups in the opportunity to enjoy socio-political rights, material resources, technologies, health, education and other ingredients of human wellbeing (Daily & Ehrlich, 1996). In the context of sustainability, the term equity has to do with fairness -- whether all people have similar rights, opportunities and access to all forms of community capital (Sustainable Measures, 2018). In this study, we treat equity as an important factor of sustainability, it relates to and addresses social aspects of sustainability (Technion, 2017; Waheed et al., 2009). It includes both the “intra-generational equity” and “inter-generational equity”. Intra-generational equity refers to fairness in the allocation of resources between competing interests at present (Waheed et al., 2009). Inter-generational equity refers to the fairness in the allocation of resources between current and future generations (Waheed et al., 2009).

1.5.6 Spatial Scales

Spatial scales specify the area for which each indicator computation make sense (Bourdic et al., 2012). These are the nation, state, region, city, district, neighbourhood, block and specific project or individual project. Indicators for different spatial scales may be different. For example, in the urban sustainability analysis, the city and district scales focus primarily on connectivity, but this loses its significance at the neighbourhood scale (Bourdic et al., 2012). Morphology plays an important role on the neighbourhood scale. And the segregation (social, residential, sectoral) that may have been masked by the indicator on the district scale may show up on the neighbourhood scale (Bourdic et al., 2012). For an American grid, the appropriate scale for neighbourhood is approximately $400m \times 400m$ to maintain the coherence of the urban fabric (Bourdic et al., 2012).

1.5.1 Urban Infrastructure

Urban infrastructure takes many forms. Engineered infrastructure, or "grey" infrastructure, is the capital investments that move or house people, goods, water, waste, and energy and the associated guidelines for their construction, operation, maintenance, and rehabilitation (Childers et al., 2015). Urban infrastructure that incorporates natural elements is often called "green" infrastructure. This includes traditional examples, such as parks, and street trees, but also includes more novel urban components such as community gardens and multi-purpose and multi-function storm water management facilities. Virtually all cities are also characterized by a variety of water features that provide a range of ecosystem services, including rivers and streams, lakes, and fountains. Collectively these are known as "blue" infrastructure. Because green and blue infrastructure features take advantage of natural structures and ecological processes, they are surprisingly adaptable to a changing future, and thus impart resilience to urban systems far more than do inertia-bound grey infrastructures (Childers et al., 2015).

1.6 Organization of the Dissertation

The dissertation is composed of six chapters. The chapters are organized as follows:

- Chapter 1 (this chapter) describes the background of the study, its purpose and scope and some important terminologies used in this dissertation.
- Chapter 2 provides an overview of the commonly used sustainability assessment methods and frameworks. It also gives a brief introduction of the methods that are used in the proposed framework which will be discussed in details in Chapter 3.
- Chapter 3 is the main content of this dissertation. It presents the proposed spatial sustainability assessment framework in detail, including indicators selection, data quality assessment, spatial analysis, Life Cycle Analysis / Assessment (LCA), Multi-Criteria Analysis (MCA) with Analytical Hierarchy / Network Process and Principle Component Analysis (PCA), assessment results analysis and visualization. The last section of this chapter demonstrates implementations of the proposed spatial sustainability assessment.
- Chapter 4 presents the proposed spatial sustainability assessment designed and implemented in Chapter 3 with the sustainability analysis of moving interstates within the Perimeter (I-285) underground. It gives a brief introduction of the study area and then presents all the data used in the sustainability assessment. After that, all the indicators and their evaluation methods are demonstrated, as well as the

discussion of evaluation results. The last section of this chapter makes conclusions from the assessment results and gives a vision of future works need to be done.

- Chapter 5 makes conclusions of the work conducted in this thesis.
- Chapter 6 provides some recommendations for future work.

CHAPTER 2. LITERATURE REVIEW

This chapter reviews the commonly used sustainability assessment methods and the sustainability assessment frameworks proposed by researchers or organizations. The methods used for sustainability assessment are first introduced and then the frameworks which may use one or multiple sustainability assessment methods are described. A summary of the reviewed methods and frameworks is provided at the end of this chapter.

2.1 Sustainability Assessment Methods

Sustainability assessment is often described as a process by which the implications of an initiative on sustainability are evaluated, where the initiative can be a proposed or existing policy, plan, program, project, piece of legislation, or a current practice or activity (Pope et al., 2004). The choice of sustainability assessment method(s) depends on the scope, assumptions, values, and precision of the assessment (Zijp et al., 2015). Because each sustainability assessment has its specific object of analysis, spatial and temporal dimensions, the required level of detail and required level of integration (Zijp et al., 2015).There are many sustainability assessment methods used by researchers or organizations, such as Life Cycle Assessment (LCA) (Finkbeiner, Schau, Lehmann, & Traverso, 2010; Heller, Keoleian, & Arbor, 2000; Onat, Kucukvar, & Tatari, 2014), Cost-Benefit Analysis (CBA) (Gasparatos, 2007; Hellström et al., 2000; Lautso, 2002) , Multi-Criteria Analysis (MCA) (Buchholz et al., 2007; Graymore et al., 2009; Singh et al., 2009), Principle Component Analysis (PCA) (Dong et al., 2015), Sustainability Impact Assessment (Ness et al., 2007; Pope et al., 2004) and so on. The most popular sustainability assessment methods are listed and reviewed in the following sections below.

2.1.1 Life Cycle Analysis

Life Cycle Analysis (LCA) is a systematic analytical method used to evaluate the resource consumption and environmental burdens associated with a product, through the whole life cycle from raw materials to final disposal of the product (or from "cradle to the grave") (Heller and Keoleian, 2000; Williams, 2009; Finkbeiner et al., 2010). LCA focuses on the biophysical impacts of a product system: resource depletion, energy consumption, water, and air pollution, human health impacts, and waste generation (Heller and Keoleian, 2000; Finnveden et al., 2009). It provides a comprehensive view of environmental aspects of the product or process and a more accurate picture of the true environmental trade-offs in product and process selection, such as carbon footprint assessment, water footprint assessment, environmental performance monitoring, and product design and eco-design (PRé Sustainability, 2017). SimaPro is a popular tool using LCA to assess the environmental impacts of a particular product. LCA typically does not address the economic or social aspects of a product, but the life cycle approach and methodologies can be applied to other aspects (International Standard Organization, 2006; Li et al., 2012).

There are four phases in an LCA study or analysis, as shown in Figure 2-1: Goal and Scope Definition, Life Cycle Inventory Analysis (LCI), Life Cycle Impact Assessment (LCIA), and Interpretation (see Figure 2-1). The Goal and Scope Definition includes subject, intended use and system boundaries of the study (Finnveden et al., 2009; International Standard Organization, 2006). LCI is an inventory of input/output data about the system being studied. LCIA aims to understand and evaluate the magnitude and significance of potential environmental impacts of the studied system (Finnveden et al., 2009; International Standard Organization, 2006). In the Interpretation phase, results from previous phases are evaluated concerning the goal and scope of the study, to reach conclusions and recommendations (International Standard Organization, 2006). The detail contents of each phase can be found in (International Standard Organization, 2006). Figure 2-2 is an example of using LCA for product system.

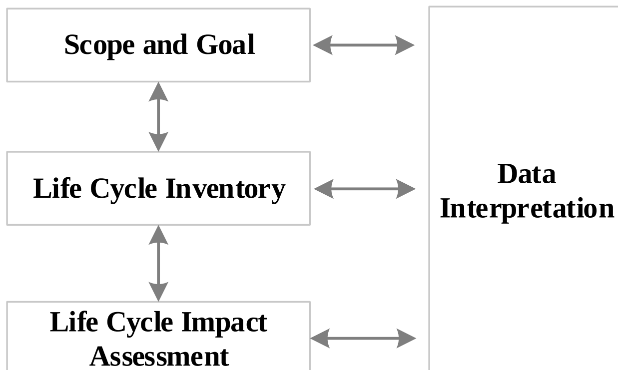


Figure 2-1 Phase of Life Cycle Analysis (Aida, 2009)

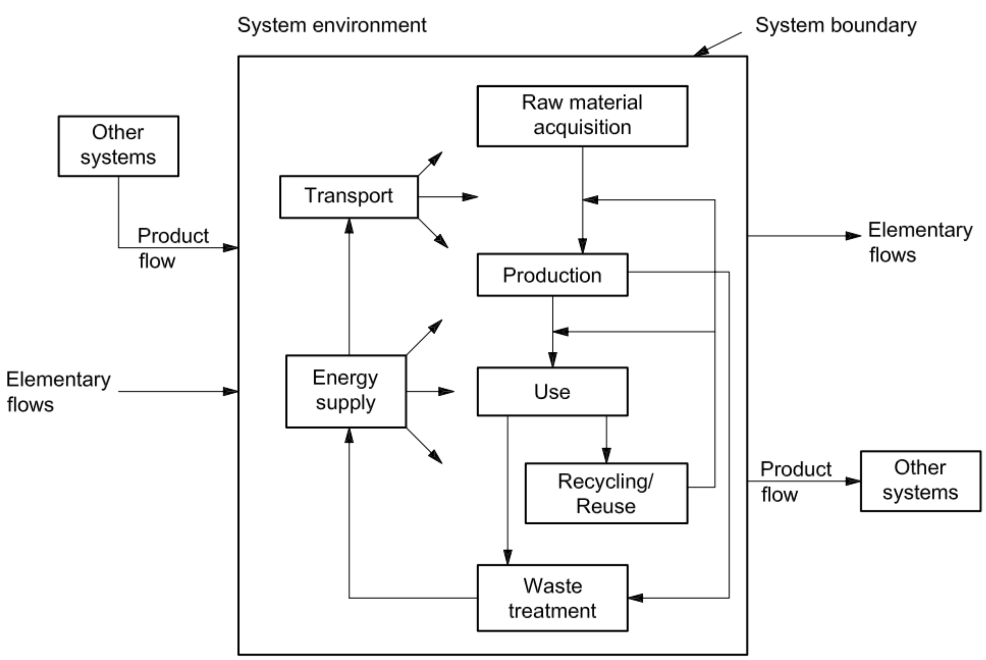


Figure 2-2 Example of LCA for a Product System for (ISO, 2006)

LCA can help researchers or users from the following aspects (Scientific Applications International Corporation, 2006; International Standard Organization, 2006):

- Quantifying environmental releases to air, water, and land concerning each life cycle stage.
- Assessing human and ecological effects of material consumption and environmental releases to the local community, region, and the world.
- Comparing the health and ecological impacts between two or more rival products/processes.
- Identify impacts to one or more specific environmental areas of concern, or opportunities to improve the environmental performance of products at various points in their life cycle.
- Selection of relevant indicators of environmental performance.

However, LCA analysis is complex and time-consuming, depending on how thorough the user wishes to conduct LCA. It requires large data, and availability of data can greatly affect the accuracy of results. Besides, its assessment results are subjective and can be affected by the definition of system boundary. Therefore, it is important to balance the availability of data, the time necessary to conduct the study, and required financial resources with projected benefits of LCA.

2.1.2 *Multi-Criteria Decision Analysis*

Multi-Criteria Analysis (MCA) typically use a structured and transparent methodology to score alternatives on each decision criterion and then aggregate the scores based on the relative weights of criteria (Golder Associates, 2011). MCA has been used effectively to develop decision support tools which could rank decision options considering multiple criteria (both qualitative and quantitative criteria) simultaneously for a range of planning and natural resource management issues to aid decision makers in making a well-informed decision (Doukas et al., 2012; Egilmez et al., 2015). It can bring together the sustainability criteria from all pillars (social, economic and environmental) to give an integrated assessment of sustainability(Graymore et al., 2009). MCA has been shown to be effective and has been used for sustainability assessment of many issues, such as agricultural systems (Lopez-Ridaura et al., 2002), policy alternatives (Nijkamp and Ouwersloot, 1997), infrastructures (EnvisionTM, 2015) , UK major ports (Asgari et al., 2015) , bioenergy systems (Buchholz et al., 2009), locating sustainable suburban centers (AbuSada & Thawaba, 2011), metropoles in the US. (Egilmez et al., 2015) and so on.

The big challenge of MCA is requiring stakeholders or decision makers to subjectively place importance (or weight) on each criterion (Graymore et al., 2009). To develop an accurate measure of sustainability for decision support, an objective method for developing weights based on the current understanding of sustainability is needed to ensure that the most accurate result is obtained. There are many methods for developing weights for indicators, such as equal weights, correlation analysis, Principal Component Analysis (PCA), multiple correspondence analysis, regression analysis, data envelopment analysis, unobserved component models, distance to targets, public opinion, budget allocation, Analytic Hierarchy Process (AHP), Analytic Network Process (ANP), conjoint analysis, etc. (Singh et al., 2007). Among them, PCA, AHP, and ANP are the most widely used (Buchholz et al., 2007) and they are more suitiable for constructing Composite Sustainability Index (CSI). The author will review these three methods in details in the following sections.

2.1.3 Principle Component Analysis

PCA was first proposed by Pearson in 1901 (Li et al., 2012). It is a multivariate statistical approach to reduce data dimensions by using linear orthogonal transforms of multiple highly correlated variables into a new set of independent variables, called “principal components” (Doukas et al., 2012). PCA can be done by eigenvalue decomposition or singular value decomposition of a data covariance matrix, usually after standardizing the attribute data (Li et al., 2012). It is effective in compressing data dimensions without too much information loss and can identify data patterns and highlight similarities and differences (Dong et al., 2015).

PCA is popularly used as a tool in exploratory data analysis and for making predictive models (Li et al., 2012). Lai (2003) used PCA to measure and to analyze the progress of human development in Chinese provinces since 1990. Jollands et al. (2004) used PCA to derive eco-efficiency indices for New Zealand. Adler et al. (2010) combined PCA and DEA to measure the relative socioeconomic performance of developing countries. Doukas et al. (2012) used PCA to assess energy sustainability of rural communities in Europe. Dong et al. (2015) used PCA to evaluate the sustainability of the natural gas industry in China.

PCA is also frequently used to construct “Sustainable Development Indicators” (SDIs) (Doukas et al., 2012) or Composite Sustainable Index (CSI). Singh et al. (2009) list many sustainability indicators constructed using PCA. But They said policy-makers do not use most of those indices due to measurement, weighting and indicator selection problems. Among them, Human Development Index (HDI), Ecological Footprint, Index for Sustainable Economic Welfare (ISEW), Genuine Progress Indicator (GPI) and Environmental Pressure Indicators (EPI) have been computed by researchers for a number of countries under different assumptions due to the variation in data quality and availability (Singh et al., 2009). When using PCA to construct composite sustainability indicators, researchers usually construct environmental, economic, and social sustainability composite indicators by PCA separately and then combine them together to form a CSI for overall sustainability performance analysis like the example demonstrated by (Li et al., 2012; Soler-Rovira and Soler-Rovira, 2009).

However, when using PCA to build the CSI, it is difficult to explain the contributions or influence of each indicator to the CSI. Besides, PCA is sensitive to relative scaling of the original variables (Dong et al., 2015), modification, revisions, updates, the presence of outliers, small sample problem and data shortage. Furthermore, PCA is not suitable for non-linear combinations (Singh et al., 2007).

2.1.4 *Analytical Hierarchy Process*

The AHP proposed by Saaty (1987) is used to derive the ratio scale from both discrete and continuous paired comparisons. These comparisons may be taken from actual measurements or from a fundamental scale which reflects the relative strength of preferences and feelings (Saaty, 1987). AHP is a nonlinear framework for carrying out both deductive and inductive thinking by taking several factors (both quantitative and qualitative measures) into consideration simultaneously and allowing for dependence and feedbacks from users (Saaty, 1987; Singh et al., 2007). It has been accepted as a leading multi-attribute/ multi-criteria decision model both by practitioners and academics (Krajnc and

Glavič, 2005). AHP enables all members of the evaluation team to visualize the problem systematically in terms of criteria and sub-criteria (Singh et al., 2007). It can also be easily understood and applied by operating managers or policy makers. Therefore it is suitable for sustainability assessment and has been used by many researchers. For example, Krajnc and Glavič (2005) applied AHP to construct Composite Sustainable Development Index for companies. Singh et al. (2007) used AHP to develop Composite Sustainability Performance Index for the steel industry. Ugwu and Haupt (2007) used AHP to determine weights of indictors for MCA to assess infrastructure sustainability of South Africa construction industry. Asgari et al. (2015) used AHP to investigate the sustainability performance of five major UK ports.

Krajnc and Glavič (2005) give a detail illustration of how to calculate composite sustainability indicators using AHP. There are three main steps to conduct AHP. First, we need to set the problem or subjective of sustainability assessment as a hierarchy of more easily comprehended sub-problems, each of which can be analyzed independently (Asgari et al., 2015). At the top of the hierarchy is the overall objective of sustainability assessment. The criteria are on the next level, which can be decomposed to the sub-criteria and further decomposed to the lower levels (Kadoić et al., 2013). On the last level are the indicators. Figure 2-3 gives an example of a generic hierarchy scheme of indicators for calculation of composite sustainable development index.

Second, we make pair-wise comparisons between each pair of indicators at the same level of the hierarchy and assign weights to reflect their relative importance (Singh et al., 2007). The intensity of preference is expressed on a factor scale from 1 to 9 (see

Table 2-1 for details). These pair-wise comparisons result in a $(N \times N)$ positive reciprocal matrix A. Where N is the number of indicators at the same level of hierarchy.

Finally, we can calculate the normalized weight of each indicator or criteria based on matrix A. Then we can use the weights to calculate the composite sustainable indicators.

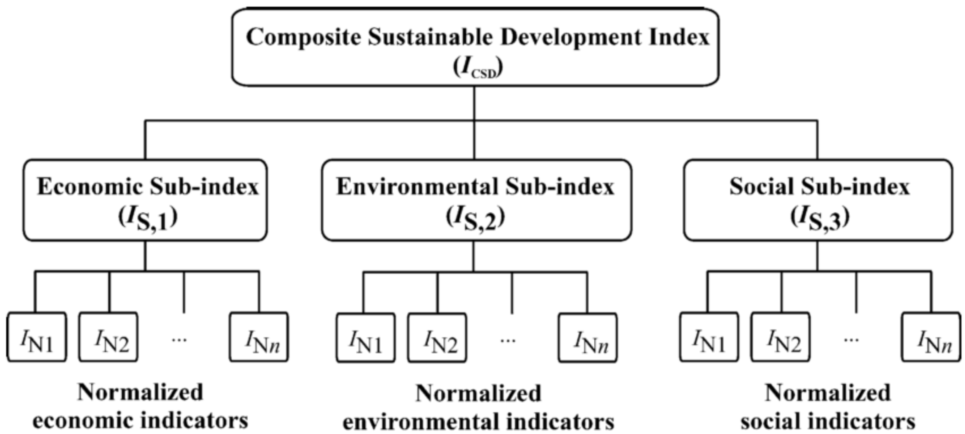


Figure 2-3 Generic Hierarchy Scheme for Calculation of Composite Sustainable Development Index (Krajnc & Glavič, 2005)

Table 2-1 Comparison Scale of AHP (R. W. Saaty, 1987; T. L. Saaty, 2004)

Factor of Preference, ρ	Importance Definition
1	Equal importance: two indicators contribute equally to the objective
3	Moderate importance of one over another
5	Strong or essential importance of one over another
7	Very strong or demonstrated importance of one over another
9	Extreme importance of one over another
2, 4, 6, 8	Intermediate values, use this when compromise is needed
Reciprocals, 1/ρ	Reciprocal for inverse comparison. If one indicator *i* has one of the above numbers (factor of preference) assigned to it when compared with indicator *j*, then *j* has the reciprocal value when compared with *i*.

Inconsistency is likely to occur when decision-maker makes careless errors or exaggerated judgments during the process of pair-wise comparison. AHP uses the consistency ratio (CR) to check the consistency of each judgment (Asgari et al., 2015; Krajnc and Glavič, 2005). If the consistency ratio is less than 0.1, then all the judgments are consistent. Otherwise, some judgements are inconsistent, and the decision-maker has to re-evaluate his or her judgments in the pairwise comparison matrix until the ratio is less than 0.1 (Krajnc and Glavič, 2005).

As illustrated in the previous steps, in order to use AHP to construct composite sustainability index or to determine the weights of indicators, all the indicators or criteria must be represented in hierarchy way (Singh et al., 2007). However, sometimes dependencies or inter-linkages exist between criteria, which AHP does not consider. In this case, using ANP is more appropriate. Besides, stakeholders must be able to provide a pair-wise comparison given any two evaluation indicators, and they should never decide that one indicator is infinitely superior to another (Singh et al., 2007). When there are many indicators, the pair-wise comparisons can be huge and takes lots of time.

2.1.5 *Analytical Network Process*

ANP provides a general framework to deal with decisions without making assumptions about the independence of higher-level elements from lower level elements and the independence of the elements within a level as in a hierarchy (like AHP) (Saaty, 2004b). ANP is a generalization of AHP. Both of them use a ‘‘ratio scale theory’’ based on a pairwise comparison of criteria and a subsequent ratio scale estimation for each criterion, usually employing a nine-point scale (see Table 2-1) (Saaty, 1996).

As mentioned before, the interactions among different dimensions or aspects of sustainability and the cause-effect relations between indicators are the fundamental basis of sustainability; thus it is very important and necessary to consider the interactions in the sustainability assessment model. From this point of view, ANP could better describe sustainability assessment problem than AHP. Many researchers have tried to use ANP in their process of evaluating sustainability. For example, Bottero et al. (2007) use ANP to assess the sustainability of an urban transformation project in Italy, compare three transformation alternatives based on results given by ANP. Wang et al. (2010) combine ANP with fuzzy Delphi method to evaluate sustainable development of housing community. Medel-González et al. (2016) combine MCA with ANP to measure the sustainability performance in Cuban organizations.

However, there are some disadvantages when applying ANP into the evaluation of sustainability, due to the difficulty of the consistence of pair-wise comparison (Kadoić et al., 2013), especially when many experts or stakeholder engaged into the assessment and number of indicators or criteria are large. Besides, when combing MAC with ANP, it is very difficult to conduct sensitivity analysis, cause ANP process is time-consuming and need the participates of experts or stakeholders (Medel-González et al., 2016)

2.2 Sustainability Assessment Frameworks

Sustainability assessment frameworks provide a way to conceptualize sustainability at a high-level (Waheed et al., 2009; Zijp et al., 2015). It helps to clarify what to measure, what to expect from measurement and what kind of indicators to use (Waheed et al., 2009) when estimating sustainability. Several conceptual frameworks have been proposed and developed in various disciplines, such as geothermal energy (Shortall et al., 2015), urban planning (Alshuwaikhat and Aina, 2006), urban 3-waters infrastructure (Kettle, 2006), forested landscape (Prato, 2015), transportation system (Richardson, 2005), European Environment Agency (EEA) (Stanners et al., 2007) and so on. Each of these frameworks has limited capability to deal with different issues of sustainability comprehensively, lacks the flexibility to be used in various disciplines with a unified interpretation (Waheed et al., 2009) and are not easy to be used or understand by policy makers. The main differences among frameworks are how they conceptualize the main dimensions of sustainability, inter-linkages between these dimensions, ways to select, measure and aggregate indicators (Waheed et al., 2009). Table 2-2 provides a brief overview and main features of commonly used frameworks. Among them, Linkage-based, impact-based, LCA and Process/Stakeholder-based are the most popular used framework. Driving force-Pressure-State-Impacts-Response (DPSIR), Triple-Bottom-Line (TBL), LCA and stakeholder-based methods (bolded ones in Table 2-2) are being considered in the proposed framework in this study. Because DPSIR described the inter-linkages between indicators or different aspects of sustainability, TBL is useful for assessing the impacts of an activity on the economy, environment, and general social well-being. LCA could prevent us under-estimate environmental impacts of the assessment object, while stakeholder-based framework involves stakeholders into the process of sustainability assessment, which can lead to more effective and enduring solutions for sustainability and present opportunities to educate the public and influence collective behaviors (Waheed et al., 2009). Combining advantages of these four frameworks into the proposed framework in this study will make it a powerful framework with systematic methods which consider all aspects of sustainability as well as the linkage between different dimensions, can be general enough to be used in many disciplines or areas easily and effective in communication with policy makers or stakeholders.

Table 2-2 - Main Features of Sustainability Framework (Waheed et al. 2009)

Frameworks	Main Features
Linkage-based	• Use concepts of causality (Jeon and Amekudzi, 2005) • Can be tied to sustainability through certain assumptions Difference forms include Pressure-State-Response (PSR), **DPSIR,** Driving force-Pressure-State-Exposure-Effect-Action (DPSEEA)
Impact-based	• Reactive in nature • Reductionist approach to sustainability • Focuses on the impacts of various actions on the sustainability of a particular system. • A typical example is **TBL analysis** (e.g., Global Reporting initiative with five dimensions, UN-CSD with four dimensions. Also used in a various engineering discipline, e.g., Transportation (Khan et al., 2002; Litman, 2009), water and sewer system (Ashley and Hopkinson, 2002), building infrastructure (Pearce and Vanegas, 2002))
Objective-based	• Proactive framework • Ensures that a particular initiative contributes to a defined state of sustainability • Form a part of the majority of present frameworks (for example, Strategic Environmental Assessment (SEA))
Influence-based	• Indicators categorized by their level of influence on the sustainability of an organization or institution
Process/Stakeholder-based	• Involves extensive planning process that engages stakeholders • Extensively used for planning of community projects • E.g., Multi-Stakeholders Process used in Environmental Sustainability Kit

Material Flow Assessment / **Life Cycle Assessment**	• Material Exchange between economy and natural environment • Cradle to the grave assessment of environmental impacts • Commonly used in the chemical industry (Khan et al., 2004) • E.g. Life cycle iNdeX (LInX) (Khan et al., 2004)

2.2.1 *Driving Force Pressure State Impact Response (DPSIR)*

DPSIR framework is an extension of the PSR framework and has been adopted by the European Environmental Agency (EEA) and the European Statistical Office in 1997 (Singh et al., 2009). The DPSIR framework has been used to structure the interplay between the environment and socioeconomic activities by most member states of the European Union (EU) and by many international organizations including the European Environmental Agency and EUROSTAT (the statistical office for the European Communities) (Waheed et al., 2009).

From the DPSIR system analysis view, social and economic developments drive changes that exert pressure on the environment; consequently, changes occur in the state of the environment. This leads to impacts on factors, such as human health, ecosystem functioning, materials (such as historic buildings), and economy, where impacts refer to information on the relevance of changes in the state of the environment. Finally, societal responses are made that can affect earlier parts of the system directly or indirectly (Stanners et al., 2007). In DPSIR framework, there is usually a chain of causal links starting with ‘driving forces’ (e.g. economic sectors, human activities) through ‘pressures’ (e.g. emissions, waste) to ‘states’ (e.g. physical, chemical and biological) and ‘impacts’ on ecosystems, human health and functions, eventually leading to political 'responses' (e.g. prioritization, target setting, policies). Many researchers explained each part of DPSIR in details and gave many examples, such as Maxim et al. (2009) and Stanners et al. (2007). They will not be repeated here. A typical example of the DPSIR framework for water issue is shown in Figure 2-4 below.

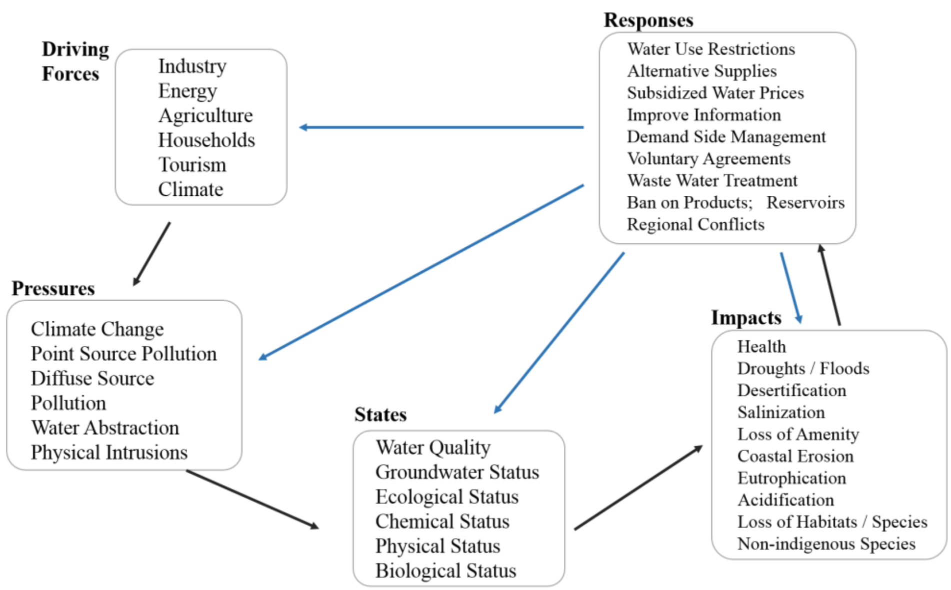


Figure 2-4 DPSIR Framework Used for Water Issue (Adopted from (Kristensen, 2004))

DPSIR is useful in describing the relationships between the origins and consequences of environmental problems (Kristensen, 2004), and can help identify the sustainability assessment indicators (Stanners et al., 2007). However, the simple causal relations cannot capture the complexity of interdependencies in the real world (Maxim et al., 2009).

2.2.2 Driving Force-Pressure-State-Exposure-Effects-Action (DPSEEA)

DPSEEA can be viewed as a modified version of DPSIR. DPSEEA framework, proposed by the World Health Organization (WHO), provides a broader approach to include impacts of macro driving forces and pressures on both health and the environment. And all sectors including government, private sector and individuals can contribute to the outcomes at all levels, and this information can be used to provide feedback at all levels. DPSEEA framework has been widely used in the environmental health sector. In combination with multi-criteria decision-making, this framework has a great potential to contribute significantly to sustainability analysis (Waheed et al., 2009). A typical example of the DPSEEA framework is shown in Figure 2-5 below.

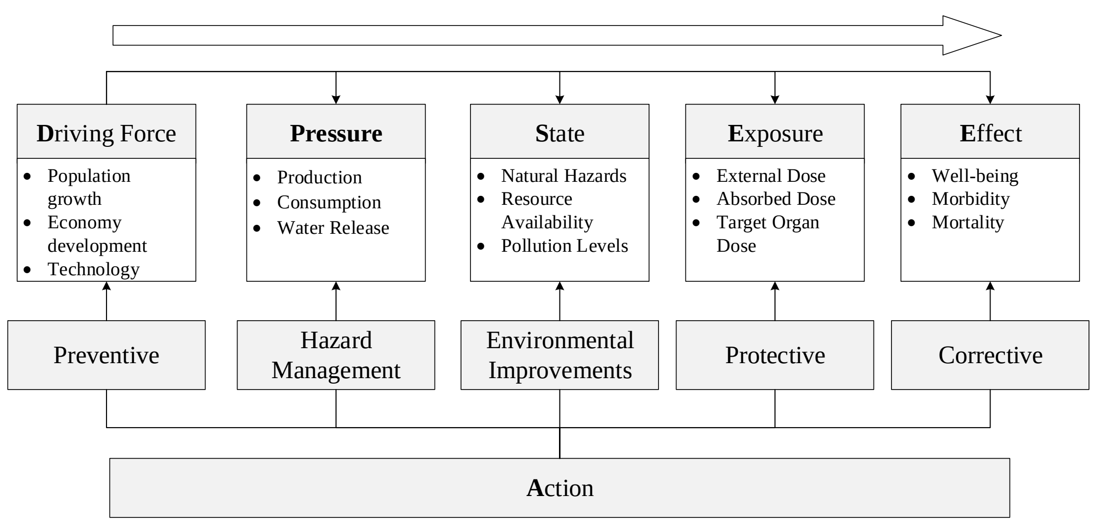


Figure 2-5 DPSEEA Framework (Waheed et al. 2009)

However, the simple unidirectional linkages among issues or factors used in DPSEEA are not very conducive to understanding and describing the complexity of the processes behind sustainability assessment. This is a limit of all linkage-based frameworks. Besides, DPSEEA cannot work effectively if the evidence for causal linkages is missing or vague. It could also lead to oversimplification of spatial and temporal interactions that results in poorly informed management decisions.

2.2.3 Triple-Bottom-Line (TBL)

TBL is a framework proposed by John Elkington to measure performance in corporate America in the mid-1990s (Slaper and Hall, 2011). This new framework can be considered a reductionist approach to sustainability that places equal importance on environmental, social and economic considerations in decision-making (Pope et al., 2004). TBL uses a bottom-up approach to define sustainability performance criteria. The criteria are generated by assuming that the state of sustainability can be defined by environmental, social and economic objectives and proposed criteria are developed under these categories. Under the TBL framework, it is much easier to use MCA to evaluate sustainability (Waheed et al., 2009). Many researchers use the TBL framework to evaluate the sustainability of their objects; the typical examples are listed in Table 2-2.

However, the concept of sustainability into the three pillars of TBL tends to emphasize potential competition for interests rather than the linkages and interdependencies between different aspects of sustainability (Pope et al., 2004). Besides, TBL divides the holistic concept of sustainability into three pillars as a starting point invariably runs the risk of “the sum of parts is less than the whole”. This is particularly true if the interrelations between the three pillars are not adequately understood and described, and therefore sustainability is reduced to a consideration of separate environmental, social and economic factors, the sum of which is less than the whole, that is, sustainability (Pope et al., 2004).

2.2.4 Stakeholder-Based Frameworks

A process-based framework involves a planning process that effectively engages stakeholders in creating their vision for sustainability. It involves all the representatives from various constituencies within a community, based on a decision aiding process for developing consensus (Environmental Defense, 1999). Engagement of stakeholders into the process of sustainability assessment is an important and critical component to achieve sustainability objectives (Waheed et al., 2009). Because stakeholders provide necessary authority, prospective, prestige, talent, and resources to develop effective solutions for sustainability or sustainable development and implement them (Environmental Defense, 1999). Velazquez et al. (2006) even proposed a model to study how people responsible for sustainability initiatives affect collective behavioral change by educating stakeholders and promoting consensus-based sustainability goals for sustainable institutions.

2.2.5 GoldSET Evaluation Framework

GoldSET (Golder's "Sustainability Evaluation Tool") framework is a combination of the impact-based and process-based framework, which is shown in Figure 2-6 below. It uses a simple, but a systematic process to evaluate the sustainability of project alternatives across three key dimensions of sustainable development (environment, society, and economics). The simplified multi-criteria analysis was used to compare project options by scoring performance indicators, based on an option's relative performance and relevance of indicators, and then presents results in a radar chart (see Figure 2-7) that clearly illustrates the strengths and weaknesses of each option across the three dimensions of sustainable development (Golder Associates, 2011). GoldSET can facilitate effective communication between stakeholders and project proponents and can help project proponents address stakeholders' concerns because its decision process is explicit and transparent. However, it does not provide many details about the data used in the evaluation and not consider the inter-linkages between different aspects or dimensions of sustainability. Besides, it has limited spatial and temporal capacities.

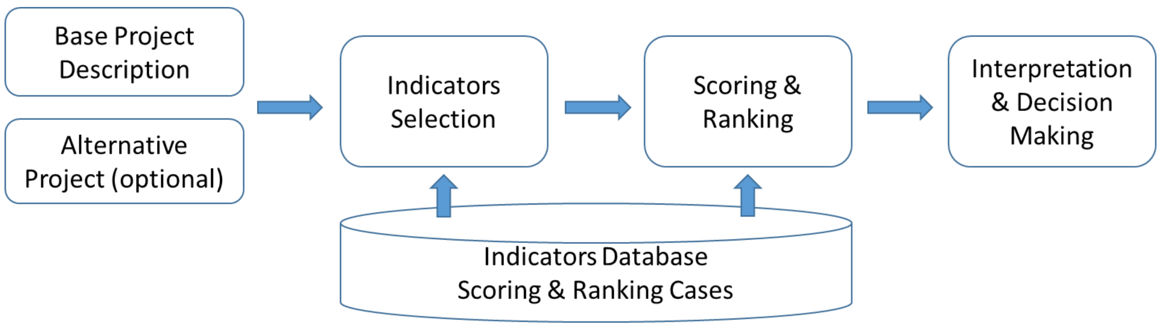


Figure 2-6 GoldSET Evaluation Framework (adapted from (Golder Associates, 2011))

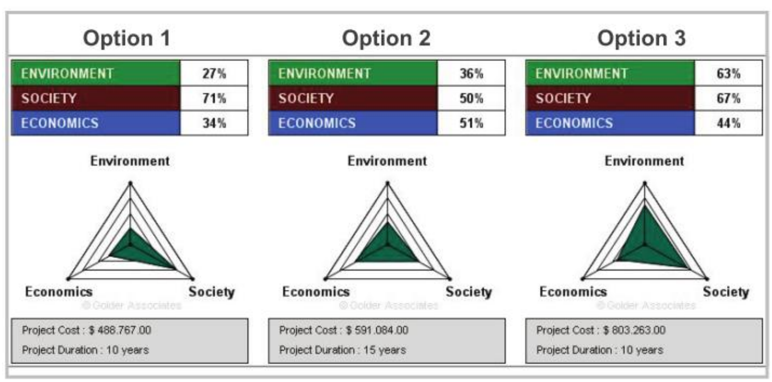


Figure 2-7 Sustainability Assessment Results for Three Options in GoldSET (from (Golder Associates, 2011))

2.2.6 Spatial Frameworks

As mentioned before, sustainability has natural spatial aspect, and its spatial and temporal scales or aspects need to be considered into the sustainability assessment framework. The spatial assessment of sustainability should consider the spatial difference of distribution or allocation of natural and humanmade resources, population, economic development level, energy consumption structure, and culture, as well as spatial patterns and spatial locations of infrastructures. Geographical Information Systems (GIS) is a very good framework for gathering, managing, analyzing and presenting spatial or geographic data. It can integrate many types of spatial data, such as geographical data, socio-economic data, and attribute data, or even combine using data from different sources, such as remote sensing images, data from Census Bureau, and data in different Coordinate Reference System (CRS) or spatial scales. It also has powerful spatial analysis functions, which can reveal a deeper insight into spatial data, such as spatial patterns, relationships and spatial distribution of economic and related activities.

GIS based sustainability assessment framework could make the sustainability assessment more accurate and presents the results in a format that can be easily used by other users or in other data analysis tools (e.g., we can present the sustainability assessment results by maps.). It can also be effective in identifying areas most in need of initiatives to progress sustainability for regional managers (Graymore et al., 2009).

As said in the last section, we can combine different sustainability assessment frameworks mentioned in Table 2-2 to build an intergraded framework. Graymore et al. (2009) proposed a GIS based framework combing MCA, AHP, and TBL to assess the sustainability of sub-catchments in the Glenelg Hopkins Catchment, Australia. Xu and Coors (2012) propose an integrated sustainability assessment framework which includes GIS analysis, System Dynamics (SD) model, DPSIR and 3D visualization. It covers sustainability assessment of urban residential development and spatial distribution analysis of residential buildings based on sustainability indicators categorized in four groups, namely housing, society, economics, and environment (Xu and Coors, 2012). Yigitcanlar and Dur (2010) developed an indicator-based comparative urban sustainability assessment model: The Sustainable Infrastructure, Land-use, Environment and Transport Model (SILENT). SILENT considers the sustainability of land-use, environment, transport systems and infrastructure with a TBL approach, and uses MCA to generate the integrated sustainability index, following four logical steps similar to the OECD's Composite Indicators Methodology (OECD, 2008). It uses a grid-based system and divides the study area into small cells (e.g., 100m*100m), and then conduct analysis and visualize the results in ArcGIS. Alshuwaikhat and Aina (2006) integrated GIS with MCA to assess urban sustainability in Dammam City, Saudi Arabia. Their analytical framework is shown in Figure 2-8.

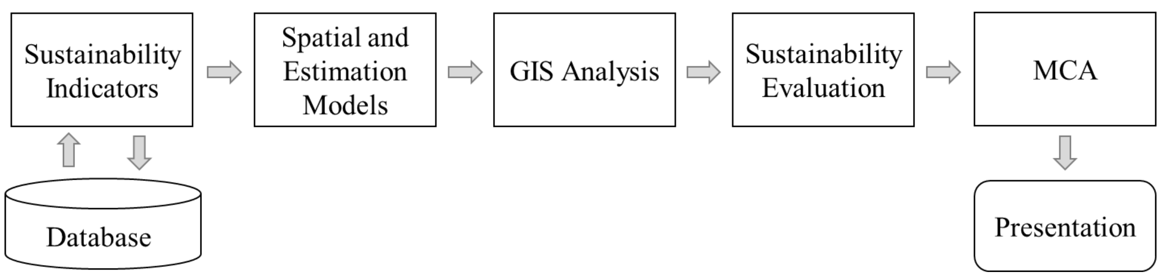


Figure 2-8 Analytical Framework for Evaluating Urban Sustainability (Adapted from (Alshuwaikhat & Aina, 2006))

However, most of them develop the spatial sustainability assessment tools based on ArcGIS, which is commercial software. It is better to use open source GIS software (e.g., QGIS) or to include the commonly used GIS functions in the framework directly and to provide an interface to use open source GIS software. Besides, most of them integrate

MCA and AHP to make the spatial sustainability assessment framework. However, the tool does not include LCA and the interlinkages between different aspects of sustainability, which may make the sustainability assessment not complete or accurate.

2.3 Summary

In this chapter, the author reviewed the most popular sustainability assessment methods and frameworks. Each method has its advantages, but MCA is considered the best suite for sustainability assessment, considering the concept of sustainability. Because MCA could consider multiple criteria (both qualitative and quantitative criteria) simultaneously for a range of planning and natural resource management issues to aid decision makers in making a well-informed decision. It requires stakeholders or decision makers to subjectively place importance (or weight) on each criterion, which may lead a not objective assessment results. Therefore, we can use PCA, AHP or ANP to make the weight of indicators more objective and consistent. LCA is widely used in the environmental impacts assessment of products. It could prevent us underestimate the environmental impacts of the sustainability assessment object, but it is very time-consuming, and assessment results can be subjective due to the definition of system boundary. We need to find a good balance of time and benefits of LCA when using it in the sustainability assessment.

Sustainability assessment frameworks provide a way to conceptualize sustainability at a high-level. It clarifies what to measure, what to expect from measurement and what kind of indicators to use when evaluating sustainability. There are many frameworks proposed by researchers, institutions or organizations. Among them, the author mainly reviewed DPSIR, DPSEEA, Stakeholder-based frameworks, GoldSET, and Spatial

Frameworks. DPSIR and DPSEEA describe the relationships between the origins and consequences of environmental problems and can help identify sustainability assessment indicators. But the simple causal relations described by them cannot capture the complexity of interdependencies in the real world. Besides, DPSEEA could also lead to oversimplification of spatial and temporal interactions, which would result in poorly informed management decisions. TBL make it easier to use MCA to evaluate sustainability, but it tends to emphasize potentially competing for interests (three pillars) rather than the linkages and interdependencies between different aspects of sustainability. The stakeholder-based framework involves stakeholders into the process of sustainability assessment, which can lead to more effective and enduring solutions for sustainability and present opportunities to educate the public and influence collective behaviors (Waheed et al., 2009). GoldSET is an integrated framework which combines impact-based and process-based framework. It uses MCA to aggregate indicators and provides a good and easy way to present the sustainability assessment results. However, GoldSET does not consider the inter-linkages between different aspects or dimensions of sustainability and has limited spatial and temporal capacities. Spatial frameworks consider the spatial part of sustainability in the assessment. But most of the spatial frameworks based on ArcGIS, which is commercial software. It is better to use open source GIS software, like QGIS or integrate the frequently used spatial analysis functions into the spatial framework directly. Even many of them use MCA and AHP in the assessment but does not consider the inter-linkages between different aspects of sustainability. Therefore, the proposed framework in this study, which will be discussed in detail in the next chapter, will combine advantages of DPSIR, LCA, TBL, Stakeholder-based framework and spatial framework. The proposed spatial sustainability assessment framework with provide systematic methods (integrate MCA, AHP/ANP) to evaluate sustainability, which consider all aspects of sustainability as well as the linkage between different dimensions, can be general enough to be used in many disciplines or areas easily and effective in communication with policy makers or stakeholders.

CHAPTER 3. PROPOSED SPATIAL SUSTAINABILITY ASSESSMENT FRAMEWORK

The proposed spatial sustainability assessment framework integrates DPSIR, TBL, LCA, and stakeholder-based methods. It can make full use of the advantages of each method and can present the sustainability assessment results from different aspects in multiple ways. The proposed spatial sustainability assessment framework is shown in Figure 3-1. There are eight components in this proposed framework, including “Define Object & Scales”, “Select Indicators”, “Select Data”, “Assess Data Quality”, “Data Exploration”, “Spatial Sustainability Assessment”, “Results Analysis”, and “Results Visualization & Interpretation”. The order of components indicates the process of spatial sustainability assessment. The first component of this framework is “Define Object & Scales”, which defines the objective and scales of the sustainability assessment, for example, defining the type of infrastructure, its location, its timeline, the dimensions of sustainability assessment, and the scales of the assessment (spatial and temporal scales). The scale determines the level of detail and complexity of the assessment work. Sustainability in the framework has four dimensions, which are environment, economy, social, and resilience. Then the user can select indicators from the spatial sustainability assessment database, define indicators by themselves, or select/define indicators with the help of DPSIR. Next, the user selects the data used to evaluate the sustainability indicators. The data can be selected from local files or the spatial sustainability assessment database in the framework. After that, the “Assess Data Quality” module will check the quality of data, before using the input data to conduct spatial sustainability assessment. If the data quality is good enough, then the user can use it to evaluate the sustainability of their object. After finishing the sustainability assessment, the user can analyze the uncertainty and sensitivity of the assessment. In the end, the user can visualize the assessment results in multiple ways and save the results into the spatial sustainability assessment database. Following sections in this chapter illustrate the design and implementation of each component in the proposed spatial sustainability assessment framework.

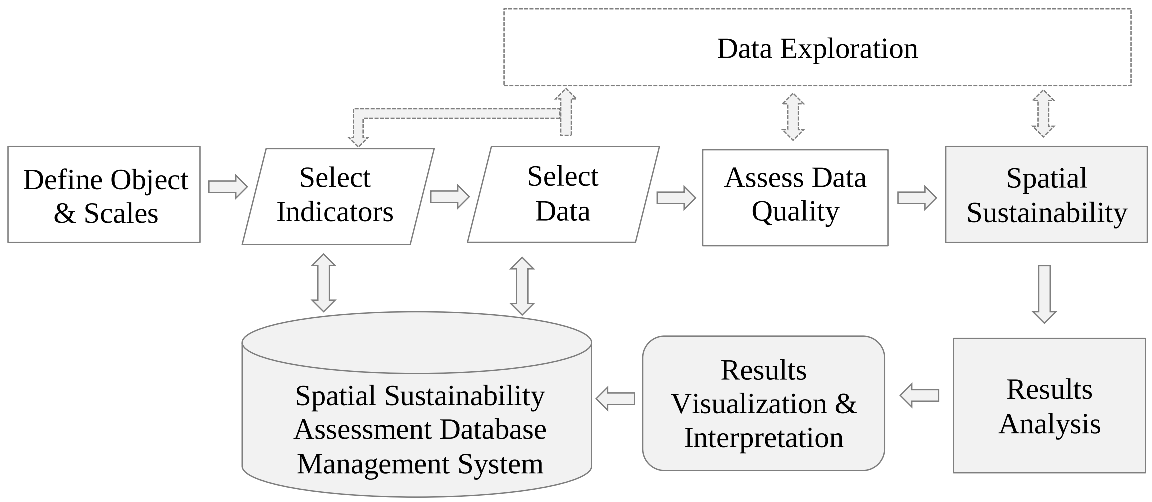


Figure 3-1 Proposed Spatial Sustainability Assessment Framework

The big advantages of the proposed sustainability assessment framework is that it combines elements of different approaches and takes advantage of these to build a better hybrid framework. The proposed framework not only considers all aspects of sustainability (environmental, economic, social, and resilience), but also includes inter-linkages between different aspects and their dynamic changes in the process of sustainability assessment. The proposed framework provides a clear way to conceptualize sustainability at a high-level. It clarifies the process of evaluating sustainability and principles or methods to select indicators, to choose data used for the assessment, to measure, analyze and visualize indicators. This framework provides integrative and systematic methods to assess the sustainability of infrastructure from multiple dimensions at different spatial and temporal scales. Thus it encourages or pushes the stakeholders or policy-makers focusing on the sustainability contributions of one project when making decisions, as opposed to only focusing on the economic benefit and cost. For example, the widely used "cost-benefit" analysis/model makes people consider more about the money benefit/cost of one project, rather than its other "potential" and not directly visible benefits, such as improvement of people's quality of life (happiness, beauty) and equity. Besides, it engages policymakers and other participants in the process of sustainability evaluation, by letting them set the weight of indicators with the help of AHP. The proposed framework also allows for the visualization of the assessment results from different perspectives in multiple ways. It also has a spatial database to save and manage the data, indicators and other information used in the process of sustainability assessment. The framework is a data-driven framework; thus it can be easily used in other disciplines and other areas.

3.1 Define Object and Scales

"Define Object & Scales" is the first component in the proposed framework and the starting point of sustainability assessment. In this component, users need to define the objective of the sustainability assessment, such as "evaluating the sustainability condition of the City of Atlanta" "assessing the sustainability performance of the transportation system in the State of Georgia". The objective describes what system or subject one wants to the sustainability of (e.g., a city, a state, a transportation system), the study area or location, timespan, the purpose of this sustainability assessment, and which dimension of sustainability is considered in the sustainability assessment. Sustainability in the proposed framework has four dimensions, which are economic, environmental, social, and resilient. It is not necessary to include all the four sustainability dimensions in one's sustainability assessment. Users need to determine which sustainability dimension they want to be included in the sustainability assessment based on the objective of their sustainability assessment.

Scales in the proposed framework includes both the spatial scales and temporal scales. The scales determine the level of detail and complexity of the assessment work. They determine the indicators and data used for the sustainability assessment and tell the accuracy of the assessment results. The spatial scale defines the extent of an area at which the sustainability indicators and CSI is evaluated (i.e., the smallest area that one can see the detail information of sustainability performance of its sustainability assessment object). It describes the level of detail that one wants to consider in the data of his or her study area. For example, one may want to evaluate the sustainability of the City of Atlanta at census tract scale; then each census tract has a CSI value and individual sustainability indicator value. The proposed framework provides 11 different scales; Global, National, Regional, Division, State / Province, County, Census Tract, Census Block Group, Census Block, Traffic Analysis Zone, and user-defined scale. Things change with time, as well as sustainability; the temporal scale defines the time span in which one wants to see the status and changes of sustainability performance. It determines how much further one wants to see for the sustainability assessment of his or her study area. For example, if the temporal scale is 10 years, then the sustainability assessment only considers the environmental impacts of one project has in 10 years.

When choosing the spatial and temporal scales, the following criteria are used in the proposed framework:

- The objective of the sustainability assessment: scales defined for the subsequent sustainability assessment must be related to the objective of the sustainability assessment. For example, if one wants to evaluate the sustainability condition of one state; then the spatial scale can be State or Counties.
- Sustainability dimensions included in the sustainability assessment: can be a user-defined scale and represent the sustainability performance for the user-selected sustainability dimensions?
- Accessibility of data: Is the data at the defined scale available? Can one obtain sufficient data for the subsequent sustainability assessment? Can one obtain most of the needed data at the defined scale? For example, one needs to get 10 different data set to evaluate the sustainability of one city, but most of the data available are at the state scale, then he or she may need to reconsider the spatial scales of the sustainability assessment because it is likely that he or she cannot get reasonable or reliable sustainability assessment results based on his/ her data.
- Amount of data: is there sufficient space to store and manage the data? Because smaller scale includes more detail information and needs more space to store and manage the data.
- Time limitation: does the user have enough time to obtain and process the data at the defined-scale? It takes more effort and time to obtain and process smaller scale data.

- Budget limitation: does the user have sufficient money to obtain the data or buy the service to process or store the data?

3.2 Select Indicators

Indicators are base to the sustainability assessment. Selection of appropriate impact indicators is the biggest challenge to sustainability assessment. It requires a balance between simplification and complication (Singh et al., 2009). There are some ways to select appropriate indicators, such as learning from experts' experience, or with the help of DPSIR (Kristensen, 2004). When selecting indicators for the spatial sustainability assessment, the following criteria are considered in the proposed framework.

- The indicators should be available and measurable (Hellström et al., 2000).
- The indicators need to be comparable across time and space (Ebert and Welsch, 2004) so that we can test future policy impacts (Lautso et al., 2002), changes of sustainability with time or compare the sustainability among different locations.
- The indicators need to be relevant to the objective of sustainability assessment (Hak et al., 2016).
- The indicators should represent all necessary sustainability domains (Kettle, 2006; Yigitcanlar and Dur, 2010), including environmental, economic, social and resilient dimension, and whole life cycle of the objective of sustainability assessment.
- The indicators need to be easily understood (or interpreted) and communicated by decision-makers and other target audiences (Li et al., 2012; Soler-Rovira and Soler-Rovira, 2009).

When using the "Select Indicator" module to select or define indicators, users need to first determine the sustainability dimension for the indicator, and then select or define other information of their indicators (e.g., the name of the indicator, category, subcategory, spatial or temporal scales, measurement reference, and so on). The organization of current indicators in the proposed framework are shown in Figure 3-2 and Figure 3-3. Indicators in one sustainability dimension are classified into different categories, and indicators in one category are further classified into multiple subcategories, which include multiple individual indicators (see Figure 3-3).

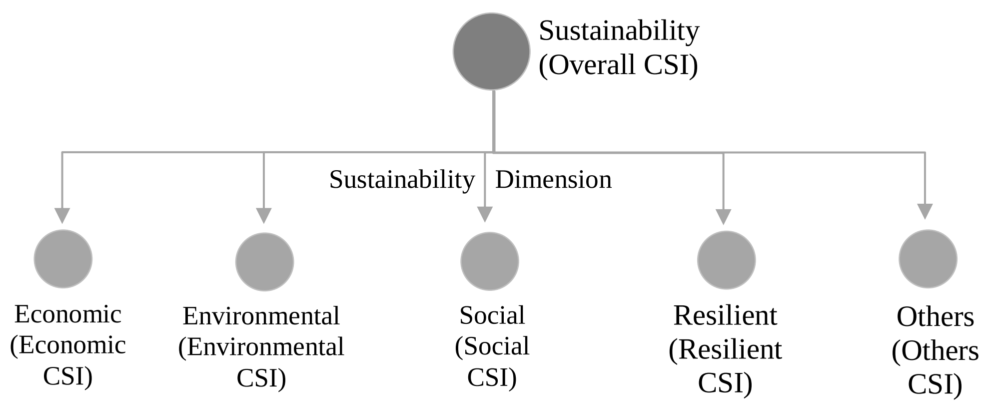


Figure 3-2 Organization of Indicators: Dimension of Sustainability

There are many other researchers who organize their sustainability indicators in this way, examples can be found in (Ochsenbein and Wachter, 2004), (Ugwu and Haupt, 2007), and (Wei et al., 2007). Sustainability in this framework can have four dimensions, which are the economic, environmental, social, and resilient dimension. However, it is very difficult to distinguish indicators based on their dimension, because these indicators describe the interactions among different aspects of sustainability. Thus, we add an "others dimension" to include all the other indicators that can not be described by each of the four dimensions, such as the indicators that include information of multiple sustainability dimensions, and the indicators which describe the interactions or linkages between dimensions of sustainability.

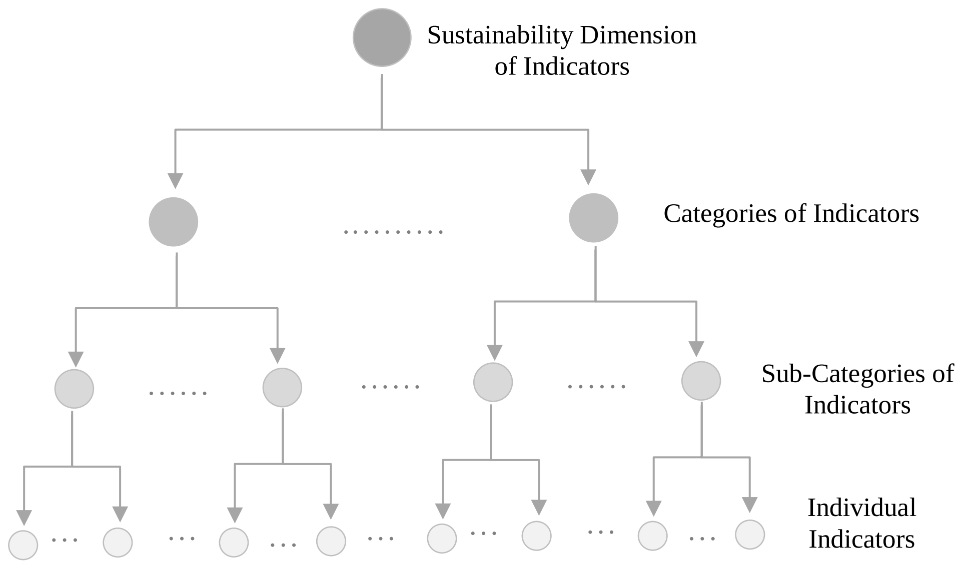


Figure 3-3 Organization of Indicators in Each Dimension of Sustainability

3.2.1 *DPSIR*

DPSIR works very well for identifying and developing indicators of sustainable development, as well as revealing the interactions between multiple indicators (Wei et al., 2007). In this framework, when using DPSIR to define or select indicators for sustainability assessment, users can put indicators in the corresponding cell in Table 3-1. Users first decide which component of DPSIR the indicators belong to: “driving force”, “pressure”, “state”, “impact” or “response” component. Then users determine the sustainability dimension of the indicator. Note that there can be multiple indicators under each dimension of sustainability, even there is only one cell for each dimension in each component of DPSIR.

Table 3-1 Indicator Selected or Defined by DPSIR in the Proposed Framework

Dimension / DPSIR	Economic	Environmental	Social	Resilient	Linkage
Driving Force					
Pressure					
States					
Impacts					
Responses					
Linkages					

3.2.2 *Life Cycle Thinking*

LCA provides a comprehensive view of the environmental aspects of sustainability objective. It can prevent us underestimating sustainability assessment in the environmental dimension. LCA of infrastructure can give us biophysical impact data of the infrastructure or infrastructure system, such as resource depletion, resource consumption (e.g., water, energy, and land), water and air pollution, human health impacts, waste generation, and so on. However, LCA analysis is very complex and time-consuming, and LCA typically does not address the economic or social aspects of a product. Therefore, we adopt the "life cycle thinking" into the selection or calculating of indicators in the proposed framework. In other words, indicators selected for sustainability assessment can not only capture the sustainability status of current objective but also the sustainability status of its whole life cycle. For example, the life cycle based sustainability indicators for U.S. food system (Heller et al., 2000) include indicators of seed production and animal breeding, indicators of agricultural growing and production, indicators of food processing, packaging and distribution, indicators of food preparation and consumption, and indicators of food waste managements.

3.3 Select Data

After selecting the indicators to be used for the sustainability assessment, the users need to determine which data they want to use to measure each indicator. The following criteria are considered when selecting data. In the proposed framework, users can select data from the database, local files, or the currently opened layers in QGIS.

- Relevance to Object: data selected for the subsequent sustainability assessment must be related to the object of sustainability assessment.
- Sustainability dimensions: data selected must represent all sustainability dimensions in which users want to evaluate CSI.
- Spatial scales: data can represent the information at the user-selected spatial scale with acceptable accuracy.
- Data quality: When selecting data, users need to consider the data sources and collection time of the data. The data used for the sustainability assessment should have good enough quality. Lower quality data may give an inaccurate or highly biased estimation of CSI. Relevance to the selected indicators: the

data should be able to measure the selected indicators and gives reliable estimation for each indicator.

3.4 Data Quality Assessment

The data used for sustainability assessment is usually from multiple data sources, such as the United States Census Bureau, Atlanta Regional Commissions (ARC), Yellow Pages, Google Earth, and so on. It is necessary to check data quality before using it to evaluate sustainability or to calculate values of each indicator, in order to make the following analysis more objective and less biased. Data quality depends on multiple factors, such as the purpose for which the data is used, users of the data, and the time to use it.

Data quality problems are present in single data collections, such as files and databases, due to misspellings during data entry, missing information or other invalid data (Rahm and Do, 2000). The data quality problems become more serious when integrating multiple data sources in the analysis because different sources of data often contain redundant data in different representations (Rahm and Do, 2000). For example, duplicated or missing information may produce incorrect or misleading statistics. Therefore, it is necessary to consolidate different data representations and eliminate duplicate information, in order to make data accurate and consistent (Rahm and Do, 2000). **Data Cleaning** is the process to detect and correct (or remove) corrupt or inaccurate records from a record set, table, or database (Wikipedia, 2017). It aims to detect incomplete, incorrect, inconsistent, inaccurate or irrelevant parts of the data and to modify or delete them, thus to improve data quality. Data cleaning can be performed interactively by the user manually or by computer program and algorithms.

In the proposed framework, data cleaning includes removing outliers, interpolating missing values, eliminating duplicate records, and removing inconsistencies in datasets. The data inconsistencies in this proposed framework include the difference of spatial CRS, spatial scales, field value units, and time. Data cleaning in the proposed framework can be finished manually (see Table 3-2) or automatically by machine learning algorithms.

Detect and Interpolate Missing Values: the framework can find "Null" values in one specified field in the data and show them to users. The user can choose to delete these null values records or to interpolate missing values by some interpolating methods. There are many spatial interpolating methods, such as Inverse Distance Weighting (IDW) and Triangulated Irregular Networks (TIN), natural neighbor interpolation, Kriging, and Artificial Neural Networks (ANN) (Sárközy, 1999).

Fix Inconsistencies among Datasets: We first check if all the data used for sustainability assessment has the same spatial CRS. If not, we will transfer the CRS which selected by users. We then check to see if all the data is at the same spatial scales. For example, we want to analyze at census tract scale, but some of our data is at the census block level (a finer spatial scale), some data is at the county level (a bigger spatial scale). In this case, we can aggregate the data at the census block scale into census tract scale, and disaggregate the data at the county scale into census tract scale. In the end, we manually check if the same attribute values from different sources have the same units based on metadata, for example, one data has income value in 'dollar per month', while another data has income value in 'dollar per year', in this case, we need to make the unit consistent. And we will try to convert all the data used for sustainability assessment in the International Unit (IU).

Table 3-2 Manually Data Quality Check Table

Data Name	Quality Level			Scores
	High (2)	Medium (1)	Low (0)	
…	…	…	…	…
Total				

3.5 Spatial Sustainability Assessment

"Spatial Sustainability Assessment" is the core component of the proposed framework. It includes all the data analysis methods need to be used during the process of spatial sustainability assessment, such as spatial analysis and multi-criteria decision analysis.

3.5.1 *Spatial Analysis*

Sustainability has spatial components or aspects by nature. Much useful spatial information can be obtained by spatial analysis functions of GIS, such as the spatial pattern of infrastructure, their location, geology conditions, the distribution of resources and precipitation. The spatial pattern of infrastructure can affect the sustainability of the whole infrastructure system. Appropriate distribution of infrastructure can improve its resilience to disasters or changes. The centralization of key infrastructure services can be a factor related to its resilience, and highly centralized infrastructure may be more vulnerable to disasters. Connectivity of infrastructure system can be a measure of its efficiency (high connectivity indicates a more efficient system), spatial and statistical distribution of people's accessibility to public services (e.g., green space, school, hospital, water, energy and so on) is a measure of social equity.

The proposed framework provides all the commonly used spatial analysis functions in GIS, including spatial pattern analysis, resource location and allocation analysis, density analysis, buffer analysis, and so on. Since the implementation of this framework is embedded in QGIS 2.18 seamlessly, all the spatial analysis tools provided in QGIS 2.18 can be directly used in the framework when evaluating sustainability, such as hotspots analysis, density analysis, location, and allocation analysis, terrain analysis, nearest neighbor analysis, spatial regression analysis, and so on. The spatial analysis functions that are designed/implemented in the proposed framework are discussed in details below.

3.5.1.1 Buffer Analysis

Buffer analysis provides methods to create buffer zones along spatial features, including point, line, and polygon. Buffer zones in the real world are often set up to protect the environment, to protect residential and commercial zones from industrial accidents or natural disasters, or to prevent violence (QGIS 2.8, 2018), such as the greenbelts between residential and commercial areas. The buffer zone in GIS is the area that within a specified distance (defined by users) to point, line or polygon features. It is represented as vector polygons enclosing these features, as shown in Figure 3-4.

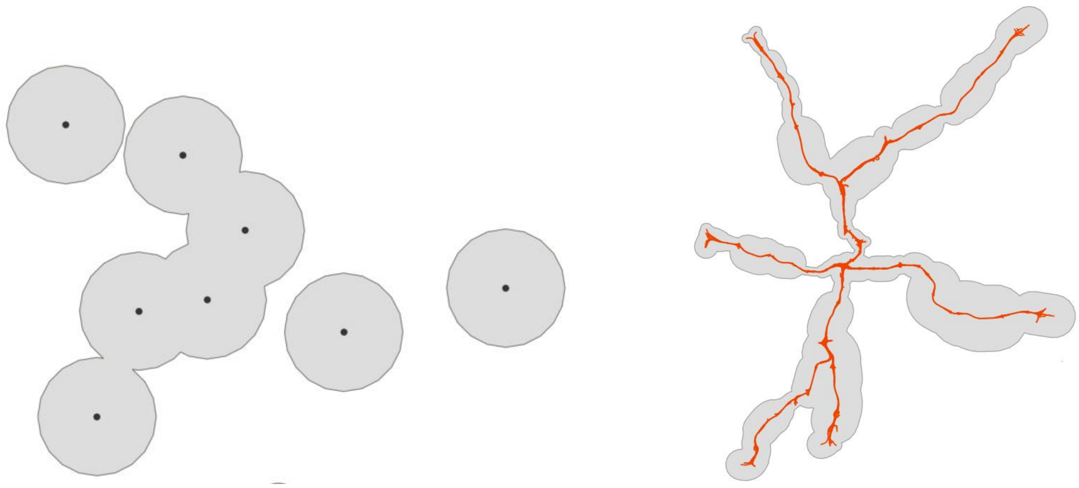


Figure 3-4 Buffer Zones Around Point (Left) and Line Features (Right)

The buffer distance can be a fixed number defined by the user, and the number has the same units as the CRS of data. It can also vary according to numerical values of an attribute in the vector layer for each feature, such as the polyline buffer zone shown on the right of Figure 3-4. By default, buffer zones are created on both side of the feature. However, sometimes buffers around polyline features, such as rivers or roads, do not have to be on both sides of lines. They can be on either the left side or the right side of the line feature. In these cases, the left or right side is determined by the direction from the starting point to the end point of the line. Besides, a feature can also have more than one buffer zone. For example, your working office may be buffered with a distance of 10, 15, 20, 25, 30, and 35 minutes of walking commuting time (Figure 3-5), thus forming multiple rings around the office as the boundary. If overlapping the buffer rings with house scores, you can easily find the best house location in different commuting distances. As shown in Figure 3-5, if you want some house within 10 minutes walking to your office, then the green dot in the dashed circle is the best option.

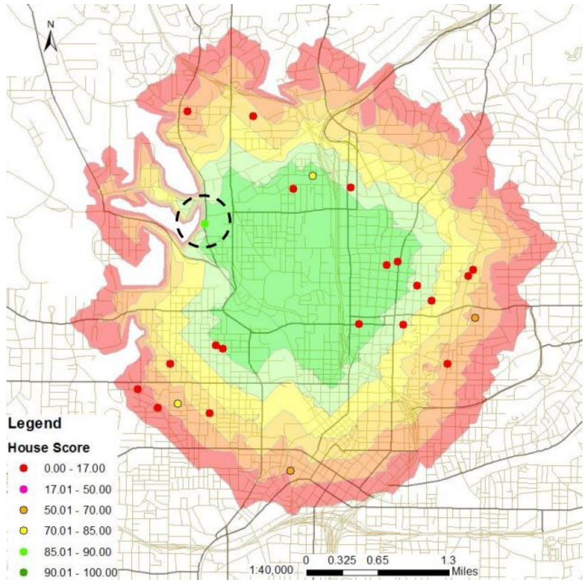


Figure 3-5 House Score in Different Commuting Circle for Walking

3.5.1.2 Spatial Auto-Correlation Analysis

According to Tobler's first law of geography, "everything is related to everything else, but near things are more related than distant things". From the perspective of geographic information science, spatial dependence (or spatial auto-correlation) is a defining characteristic of geographic data that makes many of the functions of geographic information systems possible (Goodchild, 2009). Spatial interpolation infers the value of fields from sample points, relies entirely on the positive spatial autocorrelation. Spatial autocorrelation exists in almost every spatial data. The assessment of spatial autocorrelation is generally considered to be one of the primary tasks of geographical data analysis (Hubert and Arabie, 1991).

Spatial autocorrelation analysis tests whether the observed value of a nominal, ordinal, or interval variable at one locality is independent of values of the variable at neighboring localities (Sokal and Oden, 1978; Sawada, 2018). In its most general sense, spatial autocorrelation describes the degree to which objects or activities at some place on the earth's surface are similar to other objects or activities located nearby and reflects Tobler's first law of geography (Goodchild, 1987). Usually, global spatial autocorrelation statistics give you an idea if spatial autocorrelation is present in the dataset while local spatial autocorrelation statistics allow you to see where the autocorrelation is happening or to identify spatial clusters and spatial outliers.

Spatial autocorrelation analysis requires some measure of contiguity or neighborhood relation to get the spatial weights between observations. There are three commonly used simple and intuitive neighborhood relations in the spatial autocorrelation analysis when considering continuous data in a raster format. They are rook’s case, bishop’s case, and queen’s (king’s) case, as shown in Figure 3-6.

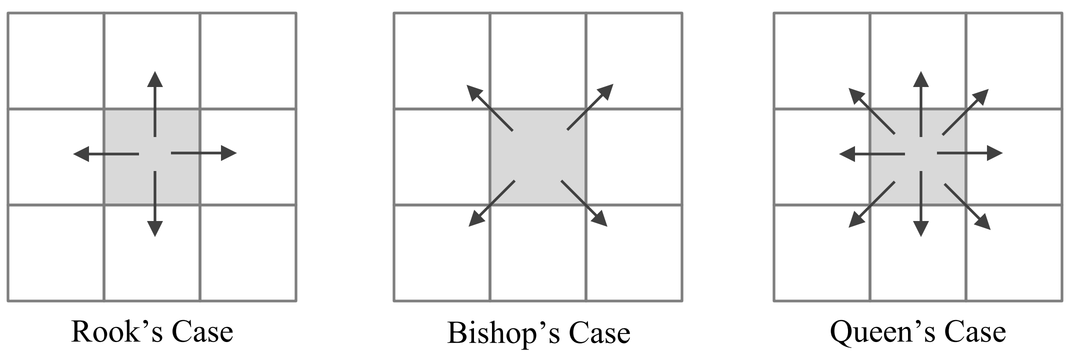


Figure 3-6 Different Forms of Contiguity in Spatial Autocorrelation Analysis

Among them, the rook’s case is the most commonly used, and most programs only compute this particular case (Sawada, 2018). However, these three types of contiguity are not sufficient for vector data or irregularly spaced points. In these cases, contiguity can be defined by distances to the *n* nearest neighbors or by the distance between a variate *x* and all its neighbors. For example, contiguity can be immediately neighboring administrative districts, or polygons that share the same line or connected nodes in a network.

The global spatial autocorrelation statistical method was used to measure the correlation among neighboring observations, to find patterns and levels of spatial clustering among neighbors (Boots and Getis, 1988; Tsai et al., 2009). Spatial autocorrelation statistics include Moran's I and Geary's C. **Moran's I** is more frequently used in the study of spatial correlation. It is a cross-product statistic between the deviation of a variable and its spatial lag and can be calculated by Equation 3.1. Its null hypothesis states that the attribute being analyzed is randomly distributed among the features in your study area.

$$I = \frac{\sum_i \sum_j w_{ij} z_i z_j / S_O}{\sum_i {z_i}^2 / N} = \frac{N}{S_O} \sum_i \sum_j w_{ij} \frac{(x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2} \tag{3.1}$$

Where N is the total number of features (or observations); x_i and x_j are observation value for feature i and feature j with mean $\mu = \sum_i x_i$; w_{ij} is the spatial weight between feature (or observation) i and j, it is element in the spatial weight matrix (or contiguity matrix) corresponding to the observation pair i, j; and S_O is the aggregate of all spatial weights, $S_O = \sum_i \sum_j w_{ij}$. The spatial weight matrix contains information about the neighborhood structure for each feature or observation. If observation i and observation j are neighbors or adjacent, then their spatial weigh $w_{ij} = 1$, otherwise, $w_{ij} = 0$. And a feature or observation is not adjacent to itself, $w_{ii} = 0$. Thus, S_O is two times of the number of neighbor pairs.

Moran's I varies from -1 to +1. In the absence of autocorrelation and regardless of the specified weight matrix, the expectation of Moran's I statistic is $-\frac{1}{N-1}$, which tends to zero as the sample size increases. A Moran's I coefficient larger than the expectation indicates positive spatial autocorrelation, and a Moran's I coefficient less than the expectation indicates negative spatial autocorrelation. A zero Moran's I coefficient implies a random pattern.

Due to spatial heterogeneity, the estimated degree of autocorrelation may vary significantly across geo-space (Tsai et al., 2009). Local spatial autocorrelation statistics, such as **Local Moran's I**, estimate the spatial autocorrelations at each location or for each observation, instead of estimating it on the whole dataset. It can detect spatial clusters and spatial outliers at different locations. Local Moran's I can be calculated by Equation 3.2.

$$I_i = \frac{\sum_j w_{ij} z_i z_j}{{S_i}^2} = \sum_j \frac{w_{ij}(x_i - \mu)(x_j - \mu)}{{S_i}^2} = \frac{(x_i - \mu)}{{S_i}^2} \sum_j w_{ij}(x_j - \mu)$$

$${S_i}^2 = \frac{\sum_{j \neq i} {z_j}^2}{N-1} = \frac{\sum_{j \neq i} (x_j - \mu)^2}{N-1} \tag{3.2}$$

The null hypotheses is also that observations are randomly distributed. A positive value for Local Moran's I indicates local spatial clustering of similar values (sometimes referred to hot spots), either high or low (Anselin, 1995). A negative value for Local Moran's I indicates local spatial clustering of dissimilar values (can be outliers), for example, a location with high values surrounded by neighbors with low values. In either instance, the p-value for the feature must be small enough for the cluster or outlier to be considered statistically significant.

3.5.1.3 Comparison Analysis

Comparison analysis in the framework calculates the difference of specified attribute field between two layers. These two layers must have the same data type (e.g., both are polygon layers), and have the specified attribute field in their attribute table. For example, if you want to compare the population change of each state in the US from 2010 to 2015. Then the comparison analysis can generate a difference layer with the specified population field, and the field value is the population difference between these two layers.

3.5.1.4 Random Points or Vectors Generator

Random points or vectors are commonly used in the spatial analysis; thus this framework provides tools to generate a specified number of random points or random vectors within some boundary. By default, they are randomly evenly distributed in the boundary. For random vectors, users can also specify the length of vectors, which is in map units according to CRS of the boundary layer. The generated random point layer or random vector layer has the same CRS with the boundary layer.

3.5.1.5 Aggregate Attributes

Aggregating attributes is very important and necessary especially when you want to convert all your data into the same spatial scale (e.g., to aggregate your data in a smaller scale to a larger scale) or to aggregate attribute values of points to polygons that contain those points. This proposed framework provides tools to aggregate attributes of smaller polygons to bigger polygons, or aggregate attributes of points or polylines to polygons. There are seven different methods to compute the aggregated value, including Sum, Min, Max, Median, Arithmetic mean, Geometric mean, and Interquartile Mean (IQM). IQM aggregation method only aggregates the data with value is in the second and third quartiles. Assuming the values have been ordered, then their IQM can be calculated by Equation 3.3.

$$x_{IQM} = \frac{2}{n}\left(\sum_{i=\left\lceil\frac{n}{4}\right\rceil+1}^{\left\lfloor\frac{3n}{4}\right\rfloor} x_i + \frac{x_{\left\lceil\frac{n}{4}\right\rceil} + x_{\left(\left\lfloor\frac{3n}{4}\right\rfloor+1\right)}}{2} \times \left(\frac{n}{2} - \left(\left\lfloor\frac{3n}{4}\right\rfloor - \left\lceil\frac{n}{4}\right\rceil\right)\right)\right) \tag{3.3}$$

$$x_{IQM} = \frac{2}{11}\left(\sum_{i=4}^{8} x_i + \frac{x_3 + x_9}{2} \times \left(\frac{11}{2} - (8-3)\right)\right) \quad n = 11$$

3.5.2 *Network Analysis*

The framework provides some measures of the network, such as connectivity measures, centrality, and accessibility. Their measurements used in the framework are explained in details as follows. To make symbols consistent, let the network be $G = (V, E)$, where V is the set of nodes in the network or graph G, E is the set of edges or links in the network or graph G. $\mathrm{n} = |V|$ is the number of nodes, $\mathrm{m} = |E|$ is the number of edges in graph G. $|E|_{max}$ is the maximum possible edges in the graph.

3.5.2.1 Connectivity of Network

"Connectivity of networks is the degree of connection between all vertices or the degree of completeness of the links between nodes" (Taaffe and Gauthier, 1973; Robinson and Bamford, 1978). A well-connected network has many short links, numerous inter-sections, and minimal dead-ends, which provide continuous and direct routes to destinations (Sreelekha et al., 2016). The greater the degree of connectivity within a network, the more efficient is the network system. There are many measures of network connectivity, such as alpha index, beta index, gamma index, detour index, and so on.

Alpha Index evaluates the number of cycles in a graph or network in comparison with the maximum number of cycles (See Equation 3.4). Values of the alpha index range from 0 to 1, with higher values representing a more connected network. A value of 1 indicates a completely connected network and very serious redundancies; thus it is very rare in reality (Rodrigue et al., 2017). Alpha index is good if your object is to provide multiple alternative routes.

$$\alpha = \frac{u}{2|V|-5} = \frac{|E|-|V|+p}{2|V|-5} = \frac{\mathrm{m}-\mathrm{n}+p}{2\mathrm{n}-5} \qquad \alpha \in [0,1] \qquad (3.4)$$

Where, u is the maximum number of independent cycles in a graph or network ($u = |E| - |V| + p$), p is the number of non-connected subgraphs.

Beta Index measures relationships between the number of links and the number of nodes (see Equation 3.5). The greater the value of beta index, the greater the connectivity. β value for tree types of structures and disconnected networks is always less than 1. It would be zero when there are no edges in the network. β value for any network structure with one circuit would always be equal to 1. β value exceeds 1 for a complicated network structure having more than one circuit. For example, a perfect grid has β value of 2.5 (Dill, 2004). However, β value does not reflect or consider the length of links in the connectivity measurement. For example, a perfect grid of 1,000-foot blocks can have the same link-node ratio as a grid with 200-foot blocks (Dill, 2004). Therefore increasing β value may not improve the shortest distance problems.

$$\beta = \frac{|E|}{|V|} = \frac{\mathrm{m}}{\mathrm{n}} \qquad \beta \geq 1 \tag{3.5}$$

Gamma Index considers relationships between the number of observed links and the number of possible links (see Equation 3.6 for Gamma index of networks without self-loop) (Rodrigue et al., 2017). The value of gamma ranges from 0 to 1. A gamma index of 0.5 means that the network is 50% connected. A value of 1 indicates a completely connected network but is very rare in reality. Gamma index is an efficient value to measure the progression of a network in time (Rodrigue et al., 2017).

$$\gamma = \frac{|E|}{|E|_{max}} = \begin{cases} \dfrac{\mathrm{m}}{3(\mathrm{n}-2)} & \text{planar network} \\ \dfrac{\mathrm{m}}{\mathrm{n}(\mathrm{n}-2)/2} & \text{nonplanar, undirected network} \\ \dfrac{\mathrm{m}}{\mathrm{n}(\mathrm{n}-2)} & \text{nonplanar, directed network} \end{cases} \quad \gamma \in [0,1] \tag{3.6}$$

Detour Index (DI) is also called Pedestrian Route Directness (PRD). DI measures the efficiency of a transport network in terms of how well it overcomes distance or the friction of distance (Rodrigue et al., 2017). It is the ratio of route distance to straight-line distance for two selected points (See Equation 3.7). The lower the detour index, the more direct is a given route. The closer the detour index to 1, the more spatially efficient the network is (Rodrigue et al., 2017). However, networks having a detour index of 1 are rarely seen in practice.

$$\mathrm{DI}_{ij} = \frac{d_R(i,j)}{d_E(i,j)} \qquad \mathrm{DI}_{ij} \geq 1 \tag{3.7}$$

Where, DI_{ij} is the detour index from point i to point j. $d_R(i,j)$ is the total 'route' distance from point i to point j. $d_E(i,j)$ is the natural Euclidean distance between point i and point j. Note that DI_{ij} may not equal to DI_{ji}. Point pair (i,j) is also referred as a vector in the thesis.

DI is often used for assessing the effects of adding or removing links in a given network (Raghav, 2018). DI or PRD is also a better measure for promoting bicycling and walking than beta index. It directly reflects the distance traveled for a trip, which is a primary factor in determining whether a person walks or bikes. However, using PRD as a measure of network connectivity, we need to generate enough randomly distributed pairs of points (or vectors) in the study area to make the DI value more reasonable and less subjective.

3.5.2.2 Accessibility of Network

Accessibility illustrates where the opportunities (activities, services, goods, facilities, and destinations) are located concerning people and the convenience to access these opportunities (Black and Conroy, 1977). It is the potential for interactions and exchanges between places or locations (Hansen, 1959). Litman (2008) reviews the meaning of accessibility and its implications in various disciplines, such as transportation planning, social planning, geography, and urban economics. Accessibility commonly refers to people's access to goods, services, activities and facilities in a given destination, and the relative ease of reaching a particular location or area (Litman, 2003). It can reflect people's ability to access or use services and opportunities.

There are many ways to measure accessibility, such as attraction-accessibility measure, location benefits accessibility measure, cumulative opportunities measures (Kong et al., 2007), utility-based accessibility measure (Gulhan et al., 2013), and so on. Among them, the attraction-accessibility considers the spatial interactions between origins and destinations. It computes a summary score for an individual or place based on the attractiveness of potential activity locations and their required travel costs (e.g., travel distance, time, or money). The most common attraction-accessibility measure is the Hansen (1959) potential measure (see Equation 3.8). It assumes that accessibility is likely to decrease as the cost of reaching it increases.

$$A_i = \sum_{j \in K_i} w_j \times f(c_{ij}) \tag{3.8}$$

Where A_i is the accessibility of travel origin i to destination j; w_j is the attractiveness of destination j, it is often estimated as the number of opportunities. $f(c_{ij})$ is a general function describing the travel cost or travel impedance between origin i and destination j. $f\left(c_{ij}\right)$ can be distance, time, or money. It is usually estimated by straight-line Euclidean distance, network distance, actual traveling time or travel distance, and so on. According to the First Law of Geography, we assume that the attractiveness of destination j to origin i will decrease with increases of the distance to destination. For simplify, we use the simplest inverse power gravity formulation (Bhat et al., 2001) as our travel cost function,$f\left(c_{ij}\right) = {d_{ij}}^{-\alpha}$, $\alpha = 2.0$, where d_{ij} is the real distance from location i to location j. K_i is the choice set available to origin point i. We also assume people's choice set will not change after reach some distance threshold, according to the First Law of Geography. In other words, people is unlikely to use services or facilities very far away, thus those far facilities are not in their choice set. For example, Alshuwaikhat and Aina (2006) assumes that services must be within 400 meters of citizens for the services to be accessible by walking. Thus services which are located more than 400 meters are not within citizens' choice set for walking.

3.5.2.3 Centrality of Network

The centrality of the network measures the importance of nodes or links (edges) in one network. It plays an important role in the analysis of many types of networks, such as social network, transportation network. There are numerous measures of the centrality of a network, including betweenness centrality, closeness centrality, eigenvector centrality, PageRank centrality, degree centrality, and so on (Borgatti, 2005). Each measure has its assumptions about how things flow in a network; thus each centrality measure works for different types of networks. For example, betweenness centrality assumes flows move only along the shortest possible paths, and the flow is indivisible (like a package). It works best for a network that represents the package delivery process. The eigenvector centrality measure is ideally suited for the network that describes influence type processes (Borgatti, 2005). Details of commonly used network centrality measures are as follows.

Betweenness Centrality can be defined as the number of a shortest path passing through one node or one edge. If there is more than one shortest path between a pair of vertices, each path is given equal weight (e.g. $\frac{1}{\sigma_{st}}$ in Equation 3.9) such that the total weight of all of the paths is unity (Girvan & Newman, 2002). **Betweenness Centrality** measures the potential of a vertex or an edge to control information flow in the network (Unnithan et al., 2014). **Betweenness centrality of vertex** or node v is defined as the number of shortest paths between pairs of other vertices that run though node v. It is a measure of the influence of a node over the flow of information between other nodes, especially in cases where information flow over a network primarily follows the shortest available path (Girvan and Newman, 2002). High betweenness centrality of vertex may have considerable influence within a network, because these vertex control information passing between others (Franceschet, 2014). Shortest-path Betweenness Centrality of node v can be calculated by Equation 3.9.

Edge betweenness centrality is defined as the number of shortest paths that go through an edge in a graph or network (Girvan and Newman, 2002). An edge with a high edge betweenness centrality value plays an important role in the communication or information flow in the network, because removal of this edge may affect the communication between many pairs of nodes through the shortest paths between them. Shortest-path Betweenness Centrality of an edge e is calculated by Equation 3.10 (Avrachenkov et al., 2013).

$$C_B(v) = \frac{1}{N_B} \sum_{\substack{s,t \in V \\ s \neq t \neq v}} \frac{\sigma_{st}(v)}{\sigma_{st}} = \frac{1}{n(n-1)} \sum_{\substack{s,t \in V \\ s \neq t \neq v}} \frac{\sigma_{st}(v)}{\sigma_{st}} \tag{3.9}$$

$$C_B(e) = \frac{1}{N_B} \sum_{\substack{s,t \in V \\ s \neq t}} \frac{\sigma_{st}(e)}{\sigma_{st}} = \frac{1}{n(n-1)} \sum_{\substack{s,t \in V \\ s \neq t}} \frac{\sigma_{st}(e)}{\sigma_{st}} \tag{3.10}$$

Where, $C_B(v)$ is the betweenness centrality of node v. $C_B(e)$ is the betweenness centrality of edge e. σ_{st} is the number of shortest paths from node s to node t. $\sigma_{st}(v)$ is the number of shortest paths from node s to node t with node v as an inner vertex along the path. $\sigma_{st}(e)$ is the number of shortest paths from node s to node t and passing edge e. V is set of nodes in the network. N_B is a normalization constant, $N_B = (n-1)(n-2)$ if v cannot be a start or end vertex of the shortest path. $N_B = \mathrm{n}(n-1)$ if v may also be a start or end vertex of the shortest path. $\mathrm{n} = |V|$ is the number of nodes in the network.

Shortest-path Closeness Centrality of a node i is defined by Equation 3.11. It measures the degree to which a vertex is close to other vertices (Brandes and Fleischer, 2005). It can only be used for connected graph. If the graph is non-connected, we need to use Harmonic Centrality (See Equation 3.12).

$$C_C(i) = \frac{N_C}{\sum_{i,t\,\in V}^{t} d_{it}} = \frac{n-1}{\sum_{i,t\,\in V}^{t} d_{it}} \tag{3.11}$$

$$C_C(i) = \frac{1}{N_C} \sum_{\substack{t \\ i,t\,\in V \\ i\neq t}} \frac{1}{d_{it}} = \frac{1}{n-1} \sum_{\substack{t \\ i,t\,\in V \\ i\neq t}} \frac{1}{d_{it}} \tag{3.12}$$

Where, $C_C(i)$ is the closeness centrality of a node i, d_{it} is the length of a shortest path between node i and node t, V is set of nodes in the network. N_C is a normalization constant, $N_C = n - 1$, $\mathrm{n} = |V|$ is the number of nodes in the network.

Eigenvector Centrality is defined as the principle eigenvector of the adjacency matrix defining the network (Borgatti, 2005). It assumes that traffic or flow in the network can move via unrestricted walks rather than being constrained by trails, paths, or geodesics. Eigenvector centrality naturally describes a mechanism in which each node affects all of its neighbors simultaneously, as in a parallel duplication process. Hence, it is ideal for measuring centrality for influence type process.

3.5.3 *Multi-Criteria Decision Analysis*

MCA considers multiple decision criteria systematically. It can incorporate both qualitative and quantitative data into the process of assessment. The most important and challenging part of MCA is assigning weights for each indicator and aggregating all indicators together. The process to evaluate sustainability by MCA in this framework is illustrated in Figure 3-7. It first conducts some pre-analysis on indicators, including correlation analysis and normalization of indicators. Then it weights each indicator manually or by AHP, and determines the sign of indicator's impact on sustainability objective. For example, an increased value of air pollution has a negative impact on sustainability objective, and then the corresponding indicator has a negative sign; while the increased value of profit has a positive impact to the economic performance of the study area, thus the corresponding indicator has a positive sign. After having the weight and signs of each indicator, the framework aggregates the indicators to get the dimensional Composite Sustainability Index (CSI) for each dimension of sustainability and the overall composite sustainability score (i.e., overall CSI) which counts the impacts of all dimensions. There are many aggregation methods used in MCA, details see (Díaz-Balteiro and Romero, 2004). This framework provides "linear-weighted-additive" and "geometric mean" aggregation model.

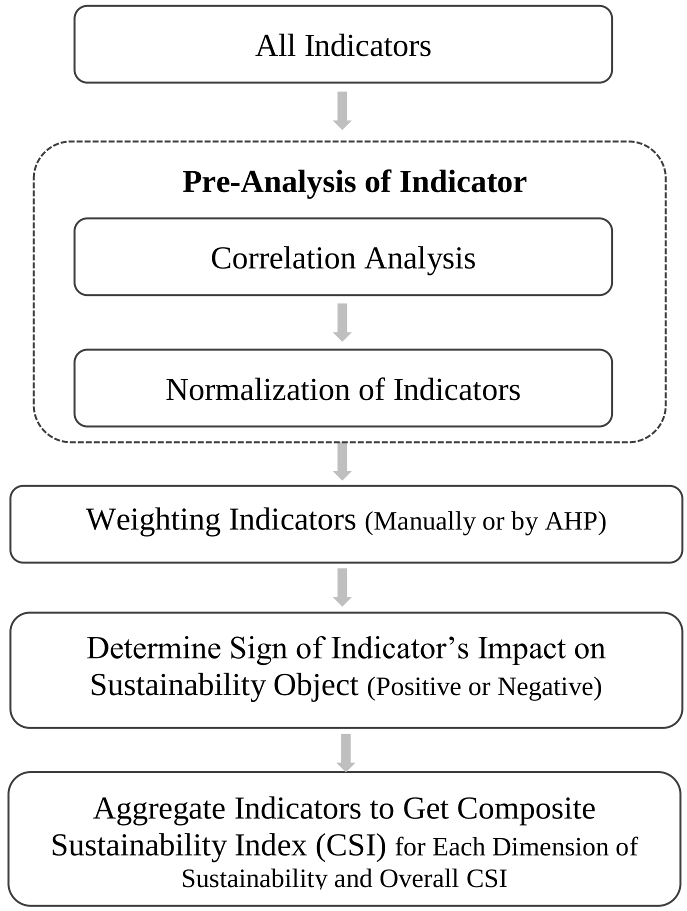


Figure 3-7 Process of Evaluating Sustainability by MCA

As discussed and mentioned in Chapter 2, indicators of sustainability are usually not independent, and there are interactions or correlations between them. There are many options to handle the correlations between indicators. Two options are provided in the proposed framework (illustrated in Figure 3-8 and Figure 3-9). The first option is to remove the highly correlated indicators. For example, if there are two indicators are highly correlated very each other (e.g., correlation coefficient is greater than 0.8), the user can keep one of them for the following sustainability analysis. Users can also use PCA to select the independent indicators for the following MCA, the process of sustainability assessment is shown in Figure 3-8. The second option is to use PCA to remove the redundancy information caused by the correlation between indicators, and derive the weight of indicators. The sustainability assessment which use PCA to process the correlated indicators is shown in Figure 3-9.

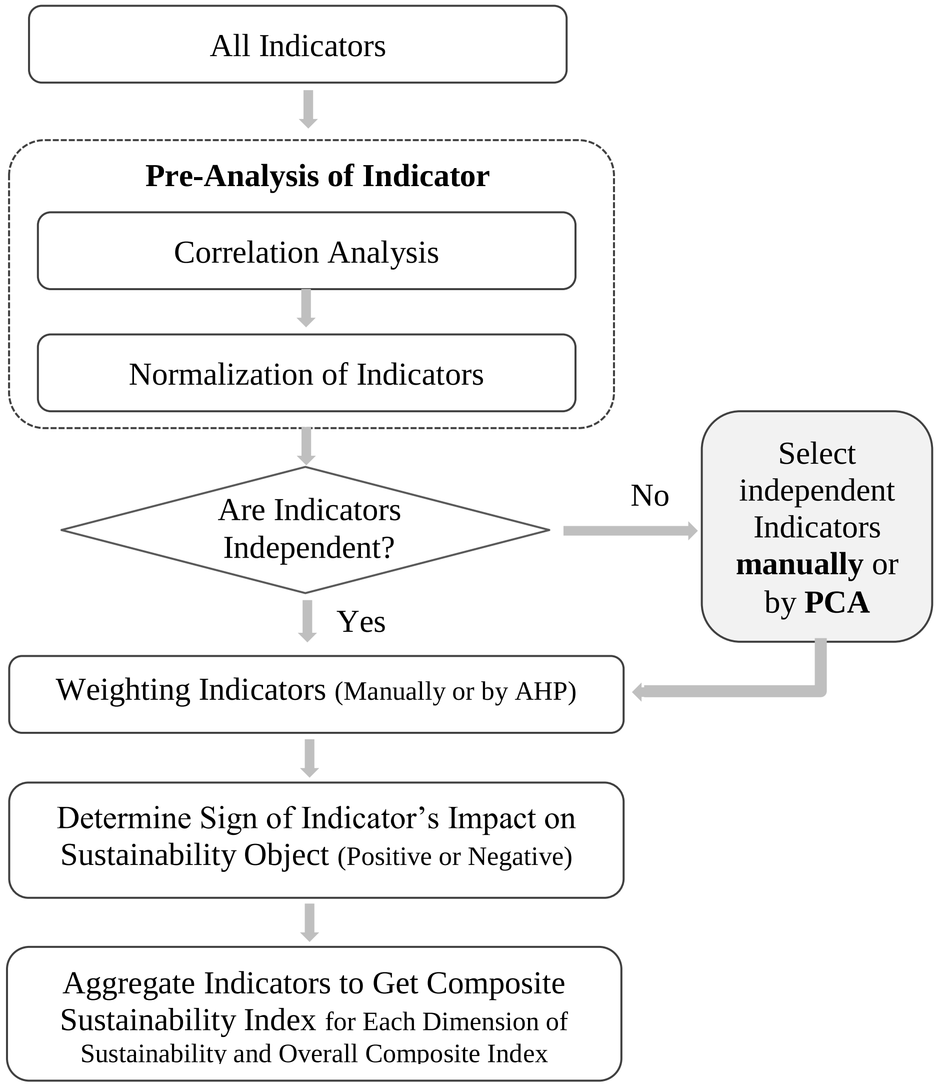


Figure 3-8 Process of Evaluating Sustainability by MCA (Option 1: Remove Highly Correlated Indicators Manually or by PCA)

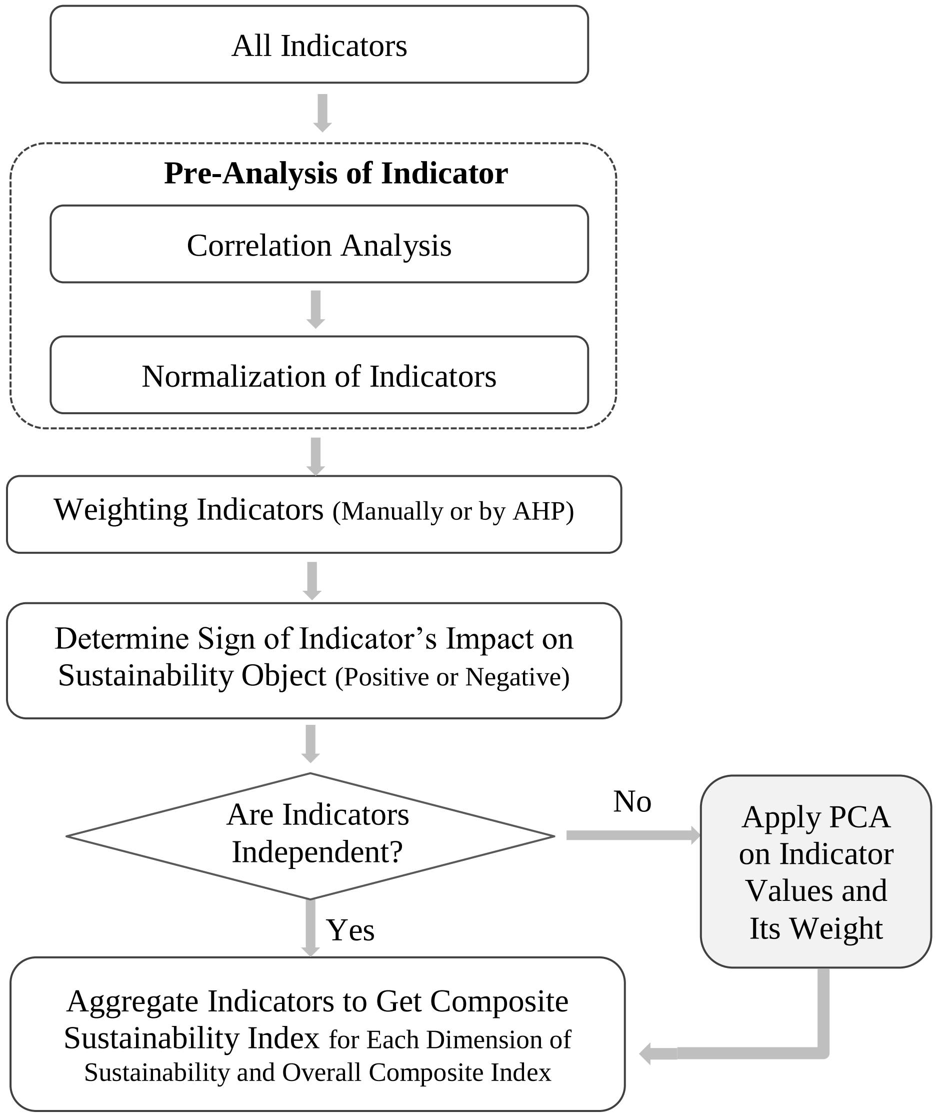


Figure 3-9 Process of Evaluating Sustainability by MCA (Option 2: Use PCA to Derive Weight of Indicators)

3.5.3.1 Indicators Pre-Analysis

Indicators selected or defined by users for the sustainability assessment may have correlations or have different units and ranges. The interrelations between four aspects of sustainability must also be considered when conducting spatial sustainability assessment. Because "combined impacts of sets of measures as a whole, are likely to be more than the simple sum of impacts of their constituent measures due to synergistic effects'' (Pope et al., 2004). And MCA requires the evaluation criteria be independent of each other, since the correlation between criteria may result in malfunction of weighting for MCA by "double counting similar variables" (Li and Yeh, 2002). Therefore, we need to analyze the correlations between indicators and make sure all the indicators used for MCA are independent.

Correlation Analysis of Indicators

In this framework, the correlation coefficient between any two indicators will be calculated and a correlation matrix will be produced and plotted. There are many measures of correlation coefficient depending on the type of data. Table 3-3 (Statistics How To, 2018) shows some of the common choices for correlation coefficients.

Table 3-3 Correlation Coefficients Measures for Different Data Types

	Quantitative	Ordinal	Nominal
Quantitative	Pearson	Biserial	Point Biserial
Ordinal	Biserial	Spearman's rho or Kendall's tau	Rank Biserial
Nominal	Point Biserial	Rank Biserial	Phi Coefficient

Among all the coefficient measure, Pearson's correlation coefficient is the best known and most commonly used measure of dependence between two quantities. It measures the strength of the linear relationship between two ratio variables and can be calculated by Equation 3.13.

$$\rho_{X,Y} = \frac{\mathrm{cov}(\mathrm{X},\mathrm{Y})}{\sigma_X \sigma_Y} = \frac{\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} = \frac{\sum_i (x_i-\mu_X)\,(y_i-\mu_Y)}{\sqrt{\sum_i (x_i-\mu_X)}\sqrt{\sum_i (y_i-\mu_Y)}} \tag{3.13}$$

Where, x_i, y_i are individual sample points in dataset X, Y. μ_X, μ_Y are their sample mean and σ_X, σ_Y are their covariance.

Spearman's rank correlation coefficient or Spearman's rho assesses how well the relationship between two variables can be described using a monotonic function (Wikipedia, 2018). Spearman's coefficient is appropriate for both continuous and discrete ordinal variables. It is defined as the Pearson correlation coefficient between the ranked variables (Wikipedia, 2018). For a sample of size n, the n raw scores X_i, Y_i are converted to ranks r_{X_i}, r_{Y_i}, then Spearman's coefficient R_s can be calculated by Equation 3.14.

$$R_s = \rho_{r_X,r_Y} = \frac{\mathrm{cov}(r_X, r_Y)}{\sigma_{r_X}\sigma_{r_Y}} \tag{3.14}$$

Where, ρ_{r_X,r_Y} is the Pearson correlation coefficient of rank variable r_X, r_Y; $\mathrm{cov}(r_X, r_Y)$ is the covariance of these two rank variables, and $\sigma_{r_X}, \sigma_{r_Y}$ are the standard deviations of the rank variables.

If all the n ranks are distinct integers, then the Spearman's coefficient can be computed using Equation 3.15 (Wikipedia, 2018). Where, $(r_{X_i} - r_{Y_i})$ is the difference between two ranks of observation X_i, Y_i, n is the number of observations.

$$R_s = 1 - \frac{6\sum_i (r_{X_i} - r_{Y_i})^2}{n(n^2-1)} \tag{3.15}$$

In the proposed framework, users can compute the Pearson, Spearman, or Kendall correlation coefficients and their significance levels for all pairs of variables. There are different ways to visualize the correlation matrix, see examples in (STHDA, 2017a; STHDA, 2017b; CRAN, 2017). This framework produces a correlation matrix plot or figure as shown in Figure 3-10.

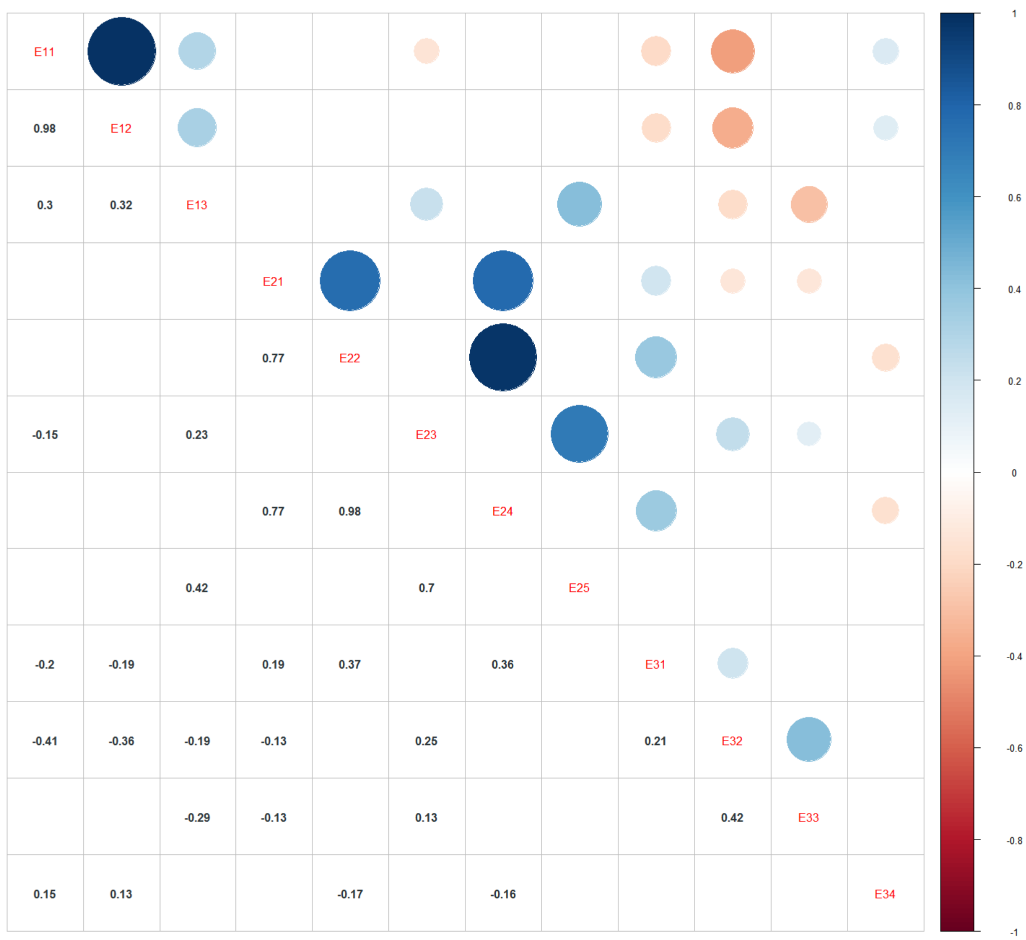


Figure 3-10 Correlation Matrix Plot

In the upper triangular, positive correlations are displayed in blue and negative correlations in red color. Color intensity and the size of the circle are proportional to the correlation coefficients. In the lower triangular, the statistically significant coefficient values are shown in the corresponding cell. The legend color shows the correlation coefficients and the corresponding colors. All the non-significant coefficients leave blank in the figure. The correlation matrix is ordered the same as the input variables by default. Users can also choose different methods to reorder the correlation matrix, such as alphabetical order, the angular order of the eigenvectors (AOE), the first principal component order (FPC), and the hierarchical clustering order (hclust).

Normalization of Indicators

Indicators used in the assessment may have different measurement units and value ranges; they need to be normalized before conducting MCA (Graymore et al., 2009). Normalization is usually applied to single variables to make them comparable, i.e., transforming the various scales of variables into one unique scale (Bohringer and Jochem, 2007). There are many normalization methods (see (OECD, 2008) for details), such as ranking, standardization (Z-Score), min-max, categorical scales, “distance to reference” and so on. The proposed framework provides the first four normalization methods, which are shaded by gray color in Table 3-4.

Ranking is the simplest normalization technique (see Equation 3.16). It is not affected by outliers and allows users to compare the sustainability performance of one area over time in terms of relative positions (ranking). However, users cannot evaluate the absolute value of sustainability performance, since the information of absolute value is lost after ranking normalization.

Standardization (or **Z-Scores**) normalization method converts indicators to a common scale with zero mean and standard deviation of one (see Equation 3.17). Z-Scores normalization method can remove the effect of outliers in indicator values. Therefore, Z-Score is not the best normalization method if an extremely good result on a few indicators is thought to be better than a lot of average scores.

Table 3-4 Normalization Methods (Modified Based on (OECD, 2008))

Method	Equation	
Ranking	$x_i' = Rank(x_i)$	(3.16)
Standardization (Z-Score)	$x_i' = \dfrac{x_i - \mu_i}{\sigma_i}$; $\overline{x_i'} = 0, S(x_i') = 1$	(3.17)
Min-Max	$x_i' = \dfrac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \in [0,1]$	(3.18)
Distance-to-Reference	$x_i' = \dfrac{x_i - x_i^{Ref}}{x_i^{Ref}}$ or $x_i' = \dfrac{x_i}{x_i^{Ref}}$	(3.19)
Categorical-Scales	$x_i' = \begin{cases} v_1 & \text{if } x_i \text{ in range } [r_0, r_1] \\ v_2 & \text{if } x_i \text{ in range } [r_1, r_2] \\ v_3 & \text{if } x_i \text{ in range } [r_2, r_3] \\ & \dots \\ v_k & \text{if } x_i \text{ in range } [r_{k-1}, r_k] \end{cases}$	(3.20)

Note: x_i is the value of indicator *i*, μ_i, σ_i are its mean and standard deviation value; $\min(x_i)$, $\max(x_i)$ are its min and max value; x_i' is its normalized value. x_i^{Ref} is the reference value of indicator *i* , it can be meaningful value defined by users, for example, the target value of indicator *i* under the corresponding sustainability objective. $v_1, v_2, v_3, \dots$, and v_k are categorical scores of indicator after Categorical-Scales normalization.

Min-Max method normalizes all indicators to have an identical range [0, 1] by Equation 3.18. However, the extreme values or outliers in indicator values could distort its normalized value. Besides, indicators whose value lies within a small interval can be stretched to a wider range, thus may increase its effect on the composite sustainability indicators.

Distance-to-Reference method is similar to the rank method, but it measures the indicator's relative value or position to the reference value (see Equation 3.19). The reference value can be a sustainability target, which is to be reached in a given time frame. For example, the Tokyo Protocol has established an 8% reduction target for CO_2 emissions by 2010 for European Union members. The reference could also be an external benchmark value. For example, the United States and Japan are often used as benchmarks for the composite indicators built in the framework of the EU Lisbon agenda (OECD, 2008). Alternatively, the reference can also be an average value of the indicator.

Categorical-Scale normalization method assigns a score (v_i) for indicator at some range (See Equation 3.20). Categorical score can be numerical (e.g. 1,, 2, or 3) or qualitative (e.g. ",fully achieved", "partly achieved" or "not achieved"). In most cases, the scores are assigned based on percentiles of the indicator's distribution. For example, the top 5% receive a score of 100, the indicator values between the 85th and 95th percentiles receive a score of 80, the indicator values between the 65th and the 85th percentiles receive a score of 60, and so on. If we use the same percentile transformation for different years, then we can compare the change of indicator's distribution over time. However, it is difficult to capture the change of its absolute value over time, since this normalization method excludes large amounts of information about indicator, such as variance of the transformed indicators and absolute values.

Select Independent Indicators by PCA

PCA can help us find the best independent sustainability indicators for the following MCA. First, we calculate the correlation coefficient among all indicators and get the sum of absolute correlation coefficients for each indicator (i.e., sum of the absolute correlation coefficients for each column in the correlation matrix). Second, we normalize all the indicators using the Z-Score method and then conduct PCA for all the normalized indicators. Third, we get the first *k* PCs, which counts for more than 90% of the indicators' variances (users can determine how many information are kept in their PCs, which determined the value of *k*). Fourth, we first select indicators with a factor loading (the corresponding element of eigenvector) that are greater than 90% of the highest absolute factor loading from each of the k PCs (PCA may select different indicators by changing the percentage of the highest absolute factor loading). Finally, we examine the correlation coefficient among the selected indicators in the same PC. If the indicators are statistically significantly correlated (i.e., correlation coefficient **≥ 0.8** and its **p-value < 0.05**), then only the indicator with the highest sum of absolute correlation coefficients is retained. Otherwise, all of the indicators are retained. Chu et al. (2018) illustrate an example to select indicators by PCA.

3.5.3.2 Weighting Indicators

The weights of indicators highly depend on the object, scales, and dimensions of the sustainability assessment. The basic principle is that important indicators should have higher weights. Since the main idea of sustainability is to achieve balanced development in each dimension of sustainability, equal weight for each dimension will be considered by default in this framework. However, users can manually set their weight for indicators in each dimension, according to their specific object and priorities. Users can also use AHP to generate weights of indicators automatically, based on the priorities provided by users. The process to determine the weights of indicators by AHP is illustrated in Figure 3-11.

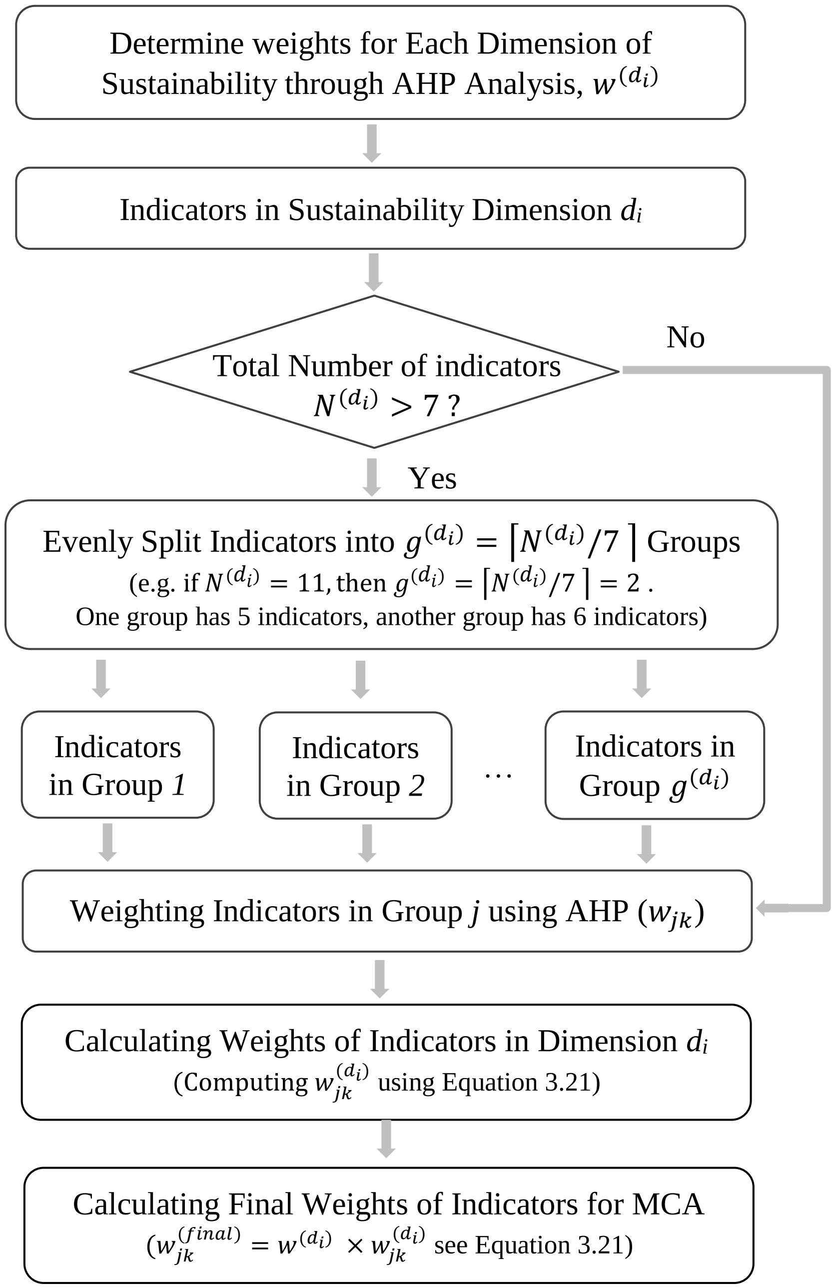


Figure 3-11 Process of Weighting Indicators Using AHP

When using AHP to determine the weight of indicators, we first determine the weight for each dimension of sustainability through AHP analysis, which is $w^{(d_i)}$. Then we conduct AHP for indicators at different dimensions of sustainability separately, and we get the weight of indicators within each dimension of sustainability, which are referred as dimension weight. However, AHP in this framework requires no more than 7 for each analysis, in order to make sure that a small error does not affect or distort the relative priorities in AHP (Saaty, 1987). If there are more than 7 sustainability indicators in one dimension, indicators must be split into different groups, thus each group has no more than 7 indicators. We then apply AHP for each group of indicators, and get the weight of indicator within the group, which are referred as group weight. After getting weighs of indicators for all groups, we then calculate the weights among groups or within dimension ($\mathrm{w}_{jk}^{(d_i)}$) by Equation 3.21, this weight is used to calculate the CSI for each dimension of sustainability by MCA, $I_{CSI}^{(di)}$ (see Equation 3.23). Finally, we calculate the final indicator weight, $\mathrm{w}_{jk}^{(final)}$, which can be used to calculate the CSI by MCA, $I_{CSI}^{(overall)}$ (see Equation 3.23).

$$\mathrm{w}_{jk}^{(d_i)} = \frac{\mathrm{N}_j^{(d_i)}}{N^{(d_i)}} \times w_{jk}$$

$$\mathrm{w}_{jk}^{(final)} = w^{(d_i)} \times \mathrm{w}_{jk}^{(d_i)} \tag{3.21}$$

Where, w_{jk} is the group weight of k-th indicator in group j, $\sum_k w_{jk} = 1$. $\mathrm{w}_{jk}^{(d_i)}$ is the dimension weight for k-th indicator in group j, $\sum_j \sum_k {w_{jk}}^{(d_i)} = 1$. $\mathrm{N}_j^{(d_i)}$ is the number of indicators in group j for sustainability dimension d_i, $N^{(d_i)}$ is the total number of indicators in sustainability dimension d_i, and $N^{(d_i)} = \sum_j \mathrm{N}_j^{(d_i)} \,.\, \mathrm{w}_{jk}^{(final)}$ is the final weight of each indicator, which can be used in the aggregation model of MCA, $\sum_j \sum_k w_{jk}{}^{(final)} = 1$. $w^{(d_i)}$ is the weight of sustainability dimension d_i, $\sum_i w^{(d_i)} = 1$.

Assume we want to weight *n* indicators using AHP ($n \leq 7$ in this framework). The detail steps to calculate weight of indicator by AHP is as follows.

Step 1: We make pair-wise comparisons between each pair of indicators, and assign comparison scale to reflect their relative importance, according to information provided in Table 2-1. Thus we obtain a $(n \times n)$ positive reciprocal matrix A, after $n \times (n-1)/2$ pairs of comparison, because the matrix is reciprocal. Where A_{ij} is the comparison scale of indicator *i* with respect to indicator *j* considering the object of sustainability assessment, A_{ji} is reciprocal to A_{ij} ($A_{ji} = {}^{1}\!/_{A_{ij}}$), and diagonal element $A_{ij} = 1$. For example, if indicator *i* is extremely more important than indicator j, then $A_{ij} = 9$, and $A_{ji} = 1/9$. If indicator *i* is moderately more important than indicator j, then $A_{ij} = 3$, and $A_{ji} = 1/3$.

Step 2: We check the consistency of matrix A. One matrix is said to be consistent if $A_{ij} \times A_{jk} = A_{ik}, \forall i, j, k$. We use Equation 3.22 to calculate Consistency Index (CI) and Consistency Ratio (CR). If the C $(n) \leq CR_t(n)$, the matrix is consistent, and can be used to derive weight of indicators. Otherwise, users need to modify pair-wise comparison matrix A until its consistency ratio meet the requirement. Usually if $n \geq 5$ $CR_t(n) = 0.1$ (Krajnc and Glavič, 2005; Saaty, 1987), $CR_t(4) = 0.08$, and $CR_t(3) = 0.05$ (Saaty, 2008).

$$CI(n) = \frac{\lambda_{max} - n}{n - 1}$$

$$CR(n) = \frac{CI(n)}{RCI(n)} \tag{3.22}$$

Where, $CI(n)$, $CR(n)$ is the consistency index and consistency ratio respectively of a matrix with size $n \times n$. n is the number of indicators. λ_{max} is the principal eigenvalue of matrix A, the corresponding principal eigenvector of matrix A is unit vector $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3, \dots, \nu_n)$. $RCI(n)$ is the random consistency index. It is obtained by calculating the average eigenvalues of a very large samples (Saaty, 2008), which are randomly generated reciprocal matrices using scales defined in Table 2-1. Values of $RCI(n)$ with $3 \leq n \leq 15$ are provided in Table 3-5.

Table 3-5 Random Consistency Index for Matrices with Different Size (Saaty, 2008)

n	3	4	5	6	7	8	9	10	11	12	13	14	15
RCI	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49	1.52	1.54	1.56	1.58	1.59

If $CR(n) > CR_t(n)$, three things are conducted (Saaty, 2008; Saaty and Tran, 2007): 1) Find the most inconsistent comparison scale in matrix A, which has the largest inconsistency error. The inconsistency error of comparison scale A_{ij} is $\varepsilon_{ij} = A_{ij} \times \nu_j / \nu_i$ (If A_{ij} is consistent, $\varepsilon_{ij} = 1$). 2). Determine the range of values to which the comparison scale can be changed so that the inconsistency would be improved. 3) Ask user to consider if he or she can change the comparison scale to a plausible value in that range. If the user is unwilling to change or the change made is inadequate to reduce inconsistency, then the second most inconsistent comparison scale in matrix A needs to be examined carefully. The process is kept going until the matrix becomes consistent.

Step 3: We derive relative weight of each indicator or criteria based on matrix A. A quick way to find the approximate normalized weight of each indicator is normalizing each column in matrix A (Krajnc and Glavič, 2005). For example, dividing the relative weight of indicator j by the sum of relative weights in column j, and then averaging the values across row j to get the normalized weight of indicator j, w_j. However, Saaty (1987) recommended that for important applications one should only use the **principal eigenvector** of matrix A to get the normalized weight of indicators, because approximations can lead to rank reversal in spite of its closeness of result to eigenvector. Assume the principal eigenvector of matrix A is unit vector $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3, \dots, \nu_n)$, and the corresponding eigenvalue is λ_{max}, then the normalized weight of indicator j is $w_j = \nu_j/(\sum_{i=1}^{n} \nu_i)$.

Step 4: Finally, we can use the normalized weights of each indicator to calculate its dimensional weight and final weight by Equation 3.21. And then use the weights of indicators to calculate CSI, by Equation 3.23.

3.5.3.3 Aggregating Indicators

To construct the CSI, we need to aggregate the indicators after analyzing each indicator and obtaining its weight. There are many models to aggregate indicators, among them additive aggregation model and geometric aggregation model are used in the proposed framework.

The additive aggregation model is the simplest aggregation method, and a linear additive model is used in the proposed framework. It sums the weighted and normalized individual indicators to construct the CSI, I_{CSI}, see Equation 3.23 below.

$$I_{CSI}^{(di)} = \sum_{j=1}^{g^{(d_i)}} \sum_{k=1}^{K_j} \mathrm{w}_{jk}^{(d_i)} \times I_{jk} \times Sgn(I_{jk})$$

$$I_{CSI}^{(overall)} = \sum_{d_i} \sum_{j=1}^{g^{(d_i)}} \sum_{k=1}^{K_j} \mathrm{w}_{jk}^{(final)} \times I_{jk} \times Sgn(I_{jk}) \tag{3.23}$$

Where, $I_{CSI}^{(di)}$ is the CSI in dimension d_i, $I_{CSI}^{(overall)}$ is the overall CSI among all dimensions of sustainability. $g^{(d_i)} = \left\lceil \frac{N^{(d_i)}}{7} \right\rceil$ is the number of groups of indicators in sustainability dimension d_i, $K_j$is the number of indicators in group j. $N^{(di)} = \sum_{j=1}^{g^{(d_i)}} K_j$ is the total number of indicators in sustainability dimension d_i. $Sgn\left(I_{jk}\right) \in \{+1, -1\}$ is the sign of indicator I_{jk}, which imply the positive or negative impact of indicatorI_{jk} to the sustainability objective. Note that $\sum_{d_i} \sum_{j=1}^{g^{(d_i)}} \sum_{k=1}^{K_j} \mathrm{w}_{jk}^{(final)} = 1$.

The linear additive aggregation model assumes that the indicators are mutually preferentially independent (OECD, 2008). This assumption implies full compensability of all indicators, such that poor performance in some indicators can be compensated by sufficiently good performance in other indicators (Saltelli et al., 2008). However, the assumption is unrealistic for sustainability indicators because there are inter-linkages between indicators. Linear additive aggregation could thus result in a biased CSI. In other words, the CSI would not entirely reflect the information of its individual sustainability indicators. The geometric aggregation model is a non-compensable model and it can overcome some shortages of the linear additive model. It aggregates sustainability indicators by Equation 3.23.

$$I_{CSI}^{(di)} = \prod_{j=1}^{g^{(d_i)}} \prod_{k=1}^{K_j} {I_{jk}}^{w_{jk}^{(final)} \times Sign(I_{jk})}$$
$$I_{CSI}^{(overall)} = \prod_{d_i} \prod_{j=1}^{g^{(d_i)}} \prod_{k=1}^{K_j} {I_{jk}}^{w_{jk}^{(final)} \times Sign(I_{jk})} \tag{3.23}$$

3.5.3.4 PCA

AHP is used to derive the weights of criteria or indicators when they are independent and outer-dependent to its upper-level criteria (Liu et al., 2014). Therefore, if some indicators correlated with other indicator or indicators, we need to remove the correlation between them to make them independent to each other. PCA is thus a method that can summarize the correlations among a set of indicators with a smaller set of linear combinations, without losing much information. In other words, PCA transforms a set of correlated indicators into uncorrelated orthogonal indicators and keep the total variance of indicators as much as possible. PCA can be used to analyze the correlations between indicators and to select a set of indicators for MCA. It can also be used to generate CSIs. The process to use PCA to construct the CSIs are illustrated in detail below.

Assume there are *n* indicators, *m* polygons or cells in the study area, and we want to calculate the CSI for each polygon or each cell in the study area. Let $\boldsymbol{x_i}$ be the indicator vector for polygon i in the study area, then $\boldsymbol{x_i} = (x_{i1}, x_{i2}, \dots, x_{in})^T \in R^{n\times 1}$, matrix X = $(\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_m}) \in R^{n\times m}$. Let $\mathbf{w} = (w_1, w_2, \dots, w_n)^T \in R^{n\times 1}$ be the weight vector of indicators defined manually by users or obtained by AHP, w_j is the weight of indicator *j*. Follows are detail steps to calculate the CSI by conducting PCA on matrix X.

Step 1: Normalize data used for PCA by Z-Score normalization method (See Equation 3.24), so that we can minimize the effect of outliers to the performance of PCA.

$$\boldsymbol{z_i} = (\boldsymbol{x_i} - \boldsymbol{\mu})/\boldsymbol{S_i}$$

$$\boldsymbol{\mu} = \frac{1}{m}\sum\nolimits_i \boldsymbol{x_i} = (\mu_1, \mu_2, \dots, \mu_n)^T \tag{3.24}$$

$$\boldsymbol{S_i} = \frac{1}{m-1}\sum\nolimits_i (\boldsymbol{x_i} - \boldsymbol{\mu})^{\mathrm{T}}(\boldsymbol{x_i} - \boldsymbol{\mu}) = (S_1, S_2, \dots, S_n)^T$$

Where, $\boldsymbol{z_i} = (z_{i1}, z_{i2}, \dots, z_{in})^T \in R^{n\times 1}$ is the normalized indicator vector for polygon i in the study area, and $z_{ij} = (x_{ij} - \mu_j)/S_j$, z_{ij} is the normalized value of indicator *j* for polygon *i*, μ_j, S_j is the mean and standard deviation value of indicator *j*.

Step 2: Estimate covariance matrix C, using Equation 3.25.

$$C = \frac{1}{m}\sum\nolimits_i (\boldsymbol{x_i} - \boldsymbol{\mu})(\boldsymbol{x_i} - \boldsymbol{\mu})^T \in R^{n\times n} \tag{3.25}$$

Step 3: Get the eigenvalue and unit eigenvector of covariance matrix C, and sort the eigenvalues in descending order. Therefore, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and the corresponding unit eigenvectors are $\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_n}$ ($\boldsymbol{e_i} \in R^{n\times 1}$). An important property of the eigenvalues is that they add up to the sum of diagonal elements of covariance matrix ($\sum_i \lambda_i = \sum_i C_{ii}$). In other words, $\sum_i \lambda_i$ is equal to the sum of variance of each indicator used in the analysis, and each component's eigenvalue represents how much variance it explains. Thus the percentage of variance explained by principle component $\boldsymbol{e_i}$ can be calculate as $\mathrm{w}_i^\lambda = \frac{\lambda_i}{\sum_i \lambda_i} \times 100\%$.

Step 4: Choose the first k eigenvectors $(\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_k})$ with the corresponding eigenvalues are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$. Where the cumulative amount of variance explained by these k eigenvalues can reach up at least 90% ($\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \geq 0.9$). Then we can form the transformation matrix $P = (\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_k})^T \in R^{k \times n}$ and the k principal components' weight vector $\mathbf{w}^{\boldsymbol{\lambda}} = (w_1^\lambda,\ \mathrm{w}_2^\lambda, \dots, \mathrm{w}_k^\lambda)^T$.

Step 5: Use Equation 3.26 to compute the coordinate of $\boldsymbol{x_i}$ (i.e. i,ndicator vector for polygon *i* in the study area) and the transformed weight of $\boldsymbol{x_i}$ in the *k* dimension new space with basis defined by the k principle eigenvectors, $\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_k}$.

$$\begin{aligned} \boldsymbol{x_i^{pc}} = P\boldsymbol{z_i} &= \left(\frac{\boldsymbol{e_1}^{\boldsymbol{T}}(\boldsymbol{x_i} - \boldsymbol{\mu})}{\boldsymbol{S_i}}, \frac{\boldsymbol{e_2}^{\boldsymbol{T}}(\boldsymbol{x_i} - \boldsymbol{\mu})}{\boldsymbol{S_i}}, \dots, \frac{\boldsymbol{e_k}^{\boldsymbol{T}}(\boldsymbol{x_i} - \boldsymbol{\mu})}{\boldsymbol{S_i}}\right)^T \\ &= \left(\frac{\boldsymbol{e_1}^{\boldsymbol{T}}(\boldsymbol{x_i} - \boldsymbol{\mu})}{\sqrt{\lambda_1}}, \frac{\boldsymbol{e_2}^{\boldsymbol{T}}(\boldsymbol{x_i} - \boldsymbol{\mu})}{\sqrt{\lambda_1}}, \dots, \frac{\boldsymbol{e_k}^{\boldsymbol{T}}(\boldsymbol{x_i} - \boldsymbol{\mu})}{\sqrt{\lambda_k}}\right)^T \end{aligned} \tag{3.26}$$

$$\mathbf{w}^{\boldsymbol{pc}} = P\boldsymbol{w} = (\boldsymbol{e_1}^{\boldsymbol{T}}\boldsymbol{w},\ \boldsymbol{e_2}^{\boldsymbol{T}}\boldsymbol{w}, \dots,\ \boldsymbol{e_k}^{\boldsymbol{T}}\boldsymbol{w})^T \in R^{k \times 1}$$

$$\boldsymbol{w_{final}^{pc}} = < \mathbf{w}^{\boldsymbol{pc}}, \mathbf{w}^{\boldsymbol{\lambda}} > = (\boldsymbol{e_1}^{\boldsymbol{T}}\boldsymbol{w} \times \mathrm{w}_1^\lambda,\ \boldsymbol{e_2}^{\boldsymbol{T}}\boldsymbol{w} \times \mathrm{w}_2^\lambda, \dots,\ \boldsymbol{e_k}^{\boldsymbol{T}}\boldsymbol{w} \times \mathrm{w}_k^\lambda)^{\boldsymbol{T}}$$

Step 6: Calculate the CSI for polygon i in the study area, using the k principle components by Equation 3.27. If all the indicators used for PCA are from the sustainability dimension d_i, then Equation 3.27 gives the CSI of dimension d_i for polygon i in the study area. If the indicators used for PCA are from all dimensions of sustainability, then Equation 3.27 gives the overall CSI for polygon i in the study area.

$$I_{CSI}(i) = \frac{1}{\sum_{j=1}^{k}\left(\boldsymbol{e_j}^{\boldsymbol{T}}\boldsymbol{w} \times \mathrm{w}_j^{\lambda}\right)} \sum_{j=1}^{k} x_{ij}^{pc} \times (\boldsymbol{e_j}^{\boldsymbol{T}}\boldsymbol{w} \times \mathrm{w}_j^{\lambda}) \tag{3.27}$$

PCA can summarize a set of linearly correlated individual indicators while preserving the maximum possible proportion of total variations in the original data set (OECD, 2008). When the indicators are not linearly dependent, kernel-PCA should be used for this analysis. However, PCA is not a propriety method when there is no clear correlation between indicators (i.e., the correlation coefficient is not very high). Also it is sensitive to modifications in the basic data, such as data revisions and updates (OECD, 2008). Besides, PCA does not have good performance when the data samples are not sufficient. It is also very difficult to explain the results obtained by PCA, such as the relationship between indicators and how much each indicator contributes to the CSI.

3.6 Assessment Results Analysis

The comprehension of our world makes it very difficult to define boundaries or thresholds between what does contribute to sustainable development and what does not. Besides, due to the limit of data, measurements and our understanding or knowledge of sustainability, uncertainties exits in the whole process of sustainability assessment. For example, the assumptions in estimating of sustainability, mechanism for including or excluding indicators in the assessment, transformation and/or trimming of indicators, normalization schemes, choice of imputation algorithm, choice of weights and aggregation system, and so on (Saltelli et al., 2008; Singh et al., 2007). Sensitivity analysis evaluates how the sustainability assessment result varies with the uncertainties in the input of the sustainability assessment model. Uncertainty analysis aims to quantify the overall uncertainty of sustainability assessment as a result of uncertainties in the model inputs (OECD, 2008). Sensitivity analysis is highly related to uncertainty analysis. Therefore, a combination of "uncertainty analysis" and "sensitivity analysis" can help guarantee both robustness and transparency of the sustainability assessment results, further help the public or policy makers know how confident they should be to the sustainability assessment result and know the risk of their polies (Ciuffo et al., 2012).

3.6.1 *Uncertainty Analysis*

Uncertainty analysis recognizes the uncertainty of each sustainability assessment. It identifies how the uncertainty in inputs and parameters of sustainability assessment model propagates into the sustainability assessment results (Ciuffo et al., 2012) and how large the uncertainty of the sustainability assessment results are (e.g., confidence interval at a given confidence level). There are many forms of uncertainty, including parameter uncertainty, scenario uncertainty, and model uncertainty (Lee et al., 2017). Parameter uncertainty includes the observation or measurement errors of inputs of the model. Scenario uncertainty includes users' choices regarding the functional unit, values, weighting factors, time horizons, geographical scales, natural contexts, allocation procedures, waste-handling scenarios, use of environmental thresholds and expected technology trends (Lee et al., 2017; Lloyd and Ries, 2007). Model uncertainty includes models for deriving outputs and characterization factors (Lee et al., 2017; Lloyd and Ries, 2007). In this framework, we focused mainly on parameter uncertainty, but readers or users can easily add other forms of uncertainty in this framework in the future. Since there are many "uncertain" parameters or variables to be evaluated in the uncertainty analysis, the efficiency of uncertainty analysis can be improved if identifying "key input variables", which contribute considerably to the uncertainty of sustainability assessment result. Global sensitivity analysis (Leamer, 1983; Saltelli, 2002) or Contribution to Variance (CTV) is an effective method to identify these 'significant variables' (Ciuffo et al., 2012).

The Intergovernmental Panel on Climate Change (IPCC) provides good practice guidance on a variety of uncertainty analysis methods (LEE et al., 2017; IPCC, 2006). Among them, "Monte Carlo Simulation" (MCS) method consider both covariance and issues of dependence and correlation in the uncertainty analysis. MCS assumes that all input variables follow some probability distributions (or statistical distribution), including parametric distribution and non-parametric distribution (e.g., lognormal and normal distributions) (Lee et al., 2017). If it is very difficult to get the probability distribution of input variables, users can try other uncertainty methods (Chen and Corson, 2014; Lee et al., 2017) or use the range of input variables or other parameters in MCS, instead of its probability distribution (Ciuffo et al., 2012). In the proposed framework, we use MCS to evaluate the uncertainty of sustainability assessment results.

MCS generate many random but plausible combinations of indicator values and weights and then evaluate how the sustainability result change with these randomly selected indicator scores and weights. MCS first select ranges and distributions for each input parameter to characterize their uncertainty (Kotek et al., 2007). It then generates samples (with N elements) for each input parameter from its selected distribution (Kotek et al., 2007). In the end, it evaluates the model using each element of the samples and then analyzes the uncertainty of this model (Kotek et al., 2007). The evaluation results of MCS may be a normal distribution of these combinations. If the assessment result falls near the median score for each sustainability dimension and has little variance in the probability distribution, then the assessment will be considered reliable (Golder Associates, 2011).

Instead of Primitive Monte Carlo (PMC) simulations, the proposed framework uses Latin Hypercube Sampling (LHS) to generate random samples for the uncertainty analysis. LHS is an extension of quota sampling (Helton and Davis, 2003). It is a stratified sampling procedure (Kotek et al., 2007; Pfister and Scherer, 2015), and ensures full coverage of all input parameter ranges. LHS is a widely used sampling technique for the propagation of uncertainty analyses and sensitivity analysis of complex systems (Helton and Davis, 2003).

In LHS, the region between 0 and 1 is uniformly divided into N non-overlapping intervals for each random variable. For each random variable, one value per interval is randomly selected without replacement in the same probability. This can be accomplished by initially generating N random numbers (u) within range $[0,1]$, and then using them to generate random numbers in the *i-th* non-overlapping intervals (u_i) for each random variable by Equation 3.28 (Khatri, 2013). If the variable is in range $[a, b]$, then we can use $u_i(b-a)+a$ as its random value in the *i-th* non-overlapping intervals.

$$u_i = \frac{u}{N} + \frac{(i-1)}{N} \quad (i = 1,2,\dots,N; \text{random number } u \in [0,1]) \tag{3.28}$$

In this framework, LHS (Helton & Davis, 2003; Khatri, 2013) is used to build $N * k$ samples with *N* elements for k input parameters of model $Y = f(X_1, X_2, \ldots, X_k)$. The range of each parameter is divided into *N* non-overlapping intervals of equal probability $1/N$. One value from each interval is random selected, thus each parameter has a sequence of *N* random samples (values). The *N* values of parameter X_1 are paired randomly without replacement with the *N* values of parameter X_2. And then these *N* pairs of (X_1, X_2) are combined in a random manner without replacement with the *N* values of parameter X_3 to form *N* triples of (X_1, X_2, X_3). This process is continued until a set of *N* k-tuples are formed. These *k*-tuples are of the form $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ik}), i = 1,2, \ldots, N$ (Helton and Davis, 2003). In other words, $\boldsymbol{x}_i$ is one possible values of these *k* parameters. Therefore, we can conduct N evaluations of the model and get N sustainability assessment results ($Y_i = f(\boldsymbol{x}_i), i = 1,2, \ldots, N$), in order to analyze its uncertainty. After getting the *N* sustainability assessment results from MCS, we can use their variance (${\sigma_Y}^2$), Coefficient of Variation (CV), and the output probability distribution (e.g., confidence interval and quantiles) to quantify uncertainty of model outputs (see Equation 3.29 and Equation 3.30). We can also use the N sustainability assessment results from MCS to calculate the average absolute rank difference (see Equation 3.31), which is also used by Saltelli et. al. (2008) for the sensitivity analysis of composite indicators.

$$\sigma_Y = \sqrt{\sum_{i=1}^{N} (Y_i - \mu_Y)^2}$$

$$CV = \frac{\sigma_Y}{\mu_Y} \qquad \mu_Y = \frac{\sum_{i=1}^{N} Y_i}{N} \tag{3.29}$$

$$CI_Y = \left[\mu_Y - \frac{\sigma_Y}{\sqrt{N}} * t_{\alpha,v}\ ,\ \mu_Y + \frac{\sigma_Y}{\sqrt{N}} * t_{\alpha,v}\right]\ ;\quad \alpha = \frac{1-C}{2}, v = N-1 \tag{3.30}$$

Where μ_Y is the mean value of these N model estimations, σ_Y is its standard deviation. CV is also known as Relative Standard Deviation (RSD). CI_Y is the confidence interval at the confidence level C, the most commonly used C is 90%, 95%, and 98%. However, users can choose their own confidence level in this framework. $t_{\alpha,v}$ is the statistical value obtained from student *t*-distribution table, with $(N-1)$ degree of freedom at confidence level C. If the sustainability assessment result Y_{ref} is within CI_Y, then we can say at confidence level C, the assessment result is reliable.

$$\overline{RS}^j = \frac{1}{N}\sum_{i=1}^{N}\left|Rank\left(Y_i{}^j\right) - Rank\left(Y_{ref}{}^j\right)\right|$$
$$\overline{RS} = \frac{1}{M}\sum_{j=1}^{M}\overline{RS}^j \tag{3.31}$$

For polygon *j* in the study area, $\overline{RS}^j$ is its scope of uncertainty (i.e. av,erage range difference), $Rank\left(Y_{ref}{}^j\right)$ is its original estimated rank over all the *M* polygons in the study area, $Rank\left(Y_i{}^j\right)$ is its rank of the *i-th* sustainability assessment results evaluated from MCS. $\overline{RS}$ is the overall average rank differences for all the *M* polygons in the study area for the *N* sustainability assessment results.

LHS requires fewer iterations to archive robust results than PMC, and it is about five times more efficient than PMC. Typically, 10,000 iterations are used for PMC (Pfister and Scherer, 2015); thus we can use 1,000 iterations (i.e., $N = 1000$) for LHS in this proposed framework.

3.6.2 Sensitivity Analysis

Sensitivity Analysis (SA) sometimes refers to vulnerability analysis or importance ranking (Ciuffo et al., 2012). SA studies how uncertainty in the output of a model can be apportioned to different sources of uncertainties in the model input (Saltelli et al., 2008). In other words, SA explores how the model output would change with variations of parameters in the model and how they interact with each other (Saltelli et al., 2010). Here in this framework, sensitivity analysis measures how the given sustainability assessment result depends upon the data or method used for sustainability assessment model and how the results change with the input of sustainability assessment model. Ciuffo et al. (2012) introduce many sensitivity analysis methods, including input/output scatter-plots, Sigma-normalized derivatives, standardized regression coefficient, elementary effects test, variance-based techniques, Monte Carlo filtering (Saltelli et al., 2008) and meta-modelling. Among them, input/output scatter-plots is the simplest way to perform sensitivity analysis, but it only analyzes the sensitivity of one parameter a time. If there are many variables in the sustainability model, the scatter-plot method becomes unpractical. Sigma-normalized derivatives is a natural way to perform sensitivity analysis, especially for analytical models. However, the derivatives' computation is very time expensive. Standardized regression coefficients are more robust and reliable than sigma- normalized derivatives because it explores the entire space of input variables (Ciuffo et al., 2012). However, its precision is dependent on the size of Monte Carlo experiment. When one is not interested at studying the specific value of sustainability assessment result but is interested in if the assessment result is above or below a certain threshold, a Monte Carlo filtering can be used (Ciuffo et al., 2012). Variance-based methods for sensitivity analysis were first employed by Cukier et al. (1973) and generalize by Sobol to provide a Monte Carlo-based implementation of the concept (Ciuffo et al., 2012). It is model-free and explores the whole range of variation of the input factors, instead of just sampling factors over a limited number of values (OECD, 2008). Besides, it can analyze many uncertain factors simultaneously instead of individually. Furthermore, it is easy to interpret the sensitivity analysis results. Therefore, the Monte Carlo-based implementation of Variance-based method is used in the proposed framework to conduct sensitivity analysis.

Given a model $= f(X_1, X_2, \ldots, X_k)$, with X_i as the *i-th* input variable (e.g., indicators and weights in sustainability assessment model), Y as a scalar. The first order sensitivity index for factor x_j can be calculated by Equation 3.32 (Saltelli, 2002; Saltelli et al., 2010). A high value of S_j implies X_j is an important factor for the model (Saltelli et al., 2008).

$$S_j = \frac{V_{X_j}(E_{\mathbf{X}_{-j}}(Y \mid X_j))}{V(Y)} \in [0,1] \tag{3.32}$$

Where $V(Y)$ is the variance of model output $V(Y) = E(Y^2) - E^2(Y)$, $V_{X_j}(E_{\mathbf{X}_{-j}}(Y \mid X_j))$ is the conditional variance of Y given factor X_j. $\mathbf{X}_{-j}$ denotes the matrix of all factors but except X_j. The mean of Y, $E_{\mathbf{X}_{-j}}(Y \mid X_j)$, taken over all possible values of $\mathbf{X}_{-j}$ while keeping X_j fixed. Variance operation $V_{X_j}(\cdot)$ takes over all possible values of X_j.

$\sum_{j=1}^{k} S_j \leq 1$, and $\sum_{j=1}^{k} S_j = 1$ only for an additive model with independent factors (Ciuffo et al., 2012; Saltelli et al., 2004). For non-additive model, higher order sensitivity indices have to be computed to account for interaction effects among sets of input factors (OECD, 2008). However, higher order sensitivity indices are usually not estimated, due to expensive computing cost. Thus total effect sensitivity index is calculated instead. Total effects sensitivity index of the input factor X_j can be calculated by Equation 3.33, details see (OECD, 2008; Saltelli et al., 2008, 2010). It accounts for the total contribution of factor X_j to the variance of Y, i.e. it equals its first-order effect plus all higher-order effects due to interactions.

$$S_{Tj} = \frac{\mathrm{E}_{\mathbf{X}_{-j}}(\mathrm{V}_{\mathrm{X}_j}(\mathrm{Y} \mid \mathbf{X}_{-j}))}{V(Y)} = \frac{V(Y) - \mathrm{V}_{\mathbf{X}_{-j}}(\mathrm{E}_{\mathrm{X}_j}(\mathrm{Y} \mid \mathbf{X}_{-j}))}{V(Y)} \tag{3.33}$$

S_{Tj} measures the synthetically interactions that involve factor X_j. $S_{Tj} \geq S_j$ if factor X_j is not involved in any interaction with other input factors (Saltelli et al., 2008). $\mathrm{V}_{\mathbf{X}_{-j}}(\mathrm{E}_{\mathrm{X}_j}(\mathrm{Y} \mid \mathbf{X}_{-j}))$ is the total contribution to the variance of Y due to all the $(\mathrm{k}-1)$ non X_j factors (i.e. $\mathbf{X}_{-j}$). It is the first order effect of $\mathbf{X}_{-j}$. Therefore, $V(Y) - \mathrm{V}_{\mathbf{X}_{-j}}(\mathrm{E}_{\mathrm{X}_j}(\mathrm{Y} \mid \mathbf{X}_{-j}))$ is the contribution of factor X_j to the variance of Y, details see (Saltelli et al., 2010). In general, $\sum_{j=1}^{k} S_{Tj} \geq 1$, and $\sum_{j=1}^{k} S_{Tj} = 1$ only for additive model. If $S_{Tj} \cong 0$, then factor X_j can be fixed at any value within its range of uncertainty without appreciably affecting the outputs' variance $V(Y)$ (Saltelli et al., 2008).

Pair (S_j, S_{Tj}) gives a fairly good description of the model sensitivities. For a given factor X_j, $(S_{Tj} - S_j)$ is a measure of how much X_j is involved in interactions with any other input factor. A significant difference between S_{Tj} and S_j implies an important role of interactions for factor X_j in model $Y = f(X_1, X_2, \ldots, X_k)$ (Saltelli et al., 2008). In this case, we can analyze the second order sensitivity index to see details of the interactions. Herman (2013) suggested that for total-order indices to be important, they usually need to be above 0.05 at the very least and the most dominant parameters can have values upward of 0.8.

There are many methods to estimate S_j and S_{Tj}, see (Saltelli et al., 2008) for details. These methods usually first generate certain number of samples to represent or describe the distribution of input factors of one model, then use these samples to estimate output of the model, and finally use these outputs to calculate the sensitivity indices. When the input factors (parameters) of model $Y = f(X_1, X_2, \ldots, X_k)$ are independent, the methodology described in (Ciuffo et al., 2012; Saltelli et al., 2008, 2010) is used in this framework. It first uses quasi-random sampling method to get $N * (k + 2)$ different pairs of input factor values, and then estimate $N * (k + 2)$ model outputs (is the number of samples generated by quasi-random sampling method) (OECD, 2008). Finally, it uses these outputs to calculate sensitivity indices. Section 3.4.2.1 illustrates detail steps to calculate sensitivity indices of model with independent input factors (Ciuffo et al., 2012). When the input factors of model $Y = f(X_1, X_2, \ldots, X_k)$ are correlated, McKay's method (McKay, 1995) is used in this framework. Its details are illustrated in section 3.4.2.2.

3.6.2.1 Model with Independent Factors

As mentioned earlier, the factors of sustainability assessment model are usually dependent or correlated. When using this method to conduct sensitivity analysis, the proposed sustainability assessment framework assumes factors or parameters of sustainability assessment model are independent, and the dependencies of these factors are treated as explicit relationships with a noise term as suggested by (Saltelli et al., 2008). The process of evaluating the sensitivity of parameters in model $Y = f(X_1, X_2, \ldots, X_k)$ is illustrated in details below.

Step 1: Generate two independent matrix of quasi-random numbers with size (N, k) as the input matrix of model $Y = f(X_1, X_2, \ldots, X_k)$. These two matrix are represent by matrix **A** and matrix **B**. They can be easily generated from a quasi-random sequence of size $(N, 2k)$: matrix **A** is the left half of the quasi-random sequence and matrix **B** is the right half of it (see Equation 3.34). *N* is the number of elements (or samples) generated by Monte Carlo Simulation, it can vary from few hundread to several thousands ($N \geq 500$ (Saltelli et. al., 2010)). k is the number of input factors in the model. Each row of matrix ***A*** or matrix ***B*** is a realization from the multivariate joint probability distribution of the set X (Saltelli et al., 2004), it is one possible input values for model $Y = f(X_1, X_2, \ldots, X_k)$. Column *i* in the matrix is a sample from the marginal distribution of the corresponding input factor X_i .

$$\boldsymbol{A} = \begin{pmatrix} z_{11},\ z_{12}, \cdots,\ z_{1k} \\ z_{21},\ z_{22}, \cdots,\ z_{2k} \\ \cdots \\ z_{N1},\ z_{N2}, \cdots,\ z_{Nk} \end{pmatrix} \quad \boldsymbol{B} = \begin{pmatrix} z_{1,k+1},\ z_{1,k+2}, \cdots,\ z_{1,2k} \\ z_{2,k+1},\ z_{2,k+2}, \cdots,\ z_{2,2k} \\ \cdots \\ z_{N,k+1},\ z_{N,k+2}, \cdots,\ z_{N,2k} \end{pmatrix} \tag{3.34}$$

Step 2: Generate a set of *k* matrices $\boldsymbol{C_i}$ $(i = 1,2,3, \ldots, k)$. Matrix $\boldsymbol{C_i}$ is obtained by replacing only the *i-th* column of matrix ***A*** with the *i-th* column of matrix ***B*** (highlighted below in matrix $\boldsymbol{C_i}$).

$$\boldsymbol{C_i} = \begin{pmatrix} z_{11}, \ z_{12}, \dots, \ z_{1,i-1}, \ \boldsymbol{z_{1,k+i}}, \ z_{1,i+1}, \dots, z_{1k} \\ z_{21}, \ z_{22}, \dots, \ z_{2,i-1}, \ \boldsymbol{z_{2,k+i}}, \ z_{2,i+1}, \dots, z_{2k} \\ \dots \\ z_{N1}, \ z_{N2}, \dots, \ z_{N,i-1}, \ \boldsymbol{z_{N,k+i}}, \ z_{N,i+1}, \dots, z_{Nk} \end{pmatrix} \quad \text{for } i = 1,2,3,\dots,k$$

Step 3: Evaluate the model for all the $N * (k+2)$ combinations of input variables as given by matrices ***A*, *B*,** and $\boldsymbol{C_i}$ $(i = 1,2,..,k)$, and produce the vector of output, $\boldsymbol{Y_A} = f(\boldsymbol{A}), \boldsymbol{Y_B} = f(\boldsymbol{B}), \boldsymbol{Y_{C_i}} = f(\boldsymbol{C_i})$ for $i = 1,2,..,k$ .Vector $\boldsymbol{Y_{A,B}} = \binom{\boldsymbol{Y_A}}{\boldsymbol{Y_B}}$ includes all the output vales from matrix ***A*** and matrix ***B***. We then use these vector of output to estimate the first-order sensitivity index (S_j) and total effects sensitivity indices (S_{Tj}) of factor X_j, by Equation 3.35 and Equation 3.36 respectively (Ciuffo et al., 2012; Saltelli et al., 2010).

$$S_j = \frac{\mathrm{V}_{\mathrm{X}_j}(\mathrm{E}_{\mathbf{X}_{-j}}(\mathrm{Y} \mid \mathrm{X}_j))}{V(Y)} = \frac{\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{Y_B^i} * (\boldsymbol{Y_{C_j}^i} - \boldsymbol{Y_A^i})}{\frac{1}{2N}\sum_{i=1}^{2N}(\boldsymbol{Y_{A,B}^i})^2 - \left(\frac{1}{2N}\sum_{i=1}^{2N} \boldsymbol{Y_{A,B}^i}\right)^2} \tag{3.35}$$

$$S_{Tj} = \frac{\mathrm{E}_{\mathbf{X}_{-j}}(\mathrm{V}_{\mathrm{X}_j}(\mathrm{Y} \mid \mathbf{X}_{-j}))}{V(Y)} = \frac{\frac{1}{2N}\sum_{i=1}^{N} (\boldsymbol{Y_A^i} - \boldsymbol{Y_{C_j}^i})^2}{\frac{1}{2N}\sum_{i=1}^{2N}(\boldsymbol{Y_{A,B}^i})^2 - \left(\frac{1}{2N}\sum_{i=1}^{2N} \boldsymbol{Y_{A,B}^i}\right)^2} \tag{3.36}$$

Where, *N* is the number of samples in matrix ***A*, *B***, and $\boldsymbol{C_j}$ $(j = 1,2,\dots,k)$, each sample (i.e. eac,h row of matrix ***A*, *B***, or $\boldsymbol{C_j}$) is a possible value set for input variables $(X_1, X_2, \dots, X_k)$. $\boldsymbol{Y_A^i}, \boldsymbol{Y_B^i}, \boldsymbol{Y_{C_j}^i}$, and $\boldsymbol{Y_{A,B}^i}$ are the *i-th* element of output vector $\boldsymbol{Y_A}, \boldsymbol{Y_B}, \boldsymbol{Y_{C_j}}, \boldsymbol{Y_{A,B}}$ respectively.

3.6.2.2 Model with Correlated Factors

Sensitivity analysis of dependent factors requires much higher sample size and need much more expensive computation time than that of independent factors. There are many methods to evaluate the sensitivity of models with correlated factors. Iman and Conover (1982) proposed a method for producing replicated LHS samples with matrix of N van der Waerden scores $\Phi^{-1}\left(\frac{i}{N}+1\right)$ ($i = 1,2,3, \dots, N$; N is the sample size , Φ^{-1} is the inverse of the standard normal distribution) as an approximation of Spearman rank correlation matrix of the samples (Helton and Davis, 2003; Stein, 1987), detail procedure of this sampling method see (Helton and Davis, 2003). But when the sample size N is very large, using rank correlations to describe correlations between factors becomes inappropriate (Stein, 1987). In this case, a method proposed by Stein (1987) can be used to create replicated LHS samples for correlated factors or variables of model $Y = f(X_1, X_2, \dots, X_k)$. It is implemented in tool SIMLAB. Users can also use this tool to evaluate sensitivity of their sustainability assessment results (Saltelli et al., 2004).

The proposed framework also uses the replicated LHS method to estimate the first order sensitivity index described in (Saltelli et al., 2004) and (McKay, 1995), when factors of model $Y = f(X_1, X_2, \dots, X_k)$ are correlated. It is refer as McKay's method in the following contents. The procedure to conduct sensitivity analysis using this method is also implemented in SIMLAB (Saltelli et al., 2004). Its detail steps are as follows.

McKay's method first generate samples by replicated LHS (McKay, 1995). When using simple LHS, the value range of each factor is divided into N non-overlapping intervals or bins with equal probability (Saltelli et al., 2004). Thus the probability for one factor to fall in any of the bins is exactly $1/N$. Suppose we generate r independent such samples using LHS, an example of r-LHS samples for two factors is shown in Figure 3-12. In most cases, $r \leq 50 \ll N$ (Saltelli et al., 2004). For factor X_j , $p_i(Y|X_j = x_{j,i}^*)$ is the conditional distribution of Y when X_j is fixed to value $x_{j,i}^*$ in the *i-th* bin ($X_j = \{x_{j,1}^* , x_{j,2}^* , x_{j,3}^* , \dots , x_{j,N}^* \}$) (Saltelli et al., 2004). $y_{j,i}^l$ $(l = 1,2, \dots , r)$ is a value of Y corresponding to replica l from the conditional distribution ($y_{j,i}^l = f(X_1, X_2, \dots, X_j = x_{j,i}^*, \dots, X_k)$. Therefore, each factor has *N* possible different values, with one fixed value in each bin. And there are $r * N$ sets of values for input factors $(X_1, X_2, \dots, X_k)$, thus in total $r * N$ values of Y can be generated. Then McKay's method estimates mean output and first order sensitivity index by Equation 3.37 and Equation 3.38 respectively (Saltelli et al., 2004). In this framework, we can use $N = 500, r = 20$, as the one used by (Saltelli and Tarantola, 2002).

$$E(Y|X_j) = \bar{y_J} = \frac{1}{rN}\sum_{i=1}^{N}\sum_{l=1}^{r} y_{j,i}^l = \frac{1}{N}\sum_{i=1}^{N}\overline{y_{J,\iota}} = \frac{1}{r}\sum_{l=1}^{r} \bar{y_J}^l \tag{3.37}$$

$$E(Y|X_j = x_{j,i}^*) = \overline{y_{J,\iota}} = \tfrac{1}{r}\textstyle\sum_{l=1}^{r} y_{j,i}^l \qquad \bar{y_J}^l = \tfrac{1}{N}\textstyle\sum_{i=1}^{N} y_{j,i}^l$$

$$S_j = \frac{\frac{1}{N}\sum_{i=1}^{N} (\overline{y_{J,\iota}} - \bar{y_J})^2}{\frac{1}{rN}\sum_{i=1}^{N} \sum_{l=1}^{r}(y_{j,i}^l - \bar{y_J})^2} \tag{3.38}$$

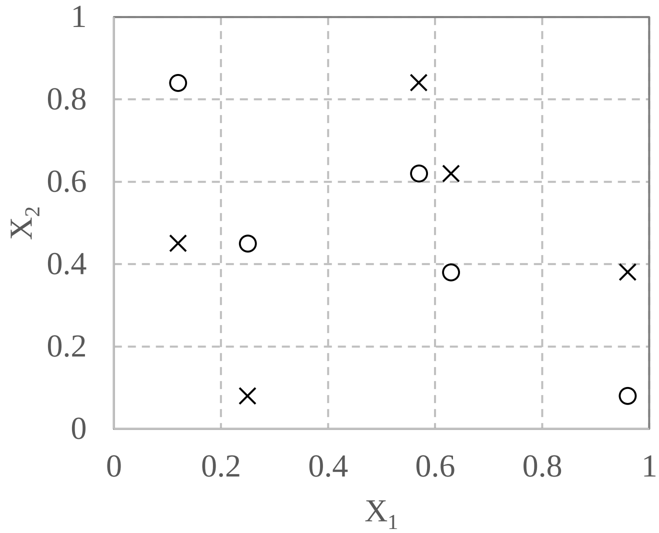


Figure 3-12 Example of r-LHS sampling for two variables (Saltelli et al., 2004)

The number of bins *N*= *5*, the number of replicates *r = 2*. In each replicate, the same values for each factor are used in each bin.

3.6.3 *Results Interpretation*

Results interpretation gives a brief interpretation of information behind sustainability assessment results in the proposed sustainability assessment framework. It includes the following information:

- CSI value of each dimension of sustainability, and overall CSI value.
- Uncertainty analysis of sustainability assessment results, such as their confidence interval and confidence level, variations, probability distribution and so on.
- Sensitivity analysis of sustainability assessment results, including the first-order sensitivity index, total effects sensitivity index and the methods to estimate these two sensitivity indices.
- Indicator weighting strategies and the working flow of calculating CSI.

- Assumptions and Limitations of the sustainability assessment.

3.7 Assessment Results Visualization

A picture or graph is worth a thousand words. Visual analytics can be a powerful tool for sharing research outcomes with diverse stakeholders. Visuals simplify the communication of complex patterns, relationships, and contexts between scientific concepts and theoretical frameworks (Community Research Connections, 2017). A good visualization could help researchers, public or policy makers to gain insights about information behind data. After finishing sustainability assessments, the proposed framework will visualize assessment results in multiple ways from different aspects, such as radar chart, sunburst chart, bar chart, line chart, histograms, and maps (geographical maps or tile grid maps). Besides, a sustainability assessment report will also be produced to summary all the data, methods and results of the assessment.

3.7.1.1 Radar Chart

Radar chart uses a polygon with black lines to represent the sustainability assessment results, which is normalized to range [0, 1]. The number of sides of the polygon equals the number of dimensions of sustainability. Each vertex of polygon indicates the sustainability assessment result of the corresponding dimension of sustainability (i.e., dimensional composite sustainability index). The confidence intervals of sustainability assessment results are the shaded area in the figure. The specified values of the confidence interval and corresponding confidence level are labeled besides the axis of each dimension of sustainability. The very outside polygon represents the ideal sustainability assessment results, with each dimension reaching its maximum value, which is one. Examples of bar chart are shown in Figure 3-13 and Figure 3-14, for four dimensions of sustainability and three dimensions of sustainability respectively.

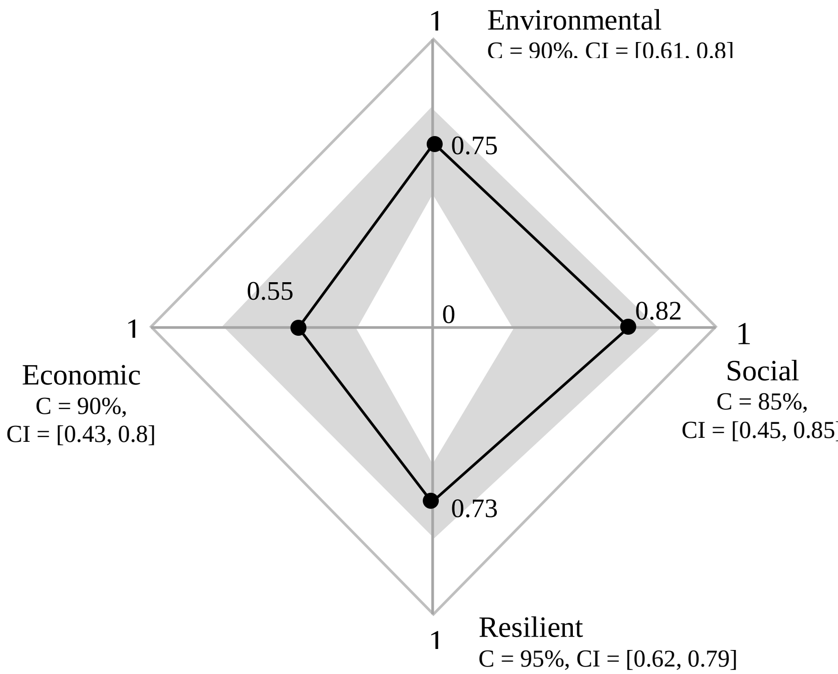


Figure 3-13 Radar Chart of Sustainability Assessment Results (Four Dimensions)

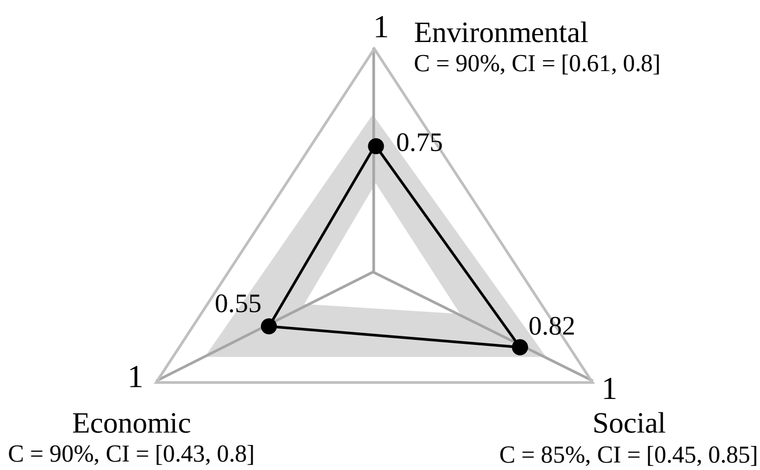


Figure 3-14 Radar Chart of Sustainability Assessment Results (Three Dimensions)

3.7.1.2 Sunburst Chart

The sunburst chart is ideal for displaying hierarchical data. Each level of the hierarchy is represented by one ring or circle with the innermost circle as the top of the hierarchy. Since the organization of sustainability indicators is in hierarchical structure， sunburst chart is a good way to represent sustainability results. Heeres et al. (2012) also use a sunburst chart to represent sustainability indicator values. In this framework, the sunburst chart can represent the sustainability assessment results, the normalized value of indicators, and weights of indicators.

An example of a sunburst chart is shown in Figure 3-15. The very inner circle marked out overall sustainability assessment result (i.e., overall composite sustainability indicator). The second level fans represent the sustainability assessment results for each dimension of sustainability (i.e., dimensional composite sustainability indicator). The width and height of fans depict for the weight and normalized value of each dimension when calculating the overall sustainability respectively. The color of each fan stands for its sustainability dimension. The third level fans represent the individual sustainability indicators in each dimension of sustainability. The width and height of fans represent the final weight and normalized value of each indicator when calculating the overall sustainability respectively. The star sign above each fan indicates the corresponding indicator is important to the sustainability assessment result, based on sensitivity analysis results. The color of each fan also indicates its sustainability dimension.

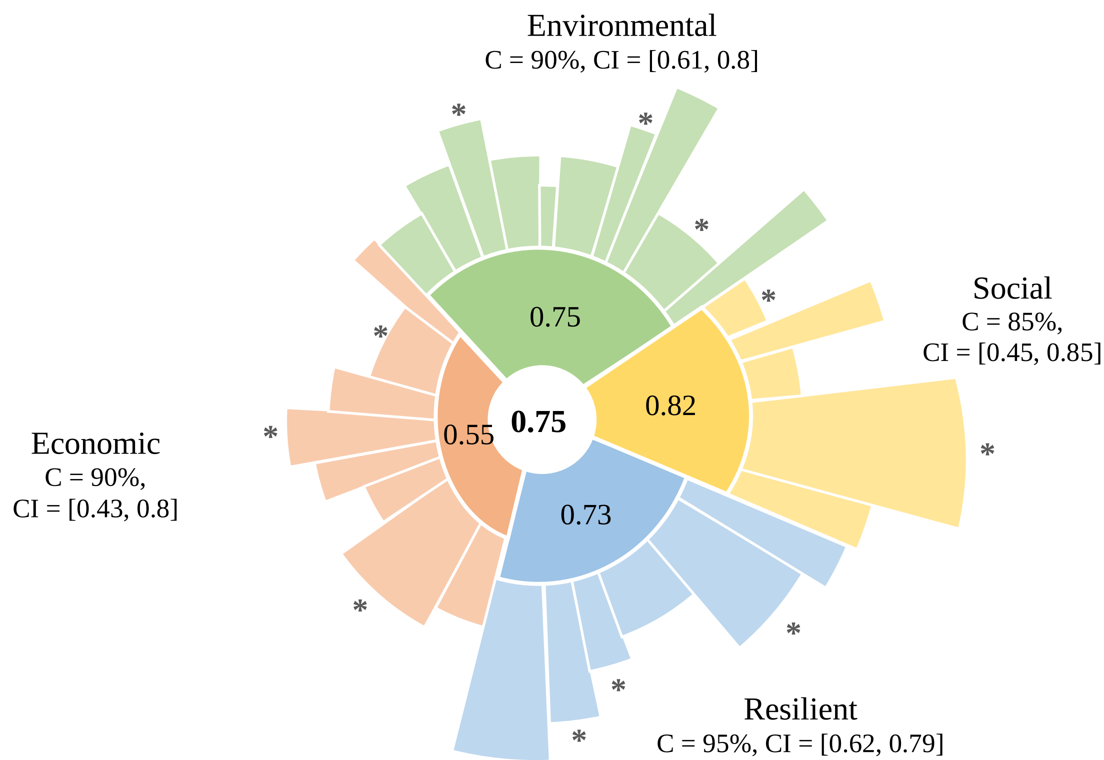


Figure 3-15 Sunburst Chart of Sustainability Assessment Results (Four Dimensions)

Sunburst chart works very well with small amounts of indicators. With the increase of the number of indicators, it may become very difficult to distinguish or recognize each fan and its information. Therefore, when the number of indicators is very big, users can use the stratified bar chart to show sustainability assessment results and its hierarchical structure, an example is shown in Figure 3-16.

3.7.1.3 Stratified Bar Chart

A stratified bar chart uses bars to represent the sustainability assessment results and organizations of indicators. The weight and height of each bar indicate the value and weight of the indicator it represents respectively. An example of a stratified bar chart is shown in Figure 3-16. The big rectangles at lower level represent the sustainability assessment results for each dimension of sustainability (i.e., dimensional CSI). The width and height of each rectangle represent weight and the normalized value of each dimension when calculating the overall CSI respectively. The color of each rectangle indicates its sustainability dimension. The rectangles (bars) at the upper level represent the individual sustainability indicators in each dimension of sustainability. The width and height of each bar represent the final weight and normalized value of each indicator when calculating the overall sustainability respectively. The star sign above each bar indicates the corresponding indicator is important to the sustainability assessment result, based on sensitivity analysis results. The color of each bar also indicates its sustainability dimension.

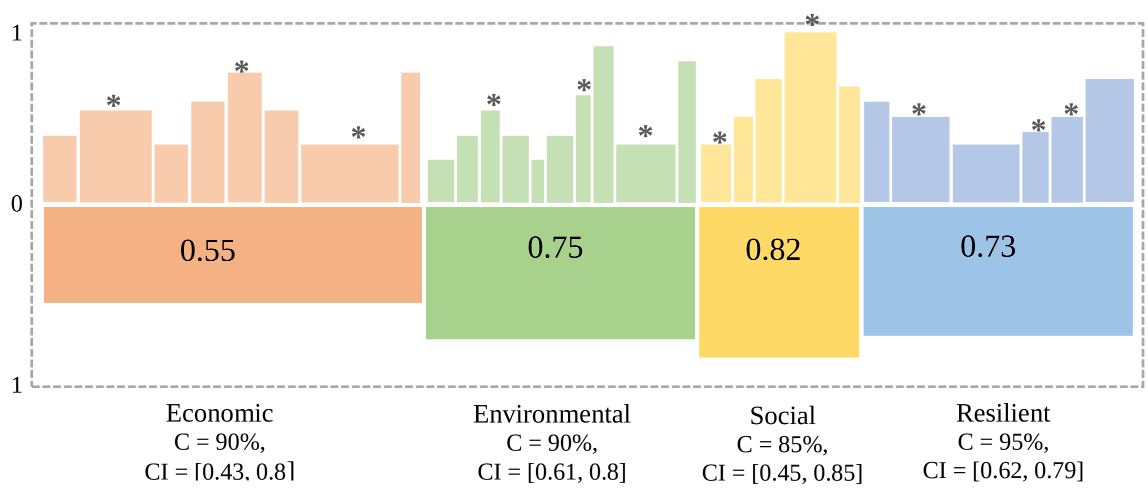


Figure 3-16 Stratified Bar Chart of Sustainability Assessment Results

The stratified bar chart can represent many indicators. However, as the limit of screens or pages, if there are too many indicators, such as more than 50, then the width of many bars could become too small to be recognized. Thus, users can create one stratified bar chart for each dimension of sustainability.

3.7.1.4 Line Chart

Line chart uses lines to represent sustainability assessment results. Length of the line represents the value of indicators, assessment results, the weight of indicator, or sensitivity results of the indicator. The color of the line indicates the dimension of sustainability. In this framework, multiple line charts can be created after finishing sustainability assessment, including the chart of sustainability assessment results, weight and value of indicators and sensitivity of indicators, examples see Figure 3-17 and Figure 3-18. In Figure 3-17, length of the line represents the dimensional CSI value, its color indicates the dimension of sustainability, and the box on each line represents the confidence interval of CSI value in the corresponding dimension. In Figure 3-18 (a), the length of the line represents value and weight of indicators, stars besides line indicate the corresponding indicators are important to sustainability assessment results, according to sensitivity results shown in Figure 3-18 (b).

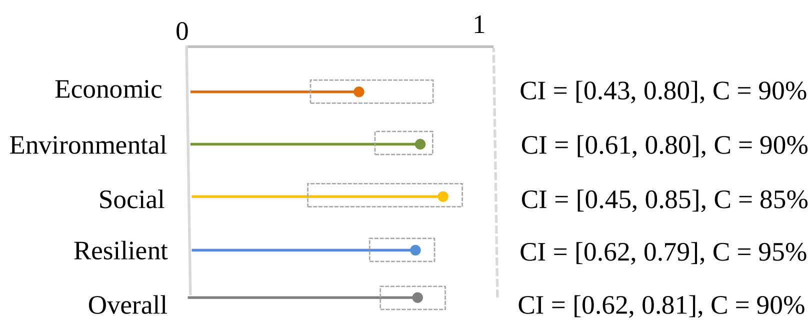
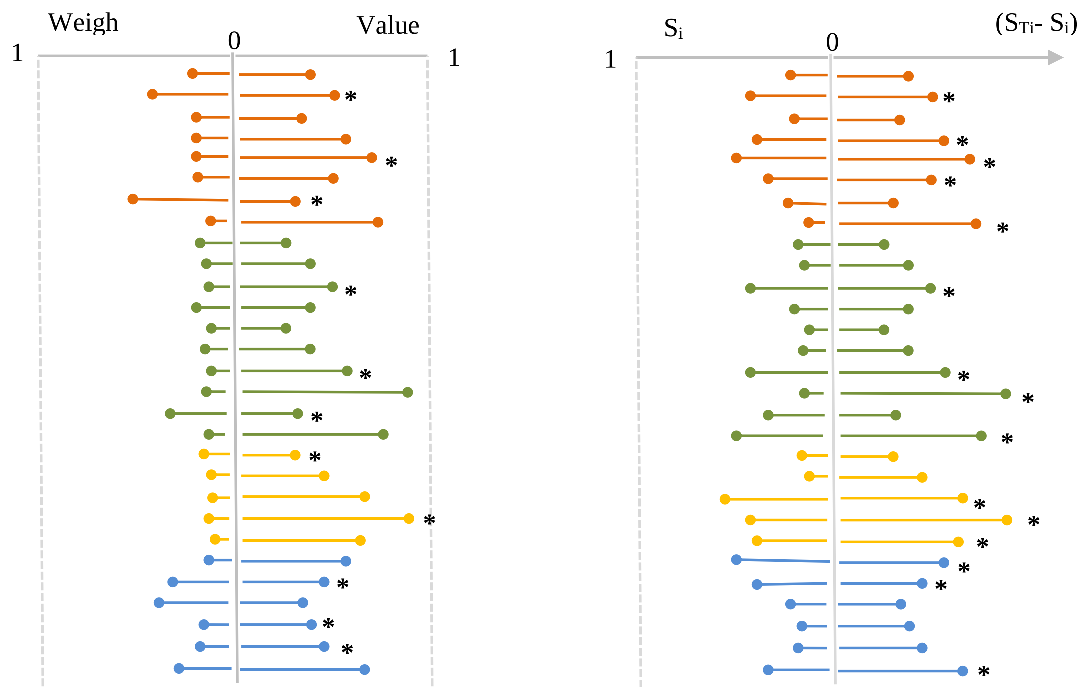


Figure 3-17 Line Chart of Sustainability Assessment Results



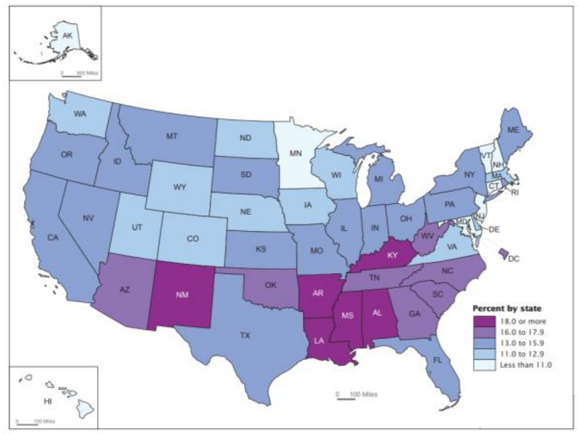
(a) Weights & Values of Indicators (b) Sensitivities of Indicators

Figure 3-18 Weights, Values, and Sensitivities of Indicators

3.7.1.5 Tile Grid Maps

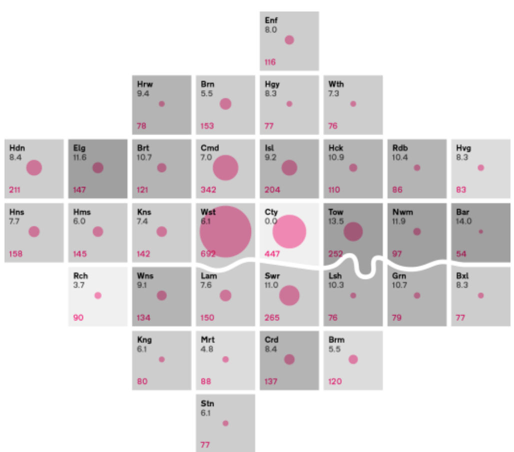
The most conventional way to visually encode spatial information is by choropleth geographical map (Figure 3-19). However, on the conventional choropleth geographical map, all the shapes are irregular with different size, sizes of shape areas are non-uniform, and shapes with large areas often receive greater visual emphasis than small areas (Tufte, 1990). Sometimes, the very small areas tend to be ignored, even they have very important attribute values. For example, dark color in a small shape could get lost amid lighter colors for a bigger shape. Besides, the choropleth map could distort the viewers' perceptions of the number of shape in different categories. However, the tile grid map does not have these problems, since all shapes on the map are in the same size and shape, i.e., each shape is visualized or treated equally. The tile grid map can include different sorts of data displays, making the map a nice small multiple visualizations (Shaw, 2016). The tile grid maps of the classic geographical map in Figure 3-19 are shown in Figure 3-20.

(a) London City (Shaw, 2016)

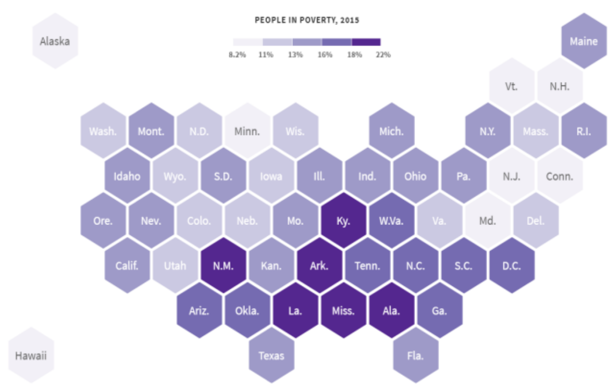


(b) People in Poverty of US States (Bishaw and Glassman, 2016)

Figure 3-19 Classic Geographical Maps



(a) London City (Shaw, 2016)



(b) People in Poverty of US States (Stiles, 2016)

Figure 3-20 Tile Grid Maps

In the proposed spatial sustainability assessment framework, each tile grid represents one spatial unit, such as census tract, city, and state, based on spatial scales of sustainability assessment. For example, if one user conducts sustainability assessment on the spatial scale of census tract, then the sustainability of each census tract will be evaluated, and each tile grid represents one census tract in the study area. To create tile grid map visualization, we need to transfer classic geographical maps to tile grids. There are many algorithms to transfer classic maps automatically (Eppstein et al., 2013). The method proposed by Eppstein et al. (2013) is used in this proposed sustainability assessment framework.

3.7.1.6 Assessment Report

The sustainability assessment report is a summary of the assessment in the proposed framework. It includes the following information:

- The objective of the assessment, and scales of sustainability assessment (e.g., spatial scales and temporal scales).
- List of indicators used for sustainability assessment and properties of these indicators, including value range of each indicator, their correlations analysis result (e.g., their data quality, the correlation matrix of indicators like the one shows in Figure 3-10), and the corresponding correlation analysis methods (e.g., correlation coefficients meausres).
- Data used for sustainability assessment and its qualities.
- A flowchart of sustainability assessment that shows the whole process of sustainability assessment.
- Sustainability assessment results, their uncertainty and sensitivity, visualizations charts or graphs, and their interpretations.
- Suggestions or recommendations are given to the stakeholder or decision-makers.

3.8 Implementation

The framework is implemented as a Plugin in QGIS. It can be enabled or disabled by enabling or disabling the respective plugin using the QGIS plugin manager. QGIS (formerly known as Quantum GIS) is one of the most popular open source GIS software with a growing user base and increasing importance in the education sector (Graser and Olaya, 2015). It can be used for spatial data creation, editing, analysis, and mapping. QGIS provides a Python Application Program Interface (API), which is used to expand its functionality. When implementing the proposed spatial sustainability assessment framework, I follow the following principles：

Easy to use: I try to make the Graphical User Interface (GUI) as simple as possible and let the users easily understand the principles or methods used in each function in the proposed framework. The proposed framework is implemented as a Plugin in QGIS, thus for people who have used GIS tools, the tool is easy to use. Besides, the design of the framework makes it easy to interpret the results of the sustainability assessment results.

General: The proposed spatial sustainability assessment framework is a general framework that can be used by users from almost all disciplines. When implementing the functions in the framework, the author tries to give users much flexibility to control the parameters of each function and the format of input/output data. Therefore, users can customize their sustainability assessment and analysis.

Adaptable and Reusable: The framework provides users space to expand or add new functions to the framework easily. Users can also reuse the implementation code for this proposed framework.

Accessibility: The author would like to share the code and let more people access to the proposed framework. Therefore, other researchers or users of this plugin can study the sustainability assessment analysis process and reproduce published results in a much more straight-forward fashion. In other words, they do not need to reproduce the individual steps based on a textual description or to implement the analysis from pseudocode (Graser and Olaya, 2015). The author also provides users as many information as possible, such as comments, tool tips, the tool helps, and other related information (e.g., reference paper of algorithms and source code referred when implementing the framework).

Homogeneous: Implementations of all functions of the framework are homogeneous. The author tries to make the implementation style, GUI behavior (e.g., the order of entries in layer selectors, the location of help buttons or dialog, and so on), and analysis tools consistent.

3.8.1 *Programing Languages and Setup*

3.8.1.1 Programming Languages

The author follows the QGIS Processing framework architecture to implement the proposed framework. QGIS processing is an open-source object-oriented Python framework, which provides seamless integration of geo-processing tools from a variety of different software libraries (Graser and Olaya, 2015). It provides a platform for the development of analysis algorithms that makes it easy to implement and use these algorithms. QGIS Processing is written in Python and connects to the QGIS API. It also connects to external applications, such as SAGA, GRASS GIS, and R Toolbox binaries, which make it more efficient to integrate the data analysis functions in these external applications to QGIS Processing. Besides, QGIS Processing is very flexible, generates GUI automatically, and provides consistent behavior across different tools (Graser and Olaya, 2015). Furthermore, all the functions implemented using the QGIS Processing framework can be directly used in the QGIS Processing Graphical Molder.

The author mainly used Python to implement the data analysis functions and modules in the proposed spatial sustainability assessment framework. The data visualization and some statistical analysis functions are implemented in Python and R. The spatial databased can be built by the open source PostgreSQL or SpatiaLite. All the processing functions related to the database are written in SQL.

3.8.1.2 Setup of the Proposed Framework

Before using the implemented framework, users need to install the software, packages or libraries that support the framework (see Table 3-6). After installing all the software and packages listed in Table 3-6, users can copy the code folder ("SustainAssess") into the following place: "..\.qgis2\python\plugins". Then users open QGIS, under the menu "Plugins", click "Manage and Install Plugins …", the following window will open. In the "Search" textbox, type in "spatial sustainability assessment", choose it and activate it (see Figure 3-21). Close the plugin management window, back to the QGIS main view, you can see a menu named "Sustainability" (see Figure 3-22). This is the implemented framework. Under the "Sustainability" menu, there are many tools and a "wizard". The following sections will illustrate each tool and the wizard in details.

Table 3-6 Software and Packages Support the Proposed Framework

Software	Detail Information
QGIS	Version 2.18.21
Python	Version 2.7
Python Packages	1) qgis Python API 2) numpy+mkl 3) scipy 4) pandas 5) sciki-learn: version 0.19.0 6) networkx 7) igraph 8) rpy2: version 2.7.8 9) psycopg2: version 2.7.3.2 10) pyodbc: version 3.0.7 11) pyQT 4
R	version 3.5.2
R packages	1) ggplot2 2) corrplot 3) spdep, sp
PostgreSQL	Need to be installed if users want to access PostGIS database from the proposed framework.
SpatiaLite	Need to be installed if users want to access SpatiaLite database from the proposed framework.

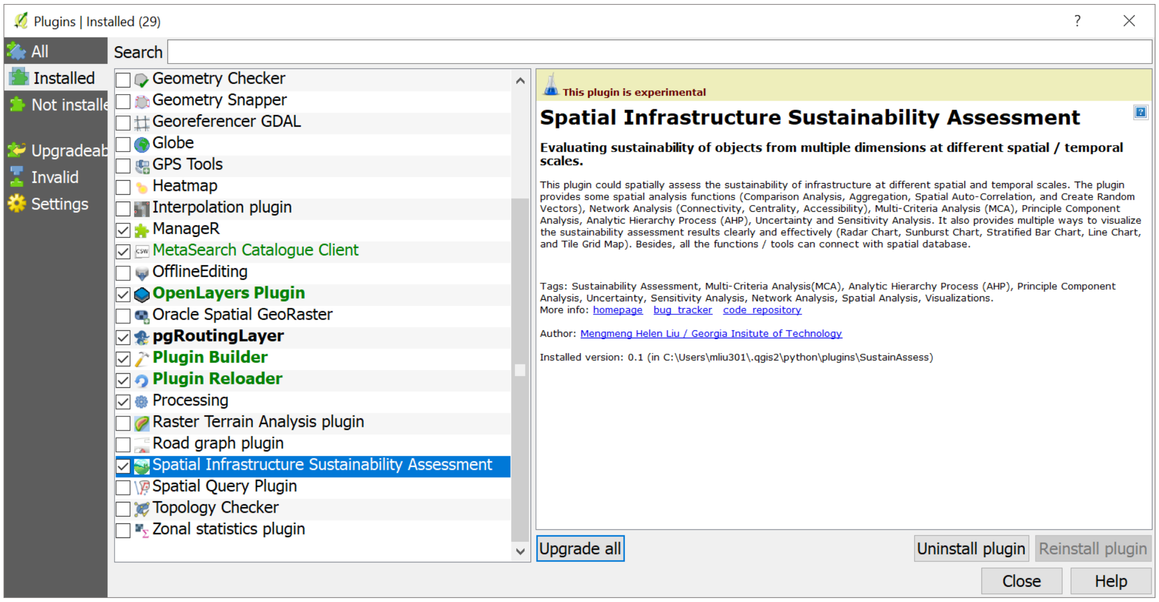


Figure 3-21 Activate “Spatial Infrastructure Sustainability Assessment” Plugin in QGIS

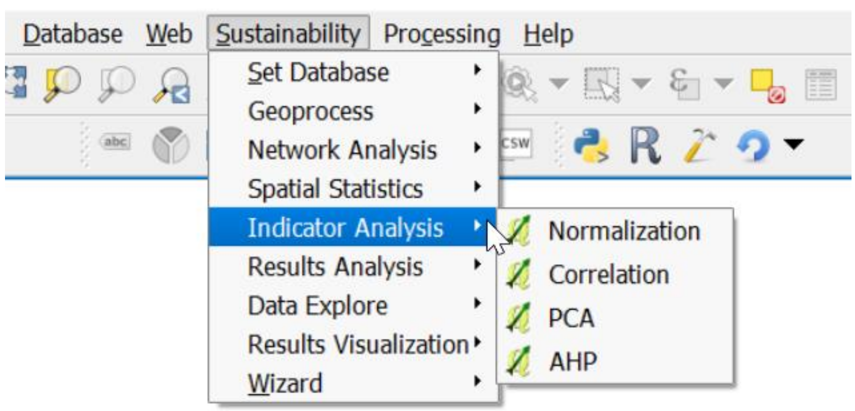


Figure 3-22 Sustainability Menu in QGIS

3.8.2 Define Object and Scales

In the “Define Object & Scale” component in the “Wizard” of the implemented “Spatial Infrastructure Sustainability Assessment” plugin. Users can define both the objective and scales (spatial and temporal scales) for the sustainability assessment. Its GUI is shown in Figure 3-23.

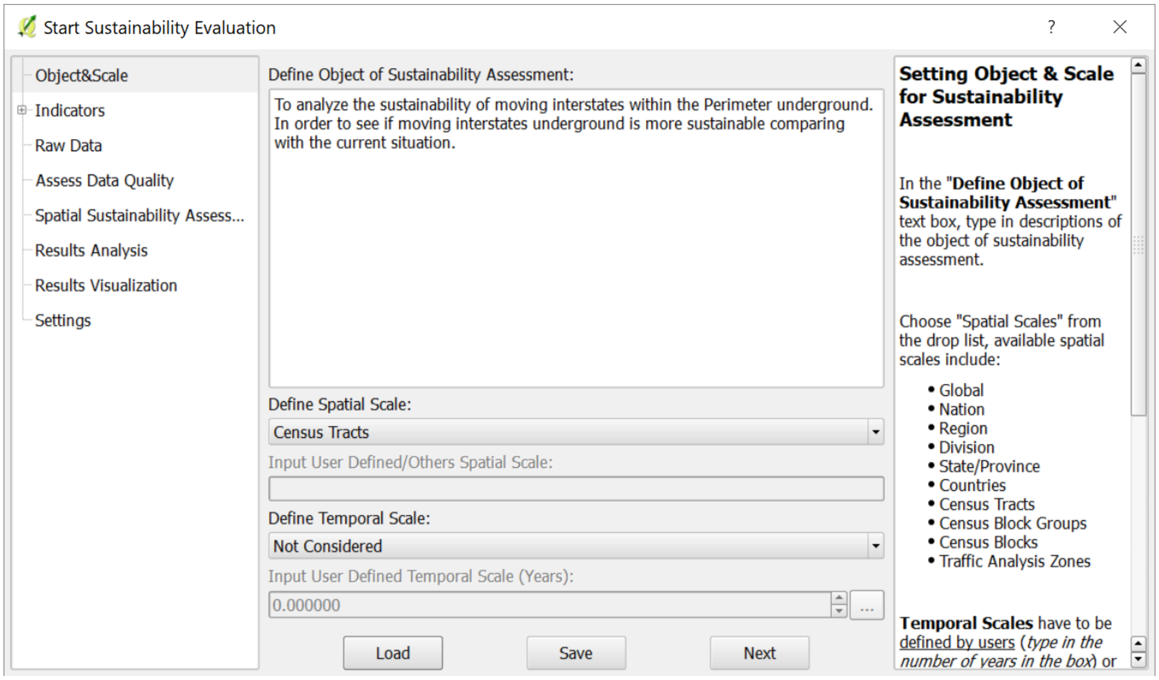


Figure 3-23 GUI of Define Object and Scales in Wizard

Users can type in the objective of the sustainability assessment in the text box. The spatial scales of the assessment can be selected from the drop list. The default available spatial scales include Global, Nation, Region, Division, State/Province, Countries, Census Tracts, Census Block Groups, Census Blocks, and Traffic Analysis Zones. Users can also define their scales if they cannot find appropriate spatial scales from the drop list. The system will add the scales defined by users to the drop list automatically so that other users can refer it or the users can select the scales from the drop list without defining new spatial scales. Users have to define their temporal scales by typing in the number of years in the box. All the information provided by users can be saved to file. Users can load the saved file to the system if they cannot finish the whole sustainability assessment process in one time. Users can see the help of this component on the right. After setting the objective and scales of the sustainability assessment, users can click the “Next” button to switch to the window of “Select Indicators”.

3.8.3 Select Indicators

Users need to select or define indicators in economic, environmental, social, and resilient dimension separately. If one indicator includes information about multiple sustainability dimensions, it can be put in the “Other Dimension”. When selecting indicators for each sustainability dimension, users can select indicators from the default indicators file, from other csv files, from the database (PostGIS table), or add new indicator manually (see Figure 3-24). When manually input the user-define indicator, users need type in the following information: “categories”, “sub-category”, “indicators” name, and notes for the indicator (see Figure 3-25). The id, dimension, spatial scale and temporal scales of the indicator are automatically filled based on the information provided in “Define Object and Scales” component in the previous step (see Figure 3-23). Users can click the “Save” button to save their selected indicators into a file, and click the “Next” button to switch to the next window to select “Environmental Indicators”. After selecting indicators for all the dimensions of sustainability, users can switch to the next component, which is “Raw Data”. Users can also see the help of this component on the right of the GUI.

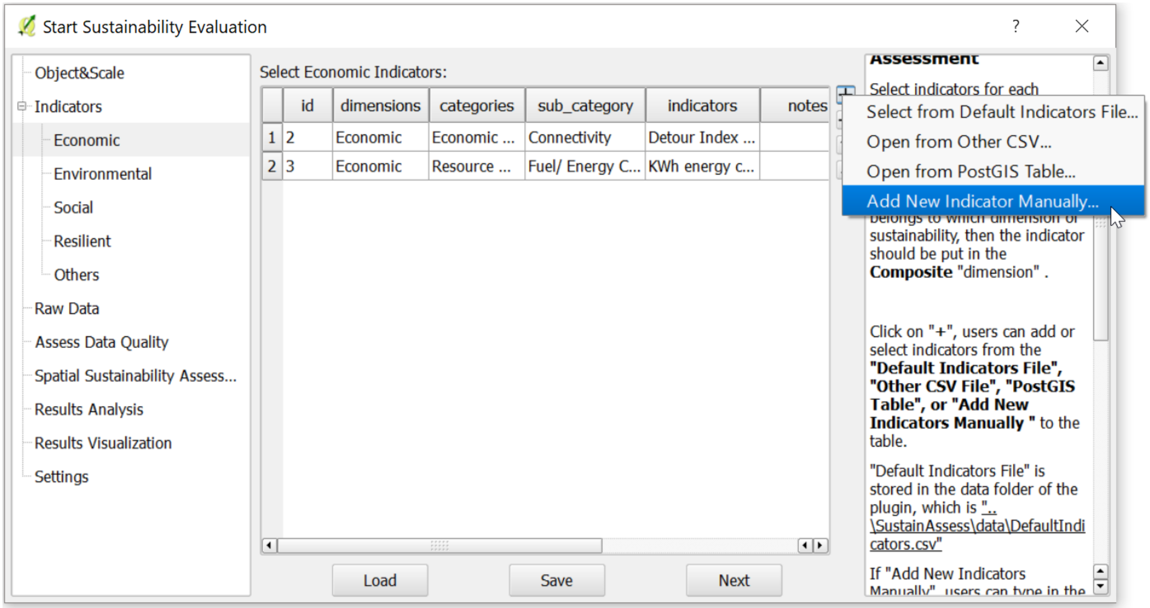


Figure 3-24 GUI of Select Economic Indicators in Wizard

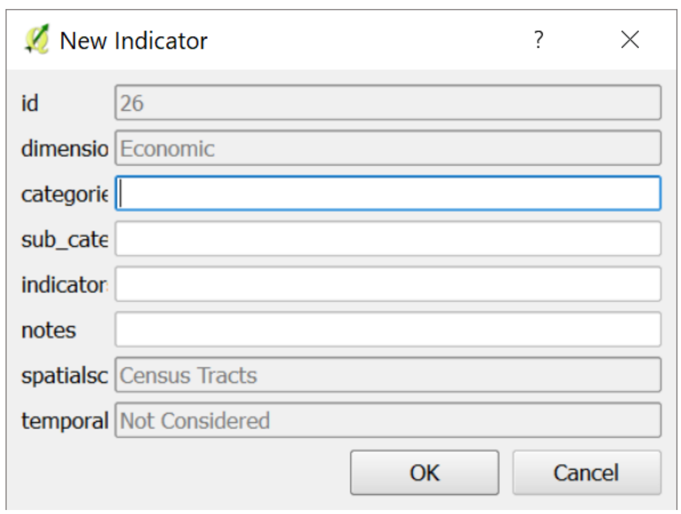


Figure 3-25 GUI of Add New Indicator Manually

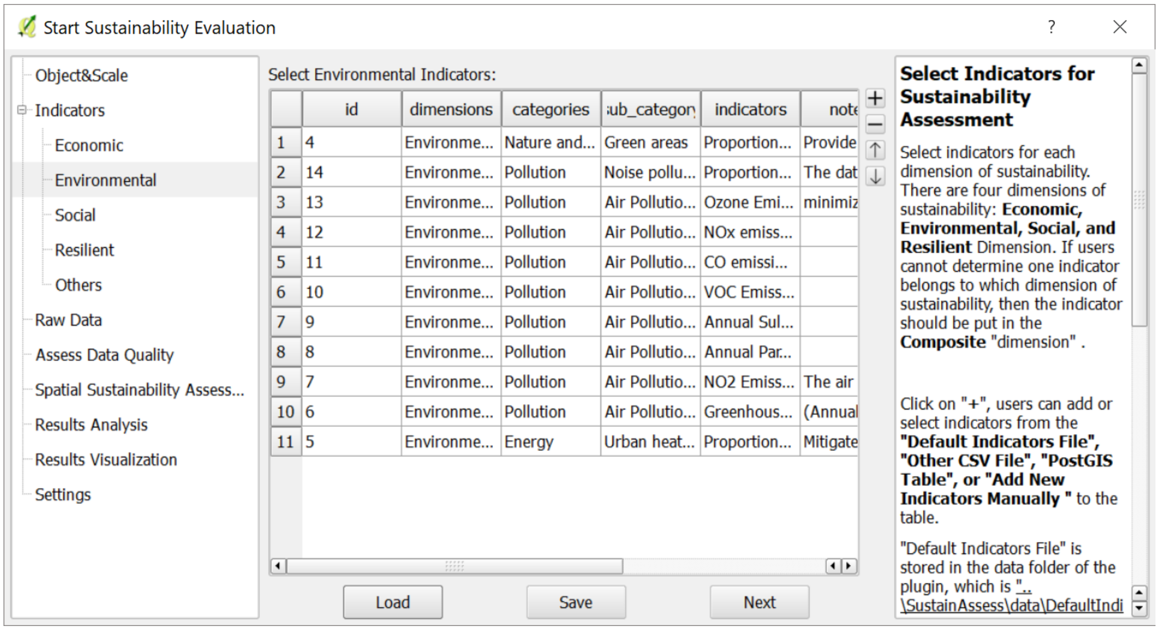


Figure 3-26 GUI of Select Environmental Indicators in Wizard

3.8.4 Select Raw Data

After finishing selecting indicators, users can start to work with the data used for the following sustainability assessment. Users can click the "+" button to select data from current layers opened in QGIS, from files, or PostgreSQL database. This component can automatically extract the filename, file type, CRS, and file path in the raw data table. In the table, users can see if all the data have the same CRS (see Figure 3-27) and their data types. Users can also easily remove data from the data by clicking Click "-" button, or change the order of data in the table by clicking the button"↑" or "↓".Users can click the "Save" button to save the data information in the table into a file. The help of this component is on the right of GUI.

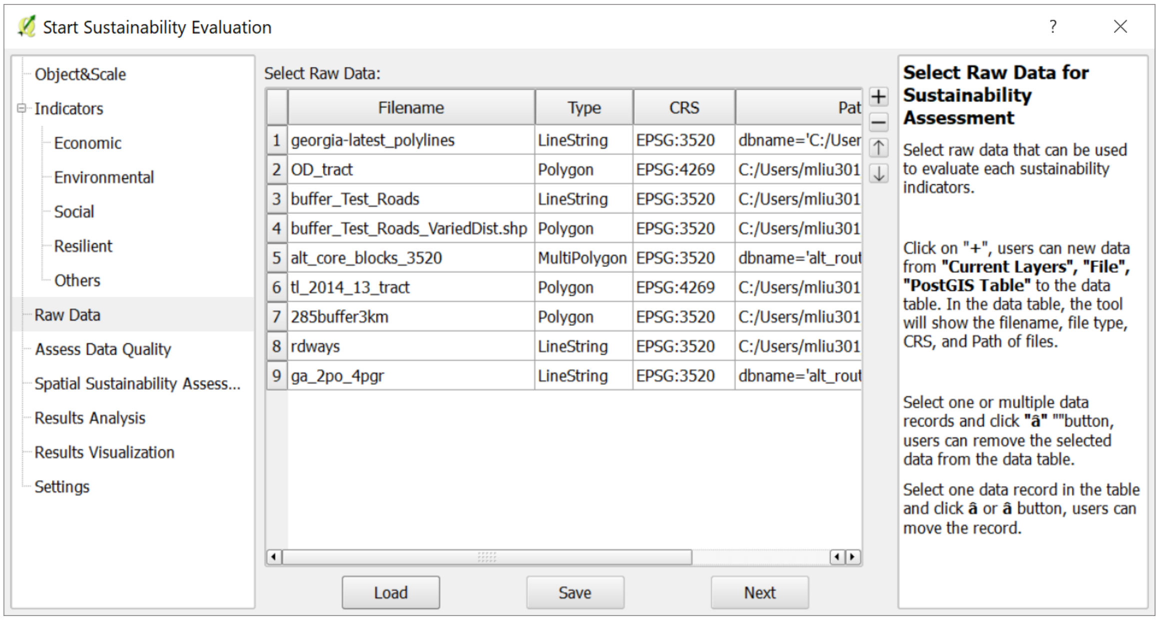


Figure 3-27 GUI of Select Raw Data in Wizard

3.8.5 *Data Quality Assessment*

Data quality assessment is a very important part of the framework. In section 3.4, the author discussed data cleaning methods in details. Users can open each data file in QGIS and clean the data by corresponding functions in QGIS. Data cleaning in the framework includes removing outliers, interpolating missing values, eliminating duplicate records, and removing inconsistencies in datasets. Due to the time limit, users can only manually determine the data quality by filling in the “Manually Data Quality Check Table” (see Figure 3-28). Users can click the “Save” button to save the “data quality check table” into a file. The help of this component is on the right of GUI. If users think the data quality is not good enough, they can go back to “Select Raw Data” component to select new data or remove the data with bad quality. Otherwise, they can click the “Next” button to start the “Spatial Infrastructure Sustainability Assessment” (see Figure 3-29).

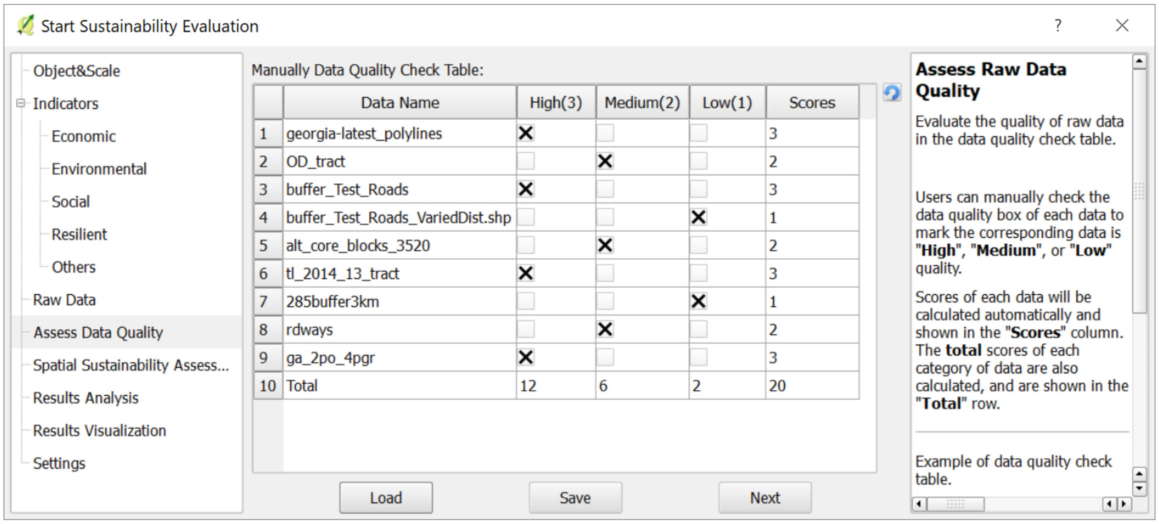


Figure 3-28 GUI of Data Quality Assessment in Wizard

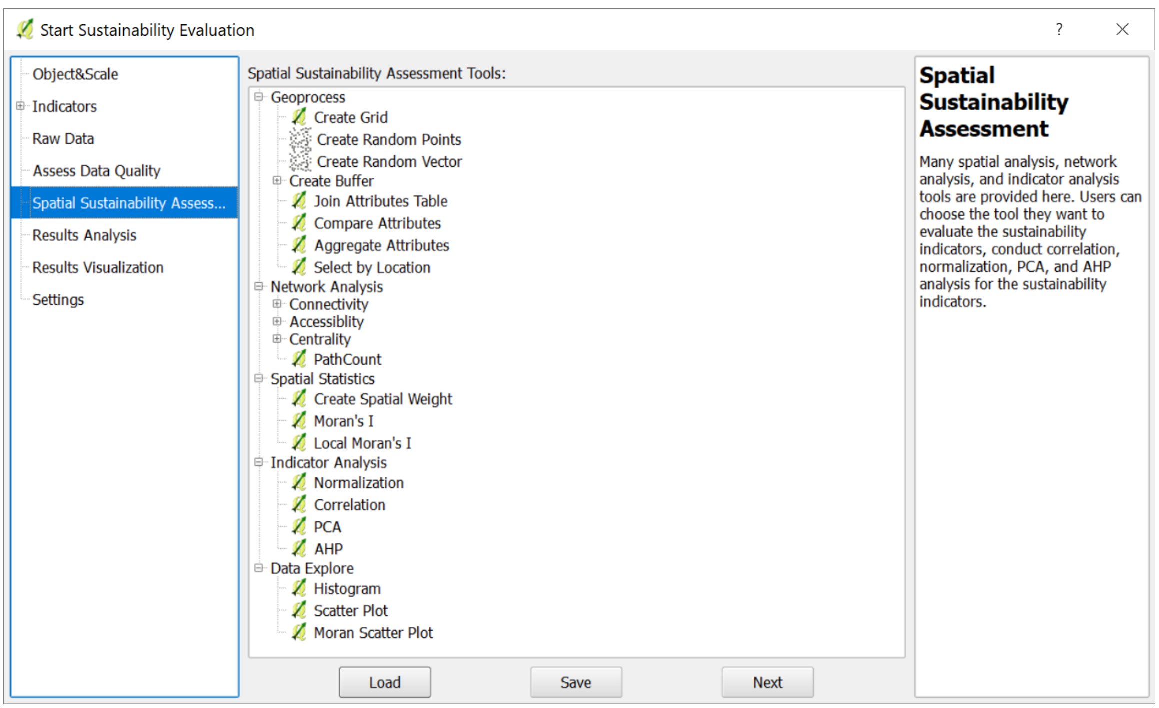


Figure 3-29 GUI of Spatial Sustainability Assessment in Wizard

3.8.6 Spatial Analysis

The spatial analysis tools implemented in the proposed framework include “Create Grid”, “Create Random Points”, “Create Random Vectors”, “Create Fixed/Varied Distance Buffer”, “Join Attribute Table”, “Compare Multiple Attributes”, “Select Features by Location”, “Spatial Statistical Analysis (Moran’s I, Local Moran’s I)”, and “Aggregate Attributes” (e.g. aggregate attributes of smaller-scale polygons to bigger-scale polygons, aggregate attributes of points to polygons). All the implemented tools support both local shapefiles and database tables (PostGIS and SpatiaLite).

3.8.6.1 Buffer Analysis

Two buffer analysis tools are provided in the plugin, they are “Fixed Distance Buffer” tool and “Variable Distance Buffer” tool. The “Fixed Distance Buffer” tool lets users create a fixed distance buffer. The “Variable Distance Buffer” tool let the buffer distance vary according to numerical values of one attribute of the vector layer. GUI and parameters of “Fixed Distance Buffer” and “Variable Distance Buffer” are shown in Figure 3-30 and Figure 3-31 respectively. The buffer distance is in map units of the “Input Layer”. The created buffer layer can be saved to file (*.shp) or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of the GUI or show helps on the web browser by clicking the “Tool Help” button. Figure 3-32 shows two example buffers created by these two buffer analysis tools.

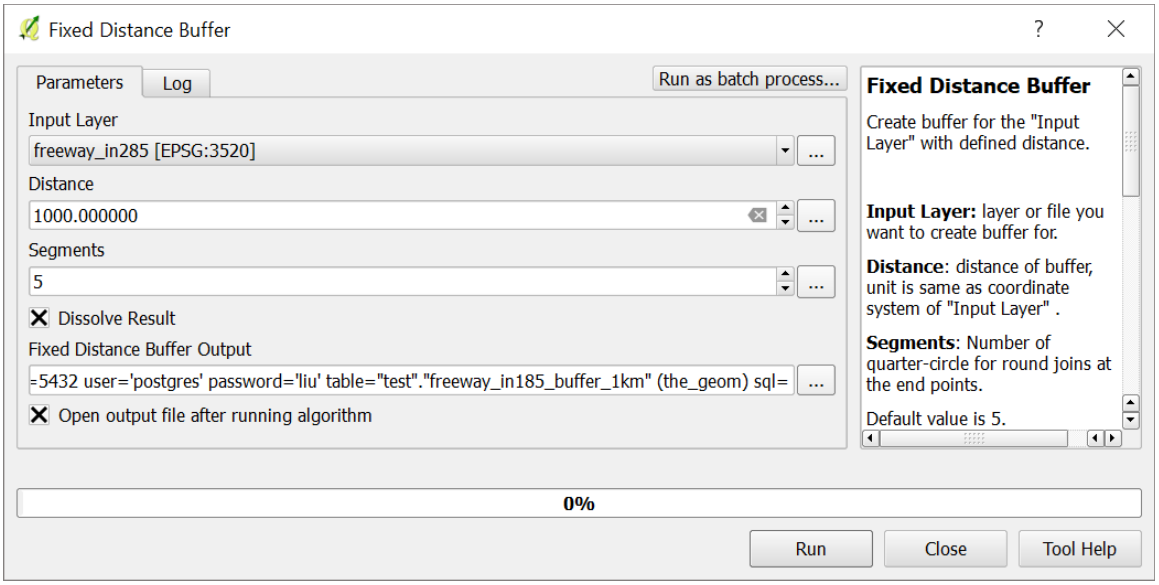


Figure 3-30 GUI and Parameters of Fixed Distance Buffer Tool

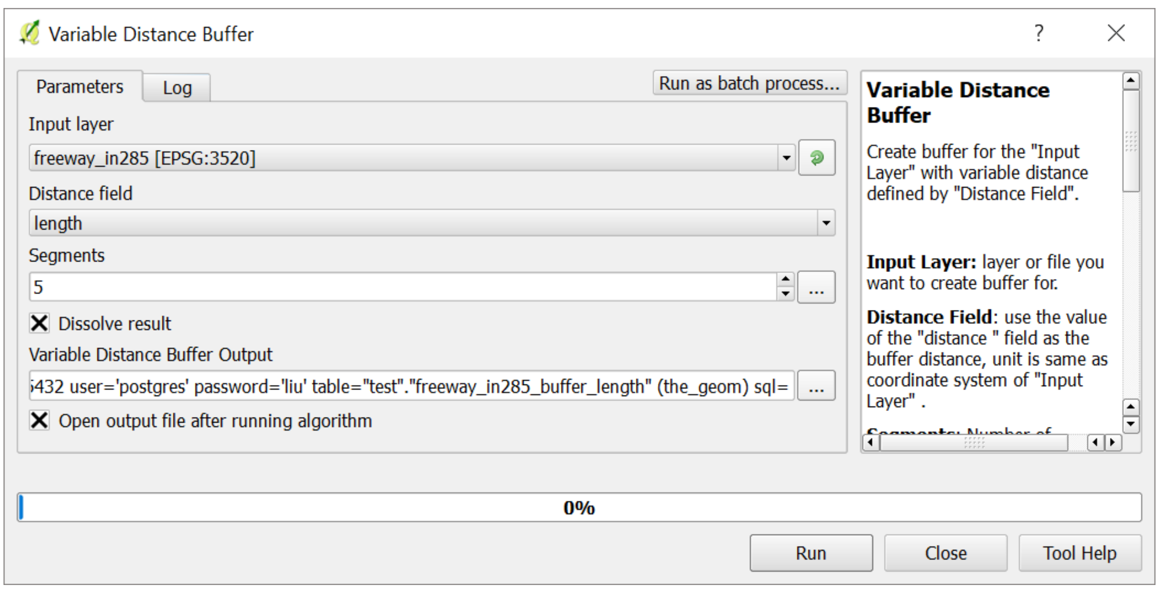


Figure 3-31 GUI and Parameters of Variable Distance Buffer Tool

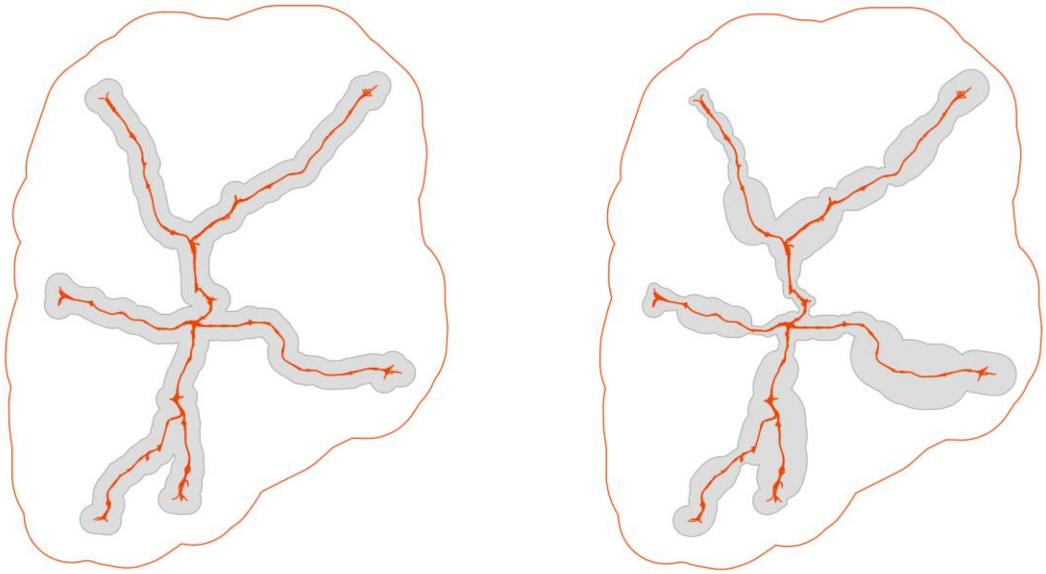


Figure 3-32 Fixed Distance Buffer (Left) and Variable Distance Buffer (Right) using Parameters in Figure 3-30 and Figure 3-31

3.8.6.2 Spatial Auto-Correlation Analysis

There are two spatial autocorrelation measures defined in the "Spatial Sustainabiltiy Assessment" plugin, they are Moran's I and local Moran's I, which measures global and local spatial autocorrelations respectively. Spatial autocorrelation analysis requires some measure of contiguity or neighborhood relation to get the spatial weights between observations. "Create Spatial Weight" tool can calculate the spatial weights between observations. GUI and parameters of "Create Spatial Weight" tool are shown in Figure 3-33.

Spatial weight between two neighbors can be binary (1, 0) or be the inverse of the distance between these two neighbors ($weight = 1/dis^{decay}$). Thus, nearby neighbors getting larger weights than neighbors farther away if the spatial weight is "inverse-distance". The neighborhood relationships and the generated spatial weights between neighbors can be saved to file. Two files with the same name will be generated, which are "*.gal" (neighborhood relationships) and "*.gal.wgt" (spatial weights between neighbors). Users can see the help of this tool on the right side or by clicking the "Tool Help" button to open help file on web browsers.

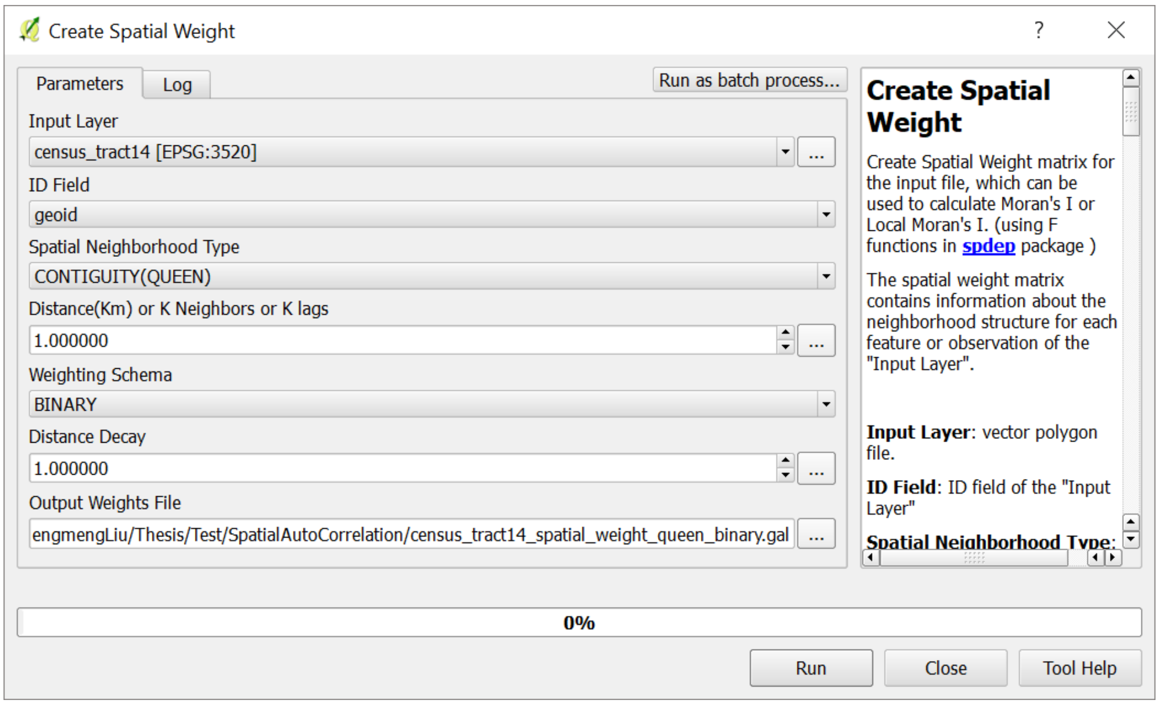


Figure 3-33 GUI and Parameters of Create Spatial Weight Tool

"Create Spatial Weight" tool supports four types of spatial neighborhood, which are "Queen's Contiguity", "Rook's Contiguity", "Fixed Distance", and "K-Nearest Neighbors".

- Queen's Contiguity: neighboring polygons are those that share a vertex with the focal polygon. This tool can create higher order neighbors (k-lags neighbors).

- Rook’s Contiguity: neighboring polygons are those that share a line segment with the focal polygon. This tool can create higher order neighbors (k-lags neighbors).
- Fixed Distance: Distance-based neighbors are those within a given proximity threshold to a focal polygon (the threshold are controlled by parameter “Distance(Km)” in Figure 3-33, in units kilometers). Distances are measured between polygon centroids, in units same as CRS of “Input Layer”.
- K-Nearest: K-Nearest Neighbors are the k closest neighbors of the focal polygon. If K (the number of neighbors) is 8, then the eight closest polygons to the target (focal) polygon are its neighbors.

After getting the spatial weight, users can use the spatial weight file to calculate the Moran’s I, and Local Moran’s I. The GUI and parameters of Moran’s I tool and Local Moran’s I tool are shown in Figure 3-35 and Figure 3-36 respectively. Users can see the help of this tool on the right side of GUI or open it on the browser by clicking the “Tool Help” button. Moran’s I results are saved into a file (*.txt), which includes the input parameters’ value, Moran’s I value and its statistical significant test results (expectation, variance, and p-value). Local Moran’s I results will be saved to four new created fields in the input layer, which are “lisa_i”, “lisa_p” (p-value of), “lisa_cl”, and “lisa_lag”. “lisa_i” is the local Moran’s I value of one feature in the input layer. “lisa_p” is the local Moran’s I significant statistics p-value, “lisa_cl” is the local spatial auto-correlation cluster types (High-High, High-Low, Low-Low, Low-High, and Not Significant), and “lisa_lag” is the average local Moran’s I value of its spatial neighbors.

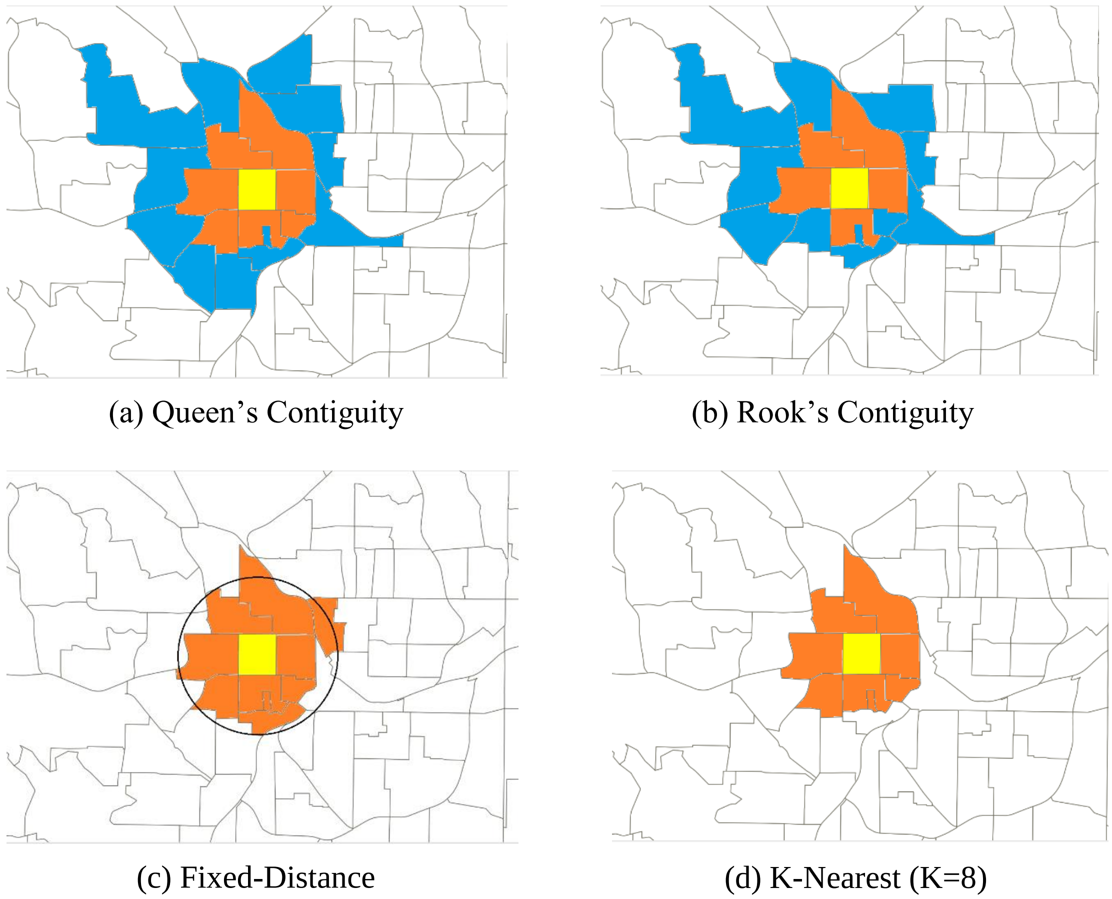


Figure 3-34 Exampls of Different Spatial Neighborhood Types

((a) 2-lags Queen’s Contiguity neighbors of the yellow polygon: both the blue and orange polygons (1-lag neighbors). (b) 2-lags Queen’s Contiguity neighbors of the yellow polygon: both the blue and orange polygons (1-lag neighbors). (c) Fixed-distance neighbors of the yellow polygon: all the orange polygons. (d) 8-Nearest neighbors of the yellow polygon: all the orange polygons.)

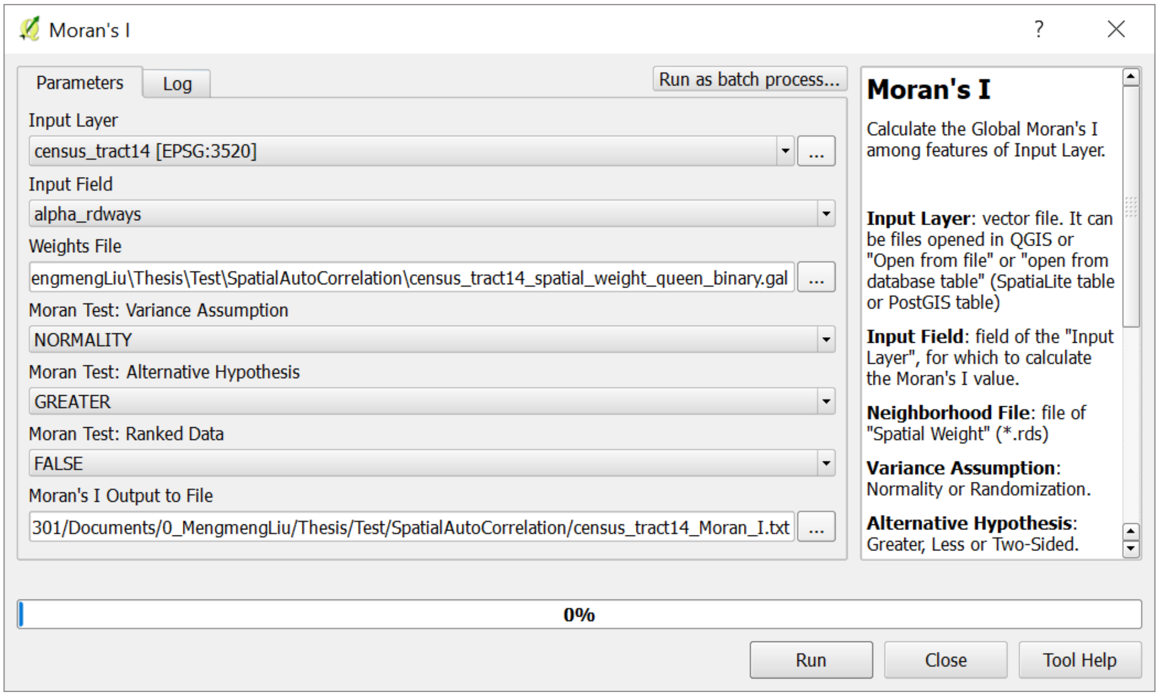


Figure 3-35 GUI and Parameters of Moran's I Tool

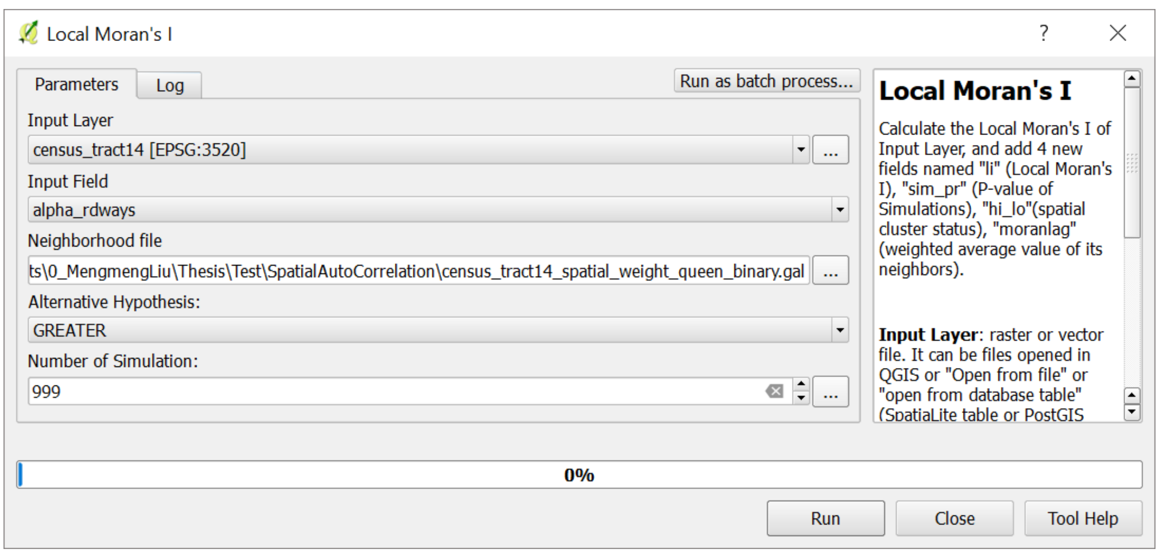


Figure 3-36 GUI and Parameters of Local Moran's I Tool

Moran's I results are saved into a file (*.txt), which includes the input parameters' value, Moran's I value and its statistical significant test results (expectation, variance, and p-value). Local Moran's I results will be saved to four new created fields in the input layer, which are "lisa_i", "lisa_p" (p-value of), "lisa_cl", and "lisa_lag". "lisa_i" is the local Moran's I value of one feature in the input layer. "lisa_p" is the local Moran's I significant statistics p-value, "lisa_cl" is the local spatial auto-correlation cluster types (High-High, High-Low, Low-Low, Low-High, and Not Significant), and "lisa_lag" is the average local Moran's I value of its spatial neighbors.

3.8.6.3 Comparison Analysis

The "Compare Attributes" tool calculates the difference of specified attribute fields between two layers. These two layers must have the same data type (e.g., both are polygon layers), and have a common key field that can connect/join the attributes of two layers together. The attribute difference value is calculated using Equation 3.37.

$$\begin{aligned} &Comparison\ result \\ &= attribute\ of\ \text{Input Layer 1}\ (operator)\ attribute\ of\ "Input\ Layer\ 2 \end{aligned} \tag{3.37}$$

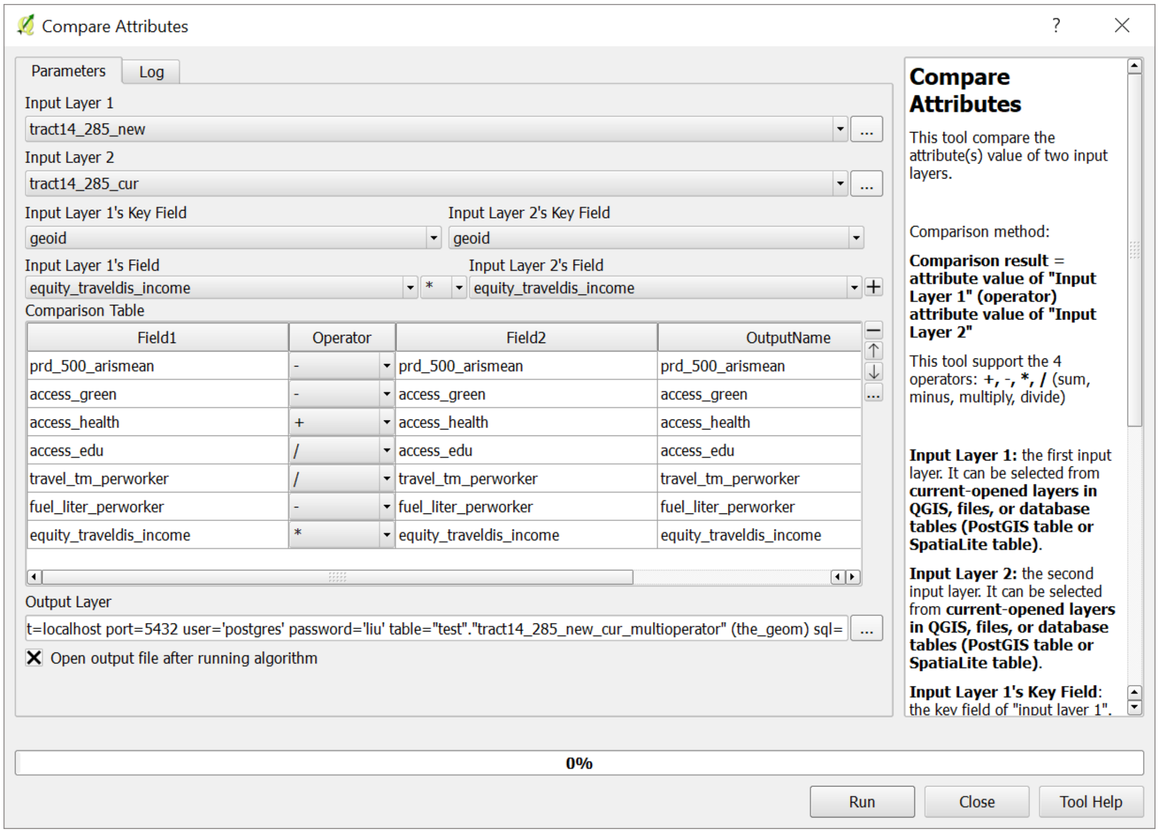


Figure 3-37 GUI and Parameters of Comparison Attributes Tool

This tool supports the four operators: +, —, *, / (sum, minus, multiply, divide). GUI and parameters of this tool are shown in Figure 3-37 below. After selecting the comparison fields for input layer 1 and 2, users can click button "+" to add the comparison fields into the "Comparison Table", and click "—" to remove the comparison fields from "Comparison Table". The comparison results can be saved to a temporary layer, to shapefile (*.shp), or database tables (SpatiaLite table or PostGIS table). The results layer can be opened automatically in QGIS after completing calculation if users check the "Open output file after running algorithm". Users can see help of this tool on the right side of the GUI or show helps on the browser by clicking the “Tool Help” button.

3.8.6.4 Random Points or Random Vectors Generator

"Create Random Points" tool generates a specified number of random points within the boundary defined by the "Input Boundary Layer". GUI and parameters of the "Create Random Points" tool are shown in Figure 3-38. By default, the generated random points are evenly distributed in the boundary. When creating random points, users can also set the minimum distance between points to avoid them too close to each other. The generated random points layer has the same CRS as the boundary layer. It can be saved as a temporary layer, shapefile (*.shp), or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of the GUI or web browser by clicking the "Tool Help" button. Random points generated using parameters defined in Figure 3-38 are shown in Figure 3-40.

"Create Random Vectors" tool generates a specified number of random vectors in specify length within the boundary defined by the "Input Boundary Layer". GUI of the "Create Random Points" tool is shown in Figure 3-39. The length of random vectors is in map units according to CRS of the boundary layer. The generated random vectors layer has the same CRS with the boundary layer. It can be saved as a temporary layer, shapefile (*.shp), or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of the GUI or web browser by clicking the "Tool Help" button. Random vectors generated using parameters defined in Figure 3-39 are shown in Figure 3-40.

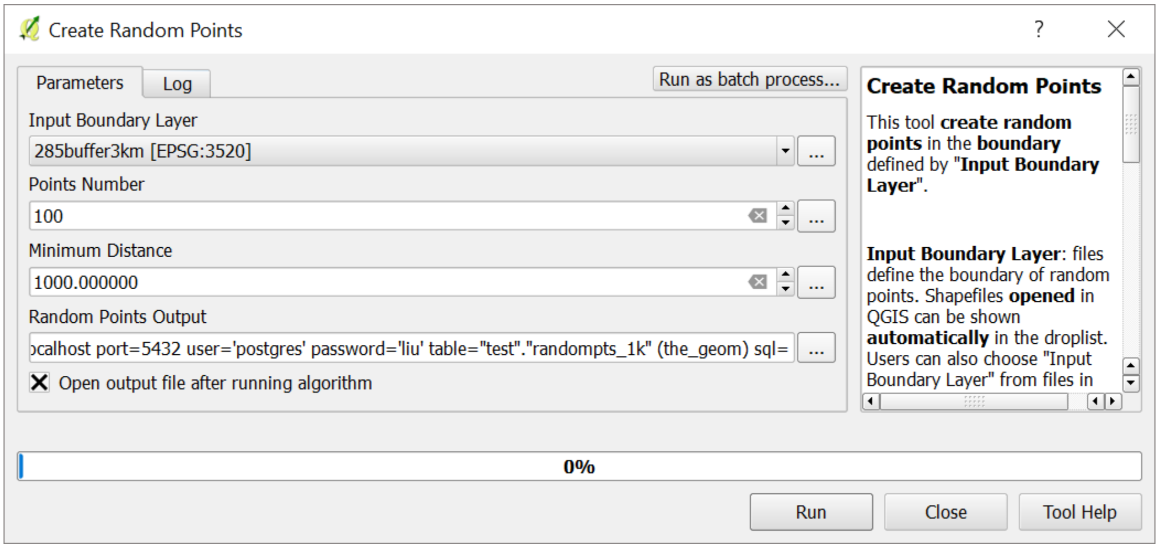


Figure 3-38 GUI and Parameters of Create Random Points Function

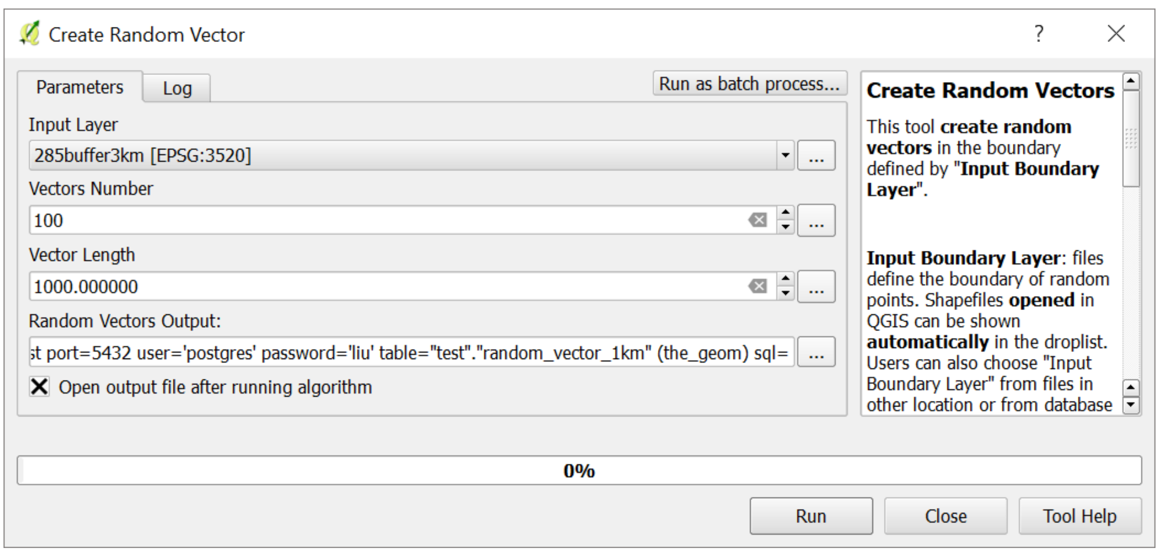


Figure 3-39 GUI and Parameters of Create Random Vectors Function

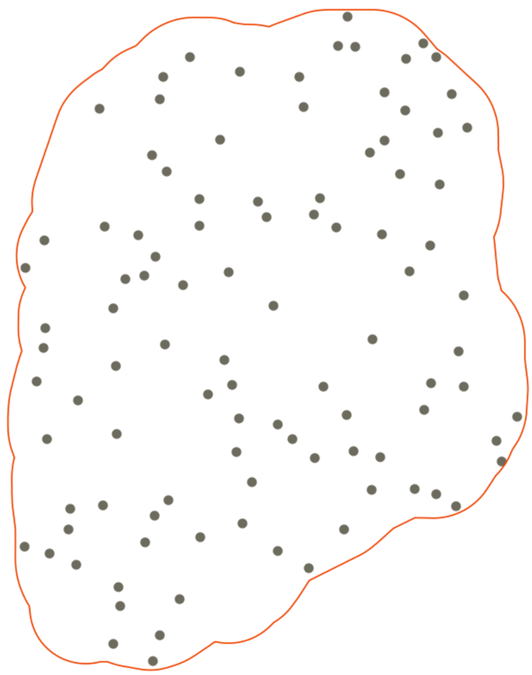


Figure 3-40 Random Points (Left) and Random Vectors (Right) Generated Using Parameters in Figure 3-38 and Figure 3-39 respectively

3.8.6.5 Create Grid

"Create Grid" function let users create three types of grid in the user-defined boundary, including rectangle grid, diamond grid, and hexagon grid (see Figure 3-41). GUI and parameters of the "Create Grid" tool are shown in Figure 3-42. Size of each grid is controlled by parameter "Horizontal Space" and "Vertical Space". If the grid type is "Hexagon", then only "Vertical Space" parameter can control the grid size. The coordinate reference system for the grid layer is the same as the layer or canvas extend which defines the extension of the grid layer. The grids can be saved as a temporary layer, shapefile (*.shp), or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of the GUI or web browser by clicking the "Tool Help" button.

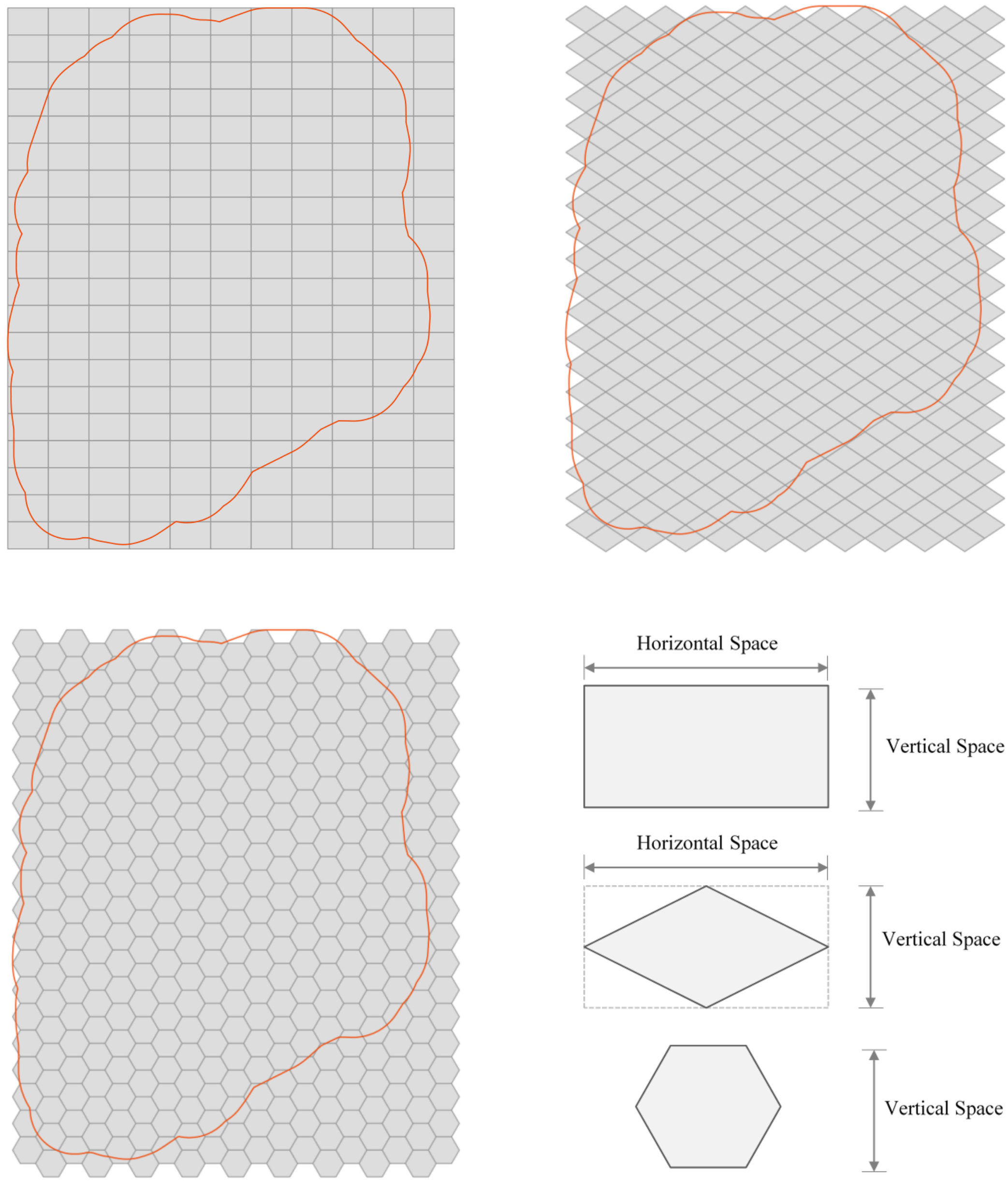


Figure 3-41 Rectangle, Diamond, and Hexagon Grid in the Extent Defined by the Orange Polygon (Using Parameters in Figure 3-42)

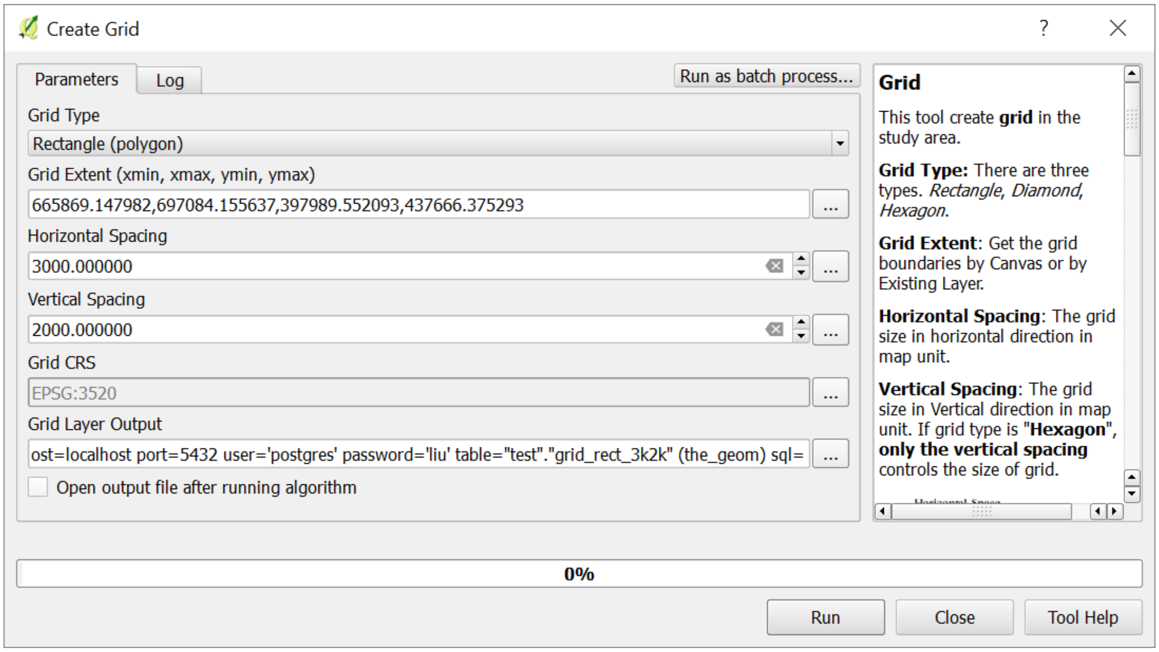


Figure 3-42 GUI and Parameters of Create Grid Function

3.8.6.6 Join Attribute Table

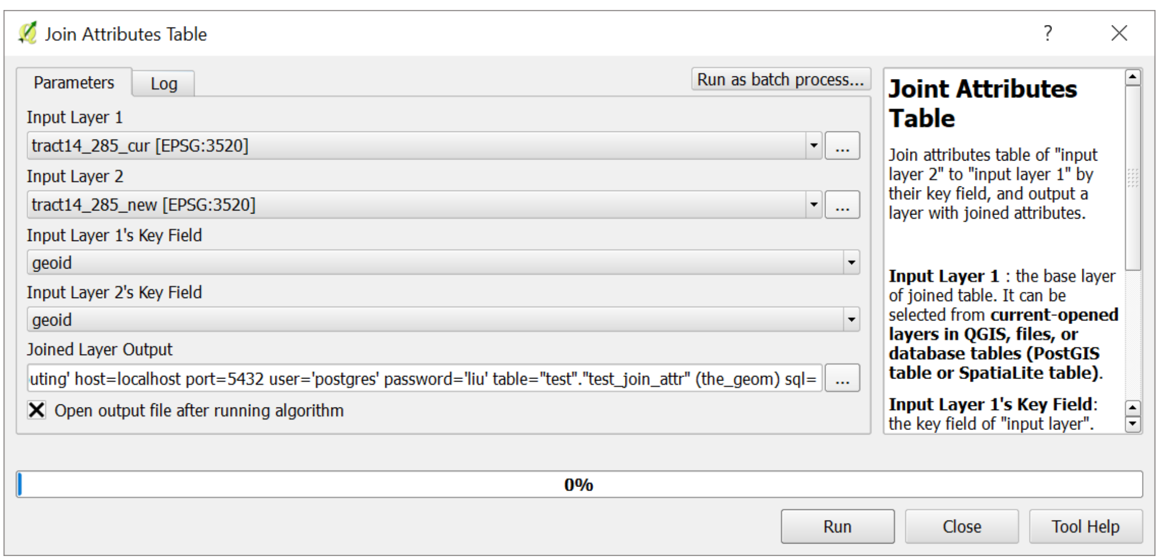


Figure 3-43 GUI and Parameters of join Attributes Table Tool

“Join Attributes Table” tool joins the attribute table of "input layer 2" to "input layer" by their key fields, and output a layer with joined attributes. GUI and parameters of this tool are shown in Figure 3-43. The output layer can be saved to temporary layer in QGIS, file (*.shp), or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of the GUI or web browser by clicking the “Tool Help” button.

3.8.6.7 Aggregate Attributes

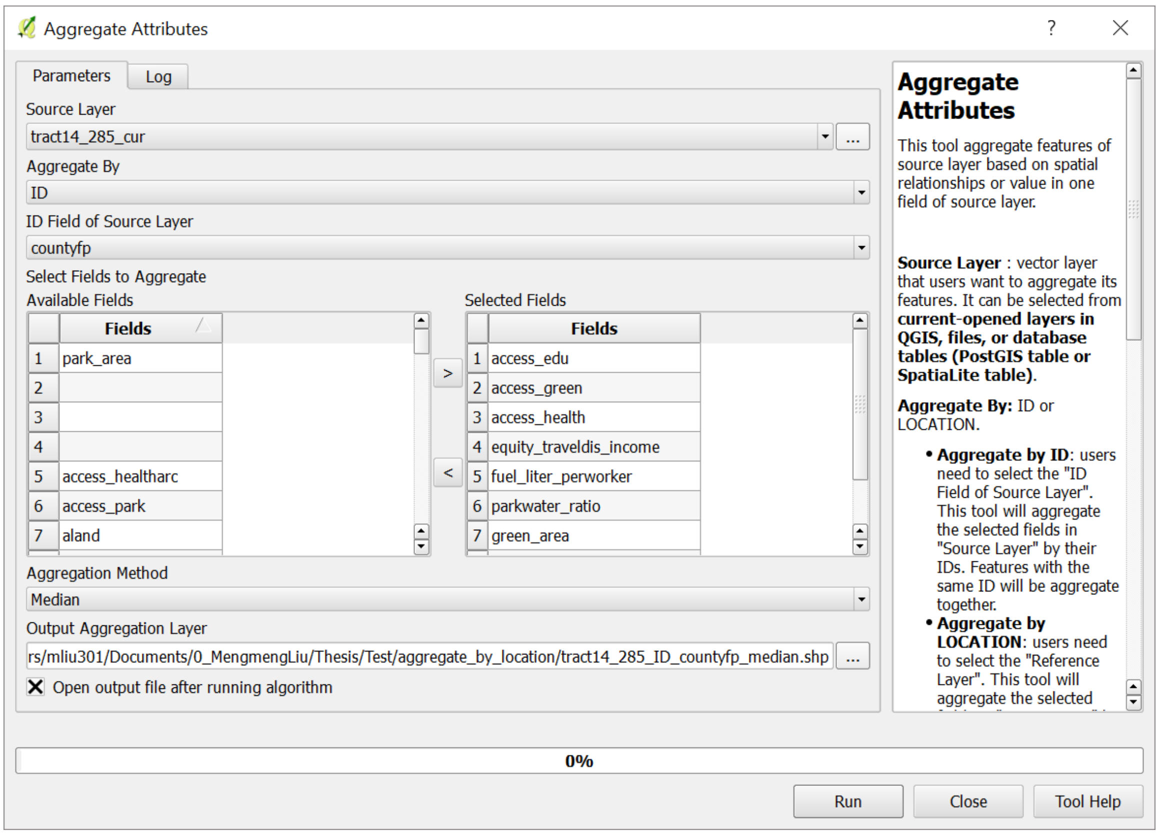


Figure 3-44 GUI and Parameters of Aggregate Attributes by ID

“Aggregate Attributes” tool can aggregate features of source layer based on the value of one field in the source layer (“aggregate by ID”) or based on its spatial relationships (intersect) to another layer (“aggregate by Location”). GUI and parameters of this tool are shown in Figure 3-44 and Figure 3-45. This tool supports seven different aggregation methods. They are Sum, Min, Max, Median, Arithmetic mean, Geometric mean, and Interquartile Mean (IQM). The aggregation results can be saved to temporary layer in QGIS, file (*.shp), or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of GUI or browser by clicking the "Tool Help" button.

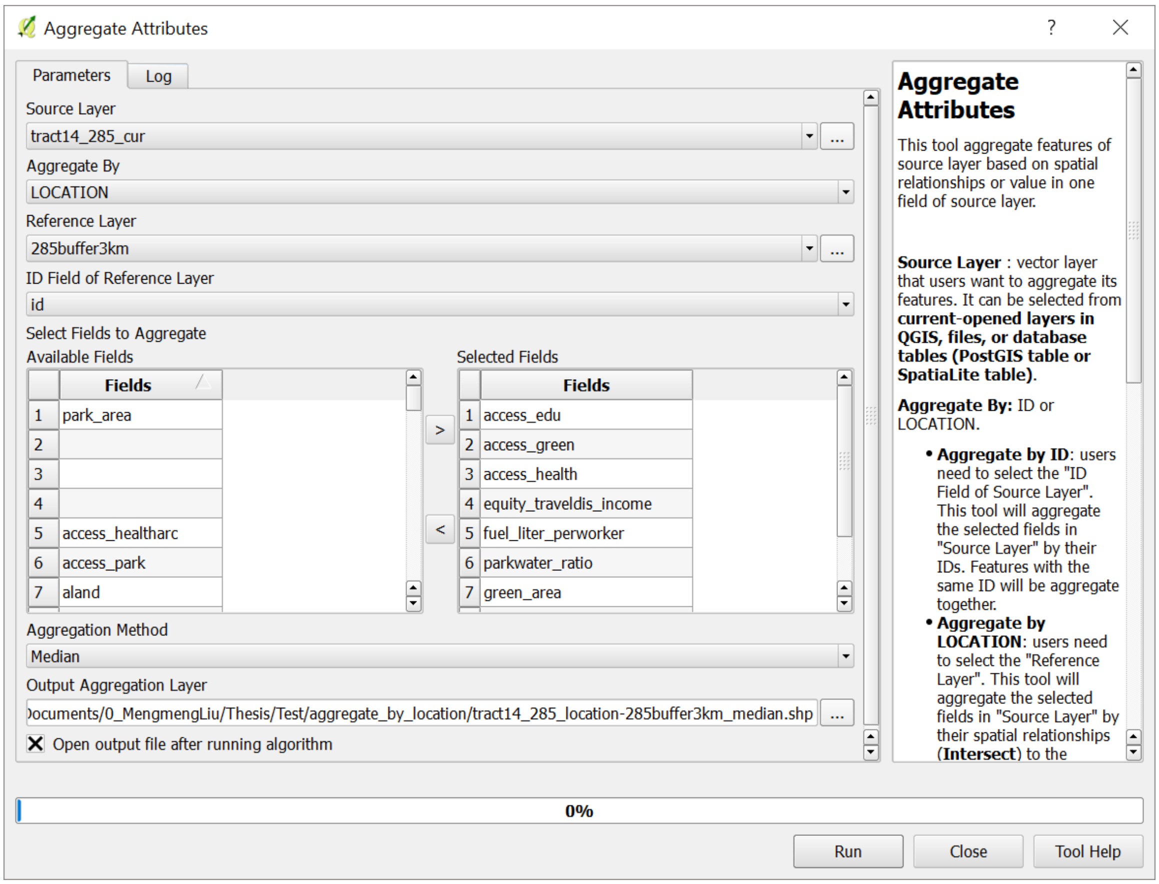


Figure 3-45 GUI and Parameters of Aggregate Attributes by Location

3.8.6.8 Select by Location

"Select by Location" tool can select features from a vector layer based on their spatial relationships to another layer. This tool supports 8 "Spatial Relations", including

Intersects, contains, disjoint, equals, touches, overlaps, within, and crosses. GUI and parameters of this tool are shown in Figure 3-46. The selected features can be saved to temporary layer in QGIS, file (*.shp), or database table (PostGIS table or SpatiaLite table). Users can see help of this tool on the right side of the GUI or web browser by clicking the "Tool Help" button.

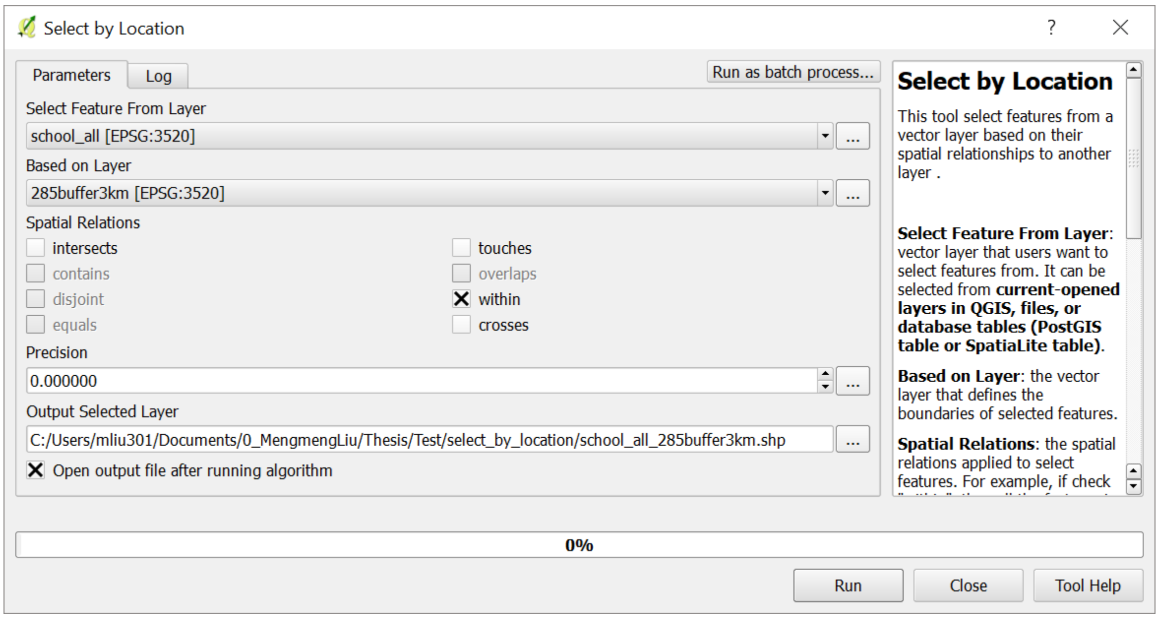


Figure 3-46 GUI and Parameters of Select by Location

3.8.7 Network Analysis

Network analysis tools in the "Spatial Infrastructure Sustainability Assessment" plugin include tools to measure connectivity, accessibility, and centrality of the network. Details of each tool are illustrated in the following sections. All the implemented tools support both local shapefiles and database tables (PostGIS and SpatiaLite).

3.8.7.1 Connectivity of Network

The measures of connectivity provided in the framework include Alpha Index, Beta Index, Gamma Index, and Detour Index. Calculation of all these measures needs the number of edges (E Field), number of nodes (V Field), and number of non-connected subgraphs (P Field) in the network, thus the author designs a tool (“Basic Network Index”) to calculate these three basic network indices, see Figure 3-47.

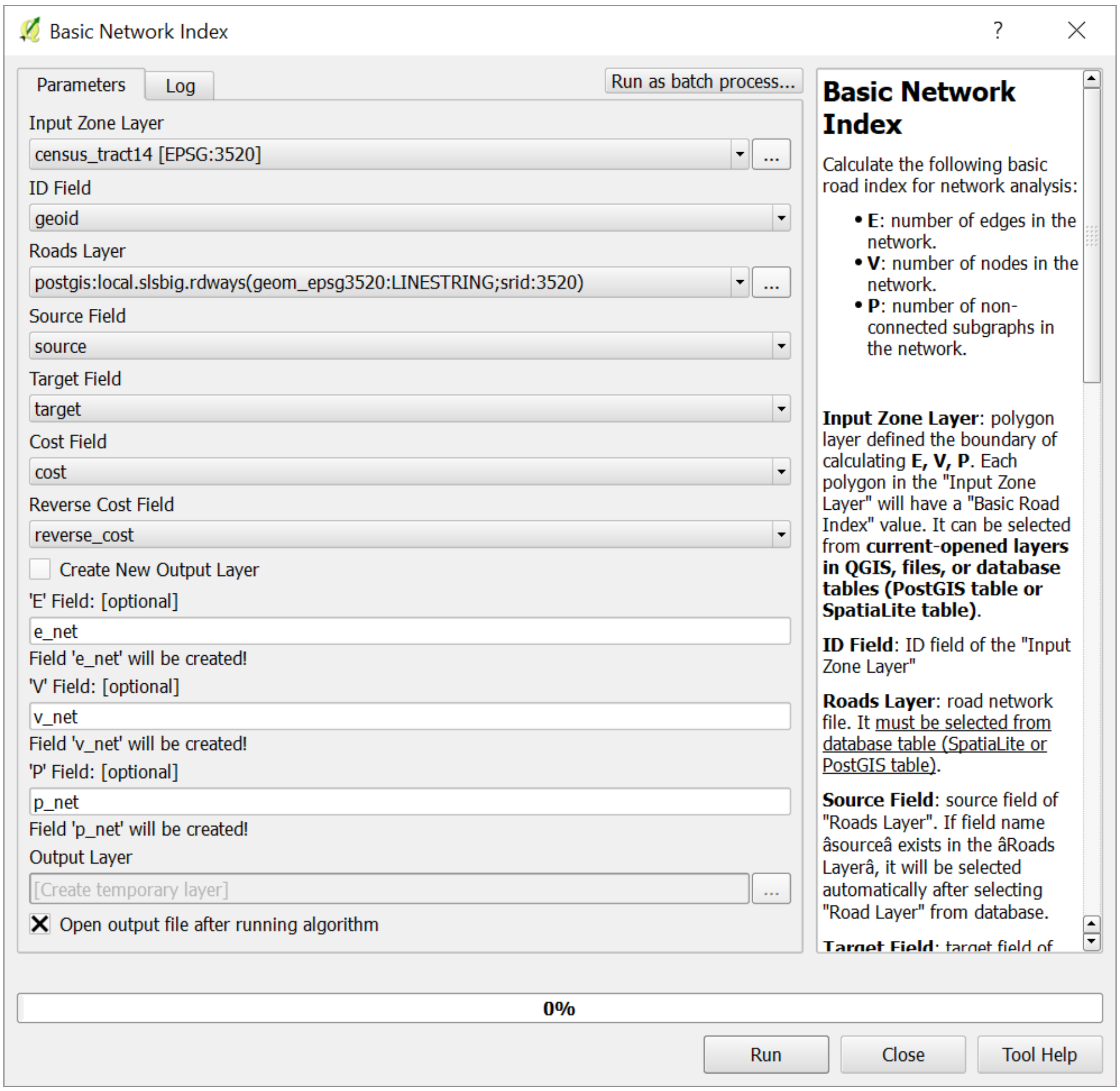


Figure 3-47 GUI and parameters of Basic Network Index Tool

GUI and parameters of “Alpha Index”, “Beta Index”, “Gamma Index”, and “Detour Index” tools are shown in Figure 3-48, Figure 3-49, Figure 3-50, and Figure 3-51 respectively. When calculating “Alpha Index”, “Beta Index”, and “Gamma Index”, the “Input Zone Layer” defines the spatial scales of network connectivity (i.e., Each polygon in the "Input Zone Layer" will have a "connectivity" value.). It can be selected from current-opened layers in QGIS, files, or database tables (PostGIS table or SpatiaLite table). The “Roads Layer” must have correct topological relationships, and must be selected from the database (PostGIS table or SpatiaLite table). After selecting the “Roads Layer”, the “Source Field”, “Target Field”, “Cost Field”, “Reverse Cost Field” will be filled automatically if the fields name ‘source’, ‘target’, ‘cost’, ‘reverse_cost’ respectively exist in the “Roads Layer”. When calculating “Alpha Index”, “Beta Index”, and “Gamma Index”, users can choose to use existing “E, V, P” fields to calculate them. Otherwise, the tool will calculate the “E, V, P” value automatically and automatically creates fields (‘e_net’, ‘v_net’, ‘p_net’) in the input layer to save their values. The connectivity results of the network can be saved to a temporary layer, to the new field in the “Input Layer”, to file (*.shp), or to the database table (PostGIS table or SpatiaLite table). Output layer can be opened automatically in QGIS if users check the box “open output file after running algorithm”. Users can see help of tools on the right side of GUI, or web browser by clicking the “Tool Help” button.

The “Detour Index” tool calculates the detour index of each random vector in the “Input Random Vector Layer” when traveling from its tail to its head on the road network defined by “Roads Layer” parameter. Users can also set the “Highway Indicator (1 or 0) [optional]” parameter to prevent projecting random vector’s tail or head to some node on highways. It must only have 1 or 0 values in the field, and 1 indicates a “highway”, 0 represents all other kinds of roads. The detour index can be saved to the “Output Field” in the “Input Layer”, “Temporary Layer”, “File” (*.shp) or database tables (PostGIS table or SpatiaLite table).

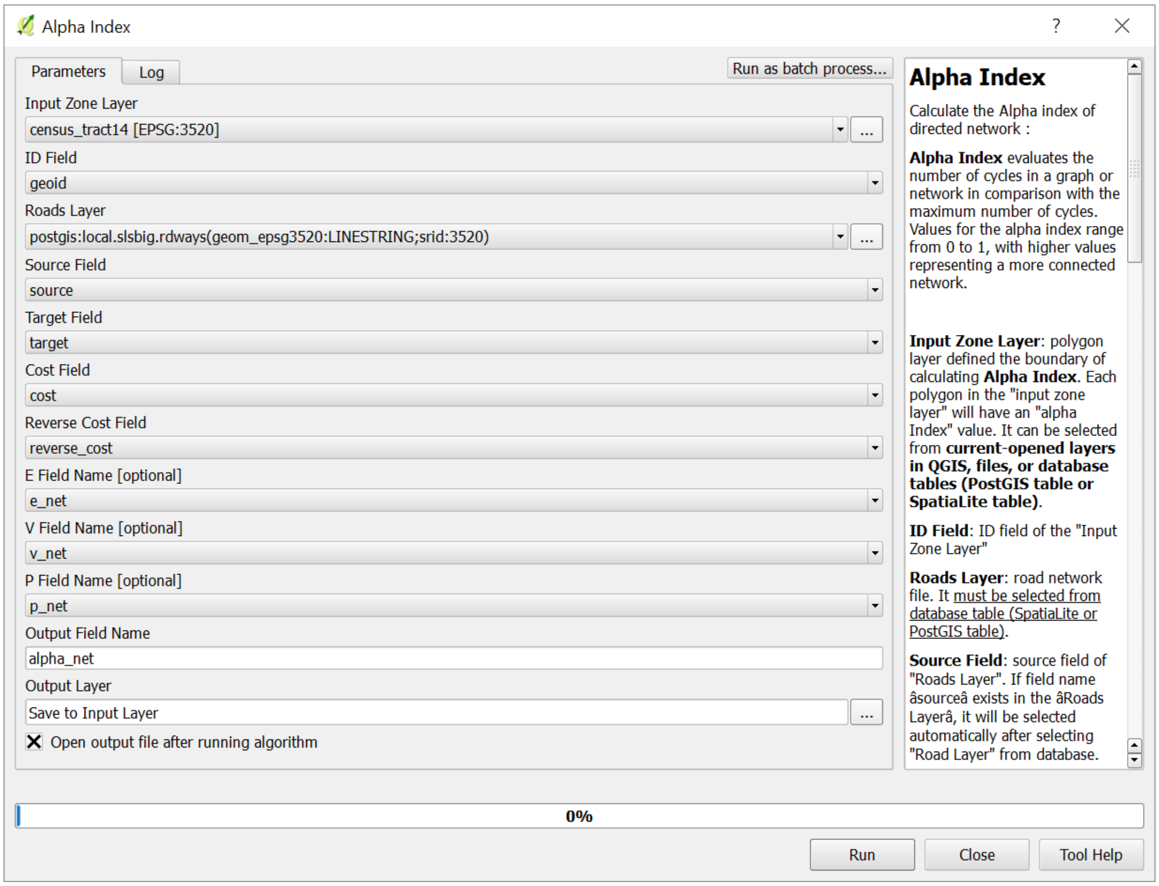


Figure 3-48 GUI and Parameters of Alpha Index Tool

(Calculating alpha index of network (rdways) for each polygon of “Input Zone Layer” (census_tract14), using existing E, V, P fields. Alpha index values are saved in the “alpha_net” field of “Input Zone Layer”.)

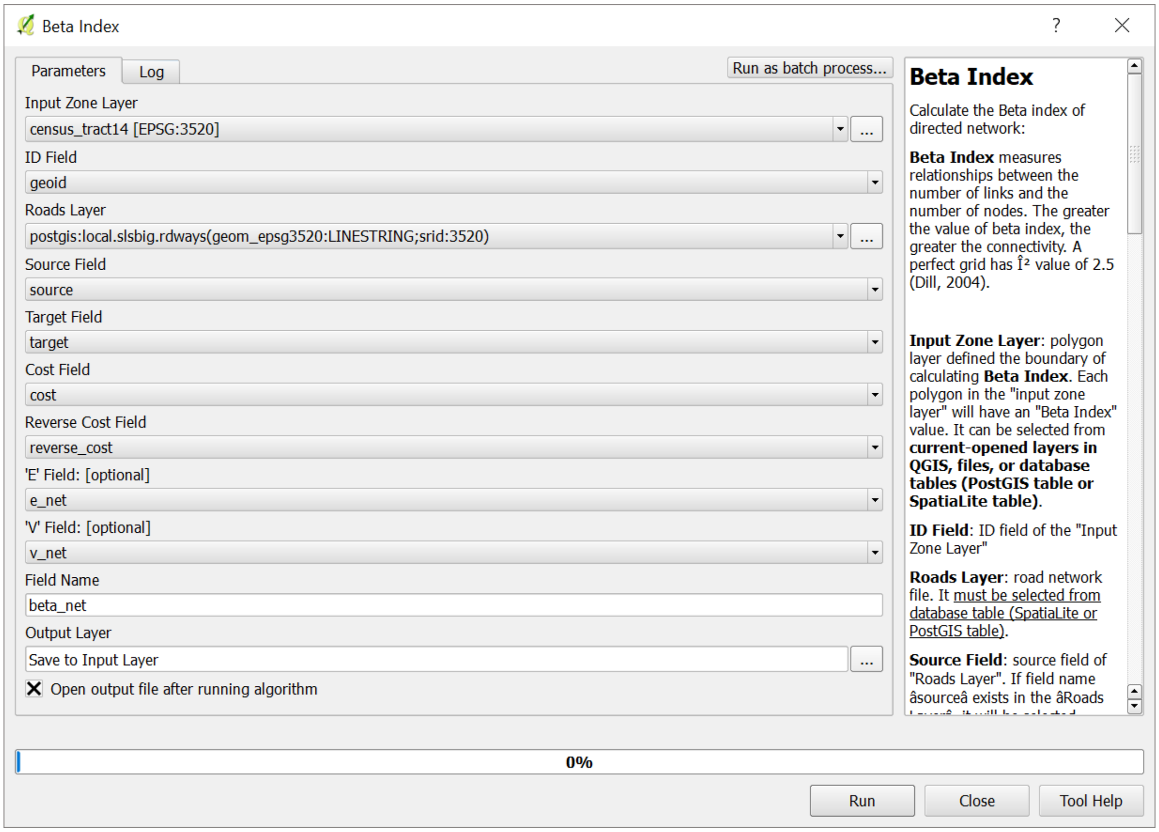


Figure 3-49 GUI and Parameters of Beta Index Tool

(Calculating beta index of network (rdways) for each polygon of “Input Zone Layer” (census_tract14), using existing E,V fields. Beta index value is saved in field ‘beta_net’ of “Input Zone Layer”.)

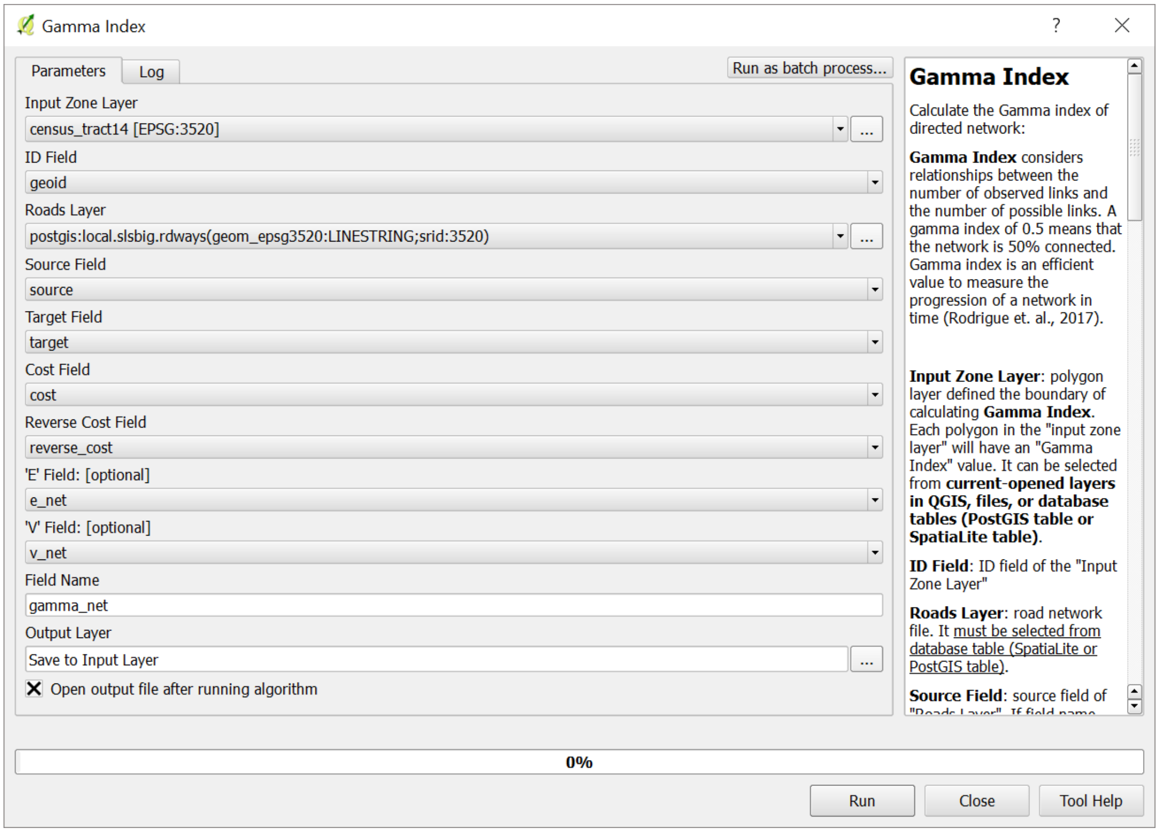


Figure 3-50 GUI and Parameters of Gamma Index Tool

(Calculating gamma index of network (rdways) for each polygon of “Input Zone Layer” (census_tract14), using existing E,V fields. Gamma index value is saved in field ‘gamma_net’ of “Input Zone Layer”)

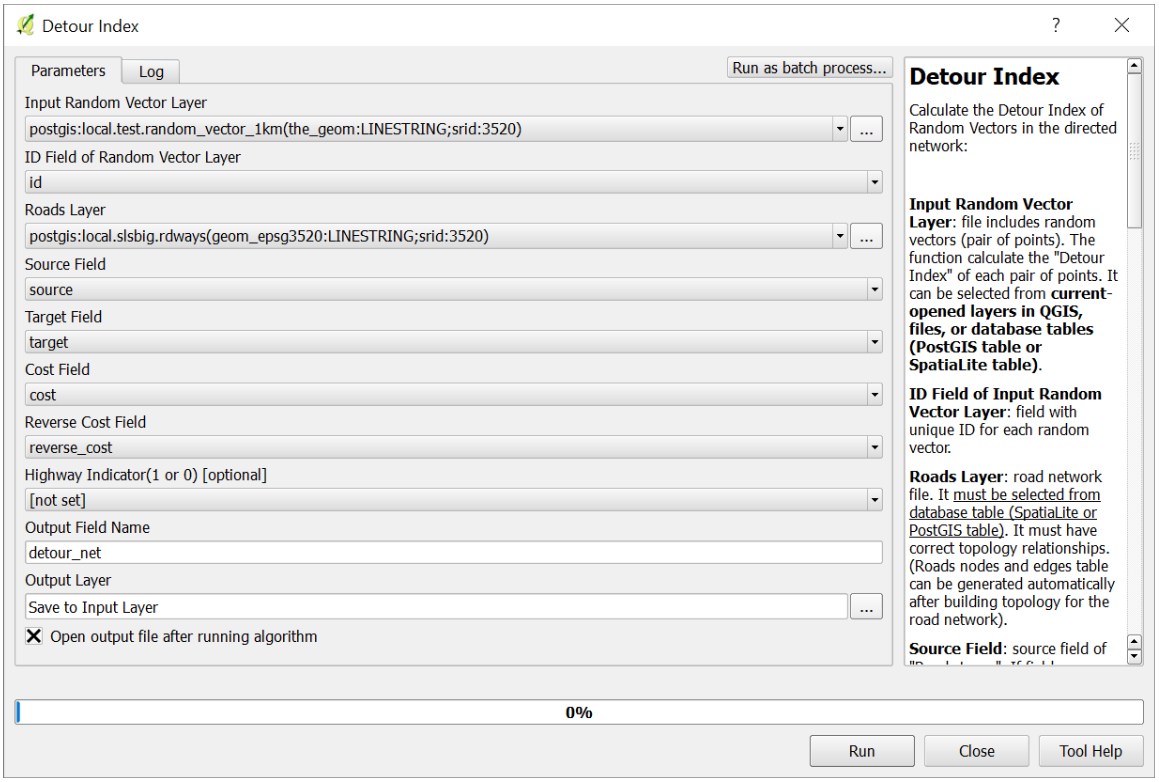


Figure 3-51 GUI and Parameters of Detour Index Tool

(Calculating detour index of each random vecotr in the “Input Random Vector Layer” when traveling on the roads network (Roads Layer). Detour index value is saved in field ‘detour_net’ of “Input Random Vector Layer”)

3.8.7.2 Accessibility of Network

There are two ways to calculate one’s accessibility in network, one is to calculate the accessibility to nearest facility in the network, and the other is to calculate the accessibility to all the facilities that can be reached within some search radius in the network. Therefore, the “Spatial Infrastructure Sustainability Assessment” plugin provides “Nearest Facility” tool and “Facilities within Distance” tool to estimate accessibility. Their GUI and parameters are shown in Figure 3-52 and Figure 3-53 respectively.

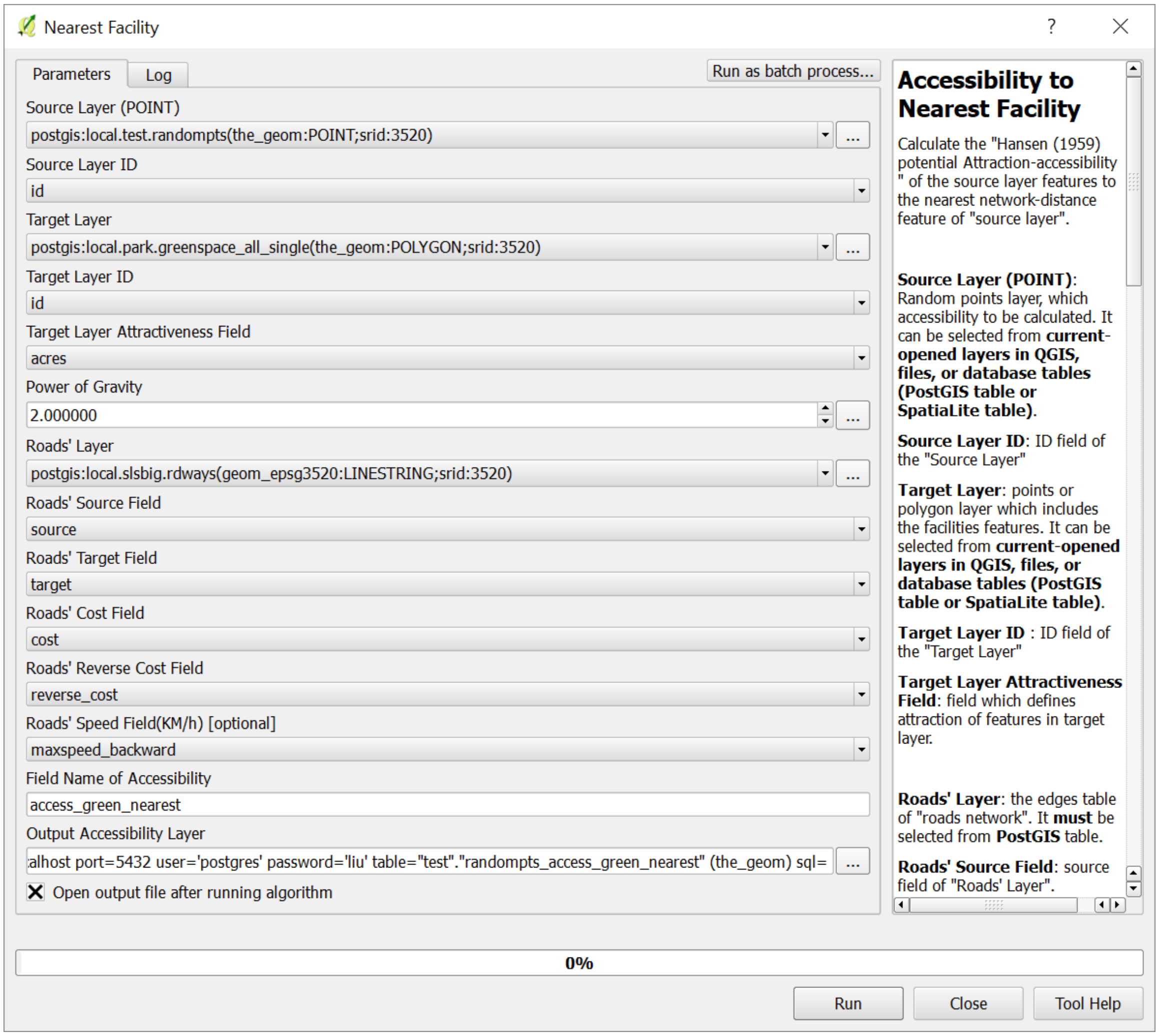


Figure 3-52 GUI and Parameters of "Nearest Facility" Tool

(For each point in "Source Layer", calculating its accessibility to the nearest features in "Target Layer" using Equation 3.8.)

The tool calculates the accessibility of each feature in "Source Layer (POINT)" to features in "Target Layer" using Equation 3.8. "Roads' Layer" defines the road network, "Output Accessibility Layer" saves the accessibility results. Users can save results to a temporary layer, to a new field in the "Input Layer", to file (*.shp), or to a database table (PostGIS table or SpatiaLite table). The output file can be opened automatically in QGIS if users check the box “open output file after running algorithm”. Users can see help of tools on the right side of GUI, or open it on the web browser by clicking the “Tool Help” button.

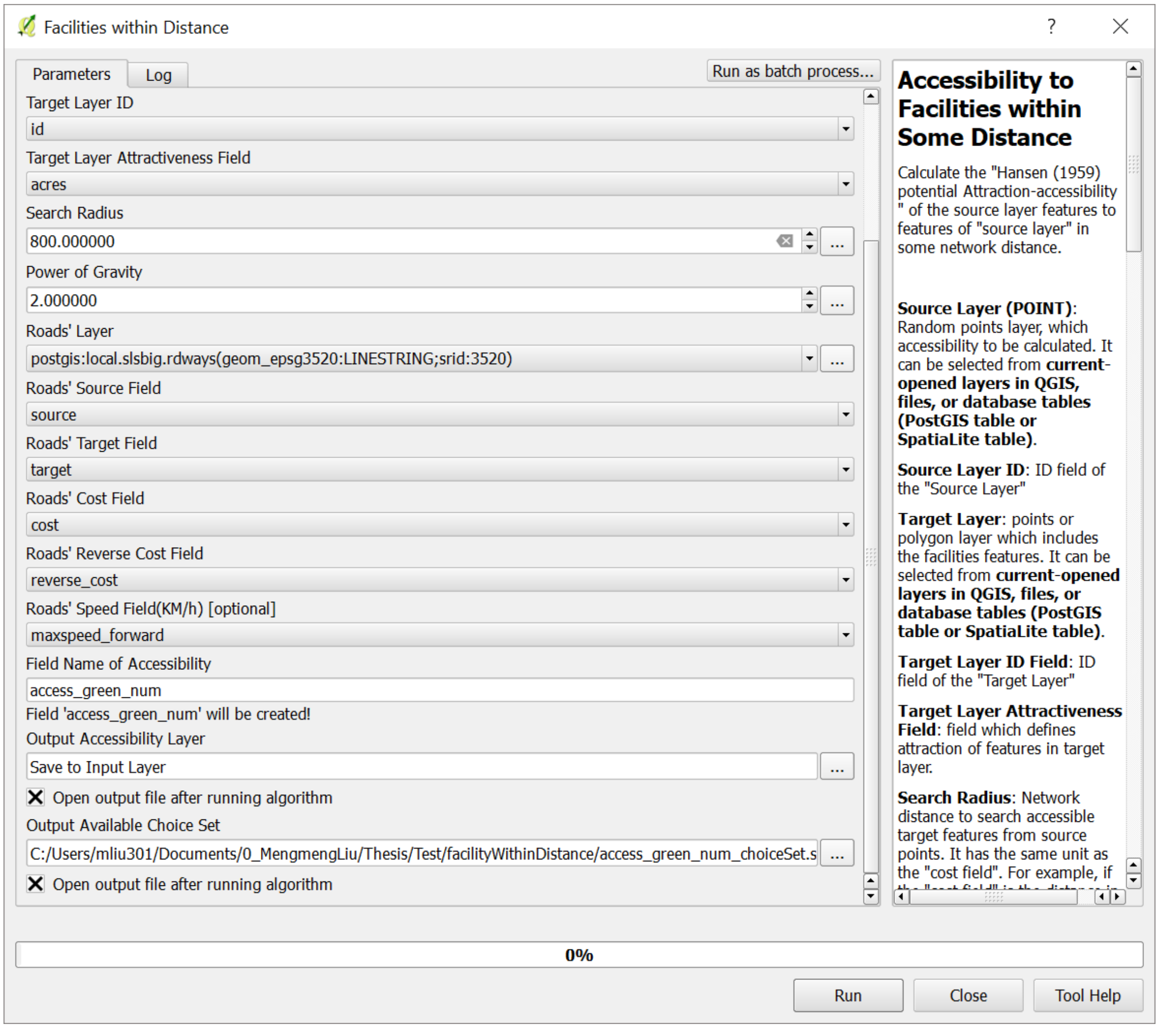


Figure 3-53 GUI and Parameters of “Facilities within Distance” Tool

(Calculating accessibility to all the target features that can be reached from each point in source layer within the “Search Radius” (800 meters))

3.8.7.3 Centrality of Network

The “Spatial Sustainability Infrastructure Assessment” plugin provides many measures of connectivity, including betweenness centrality, closeness centrality, PageRank centrality, eigenvector centrality, and Katz centrality. Their GUI and parameters are shown in Figure 3-54, Figure 3-55, Figure 3-56, and Figure 3-57 respectively.

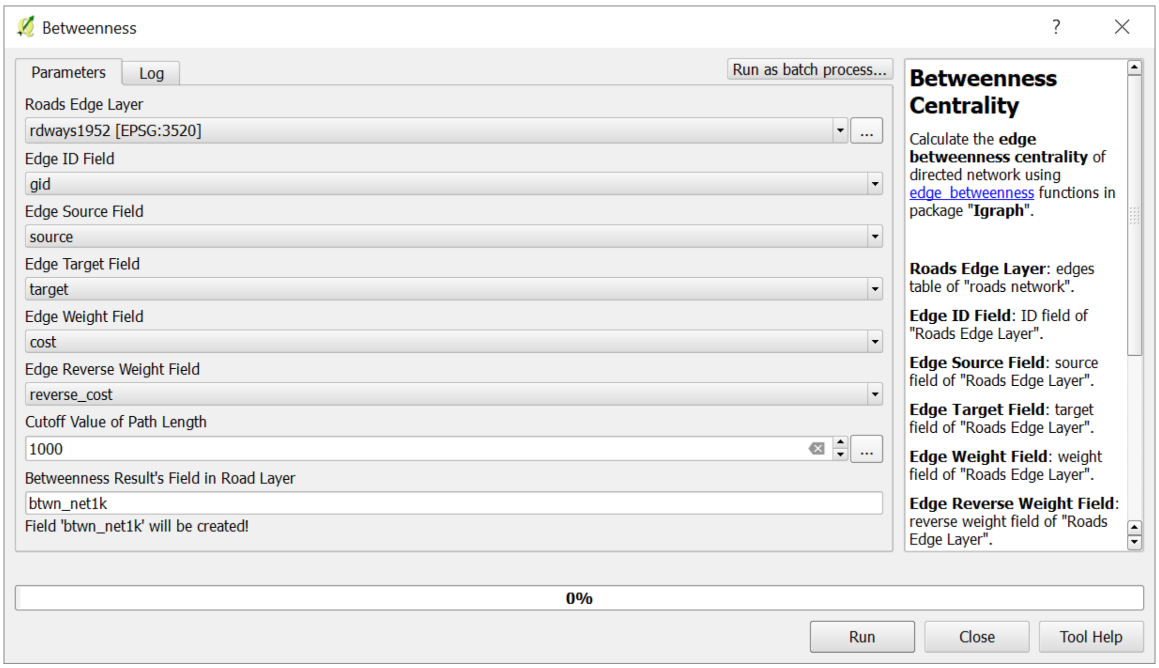


Figure 3-54 GUI and Parameters of Betweenness Centrality Tool

Users can see the help of these tools on the right side or open it on the browser by clicking the “Tool Help” button. Users can set roads node layer, roads edge layer, and specify the source, target, weight, reverse weight field, and “Cutoff value of path length” (default is zero). If users set the “Cutoff value of path length” to be an integer greater than 0, then only paths less than or equal to this length are considered in the centrality calculation (i.e., only calculate an estimation of the centrality values.). If the cutoff value is zero, the exact centrality values are returned.

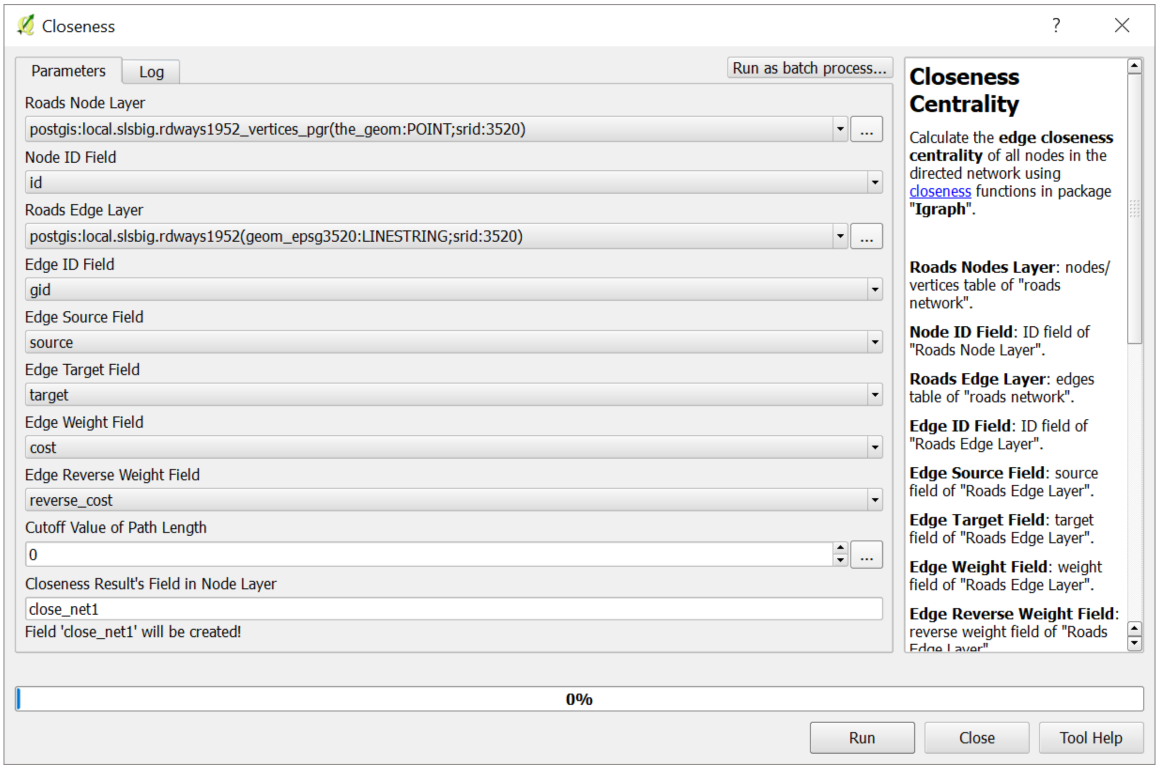


Figure 3-55 GUI and Parameters of Closeness Centrality Tool

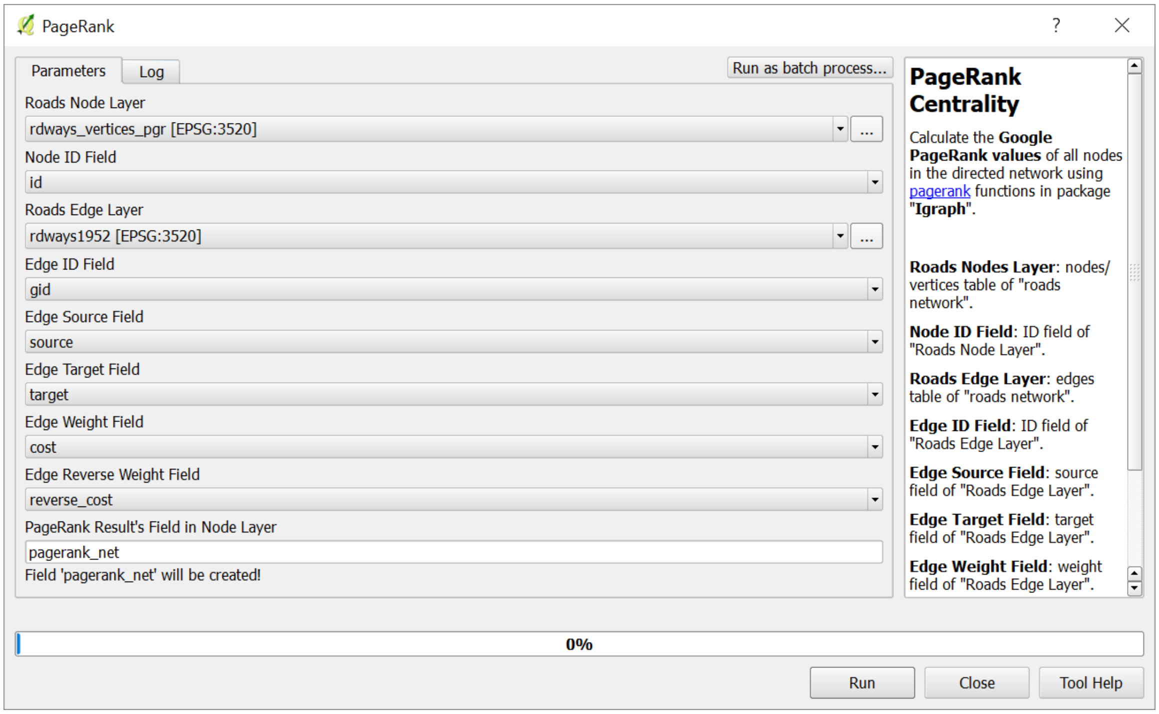


Figure 3-56 GUI and Parameters of PageRank Centrality Tool

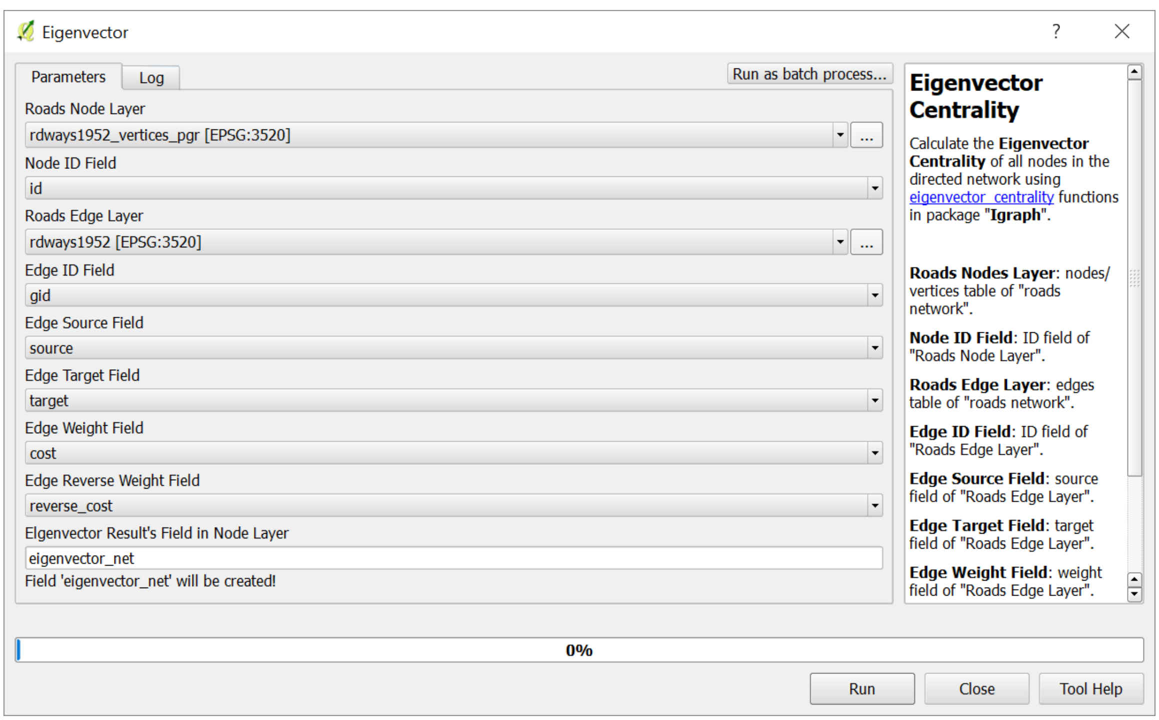


Figure 3-57 GUI and Parameters of Eigenvector Centrality Tool

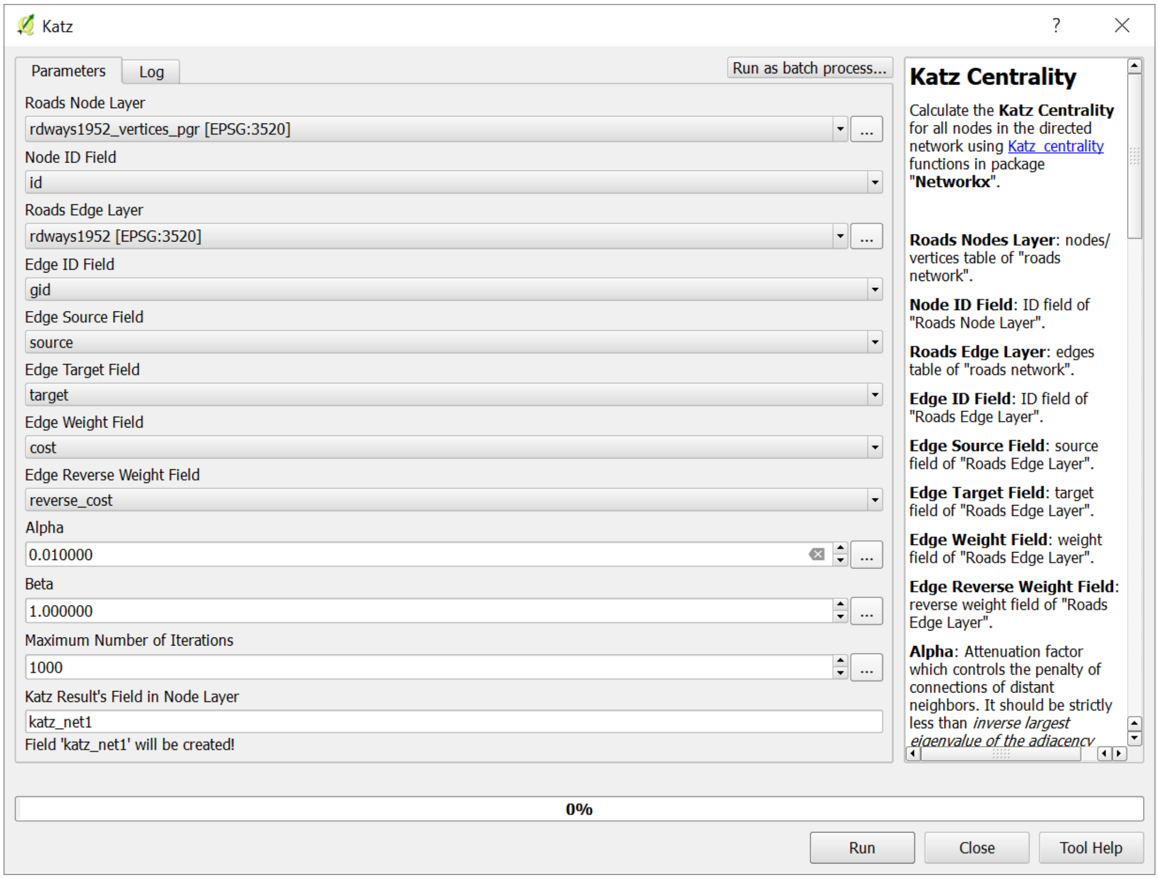


Figure 3-58 GUI and Parameters of Katz Centrality Tool

3.8.7.4 Path Count

"PathCount" tool estimates the weighted traffic flows passing through each edge in the road layer, using Equation 3.39. GUI and parameters of "PathCount" tool are shown in Figure 3-59.

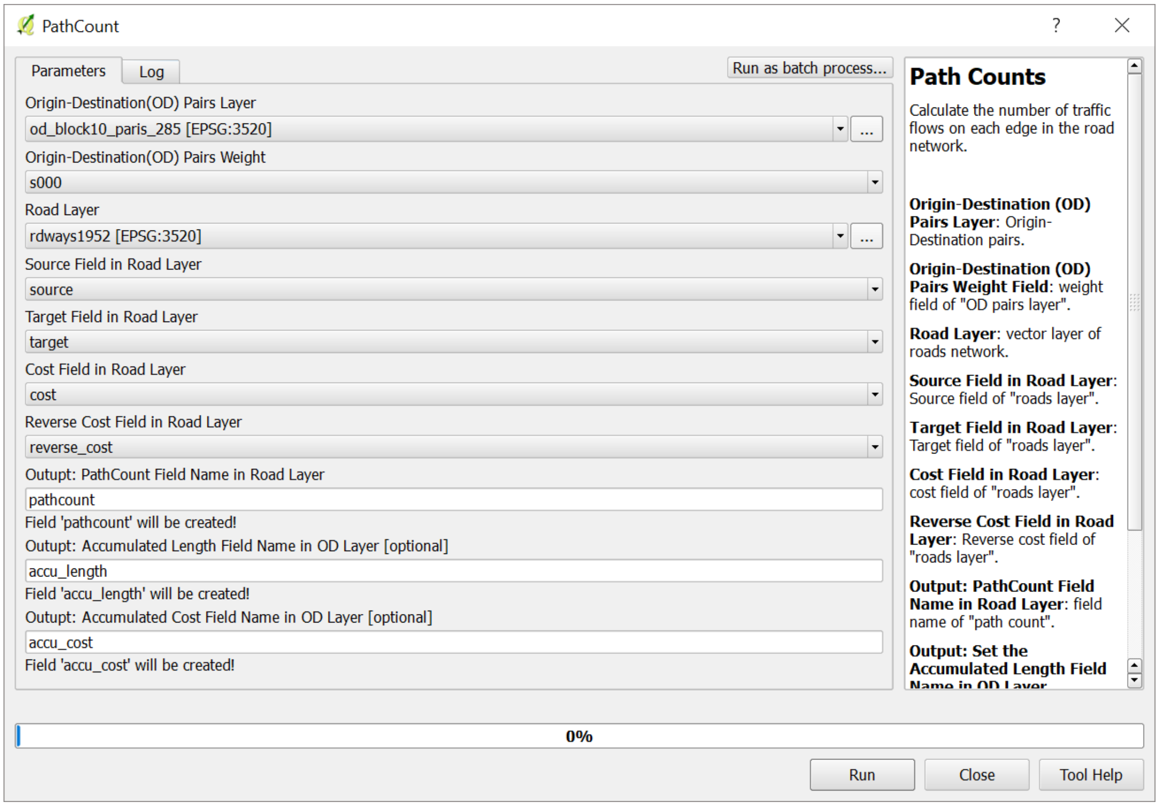


Figure 3-59 GUI and Parameters of PathCount Tool

The “OD Pairs Weight Field” defines the weight of OD pairs (w_{st} in Equation 3.39), the “cost and reverse cost field” in the road layer determines the shortest cost roads for each OD pair. If the cost is the time cost on each edge in the road layer, then this tool will give users the shortest time roads for each OD pair and then counts the weighted traffic flows on each edge in the road layer.

$$\mathrm{flow}(e) = \sum_{s,t \in OD} w_{st} * \mathbf{1}(e \in p_{st})$$

$$\mathbf{1}(e \in p_{st}) = \begin{cases} 1, & \text{if } e \text{ is on path } p_{st} \\ 0, & \text{otherwise} \end{cases} \tag{3.39}$$

Where $\text{flow}(e)$ is the weighted traffic flows passing through edge e, w_{st} is the weight of OD pairs.

3.8.8 *Multi-Criteria Decision Analysis*

3.8.8.1 Indicator Pre-Analysis

The pre-Analysis of indicators include normalization, correlation analysis, and PCA of indicators. Their details are illustrated in the following paragraphs.

Normalization of Indicators

The “Normalization” tool provides many normalization methods, including “Min-Max”, Standardization (Z-Score), Ranking, “Distance-to-Reference”, and “Categorical Scales”. Users can normalize multiple fields of the input vector layer at the same time. GUI and parameters of this tool are shown in Figure 3-60. Users can choose indicators from the “Available Fields” table by double-click it or select it and click “>” button to move the selected field into “Selected Fields” table. In this table, users can type “alias” for each field, which will show in the results layer. The normalization results can be saved to file (*.shp), to database table (SpatiaLite or PostGIS table), or a temporary layer opened in QGIS. Users can see helps of this tool on the right side of GUI or open the help on browser by clicking “Tool Help” button.

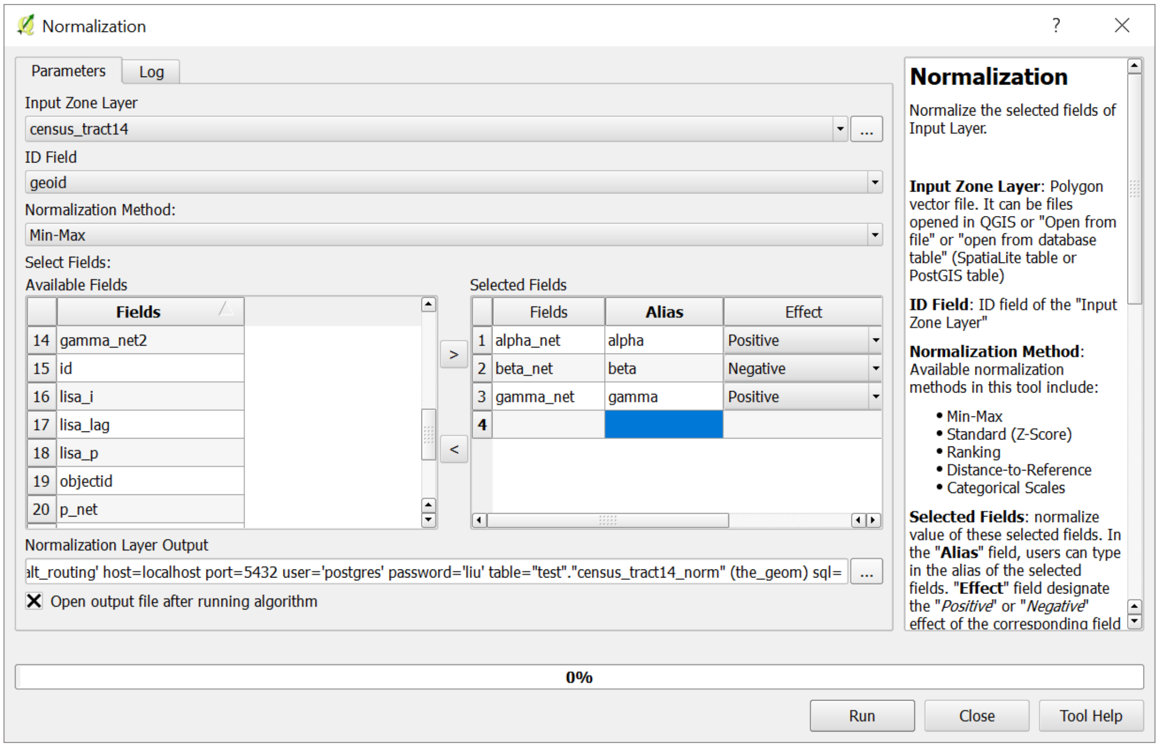


Figure 3-60 GUI and Parameters of Normalization Tool

Correlation of Indicators

The "Correlation Analysis" tool in this framework can calculate the correlations between multiple selected fields of the input zone layer. Its GUI and parameters are shown in Figure 3-61. Users can select the correlation coefficient measures from the dropdown list, which includes Person, Spearman, and Kendall. Users can also define the order of correlation matrix, and 5 order methods are provided in this tool ("Order Method" parameter).

- Original: order same as the selected fields table.
- AOE: order by the angles of the eigenvectors.
- FPC: the first principal component order.

- Hclust: the hierarchical clustering order.
- Alphabet: alphabetical order of the selected fields name, if alias name is provided, then the alias name is used in the order.

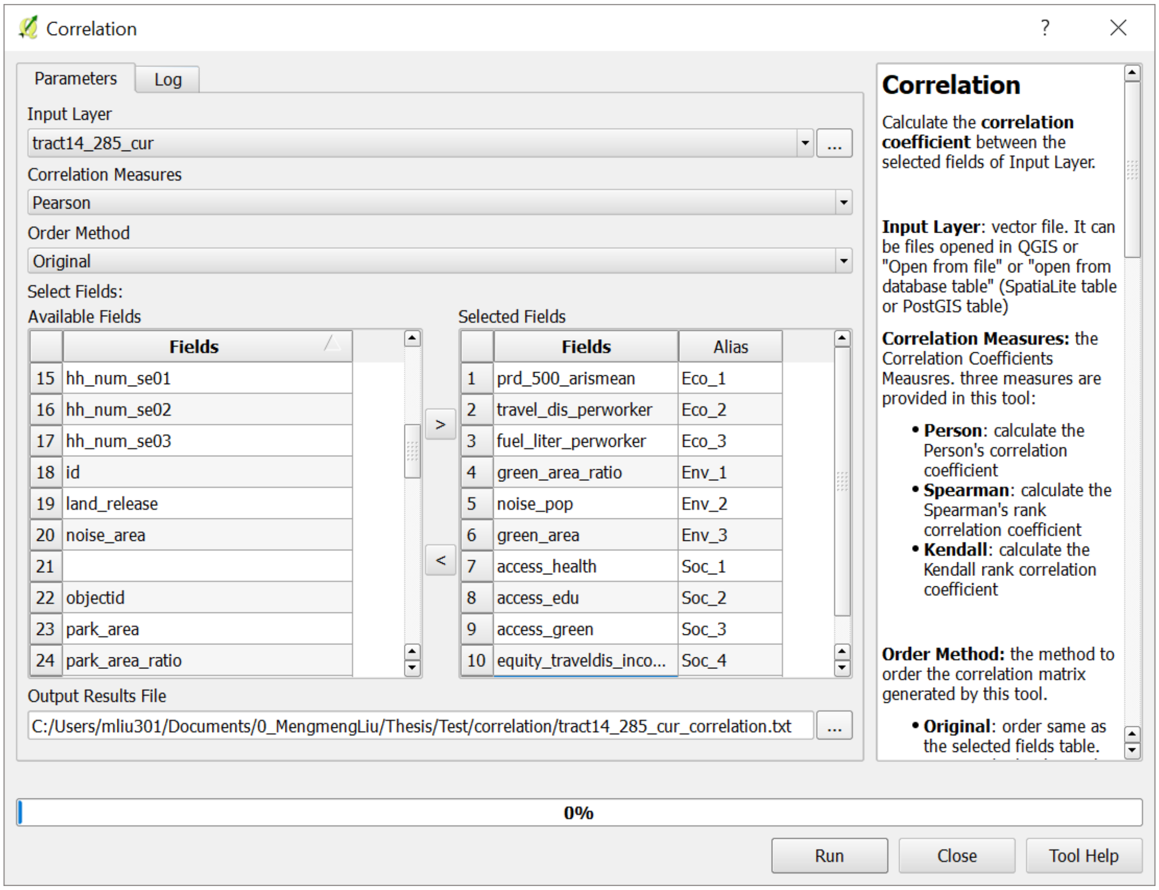


Figure 3-61 GUI and Parameters of Correlation Analysis Tool

After selecting the correlation measurement and the order method, users can choose indicators from the "Available Fields" table by double-click or select it and click ">" button to move the selected field into "Selected Fields" table. In this table, users can type "alias" for each field, which will show on the correlation matrix. The "Output Results File" includes the "Selected Fields" table, the correlation matrix, the p-value of the correlation matrix, the correlation measures, and the order method. Users can see helps of this tool on the right side of GUI or by clicking “Tool Help” button to show helps on web browser. After clicking “Run” button, a new window will open to show the correlation coefficients matrix between the selected fields (see Figure 3-62). In this window, users can save the plot to files (Postscript, PDF, Png, Bmp,Tiff, and Jpeg), copy it to clipboard, or print.

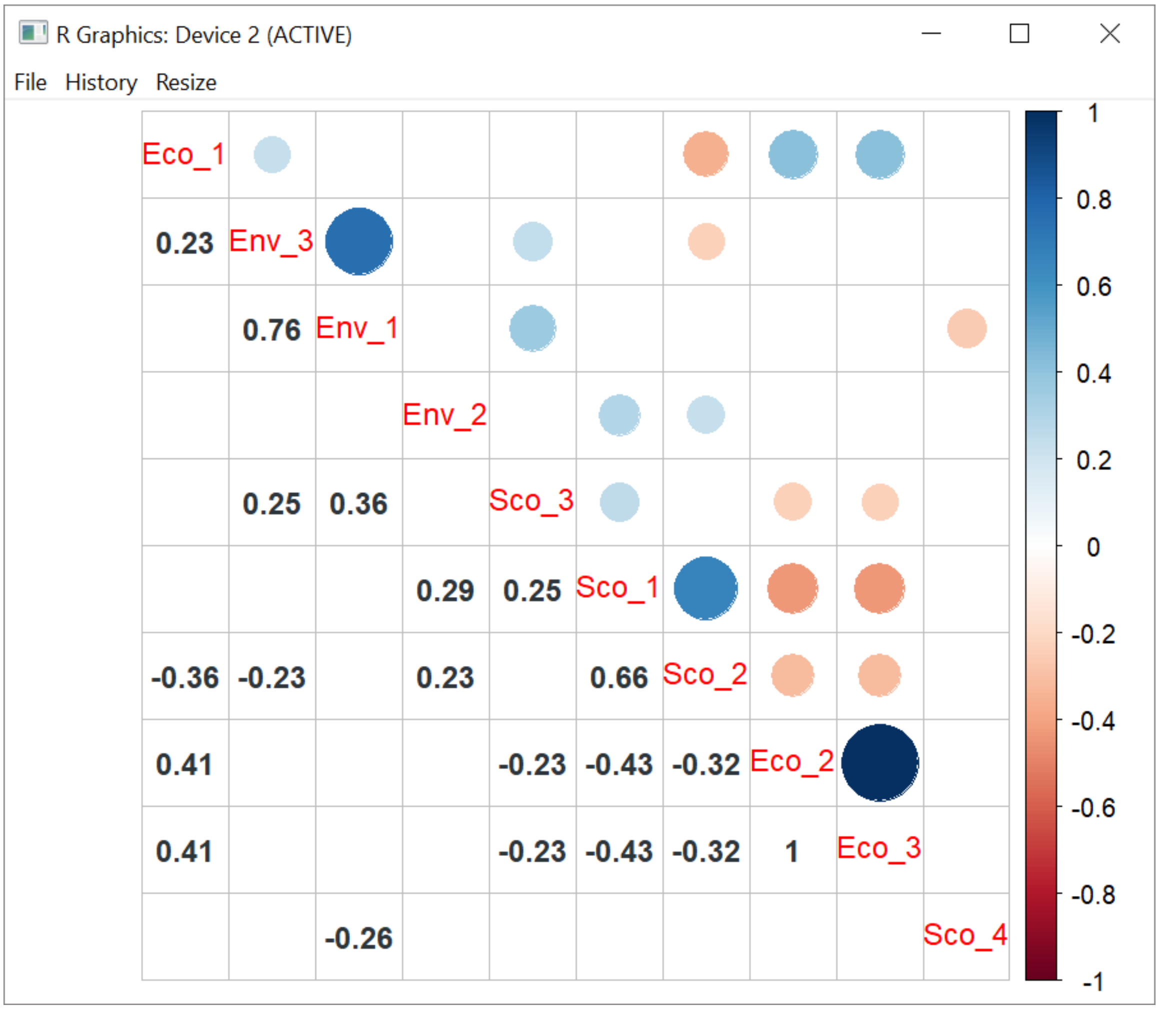


Figure 3-62 Correlation Matrix Between the Selected Fields in Figure 3-61

PCA of Indicators

The “PCA” tool in this framework select relatively independent indicators using PCA. Its GUI and parameters are shown in Figure 3-63. Users need first define the “Input

Zone Layer", which is a polygon vector file. Users can select from the dropdown list (currently opened layers in QGIS), or choose file from a local file or from "database" (SpatiaLite table or PostGIS table). Users can set the value of "Threshold of PCA Accumulation Variance" which is the ratio of variance kept by PCs. The default value is 0.9, meaning the selected PCs keep 90% of variance. This parameter controls the number of PCs kept for the following analysis. The "Factor Loading Ratio of Principle Component (PC)" controls how many fields are selected in each PC. Its default value is 0.9, meaning that only fields with absolute factor loading greater than 0.9 of the largest absolute factor loading in the PC are retained. If these fields are statistically significantly correlated (i.e., correlation coefficient ⩾ 0.8 and its p-value < 0.05), then only the fields with the highest sum of absolute correlation coefficients is retained. Otherwise, all of them are retained.

This tool can save the PCA results to a temporal file or local file. The file records the "correlation matrix" of the selected fields (as shown in Figure 3-64), the information each PC (eigenvalues, explained variance ratio, and eigenvectors (factor loadings of each selected field in each PC)), and the relatively independent indicators/ fields selected by PCA (as shown Figure 3-65). Users can see the help of this tool on the right of GUI or open help on the browser by clicking the "Tool Help" button.

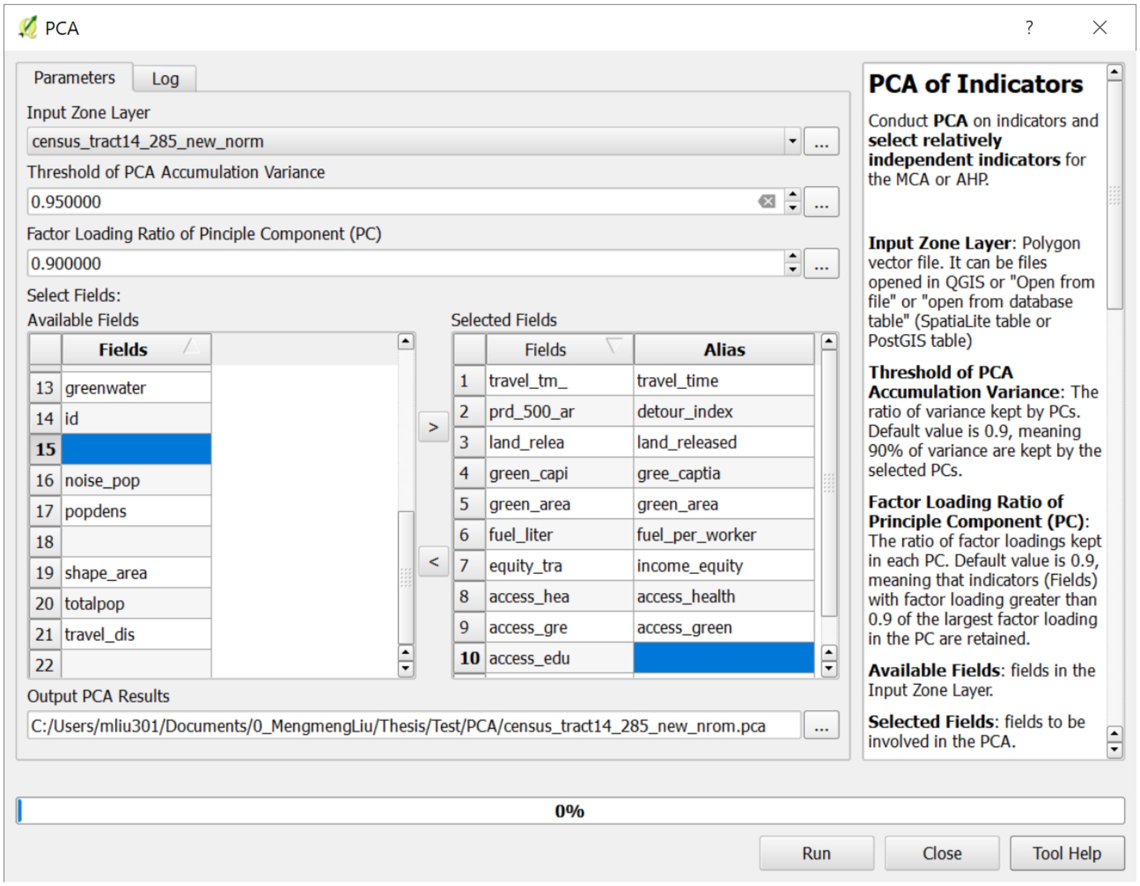


Figure 3-63 GUI and Parameters of PCA Tool (Select Indicators)

	prd_500_ar	access_gre	access_hea	access_edu	green_area	green_capi	travel_tm_	fuel_liter	equity_tra	land_relea
prd_500_ar	1	-0.092	-0.201	-0.303	0.265	0.129	0.3	0.326	0.09	0.386
access_gre	-0.092	1	0.234	0.048	0.174	0.189	-0.208	-0.196	0.039	0.046
access_hea	-0.201	0.234	1	0.433	-0.141	-0.121	-0.415	-0.362	0.05	-0.017
access_edu	-0.303	0.048	0.433	1	-0.158	-0.118	-0.126	-0.084	-0.053	-0.033
green_area	0.265	0.174	-0.141	-0.158	1	0.811	0.12	0.097	-0.052	0.282
green_capi	0.129	0.189	-0.121	-0.118	0.811	1	0.048	0.027	-0.117	0.165
travel_tm_	0.3	-0.208	-0.415	-0.126	0.12	0.048	1	0.983	0.142	-0.034
fuel_liter	0.326	-0.196	-0.362	-0.084	0.097	0.027	0.983	1	0.124	-0.008
equity_tra	0.09	0.039	0.05	-0.053	-0.052	-0.117	0.142	0.124	1	0.009
land_relea	0.386	0.046	-0.017	-0.033	0.282	0.165	-0.034	-0.008	0.009	1
Sum(abs)	3.091	2.225	2.974	2.356	3.1	2.725	3.375	3.207	1.675	1.98

Figure 3-64 Results of PCA Tool: Correlation Matrix

	PC0	PC1	PC2	PC3	PC4	PC5	PC6	
EigenValue	3.214	2.62	2.315	1.791	1.656	1.303	1.102	
Explained Variance Ratio	0.319	0.212	0.165	0.099	0.085	0.052	0.037	
Explained Variance	0.043	0.029	0.023	0.013	0.012	0.007	0.005	
	PC0	PC1	PC2	PC3	PC4	PC5	PC6	
prd_500_ar	0.489*	-0.217	0.432	0.122	0.653*	-0.089	0.274	
access_gre	-0.002	0.24	-0.038	0.094	0.095	0.918*	0.28	
access_hea	-0.282	0.424*	0.253	0.639*	0.242	-0.056	-0.451	
access_edu	-0.17	0.139	0.056	0.442	-0.3	-0.279	0.762*	
green_area	0.496*	0.363	-0.317	0.091	0.016	-0.096	-0.061	
green_capi	0.407	0.390*	-0.42	0.047	0.008	-0.133	0.022	
travel_tm_	0.218	-0.419*	-0.144	0.388	-0.225	0.126	-0.151	
fuel_liter	0.217	-0.42	-0.108	0.44	-0.22	0.13	-0.134	
equity_tra	0.002	-0.024	0.02	0.037	0.016	0.052	-0.052	
land_relea	0.383	0.244	0.659*	-0.127	-0.563	0.073	-0.124	
Fields Selected by PCA	prd_500_ar	access_gre	access_hea	access_edu	green_area	green_capi	travel_tm_	land_relea

Figure 3-65 Results of PCA Tool: PCs and Fields Selected by PCA

3.8.8.2 AHP

AHP tool can generate weights of indicators automatically, based on the pairwise comparison priorities provided by users. Users first select the dimensions of indicators, which includes “Economic”, “Environmental”, “Social”, “Resilient”, and “Composite” dimension. GUI of this step is shown in Figure 3-66 (the first three dimensions are selected).

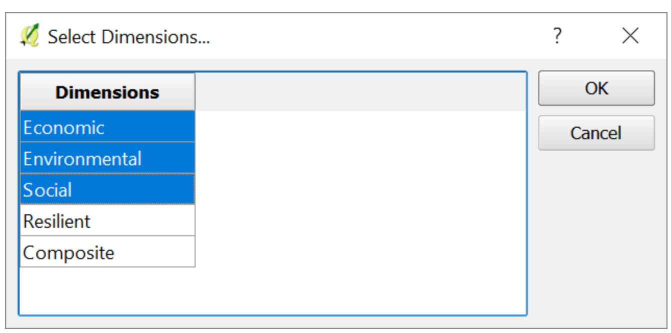


Figure 3-66 GUI of AHP: Select Dimensions

After clicking the button “OK”, users can see a new window which as shown in Figure 3-67. In this window, users can type in the pairwise comparison priorities in the “Pairwise Comparison Matrix A”, then can click “Calculate” button to derive weight of each dimension and the consistency ratio of the “Pairwise Comparison Matrix A”.

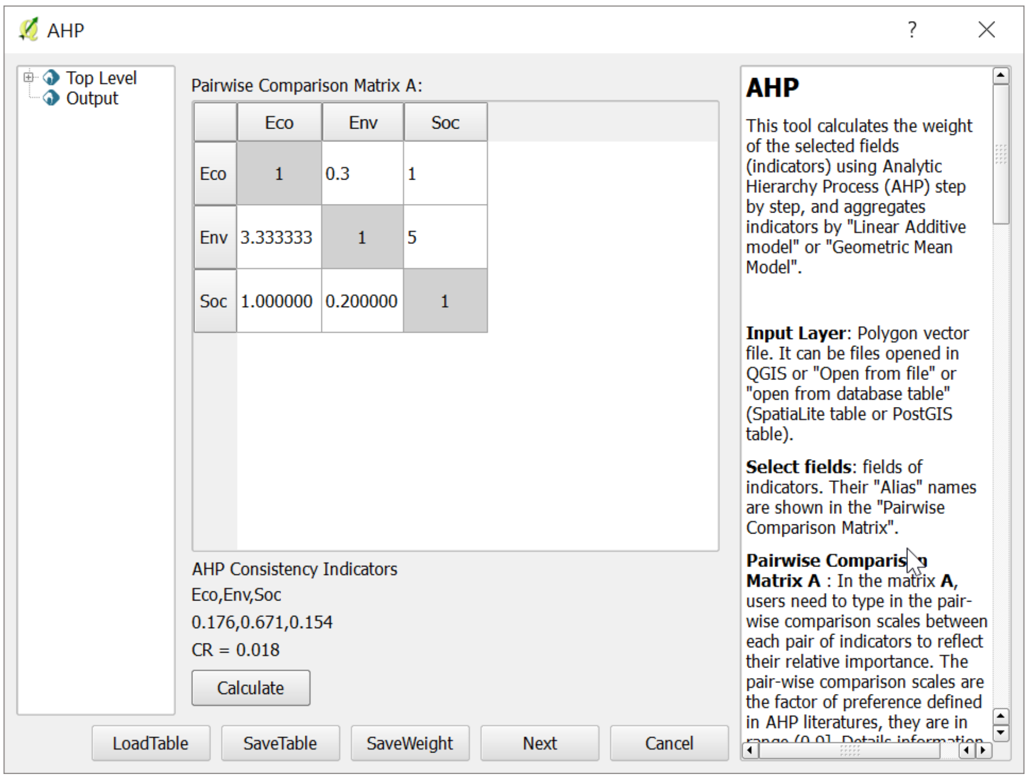


Figure 3-67 GUI of AHP: Pairwise Comparison Matrix and AHP Consistency Indicators (weights of Economic, Environmental, and Social dimension are 0.176, 0.671, and 0.154 respectively; Consistency Ratio is 0.018)

In the example shown in Figure 3-67, the “Environmental” dimension is 3.3 times more important than “Economic” dimension, and five times more important than “Social” dimension. “Economic” and “Social” dimension have the same importance. The derived weights of Economic, Environmental, and Social dimension are 0.176, 0.671, and 0.154

respectively; Consistency Ratio (CR) of comparison matrix A is 0.018. CR <0.1 means the weights derived by AHP are reasonable and can be used in the following analysis.

Users can also see the help of AHP tool on the right side of this window. Click "SaveTable" and "SaveWeight" button, users can save the "Pairwise Comparison Matrix A", and the derived weight into files. Click "LoadTable", users can load previously saved "Pairwise Comparison Matrix" into the current table. Click the "Next" button, users can then conduct AHP analysis on "Economic" dimension indicators. This tool first asks users to select indicators from "Input Layer for Economic" (GUI see Figure 3-68). Users must make sure that all the economic indicators are saved in the same file. Double click field name in the "Available Fields" table, the corresponding field name will be added in the "Selected Fields" table. Users can name alias for each field, the default value is "Eco_*i*" (*i* is the id of selected fields), examples see Figure 3-68.

After selecting the indicators, users can click the "Next" button to enter the pairwise comparison value in the "Pairwise Comparison Matrix A" in the new window (Figure 3-69). Users only need to enter the value in the upper triangle of the comparison matrix A, values in the lower triangle will be filled automatically. After filling matrix A, users need to click "Calculate" button to estimate its consistency (i.e., AHP Consistency Indicators below the matrix). If the CR < 0.1, users can click "Next" to continue work on indicators in other dimensions. Otherwise, they need to modify the value in the pairwise comparison matrix until its CR < 0.1. Figure 3-70 shows an example of inconsistency pairwise comparison matrix.

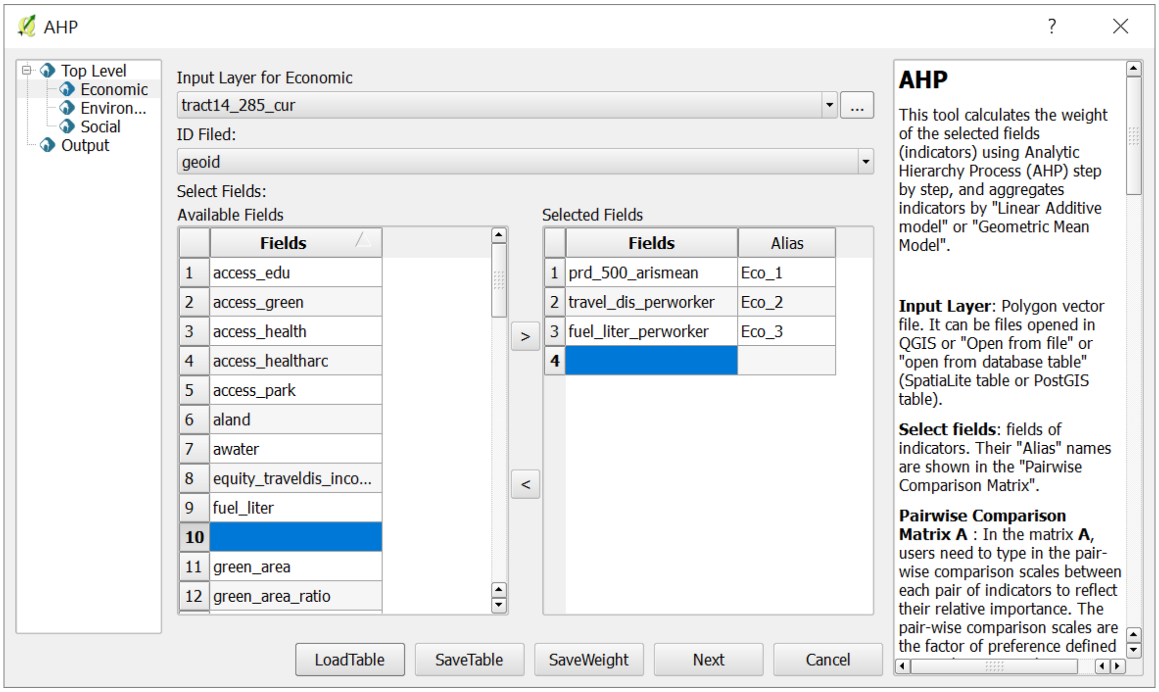


Figure 3-68 GUI of AHP: Select Economic Indicators

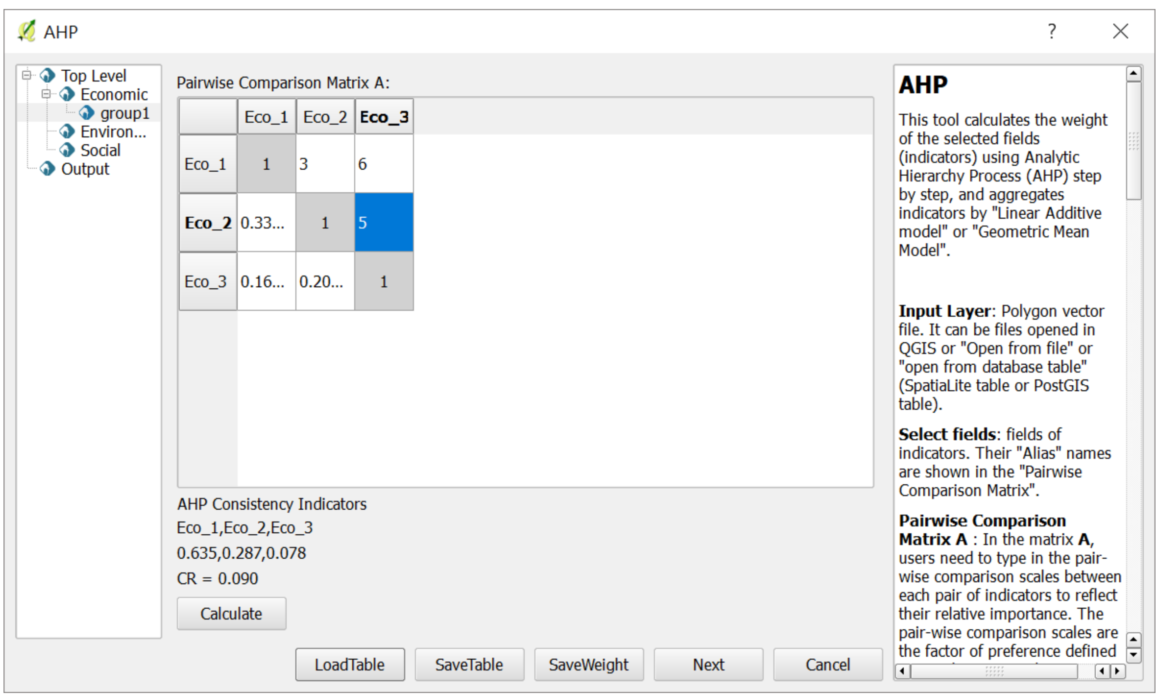


Figure 3-69 GUI of AHP: Pairwise Comparison Matrix A for Economic Indicators

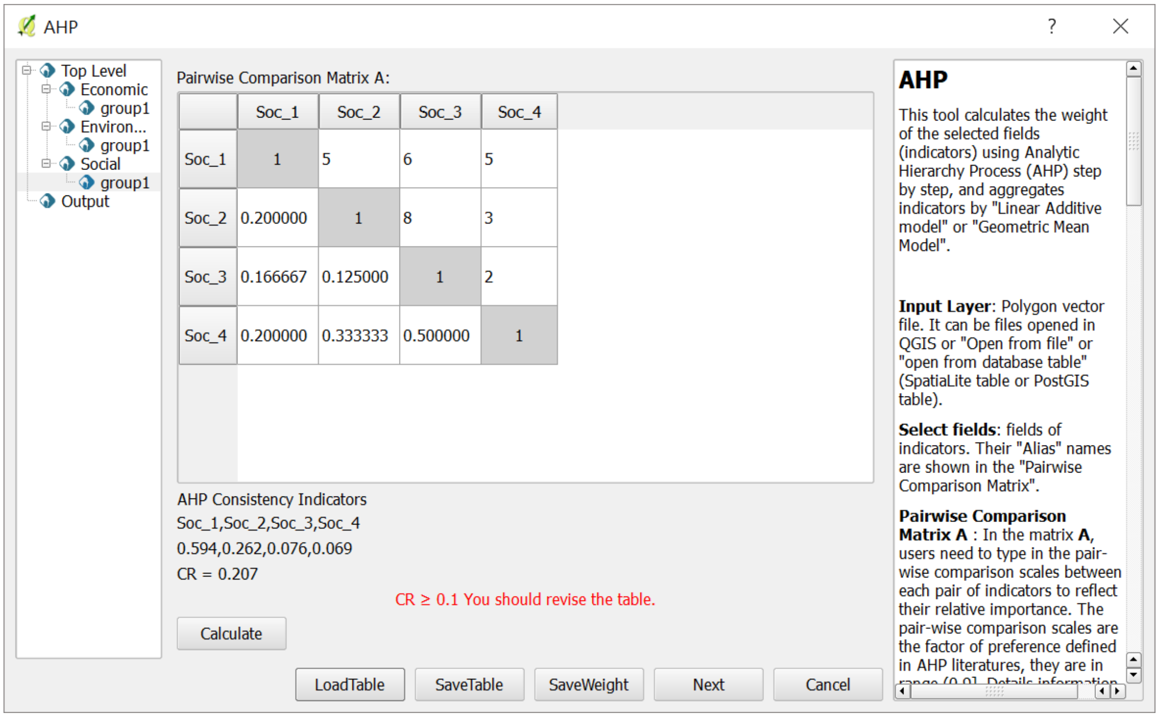


Figure 3-70 GUI of AHP: Pairwise Comparison Matrix for Social Indicators

(The pairwise comparison matrix is not consistent, because its CR ≥ 0.1. In this case, users need to modify the value in the pairwise comparison matrix.)

After extracting the weight of indicators, AHP function can also aggregate indicators (see Figure 3-71). There are two models available in the AHP function, linear additive model and geometric mean model. The aggregation results can be saved to file or save to database table (PostGIS or SpatiaLite). When the AHP process is complete, users can see a message window showing that “AHP Process Complete!”.

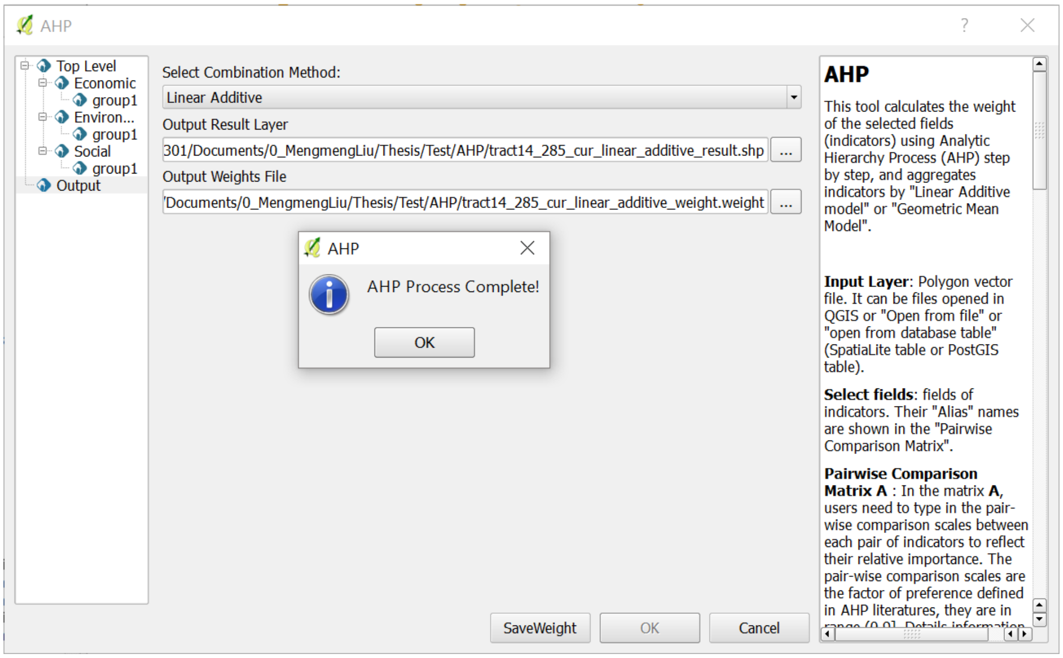


Figure 3-71 GUI of AHP: Indicators Weights and Indicators Aggregation Results

3.8.9 Assessment Results Analysis

After finishing AHP, users can use the “Uncertainty Analysis” tool and “Sensitivity Analysis” tool to evaluate the uncertainty and sensitivity of sustainability assessment results. Details of these two tools are discussed in the following paragraphs.

3.8.9.1 Uncertainty analysis

“Uncertainty Analysis” tool can determine the uncertainty of indicators’ weights to the sustainability assessment results. GUI and parameters of this tool are shown in Figure 3-72. Users can change the weight of indicators in some range (controlled by parameter “Uncertainty Range”). Users can also set the number of samples (default is 1000) and confidence level (90%, 95%, and 98%) for the confidence interval. This tool will calculate four uncertainty measures, including variance, CV, mean, and confidence interval of the sustainability assessment results (dimensional CSI and overall CSI). Uncertainty results can be saved to a temporary layer, file (*.shp) or PostGIS table. Figure 3-73 shows an example of uncertainty results.

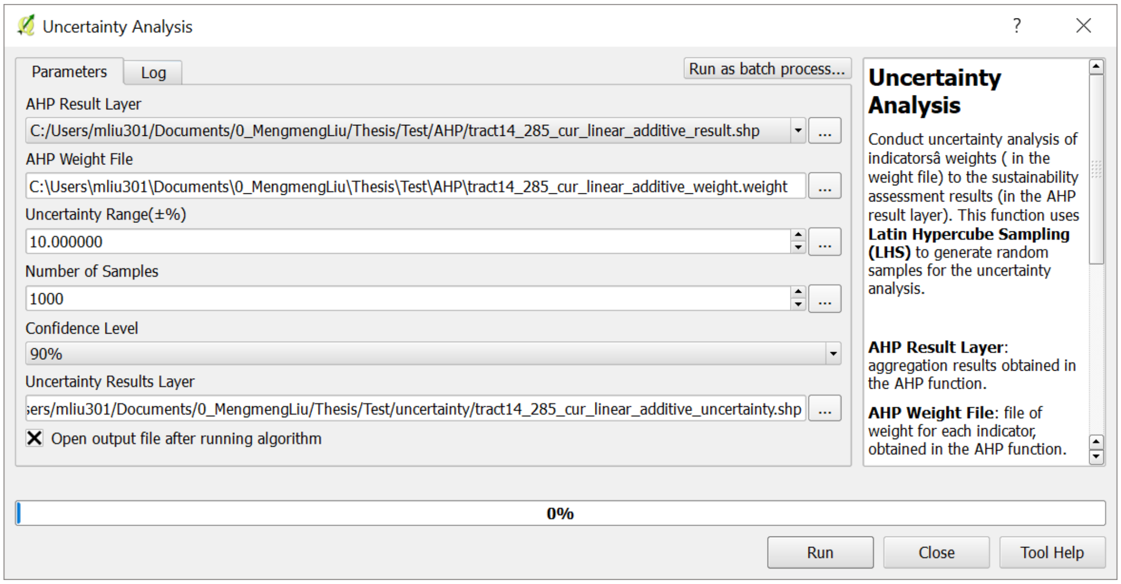
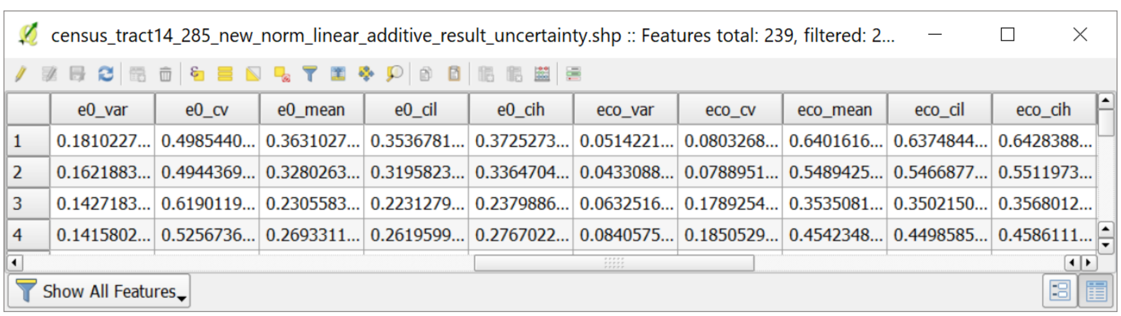


Figure 3-72 GUI and Parameters of Uncertainty Analysis Tool



	e0_var	e0_cv	e0_mean	e0_cil	e0_cih	eco_var	eco_cv	eco_mean	eco_cil	eco_cih
1	0.1810227...	0.4985440...	0.3631027...	0.3536781...	0.3725273...	0.0514221...	0.0803268...	0.6401616...	0.6374844...	0.6428388...
2	0.1621883...	0.4944369...	0.3280263...	0.3195823...	0.3364704...	0.0433088...	0.0788951...	0.5489425...	0.5466877...	0.5511973...
3	0.1427183...	0.6190119...	0.2305583...	0.2231279...	0.2379886...	0.0632516...	0.1789254...	0.3535081...	0.3502150...	0.3568012...
4	0.1415802...	0.5256736...	0.2693311...	0.2619599...	0.2767022...	0.0840575...	0.1850529...	0.4542348...	0.4498585...	0.4586111...

Figure 3-73 Example of Uncertainty Analysis Results

3.8.9.2 Sensitivity analysis

“Sensitivity Analysis” tool analyzes the sensitivity of each indicator value to the sustainability assessment results. This tool will calculate the first order sensitivity index

(Equation 3.32) and total effects sensitivity index (Equation 3.33) for each sustainability indicator, using the Sobol sensitivity analysis functions provided in SALib (Herman et al., 2017). GUI and parameters of this tool are shown in Figure 3-74.

The “AHP Weight File” is the indicators weight files generated by AHP tool. After selecting the “AHP Weight File”, the sustainability indicators used in the AHP process are automatically listed in the “AHP Indicators” table. “Sensitivity Results File” includes the first order sensitivity index, total effects sensitivity index, and their confidence interval for each indicator. If the confidence intervals of the indicators are larger than roughly 10% of the value itself, users may need to increase the sample size if computation permits.

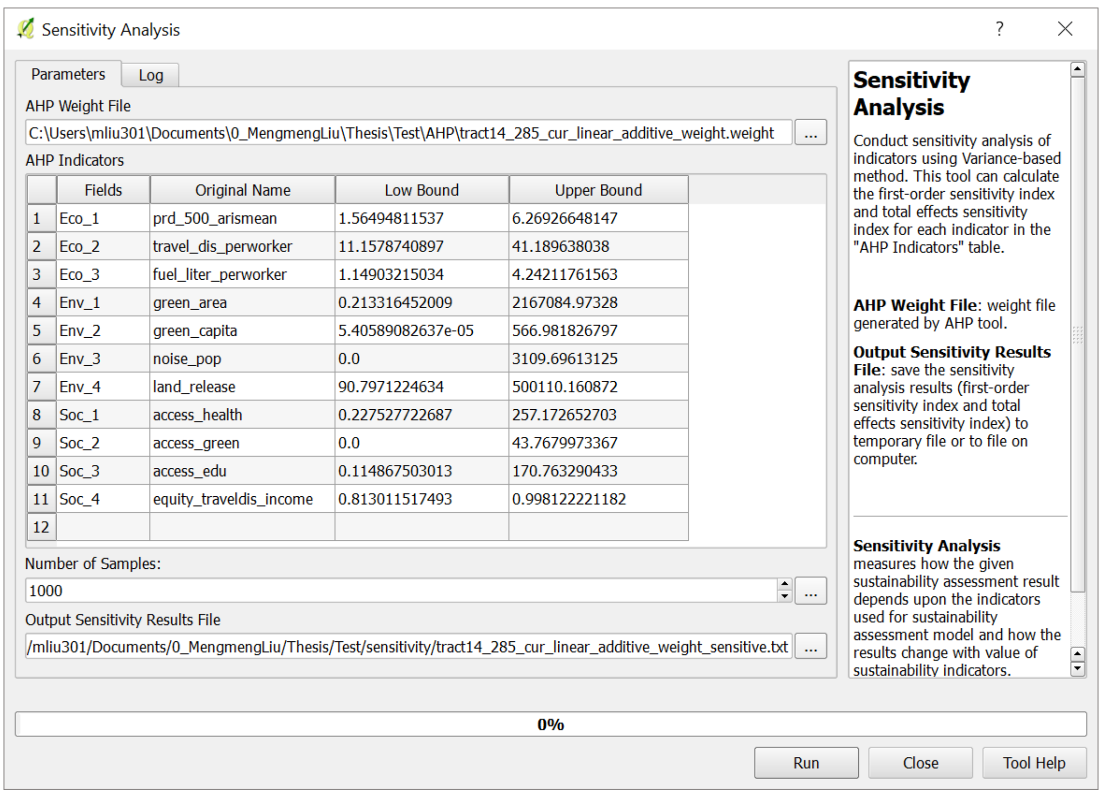


Figure 3-74 GUI and Parameters of Sensitivity Analysis Tools

3.8.10 *Assessment Result Visualization*

After finishing the "spatial sustainability assessment" and "Results Analysis" in the "Wizard", users can create their visualizations in the "Results Visualization" window (as shown in Figure 3-75). This component generates the selected visualization charts for each polygon in the "Uncertainty Result Layer", and save these charts to the folder defined in the "Directory to Save Charts" parameter. Four visualization charts are provided, including radar chart, sunburst chart, stratified bar chart, and line chart. This plugin also provides tools to generate each type of charts separately, which are illustrated in details as follows.

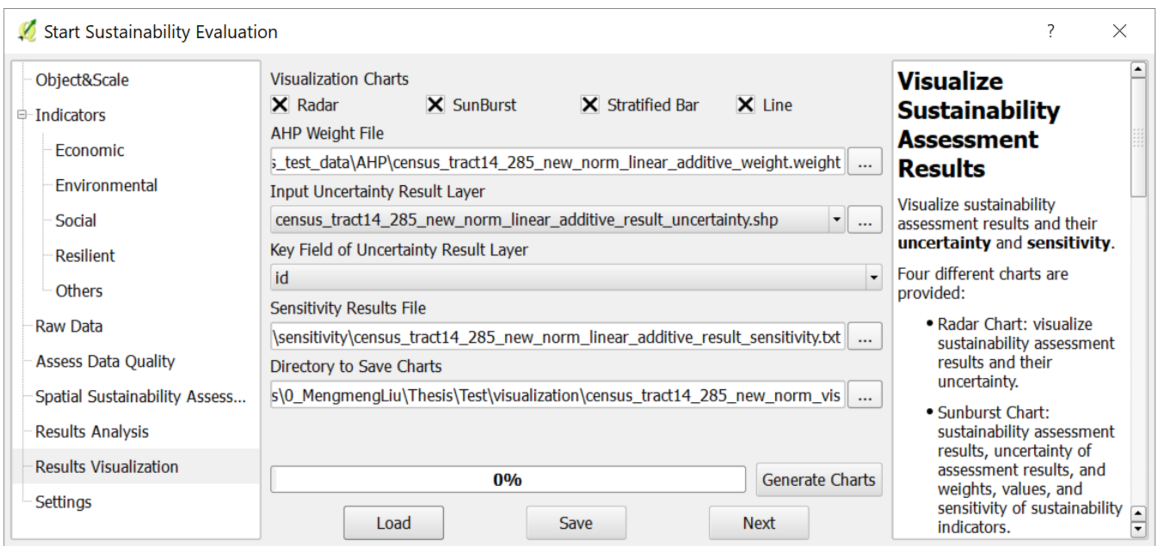


Figure 3-75 Results Visualization Component in Wizard

After creating these charts, users can open the "Check Visualization Results" tool (Figure 3-76) to check the charts for each polygon in the Input Layer by clicking on the corresponding polygon in QGIS window.

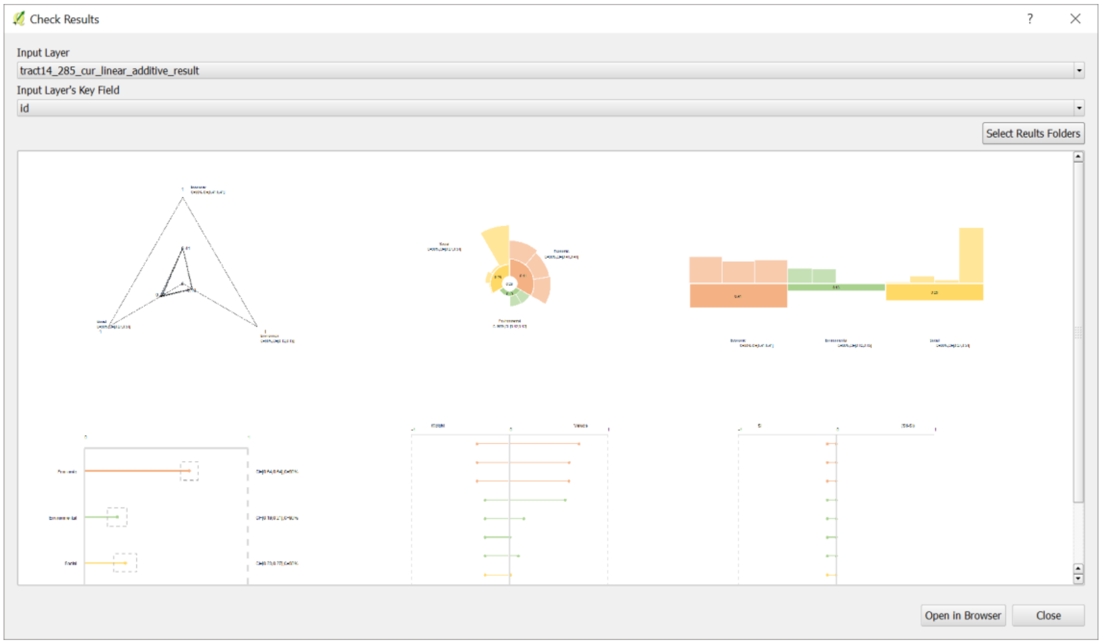


Figure 3-76 GUI of “Check Visualization Results” Tool

3.8.10.1 Radar Chart

“Radar Chart” tool creates a radar chart for each feature in the input layer. The radar chart can show the sustainability assessment results (CSI values) and uncertainty analysis result. GUI of “Radar Chart” tools is shown in Figure 3-77.

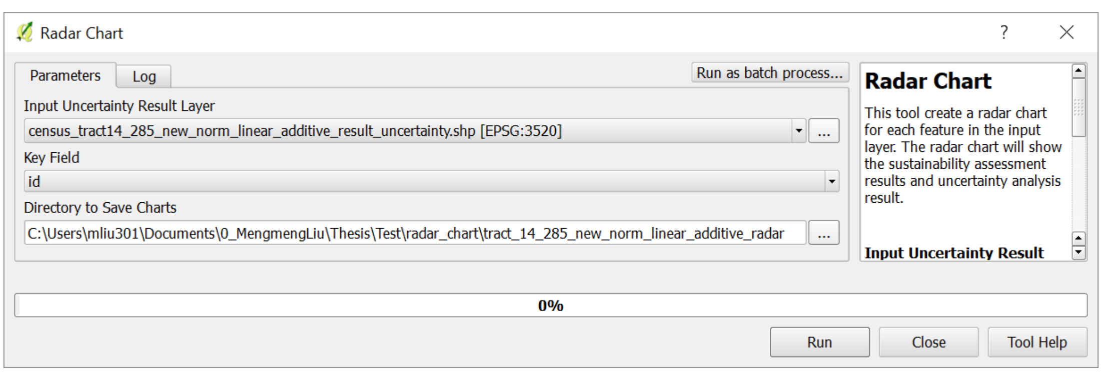


Figure 3-77 GUI and Parameters of Radar Chart Tool

3.8.10.2 Sunburst Chart

“Sunburst Chart” tools creates Sunburst chart for each feature in the input layer. The Sunburst chart shows the sustainability assessment results (CSI values), values and weight for each indicator, weight of each dimension, value of dimensional CSI and overall CSI, and uncertainty analysis results. GUI of “Sunburst Chart” tools is shown in Figure 3-78. Figure 3-79 shows some example sunburst charts generated by this tool.

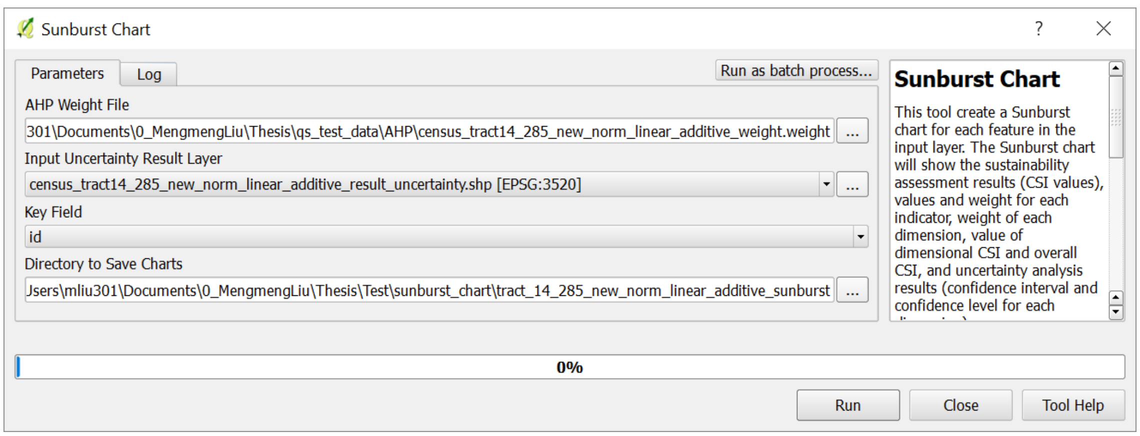


Figure 3-78 GUI and Parameters of Sunburst Chart Tool

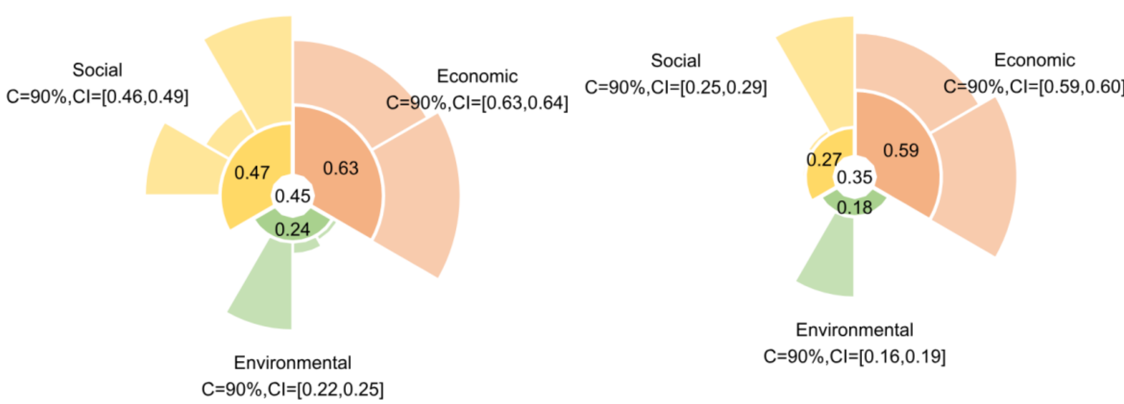


Figure 3-79 Examples Sunburst Charts Generated by Sunburst Chart Tool

3.8.10.3 Stratified Bar Chart

"Stratified Bar Chart" tool creates a stratified bar chart for each feature in the input layer. The bar chart will show the sustainability assessment results (CSI values) and weight of each dimension in the first layer, and show the values and weight for each indicator in each dimension in the upper layer. The uncertainty analysis results are labelled under the corresponding bar charts. GUI of "Sunburst Chart" tools is shown in Figure 3-80.

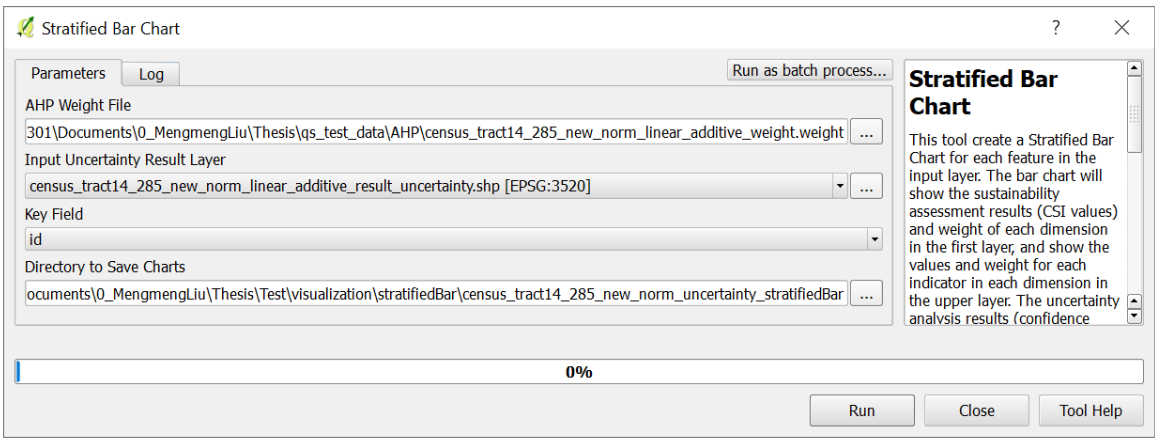


Figure 3-80 GUI and Parameters of Stratified Bar Chart Tool

3.8.10.4 Line Chart

"Line Chart" tool creates three line charts for each feature in the input layer. One shows the sustainability assessment results (CSI values) and their uncertainties (confidence level and confidence interval). One shows the values and weight of each indicator in each dimension. One shows the sensitivity analysis results for each indicator. GUI and parameters of this tool are shown in Figure 3-81.

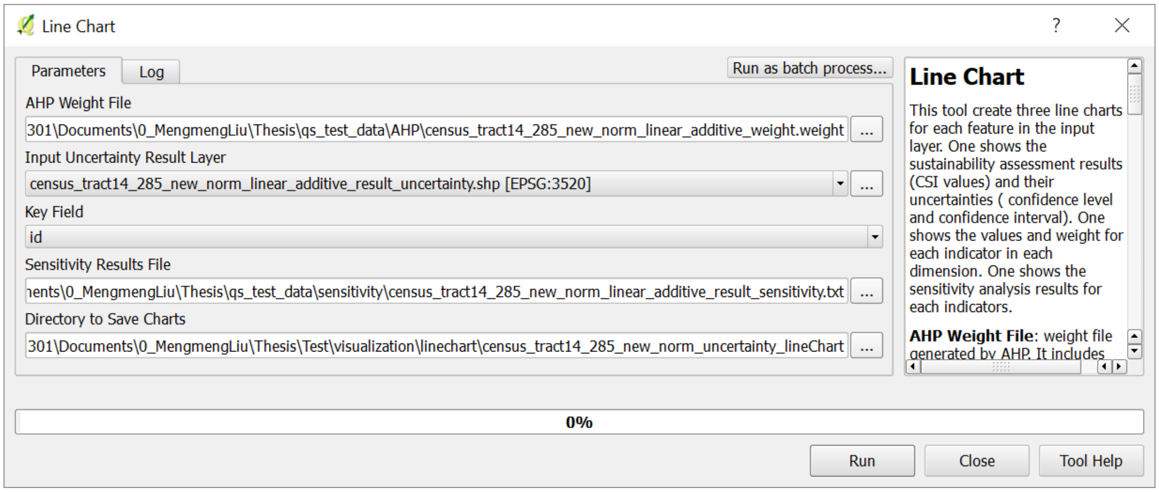


Figure 3-81 GUI and Parameters of Line Chart Tool

3.8.11 Data Exploration

The “Spatial Infrastructure Sustainability Assessment” plugin also provides tools to explore data before or during the process of spatial sustainability assessment.

3.8.11.1 Histogram

The “Histogram” tool plots the histogram of “Input Field” of the “Input Layer”. GUI and parameters of this tool are shown in Figure 3-82. After setting parameters and click the “Run” button, this tool will open a new window to show histogram (see Figure 3-83). In this window, users can save the plot to other formats (PDF, Png, Bmp,Tiff, and Jpeg).

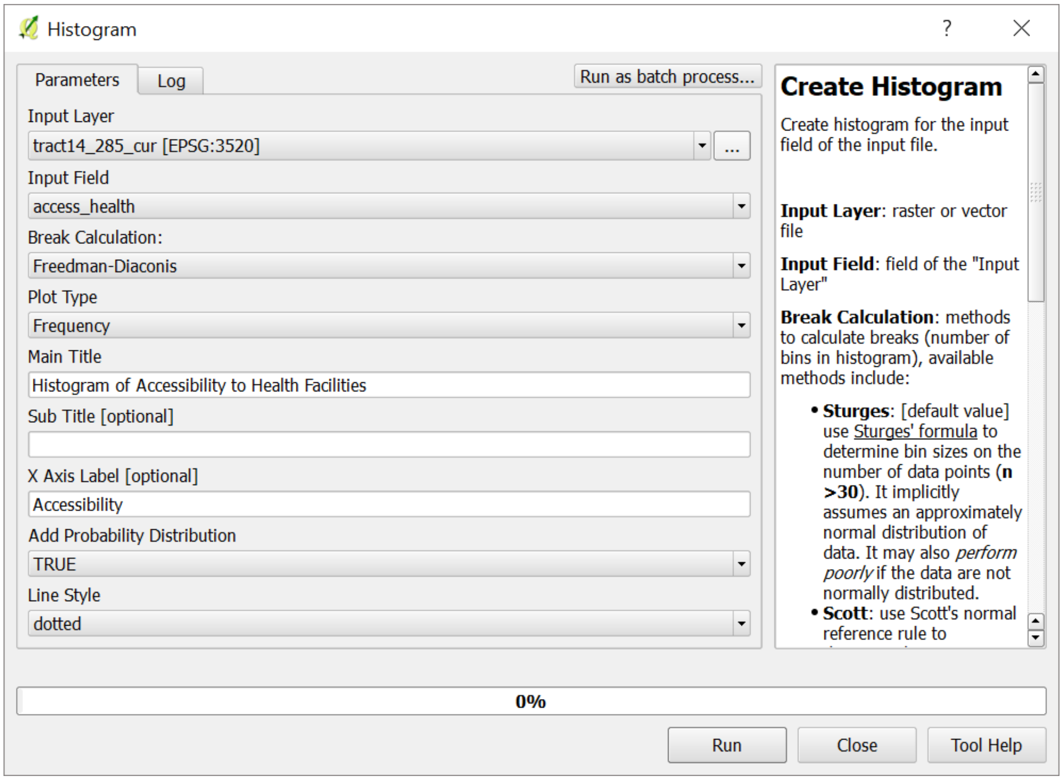


Figure 3-82 GUI and Parameters of Histogram Plot Tool

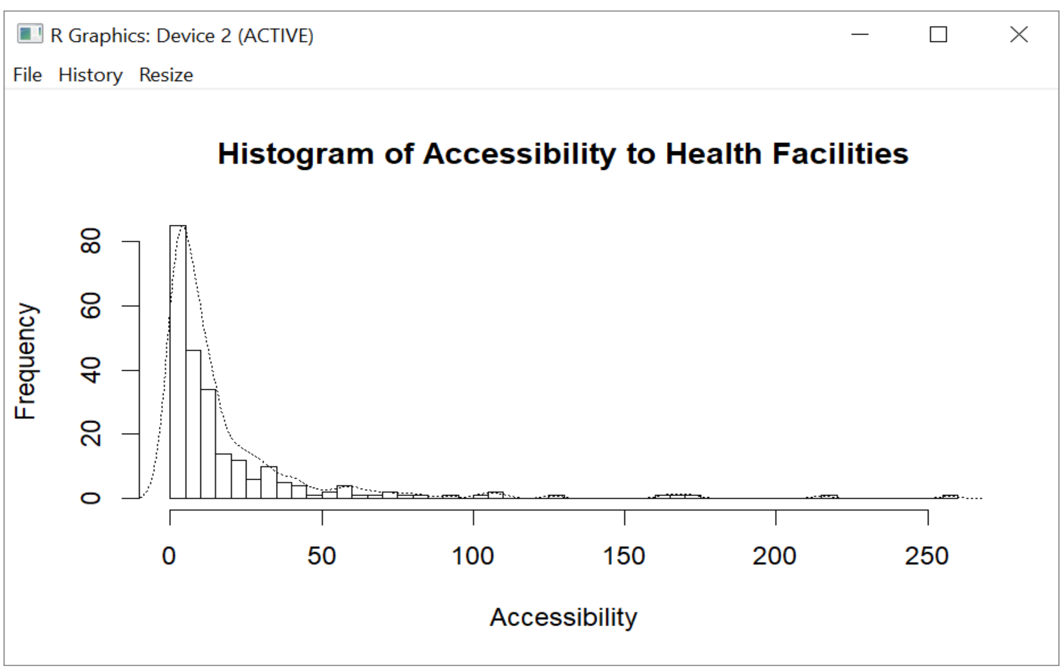


Figure 3-83 Example of Histogram Plot Using Parameters in Figure 3-82

3.8.11.2 Scatter Plot

The “Scatter Plot” tool creates scatter plot for the two “Input Fields” of the “Input Layer”. GUI and parameters of this tool are shown in Figure 3-84.

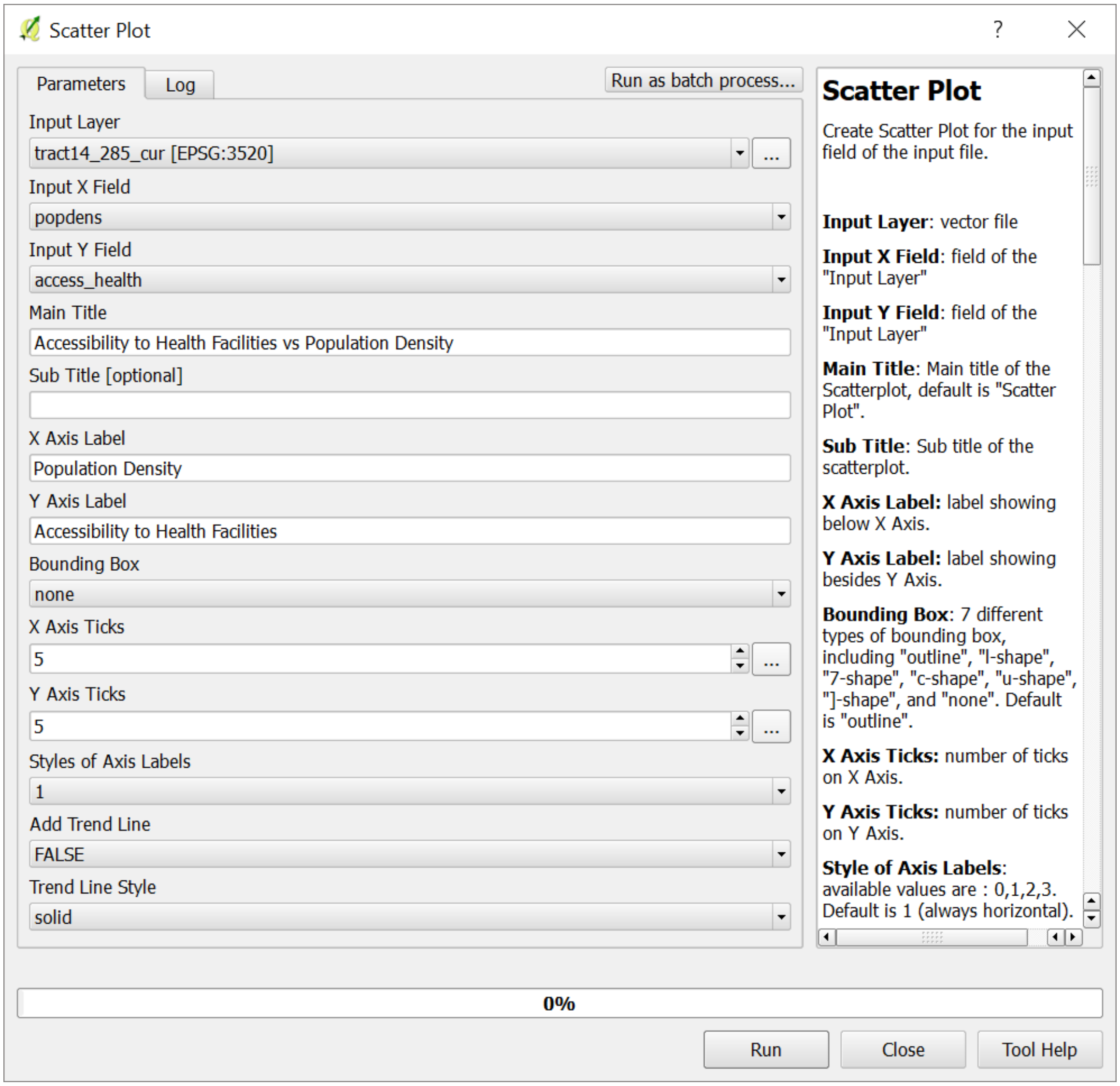


Figure 3-84 GUI and Parameters of Scatter Plot Tool

This tool gives users the flexibility to set the title, subtitle, X/Y axis labels, bounding box, X/Y axis ticks, styles of axis labels, and adding a trend line for the scatter plot. Users can see the help of this tool on the right side, or open help on the web browser by clicking the “Tool Help” button. After setting parameters and click the “Run” button, this tool will open a new window to show created scatter plot (see Figure 3-85). In this window, users can save the plot to other formats (PDF, Png, Bmp,Tiff, and Jpeg).

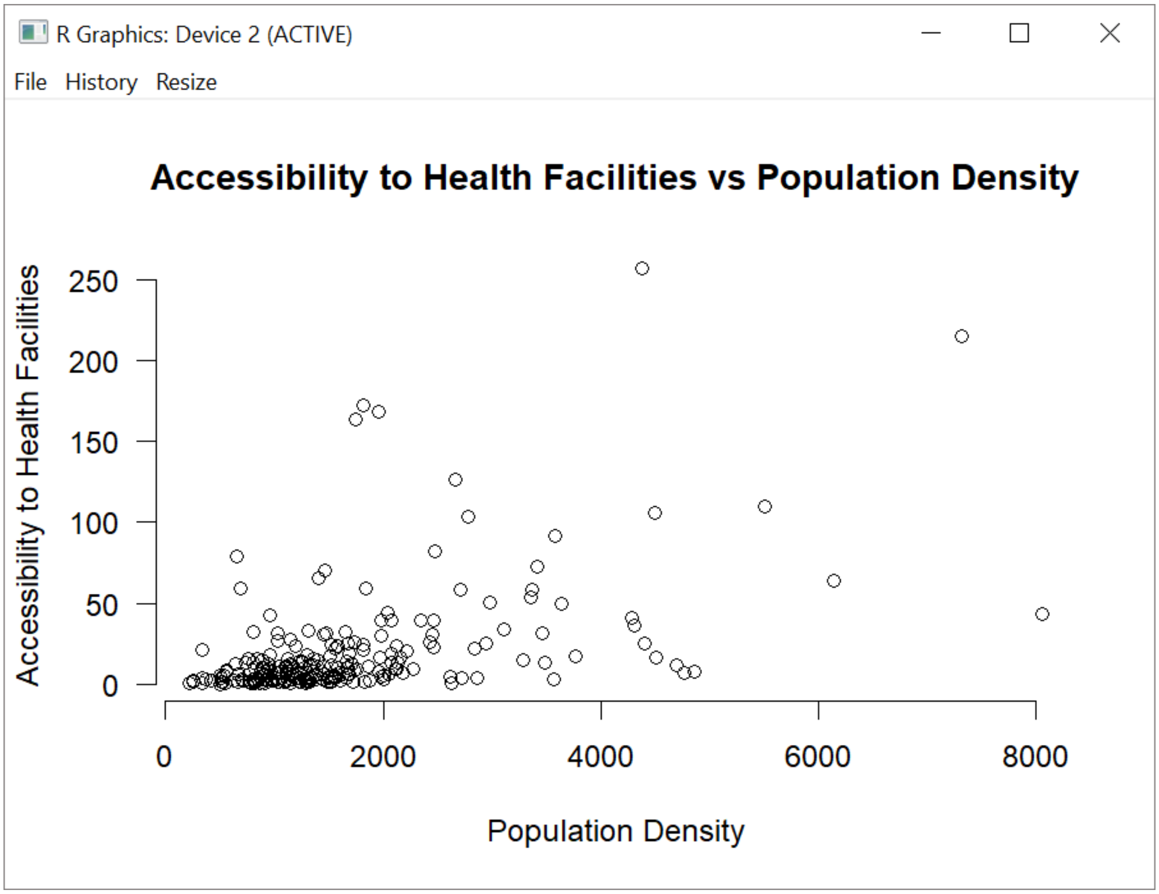


Figure 3-85 Example of Scatter Plot Using Parameters in Figure 3-84

3.8.11.3 Moran Scatter Plot

“Moran Scatter Plot” tool creates Moran scatter plot for one input field of the input layer. GUI and parameters of this tool are shown in Figure 3-86. The “Input Layer” can be layers currently opened in QGIS, shapefile (*.shp), or PostGIS table. “Weights file” describe the neighborhood relationship between features of the input layer (*.gal). It can be generated by "Create Spatial Weight" tool. Users can see help of this tool on the right side, or open help on web browser. After setting parameters and click “Run” button, this tool will open a new window to show created Moran scatter plot (see Figure 3-87). In this window, users can save the plot to other formats (PDF, Png, Bmp,Tiff, and Jpeg). An example of Moran Scatter Plot is shown in Figure 3-87.

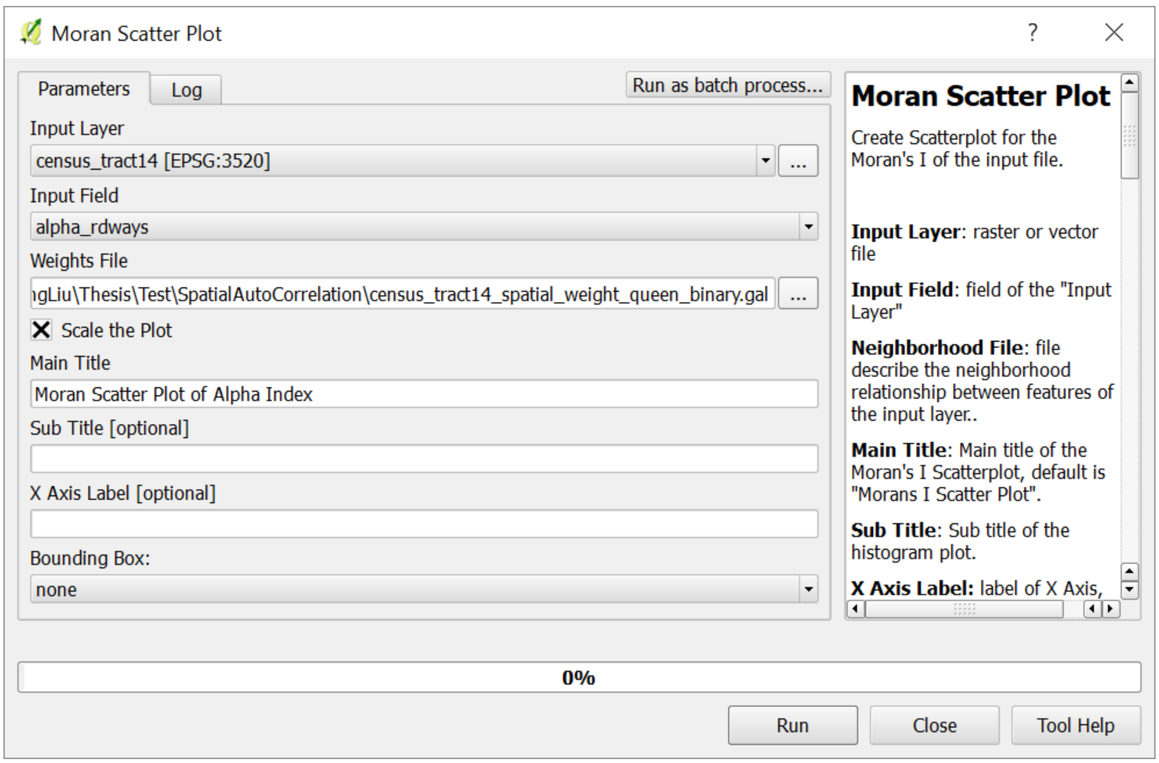


Figure 3-86 GUI and Parameters of Moran Scatter Plot Tool

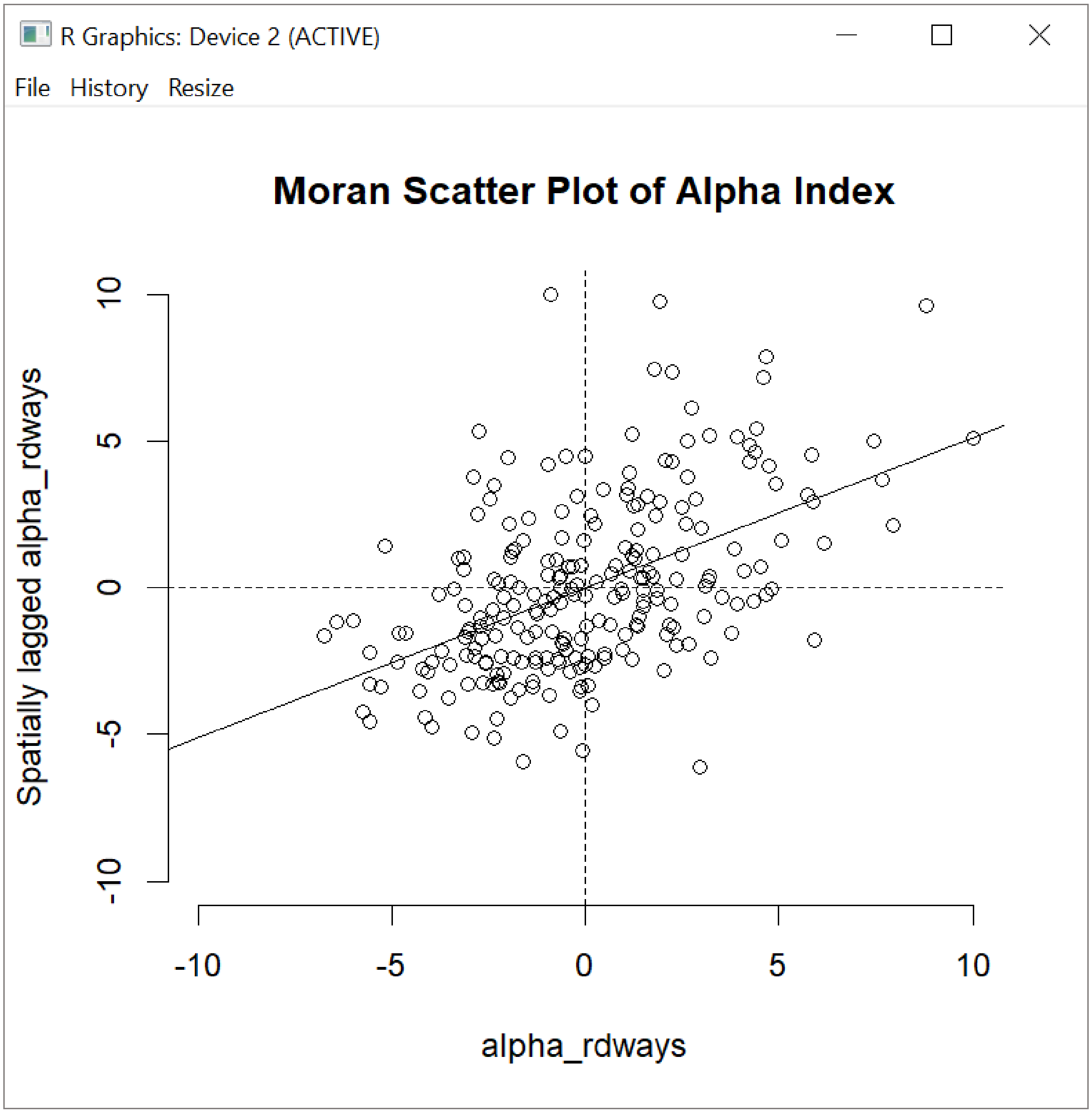


Figure 3-87 Example of Moran Scatter Plot

3.9 Summary

This chapter explained the proposed spatial sustainability assessment framework (Figure 3-1) in details. The proposed spatial sustainability assessment framework integrates DPSIR, TBL, LCA, and stakeholder-based methods. It can make full use of the advantages of each method and can present the sustainability assessment results from different aspects in multiple ways. It not only considers all aspects of sustainability (environmental, economic, social, and resilience), but also includes inter-linkages between different aspects and their dynamic changes in the process of sustainability assessment. This framework provides integrative and systematic methods to assess the sustainability of infrastructure at different spatial and temporal scales. It engages policymakers and other participants in the process of sustainability evaluation, by letting them set the weight of indicators with the help of AHP. The proposed framework can visualize the results in multiple ways (e.g., radar chart, sunburst chart, stratified bar chart, line chart, and tile grid maps), so that it can assist decision-makers to determine which actions should or should not be taken in an attempt to make their objective sustainable. It also has a spatial database to save and manage the data, indicators and other information used in the process of sustainability assessment. The proposed framework is a very general data-driven framework and can be easily used in other disciplines and other areas.

There are eight components in this proposed framework, including "Define Object & Scales", "Select Indicators", "Select Data", "Assess Data Quality", "Data Exploration", "Spatial Sustainability Assessment", "Results Analysis", and "Results Visualization & Interpretation". The order of components indicates the process of spatial sustainability assessment. The first component of this framework is "Define Object & Scales", for example define type of infrastructure, its location, its timeline, and the spatial and temporal scales of sustainability assessment. The scale determines the level of detail and complexity of the assessment work. Sustainability in the framework has four dimensions, which are environment, economy, social, and resilience. After the definition of sustainability assessment object and scales, the user can choose sustainability indicators to be used for the following sustainability assessment. In this framework, the users can use DPSIR, PCA, or life cycle thinking to select indicators from the spatial sustainability assessment database or to define indicators by themselves.

After selecting indicators, users can choose the data they used to evaluate each indicator and to check the quality of each data. The framework provides functions to interpolate missing values, to detect duplicate records automatically and to detect or fix the inconsistency among datasets. If the data quality is good enough, the users can then use the data to evaluate each sustainability indicators. The “spatial sustainability assessment” component includes all the functions that can be used by users, including spatial analysis, network analysis, indicator pre-analysis (normalization, correlation, and PCA), weighting indicator (by AHP or PCA), and aggregating indicators (to build the CSI). After evaluating each sustainability indicator or build the CSI, the framework provides systematic tools to analyze the uncertainty and sensitivity of the sustainability assessment results. The last component of this framework is “Assessment Results Visualization”. In this framework, users can visualize the sustainability assessment results in the radar chart, sunburst chart, stratified bar chart, line chart, or tile grid maps. Finally, a sustainability assessment report will generate including all the information about the sustainability assessment. It includes object, scales and dimensions of the sustainability assessment, the list of indicators, data used for the sustainability assessment and its qualities, correlation matrix of indicators, normalization method of indicators, weighting of indicators, weighting and aggregation methods, sustainability assessment results, uncertainty, sensitivity, and visualizations of sustainability assessment results, and the flow chart of sustainability assessment.

The proposed spatial sustainability assessment framework is implemented as a plugin in QGIS (namely “Spatial Infrastructure Sustainability Assessment”). All the spatial analysis and other data editing and analysis functions provided in QGIS can be seamlessly used in the framework. The Spatial Infrastructure Sustainability Assessment” plugin is easy to use, adaptable, general, reusable, and has homogenous GUI styles. The author also provides users as many information of this plugin as possible, such as comments, tool tips, tool helps, and other related information (e.g., reference paper of algorithms and source code referred when implementing the framework), therefore users can easily understand and use this plugin to evaluate the sustainability of their project. In section 3.8, the author illustrates all the tools provided in the “Spatial Infrastructure Sustainability Assessment” plugin in details. Most of the designed functions in the proposed spatial sustainability assessment framework are implemented, except DPSIR, LCA, automatic data quality check and fix, ANP, sensitivity analysis of models with correlated factors, and converting conventional geographic map to tile grid map. However, there is still much space to improve the implemented tools in the “Spatial Infrastructure Sustainability Assessment” plugin.

CHAPTER 4. PILOT STUDY: SUSTAINABILITY ASSESSMENT OF MOVING INTERSTATES UNDERGROUND

Transportation infrastructure (e.g., roads, embankments, bridges, railroads, airports, docks, and canals) serves as pathways for the mobility of people, good and services, and forms the backbone of cities and communities. The construction, operation, and rehabilitation of these transportation infrastructure facilities put significant strain on the local, state and federal agencies (Puppala and Bheemasetti, 2018). For example, they can affect land-use patterns and the demography of communities. They can also deplete natural resources and contribute to the global carbon footprint and greenhouse gas emissions. Onat et al. (2014) claimed that the U.S. transportation sector is an important source of GHG emissions and energy consumption, it is responsible for 67% of the total U.S. petroleum consumption.

The transportation infrastructure in the U.S. faces several pressures, including obsolescence, rapid industrialization, population growth, human-induced disturbances, durability issues, age and so on (Puppala and Bheemasetti, 2018). It is necessary to promote the sustainable development of transportation infrastructure. There are many definitions of "Sustainable transportation" (Duncan and Hartman, 1996; Jordan and Horan, 1997; Jabareen, 2006). In this chapter, a sustainable transportation system makes full use of resources (e.g., energy, land use, and so on) and limits emissions and waste within the area's ability to absorb (Duncan and Hartman, 1996). It operates transportation services at maximum efficiency and supports a vibrant economy. It also provides equitable access for people and their services (e.g., schools, hospitals, good quality of the environment, and so on) and helps achieve a healthy and desirable quality of life in each generation (Jordan and Horan, 1997; Jabareen, 2006).

The interstate highway system within the Atlanta Perimeter plays an important role in the residents' daily lives. They are the most heavily traveled roadways in the city. However, the interstate highway system greatly disfigures and disconnects the city. Moving all interstate highways within the Perimeter underground can increase the connections and communications between different sub-communities, especially for those that are separated by current interstate highways. With Atlanta's great granite geological condition (Galloway, 2017), it is reasonable and practical to build tunnels for interstates underground. There are Atlanta transportation plans for underground transportations, such as Atlanta North-South Tunnel, which would connect GA 400 to I-675 S at I-285 Perimeter south, and the Northwest and Northeast Tunnels to Buckhead (Georgia Globe Design News, 2017). "Over the 75 years between 2023 and 2097, the Atlanta North-South Tunnel provides cumulative (traffic) delay savings of over 2.8 billion hours." (Galloway, 2017). The Stitch project plan to construct a ¾-mile platform capping over the downtown connector, extending from the Civic Center MARTA station at West Peachtree Street to Piedmont Avenue (Kelley, 2019).The Stitch project requires around four years of pre-development and six-years to buildout and will cost $452 million (Kelley, 2019). Buckhead Community Improvement District plans to build a "Park over GA400", which creates nine acres of elevated green space in Buckhead.

Moving interstate highways underground provides many positive benefits. It can reduce energy costs, environmental pollution and noise, and the maintenance costs of vehicles (Daily Sabah, 2016). It can also free-up surface space for recreation and other purposes. For example, in the Stitch project, after capping the interstates connector at downtown Atlanta, around 14 acres of new buildable space will be created on top of the platform after capping interstates; space can be used for five-acre park, homes, office buildings, hotels and more (Kelley, 2019). Moving interstates underground or capping the interstates can also increase the real estate values of the areas close to interstates, and also catalyze the redevelopment of underutilized properties. For example, the Stitch project is estimated to create $1.1 to $3.1 billion values, generate $21 to $58 million in new revenue, and increase the city's bonding capacity by $308 to $847 billion (Atlanta Downtown, 2019). Besides, underground highways produce less noise and can keep or improve the connections between communities adjacent to interstate highways. Thus it can improve the quality of life for people who live in these communities. For example, the "Park over GA400" will provide vast and unique regional connectivity, and improve livability and economic viability in the Buckhead Area (Buckhead Community Improvement District, 2019). Besides, there are fewer limits on lane width and in transforming underground highways. For example, it is possible to increase the number of lanes to meet current traffic needs and help relieve future congestion. It is also feasible to add new MARTA line or light rail in the tunnel build for underground interstates, or even to build a 3-D tunnel network. Furthermore, it provides uninterrupted travel even in bad weather conditions such as fog, snow, and ice (Daily Sabah, 2016). Moreover, the construction of tunnels can also protect Atlanta's historic neighborhoods and magnificent tree canopy.

However, with these benefits of moving interstates underground, it is necessary to evaluate its impact on the neighborhoods or communities and its sustainability conditions. Therefore, in the case study, we analyze the sustainability of moving interstates within the

Perimeter underground, to see if moving interstates underground is more sustainable comparing with the current situation.

4.1 Study Area

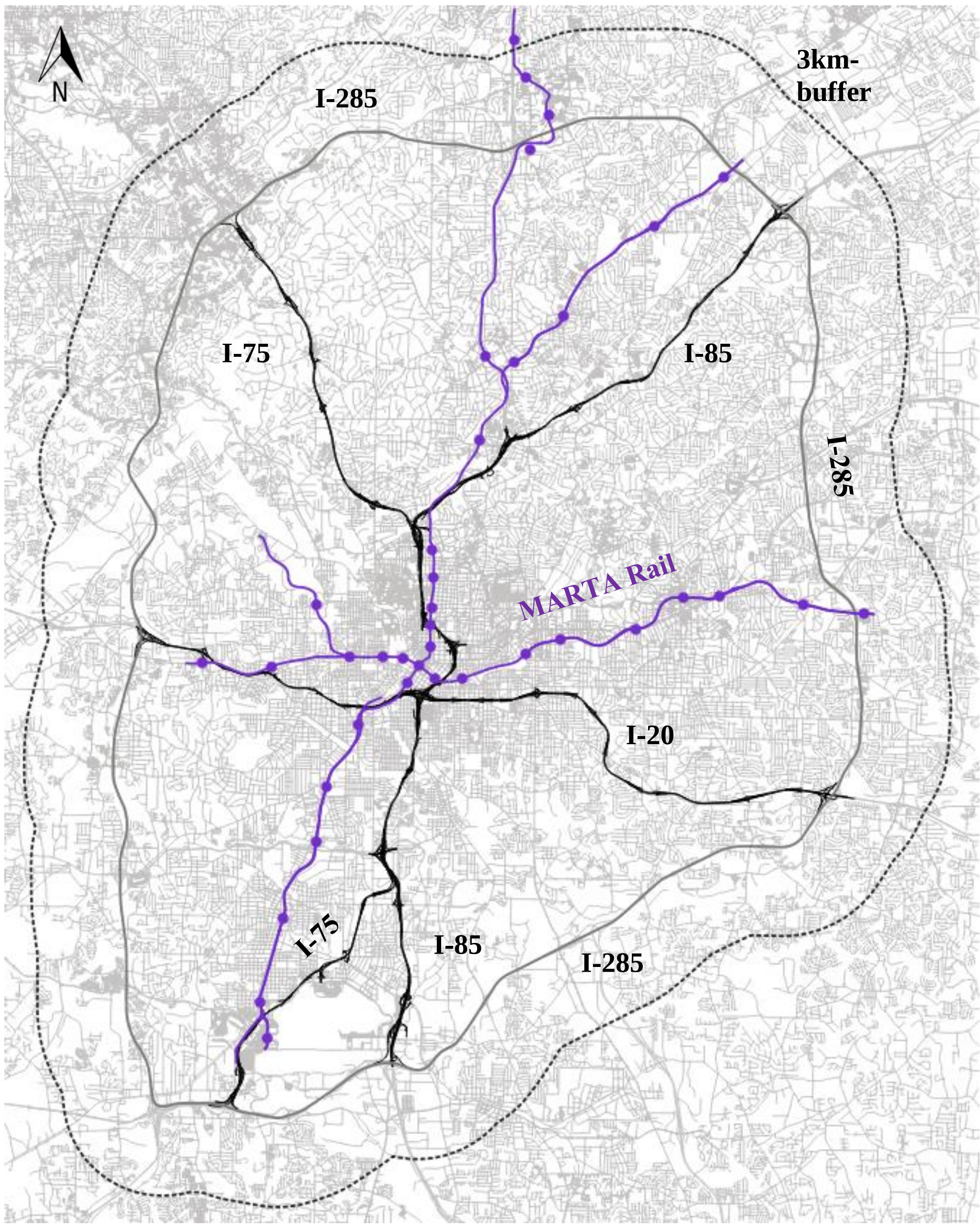


Figure 4-1 Current Road Network in Study Area

Study Area is the region within 3 km buffer of the Atlanta Perimeter (I-285), which is the dotted polygon on the map in Figure 4-1. The black lines are Interstate highways within the Perimeter, which are part of I-75, I-85, and I-20. Purple lines are the MARTA rail lines, the purple dots on the line are the rail stations.

4.2 Data and Pre-Process

All spatial data used in the study are listed in Table 4-1. Details of each data set are discussed as follows. After preprocessing, all the data use the same CRS (i.e., EPSG:3520).

Table 4-1 Data Sets for Pilot Study

Data Set	Sources	Date
Road Network	Open Street Map	2016
	DeKalb-Fulton Metropolitan Traffic Ways Plan Map	1952
	Aerial Images	1952
Social Economic Data	United States Census Bureau	2014
Green Space	ARC (Atlanta Regional Commission, 2017[a])	2016
	US parks (ESRI National Park Service, 2017)	2017
	Google Satellite Images (Google, 2017)	2017
Water Surface	ARC (Atlanta Regional Commission, 2017[b])	2001
	Google Satellite Images (Google, 2017)	2017
Health Facilities	Yellow Pages (Yellow Pages, 2017)	2017
Education Facilities	Yellow Pages (Yellow Pages, 2017)	2017
Traffic OD Flow	United States Census Bureau	2014
Household Travel Survey	Atlanta Regional Commission (ARC)	2011
Census Tract Boundary	Census Bureau	2014

4.2.1 Road Network

Road network data (see Figure 4-1) is downloaded from Open Street Map (2016). It provides the number of lanes of each road and its categories, which includes interstate highways (freeways), arterials, collectors and local roads. The data need to be pre-processed (see Figure 4-2) before using it for the following analysis.

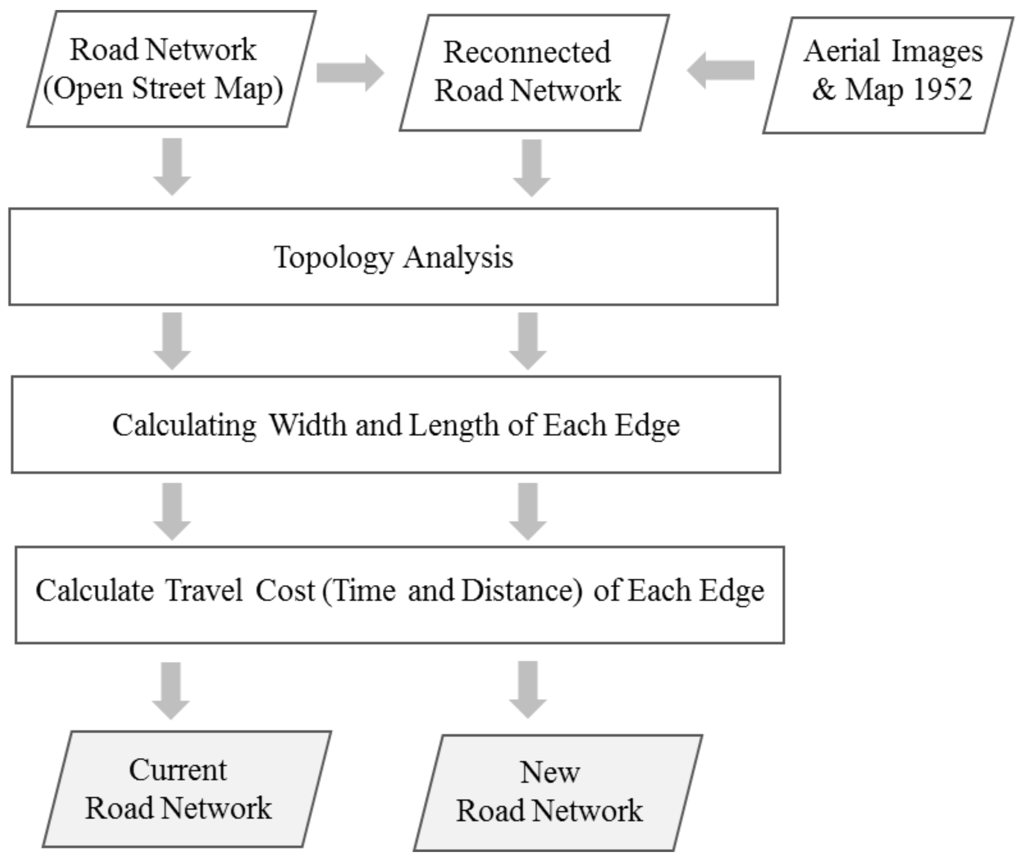


Figure 4-2 Pre-Processing of Road Network Data

The author first conduct topology analysis using PostGIS topology extension for the current road network. After topology analysis, intersections of the road are extracted as nodes in the network, road sections connecting the intersections are extracted as edges in the network. The relationship between edges and nodes are also created. Each edge has attributes that indicate its length, number of lanes, width, geometry, and direction (double way or one-way direction). Therefore, the network with topology now can represent the real road network in Atlanta. Then we calculate the width and length of each edge (see Equation 4.1). Finally, we estimate the travel cost (including time cost and distance cost) of each edge (see Equation 4.2) and generate the "current road network" which can be directly used for the following analysis (e.g., network analysis, traffic flow and commute time estimation).

$$\text{Width of Road Segment } e \; width(e) = \begin{cases} (\text{lanes} + 1) \times 3.7 + 3.0 + 1.2 & \text{if } e \text{ is interstates and lanes} < 3 \\ (\text{lanes} + 1) \times 3.7 + 3.0 + 3.0 & \text{if } e \text{ is interstates and lanes} \geq 3 \\ \text{lanes} \times 3.0 & \text{if } e \text{ is not interstates and lanes } \neq NULL \\ 2 \times 3.0 & \text{if } e \text{ is not interstates and lanes} = NULL \end{cases} \tag{4.1}$$

$$\begin{gathered} \text{DistCost}(e) = Length(e) = length\ of\ Road\ Segment\ e \\ \text{TimeCost}(e) = Length(e)/SpeedLimt(e) \end{gathered} \tag{4.2}$$

According to the interstate highway standards, the minimum lane width for most US and state highways is 12 feet (3.7 meters). Minimum outside paved shoulder width of 10 feet (3.0 meters) and inside shoulder width of 4 feet (1.2 meters). With three or more lanes in each direction, the inside paved shoulder should be at least 10 feet (3.0 meters) wide. According to the urban street design guide (National Association of City Transportation Officials, 2017), the lane width of local streets is 3 meters (10 feet). Therefore, we can calculate the width of each road segment using Equation 4.1. Note that here we only get the lower bound of interstate width, the real width is larger than the estimation given by Equation 4.1. The minimum width (for two directions) of interstates is

23.2 meters, maximum width is 86 meters, medium width is 30.6 meters, and the average width is 37.9 meters.

To examine the benefits of moving the interstate highways below ground, we also create a “new road network” using process in Figure 4-2. We first use the 1952 aerial images (Hampton, 2014) for the Vine City district of Atlanta (see Figure 4-3) and the DeKalb-Fulton Metropolitan Traffic Ways Plan map in 1952 (see Figure 4-4) to reconnect the surface roads in the present day layout to match the historical connections. This action is feasible if the interstate highways are hypothetically placed below ground. The newly added roads are mainly in the downtown and midtown Atlanta, and the airport area (see the red lines in Figure 4-5). After manually reconnecting these roads, a new road network is obtained, it is referred to as “reconnected road network” (see Figure 4-6, the red lines are the newly added roads).

We then conduct a topology analysis of the “reconnected road network”. After fixing all topology errors in the road network, we calculate the width and length of each edge in the road network, using Equation 4.1. We assume that all of the new connected roads have the same number of lanes, which is two lanes, one lane in each direction. And we assume all the new connected roads belong to "Local Roads", which has an average design speed of 40 km/h (US Department of Transportation Federal Highway Administration, 2014). Finally, we estimate the travel cost (including time cost and distance cost) of each edge (see Equation 4.2) and then get the “New Road Network” which can be used directly for the following analysis. The “New Road Network” we obtained for this study is shown in Figure 4-6.

Figure 4-3 Aerial Image for the Vine City District of Atlanta

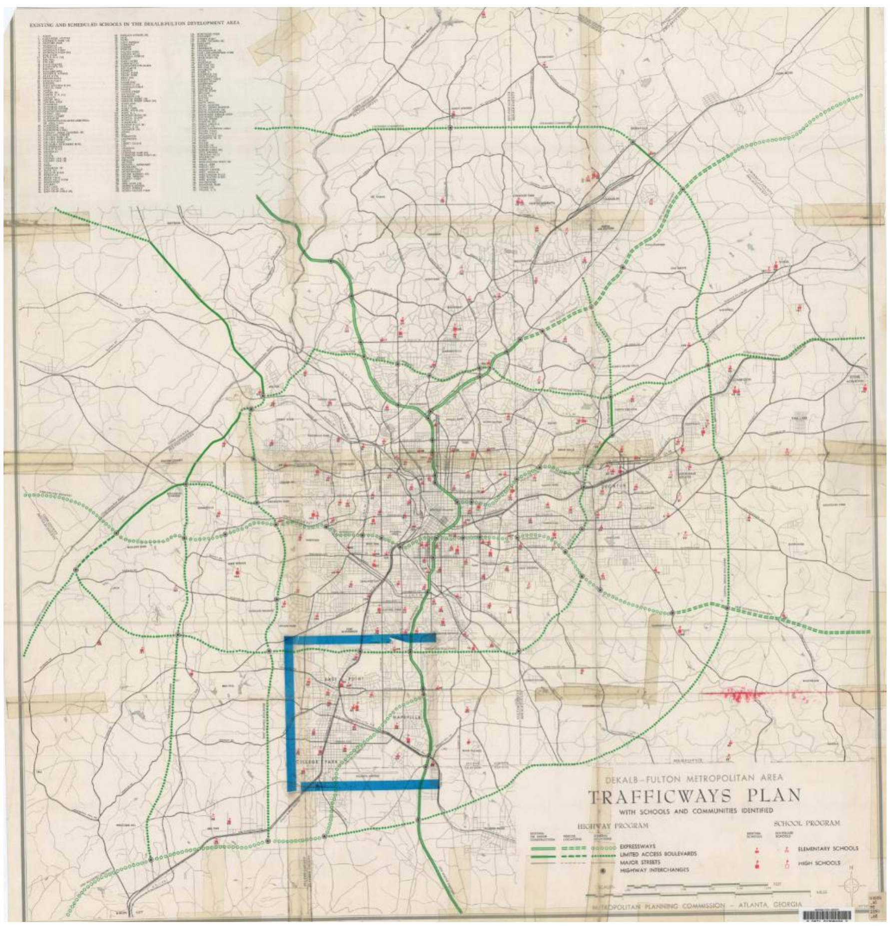


Figure 4-4 DeKalb-Fulton Metropolitan Traffic Ways Plan Map in 1952

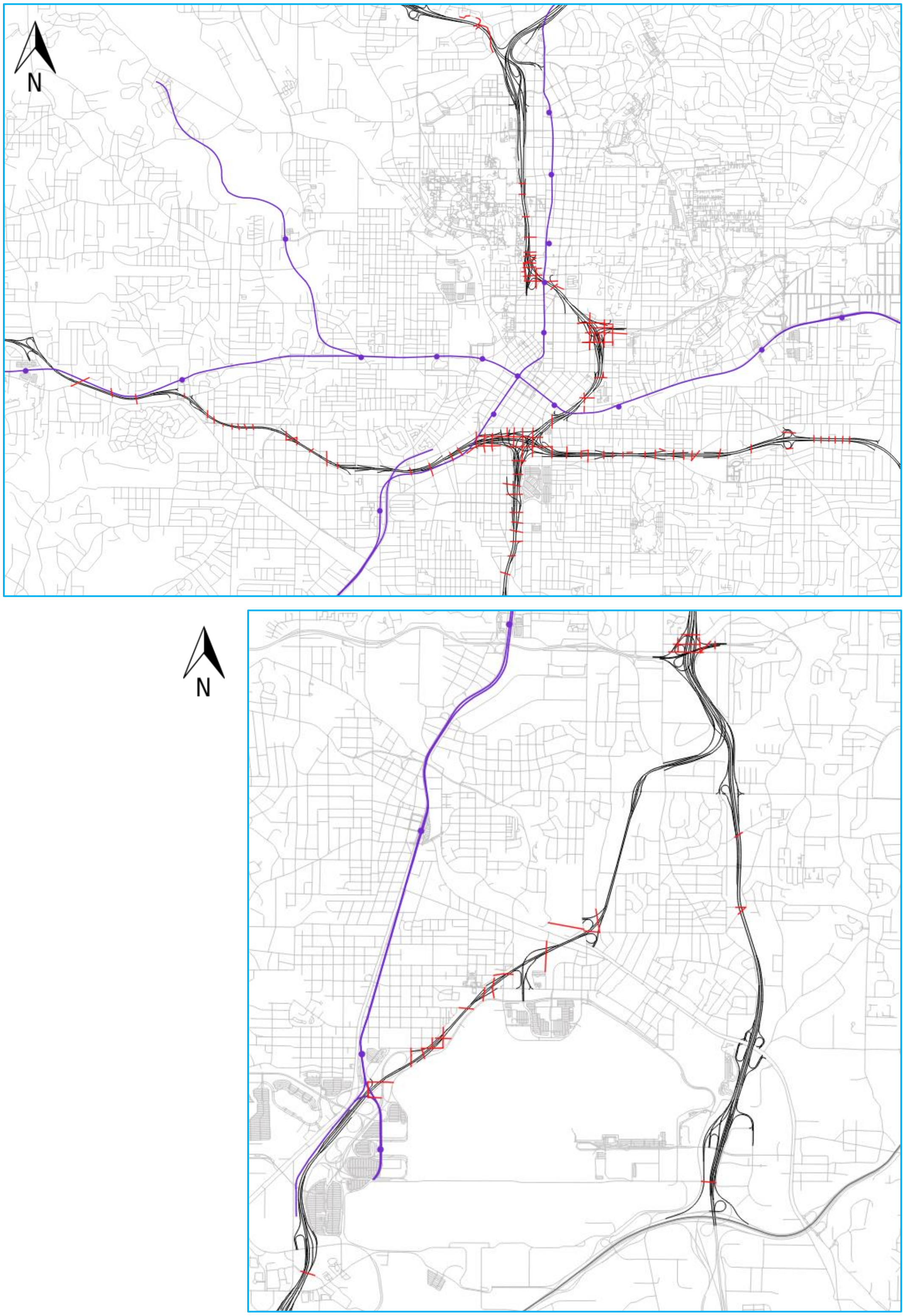


Figure 4-5 New Added Roads in Downtown, Midtown, and Airport Areas

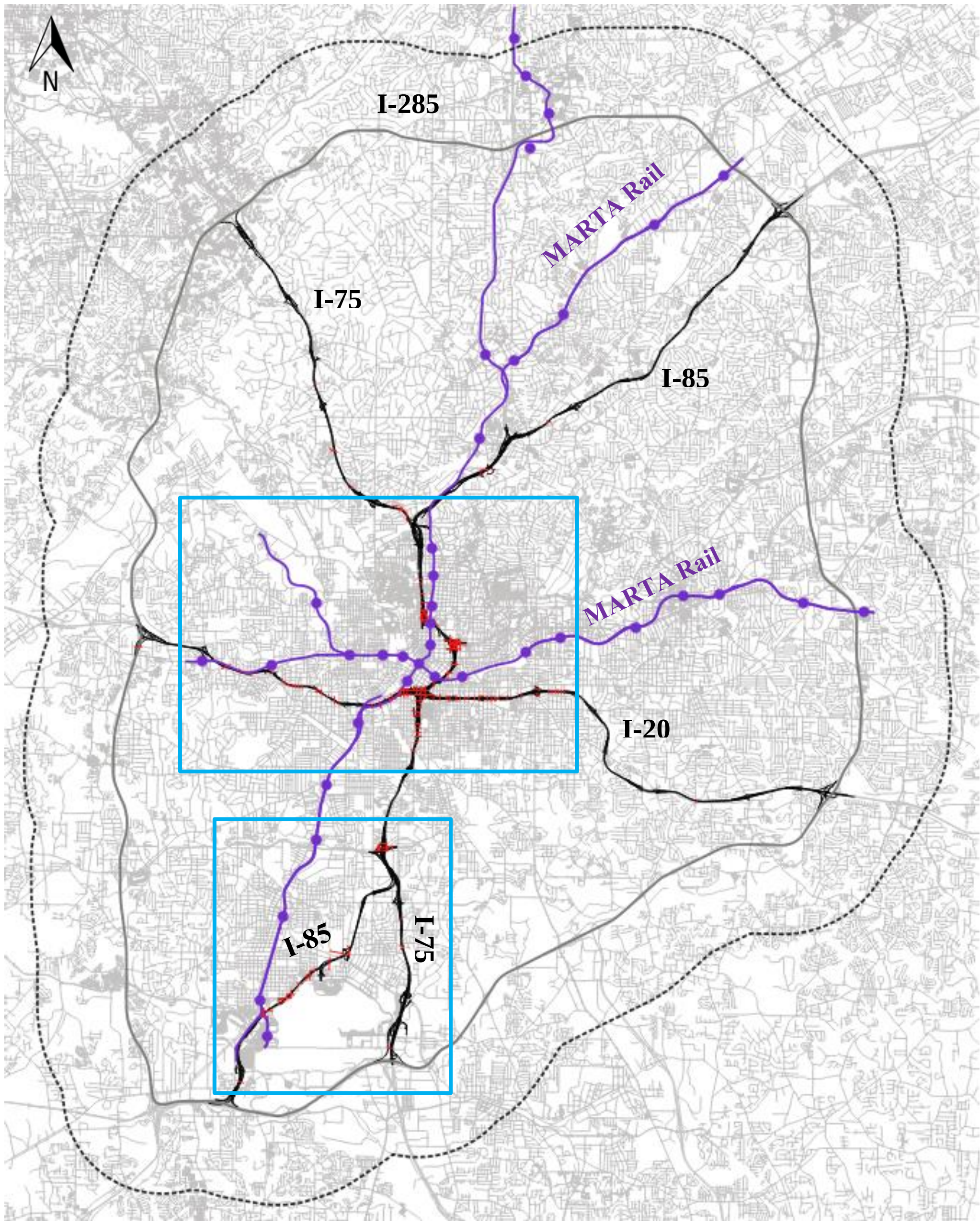


Figure 4-6 New Road Network

4.2.2 Social Economic Data

The population and income data for each census tract in the study area are downloaded from "Census Bureau". The spatial distributions of the total population are shown on the left map in Figure 4-7.

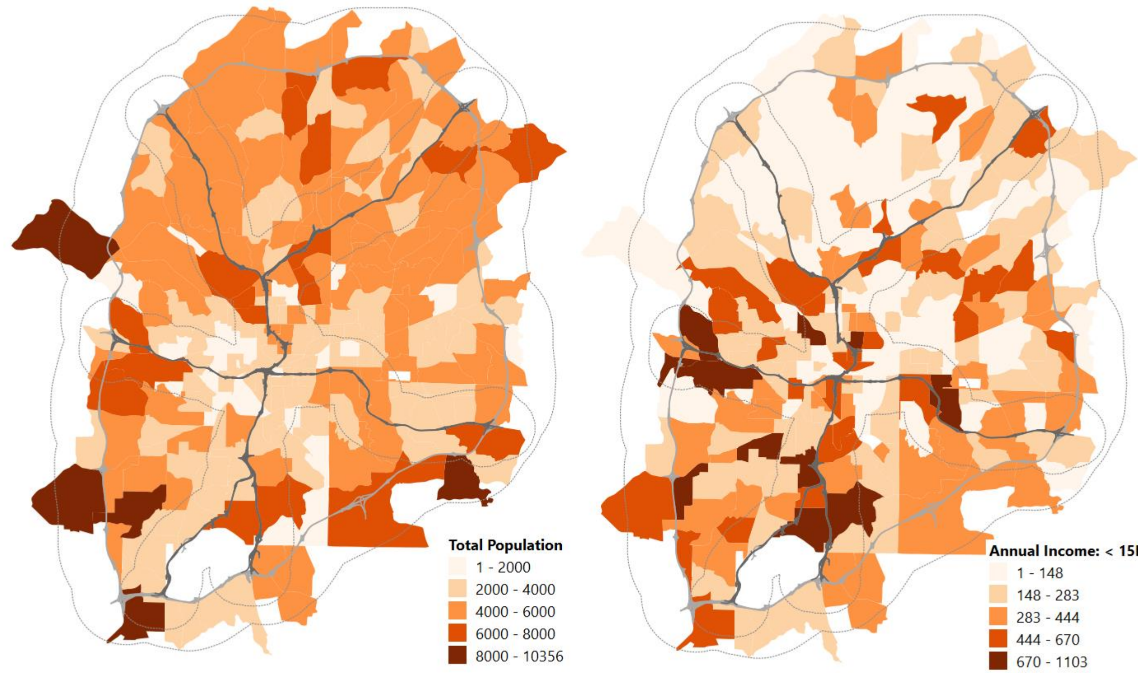


Figure 4-7 Total Population (Left) and Population with Annual Income < 15K (Right)

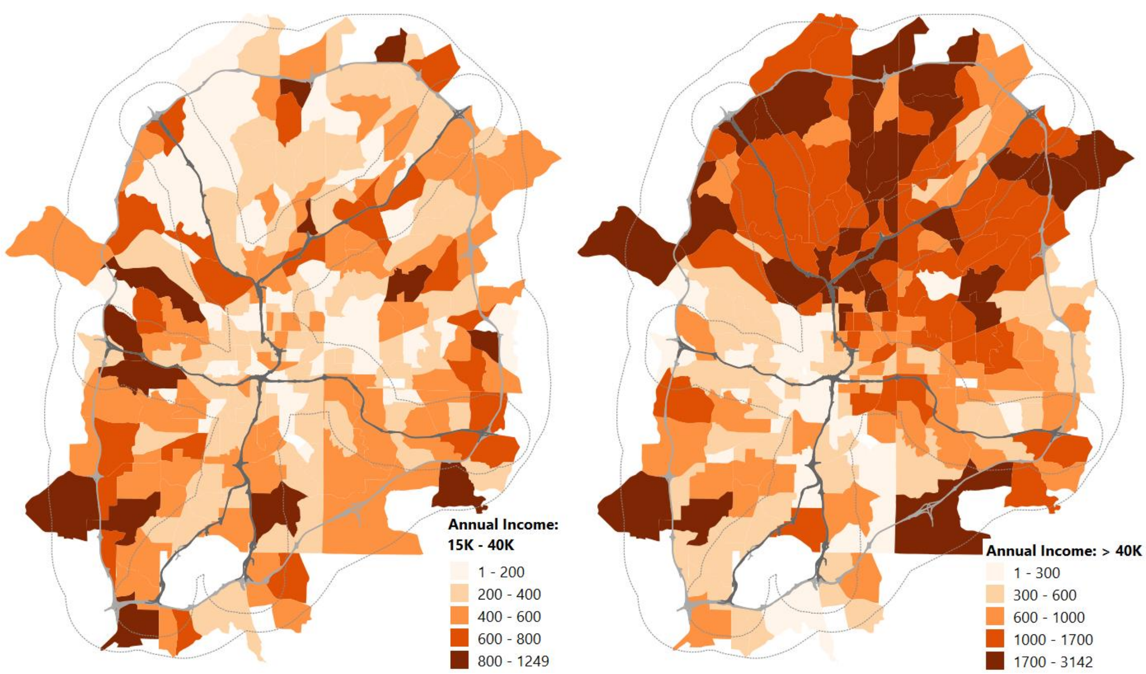


Figure 4-8 Population with Different Annual Income

There is more population to the north of I-20 within 1-285. There is less population living in the 2-km buffer of I-20 and south section of I-85 comparing with the north section of I-75 and north section of I-85. The right map in Figure 4-7 shows that most low-income population are located in the south of I-20. However, most of the high-income population are located in the north of I-20 (see the right map in Figure 4-8). There is no obvious pattern of the middle-income population, but they tend to be located around I-285.

4.2.3 *Green Space*

Green space provides many environmental benefits (Haq, 2011; Lee et al., 2015), such as negating urban heat, offsetting greenhouse gas emissions through CO_2 absorption, reducing the energy costs of cooling buildings, minimizing air, water, and noise pollution, attenuating stormwater, and preservation of biodiversity and nature conservations. Besides green spaces can also provide economic benefits, such as maintaining or increasing property values and financial returns for land developers (Heidt and Neef, 2008), attracting tourists and creating new jobs (Haq, 2011), and reducing the energy costs of cooling buildings (details see (Heidt and Neef, 2008)). Furthermore, it also brings us social benefits, like providing urban residents spaces for physical activity, social interactions and cultural purpose activities (Haq, 2011; Lee et al., 2015), increasing physical and psychological wellbeing, decreasing stress level of urban citizens, and so on (Haq, 2011).

Green space data is downloaded from the ARC (Atlanta Regional Commission, 2017[a]). It includes all parks and protected green space in the 20-county Atlanta Region, which is updated in 2016. The author verifies and updates the ARC green space data by Parks data obtained from US parks (ESRI National Park Service, 2017) and Google satellite images (Google, 2017). Distribution of green space in the study area is shown in Figure 4-9. Their total areas are 37.9 km^2 (square kilometers). There are many small green spaces in the places with a population less than 1000, and there are relatively large green spaces in the census tracts with more than 400 0population. Overall, the green space is evenly distributed in the study area.

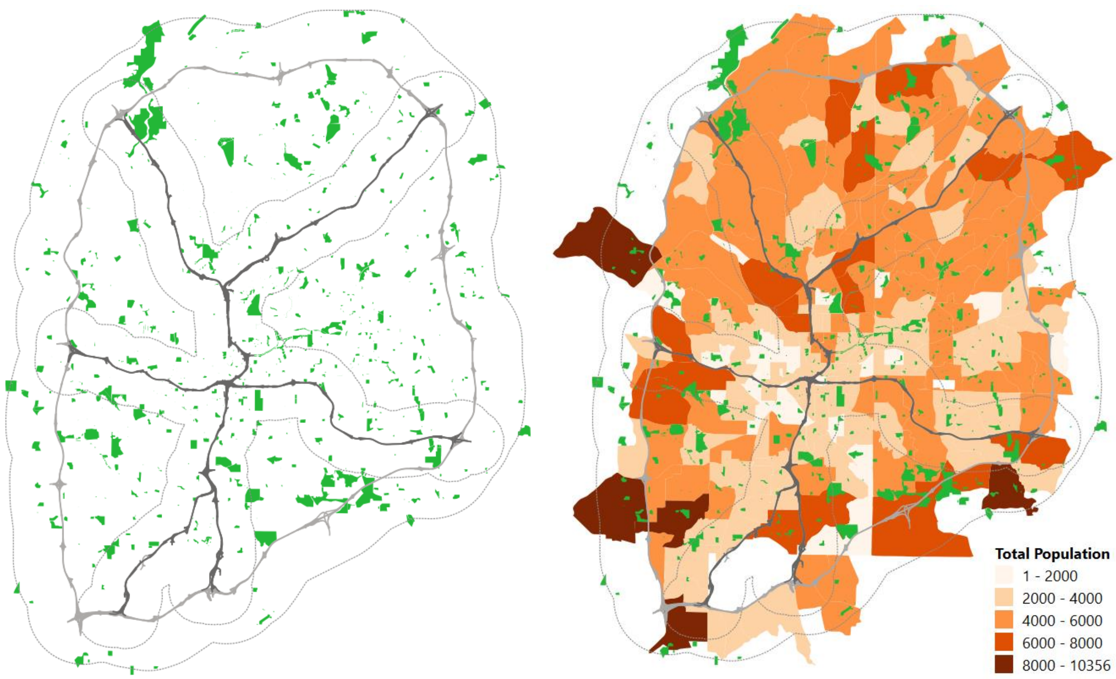


Figure 4-9 Current Distribution of Green Space

After moving interstates underground, the surface areas occupied by them will be released. However, the planning of the newly released space is a very complex city planning problem and need to consider many factors. Here we assume that all the released spaces will turn into green space except the space covered by the new connected roads. The map of new green space in the study area is shown in Figure 4-10. Their total areas are 45.5 km^2, increased by 7.6 km^2 comparing the original green space in Figure 4-9. And the released areas of interstates work as green corridors, which play important roles in maintaining high biodiversity in urban areas, by providing mobility axes for species (Heidt and Neef, 2008).

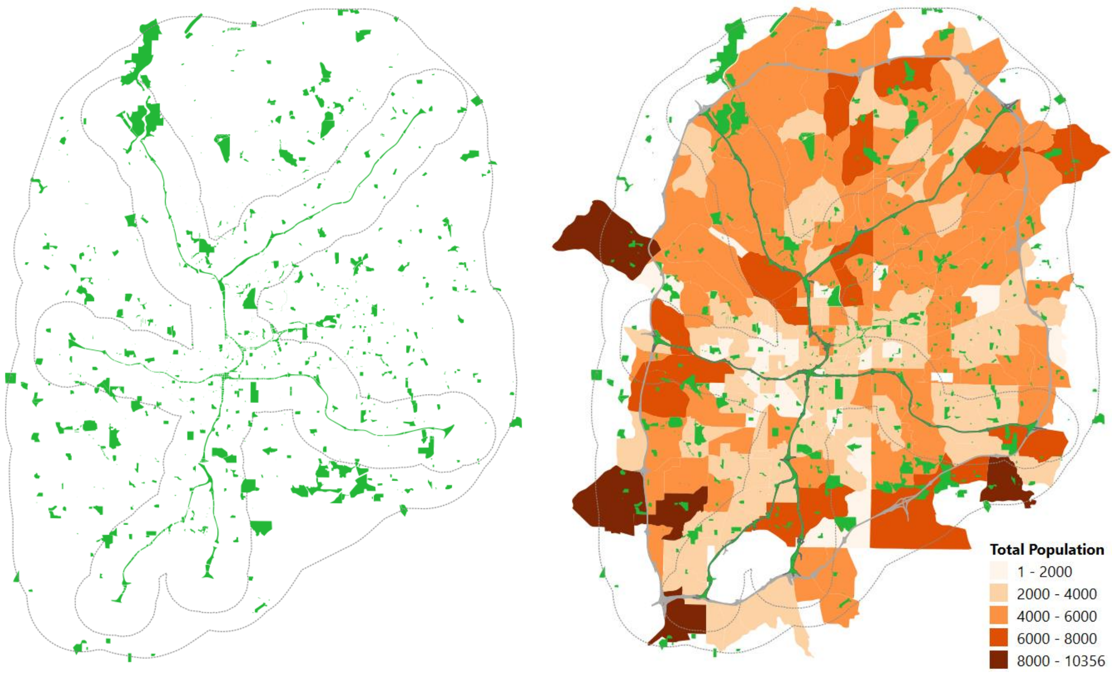


Figure 4-10 New Distribution of Green Space After Moving Interstates Underground

4.2.4 Water Surface

The water surface data is downloaded from the ARC (Atlanta Regional Commission, 2017[b]). The dataset contains polygonal hydrographic features including rivers, lakes, ponds, reservoirs, swamps, and islands. The author modifies or adds some new water surface polygons based on Google Satellite Image. The map of the water surface in the study area is shown in Figure 4-11. The total water surface area is 9.87 km^2. Most water surfaces are located outside of I-285, and the water surfaces that are within I-285 tends to located in the north of I-20.

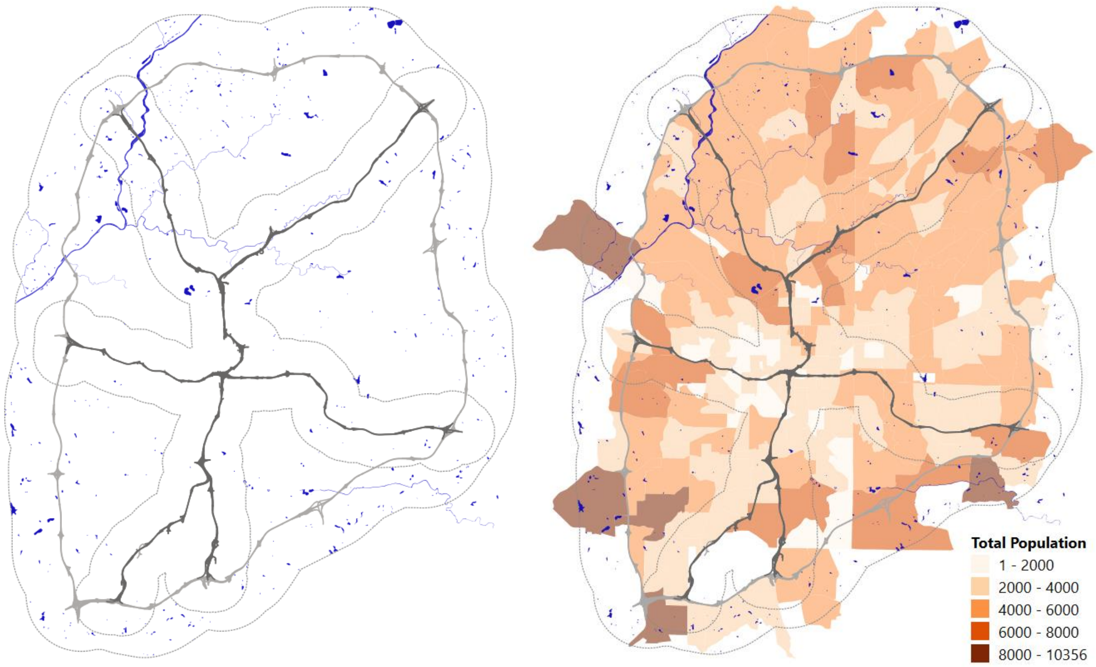


Figure 4-11 Distribution of Water Surface in the Study Area

4.2.5 Health Facilities

The health facilities used in the pilot study is download from Yellow Pages (2017). It includes all the hospitals, clinics, and emergency rooms in the study area. Map of health facilities in the study area is shown in Figure 4-12. There are 1861 health facilities in the study area. Most health facilities are located within a 2-km buffer of interstates. There are more health facilities in the downtown and mid-town Atlanta, Decatur county, and Sandy Springs, and they tend to cluster together.

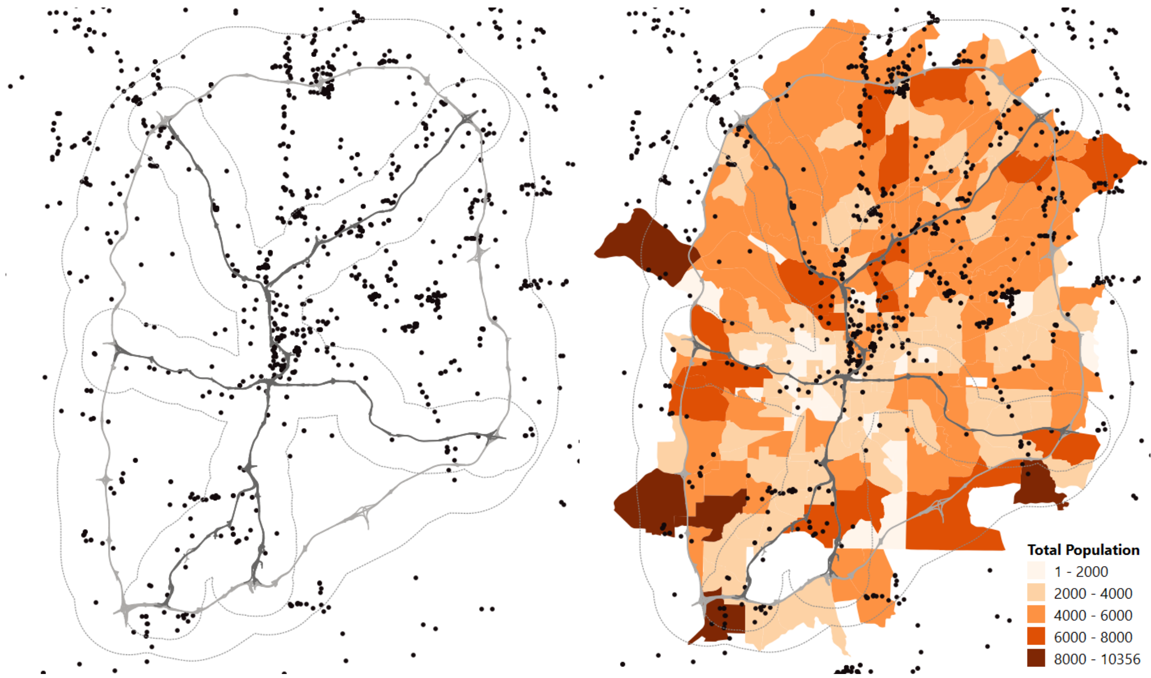


Figure 4-12 Distribution of Health Facilities in the Study Area

4.2.6 Education Facilities

The education facilities used in this study is download from Yellow Pages (2017). It includes all the elementary schools, middle school, high schools, colleges, universities and institutes in the study area. Map of education facilities in the study area is shown in Figure 4-13. There are 1190 education facilities in total in the study aera. Education facilities have a similar pattern with the heath facilities. But they are more evenly distributed in the study area comparing to heath facilities.

Figure 4-13 Distribution of Education Facilities in the Study Area

4.2.7 OD Flow

The author downloaded the Longitudinal-Employer Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) datasets in raw form as a set of comma separated variables (CSV) text file from United States Census Bureau (2017). The data set is in version LODES7. It includes all jobs with both workplace and residence for each 2010 census block in the state of Georgia in 2014. Detail information about the data is shown in Table 4-2. For the analysis in the pilot study, the author only considers the data which has workplace within 3km buffer of the region in -285 and homeplace within 80 km radius region of the center of the study area (see Figure 4-14). Finally, there are 891,519 OD pairs at the census block level in the study area, and they are referred to as “OD flow in the study area”. The total number of workers in each census block is shown in Figure 4-15. We can see that most workplaces are located along or near interstates, and the north of I-20 has more work opportunities than south of I-20. The total number of workers whose homeplace in each census block is shown in Figure 4-16, and most of them are located within the 80km radius region. Therefore, the "OD flow in the study area" can represent the real home to workplace flow in the study area. The number of workers live in each census block within I-285 are shown in Figure 4-17. We can see that most workers live outside or near I-285.

Table 4-2 Information Provided by LODES Dataset

Variables	Type	Explanation
w_geocode	Char15	Workplace Census Block Code, e.g., "131210094022007"
w_geocode	Char15	Homeplace Census Block Code, e.g., "131350507253000"
S000	Num	Total number of jobs
SA01	Num	Number of jobs of workers age 29 or younger
SA02	Num	Number of jobs of workers age 30 to 45
SA03	Num	Number of jobs of workers age 55 or older
SE01	Num	Number of jobs with earnings $1250/month or less
SE02	Num	Number of jobs with earnings $1251/month to $3333/month
SE03	Num	Number of jobs with earnings greater than $3333/month
SI01	Num	Number of jobs in Goods Producing industry sectors
SI02	Num	Number of jobs in Trade, Transportation, and Utility industry sectors
SI03	Num	Number of jobs in All Other Services industry sectors
createdate	Char	Date on which data was created, formatted as YYYYMMDD

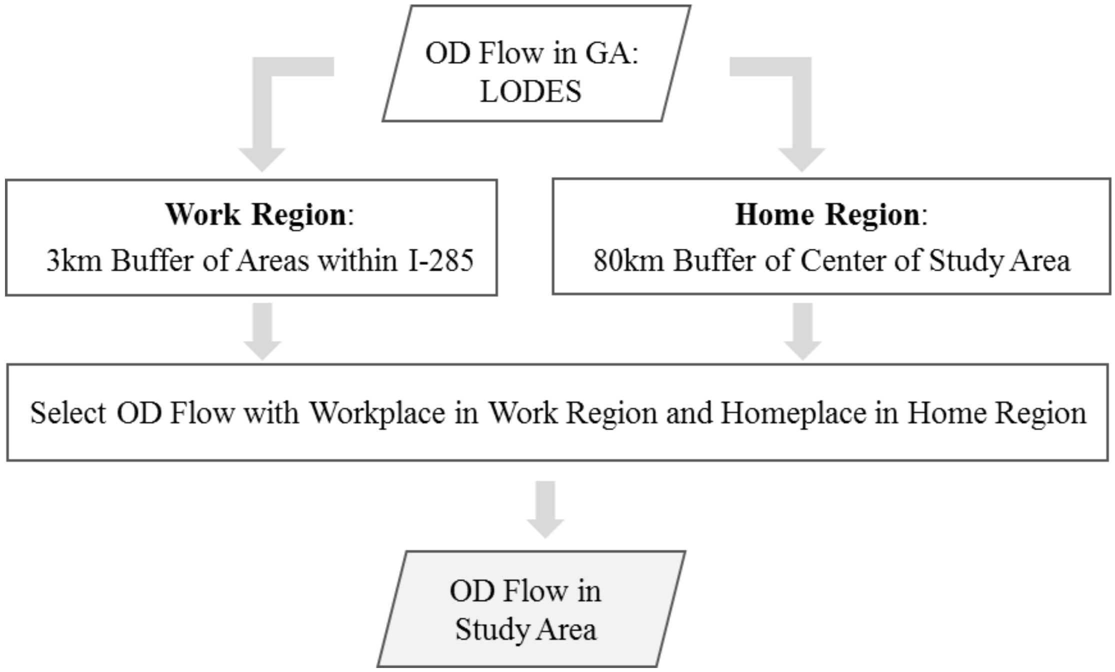


Figure 4-14 Pre-Processing of OD Flow Data

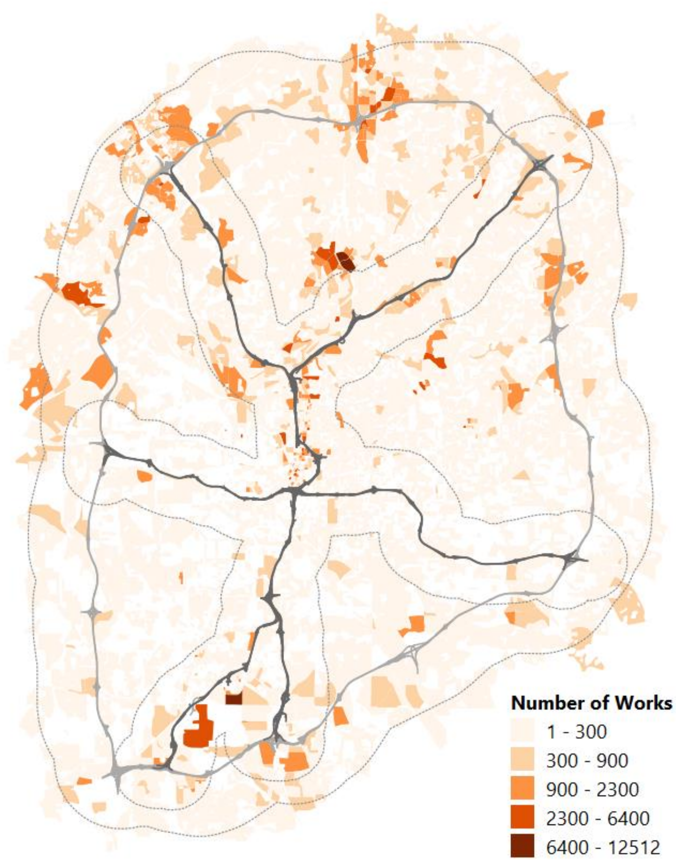


Figure 4-15 Work Places Distribution within 3km Buffer of I-285 (Census Blocks)

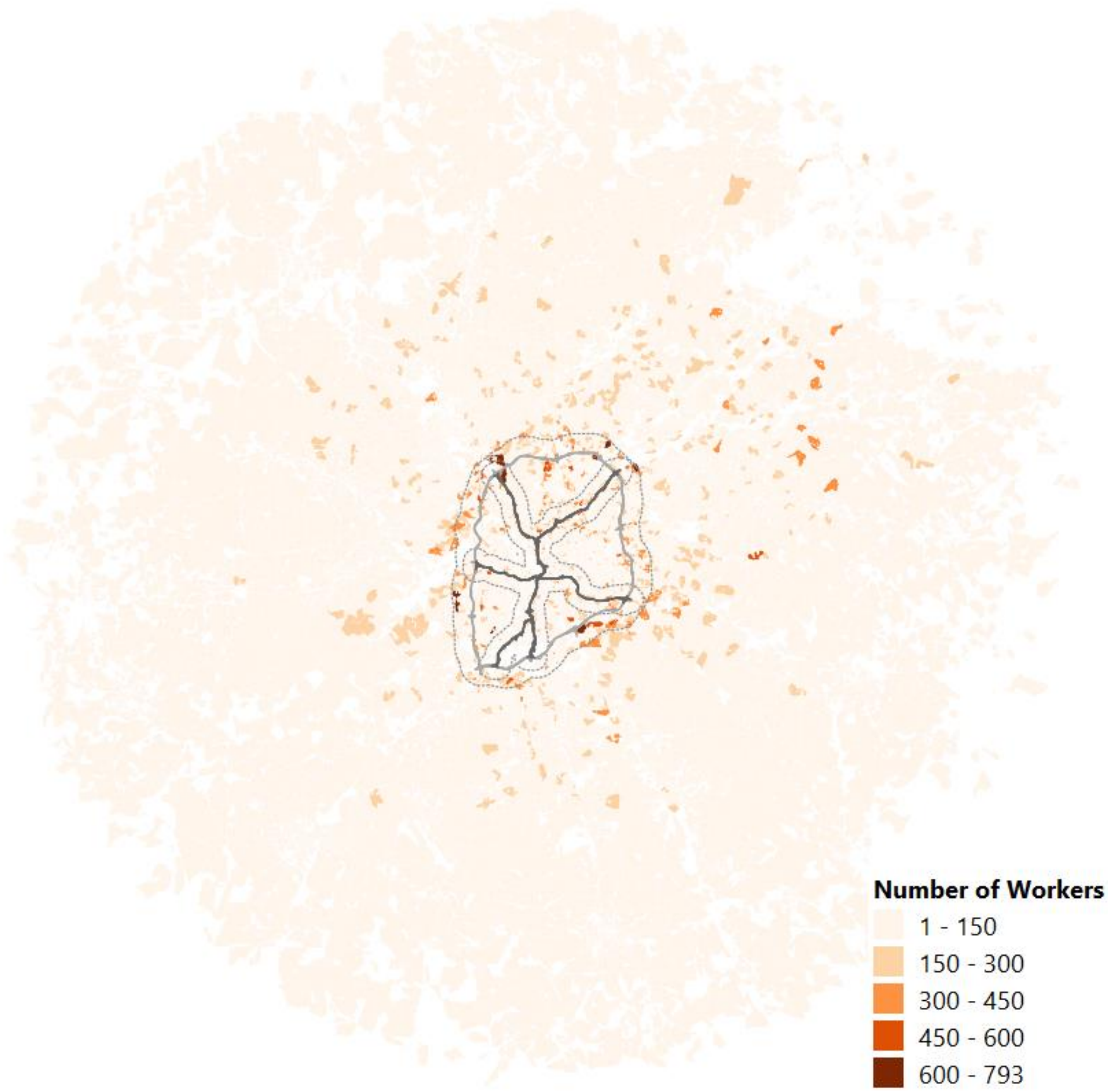


Figure 4-16 Home Places Distribution in Radius of 80km of Study Area

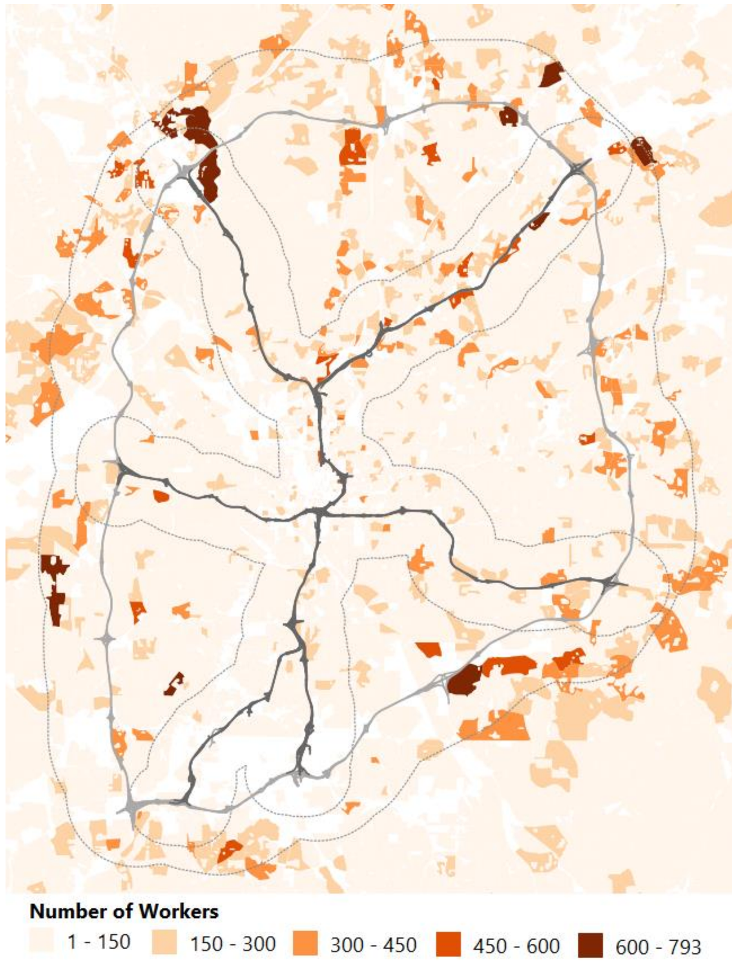


Figure 4-17 Home Places Distribution in Study Area

4.2.8 Vehicle Distribution

Atlanta Regional Commission (ARC) provides the household travel survey data in 2011 (Atlanta Regional Commission, 2011). It includes detail information about vehicles of each sample, such as year of the vehicle, make of vehicle, the model of vehicle, the body of the vehicle, type of fuels used by the vehicle, and location of vehicle owners (i.e., Traffic Analysis Zone (TAZ) and Census Tract). However, the fuel efficiency values provided by the U.S. Department of Energy (Alternative Fuels Data Center, 2015) is limited to some major vehicle categories, which are listed in Table 4-3. Therefore, the author aggregates

ARC vehicle data into four categories. Even ARC provides TAZ and Census Tract information for each vehicle; they cannot represent the vehicle distribution of each census tract due to insufficient sample size. Therefore, we use the vehicle distribution of each county to represent the vehicle distribution at each census tract. We assume that all census tracts at the same county follow the same vehicle distribution. There are four counties in our study area. The study area has 9846 samples, 56236 vehicles in total. The sample can represent the estimated 1,810,729 vehicles in the study area.

In the future, as additional vehicle survey data become available, market penetration of alternative fuel vehicles increases, and more disaggregate fuel efficiency information is acquired, we can estimate fuel consumption more accurately.

Table 4-3 Average Fuel Economy of Vehicle Categories in Study Area (Gasoline)

Vehicle Type	MPG (km/liter)	Vehicle Body in ARC Household Travel Survey
Light Duty Vehicle (C1)	21.6 (9.2)	Vans, SUV, Station Wagon and Pickup Trucks
Cars (C2)	23.4 (9.9)	Auto Sedan and Auto 2-Seat
Motorcycle (C3)	43.5 (18.5)	Motorcycle and Moped / Scooter (e.g. VESPA)
Recreational Vehicle (C4) (Alternative Fuels Data Center, 2015)	16 (6.8)	Class A, B, and C Recreational Vehicle

4.3 Indicators and Evaluation Method

4.3.1 Indicators

Jeon and Amekudzi (2005) conducted a comprehensive literature review on sustainability indicators from 16 different initiatives around the world, including North

America, Europe, Australia, and New Zealand (Jeon et al., 2013). Lautso et al. (2002) provide 35 key indicators for urban sustainability. All the sustainable indicators reviewed may be classified into four categories: economic, environmental, social and resilient. Ideally, sustainability evaluation should incorporate broader impacts of transportation systems and model the necessary interactions among these multi-dimensions (Jeon et al., 2013). However, it is very difficult to measure and to represent the interaction between multi-dimensions. We do not consider the interactions in this pilot study. Zegras (2006) presents the sustainability indicator prism that innovatively represents the hierarchy of goals, indexes, indicators, and raw data as well as the structure of multidimensional performance measures (Jeon et al., 2013). Zegras's sustainability indicator prism considered not only different dimensions of sustainability but also the community goals or various sustainability issues of the community. This framework is especially helpful when we explain the sustainability evaluation result to the public or policymakers. It makes them more easy to understand the sustainability evaluation results, and to be engaged in the process of sustainability evaluation.

Using the same strategy, the author first gets a list of indicators used by other researchers from literature that can be accessed. Second, the user classifies these indicators based on their relations to the four dimension of sustainability and the sustainability goals in the study area. In the end, the user selects relevant indicators based on the criteria proposed in section 3.2. Because the objective of this study is to evaluate the sustainability of moving interstates underground and to see how the sustainability performance changes after moving interstates underground, the author only considers indicators whose value may change after moving interstates underground. For example, people's commute time/distance, transportation fuel consumptions, noise pollution, people's accessibility to public facilities, equity, urban heat island conditions, nature and biodiversity, air pollution emissions, and so on. However, with the limit of data availability, quality of data, errors in the measure of indicators, the author only evaluated indicators listed in Table 4-4 below. For example, in the sustainability assessment, the author does not consider air pollution emissions (including different air emissions, such as CO2, CO, VOC, NOx, PM10, PM2.5, SOx, and so on) because the air emissions estimated by the MOVES model have too large errors and are highly biased at the census tract scale.

Table 4-4 Sustainable Indicators in Each Dimension and Its Measures

Dimension	No.	Name	Measures
Economic Sustainability	E11	Total Time Spent in Traffic	Average commute time from home to workplace per worker every day (minutes/worker/day)
	E12	Fuel Consumption	Average fuel consumption of traveling from home to work each day per worker (liters/worker/day)
	E13	Connectivity of Transportation Network	Detour Index of 500 meters trips (>=1)
Environmental Sustainability	E21	Nature & Biodiversity	Green Area per Capita (m^2/person)
	E22	Urban Heat Island	Green & Water Space Ratio (0 to 1)
	E23	Noise Pollution Generated	Population Affected by Interstate Highway Noise (population)
	E24	Land Use	Green Space Ratio (0 to 1)
	E25		Land Released by Moving Interstates Underground (m^2)
Social Sustainability	E31	Accessibility to Green Space	Attraction accessibility to green space (800 m)
	E32	Accessibility to Health Facilities	Attraction Accessibility to Health Facilities (3200 m)

E33	Accessibility to Education Facilities	Attraction Accessibility to Education Facilities (1600 m)
E34	Equity	Income Equity Index of Commute Distance (from home to workplace) (<= 1)

4.3.2 *Indicators Evaluation Methods*

4.3.2.1 Economic Dimension: Connectivity

Connectivity refers to the directness of travel between origin and destination points. Among all the connectivity measures, DI measures the efficiency of a connection in the network and can directly reflect the travel distance for a trip (Dill, 2004). It is the ratio of actual route distance to straight-line distance for two selected points (see Equation 3.7). DI tending to a value of one indicates a more spatially efficient network.

The process to calculate the detour index is illustrated in Figure 4-18 below. We first use a random vector generator (details see section 3.5.1.4) provided in the proposed framework to generate 10,000 random vectors. Each random vector (a pair of points) has a length of 500 meters since most walking trips are under one mile (Dill, 2004). Congalton and Green (2009) suggest that a minimum of 100 samples per class are necessary for a large area. Therefore, 10,000 random vectors are considered sufficient for an analysis to yield unbiased results. Next, we calculate DI of each random vector, thus obtain 10,000 DIs. Since the spatial scale of sustainability analysis in this pilot study is census tract level, we aggregate DI of each random vector to each census tract. For census tract *k*, as long as its boundary intersects with random vector RV_i $(i = 1,2,3, \ldots ,10{,}000)$, the DI of RV_i belongs to census tract *k*. DI of census tract *k* is the average DI of all RV_i that intersects with its boundary. In order to reduce the biased caused by the random vectors, we run the program about 200 time and each time we generated 10,000 new random vectors. Therefore, we got 200 different the average DI for each census tract in the study area. The mean value of these 200 average DIs is the final average DI for each census tract. Finally, we can analyze the spatial patterns of DI in the study area and its difference between two road networks. Besides, the DI of each census tract can also be used in the following MCA of sustainability analysis.

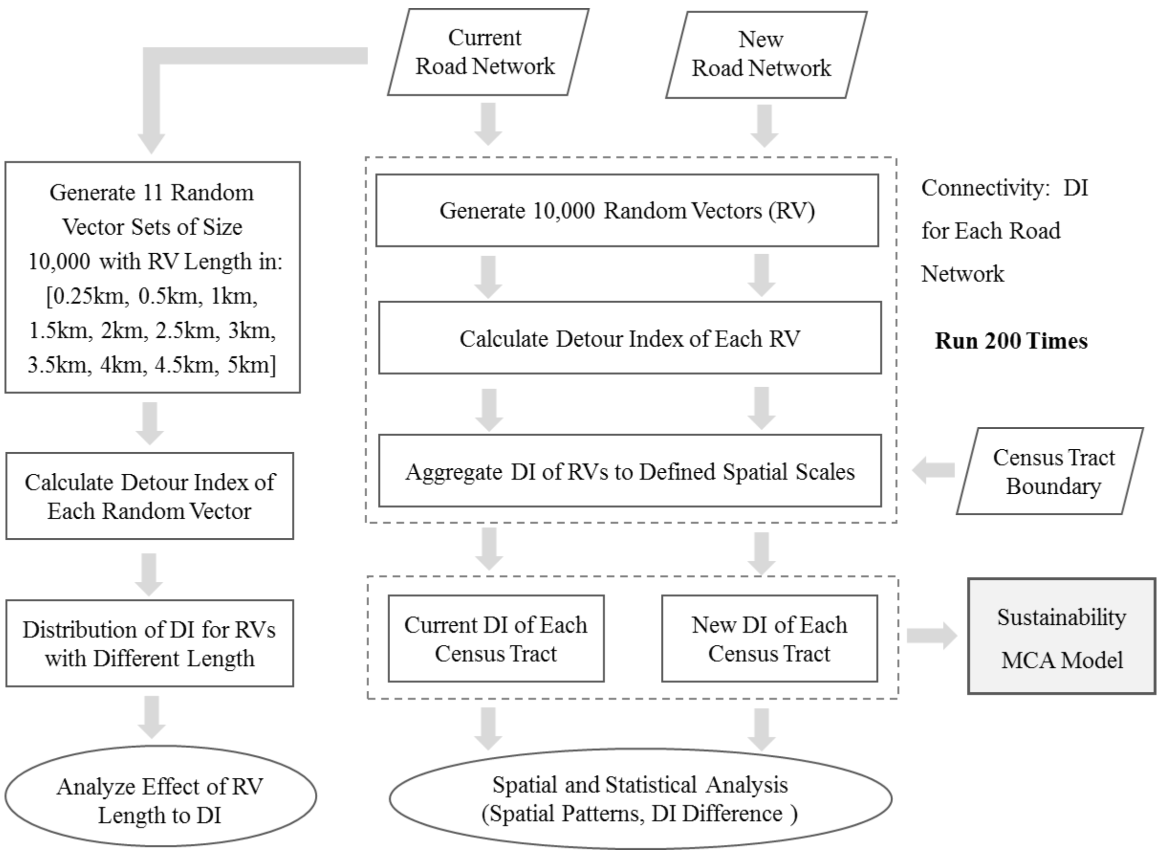


Figure 4-18 Flowchart of Calculating Detour Index

This author also studied how the length of the random vector affected the distribution of DI in the study area (see left part of Figure 4-18). The length of random vector changes from 0.25km, 0.5km, 1km, 1.5m, 2km, 2.5km, 3km, 3.5km, 4km, 4.5km, to 5km. For the random vector with each length, we generate 10,000 random vectors, calculate its DI and analyze its distribution (including, min, max, median, quantiles, and cumulative percentage distribution). In the end, we analyze and compare the changes of DI distributions with random vector length.

4.3.2.2 Economic Dimension: Total Commute Time

OD flow data provides workers' travel information between home and workplaces at census blocks level, which is how many workers in total travel from one census block to another census block. When calculating travel time, the author makes the following assumptions:

- People always travel from the central point of home place census blocks to the central point of workplace census blocks.
- All workers drive from home place to the workplace each day.
- People drive directly from home place to the workplace, without stopping by other places (e.g., restaurant, coffee store, day-care center, and so on) between home and workplace.
- People drive at the speed limit of each road link without considering waiting time for traffic signals, traffic congestions caused by over-capacity traffic flows, and other factors which may affect driving speed.
- People always choose the route with the **shortest travel time** when driving from home to their work place.

In this study, the author uses "Vehicle Hours Traveled Per Capita" from home to workplace each day to measure the "total commute time". It is the average commute time each worker who live in census tract *i* spent each day from home to workplaces. Its calculating process is illustrated in Figure 4-19. First, the shortest travel time for each OD pair is calculated using Dijkstra's shortest path algorithm provided in the Python igraph library. Second, the travel time and a number of workers are aggregate to each census tract. Third, "Vehicle Hours Traveled Per Capita" ($VHpCap_k$) can be calculated using Equation 4.1 below. Finally, we can analyze the spatial patterns of $VHpCap_k$ in the study area and its difference between two road networks. Besides, the $VHpCap_k$ of each census tract can also be used in the following MCA of sustainability analysis.

$$VHpCap_k = \sum_{i}^{n}\sum_{j}^{n}\frac{OD_{ij} \times T_{ij} \times \mathbf{1}(i \in \mathrm{track}_k)}{WorkerPop_k}$$

$$WorkerPop_k = \sum_{i}^{n}\sum_{j}^{n} OD_{ij} \times \mathbf{1}(i \in \mathrm{track}_k) \tag{4.1}$$

$$\text{Indicator Function } \mathbf{1}(i \in \mathrm{track}_k) = \begin{cases} 1 & \text{point } i \text{ in census track } k \\ 0 & \text{otherwise} \end{cases}$$

Where, $VHpCap_k$ is Vehicle Hours Traveled Per Capita for people who lives in census tract *k* going from home to workplace every day. OD_{ij} is number of people who live in census block *i*, but work in census block *j*. $WorkerPop_k$ is the total population of workers who live in census tract *k*. T_{ij} and D_{ij} is the total travel time and total travel distance from central point of census block *i* to central point of census block *j*. *n* is the number of census blocks in the study area.

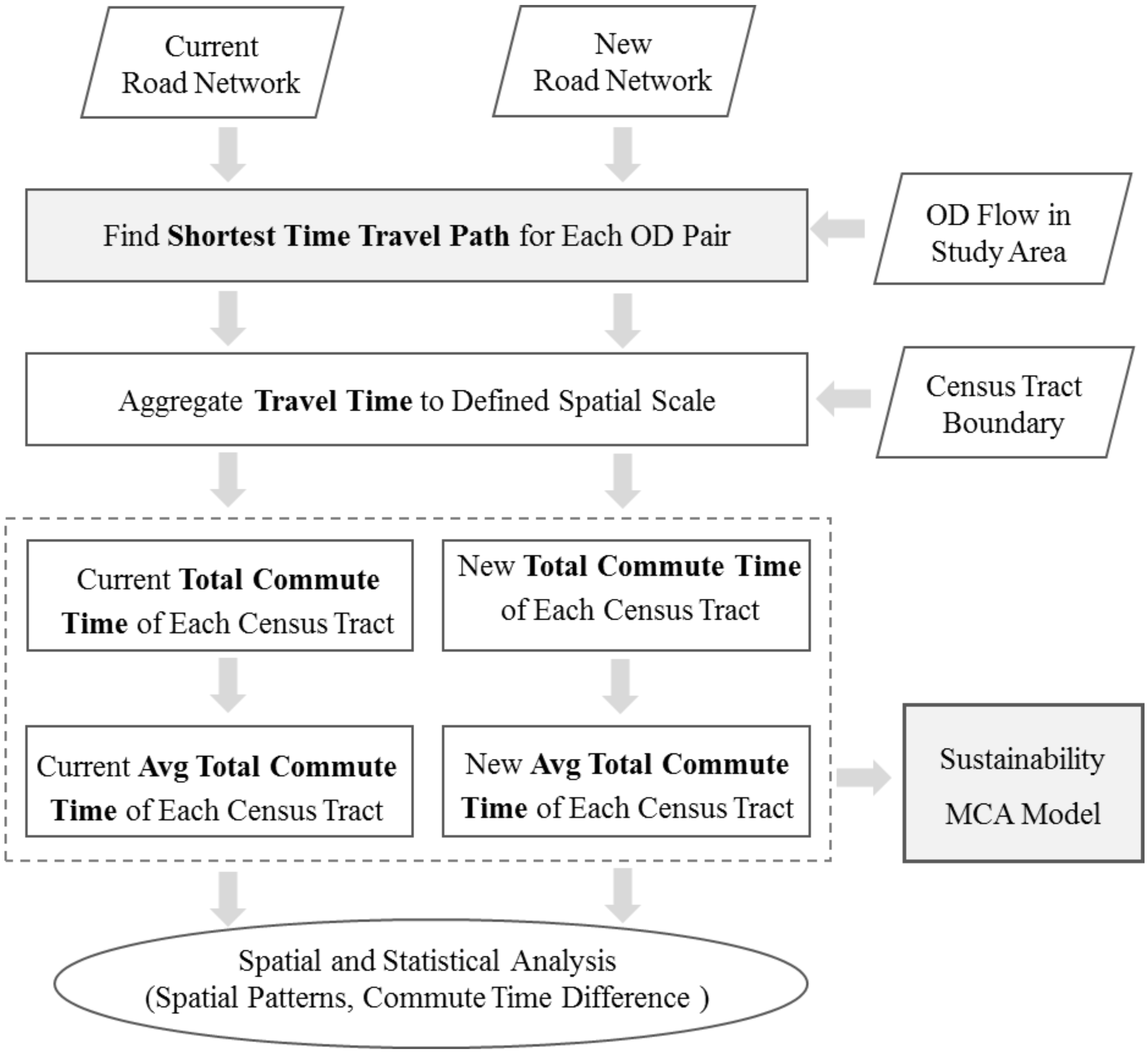


Figure 4-19 Flowchart of Calculating Total Commute Time

4.3.2.3 Economic Dimension: Fuel Consumption

Energy or fuel consumption of one person should include all fuel consumption in his or her daily life, such as cooking, heating or cooling the building, lighting, charging a computer or other digital devices, transportation (drive or take the bus to work, school, shopping et al.) and so on. However, due to data and time limit, this study only considers the fuel consumed by driving directly from home to workplace each day for workers who live in census tract *k*. The average fuel consumption per capita can be estimated based on the vehicle distance traveled per capita in census tract *k*, vehicle-type specified fuel efficiency value (*km/liter*), and vehicle distribution in this tract. Ideally, consideration of a very disaggregate representation of vehicle types in the computation of travel energy footprint would be advantageous (Garikapati et al., 2017). However, due to the limitation of fuel efficiency data, we use aggregated classification of vehicle types in this study. A summation of fuel consumption over all vehicle types in census tract *k* yields the total fuel consumption (i.e., the footprint of travel energy) for census tract *k*.

The process to calculate average fuel consumption is illustrated in Figure 4-20. First, the shortest travel time for each OD pair is calculated using Dijkstra's shortest path algorithm provided in the Python igraph library. Second, the travel distance and the number of workers are aggregate to each census tract. Third, total fuel consumption and average fuel consumption per capita are estimated by Equation 4.2. Finally, we can analyze the spatial patterns of average fuel consumption in the study area and the difference between two road networks. The average fuel consumption of each census tract can also be used in the following MCA of sustainability analysis.

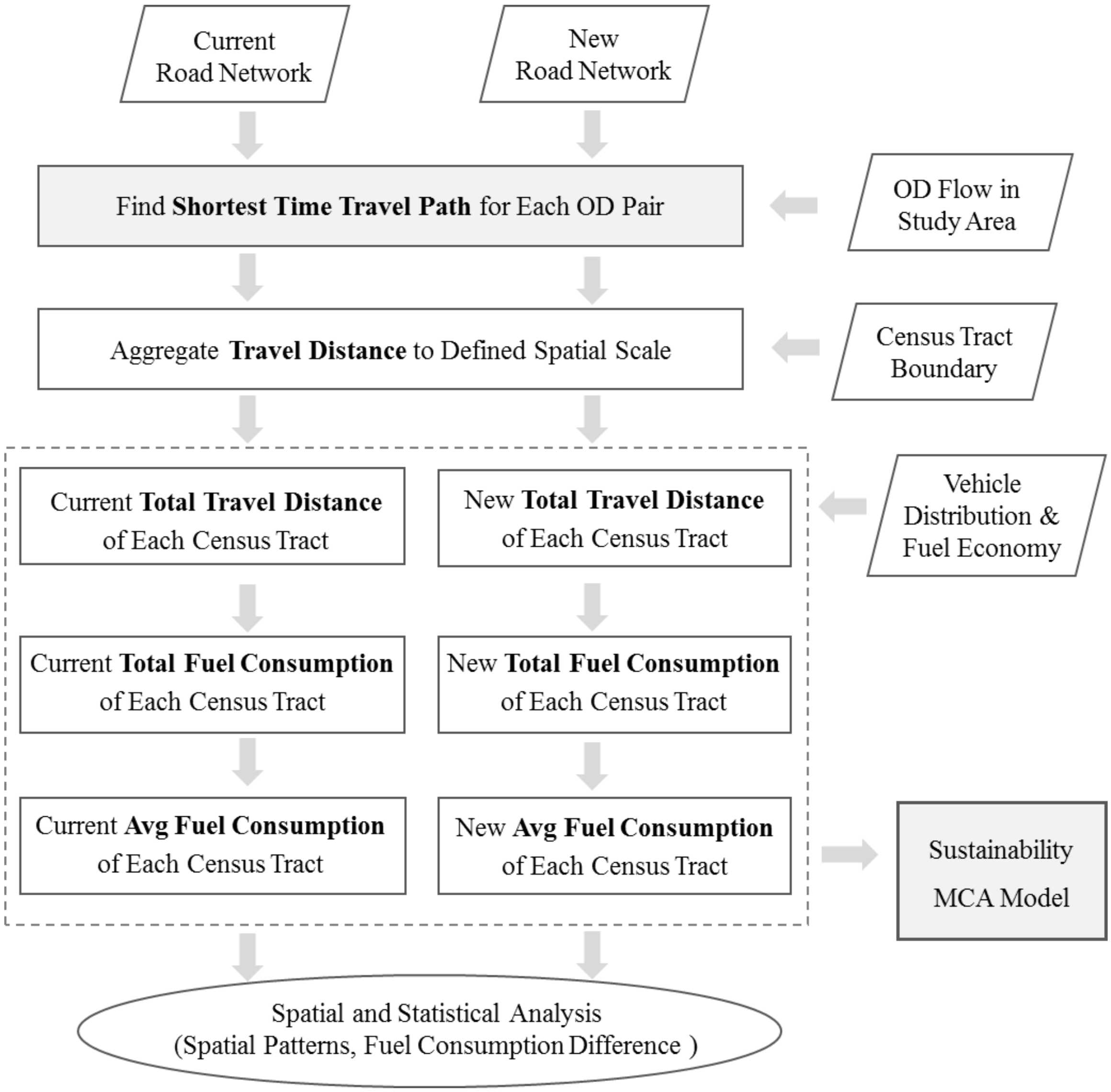


Figure 4-20 Flowchart of Calculating Fuel Consumption

$$\text{FuelTotal}_k = \sum_{t}^{N_{VehTyp}^{k}} \frac{\text{VDT}_k \times \text{Rveh}_{kt}}{\text{FuelEff}_j}$$

$$\text{FuelCap}_k = \sum_{j}^{N_{vehTyp}^{k}} \frac{\text{VDpCap}_k \times \text{Vp}_{kj}}{\text{FuelEff}_j}$$

(4.2)

$$VDT_k = \sum_{i}^{n}\sum_{j}^{n} OD_{ij} \times D_{ij} \times \mathbf{1}(i \in \text{track}_k)$$

$$VDpCap_k = \sum_{i}^{n}\sum_{j}^{n} \frac{OD_{ij} \times D_{ij} \times \mathbf{1}(i \in \text{track}_k)}{WorkerPop_k}$$

Where, FuelTotal_k and VDT_k are the total fuel consumption and total commute distance of all works who live in census tract *k* for traveling from home to workplace each day respectively. FuelCap_k is the average fuel consumption of workers who live in census tract *k* for traveling from home to workplace each day. $VDpCap_k$ is vehicle distance traveled per capita for people who lives in census tract *k* going from home to workplace every day. $WorkerPop_k$ is the total number of works who live in census tract k. Vp_{kt} is the ratio of vehicle type *t* at census tract *k*, FuelEff_j is the fuel efficiency value of vehicle type *t*. N_{VehTyp}^{k} is the total number of vehicle types at census tract *k*.

4.3.2.4 Environmental Dimension: Natural and Biodiversity

Natural and biodiversity provide health and social benefits (Ten Brink et al., 2016), such as improving air quality, mitigating pollutions and natural hazards, promoting healthier lifestyles, outdoor recreation, and physical activities, increasing social cohesion and reducing violence and crime rates, and so on. However, they are under pressure from urban sprawl, intensive agriculture, pollution, invasive species and climate change (European Commission, 2018). Many variables can measure natural and biodiversity, including species traits and population, ecosystem function and structure (e.g., ecosystem distribution, fragmentation and heterogeneity, land cover, leaf area index and so on) (Skidmore et al., 2015).

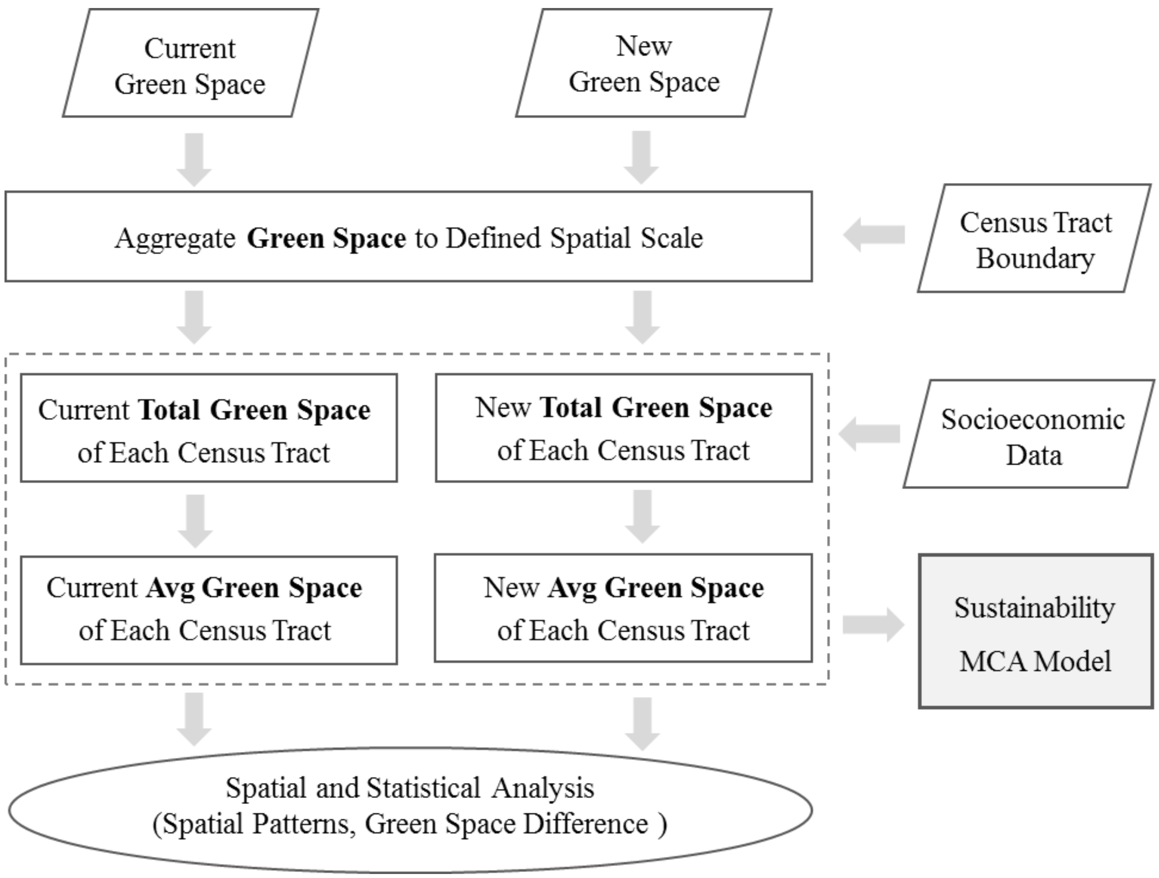


Figure 4-21 Flowchart of Calculating Natural and Biodiversity Index

Green space can help preserve and support urban biodiversity and nature conservation, as protection center for reproduction of species and conservation of plants, soil and water quality (Haq, 2011). The ability of green space to support biodiversity varies with landscape configuration (i.e., patch size, shape, and connectivity), biotic interactions, land use history, human population density of the surrounding urban matrix, economic input, and management activities (Aronson et al., 2017). In this study, natural and biodiversity are measured by “average green space areas per capita”, which can be calculated by Equation 4.3. Pop_k is the total population at census tract k. Its spatial and statistical patterns and its difference between two road networks are analyzed. And average green space value can also be used in the following MCA of sustainability analysis.

$$\text{GreenSpaceCap}_k = \frac{\text{Total Green Areas in Census Tract } k}{Pop_k} \tag{4.3}$$

4.3.2.5 Environmental Dimension: Urban Heat Island

Green space and water surfaces in the area can help reduce the rise in surface temperature, thus this study uses “proportion of green space and water surfaces” as one measure of urban heat island. Its calculating process is illustrated in Figure 4-22. The green & water ratio at census tract k, GreenWaterRatio_k, can be calculated by Equation 4.4.

$$\text{GreenWaterRatio}_\text{k} = \frac{\text{Green \& Water Area in Census Tract } k}{\text{Total Area of Census Tract } k} \tag{4.4}$$

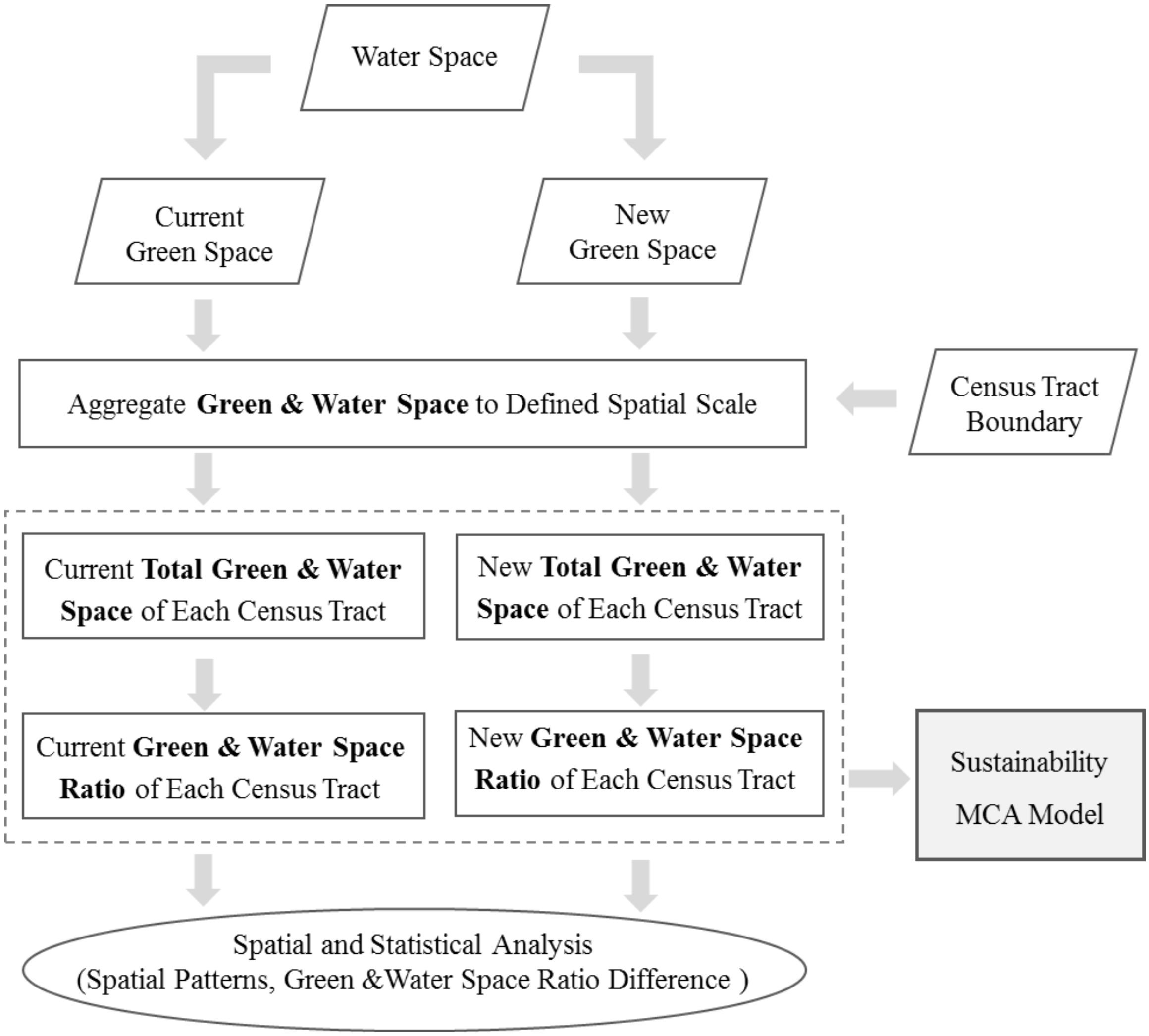


Figure 4-22 Flowchart of Calculating Urban Heat Island Indicator

4.3.2.6 Environmental Dimension: Noise Pollution

Noise pollution is an important indicator of environmental sustainability and can be affected by moving interstates underground. But the data of noise pollution is very difficult to get, and it is complex to estimate noise pollution. Due to the limitation of data availability and time, this study uses a proportion of people who live within *200* meters (USDA National Agroforestry Center, 2017) buffer of highways as one measure of the noise pollution produced by interstates. Because the study shows that levels of highway traffic noise typically range from *70* to *80 dB(A)* at a distance of *15 meters* (*50 feet*) from the highway (Chris Corbisier, 2003). The noise pollution of census tract k, $\mathrm{NoisePollut}_k$, can be calculated by Equation 4.5.

$$NoisePollut_k = \#(\text{people living in 200 meters buffer of interstates at census tract } k) \tag{4.5}$$

4.3.2.7 Environmental Dimension: Land Use

The "Green Space Ratio" and "Land Released by moving interstates underground" are used to measure the land use indicators in this study. The process to estimate green space ratio is illustrated in Figure 4-23. The green space ratio of census tract k can be calculated by Equation 4.6. Land released at census tract k is the total area of land released after moving interstates underground at census tract k.

$$GreenRatio_k = \frac{\text{Areas of Green Spaces at Census Tract } k}{\text{Total Area of Census Tract } k} \tag{4.6}$$

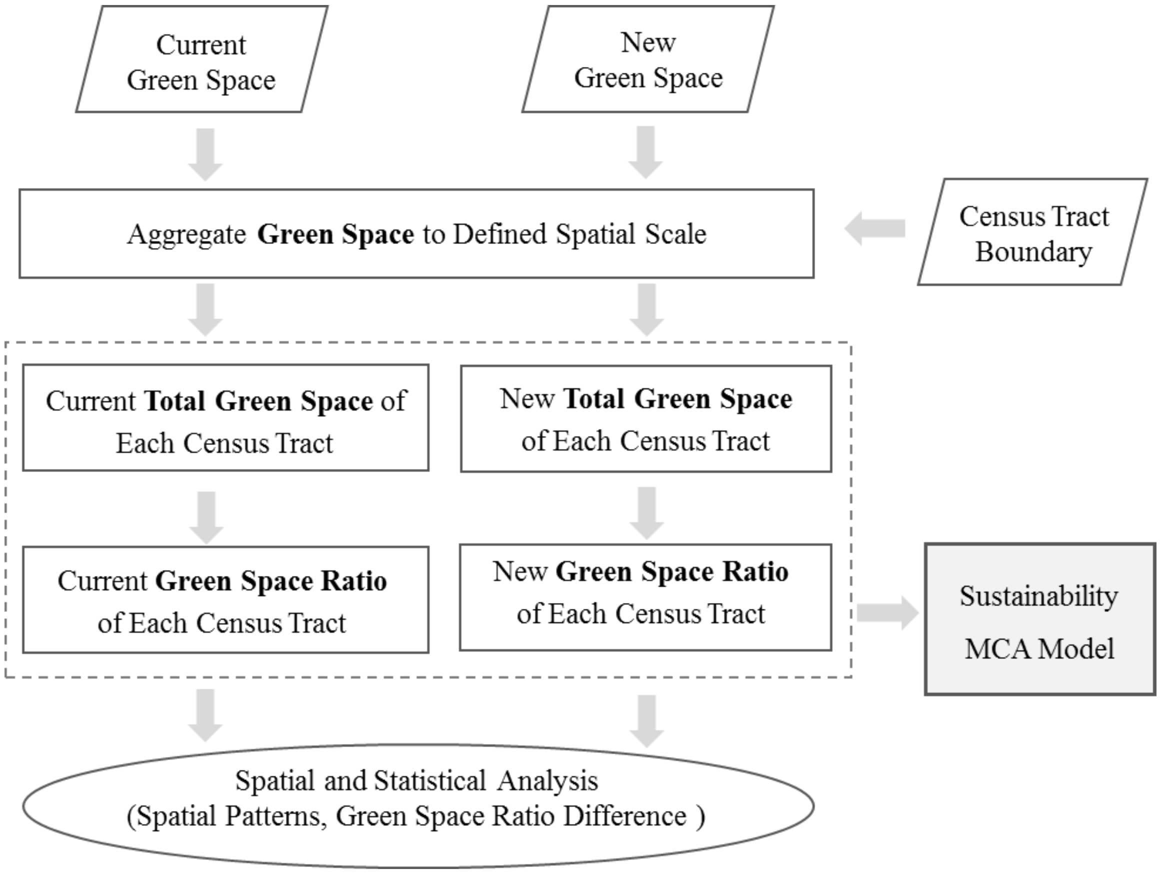


Figure 4-23 Flowchart of Calculating Land Use Indicator

4.3.2.8 Social Dimension: Accessibility

The transportation system plays a crucial role in joining locations to each other, easing people and goods movement from/to various locations and increasing people's participation in activities. Measuring the accessibility it provides for people is a good measure of transportation system performance and social equity. As shown in Equation 3.8, we need to have origins and destination and transportation network to calculate the accessibility. In this study, we first generate *10,000* random points as travel origins in the study area and then calculate accessibility to destinations from each random point (see Figure 4-24). The destinations are the parks, health facilities (including hospitals, clinics, and emergency rooms) and education facilities (including elementary schools, middle school, high schools, colleges, universities, and institutes). Two transportation networks are used, one is the current transportation network, and the other is the new transportation network with interstates moved underground (referred as reconnected road network in the following section), details see section 4.2.1. After getting accessibility to destinations from each random point, we aggregate them into defined space scales. In this study, we aggregate them into the census tract level. The average accessibility of census tract k, $Access_k$, can be calculated by Equation 4.7. Finally, we analyze the spatial and statistical patterns of accessibility and its difference between current and new road network. The accessibility indicator can also be used in the following MCA model for sustainability analysis.

$$Access_k = \frac{\sum_{\boldsymbol{i=1}}^{N_{pt}} A_i \times \mathbf{1}(i \in \text{track}_k)}{\sum_{\boldsymbol{i=1}}^{N_{pt}} \mathbf{1}(i \in \text{track}_k)} \tag{4.7}$$

$$\text{Indicator Function: } \mathbf{1}(i \in \text{track}_k) = \begin{cases} 1 & \text{if point } i \text{ in census tract } k \\ 0 & \text{otherwise} \end{cases}$$

Where, N_{pt} is the total number of origin points used to calculate accessibility in the study area and $N_{pt} = 10{,}000$ in this study. A_i is the accessibility to each destination from point i. If A_i is the accessibility to green space from point i, then $Access_k$ is the average accessibility to green space of census tract k.

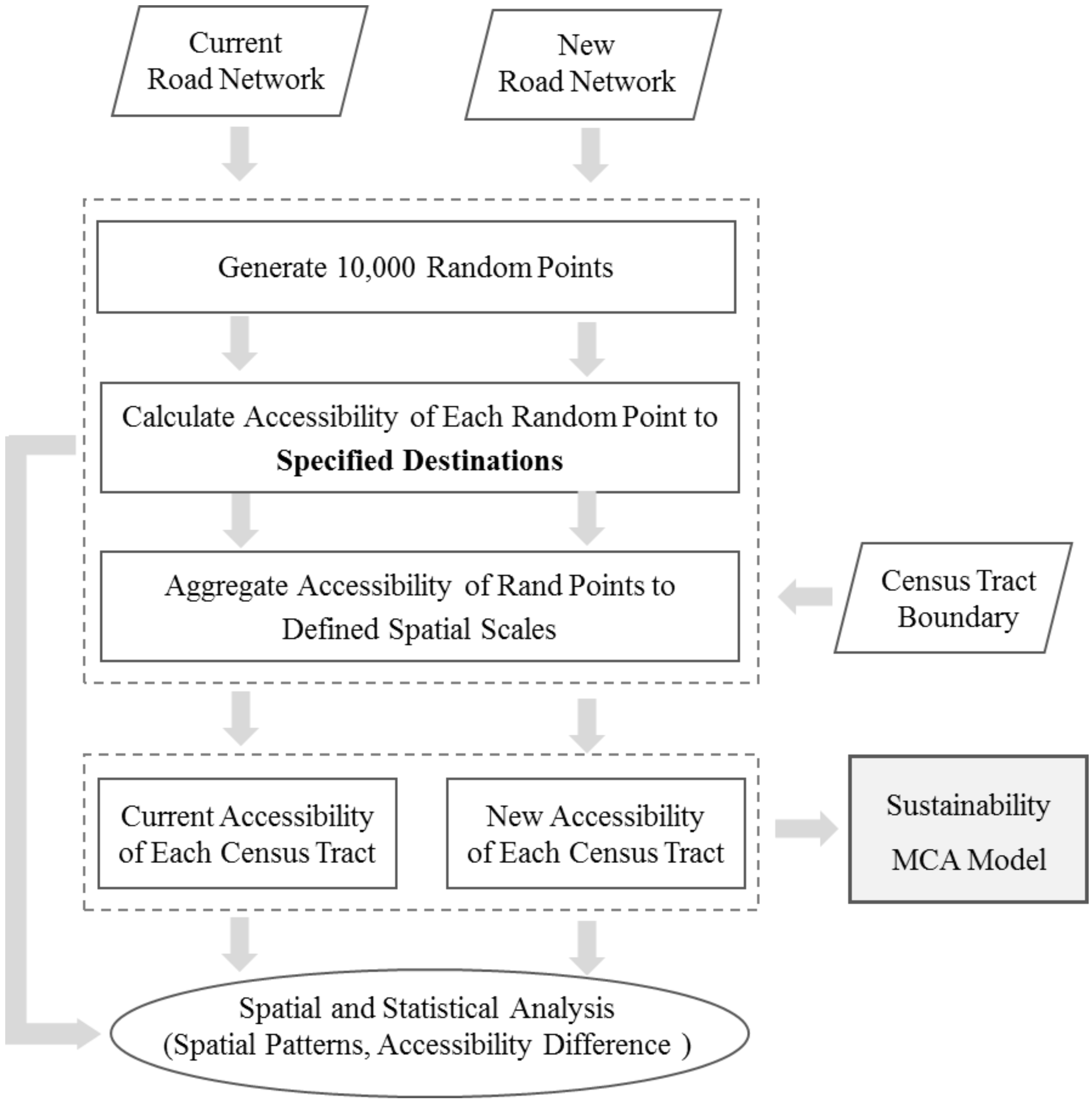


Figure 4-24 Flowchart of Calculating Accessibility to Specified Destination

1) <u>Accessibility to Green Space</u>

Urban parks and green spaces can provide social services for good quality of life and are considered a key component of sustainability (Lee and Kim, 2015). Citizen's accessibility to green space such as parks is an important indicator of a higher quality of life (Oh and Jeong, 2007). The optimal distance from home to park is less than *0.5* km or less than *5* minutes' walking time (Lee et al., 2015). The National Recreation and Park Association (NRPA) in North America recommended the maximum distance for park accessibility be *0.5* mile or *800* meters (Oh and Jeong, 2007). Therefore, we also use *800* meters as the service distance of green space in this study.

The author assumes that a larger green space is more attractive than smaller green space and can provide greater benefits to human beings. Therefore when calculating the accessibility to green space, this study uses the size of green space GA_j as its attractiveness. The accessibility to green space from origin point *i*, A_i can be calculated by Equation 4.8. The average accessibility to green space for census tract *k* can be calculated by Equation 4.7.

$$A_i = \sum_{j \in K_i} w_j \times f(c_{ij}) = \sum_{j \in K_i} GA_j \times {d_{ij}}^{-2}$$

$$= \sum_{j=1}^{N_{green}} GA_j \times {d_{ij}}^{-2} \times \mathbf{1}(dd_{ij} \leq 800\ m)$$

$$N_{green}^{i} = \sum_{j=1}^{N_{green}} \mathbf{1}(dd_{ij} \leq 800\ m) \tag{4.8}$$

$$GAaccess_i = \sum_{j=1}^{N_{green}} GA_j \times \mathbf{1}(dd_{ij} \leq 800\ m)$$

$$\text{Indicator Function: } \mathbf{1}(d_{ij} \leq 800\ m) = \begin{cases} 1 & \text{if } dd_{ij} \leq 800\ m \\ 0 & \text{otherwise} \end{cases}$$

Where, K_i is the set of parks that can be accessed from location *i* within *800* meters of Euclidean Distance ($i = 1, 2, \ldots, N_{pt}$). N_{green} is the total number of green space in the study area. GA_j is the area of green space j, d_{ij} is the route distance (real travel distance) from location *i* to the nearest boundary of green space j, dd_{ij} is the Euclidean Distance from location *i* to the nearest boundary of green space j, N_{green}^{i} is the number of green space one can access from location i, $GAaccess_i$ is the total area of green space one can access from location *i*.

2) <u>Accessibility to Health Facilities</u>

As we mentioned in the last section, to calculate accessibility to health facilities, we need to know the attractiveness of each health facility. However, with the limitation of data, we cannot estimate the attractiveness of each health facilities. In this study, we assume each health facility has the same attractiveness. In other words, $HLT_j = \text{constant for all } j$. For simplicity, we set $HLT_j = 1$. Love and Lindquist (1995) use 20 miles (32 km) threshold to evaluate the accessibility to health facilities in Illinois State. But since our study area is very small, we set our distance threshold to be 3200 meters (2 miles), even some researcher use 1 mile threshold (Kane et al., 2017). Therefore, the accessibility to health facilities from origin point *i*, A_i can be calculated by Equation 4.9. The average accessibility to health facilities for census tract *k* can be calculated by Equation 4.7.

$$A_i = \sum_{j \in K_i} w_j \times f(c_{ij}) = \sum_{j \in K_i} HLT_j \times {d_{ij}}^{-2}$$

$$= \sum_{j=1}^{N_{hlt}} {d_{ij}}^{-2} \times \mathbf{1}(dd_{ij} \leq 3200\ m)$$

$$N_{hlt}^i = \sum_{j=1}^{N_{hlt}} \mathbf{1}(dd_{ij} \leq 3200\ m) \tag{4.9}$$

Indicator Function: $\mathbf{1}(d_{ij} \leq 3200\ m) = \begin{cases} 1 & \text{if } dd_{ij} \leq 3200\ m \\ 0 & \text{otherwise} \end{cases}$

Where, K_i is the set of health facilities that can be accessed from location *i* within *3200* meters of Euclidean distance ($i = 1, 2, \ldots, N_{pt}$). HLT_j is the attractiveness of health facility j, d_{ij} is the route distance (real travel distance) from location *i* to health facility *j*, dd_{ij} is the Euclidean distance from location *i* to health facility *j*, N_{hlt}^i is the number of health facilities one can access from location *i*. N_{hlt} is the total number of health facilities in the study area.

3) <u>Accessibility to Education Facilities</u>

We use the same method to calculate accessibility to education facilities but set a distance threshold to be 1600 meters (1 mile) (Kane et al., 2017). In this study, we use Equation 4.10 to calculate one's accessibility to education facilities. The average accessibility to education facilities for census tract *k* can be calculated by Equation 4.7.

$$A_i = \sum_{j \in K_i} w_j \times f(c_{ij}) = \sum_{j \in K_i} SCH_j \times {d_{ij}}^{-2}$$

$$= \sum_{j=1}^{N_{edu}} {d_{ij}}^{-2} \times \mathbf{1}(dd_{ij} \leq 1600\ m)$$

$$N_{edu}^i = \sum_{j=1}^{N_{edu}} \mathbf{1}(dd_{ij} \leq 1600\ m) \tag{4.10}$$

Indicator Function: $\mathbf{1}(d_{ij} \leq 1600\ m) = \begin{cases} 1 & \text{if } dd_{ij} \leq 1600\ m \\ 0 & \text{otherwise} \end{cases}$

Where, K_i is the set of education facilities that can be accessed from location *i* within 1600 meters Euclidean distance ($i = 1, 2, \ldots, N_{pt}$). SCH_j is the attractiveness of education facility *j*, d_{ij} is the route distance (real travel distance) from location *i* to education facility *j*, dd_{ij} is the Euclidean distance from location *i* to education facility *j*. N_{edu} is the total number of education facilities in the study area. N_{edu}^i is the number of education facilities one can access from location *i*.

4.3.2.9 Social Dimension: Equity of Commute Distance

This study uses the "Income Equity Index" (Jeon et al., 2013) of commute distance per worker traveled each day as a measure of the equity of commute distance. As shown in Figure 4-25, we first calculate the shortest travel time path for each OD pair using Dijkstra's shortest path algorithm. Second, we aggregate the travel time of people in each income class to defined spatial scale. In this study, all data are aggregated to the census tract level. According to the income categories used in OD flow data (i.e., SE01, SE02, and SE03 in

Table 4-2), three categories of income are considered in this study. They are low income (annual income is less than 15,000 dollars), medium income (annual income is from 15,000 dollars to 40,000 dollars), high income (annual income is greater than 40,000 dollars). Third, the income equity index of commute distance at census tract k is be calculated by Equation 4.11. The total commute distance of people in income class m, VDT_k^m, is obtained by aggregating all the OD pair commute distance by the commuter's income class.

$$EquityCommute_k = 1 - \sum_{m=1}^{3} |X_k^m - D_k^m| = 1 - \sum_{m=1}^{3} \left| X_k^m - \frac{VDT_k^m}{VDT_k} \right|$$

$$VDT_k^m = \sum_{i}^{n} \sum_{j}^{n} OD_{ij} \times D_{ij} \times \mathbf{1}(i \in track_k) \times \mathbf{1}(Income_i \in \text{Class } m)$$

$$\mathbf{1}(i \in track_k) = \begin{cases} 1 & \text{if start point of } OD_{ij} \text{ in census track } k \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

$$\mathbf{1}(Income_i \in \text{Class } m) = \begin{cases} 1 & \text{if income of start point of } OD_{ij} \text{ belongs to income class } m \\ 0 & \text{otherwise} \end{cases}$$

Where, $EquityCommute_k$ is the Income Equity Index of commute distance at census tract k, X_k^m is the ratio of income class *m* households (three classes: low, medium and high income) at census tract *k*, D_k^m is the ratio of total commute distance at income class *m* for people who live in census tract *k*, $D_k^m = \frac{VDT_k^m}{VDT_k}$. VDT_k is the total commute distance of all people who live in census tract *k*. OD_{ij} is number of people who live in census block *i*, but work in census block *j*. D_{ij} is the total travel distance from central point of census block *i* to central point of census block *j*. $Income_i$ is the income of start point *i* in OD pair. n is the number of census blocks in the study area.

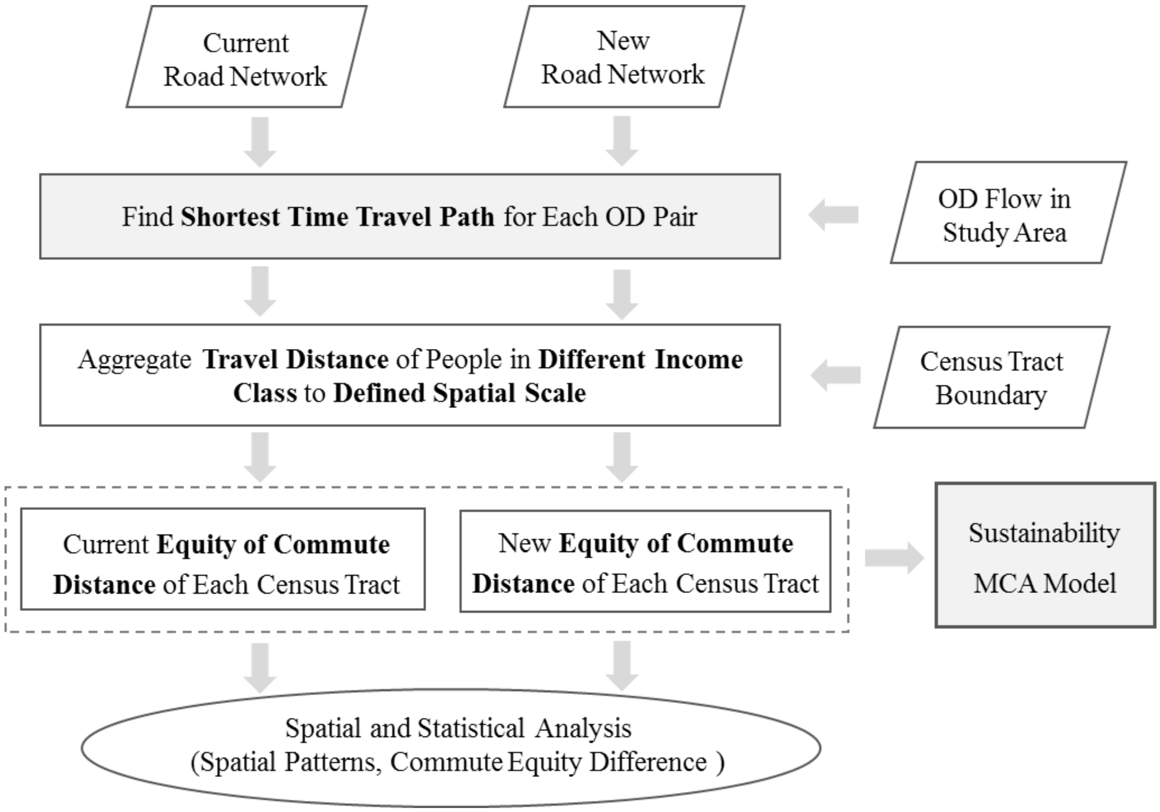


Figure 4-25 Flowchart of Calculating Equity of Commute Distance

4.3.3 Overall Sustainability Estimation

After getting the value for each indicator, we use MCA to integrate them into one measure to represent the sustainability in one region. First, we need to normalize each indicator before using them to estimate overall sustainability. There are many normalization methods, the author uses Min-Max normalization method in this study (see Equation 4.12), to make sure all the indicator values are non-negative, and larger indicator value representing better sustainability performance.

$$\mathrm{I}_i{}' = \begin{cases} \dfrac{I_i - \min_{\forall j}(I_j{}^{Cur}, I_j{}^{New})}{\max_{\forall j}(I_j{}^{Cur}, I_j{}^{New}) - \min_{\forall j}(I_j{}^{Cur}, I_j{}^{New})}, & if\ sgn(I_i) = 1 \\ 1 - \dfrac{I_i - \min_{\forall j}\left(I_j{}^{Cur}, I_j{}^{New}\right)}{\max_{\forall j}\left(I_j{}^{Cur}, I_j{}^{New}\right) - \min_{\forall j}\left(I_j{}^{Cur}, I_j{}^{New}\right)}, & if\ sgn(I_i) = -1 \end{cases} \tag{4.12}$$

Where, $\mathrm{I}_i{}'$ is the normalized indicator value, it belongs to *[0, 1]*. I_i is the *i-th* indicator value we calculated use the method presented in previous sections. $I_j{}^{Cur}$ is the *j-th* indicator value in the current road network, $I_j{}^{New}$ is the same *j-th* indicator value in the new road network. $sgn(I_i)$ is sign of indicator I_i, which imply the positive ($gn(I_i) = 1$) or negative ($gn(I_i) = -1$) impact of indicator I_i to the sustainability of the study area.

Secondly, we analyze the relationship between all indicators by calculating their Pearson correlation coefficient and conducting a statistical test of significance of the correlation-ship between two indicators. In this study, the author sets p-value to be 0.05, the confidence level is 95%. If there are multiple indicators are highly correlated, only one of them is used in the MCA model for final overall estimation of sustainability. The correlation matrix of all indicators is shown in Figure 4-29.

Thirdly, we use the remaining relatively independent indicators to estimate the composite sustainability score for each dimension of sustainability and the overall composite sustainability score by Equation 4.13. The weight of each indicator is derived by AHP tool in the implemented “Spatial Infrastructure Sustainability Assessment” Plugin in QGIS. In this study, all indicators of one dimension are treated equally when calculating the dimensional composite sustainability score (i.e., all value in the pair-wise comparison matrix are one). Each sustainability dimension is also equally treated when calculating the overall composite sustainability score in Equation 4.13.

$$\mathrm{SusEnv}_k = \sum_i \mathrm{I}_{ki}{}' \times W_{ki}^{d1} \times \mathbf{1}(\mathrm{I}_{ki}{}' \in \mathrm{SusDim}_{\mathrm{Env}})$$

$$\mathrm{SusEco}_k = \sum_i \mathrm{I}_{ki}{}' \times W_{ki}^{d2} \times \mathbf{1}(\mathrm{I}_{ki}{}' \in \mathrm{SusDim}_{Eco})$$

$$\mathrm{SusSoc}_k = \sum_i \mathrm{I}_{ki}{}' \times W_{ki}^{d3} \times \mathbf{1}(\mathrm{I}_{ki}{}' \in \mathrm{SusDim}_{Soc})$$

$$\mathrm{SusTotal}_k = \mathrm{SusEnv}_k \times W_{Env} + \mathrm{SusEco}_k \times W_{Eco} + \mathrm{SusSoc}_k \times W_{Soc}$$

$$\mathbf{1}(\mathrm{I}_{ki}{}' \in \mathrm{SusDim}_{\mathrm{Env}}) = \begin{cases} 1 & \text{if Indicator } \mathrm{I}_{ki}{}' \text{ belongs to Environmental Dimension} \\ 0 & \text{otherwise} \end{cases} \tag{4.13}$$

$$\mathbf{1}(\mathrm{I}_{ki}{}' \in \mathrm{SusDim}_{Eco}) = \begin{cases} 1 & \text{if Indicator } \mathrm{I}_{ki}{}' \text{ belongs to Economic Dimension} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{1}(\mathrm{I}_{ki}{}' \in \mathrm{SusDim}_{Soc}) = \begin{cases} 1 & \text{if Indicator } \mathrm{I}_{ki}{}' \text{ belongs to Social Dimension} \\ 0 & \text{otherwise} \end{cases}$$

Where, SusEnv_k, SusEco_k , and SusSoc_k is the composite sustainability score in the environmental, economic and social dimension of sustainability for census tract *k*. $W_{ki}^{d1}, W_{ki}^{d2}, W_{ki}^{d3}$ is the weight of indicator $\mathrm{I}_{ki}{}'$ in environmental, economic and social dimension of sustainability for census tract *k*, and $\sum_i W_{ki}^{d1} = 1, \sum_i W_{ki}^{d2} = 1, \sum_i W_{ki}^{d3} = 1$. $\mathrm{SusTotal}_k$ is the overall sustainability assessment result for census tract *k*. $W_{Env}, W_{Eco}, W_{Soc}$ is the weight of environmental, economic and social dimension of sustainability respectively, and $W_{Env} + W_{Eco} + W_{Soc} = 1$.

4.4 Indicator Evaluation Results

4.4.1 *Statistical Measures*

We compare the difference between the new value and the current value of each indicator. The difference of each indicator equals its new value minus its current value. Table 4-5 shows some statistics of the difference of each indicator. The meaning and unit of each indicator can be found in Table 4-4. The negative value means the indicator values decreased after moving interstates underground. The “counts” column records the number of census tracts which are affected by moving interstates underground. For example, the average commute time of people who live in 227 census tracts is affected by moving interstates underground. And the average commute time from home to the workplace is decreased by at most 0.55 minutes and 0.05 minutes on average for each work every day. 85 census tracts can decrease their population affected by noise pollution of interstates by 725 population on average after moving interstates underground. One of the 85 census tracts decreases its population affected by noise pollution by 3110. In other words, after moving interstates underground, fewer people would be bothered by noise pollutions produced by interstates near their home.

Table 4-5 Change of Indicators after Moving Interstates Underground

No.	Name	Min	Max	Median	Average	Counts
E11	Avg Time Spent in Traffic	-0.546	0.034	-0.033	-0.046	227
E12	Fuel Consumption	-0.073	0.023	-0.005	-0.006	192
E13	Connectivity of Transportation Network	-0.980	-0.002	-0.061	-0.109	71
E21	Nature & Biodiversity	0.034	134.254	21.645	26.746	69
E22	Urban Heat Island	0.004	0.170	0.026	0.039	70
E23	Noise Pollution Generated	-3110	-12	-585	-725	85
E24	Land Use: Green Space Ratio	0.004	0.170	0.026	0.039	70
E25	Land Use: Land Released	90.8	500110.2	77691.7	107600.9	71
E31	Accessibility to Green Space	0.002	3.605	0.112	0.289	107
E32	Accessibility to Health Facilities	0.001	24.410	0.038	0.438	120
E33	Accessibility to Education Facilities	0.001	7.575	0.049	0.321	73
E34	Equity	-0.004	0.010	0.001	0.001	48

The author also compares the change of CSI and the normalized individual indicators, their boxplots are shown in Figure 4-26. The black boxplots show the change of CSI, orange boxplots show the change of economic indicators (E11 and E13), green boxplots show the change of environmental indicators (E21, E22, E23, and E25), yellow boxplots show the change of social indicators (E31, E32, E33, and E34). We can see that CSI is improved after moving interstates underground. Environmental CSI has the largest improvement (i.e., CSI increased by 0.13 at most), economic and social CSI has very small improvement (i.e., < 0.1). Due to the property of the linear additive model, we can conclude that environmental indicators mainly contribute to the improvement of overall CSI. The individual normalized indicators tell the same story, the environmental indicators changes a lot, but economic and social indicators do not change that much.

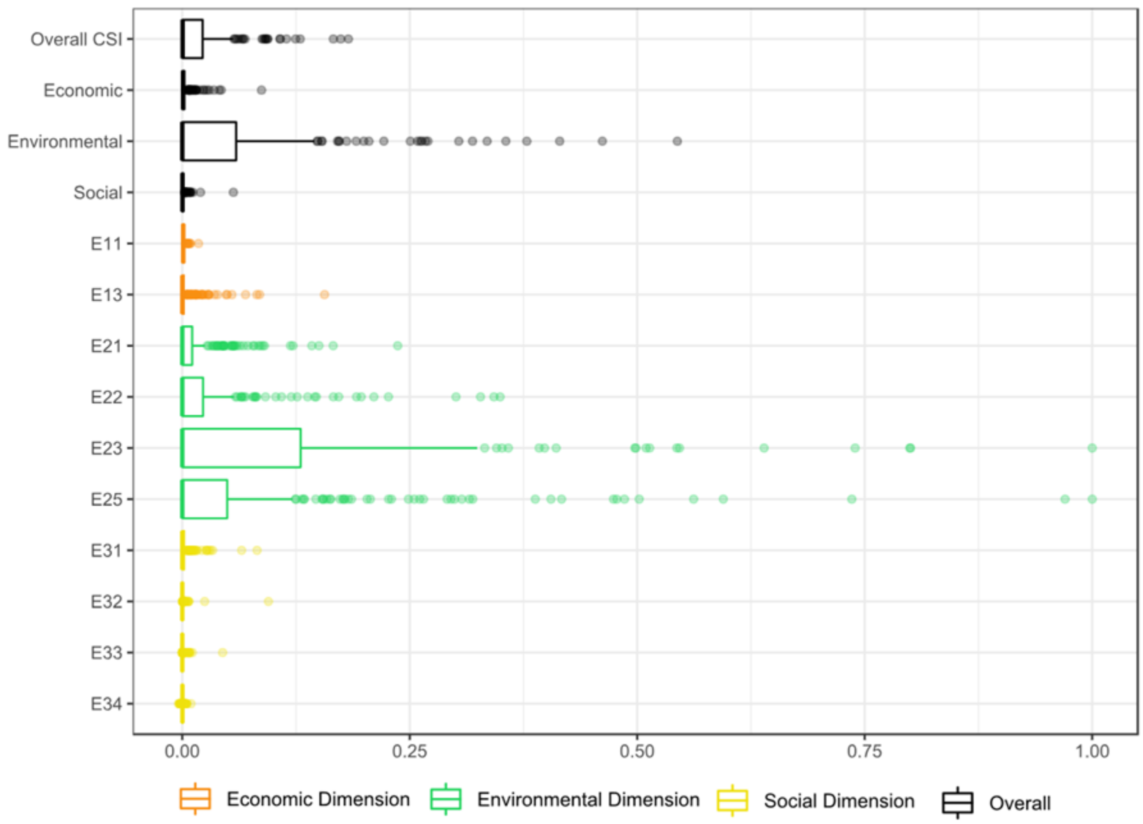


Figure 4-26 Boxplot of CSI and Normalized Individual Indicators (Difference)

Figure 4-27 and Figure 4-28 shows the statistical distribution of CSI and normalized individual indicators for the current road network and the new road network (after moving interstates underground.) These two figures show that in the linear additive aggregation model the poor performance in some indicators is compensated by sufficiently good performance in other indicators. For example, the low values of environmental and social CSI are compensated by the high value of economic CSI when using them to calculate the overall CSI. The very low value of environmental indicator E25 (with value zeros) is compensated by the very high value of another environmental indicator E23 (with a value larger than 0.8). The lower value of three social indicators (E31, E32, E33) are compensated by another one social indicator (E34), thus making the overall social CSI greater than 0.25. Therefore, the author suggests using weighted geo-mean aggregation model to generate CSI.

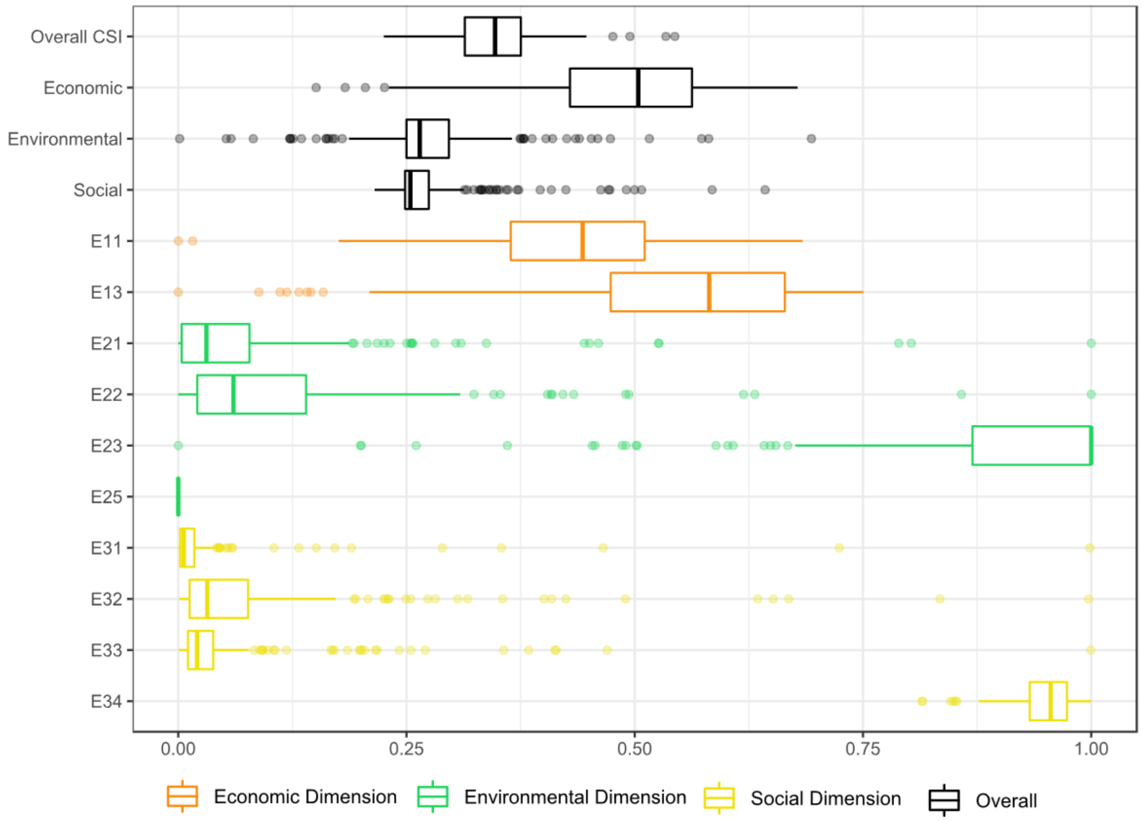


Figure 4-27 Boxplot of CSI and Normalized Individual Indicators (Current)

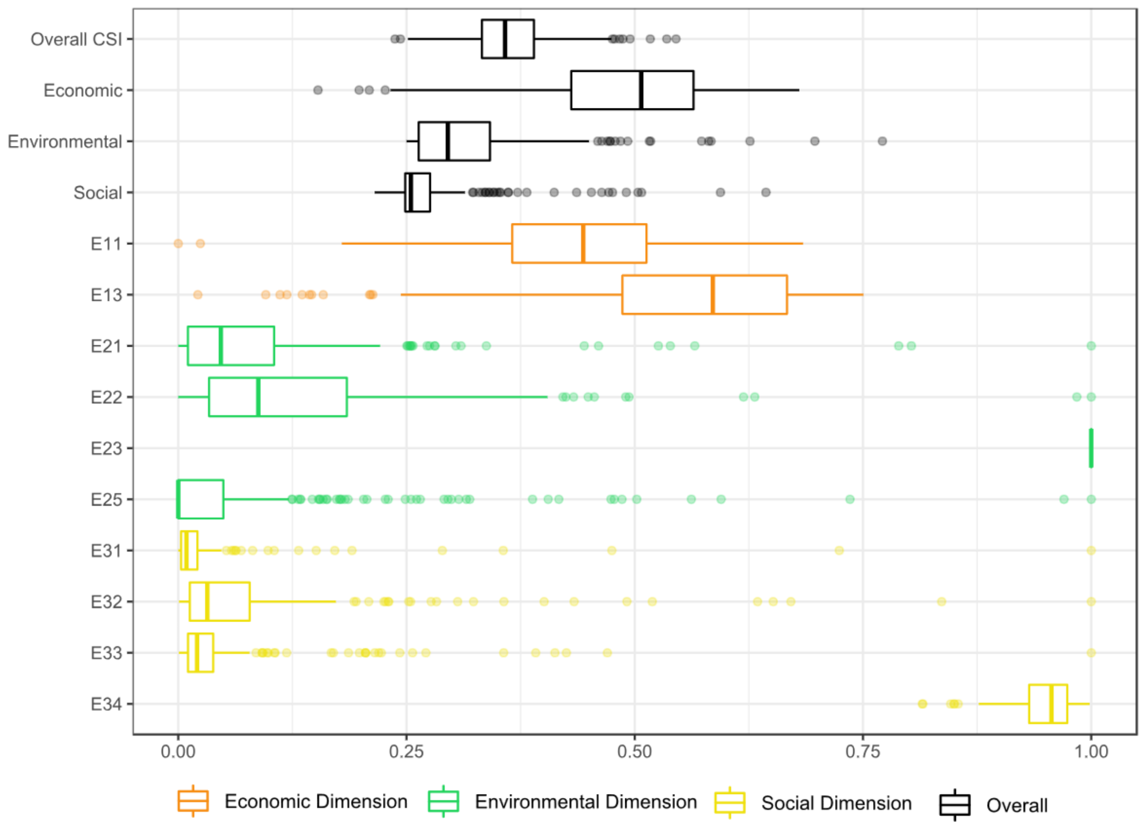


Figure 4-28 Boxplot of CSI and Normalized Individual Indicators (New: After Moving Interstates within I-285 Underground)

Figure 4-27 and Figure 4-28 also shows that Economic indicators (E11 and E13) has larger values than and Environmental social indicators, but also has bigger IQR than other indicators (i.e., values of economic indicators are more variable than the value of other indicators in the study area). Economic CSI also has larger IQR than environmental CSI and social CSI (i.e., values of economic CSI varies more than environmental CSI and social CSI). Besides, these two figures show the distribution of environmental CSI, social CSI, E22, E31, E32, and E33 are positively skewed in the study area, meaning that these indicators have relatively small values in most census tracts and has larger values only in a few census tracts. In other words, the people in most census tracts have small "green and water space ratio” (E22), and lower accessibility to green space, health facilities, and education facilities. That is to say, these indicators have much improvement in space in most census tracts of the study area.

However, indicator E13 and E23 are negatively skewed in the study area, meaning that these two indicators have relatively large values in most census tracts and has smaller values only in a few census tracts. In other words, the people in most census tracts have a larger detour index (E13) when they want to walk or drive from one place to another place 500 meters away. This may indicate that the road network in the study area has much improving space and the people’s walkability in the study area need to be improved concerning the sustainability of this study area. Indicator E23 represents the number of people affected by the noise pollutions generated by the interstates, its negative skewed distribution implies that most of the people in the study area are not affected by the noise pollution of interstates, and they live beyond the 200 meters buffer of interstates in the study area. This is consistent with the spatial distribution of population in study area in Figure 4-7. Since the skewed distribution of indicators, if one wants to use one value to represent the CSI of the whole study area by aggregating the CSI of all census tracts, the author suggests using median or inter-quantile mean to aggregate the values, and the aggregation can be achieved by “Aggregate Attributes” tool in the “Spatial Sustainability Assessment” plugin in QGIS.

4.4.2 *Correlation between Indicators*

Figure 4-29 shows the correlation matrix between all the normalized indicators. The indicators at the same dimension were put together, which was outlined by the box in

Figure 4-29. The figure shows that indicators among different sustainable dimensions are not correlated with each other. There are no correlated indicators in the social dimension. There are some correlated indicators in the economic and environmental dimension.

First, there are very strong correlations between E11 (Total travel time per capita) and E12 (fuel consumption per capita). Their correlation coefficient is 0.98, and the correlation is significant at the confidence level of 95%. Therefore, we can keep only one of them for the following assessment of sustainability. Because there are many uncertainties in the estimation of fuel consumptions, total travel time per capita (E11) is retained for the sustainability assessment. Therefore, E11 and E13 are used as economic indicators for the following sustainability assessment.

Second, E24 (Land Use: Green Space Ratio) also a very strong correlation with E22 (Urban heat island: Green & Water Ratio). Their correlation coefficient is 0.98, and the correlation is significant at the confidence level of 95%. E24 also correlates with E21 (Nature & Biodiversity), with a correlation coefficient is 0.77. But another land use indicator, land consumption (E25), has no relation with E22 and E21. Thus land consumption (E25) is retained to represent land use. Even E25 has a correlation with E23 (Noise Pollution), but they represent different aspects of environmental conditions, and their correlation coefficient is less than 0.8. Thus both of them are retained in the following sustainability assessment. Therefore, E21, E22, E23, and E25 are used as the environmental indicators for the following sustainability assessment.

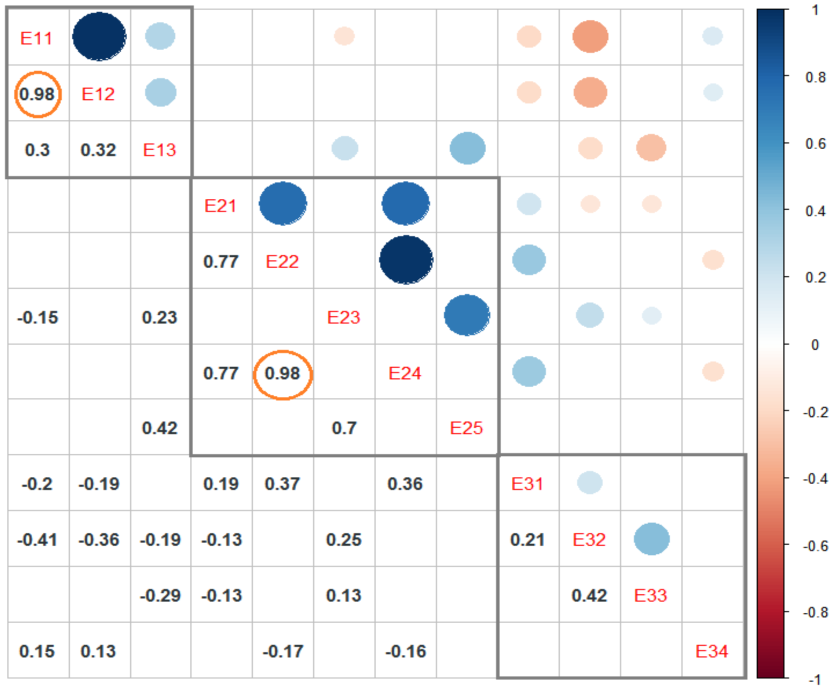


Figure 4-29 Correlation Matrix Between All the Normalized Indicators

After correlation analysis, ten indicators are retained for the MCA model. They are E11, E13, E21, E22, E23, E25, E31, E32, E33, and E34. Using the weight in Table 4-6, the author uses the linear additive model (see Equation 4.14) to estimate the overall sustainability or composite sustainability index in each dimension of sustainability. Their results are discussed in detail as follows.

$$SI_{Overall} = CSI_{Eco} \times \frac{1}{3} + CSI_{\mathrm{Env}} \times \frac{1}{3} + CSI_{\mathrm{Soc}} \times \frac{1}{3}$$

$$CSI_{Eco} = E11 \times (-0.5) + E13 * (-0.5) \tag{4.14}$$

$$CSI_{\mathrm{Env}} = E21 \times (0.25) + E22 * (0.25) + E23 * (-0.25) + E25 * (0.25)$$

$$CSI_{\mathrm{Soc}} = E31 \times (0.25) + E32 * (0.25) + E33 * (-0.25) + E34 * (0.25)$$

Table 4-6 Selected Sustainable Indicators for MCA Model and Its Weights

Dimension	Dimension Weight	No.	Name	Weight
Economic Sustainability	1/3	E11	Total Time Spent in Traffic	-0.50
		E13	Connectivity of Transportation Network	-0.50
Environmental Sustainability	1/3	E21	Nature & Biodiversity	0.25
		E22	Urban Heat Island	0.25
		E23	Noise Pollution Generated	-0.25
		E25	Land Use	0.25
Social Sustainability	1/3	E31	Accessibility to Green Space	0.25
		E32	Accessibility to Health Facilities	0.25
		E33	Accessibility to Education Facilities	0.25
		E34	Equity	0.25

4.4.3 *Economic Dimension Indicators*

There are two indicators in economic sustainability dimension, E11, and E13. Figure 4-30 shows the spatial distribution of current commute time (left) and difference of commute time difference (right). The current commute time has concentric rings style distribution. People who live downtown and mid-town have the shortest commute time (less than 15 minutes), and people who live around the south section of I-285 have the longest commute time (i.e., more than 20 minutes). However, the commute time can be reduced after moving interstates underground and reconnecting some streets that were cut or eliminated during the construction of interstates. The average commute time each worker spent each day (by driving) can be reduced by 0.55 minutes at most. The workers who live far from downtown and mid-town, especially those who live around I-285 are affected most by moving the interstates underground (e.g., the dark orange tracts around I-285). These workers usually have a longer commute time than others. Note that the author makes some assumptions when calculating the commute time, which is too simple to represent the real travel commute time, and underestimates the commute time, especially for people who do not own a personal vehicle. Therefore, after moving interstates underground, the commute time of these people may be improved a lot, and we can see a larger decrease of commute time than the right map shown in Figure 4-30.

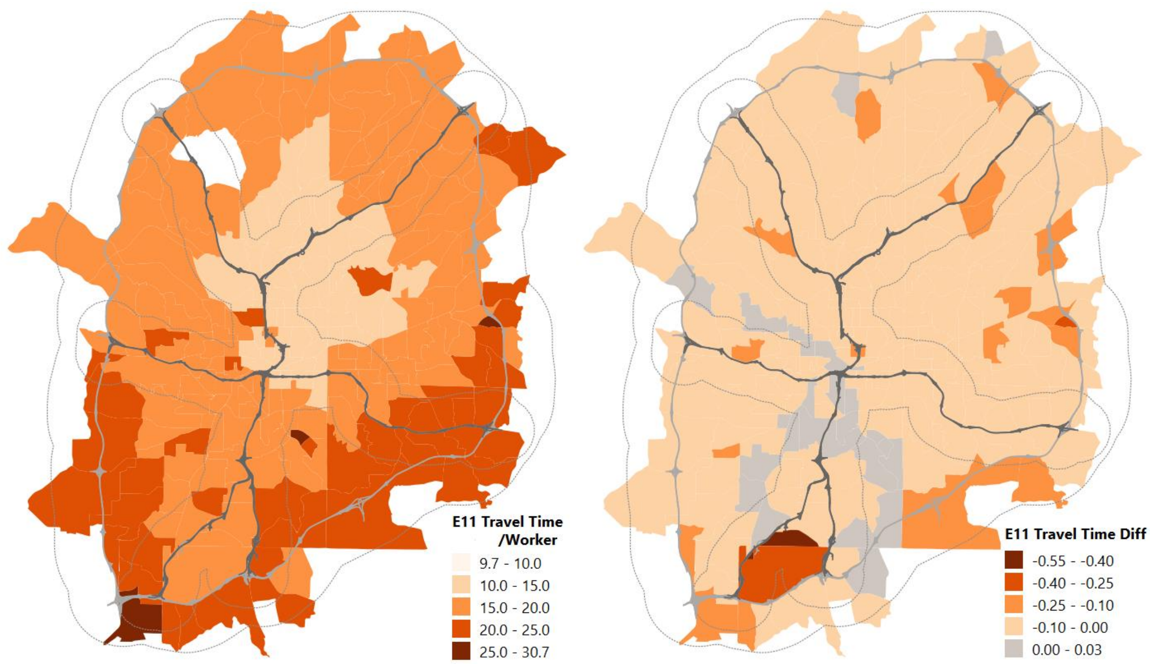


Figure 4-30 Current Average Commute Time Each Worker Spent Everyday (left) and Difference of E11 After Moving Interstates Underground

As mentioned in the last section, DI value between two points is affected by their straight-line distance. The author analyzed how the DI value changes with the increase of straight-line distance and how the interstates affect DI values. From Figure 4-31 below, we can see that DI decreases with the increase in traveling distance. When the straight-line distance longer than 3500 meters, the DI distribution is almost the same, and DI difference between cross-interstates travel and non-cross-interstates travel becomes very small. In other words, long distance traveling is less affected by the disconnect caused by interstates, especially when the traveling distance is longer than 3.5km. This also means interstate highways mainly affected the people who lived within 3.5km of the interstates and mainly affect their short distance travel (< 3.5km) which is from one side of an interstate to the other side of an interstate. DI difference between cross-interstates travel and non-cross-interstates travel is very large when the straight-line distance is less than 1km. The author set the straight-line distance to be 500 meters and analyzed the change of DI in details, the results are shown in Figure 4-32.

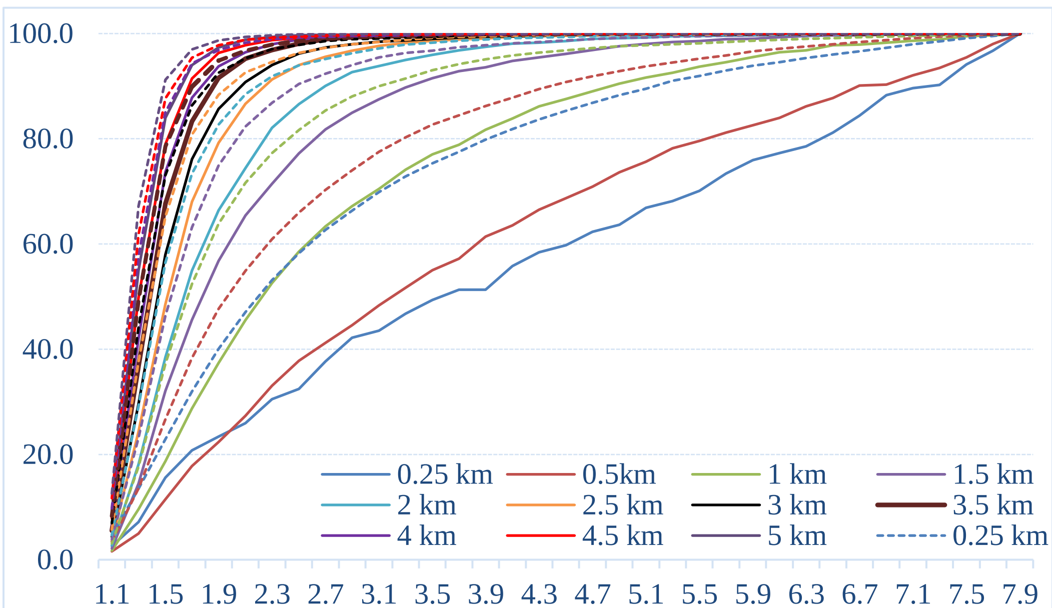


Figure 4-31 Cumulative Distribution for DI of Cross-interstates (solid lines) and Non-cross-interstates (dash lines) Trips with Different Length in Current Network

Figure 4-32 shows the current detour index (left) and difference of detour index (right). The left map in Figure 4-32 shows that there are larger DI in the census tracts that are close to interstates than other places that are far from interstates. This may indicate that the interstates indeed have impacts to DI and make people travel a longer distance to move from one place to another place that is 500 meters away. And midtown and downtown areas have smaller DI than other places (i.e., their DIs are less than 2.0). This makes sense because the downtown and midtown area have more density road network than other places. The right map in Figure 4-32 shows that DI of all census tracts decreased after moving interstates underground. In other words, people travel a shorter distance when traveling from one place to another place which is 500 meters away from or to the affected census tracts, after moving interstates underground. The DI value decreases more in census tracts that are located within the 2km buffer of interstates. This makes sense because we set the Euclidean Distance of each trip to be 500 meters when calculating the detour index. The reduction of DI around interstates can help increase the equality of connectivity to the road network. From this point of view, moving interstate highways underground can yield more equitable transportation systems, and may lead to a more sustainable highway system than the current one. Note that the change of DI values depends on the new road network, but the author only recovered the road that is cut or eliminated during the construction of interstates, even though it is likely we would add or reconnect more roads after moving interstates underground. Thus the DI can be reduced further. Therefore, we will have a better improvement in DI than the right map shown in Figure 4-32.

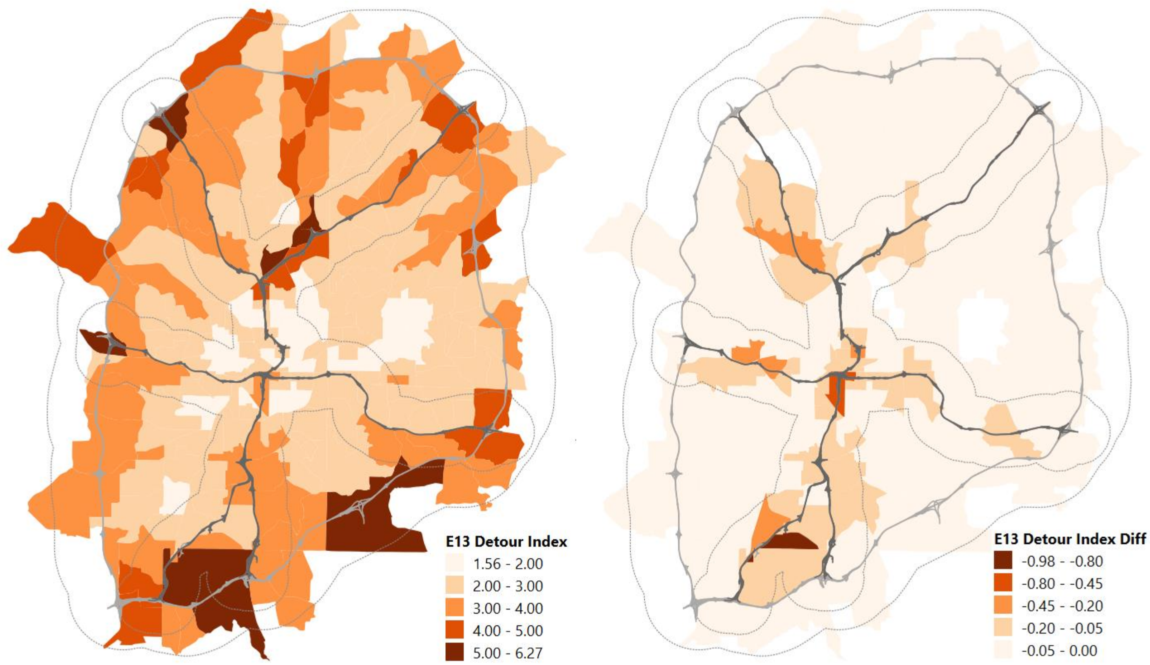


Figure 4-32 Current Detour Index for 500 Meters-Trips (left) and Difference of Detour Index (right) After Moving Interstates Underground

4.4.4 *Environmental Dimension Indicators*

There are four indicators in the environmental dimension of sustainability, including green area per capita (E21), green water ratio (E22), noise population (E23), and land release (E25). Their value change after moving interstates underground can be found in Figure 4-33, Figure 4-34, and Figure 4-35. From the left map in Figure 4-33, we can see the green space area per capita is very unevenly distributed in the study area, and its value is less than 100 (m^2/capita) in most census tracts. However, after moving interstates underground and turning the released land into green space, the green space per capita in each census tract was improved a lot, especially for the census tract who are close to interstates (see the right map in Figure 4-33).

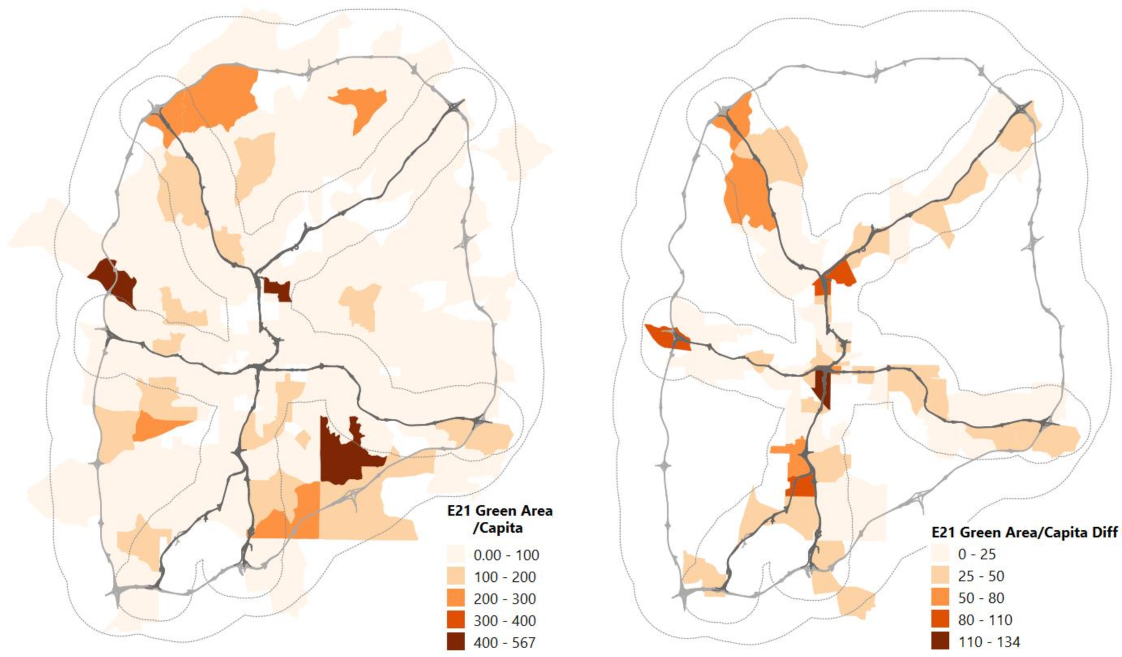


Figure 4-33 Current Green Areas per Capita (left) and Changes of E21 (right) After Moving Interstates Underground

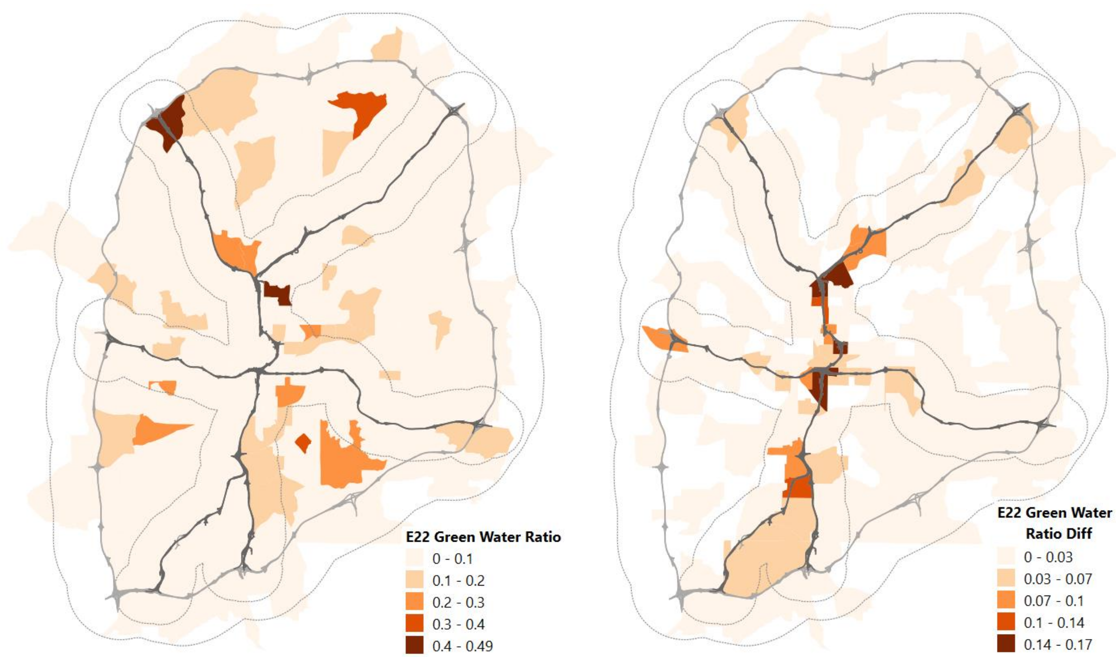


Figure 4-34 Current Green Water Ratio (left) and Changes of E22 (right) After Moving Interstates Underground

Green & Water Ratio shows the same pattern as the green space per capita; In most census tracts, the green & water ratio is less than 0.1. In other words, most census tracts have less than 10% of their areas are covered by green space or water (see the left map in Figure 4-34). After moving interstates underground, the green & water ratio in census tracts that are located within a 2-km buffer of interstates are improed by up to 17% (see the right map in Figure 4-34). The change of population affected by interstate noise is shown in the left map in Figure 4-35. People who lived in the census tracts that are located in the 2-km buffer of interstates are the ones mainly affected by the noise pollution of interstates. The affected population is unevenly distributed in the 2-km buffer of the interstates (most census tracts have less than 1000 people affected by noise pollution, but some have more than 2000 or even 3000 people affected by noise pollution).Therefore, from the noise pollution aspects, residents who live nearby interstates benefit a lot from moving interstates underground. The right map in Figure 4-35 shows that the census tracts which are located along interstates (most in the 2-km buffer of the interstates) have many lands released after moving interstates (some tracts have more than 0.4 km^2 surface areas are leased for green spaces or other purposes). Overall, the census tracts that are close to interstates (mainly in the 2-km buffer of the interstates) are more likely to benefit from the action of moving interstates underground, from the environmental sustainability aspect. However, since the newly released land after moving interstates underground can also be used for other purposes besides green space, the estimation of indicator E21, E22 may be over-estimated. The estimation of E23 may be over-estimated, because it is the size of the population living in the 200 meters of buffer of the interstates, assuming population in each census tract is evenly distributed. But the real population is not evenly distributed, and there is the spatial difference in the population distribution at lower spatial scale (e.g., census blocks). Further, there may be less population than other places in one census tract. The estimation of E25 is under-estimated because the estimated width of interstates is its lower-bound.

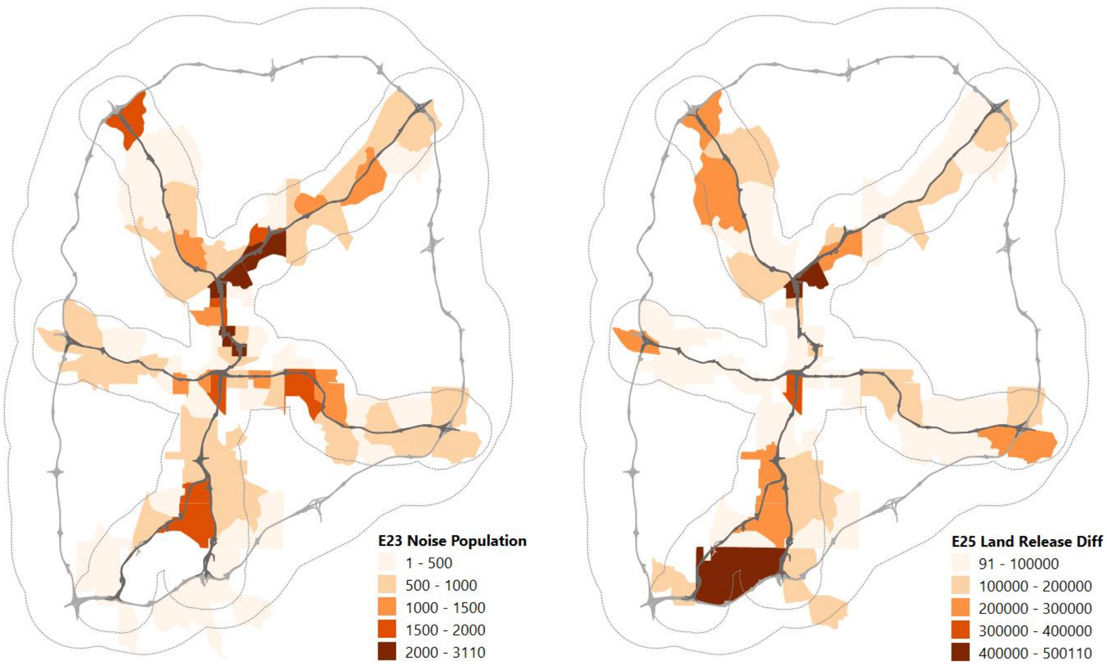


Figure 4-35 Difference of Noise Pollution (left) and Land Released After Moving Interstates Underground

4.4.5 Social Dimension Indicators

There are also four indicators in the social dimension of sustainability, including accessibility to green space (E31), accessibility to health facility (E32), accessibility to education facilities (E33), and income equity index of commute distance (E34). Their current value and change of values after moving interstates underground can be found in Figure 4-36, Figure 4-37, Figure 4-38, and Figure 4-39.

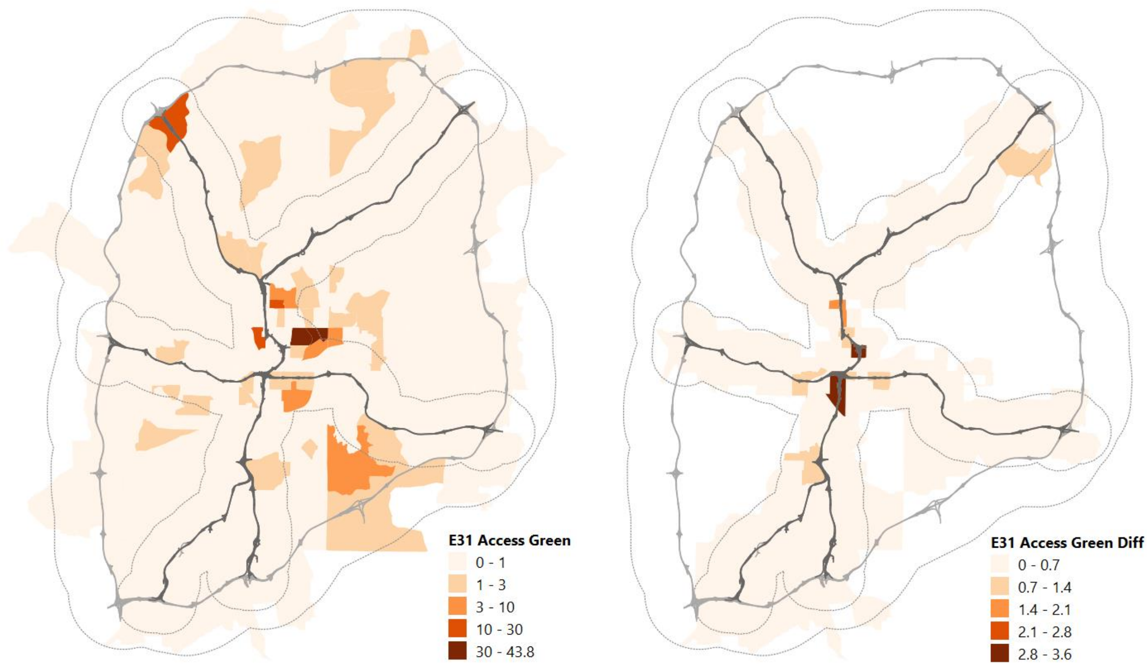


Figure 4-36 Difference of Accessibility to Green Space

First, people's accessibility to green space is unevenly distributed in the study area (see the left map in Figure 4-36). It has similar spatial distribution as indicator E21 (green area per capita), except that for census tracts that are on the east of the Connecter people still has relatively high accessibility to green space even they do not have green space or has very less green space in their census tract. In other words, people who live in these census tracts (east of the Connecter) can easily reach a green space in other census tracts within 800 meters. On the contrary, some census tracts have relatively low accessibility to green space even their green space per capita is more than 300 m^2/capita. For example, the census tract that are located in the intersection of I-75 and I-285 (i.e., the big orange polygon at the left corner on the left map in Figure 4-33). This may be caused by their high detour index (more than 3). The right map in Figure 4-36 shows that people's accessibility to green space is improved, especially those who live in the 2-km buffer of interstates. This makes sense and corresponds to our anticipation, because we have more green space at these census tracts and less detour index after moving interstates underground.

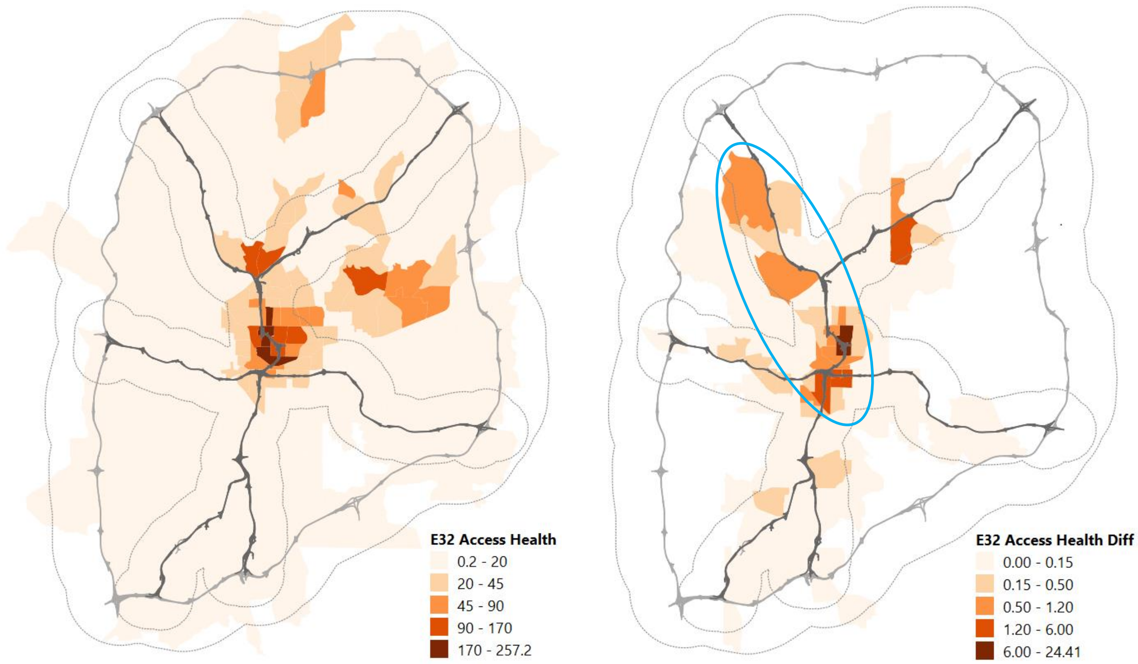


Figure 4-37 Difference of Accessibility to Health Facilities

Second, people's accessibility to health facilities is also unevenly distributed in the study area (see the left map in Figure 4-37). People who live in downtown and midtown Atlanta, Decatur county, and Sandy Springs tend to have higher accessibility to health facilities than people who live in other places. However, most of the people in the study area do not live in thse regions (i.e., midtown and downtown Atlanta, Decatur, and Sandy Springs). That is to say, there is still much space to improve people's accessibility to health facilities. The right map in Figure 4-37 shows that people's accessibility to health facilities is improved in the 2-km buffer of interstates, especially for the tracts besides the Connector (where I-75 and I-85 are merged) and north section of I-75. This is because moving interstates underground decreased people's driving distance from home to health facilities within 3200 meters, especially for people who live along the Connector and north section of I-75. And the improvement can be much more than the right map shown in Figure 4-37, because people are likely to have shorter driving distance as mentioned in the previous section.

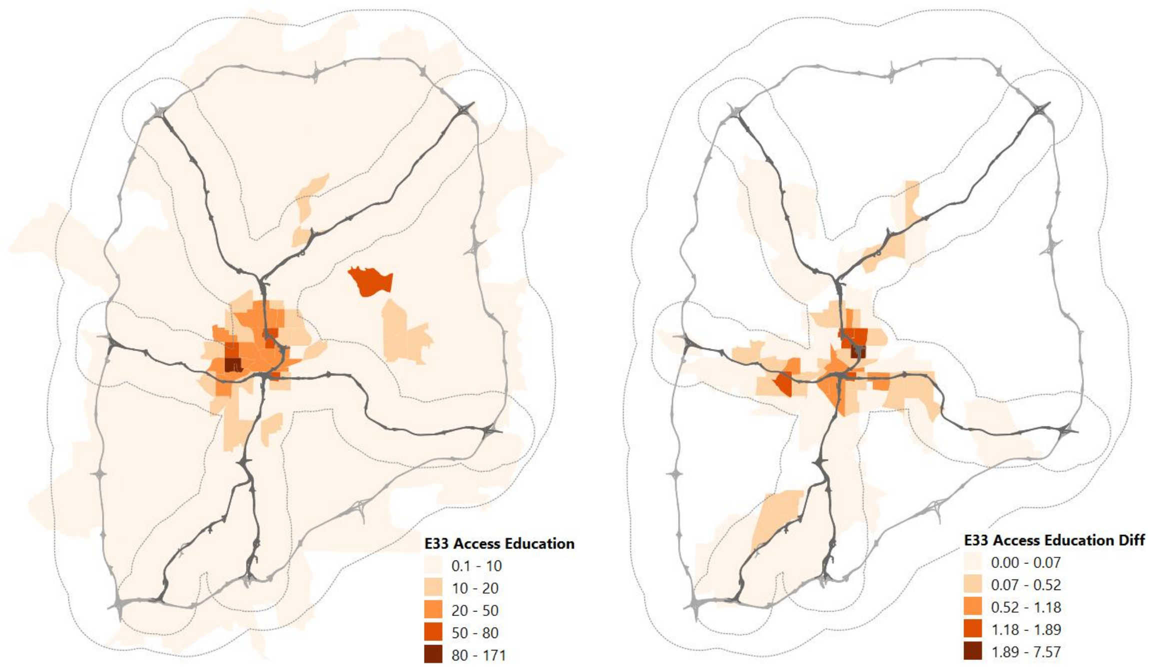


Figure 4-38 Difference of Accessibility to Education Facilities

Third, people's accessibility to education facilities is much skewed distributed in the study area (see the left map in Figure 4-38). The accessibility of most census tracts is less than 10. However, for census tracts in the downtown and midtown Atlanta, the accessibility can reach up to 171. However, there is less population in these areas (see population distribution in Figure 4-7). In other words, the education facilities in the study area do not fully service the residents, because most people need to travel a long distance to school. The right map in Figure 4-38 shows that people's accessibility to education facilities of census tracts that are within the 2-km buffer of interstates are changed after moving interstates underground, especially those that are located in downtown and midtown Atlanta. This is because moving interstates underground decreased the driving distance from home to health facilities within 1600 meters for people who live in these areas. And the improvement can be much more than the right map shown in Figure 4-38, because people are likely to have shorter driving distance as mentioned in the previous section.

The left map in Figure 4-39 shows the current income equity index of commute distance in each census tract. The equity condition in the study area is good and all census tract has equity index greater than 0.8. There is no obvious spatial pattern, but the census tract with lots of high income population (i.e., the census tracts with more than 1700 population with annual income > 40K, detail see the right map in Figure 4-8) tends to have relatively lower income equity index (< 0.9). The right map in Figure 4-39 shows that some census tract's equity decreased, some census tract's equity increased. And the equity of census tracts in the north of I-20 tends to be increased. But equity of census tracts located near the Airport was increased most, which is up to 0.01. The estimation of income equity index only considered three income classes (low, medium, high) and the commute distance does not consider the difference of people’s transportation mode (e.g., drive, bus, rail, by bike, or walk). And Further, the author assumes everyone drives to work every day. However, it is likely that people who take the bus or commute by bicycle or walk have longer travel commute distance than others. Therefore, the indicator E34 may be over-estimated, and its value may have larger spatial variance, and the current equity condition is worse than the left map in Figure 4-39, its improvement is larger than the right map in Figure 4-39. Overall, social sustainability is improved after moving interstates underground, and the real improvement can be larger than the estimation in this study if based on more data.

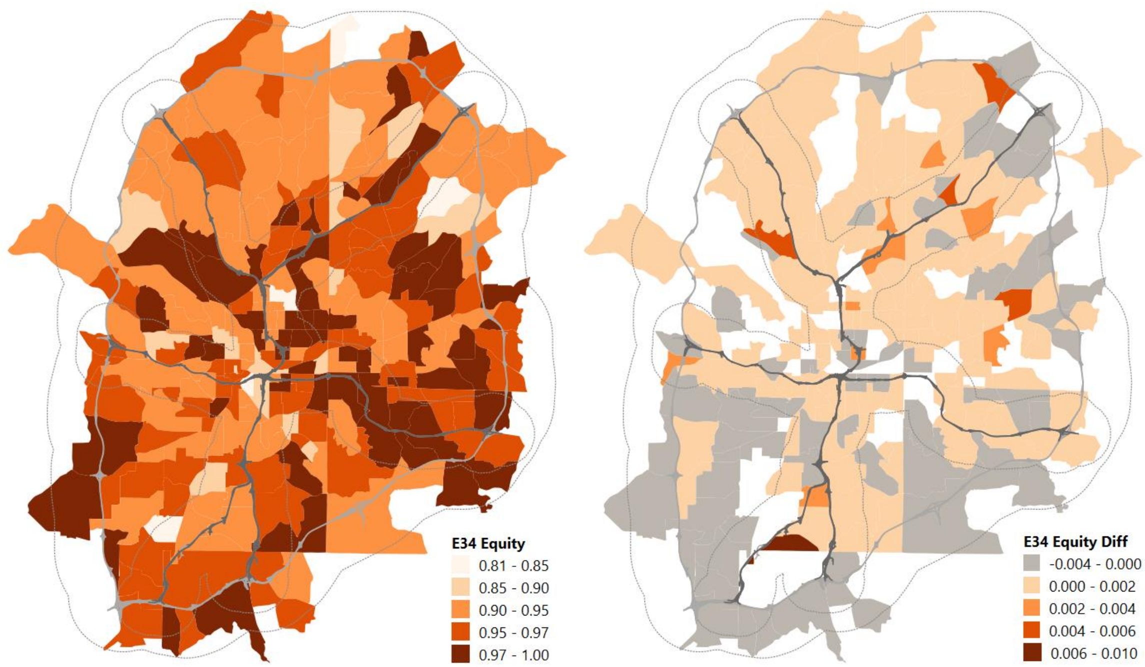


Figure 4-39 Difference of Income Equity to Commuting Distance

4.4.6 Composite Sustainability Index (CSI)

The difference of Composite Sustainability Index is shown in Figure 4-40, Figure 4-41, Figure 4-42 and Figure 4-43. They show the change of overall CSI, Economical CSI, Environmental CSI, and Social CSI respectively. These four figures show the same pattern and tell the same story. From these maps, we can see that the overall sustainability was improved in the study area, especially for the census tracts that are located within the 2-km buffer of interstates. But each census tract has different improvement, the darker the color, the higher the improvement. Among all the tracts within 2-km buffer of interstates, the dark-red tracts (outlined by the bright blue box in Figure 4-40) have the highest improvement of overall CSI; their overall CSI values are increased by more than 0.15. These results suggest that we can move those sections of interstates underground first, and we can conduct more detail sustainability analysis in these highly changed tracts, using lower scale data (e.g., census blocks or census block groups).

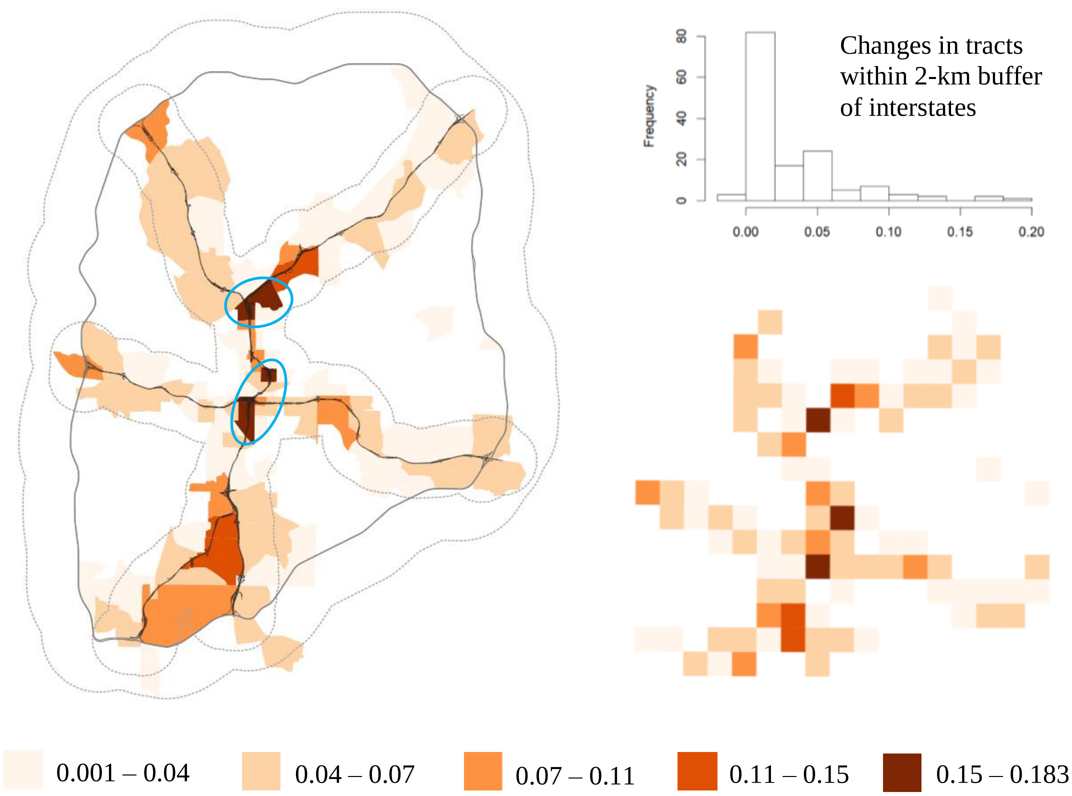


Figure 4-40 Difference of Overall Composite Sustainability Index

Besides, the sustainability at each dimension is also improved. The change of economic CSI is less than 0.09, and the changes in economic CSI in most tracts is less than 0.02. Only one tract (in dark red color) that in the south of the airport has larger than 0.07 improvement and another two tracts have improvement more than 0.03 and less than 0.05. So the economic CSI does not improve a lot in the study area, based on the current estimation of economic indicators. However, the economic CSI improvement should be much larger than the current value, because we overestimated the new value of E11 and E13. For example, we assume that the traffic capacity is the same as the current situation (i.e. with interstates above ground). However, after moving interstates we can easily add new lanes, increase the traffic capacity of interstates, and connect or build more surface roads to connect the roads that are cut or split by current interstates. Therefore, the new real value of E11 and E13 will be smaller than their current new value that estimated in the pilot study.

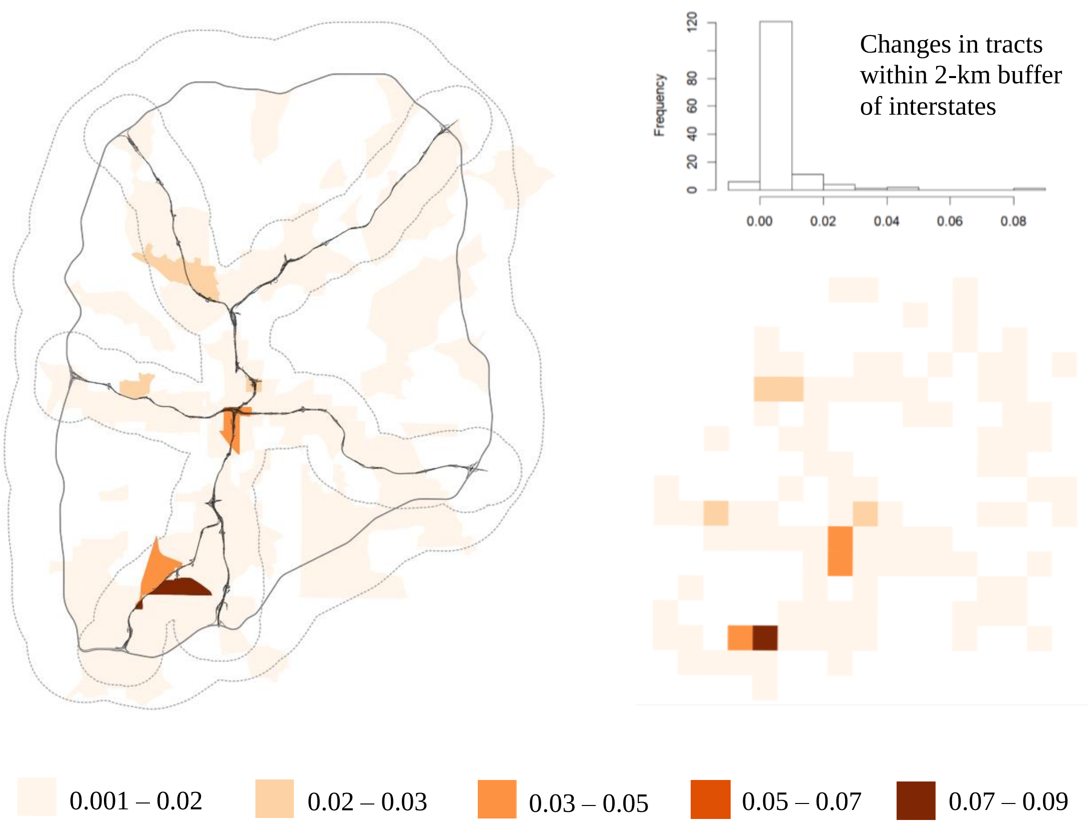


Figure 4-41 Difference of Economic Composite Sustainability Index

Changes of environmental CSI has the same spatial pattern as overall CSI. But the change values of the environmental sustainability indicators may be over-estimated because we assume all the surface lands released after moving interstates are simply converted to green space, but it is likely that they will also be converted to other types of land-uses. However, the area of land released after moving interstates underground is under-estimated, because the estimation of the width of interstates is lower than its real width.

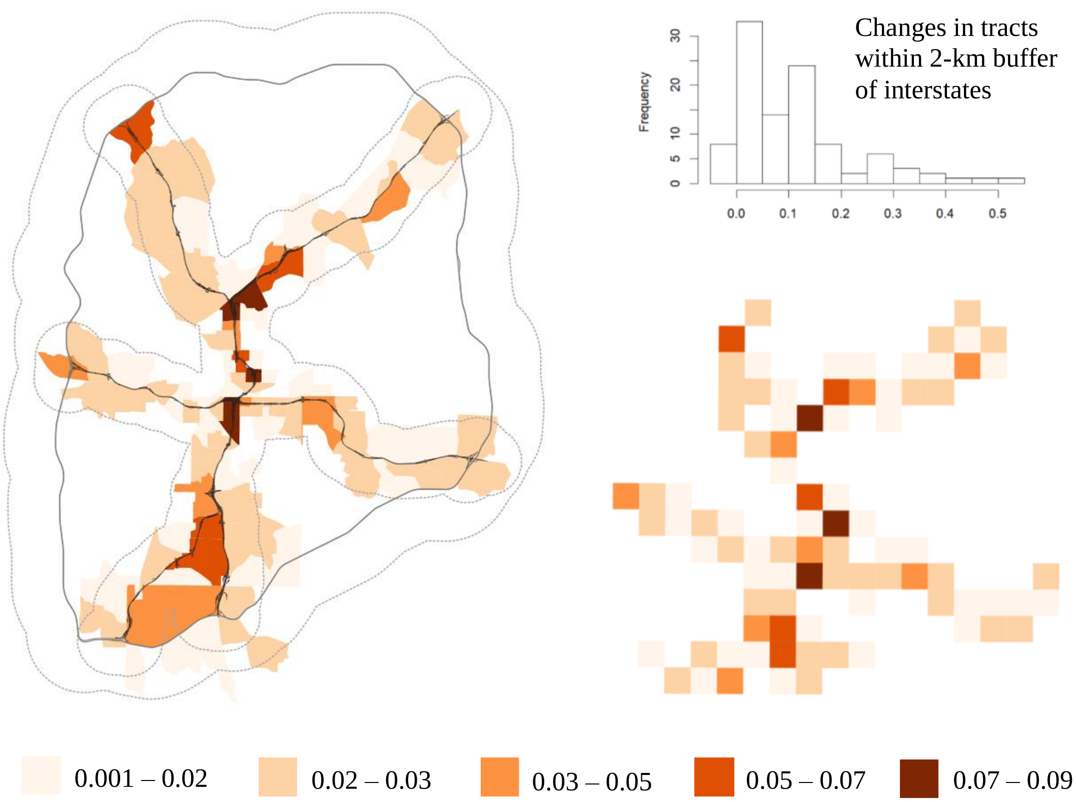


Figure 4-42 Difference of Environmental Composite Sustainability Index

Social CSI does not change very much. But if we have more detail data about health facilities, education facilities, and people's transportation mode (e.g., driving by car, taking bus, taking train, riding bicycles, or walking), and include more social indicators to estimate social CSI, then the change of social CSI will be much larger than the estimation results in this initial study.

In summary, this sustainability assessment results are promising and may indicate that moving interstates underground make the study area more sustainable, especially moving the sections within tracts that are outlined by the bright blue box in Figure 4-40. However, further studies are needed to evaluate the sustainability of moving interstates underground.

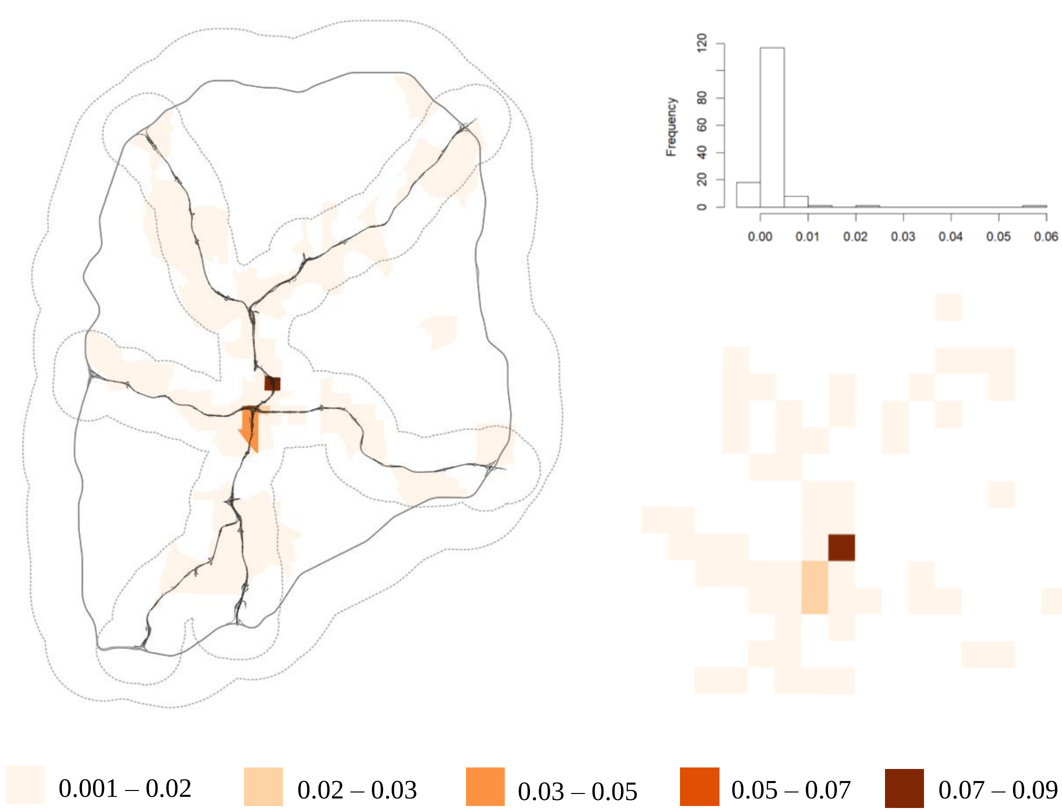


Figure 4-43 Difference of Social Composite Sustainability Index

4.4.7 Uncertainty and Sensitivity Analysis

After constructing the CSIs, the author conducts the uncertainty and sensitivity analysis of the sustainability assessment results (CSIs), using the "Uncertainty Analysis" tool in the implemented "Spatial Infrastructure Sustainability Assessment" plugin. The author evaluates the uncertainty of CSI at 90% confidence level when changing the weight of indicators within ±10% of its original value. The boxplot of uncertainty analysis results is shown in Figure 4-44 and Figure 4-45 respectively for the current network and the new road network (after moving interstates underground). From these two figures, we can see that both overall CSI and dimensional CSI have very narrow confidence interval, which indicates there is less uncertainty in the original sustainability assessment results. However, their coefficient of variance (i.e., e0_cv, eco_cv, env_cv, and soc_cv) are different, meaning they have different variability. Overall CSI and economic CSI has less variance than environmental CSI and social CSI. Also after moving interstates underground, the coefficient of variance of environmental CSI is reduced in some census tracts. In other words, there is less uncertainty in the environmental CSI of new road network compared with the current road network (both env_ci and env_cv are reduced after moving interstates underground, comparing the corresponding values in Figure 4-44 and Figure 4-45).

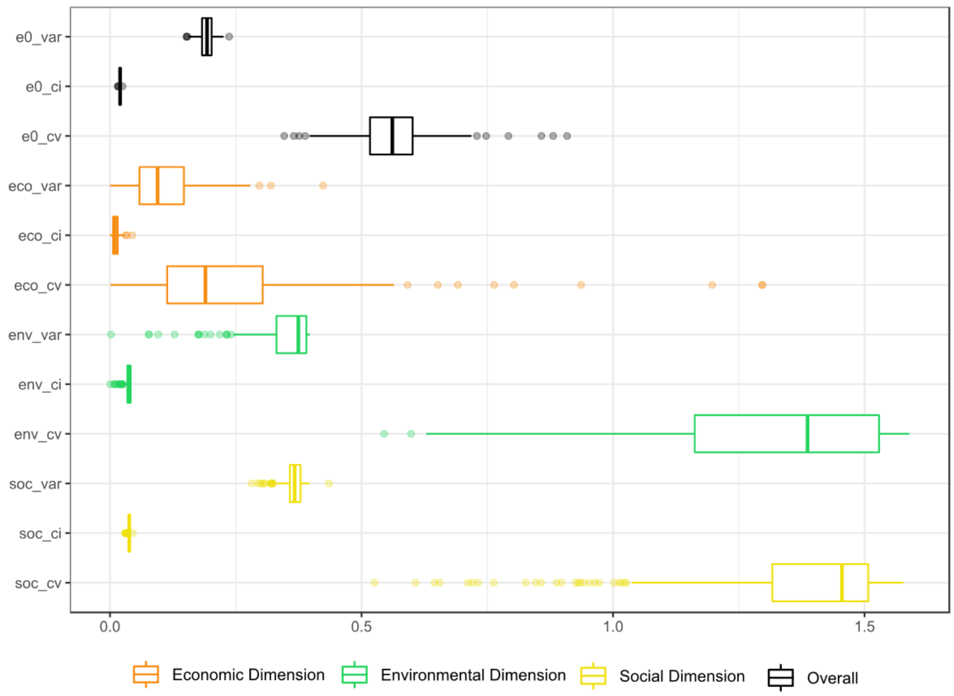


Figure 4-44 Boxplot of Uncertainty Analysis Results (Current)

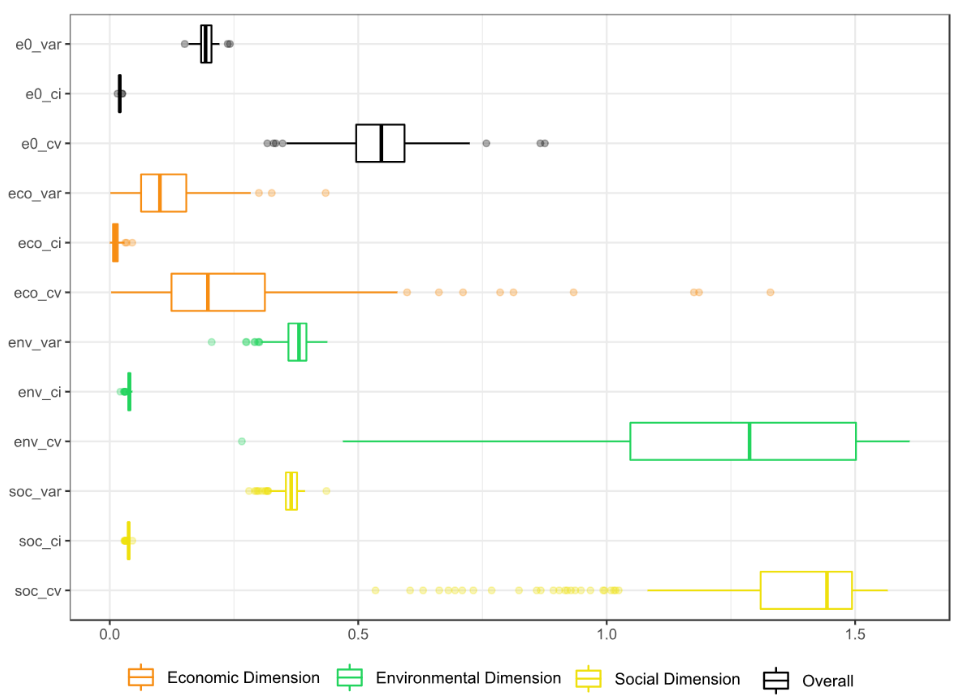


Figure 4-45 Boxplot of Uncertainty Analysis Results of CSI (New)

The width of the confidence interval (at 90% confidence level) in each census tract for current and new road network is shown in Figure 4-46 and Figure 4-47 respectively. The width of the confidence interval for overall CSI of each census tract in the study area is mainly located in range [0.01, 0.03]. The width of the confidence interval for Economic CSI of each census tract in the study area is in range [0.0, 0.02]. The width of the confidence interval for environmental CSI and social CSI of each census tract in the study area is in range [0.03, 0.045]. In other words, if we change the weight of each indicator within ±10% of its original value, at 90% confidence level the change of overall CSI is less than ±0.03, the change of Economic CSI is less than ±0.02, the change of environmental CSI and social CSI is bigger than ±0.03 and is less than ±0.045. From maps in Figure 4-46 and Figure 4-47, we can also see that the economic CSI and overall CSI for each census tract have narrower confidence interval than its environmental CSI and social CSI.

The coefficient of variance (CV) in each census tract for current and new road network is shown in Figure 4-48 and Figure 4-49 respectively. Overall CSI and Economic CSI has a very small CV, environmental CSI and social CSI have a relatively large CV. There is some change in CV after moving interstates underground, for example, the CV of new overall CSI and new environmental CSI in some census tracts are reduced, and these census tracts are mainly located within the 2-km buffers of interstates.

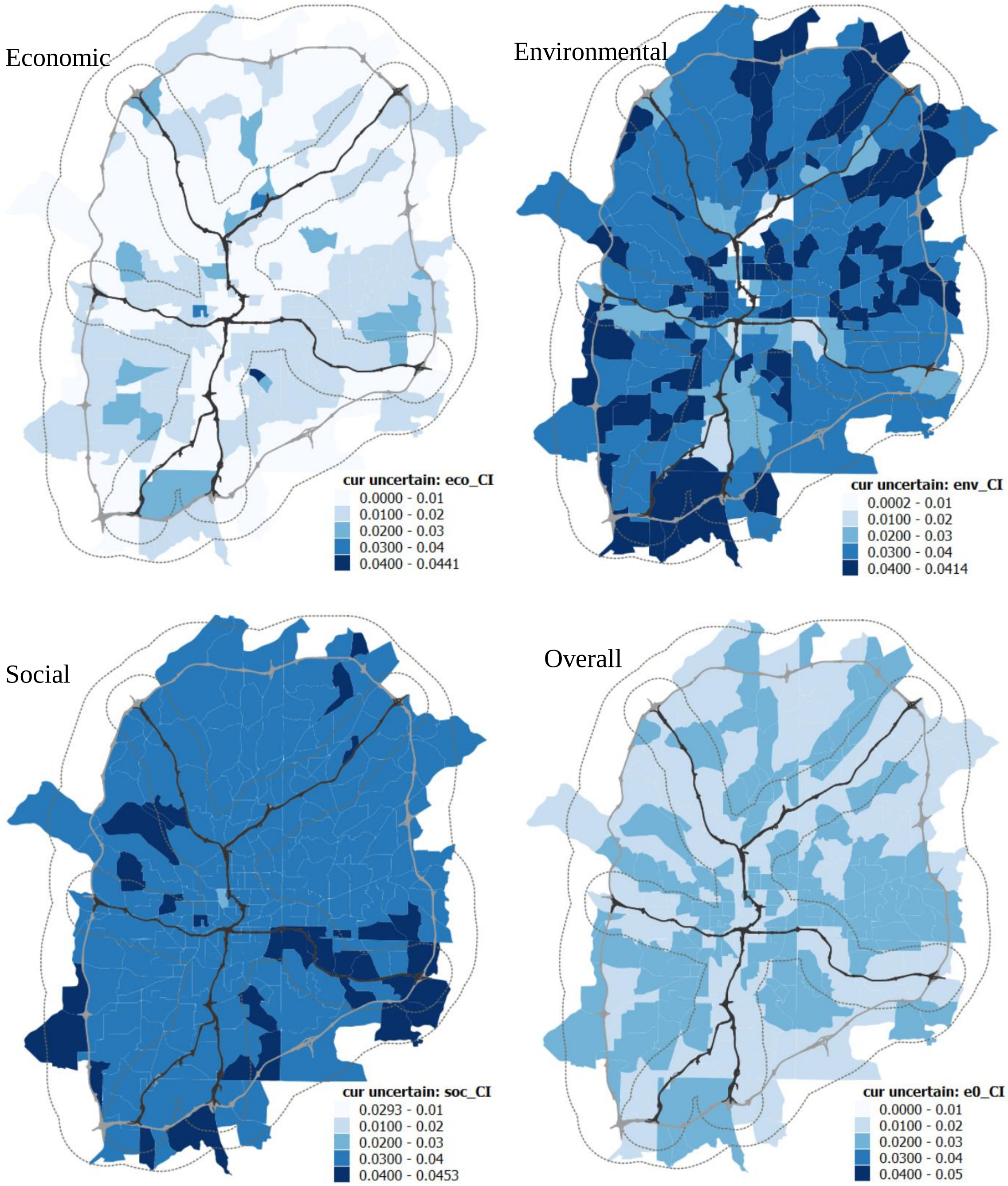


Figure 4-46 Confidence Interval of the CSI (Current)

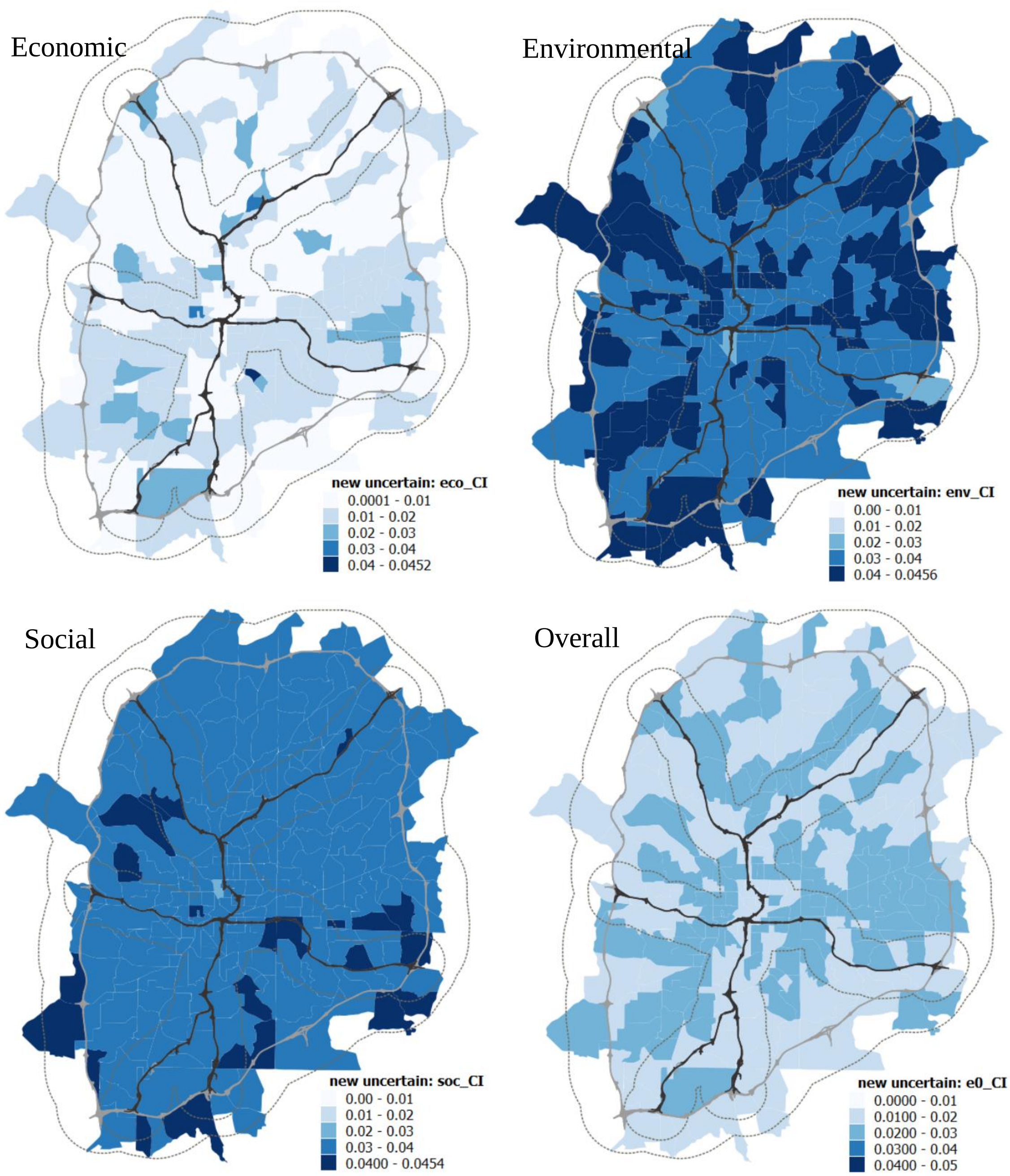


Figure 4-47 Confidence Interval of the CSI (New)

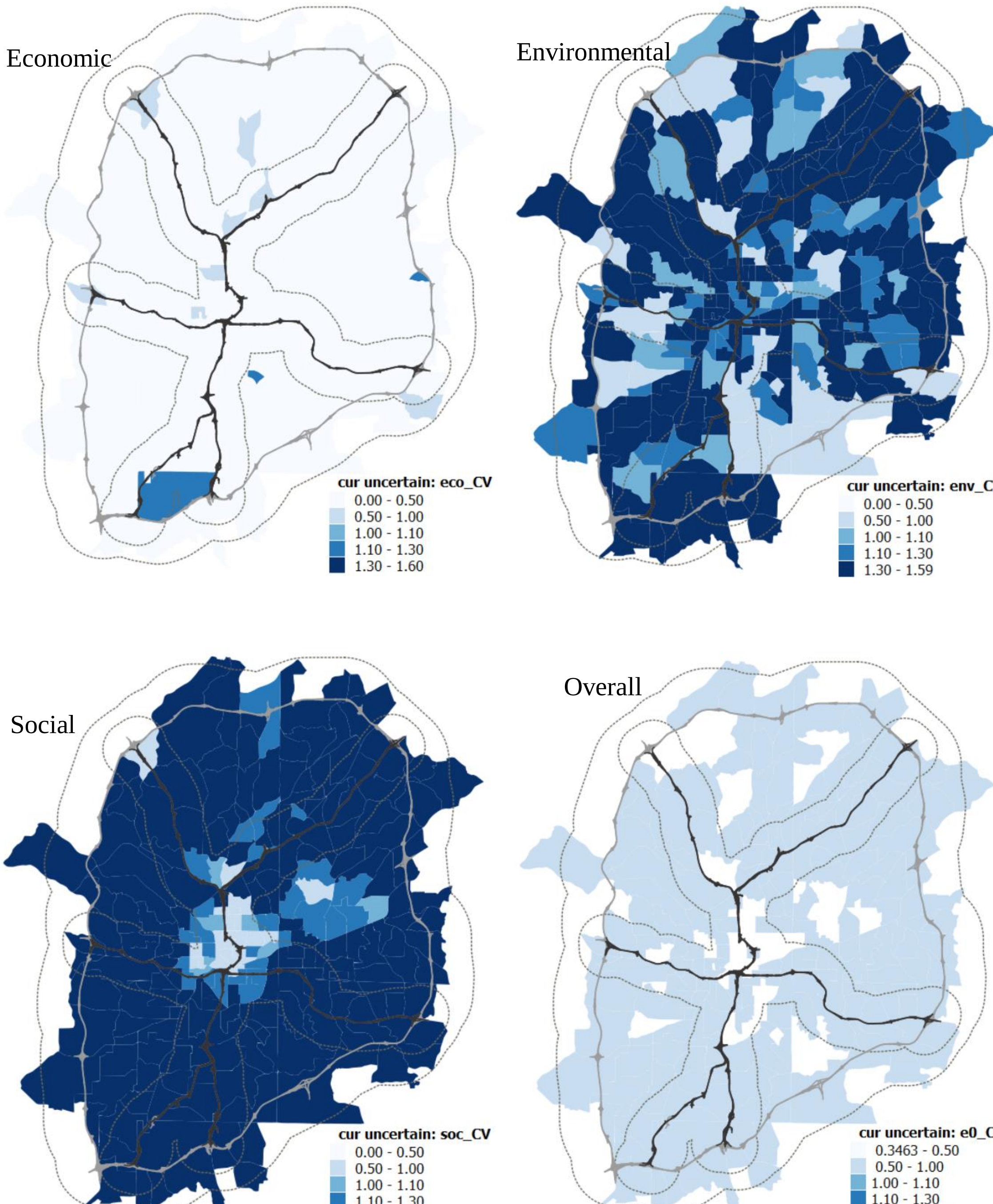


Figure 4-48 Coefficient of Variance of CSI (Current)

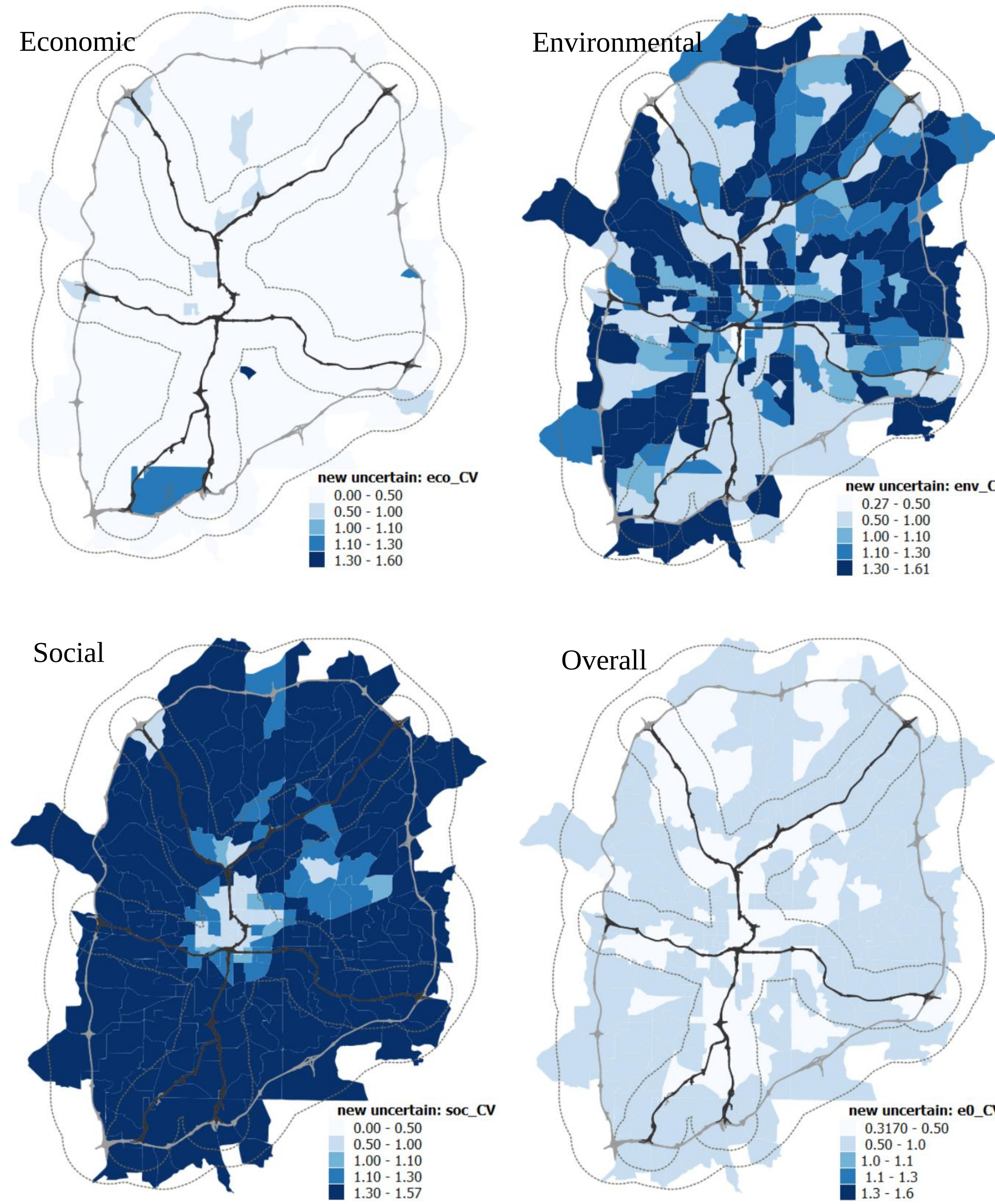


Figure 4-49 Coefficient of Variance of CSI (New)

Sensitivity analysis results are shown in Table 4-7. The sensitivity analysis results for both the current and new indicator's values are almost the same. From the first-order sensitivity index (Si), economic indicators (E11 and E13) contributes 18% and 22% to the overall CSI value respectively. Indicator E34 contribute less than 0.3% to the overall CSI. The other indicators contribute almost the same, which is about 9.8% of the overall CSI. From the total-effect sensitivity index (STi), the economic indicators (E11 and E13) also play an important role in the estimation of overall CSI. Because of the difference between first-order sensitivity index and total-effect sensitivity index, one can conclude that there are no interactions between these indicators. From the sensitivity analysis results, if there is more improvement in the economic indicators, which are very likely to happen, then the overall CSI value will be affected most. However, things could change if we introduce more economic indicators into this sustainability assessment, or use different aggregation models to build the CSI.

Table 4-7 Sensitivity Analysis Results for Current and New Values of Indicators

	Sensitivity	E11	E13	E21	E22	E23	E25	E31	E32	E33	E34
Current	S_i (%)	**17.6**	**22.1**	9.8	10.1	9.7	0.0	9.8	9.6	9.8	0.3
	S_{Ti}	**.183**	**.223**	.099	.099	.096	.000	.098	.098	.100	.003
New	S_i (%)	**18.1**	**21.3**	10.1	10.3	0.0	10.0	10.0	9.8	9.9	0.3
	S_{Ti}	**.186**	**.213**	.100	.101	.000	.100	.100	.100	.101	.003

4.5 Conclusion

In this chapter, the author evaluates the sustainability of moving interstates within the Perimeter in Atlanta underground and compare it with the sustainability performance of the current road network. Based on the objective of this pilot study, the author chooses twelve indicators for the sustainability assessment, and use spatial analysis and indicator analysis tools in the "Spatial Infrastructure Sustainability Assessment" plugin to evaluate the value of each indicator. After correlation analysis of the sustainability indicators, ten relatively independent indicators are used for the MCA. The author uses AHP tools in the "Spatial Infrastructure Sustainability Assessment" plugin to derive the weight of indicators and aggregate the individual indicators to construct the CSIs (including overall CSI, economic CSI, environmental CSI, and social CSI). After getting the CSIs, the author evaluates the "uncertainty and sensitivity" of these CSIs, using "uncertainty analysis" tool and "sensitivity analysis" tool in the "Spatial Infrastructure Sustainability Assessment" plugin. In the end, the author visualizes the sustainability assessment results in the radar chart, sunburst chart, stratified bar chart, and line chart, using the visualization charts in the "Spatial Infrastructure Sustainability Assessment" plugin.

Based on the statistical analysis and spatial distribution analysis of the CSIs of the new road network and the CSIs of the current road network, the author reaches the following conclusions. First, after moving interstates underground, both the overall CSI and dimensional CSIs in the study area are improved, especially for the census tracts that are close to interstates (i.e., located within the 2-km buffer of interstates). While each tract has different improvement, among all the tracts within 2-km buffer of interstates, three tracts have the highest improvement of overall CSI, their overall CSI values are increased by more than 0.15. These results suggest that we can move those sections of interstates underground first, and we can conduct more detail sustainability analysis in these highly changed tracts, using lower scale (higher resolution) data (e.g., census blocks or census block groups).

Second, there is small uncertainty in the CSIs, and overall CSI and economic CSI has much smaller uncertainty than environmental CSI and social CSI. Third, environmental indicators contribute more to the improvement of overall CSI than environmental and social indicators. Fourth, the spatial distribution of each individual indicators and CSIs are not evenly among census tracts in the study area, which indicates the inequity among census tracts, and after moving interstates underground, the equity conditions among different census tracts are improved.

The results indicate that moving interstates underground can contribute to the sustainability of the study area and improve the sustainability performance of the census tracts that are located within the 2km-buffer of the interstates. However, further studies are needed to evaluate the sustainability of moving interstates underground and its contributions to the sustainability of the study area.

4.6 Future Work of Pilot Study

Even though we conducted a systematic sustainability analysis in the pilot study, there are limitations of the study, and several additional tasks need to be done in the future. First, we need to improve the measurement of some indicators used in this study. For example, combining DI results with population density data to re-calculate the change of DI after moving interstates underground. The estimation of energy cost needs to use more detailed vehicle classification data and vehicle distribution data, as well as use a better energy estimation model. The estimation of equity should consider more information, such as race, accessibility to public services, and vulnerability to pollutions or disasters, affordability of public transit service for lower-income residents, proportion of residents access to public transit service within 500 meters (Jeon and Amekudzi, 2005), mobility and transport for older and disabled persons (Shiau and Liu, 2013), residential population exposed to interstate highways, and so on. The assessment of land use should also include more land use information. Besides, we need to run some of the evaluation programs multiple times to eliminate the randomness of the results. Because during the process of evaluation of some indicators, we used random points, such as accessibility.

Second, sustainability is a complex and comprehensive measure affected by many factors. The indicators used for sustainability assessment in this study are fairly not sufficient. Thus the results may be biased. Therefore we need to introduce more indicators and in the process of sustainability assessment. For example, the ecological footprint can reflect living standards or quality of life. People in a spatially dispersed city would have an even larger footprint than people in the compacted city with the same style and standard of living (Hillier, 2009). Quality of Life can also be measured by average time of non-recreational travel, accessibility to green spaces and public services or infrastructures, noise pollution level of residential places, traffic congestion levels, and main mode for the journey to work. Bio-capacity Ratio is a good indicator when comparing the sustainability of different areas. Safety Indicator can be measured by the number of accidents, vulnerable user accident ratios (bike, motorcycle), or number of crimes within a 500 meters buffer of interstates. Resilience or vulnerability of the transportation system to disasters, bad weathers, disruptions or changes. We will need to perform additional experiments to analyze how people's accessibility to public services or commute time changes if we remove each section of interstates from the current transportation system by the order of their edge betweenness centrality. We will also explore other measures of resilience and vulnerability.

To measure these indicators, we need to collect more data. Such as traffic capacity of each road in the transportation network, locations of underground infrastructure systems or utilities, aboveground telecommunications networks, locations of traffic ingress and egress points, traffic volumes or demand (including both people commuting flow and transporting of goods), MARTA rail data, socio-economic information (e.g., population, income, race, education level, and so on), land use information, crime data, accidents data, and other related information.

Third, we need to make more reasonable assumptions of the scenario after moving interstates underground. For example, after moving interstates underground, the newly released land may be turned into commercial buildings or residential buildings, not just green space as we assumed in this study. And when reconnecting the current roads, we may reconnect more roads or redesign current roads (e.g., increase traffic capacity) after moving interstates underground. In the future, we need to work together with ARC or city planning agencies to make our assumptions more reasonable.

Fourth, moving all the interstates within the perimeter (interstates I-285) underground may be unrealistic in the next 100 years. According to the sustainability assessment results in this study, the sustainability impacts of each section of interstates are very different. We can first move some sections of interstates underground, for example, the sections within those three census tracts with the highest CSI improvement. But we need to conduct more detail sustainability analysis in these highly changed tracts, using lower scale data (e.g., census blocks or census block groups) in the future.

4.7 Summary

In this chapter, we evaluated the sustainability of moving interstates within the Perimeter in Atlanta underground using the proposed spatial sustainability assessment framework discussed in Chapter 3. The sustainability assessment results show that after moving interstates underground, both the overall sustainability and sustainability in economic, environmental, and social dimension in the study area are improved, especially for the census tracts that are close to interstates. But these tracts have different sustainability improvement, three of them have the highest improvement. In other words, moving interstates underground could make the study area more sustainable, especially moving the interstates sections that are located in the three tracts with the highest sustainability improvement. However, sustainability is a complex and comprehensive measure affected by many factors further studies are needed to evaluate the sustainability of moving interstates underground. Some recommendations for future works are also indicated in the chapter.

CHAPTER 5. CONCLUSIONS

The world's natural resources are presently under increasing pressure and humanity is facing multiple sustainable challenges, such as environmental degradation, pollution, water shortage, climate change, poverty, social inequalities, carbon-intensive lifestyles, and so on. Sustainability is the only way to resolve the economic, environmental and social challenges simultaneously. It has become a prominent concept in societal and political discourses around the world and serves as a major guideline for political actions and future societal development. However, to help decision makers (policy makers) determine which actions should or should not be taken in an attempt to make our society sustainable or to assess the efficiency of actions engaged, a sustainability assessment framework is needed to evaluate the sustainability of our proposed project or plans in short-term and long-term perspectives. Different methods have been proposed for sustainability assessment. However, only a few of them take into account all the aspects of sustainability when evaluating the sustainability of one project or one system. Further, they seldom consider the inter-linkages and dynamics changes happening in one system. Among all the sustainability assessment methods, MCA is considered the best suited for sustainability assessment. Because MCA can consider multiple criteria (both qualitative and quantitative criteria) simultaneously to help stakeholders make a well-informed decision. However, it requires stakeholders or decision makers to subjectively place importance (or weight) on each criterion, which may lead a not objective assessment results. Therefore, we use AHP or ANP to generate or derive more objective and consistent weight for each indicator.

Sustainability assessment frameworks provide a way to conceptualize sustainability at a high-level. It clarifies what to measure, what to expect from measurement and what kind of indicators to use when evaluating sustainability. There are many frameworks proposed by researchers, institutions or organizations. Among them, the thesis mainly reviewed DPSIR, DPSEEA, Impact-based frameworks, Stakeholder-based frameworks, GoldSET, and Spatial Frameworks. DPSIR and DPSEEA describe the relationships between the origins and consequences of environmental problems and can help identify sustainability assessment indicators. But the simple causal relations described by them cannot capture the complexity of interdependencies in the real world. Besides, DPSEEA could also lead to oversimplification of spatial and temporal interactions, which would result in poorly informed management decisions. TBL make it easier to use MCA to evaluate sustainability, but it tends to emphasize potentially competing for interests (three pillars) rather than the linkages and interdependencies between different aspects of sustainability. The stakeholder-based framework involves stakeholders into the process of sustainability assessment, which can lead to more effective and enduring solutions for sustainability and present opportunities to educate the public and influence collective behaviors (Waheed et al., 2009). GoldSET is an integrated framework which combines impact-based and process-based framework. It uses MCA to aggregate indicators and provides a good and easy way to present the sustainability assessment results. However, GoldSET does not consider the inter-linkages between different aspects or dimensions of sustainability and has limited spatial and temporal capacities. Spatial frameworks consider the spatial part of sustainability in the assessment. But most of the spatial frameworks are based on ArcGIS, which is commercial software. It is better to use open source GIS software, like QGIS or integrate the frequently used spatial analysis functions into the spatial framework directly. Even many of them use MCA and AHP in the assessment but does not consider the inter-linkages between different aspects of sustainability.

The spatial infrastructure sustainability assessment framework proposed in this thesis not only considers all aspects of sustainability (environmental, economic, social, and resilience), but also include inter-linkages between different aspects and their dynamic changes in the process of sustainability assessment. This framework provides a scientific and comprehensive method to spatially assess the sustainability of infrastructure at different spatial and temporal scales. The framework provides components to define sustainability assessment objective, to select indicators, to calculate the value of some indicators, to conduct multi-criteria decision analysis (pre-analysis, weighting, and aggregation of indicators), to evaluate the uncertainty and sensitivity analysis of the assessment results, and to visualize sustainability evaluation results in multiple ways. The framework engages policymakers and other participants in the process of sustainability evaluation, by letting them set the weight of indicators with the help of the AHP or PCA in the framework. Besides, there is a spatial database to manage the data, indicators and other information used during the process of sustainability assessment. The framework is fully data-driven, its assessment results are dependent on the input data. It is a very general framework, which can be easily used in other disciplines and other areas.

The framework is implemented as a plugin in QGIS. All the functions provided in QGIS can be used seamlessly in the framework. The framework provides various tools to help conduct sustainability assessment, including network analysis, spatial statistical analysis, comparison analysis, aggregation attributes to larger scale, Pre-analysis of indicators (normalization, correlation, PCA), AHP, aggregating indicators, uncertainty and sensitivity analysis of sustainability assessment results, data explorations, and multiple visualizations of sustainability assessment results. It also provides users a “wizard” to guide users conducting a sustainability assessment of their projects step by step.

To date, the author only implemented part of the functions in the proposed framework, with others to be added in the future. First, the current functions in the framework need to be optimzied by decreasing their processing time. Second, more function need to added to the framework to help evaluate sustainability indicators, such as measurement of resilience or vulnerability (e.g., vulnerability and resilience of infrastructure to natural disasters, vulnerabilities of communities of different social class, race or age.), measures of equity, energy consumption, waste and pollutions, Gini Coefficient, and so on. Third, life cycle assessment processes and DPSIR need to be integrated into the framework. Fourth, the framework need to address the multi-scale problems to make the sustainability assessment at different spatial scales consistent. Fifth, a spatial database of sustainability evaluation indicators need to be well designed and constructed, so that it better supports the spatial sustainability assessment framework.

As an application of the proposed spatial sustainability assessment framework, the author uses the framework to evaluate the sustainability performance of moving interstates within the Atlanta Perimeter (I-285) underground. First, twelve indicators are selected from economic, environmental, and social dimensions of sustainability for the sustainability assessment. After analyzing the correlations between indicators, ten relatively independent indicators are retained for MCA. Economic sustainability indicators include “total time spend in traffic” and “connectivity of transportation network”; Environmental sustainability indicators include “nature and biodiversity”, “urban heat island”, “noise pollution generated”, and “land use”; Social sustainability indicators include “accessibility to green space ”, “accessibility to health facilities ”, “accessibility to education facilities”, and “equity”. AHP is used to derive the weight of each indicator and “linear additive model” is used to aggregate indicators to build CSI for economic, environmental and social dimension, as well as the overall CSI. Finally, uncertainty and sensitivity analysis are conducted for the sustainability assessment. Based on the sustainability assessment results, after moving interstates underground, both the overall sustainability and the sustainability at each dimension in the study area are improved, especially for the census tracts that are close to interstates, and they are mainly located within the 2-km buffer of interstates.

However, further studies are needed to evaluate better the sustainability of moving interstates underground. First, the author plans to improve the measurement of some indicators used in this study. For example, the estimation of energy cost needs to use more detail vehicle classification data and vehicle distribution data, as well as use a better energy estimation model. The estimation of equity should consider more information, such as race, accessibility to public services, vulnerability to pollutions or disasters, affordability of public transit service by lower-income residents, the proportion of residents access to public transit service within 500 meters, the residential population exposed to interstate highways, and so on. Besides, the assumptions of the scenario after moving interstates underground need to be adjusted or updated. For example, after moving interstates underground, the newly released land may be turned into commercial buildings or residential buildings, not just green space as assumed in this study. And when reconnecting the current roads, more roads can be reconnected, and some current roads can be redesigned (e.g., increase traffic capacity) after moving interstates underground. In the future, the author plans to work together with ARC or city planning agencies to make these assumptions more reasonable.

Second, the author plans to add more indicators in the process of sustainability assessment. For example, “Quality of Life” can be measured by ecological footprint, the average time of non-recreational travel, accessibility to green spaces and public services or infrastructures, noise pollution level of residential places, traffic congestion levels, and main mode for the journey to work. Resilience or vulnerability of the transportation system to disasters, bad weathers, disruptions or changes. To measure these indicators, the author needs to collect more data. Such as traffic capacity of each road in the transportation network, locations of underground infrastructure systems or utilities, aboveground telecommunications networks, traffic volumes or demand (including both people commuting flow and transporting of goods), MARTA rail data, socio-economic information (e.g., population, income, race, education level, and so on), land use information, crime data, accidents data, and other related information.

CHAPTER 6. PLAN OF FUTURE RESEARCH

The proposed spatial sustainability assessment framework integrates DPSIR, TBL, LCA, and stakeholder-based methods. It can make full use of the advantages of each method and can present the sustainability assessment results from different aspects in multiple ways. This framework provides integrative and systematic methods to assess the sustainability of infrastructure at different spatial and temporal scales. It engages policymakers and other participants in the process of sustainability evaluation, by letting them set the weight of indicators with the help of AHP. However, due to the limit of time, not all the proposed methods are implemented in the "Spatial Sustain Assess" QGIS Plugin, and much work need to do in the future to expand its functionality and efficiency.

- First, improve and optimize current tools implemented in the framework by decreasing their processing time and making it more stable and reliable. The following works need to do in the future:
    - Use GPU and parallel computing to accelarate the computing of network analysis, spatial analysis, sustainability assessment results analysis, data exploration, and so on.
    - Improve the current spatial analysis functions or tools, such as the let the buffer analysis tool create multiple ring buffers. Improve the random points and random vector generation tools so that they can generate random vector or rando pints following any user-defined distributions. For example, generating the random points follow the same distribution of population

density in another layer, or generating random vectors following a normal distribution.

- Improve AHP, by considering the spatial difference of one indicator (i.e., geolocations). For example, a major bridge project that is located in an environmentally sensitive area would have greater weightings assigned to environmental indicators than other places (Ugwu & Haupt, 2007). Thus in the future, the spatial difference should be considered in the indicator weighting process.
- Make the uncertainty and sensitivity analysis tools more flexible, such as its input can be any sustainabiltiy assessment results instead of the results obtained from the AHP tool in the framework.
- Find a better way to organize and store historical assessment information in the spatial database, including the assessment object, scales, dimensions, indicators, data used, assessment methods, and results.
- Improve the sustainability evaluation indicators database. Even the present study built a database for the current commonly used indicators in literature, it is not enough. In the future, the author will continue to collect commonly used sustainability indicators, and identify a better way to store and organize them in the spatial database.
- Improve the dynamic links between the components in the wizard, so that if the user changes something in one component, the value in other components can be updated automatically. For example, if the user modifies the indicators in the environmental dimension, then all the following

components that depend on the indicators should be updated automatically, such as the AHP comparison matrix.

  - The framework need to address the multi-scale problems to make the sustainability assessment at different spatial scales consistent. This framework provides tools to aggregate sustainability assessment results in a smaller spatial scale to a bigger scale. Current aggregation methods implemented in the framework do not consider the complex interactions within a system (e.g., city, state), which can significantly affect the aggregation results. Therefore, in the future, a better aggregation method which considers the spatial interactions should be implemented in the framework.
  - The author will keep updating the "Spatial Sustain Assess" plugin so that it can also work in the latest version of QGIS.
- Second, implement the other functions or tools designed in chapter 3. The functions need to be implemented in the future, including:
  - Improve the indicator selection module. For example, integrate life cycle thinking, DPSIR, and PCA into the framework. Add new functions to recommend users indicators based on the object, scales, and sustainability dimension of their project.
  - Automatic data quality evaluation, including detecting and interpolate missing values (interpolate values based on similarity, median, mean, or distance between points, IDW, and so on).
  - Spatial analysis functions: add density analysis tools.

    - Data exploration analysis: add clustering analysis (k-means), basic statistical analysis, and boxplot analysis.
    - Network anlysis functions: add functions to calculate the measurement of resilience or vulnerability (e.g., vulnerability and resilience of infrastructure to natural disasters, vulnerabilities of communities of different social class, race or age.)
    - Other commonly used functions to evaluate sustainability indicators, such as measures of equity, energy consumption, waste and pollutions, Gini Coefficient, quality of life, and so on.
    - Add more indicator aggregation methods to the "Spatial Sustain Assess" plugin, such as the aggregation methods introduced by Díaz-Balteiro & Romero (2004).
    - Add functions or tools to interprate the sustainability assessment automatically.
    - Implement the tool to convert the conventiaonal geographical map into tile grid map (see section 3.7.1.5 for the algorithms).
    - Add tools to collect or derive data from the commonly used public data sources, including United Census Bureau, Open Street Map, ARC, ESRI Data & Maps, Bureau of Transportation Statistics, Yellow Pages, and some social media data (including Twitter and Facebook).
- Third, improve the design of the proposed spatial sustainability assessment framework with the development of knowledge and technology. For example, use D3 to visualize the sustainability assessment results, use some big data technique

to improve the sustainability assessment framework. Add big data analysis tools in the framework, so that users can take advantage of big data when evaluating the sustainability of their project. Nowadays, we live in a “Big Data” Era, and there are many data that we can use to study sustainable developments. For example, we can use the “India Night Lights” data to estimate the energy consumption in India, we can use big social media data to estimate the human nobilities between two places, to estimate people’s quality of life, or to predict the impacts of some natural disasters. We can use big data to optimize the current public transportation routes or schedules or to optimize the delivery of packages, electrics, gases, water, internet data, telecommunications, and so on.

REFERENCES

AbuSada, J., & Thawaba, S. (2011). Multi Criteria Analysis for Locating Sustainable Suburban Centers: A Case Study from Ramallah Governorate, Palestine. Cities, 28(5), 381-393.

Achour, N., Pantzartzis, E., Pascale, F., & Price, A. D. F. (2015). Integration of Resilience and Sustainability: from Theory to Application. International Journal of Disaster Resilience in the Built Environment, 6(3), 347-362.

Adler, N., Yazhemsky, E., & Tarverdyan, R. (2010). A Framework to Measure the Relative Socio-economic Performance of Developing Countries, Socio-Economic Planning Sciences, 44(2), 73-88.

Ali-Toudert, F., & Ji, L. (2017). Modeling and Measuring Urban Sustainability in Multi-Criteria Based Systems ---A Challenging Issue. Ecological Indicators, 73, 597-61.

Alshuwaikhat, H. M., & Aina, Y. A. (2006). GIS-based Urban Sustainability Assessment : The Case of Dammam city, Saudi Arabia. Local Environment, 11(October), 141-161.

Alternative Fuels Data Center. (2015). Maps and Data — Average Fuel Economy of Major Vehicle Categories. U.S. Department of Energy. https://www.afdc.energy.gov/data/10310. (Accessed Sep. 25th, 2017)

Anselin, L. (1995). Local Indicators of Spatial Association — LISA. *Geographical Analysis*, *27*(2), 93-115.

Asgari, N., Hassani, A., Jones, D., & Nguye, H. H. (2015). Sustainability Ranking of the UK Major Ports: Methodology and Case Study. Transportation Research Part E: Logistics and Transportation Review, 78(0), 19-39.

Ashley, R. & Hopkinson, P. (2002). Sewer Systems and Performance Indicators — Into the 21st Century. Urban Water. 4, 123-135.

Atlanta Regional Commission. (2017a). Green Space. http://opendata.atlantaregional.com/datasets/greenspace. (Accessed Nov. 14th, 2017)

Atlanta Regional Commission. (2017[b]). Lakes, Ponds, Reservoirs, and Swamps Georgia. http://opendata.atlantaregional.com/datasets/lakes-ponds-reservoirs-and-swamps-georgia?selectedAttribute=CLASS (Accessed Nov. 14th, 2017)

Atlanta Regional Commission. (2011). Household Travel Survey. https://www.afdc.energy.gov/data/. (Accessed June. 14th, 2015)

Avrachenkov, K., Litvak, N., Medyanikov, V., & Sokol, M. (2013). Alpha Current Flow Betweenness Centrality. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8305 LNCS(1), 106-117.

Bhat, C. R., Handy, S., Kockelman, K., Mahmassani, H., Srour, I., & Weston, L. (2001). Assessment of Accessibility Measures. Center for Transportation Research, Bureau of Engineering Research, the University of Texas, Austin, Research Report No. 4938-3.

Black, J. & Conroy, M. (1977). Accessibility Measures and Social Evaluation of Structure. Environment and Planning A, 9, 1013-1031.

Bishaw, A. & Glassman, B. (2016). Poverty: 2014 and 2015, American Community Survey Briefs. https://www.census.gov/library/publications/2016/acs/acsbr15-01.html. U.S. Department of Commerce Economics and Statistics Administration, U.S. Census Bureau. (Accessed Jul 12th, 2018).

Bohringer, C., & Jochem, P. E. P. (2007). Measuring the Immeasurable — A Survey of Sustainability Indices. Ecological Economics, 63(1), 1-8.

Borgatti, S. P. (2005). Centrality and Network Flow. Social Networks, *27*(1), 55-71.

Bottero, M., Mondini, G., & Valle, M. (2007). The Use of the Analytic Network Process for the Sustainability Assessment of an Urban Transformation Project. International Conference on Whole Life Urban Sustainability and Its Assessment.

Bourdic, L., & Salat, S. (2012). Building Energy Models and Assessment Systems at the District and City Scales: A Review. Building Research and Information, 40(4), 518-526.

Bourdic, L., Salat, S., & Nowacki, C. (2012). Assessing Cities: A New System of Cross-Scale Spatial Indicators. Building Research and Information, 40(5), 592-605.

Boots, B. N., & Getis A. (1998). Point Pattern Analysis. Newbury Park, CA: Sage Publications.

Brandes, U., & Fleischer, D. (2005). Centrality Measures Based on Current Flow. Lecture Notes in Computer Science, 533-544.

Brundtland, G. (1987). Report of the World Commision on Environement and Development: Our Common Future. Oxford Paperbacks, Report of, 400.

Buchholz, T., Rametsteiner, E., Volk, T. A., & Luzadis, V. A. (2009). Multi Criteria Analysis for Bioenergy Systems Assessments. Energy Policy, 37(2), 484-495.

Chan, K., Saltelli, A., & Tarantola, S. (1997). Sensitivity Analysis of Model Output: Variance-Based Methods Make the Difference. Proceedings of the 29th Conference on Winter Simulation, (ii), 261-268.

Chen, X., & Corson, M.S. (2014). Influence of Emission-Factor Uncertainty and Farm-Characteristic Variability in LCA Estimates of Environmental Impacts of French Dairy Farms. J. Clean. Prod. 81, 150-157.

Childers, D., Cadenasso, M., Grove, J., Marshall, V., McGrath, B., & Pickett, S. (2015). An Ecology for Cities: A Transformational Nexus of Design and Ecology to Advance Climate Change Resilience and Urban Sustainability. Sustainability, 7(4), 3774-3791.

Chu, K., Liu, W., She, Y., Hua, Z., Tan, M., Liu, X.,Gu, L., Jia, Y. (2018). Modified Principal Component Analysis for Identifying Key Environmental Indicators and Application to a Large-Scale Tidal Flat Reclamation. Water, 10(1), 69.

Ciuffo, B., Miola, A., Punzo, V., & Sala, S. (2012). Dealing with Uncertainty in Sustainability Assessment. JRC Technical Reports. European Union, Ispra, Italy.

Community Research Connections (2017). Visualizations: A Picture is Worth a Thousand Words. https://crcresearch.org/visualizations/visualizations-static-page (Accessed April 15th, 2017).

Comprehensive R Archive Network (CRAN) (2017). An Introduction to Corrplot Package. https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html. (Accessed Dec 5th, 2017).

Congalton, R. G. & Green, K. (2009). Assessing the Accuracy of Remotely Sensed Data. Principles and Practices (2nd Ed.). Boca Raton, FL: CRC Press.

Corbisier, C. (2003). Living With Noise. US Department of Transportation Federal Highway Administration. https://www.fhwa.dot.gov/publications/publicroads/03jul/06.cfm. (Accessed February 15th, 2017)

Costanza, R. & Patten, B. C. (1995). Defining and Predicting Sustainability. Ecological Economics 15(3), 193-196.

Daily, G. C., & Ehrlich, P. R. (1996). Socioeconomic Equity, Sustainability, and Earth´s Carrying Capacity. Ecological Applications.

Daily Sabah (2016). Eurasia Tunnel Opens Today to Ease Istanbul's Traffic Congestion. Daily Sabah Istanbul. http://www.dailysabah.com/istanbul/2016/12/20/eurasia-tunnel-opens-today-to-ease-istanbuls-traffic-congestion-1482188059 (Accessed January 10th, 2017)

Derissen, S., Quaas, M. F., & Baumgärtner, S. (2011). The Relationship between Resilience and Sustainability of Ecological-Economic Systems. Ecological Economics, 70(6), 1121-1128.

Díaz-Balteiro, L., & Romero, C. (2004). In Search of a Natural Systems Sustainability Index. Ecological Economics, 49(3), 401-405.

Dill, J. (2004). Measuring Network Connectivity for Bicycling and Walking. Washington DC: Transportation Research Board 2004 Annual Meeting (CD-ROM), 131-139.

Dong, X., & Guo J. et al. (2015). Sustainability Assessment of the Natural Gas Industry in China Using Principal Component Analysis. Sustainability 7(5), 6102-6118.

Doukas, H., Papadopoulou, A., Savvakis, N., Tsoutsos, T., & Psarras, J. (2012). Assessing Energy Sustainability of Rural Communities using Principal

Component Analysis. Renewable and Sustainable Energy Reviews, 16(4), 1949-1957.

Duncan, B., & John, H. (1996). Sustainable Urban Transportation Initiatives in Canada. APEC Forum on Urban Transportation. Seoul, Korea.

D. Love, & P. Lindquist. (1995). The Geographical Accessibility of Hospitals to the Aged: a Geographic Information Systems Analysis within Illinois. Health Services Research, 29(6):629651.

Ebert, U., & Welsch, H. (2004). Meaningful Environmental Indices: A Social Choice Approach. Journal of Environmental Economics and Management, 47(2), 270-283.

Egilmez, G., Gumus, S., & Kucukvar, M. (2015). Environmental Sustainability Benchmarking of the U.S. and Canada Metropoles: An Expert Judgment-based Multi-Criteria Decision Making Approach. Cities, 42, Part A (0), 31-41.

Envision TM. (2015). ENVISION Manual: Envision Rating System for Sustainable Infrastructure. http://research.gsd.harvard.edu/zofnass/menu/envision/ (Accessed April 18th, 2017)

Environmental Defense. 1999. Environmental Sustainability Kit. Pollution Prevention Alliance: New York, NY, USA.

Eppstein, D., Van Kreveld, M., Speckmann, B., & Staals, F. (2013). Improved Grid Map Layout by Point Set Matching. IEEE Pacific Visualization Symposium, 25-32.

ESRI National Park Service. (2017). USA Parks. ArcGIS Hub https://hub.arcgis.com/items/f092c20803a047cba81fbf1e30eff0b5. (Accessed May 24th, 2018).

Finkbeiner, M., Schau, E. M., Lehmann, A., & Traverso, M. (2010). Towards Life Cycle Sustainability Assessment. Sustainability, 2(10), 3309-3322.

Finnveden, G., Hauschild, M. Z., Ekvall, T., Guinée, J., Heijungs, R., Hellweg, S., Suh, S. (2009). Recent Developments in Life Cycle Assessment. Journal of Environmental Management, 91(1), 1-21.

Franceschet, M. (2014). Betweenness Centrality. https://www.sci.unich.it/~francesc/teaching/network/betweeness.html (Accessed June 14th, 2018).

Galloway, J. (2017). A New Campaign Topic: Sending Metro Atlanta's Traffic Underground. https://www.ajc.com/blog/politics/new-campaign-topic-sending-metro-atlanta-traffic-underground/6DXMGeK8IUltPenuDRgeqJ/ (Accessed August 6th, 2017)

Garikapati, V. M., You, D., Zhang, W., Pendyala, R. M., Guhathakurta, S., Brown, M. A., & Dilkina, B. (2017). Estimating Household Travel Energy Consumption in Conjunction with a Travel Demand Forecasting Model. Transportation Research Record: Journal of the Transportation Research Board, 2668, 1-10.

Gasparatos, A. (2007). The Argument Against a Reductionist Approach for Assessing Sustainability. International Conference on Whole Life Urban Sustainability and Its Assessment, 27-29.

Gasparini, P., Manfredi, G., & Asprone, D. (2014). Resilience and Sustainability in Relation to Natural Disasters A Challenge For Future Cities, Springer.

Golder Associates. (2011). GoldSET- Golder's Approach to Sustainable Project Planning and Design: Methodology and Underlying Principles.

Goodchild, M.F. 1987. Spatial Autocorrelation. CATMOG 47. Geo Books, Norwich, Hutchins C Sons, Norwich.

Goodchild, M. F. (2009). What Problem? Spatial Autocorrelation and Geographic Information Science. Geographical Analysis, 41(4), 411-417.

Graymore, M. L. M., Wallis, A. M., & Richards, A. J. (2009). An Index of Regional Sustainability: A GIS-based Multiple Criteria Analysis Decision Support System for Progressing Sustainability. Ecological Complexity, 6(4), 453-462.

Gulhana, G., Ceylanb, H., Özuysalc, M., & Ceylanb, H. (2013). Impact of utility-based accessibility measures on urban public transportation planning: A case study of Denizli, Turkey. Cities, 32, 102-112.

Hak, T., Janouskova, S., & Moldan, B. (2016). Sustainable Development Goals: A Need for Relevant Indicators. Ecological Indicators, 60, 565-573.

Hampton, S. (2014). 60 Years of Urban Change: Southeast. Institute for Quality Communities. http://iqc.ou.edu/2014/12/18/60yrssoutheast/ (Accessed January 16th, 2017).

Hansen, W. (1959) How Accessibility Shapes Land Use. Journal of the American Institute of Planners, 25, 73-76.

Heller, M. C., Keoleian, G. A., & Arbor, A. (2000). Life Cycle Based Sustainability Indicators for Assessment of the U. S. Food System. University of Michigan: Ann Arbor.

Hellström, D., Jeppsson, U., & Kärrman, E. (2000). A Framework for Systems Analysis of Sustainable Urban Water Management. Environmental Impact Assessment Review, 20(3), 311-321.

Helton, J. C., & Davis, F. J. (2003). Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems. Reliability Engineering and System Safety, 81(1), 23-69.

Hillier, B. (2009). Spatial Sustainability in Cities Organic Patterns and Sustainable Forms. In Symposium A Quarterly Journal in Modern Foreign Literatures, 1-20.

Holling, C. S. (1973). Resilience and Stability of Ecological Systems. Annual Review of Ecology and Systematics, 4, 1-23.

Hossain M., & Alam S. et al. (2013). Australian Airport Network Robustness Analysis: A Complex Network Approach. Australian Transport Research Forum 2013 Proceedings, Oct. 2013, Brisbane, Australia.

Hubert, L., & Arabie, P. 1991. The Assessment of Spatial Autocorrelation through Constrained Multiple Regression. Geographical Analysis, 23(2), 95-111.

Hugé, J., Waas, T., Eggermont, G., & Verbruggen, A. (2011). Impact Assessment for a Sustainable Energy Future —Reflections and Practical Experiences. Energy Policy 39(10), 6243-6253.

Iman, R. L., & Conover, W. J. (1982). A Distribution Free Approach to Inducing Rank Correlation Among Input Variables. Communications in Statistics, Part B — Simulation and Computation, B11, 311-334.

International Standard Organization (ISO). (2006). Environmental Management — Life Cycle Assessment - Principles and Framework. ISO 14040:2006, 3(1).

Intergovernmental Panel on Climate Change (IPCC). Guidelines for National Greenhouse Gas Inventories Chapter 3 of Volume 1; Intergovernmental Panel ON Climate Change (IPCC): Hayama, Japan, 2006.

Jabareen, Y. R. (2006). Sustainable Urban Forms Their Typologies, Models, and Concepts. Journal of Planning Education and Research, 26(1), 38-52.

Jabareen, Y. (2011). Teaching Sustainability: A Multidisciplinary Approach. Creative Education, 2(4), 388-392.

Jabareen, Y. (2014). An Assessment Framework for Cities Coping with Climate Change: The Case of New York City and its plaNYC 2030. Sustainability (Switzerland), 6(9), 5898-5919.

James, P., Magee, L., Scerri, A., & Steger, M. (2015). Urban Sustainability in Theory and Practice Circles of Sustainability. New York: Routledge.

Jeon, C.M. & Amekudzi S. (2005). Addressing Sustainability in Transportation System: Definitions, Indicators, and Metrics. J. Infrastruct. Syst., 11, 31-50.

Jeon, C. M., Amekudzi, A. A., & Guensler, R. L. (2013). Sustainability Assessment at the Transportation Planning Level: Performance Measures and Indexes. Transport Policy, 25, 10-21.

Jollands, N., Lermit, J., & Patterson, M. (2004). Aggregate Eco-Efficiency Indices for New Zealand — A Principal Components Analysis. Journal of Environmental Management, 73(4), 293-305.

Jordan, D., & Thomas H. 1997. Intelligent Transportation Systems and Sustainable Communities Findings of a National Study. Paper presented at the Transportation Research Board 76th annual meeting, Washington, DC, January 12-16.

Kadoić, N., Ređep, N. B., & Divjak, B. D. (2013). Decision Making with the Analytic Network Process, (Section 2).

Kane, K., Hipp, J. R., & Kim, J. H. (2017). Analyzing Accessibility Using Parcel Data: Is There Still an Access-Space Trade-Off in Long Beach, California?. Professional Geographer, 69(3), 486-503.

Kettle, D. (2006). An Urban 3-Waters Infrastructure Sustainability Decision-Making Process. (October), 200.

Kenworthy, J. R. (2006). The Eco-city: Ten Key Transport and Planning Dimensions for Sustainable City Development. Environment and Urbanization, 18(1), 67-85.

Khan, F.I.; Sadiq, R.; Husain, T. (2002). GreenPro-I: A Risk-based Life Cycle Assessment and Decision Making Methodology for Process Plant Design. Environ. Model. Software.17, 669-691.

Khan, F.I.; Sadiq, R.; Veitch, B. (2004). Life Cycle Index (LInX): a New Indexing Procedure for Process, and Product Design and Decision-making. J. Cleaner Prod. 12, 59-76.

Khatri, K. B. (2013). Risk and Uncertainty Analysis for Sustainable Urban Water Systems. CRC Press/Balkema, Leiden, The Netherlands, 29-133.

Krajnc, D., & Glavič, P. (2005). A Model for Integrated Assessment of Sustainable Development. Resources, Conservation and Recycling, 43(2), 189-208.

Krishnamurthy K. Sreelekha, M. G. and M. V. L. R. Anjaneyulu. (2016). Interaction between Road Network Connectivity and Spatial Pattern. Procedia Technology, 24, 131-139.

Kristensen, P. (2004). The DPSIR Framework. 27-29 September 2004 Workshop on a Comprehensive / Detailed Assessment of the Vulnerability of Water Resources to Environmental Change in Africa using River Basin Approach. UNEP Headquarters, Nairobi, Kenya.

Kong, F, Yin, H. , & Nakagoshi, N. (2007) Using GIS and Landscape Metrics in the Hedonic Price Modelling of the Amenity Value of Urban Green Space: A Case Study in Jinan city, China. Landscape and Urban Planning, 79, 240-252.

Kotek, P., Jordán, F., Kabele, K., & Hensen, J. L. M. (2007). Technique of Uncertainty and Sensitivity Analysis for Sustainable Building Energy Systems Performance Calculations. 10th International IBPSA Conference, 629-636.

Lai, D. (2003). Principal Component Analysis on Human Development Indicators of China. Social Indicators Research, 61(3), 319-330.

Lautso, K., Spiekermann, K., & Wegener, M. (2002). Modelling Policies for Urban Sustainability. 42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany, European Regional Science Association (ERSA), Louvain-la-Neuve.

Leamer, E. E. (1983). Let's Take the Con out of Econometrics. American Economics Review, 73(1), 31-43.

Lee, Y.C., & Kim, K.H. (2015). Attitudes of Citizens Towards Urban Parks and Green Spaces for Urban Sustainability: The Case of Gyeongsan City, Republic of Korea. Sustainability, 7, 8240-8254.

Lee, M. H., Lee, J. S., Lee, J. Y, Kim, Y. H., Park, Y. S., & Lee, K. M. (2017). Uncertainty Analysis of a GHG Emission Model Output Using the Block Bootstrap and Monte Carlo Simulation. *Sustainability*, *9*(9), 1522.

Li, T., Zhang, H., Yuan, C., Liu, Z., & Fan, C. (2012). A PCA-based Method for Construction of Composite Sustainability Indicators. International Journal of Life Cycle Assessment, 17(5), 593-603.

Li, X., & Yeh, G.-O. (2002). Integration of Principal Components Analysis and Cellular Automata for Spatial Decision Making and Urban Simulation. Science in China Series D, 45(6), 521-529.

Liu, H. H., Yeh, Y.Y., & Huang, J. J. (2014). Correlated Analytic Hierarchy Process. Mathematical Problems in Engineering, 2014, 1-7.

Liu, M., & Frost, J. D. (2017). Sustainability of Moving Atlanta Interstate Highways within the Perimeter Underground: Preliminary Detour Analysis. International Conference on Sustainable Infrastructure 2017, New York City, New York, U.S..

Lindquist, P., & Love, D. (1995). The Geographical Accessibility of Hospitals to the Aged: A Geographic Information Systems Analysis within Illinois. Health Services Research, 29(6), 629-651.

Litman, T. (2003). Measuring Transportation: Traffic, Mobility and Accessibility. ITE Journal, 73, 28-32.

Litman, T. (2008). Evaluating Accessibility for Transport Planning — Measuring People's Ability to Reach Desired Goods and Activities. Online TDM Encyclopedia, Victoria Transport Policy Institute, 1-61.

Litman, T. (2009). Sustainable Transportation Indicators — a Recommended Research Program for Developing Sustainable Transportation Indicators and Data. Transportation Research Board 88th Annual Meeting, Washington DC, United States, January, 2009.

Lloyd, S. M., & Ries, R. (2007). Characterizing, Propagating, and Analyzing Uncertainty in Life Cycle Assessment: A Survey of Quantitative Approaches. Journal of Industrial Ecology, 11(1), 161-179.

Malczewski, J., (1999). GIS and Multicriteria Decision Analysis. John Wiley & Sons, New York.

Marchese, D., Reynolds, E., Bates, M. E., Morgan, H., Clark, S. S., & Linkov, I. (2018). Resilience and Sustainability: Similarities and Differences in Environmental Management Applications. Science of the Total Environment, 613-614, 1275-1283.

Maxim L., Spangenberg J. H., & Connor M. (2009). An Analysis of Risks for Biodiversity under the DPSIR Framework. Ecological Economics, 69(1), 12-23.

McKay, M. D. (1995). Evaluating Prediction Uncertainty. NUREG/CR-6311, LA-12915-MS, 1-65, Los Alamos National Laboratory, Los Alamos, NM.

Medel-González, F., García Ávila, L. F., Salomon, V. A. P., Marx Gómez, J., & Hernández, C. T. (2016). Sustainability Performance Measurement with Analytic Network Process and Balanced Scorecard: Cuban practical case. Production, 26(3), 527-539.

Mori, K., Fujii, T., Yamashita, T., Mimura, Y., Uchiyama, Y., & Hayashi, K. (2015). Visualization of a City Sustainability Index (CSI): Towards Transdisciplinary Approaches Involving Multiple Stakeholders. Sustainability (Switzerland), 7(9), 12402-12424.

National Association of City Transportation Officials. (2017). Urban Street Design Guide. https://nacto.org/publication/urban-street-design-guide/street-design-elements/lane-width/ (Accessed Nov. 14th, 2017).

Ness, B., Urbel-Piirsalu, E., Anderberg, S., & Olsson, L. (2007). Categorising Tools for Sustainability Assessment. Ecological Economics, 60(3), 498-508.

Ochsenbein, G., & Wachter, D. (2004). Sustainability Assessment: Conceptual Framework and Basic Methodology. Federal Office for Spatial Development (ARE), 1-62.

OECD. (2008). Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD Publishing, Paris.

Oh, K. & Jeong, S.(2007). Assessing the Spatial Distribution of Urban Parks using GIS. Landscape and Urban Planning. Landscape Urban Plan, 82. 25-32.

Onat, N., Kucukvar, M., & Tatari, O. (2014). Towards Life Cycle Sustainability Assessment of Alternative Passenger Vehicles. Sustainability, 6(12), 9305-9342.

Open Street Map (2016). http://planet.openstreetmap.org/ (Accessed January 8th, 2017)

Pearce, A.R., & Vanegas, J.A. (2002). Defining Sustainability for Built Environment Systems. Int. J. Environ. Technol. Manage. 2, 94-113.

Perrings, C. (2006). Resilience and Sustainable Development. Environment and Development Economics, 11(4), 417-427.

Pfister, S., & Scherer, L. (2015). Uncertainty Analysis of the Environmental Sustainability of Biofuels. Energy, Sustainability and Society, 5(1), 1-12.

Pope, J., Annandale, D., & Morrison-Saunders, A. (2004). Conceptualizing Sustainability Assessment. Environmental Impact Assessment Review 24(6), 595-616.

Prato, T. (2015). Conceptual Framework for Assessing the Sustainability of Forest Fuel Reduction Treatments and Their Adaptation to Climate Change. Sustainability, 7(4), 3571-3591.

PRé Sustainability. Sustainability Software for Fact-based Decision Making. https://www.pre-sustainability.com/simapro/ (Accessed April 10th, 2017)

Puppala, A.J. and Bheemasetti, T. V. (2018). Sustainability & Resilience in Transportation Infrastructure Geotechnics: Integrating Advanced Technologies for Better Asset Management. Geostrata: Addressing Infrastructure Needs. May/June, 2018. Published by Geo-Institute of ASCE.

QGIS 2.8 (2018). Vector Spatial Analysis (Buffers). https://docs.qgis.org/2.8/en/docs/gentle_gis_introduction/vector_spatial_analysis_buffers.html (Accessed June. 8th, 2018).

Raghav (2018). 4 Important Measures of Transport Networks (With Diagram). http://www.geographynotes.com/articles/4-important-measures-of-transport-networks-with-diagram/165 (Accessed June. 12th, 2018).

Rahm, E., & Do, H. (2000). Data Cleaning: Problems and Current Approaches. IEEE Data Eng. Bull., *23*(4), 3-13.

Richardson, B. C. (2005). Sustainable Transport: Analysis Frameworks. Journal of Transport Geography, 13(1), 29-39.

Riedman, D. (2016). Questioning the Criticality of Critical Infrastructure: A Case Study Analysis. Homeland Security Affairs, 12(May).

Robinson, J. (2004). Squaring the Circle? Some Thoughts on the Idea of Sustainable Development. Ecological Economics, 48(4), 369-384.

Robinson, H., & Bamford, C. G. (1978): Geography of Transport, MacDonald and Evans Publisher. London, 80 pages.

Rodrigue, J. P., Comtois, C., & Slack, B. (2017). The Geography of Transport Systems. Fourth Edition. New York: Routledge, 440 pages. ISBN 978-1138669574.

RVshare. (2017). The Most Fuel Efficient RV Rentals: Class A, B, and C. https://rvshare.com/rv/fuel-efficient-rv-rentals (Accessed Dec. 8th, 2017).

Saaty, R. W. (1987). The Analytic Hierarchy Process — What It is and How It is Used. Mathematical Modelling, 9(3-5), 161-176.

Saaty, T. L. (2004[a]). Fundamentals of the Analytic Network Process — Multiple Networks with Benefits, Costs, Opportunities and Risks. Journal of Systems Science and Systems Engineering, 13(3), 348-379.

Saaty, T. L. (2004[b]). Fundamentals of the Analytic Network Process — Dependence and Feedback in Decision-Making with a Single Network, 13(2), 129-157.

Saltelli, A. (2002). Making Best Use of Model Valuations to Compute Sensitivity Indices. Computer Physics Communications, 145, 280-297.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based Sensitivity Analysis of Model Output: Design and Estimator for the Total Sensitivity Index. Computer Physics Communications, 181(2), 259-270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). Global Sensitivity Analysis — The Primer. Chichester, UK: John Wiley & Sons, Ltd.

Saltelli, A., & Tarantola, S. (2002). On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal. Journal of the American Statistical Association, 97(459), 702-709.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. Halsted Press New York. New York, US: John Wiley & Sons, Ltd.

Sárközy, F. (1999). GIS Functions --- Interpolation. Periodica Polytechnica Civil Engineering, 43(1), 63-86.

Sawada, M. (2018). Global Spatial Autocorrelation Indices - Moran's I, Geary's C and the General Cross-Product Statistic.

http://www.lpc.uottawa.ca/publications/moransi/moran.htm. Accessed on Jul 27, 2018.

Scientific Applications International Corporation (SAIC). (2006). Life Cycle Assessment: Principles and Practice.

Shaw, T. (2016). Good Data Visualization Practice: Tile Grid Maps. https://forumone.com/ideas/good-data-visualization-practice-tile-grid-maps-0. (Accessed July 12th, 2018).

Shiau T.A., & Liu J.S. (2013). Developing an Indicator System for Local Governments to Evaluate Transport Sustainability Strategies. Ecological Indicators, 34, 361-371.

Shortall, R., Davidsdottir, B., & Axelsson, G. (2015). Development of a Sustainability Assessment Framework for Geothermal Energy Projects. Energy for Sustainable Development, 27(0), 28-45.

SimaPro Software. (2018). About SimaPro. https://simapro.com/about/. Accessed June 5th, 2018.

Singh, R. K., Murty, H. R., Gupta, S. K., & Dikshit, A. K. (2007). Development of Composite Sustainability Performance Index for Steel Industry. Ecological Indicators, 7(3), 565-588.

Singh, R. K., Murty, H. R., Gupta, S. K., & Dikshit, A. K. (2009). An Overview of Sustainability Assessment Methodologies. Ecological Indicators 9(2), 189-212.

Slaper, T., & Hall, T. (2011). The Triple Bottom Line : What Is It and How Does It Work? Indiana Business Review, 4-8.

Sokal, R.R. & Oden, N.L. 1978. Spatial Autocorrelation in Biology: 1 Methodology. Biological Journal of the Linnean Society, 10, 199-228.

Soler-Rovira, J., & Soler-Rovira, P. (2008). Assessment of Aggregated Indicators of Sustainability using PCA : the Case of Apple Trade in Spain. Proceedings of the 6th International Conference on Life Cycle Assessment in the Agri-Food Sector (LCA Food 2014), 12-14 November 2008, Zurich, Switzerland, 133-143.

Sreelekha, M. G., Krishnamurthy, K., & Anjaneyulu, M. V. L. R. (2016). Interaction between Road Network Connectivity and Spatial Pattern. Procedia Technology, 24, 131-39.

Stanners, D.; Bosch, P., Dom, A., Gabrielsen, P., Gee, D., Martin, J., & Weber, J. L. (2007). Frameworks for Environmental Assessment and Indicators at the EEA. Sustainability Indicators, 127-144.

Statistics How To (2018). Correlation Matrix: Definition. http://www.statisticshowto.com/correlation-matrix/ (Accessed June 15th, 2018)

Statistical Tools for High-throughput Data Analysis (STHDA). (2017a). Correlation Matrix: A Quick Start Guide to Analyze, Format and Visualize a Correlation Matrix using R Software. http://www.sthda.com/english/wiki/correlation-matrix-a-quick-start-guide-to-analyze-format-and-visualize-a-correlation-matrix-using-r-software#compute-correlation-matrix. (Accessed Dec 5th, 2017)

Statistical Tools for High-throughput Data Analysis (STHDA). (2017b). Visualize Correlation Matrix using Correlogram. http://www.sthda.com/english/wiki/visualize-correlation-matrix-using-correlogram. (Accessed Dec 5th, 2017).

Stein, M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. Technometrics, 29(2), 143-151.

Stiles, M. (2016). New Poverty Data Show Improving Economic Conditions in States. http://thedailyviz.com/2016/09/29/new-poverty-data-show-improving-economic-conditions-in-states/#more-2831. (Accessed July 12th, 2018).

Stumpf, K., Baumgärtner, S., Becker, C., & Sievers-Glotzbach, S. (2015). The Justice Dimension of Sustainability: A Systematic and General Conceptual Framework. Sustainability, 7(6), 7438-7472.

Sustainable Measures (2018). Sustainability Indicators 101 — Key Term; Equity. http://www.sustainablemeasures.com/node/34 (Accessed May 10th, 2018).

Taaffe, E. J.， & Gauthier, H. L. (1973), Geography of Transportation First Edition. Prentice Hall.

Technion, Y. J. (2017). A New Evaluation Method Applying Sustainability and Climate Change Concepts : The Case of Planning New York City 2030.

Tobler, W. R. (1970). A Computer Movie Simulation Urban Growth in Detroit Region. Economic Geography, 46, 234-240.

Tsai, P. J., Lin, M. L., Chu, C. M., & Perng, C. H. (2009). Spatial Autocorrelation Analysis of Health Care Hotspots in Taiwan in 2006. BMC Public Health, 9, 1-13.

Turner, B. L., Kasperson, R. E., Matson, P. A., McCarthy, J. J., Corell, R. W., Christensen, L., Eckley, N., Kasperson, J. X., Luers, A., Martello, M. L., Polsky, C., Pulsipher, A., & Schiller, A. (2003). A Framework for Vulnerability Analysis in Sustainability Science. Proc National Academy of Sciences US, 100(14), 8074-8079.

Ugwu, O. O., & Haupt, T. C. (2007). Key Performance Indicators and Assessment Methods for Infrastructure Sustainability — A South African Construction Industry Perspective. Building and Environment, 42(2), 665-680.

United States Census Bureau (2014). Tiger Products. https://www.census.gov/geo/maps-data/data/tiger.html (Accessed January 15th, 2017)

United States Census Bureau (2017). LEHD Origin-Destination Employment Statistics. https://lehd.ces.census.gov/data/ (Accessed January 20th, 2017)

Unnithan, S. K. R., Kannan, B., & Jathavedan, M. (2014). Betweenness Centrality in Some Classes of Graphs. International Journal of Combinatorics, 2014, 1-12.

US Department of Transportation Federal Highway Administration. (2014).Design Speed. https://safety.fhwa.dot.gov/geometric/pubs/mitigationstrategies/chapter3/3_designspeed.cfm (Accessed Oct. 18th, 2017)

Ulhoi, J. P., & Madsen, H. (1999). Sustainable Development and Sustainable Growth : Conceptual Plain or Points on a Conceptual Plain ?. The 17th International Conference of the System Dynamics Society & 5th Australian New Zealand Systems Conference, Wellington, New Zealand. The System Dynamics Society.

USDA National Agroforestry Center. (2017). Conservation Buffers: 6.4 Buffers for Noise Control. https://nac.unl.edu/buffers/guidelines/6_aesthetics/4.html. (Accessed March 6th, 2017)

Velazquez, L.; Munguia, N.; Platt, A.; Taddei, J. (2006). Sustainable University: What Can Be The Matter? Journal of Cleaner Production, 14 (9-11), 810-819.

Waas, T., Hugé, J., Verbruggen, A., & Wright, T. (2011). Sustainable Development: A Bird's Eye View. Sustainability, 3(10), 1637-1661.

Waas, T., Hugé, J., Block, T., Wright, T., Benitez-Capistros, F., & Verbruggen, A. (2014). Sustainability Assessment and Indicators: Tools in a Decision-Making Strategy for Sustainable Development. Sustainability, 6(9), 5512-5534.

Waheed, B., Khan, F. & Veitch B. (2009). Linkage-Based Frameworks for Sustainability Assessment: Making a Case for Driving Force-Pressure-State-Exposure-Effect-Action (DPSEEA) Frameworks. Sustainability, 1(3), 441-463.

Wang, W. M., Lee, A. H. I., & Chang, D. T. (2010). An Integrated FDM-ANP Evaluation Model for Sustainable Development of Housing Community. Optimization Letters, 4(2), 239-257.

Wei, J., Zhao, Y., Xu, H., & Yu, H. (2007). A framework for selecting indicators to assess the sustainable development of the natural heritage site. Journal of Mountain Science, 4(4), 321-330.

Wikipedia (2014). Transportation in Atlanta. Wikipedia. https://en.wikipedia.org/wiki/Transportation_in_Atlanta#Private_automobiles (Accessed January 10th, 2017)

Wikipedia (2018). Spearman's rank correlation coefficient. https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient

Williams, A. S. (2009). Life Cycle Analysis: A Step by Step Approach. ISTC Reports, Illinois Sustainable Technology Center.

Xu, Z., & Coors, V. (2012). Combining System Dynamics Model, GIS and 3D Visualization in Sustainability Assessment of Urban Residential Development. Building and Environment, 47(1), 272-287.

Yigitcanlar, T., & Dur, F. (2010). Developing a Sustainability Assessment Model: The Sustainable Infrastructure, Land-Use, Environment and Transport Model. Sustainability, 2(1), 321-340.

Zegras, C. (2006). Sustainable Transport Indicators and Assessment Methodologies. Biannual Conference and Exhibit of the Clean Air Initiative for Latin American Cities: Sustainable Transport — Linkages to Mitigate Climate Change and Improve Air Quality. São Paulo, Brazil, July 25-27, 1-17.

Zijp, M., Heijungs, R., Van der Voet, E., Van de Meent, D., Huijbregts, M., Hollander, A., & Posthuma, L. (2015). An Identification Key for Selecting Methods for Sustainability Assessments. Sustainability, 7(3), 2490-2512.