



**UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA**

**Administrative unit: University of Padova**

Department: **Land, Environment, Agriculture and Forestry (LEAF)**

---

PhD Program: **Land, Environment, Resources and Health (LERH)**

Batch: XXXIV

**The complex mechanism of urban waterlogging and the mitigation effect provided by urban green infrastructure in metropolitan coastal cities**

**PhD Program Coordinator:** Prof. Marco Borga

**Supervisor:** Prof. Paolo Tarolli  
**Co-supervisor:** Prof. Zhifeng Wu

**External evaluators:** Prof. Giulia Sofia (University of Connecticut)  
Prof. Yunpeng Wang (Chinese Academy of Sciences)

**PhD candidate:** Qifei Zhang



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

**Sede Amministrativa: Università degli Studi di Padova**

**Dipartimento Territorio e Sistemi Agro-Forestali (TESAF)**

---

**CORSO DI DOTTORATO DI RICERCA: Land, Environment, Resources, Health (LERH)**

Ciclo: XXXIV

**Il complesso meccanismo del ristagno idrico urbano e l'effetto di  
mitigazione fornito dalle infrastrutture verdi urbane nelle città costiere  
metropolitane**

**Coordinatore:** Prof. Marco Borga

**Supervisore:** Prof. Paolo Tarolli  
**Co-supervisore:** Prof. Zhifeng Wu

**Valutatori esterni:** Prof. Giulia Sofia (University of Connecticut)  
Prof. Yunpeng Wang (Chinese Academy of Sciences)

**PhD candidate:** Qifei Zhang

## Acknowledgments

After these three years of research work, there was sweat, tears, and sweetness, but more harvests allowed me to feel my growth from the bottom of my heart. Challenges, busyness, and opportunities have enriched my three-year Ph.D. life.

First of all, I would like to express my deep appreciation and gratitude to my supervisor Prof. Paolo Tarolli and co-supervisor Prof. Zhifeng Wu, for their guidance, time, patience, and encouragement to make my research more reliable and precise. Without their insightful guidance and persistent help, these researches would not have materialized. I am deeply honored to have Prof. Paolo Tarolli and Prof. Zhifeng Wu as my life mentor. Their rigorous academic attitude and generous mind will benefit me throughout my life. My heartfelt thanks also go to the external reviewers, Prof. Giulia Sofia and Prof. Yunpeng Wang for their patience in reviewing and providing constructive comments and suggestions on my thesis.

I would like to extend enormous thanks to Prof. Davide Pettenella, Prof. Marco Borga, Prof. Alessio Cecchinato, Prof. Roberto Rossi for giving us courses on “Choose, design, write and present Ph.D. research projects and results”, “Basic statistics”, “Applied statistics with applications in R”, and “Introduction to GIS”. These courses allow me to have a relatively solid foundation for my research and improve my professional skills. Simultaneously, I would also like to express my gratitude to all professors and staff in the TESAF Department for their great assistance and guiding suggestions during my three-year Ph.D. study. I am particularly grateful to our research group members, Wenfang Cao, Sara Cucchiaro, Anton Pijl, Luca Mauri, Eugenio Straffelini, and Wendi Wang, we spent most of these three years together, supporting each other and giving warm encouragement at all times. Special thanks go to my colleagues, Qianlin Ni, Erico Kutchartt, Fenya Chen, Xintian Jiang, Marco Sozzi, Ahmed Galal, Hongwei Liu, Giorgia Bottaro, Eleonora Cagliero, Lorenzo Martini, Susen Shrestha, and Handriyanti Diah P. We share a great time together in the Agripolis. I would also like to thank the assistance and technical support given by Antonella Tosatto and Antonio Cecchinato.

I would like to express my gratitude to the joint Ph.D. program of Guangzhou University and the University of Padova for giving me the opportunity to study at the University of Padova. This project greatly supports my study, research, and life in Italy. I am grateful that working in a kind and friendly office atmosphere (both in Padova, Italy, and Guangzhou, China) and everyone is ready to be helpful. I am also lucky to be a member

of the two research groups (Earth Surface Processes & Society research group and South China Urban Ecology and Human Settlement Research Center), where you can feel a sense of belonging and also motivation to be a better researcher. Thanks to the working environment provided by the LERH program, University of Padova, and Guangzhou University, I can think deeply in a comfortable and quiet environment.

Last but not least, my heartfelt thanks to my parents who always give me spiritual support and encouragement. This unconditional love gives me so much courage to move on without fear. Moral support and generous encouragement from my family and all my friends will never be forgotten.

# Table of Contents

<b>Acknowledgments</b> .....	I
<b>Summary</b> .....	V
<b>Sommario</b> .....	VII
<b>1. Introduction</b> .....	1
1.1 Background and justification .....	1
1.2 Research significance .....	13
1.3 Literature review .....	15
1.4 Research questions and objectives .....	36
1.5 General organization .....	40
<b>2. Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China<sup>1</sup></b> .....	44
2.1 Abstract .....	45
2.2 Introduction .....	46
2.3 Study area .....	49
2.4 Data and methods .....	50
2.5 Result .....	61
2.6 Discussion .....	71
2.7 Conclusion .....	75
Supplementary data .....	77
References .....	79
<b>3. A new approach to investigating the spatially heterogeneous driving forces of urban waterlogging<sup>2</sup></b> .....	86
3.1 Abstract .....	87
3.2 Introduction .....	88
3.3 Study area .....	91
3.4 Data and methodology .....	93
3.5 Results .....	104
3.6 Discussion .....	115
3.7 Conclusion .....	120
Supplementary data .....	123
References .....	126
<b>4. Explicit the urban waterlogging spatial variation and its driving factors: the stepwise cluster analysis model and hierarchical partitioning analysis approach<sup>3</sup></b> .....	132
4.1 Abstract .....	133

4.2 Introduction .....	133
4.3 Study area .....	138
4.4 Data and methodologies.....	139
4.5 Results.....	152
4.6 Discussion .....	164
4.7 Conclusion .....	167
Supplementary data .....	169
References .....	175
<b>5. Investigating the Role of Green Infrastructure on Urban Waterlogging: Evidence from Metropolitan Coastal Cities<sup>4</sup></b> .....	<b>183</b>
5.1 Abstract.....	184
5.2 Introduction .....	185
5.3 Materials and Methods.....	188
5.4 Results.....	199
5.5 Discussion .....	207
5.6 Conclusion .....	212
References .....	215
<b>6. Discussion</b> .....	<b>221</b>
<b>7. Conclusions</b> .....	<b>226</b>
<b>8. References</b> .....	<b>230</b>

# Summary

This thesis aimed to reach a thorough understanding of the complex mechanism of urban waterlogging and the mitigation effect provided by urban green infrastructure in metropolitan coastal cities. Through four articles, this thesis respectively reveals the scale effects of urban waterlogging influencing factors, identifies the spatial heterogeneous driving forces of waterlogging, assesses urban waterlogging susceptibility under different future scenarios, and clarifies the mitigation effect of urban green infrastructure on urban waterlogging.

It is well known that urban waterlogging is affected by environmental conditions and human activities. Few, however, have comprehensively revealed the relative contributions of the influencing factors to urban waterlogging concerning different scales of analysis. This thesis first offers the identification of dominant factors of urban waterlogging at different scales of analysis and the interpretation of scale effect. The result provides additional insights that the relative contributions of the influencing factors vary with the scales, underlining a strong scale effect. Since the dominant factors vary across different scales of analysis, it is unrealistic to determine a universal “optimal” analysis scale. Hence, the appropriate analysis scale should be chosen according to the specific influencing factors and the characteristics of study areas.

In highly urbanized areas, the distribution of urban landscape elements and their attributes are characterized by spatial heterogeneous. However, less effort focus on the spatial heterogeneity driving forces at the local scale, which hinders the implementation of target-specific urban waterlogging mitigation strategies. To shed some light on this topic, this research proposes an innovative method that integrated the cubist regression tree and geographical detector model to spatially clarify the local driving forces and dominant factors with different local conditions. The results denote that the waterlogging’s dominant driving factors and their contribution vary with the local site conditions, which can help to promote more targeted and effective mitigation strategies, rather than a “one-size-fits-all” policy.

Moreover, due to the spatial heterogeneity in urban areas and the non-stationary complex mechanism of urban waterlogging, a novel approach is proposed that combined the stepwise cluster analysis model (SCAM) and hierarchical partitioning analysis (HPA) to characterize the waterlogging variation and assess waterlogging susceptibility. The result indicates that the SCAM-HPA provides an effective and feasible solution for waterlogging variation simulation. Under different urbanization and rainfall change scenarios, the urban

waterlogging susceptibility has a considerable variation. The watershed spatial location and watershed characteristics are relevant aspects to be considered in identifying and assessing waterlogging susceptibility.

Urban green infrastructures (UGI) can effectively reduce surface runoff, thereby alleviating the pressure of urban waterlogging. Given the shortage of urban land resources, it is unrealistic to reduce urban waterlogging by considerably increasing the UGI area. Less attention has been paid to investigating the threshold level of waterlogging mitigation capacity. The results indicate that the mitigation capacity of green infrastructure on waterlogging presents a threshold phenomenon. The excessive area proportions of UGI within the watershed or an oversized UGI area may lead to a waste of its mitigation effect. Therefore, the area proportion of UGI and its mitigation effect should be considered comprehensively when planning UGI.

Given the growing concerns of global warming and rapid urbanization, this thesis helps to expand our understanding of the complex mechanism of waterlogging in metropolitan coastal cities and provides a theoretical and practical reference for waterlogging prevention and management.



# Sommario

Scopo di questa tesi è fornire una comprensione approfondita del complesso meccanismo del ristagno idrico e dell'effetto di mitigazione dovuto alle infrastrutture verdi nelle città costiere metropolitane. Attraverso quattro articoli, vengono mostrati rispettivamente gli effetti di scala dei fattori che influenzano il ristagno urbano, identificate le forze motrici eterogenee spaziali, valutate le aree suscettibili per diversi scenari di sviluppo e chiarito l'effetto della mitigazione dovuto alle infrastrutture verdi.

È ben noto che il ristagno idrico urbano è influenzato dalle condizioni ambientali e dalle attività umane. Pochi, tuttavia, hanno rivelato in modo esaustivo i contributi relativi dei fattori ambientali e antropici al ristagno urbano riguardanti diverse scale di analisi. I risultati indicano che tali fattori possono variare con il cambiamento della scala, evidenziando un significativo effetto di scala nelle analisi di questi processi. La scala più appropriata per gli studi di ristagno urbano può funzionare solo per specifici fattori, e quindi la migliore scala di analisi dovrebbe essere determinata dalle caratteristiche delle aree di studio.

Nelle aree fortemente urbanizzate, la distribuzione spaziale degli elementi del paesaggio urbano e i loro attributi sono caratterizzati da eterogeneità. Tuttavia, in letteratura poca attenzione è stata data sull'eterogeneità spaziale delle forzanti a scala locale. I risultati ottenuti in questa tesi dimostrano che i fattori dominanti del ristagno idrico e il loro contributo variano con le condizioni locali, informazione che dovrebbe facilitare l'attuazione di strategie di mitigazione più mirate ed efficaci.

A causa dell'elevata eterogeneità spaziale e del complesso meccanismo non stazionario del ristagno urbano, viene proposto un nuovo approccio che combina il modello di analisi cluster "SCAM" e l'analisi di partizione gerarchica (HPA) al fine di comprendere al meglio i processi di ristagno e analizzarne la suscettibilità. I risultati mostrano come il modello SCAM-HPA proposto possa fornire una soluzione efficace per la simulazione delle variazioni di ristagno urbano. È interessante notare che sotto diversi scenari di urbanizzazione e di precipitazioni, la suscettibilità al ristagno ha una notevole variazione. Le caratteristiche spaziali del bacino idrografico sono pertanto aspetti rilevanti da considerare nell'identificazione e nella valutazione della suscettibilità dei processi di ristagno.

Le infrastrutture verdi urbane (*Urban Green Infrastructures* - UGI) possono ridurre efficacemente il deflusso superficiale, alleviando i processi di ristagno idrico. A causa del consumo di suolo nelle aree metropolitane, è necessario capire come utilizzare la limitata area delle UGI per massimizzare la mitigazione del waterlogging.

Meno attenzione, tuttavia, è stata dedicata allo studio del livello di soglia della capacità di mitigazione del waterlogging. I risultati indicano che l'effetto dell'infrastruttura verde sul ristagno presenta un fenomeno di soglia. Un'eccessiva proporzione delle UGI all'interno di un bacino idrografico o un'area UGI sovradimensionata possono portare anche ad una riduzione del suo effetto di mitigazione. Pertanto, quando si pianifica la progettazione del verde urbano bisognerebbe considerare al meglio la sua distribuzione spaziale.

Alla luce delle problematiche legate al riscaldamento globale e la continua e rapida urbanizzazione, questa tesi contribuisce ad estendere la nostra comprensione del complesso meccanismo del ristagno idrico nelle città costiere fortemente urbanizzate e fornisce un riferimento teorico e pratico per la prevenzione e il controllo di questo processo e la progettazione di infrastrutture verdi urbane.

# CHAPTER 1

## 1. Introduction

### 1.1 Background and justification

Urbanization is the inevitable outcome along with the rapid socio-economic development of human society. The proportion of the world's urban population is expected to continue to increase to more than 60% by 2030 (United Nations, 2017). For the world's second-largest economy, China has observed rapid urbanization in the past decades (Fang and Wang, 2011; Wang et al., 2011). According to the UN World Urbanization Prospects (2018), China's urbanization rate has reached 56.74%, with a population of 782.20 million in 2016 (Figure 1-1). Furthermore, the "China Urban Development Report No. 12" (2019) also pointed out that China's urbanization rate will reach 70% by 2030 and around 80% by 2050. Urbanization, to some extent, increases the interaction between human society and ecosystem. However, such dramatic urbanization has inevitably altered the urban land cover features, urban hydrological process, land surface energy balance, urban microclimate, and air circulation, resulting in a series of social, environmental, and ecological problems, such as urban waterlogging, water quality deterioration, urban heat island effect, air pollutants, etc. These social, environmental, and ecological problems not only have a direct impact on the quality of urban habitat and public health, but also have a profound impact on the sustainable development of cities.

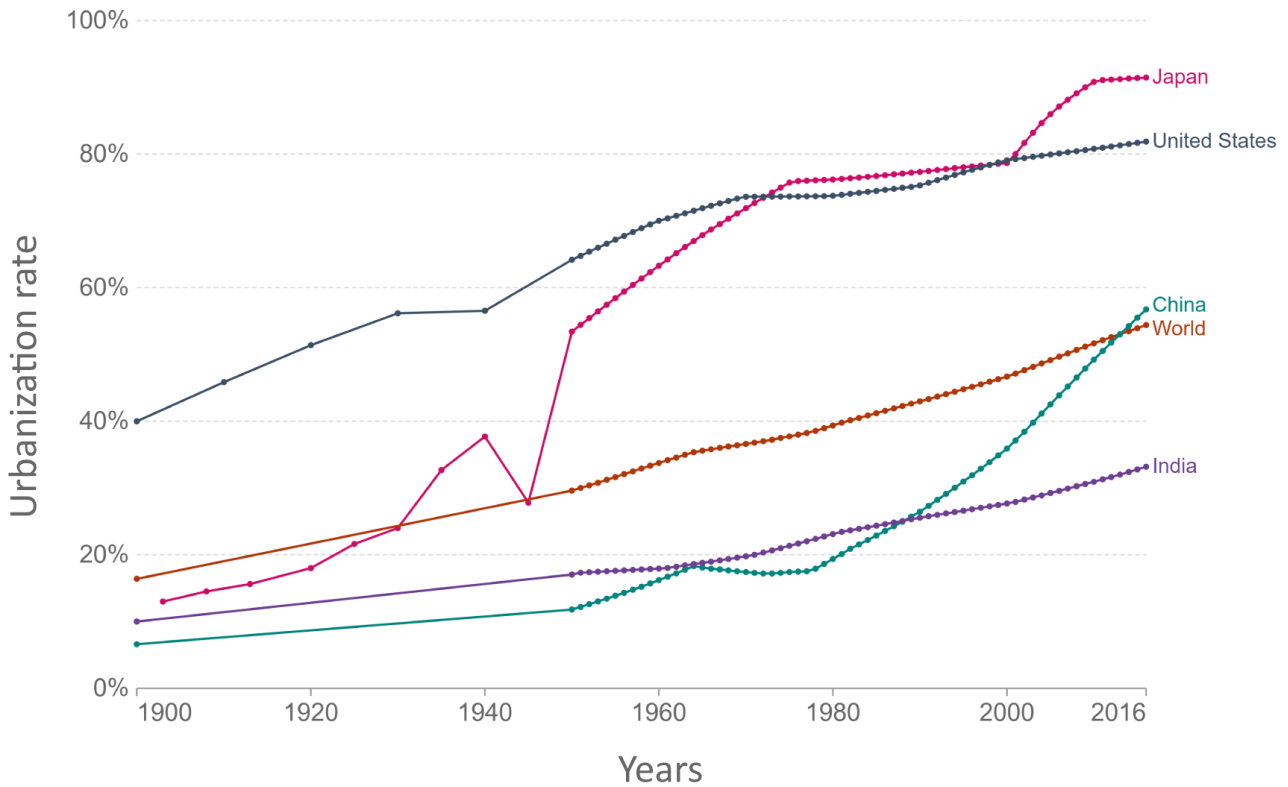


Figure 1-1 The global urbanization rate from 1900 to 2016. (Source: UN World Urbanization Prospects 2018; United Nations, Department of Economic and Social Affairs, Population Division 2018)

Urban waterlogging refers to the phenomenon of stagnant water disasters in low-lying areas where surface runoff exceeds the drainage capacity when encounters heavy or continuous precipitation (Zhang 2020; Chen et al., 2015; Yu et al., 2018). The frequent occurrence of urban waterlogging leads to serious social and environmental problems such as enormous economic losses, traffic paralysis, and water pollution. For example, when the depth of the stagnant water reaches 30 cm, it is difficult for pedestrians to walk; when the stagnant water exceeds 80 cm, the traffic is completely paralyzed. In addition, urban waterlogging also cause damage to houses and goods, affect people's daily lives and even threaten their safety. Last but not least, urban sanitation and human health are also affected. After the occurrence of urban waterlogging, a large amount of domestic waste and refuse is washed into the water, causing the surrounding water bodies to stink and pollution. At the same time, industrial waste that is not disposed of in time may also cause chemical pollution. Overall, urban waterlogging, one of the most common and serious "urban diseases", has become a huge challenge to the healthy and sustainable development of cities.

Although both urban waterlogging and floods have certain similarities, there are certain differences between them. Firstly, the causes are different. The main cause of urban waterlogging is the convergence of surface

runoff from local heavy rainstorms or long-duration rainfall in certain low-lying areas, resulting in water accumulation to a certain extent and affecting people's production and life. However, floods are generally caused by heavy rainfall or long-duration rainfall in the upper reaches of the river basin, resulting in large flows or high water levels in rivers or lakes, which pose a threat to urban or rural areas. In short, urban waterlogging is generally caused by local rainfall, while floods are generally caused by upstream rainfall. Secondly, the scope of impact is different. Urban waterlogging mainly affects local areas, with a relatively small scope of impact and short duration, whereas floods have a greater impact and long duration. Thirdly, the governance measures are different. The main measures to manage urban waterlogging are to reasonably reserve blue and green space for rainwater storage and stagnation and to build drainage networks, drainage pumping stations, cisterns, and other engineering facilities. Conversely, the measurements for flood control include the construction of river and lake embankments, upstream reservoirs, and flood storage and detention area.

In the context of global climate change, it has triggered the variation in spatial precipitation patterns. Simultaneously, the intensity and frequency of extreme precipitation events have increased significantly. Therefore, the regions affected by extreme precipitation increased significantly globally, creating background conditions for urban waterlogging (Tan et al., 2021; Berghuijs et al., 2017; Du et al., 2019; Blöschl et al., 2019). Numerous observed data indicate that the middle-high latitudes and tropical regions generally show an increasing trend of precipitation (Min et al., 2011). As shown in Figure 1-2, the frequency of extreme precipitation events showed an increasing trend in about 66.4% of global grid cells, and 56.7% of grid cells (Papalexiou et al., 2019). In detail, most parts of the Mediterranean, North America, and South Asia all experience an increase in the intensity of extreme precipitation events. On a regional scale, urbanization has further led to changes in the local atmospheric circulation. Urban microclimates such as urban rain islands or urban heat islands further increase the probability and intensity of precipitation in urban areas (Xu et al., 2019; Kim et al., 2017; Liu et al., 2020a; Miller and Hutchins, 2017). Therefore, changes in the global climate and urban microclimates together lead to an increase in the risk of urban waterlogging.

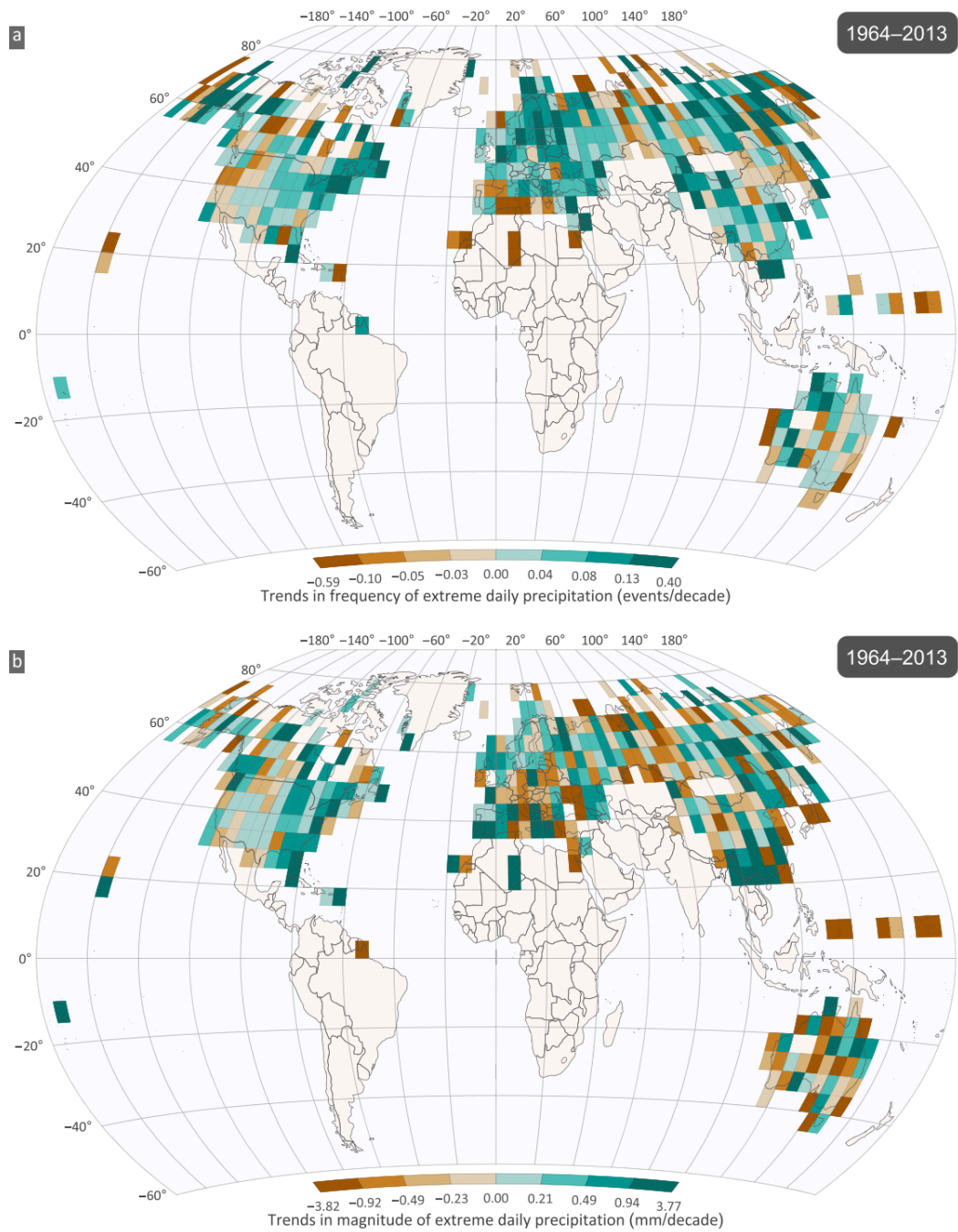


Figure 1-2 Mean trend values of the frequency (a) and magnitude (b) of extreme daily precipitation over the period 1964–2013. (Source: Papalexiou et al., 2019)

If the increase in extreme precipitation (global climate change) is the “environmental factor” leading to urban waterlogging, whereas a series of consequences arising from the process of urbanization (land cover change, modifying surface elevation) can be considered as the “anthropogenic factor” that causes urban waterlogging. Firstly, with the acceleration of urbanization, the area of impervious surfaces has increased sharply, leading to the decrease of vegetation coverage. Numerous studies have shown that vegetation coverage, as a

permeable surface, can effectively absorb rainwater and intercept surface runoff (Du et al., 2019; Kim and Park, 2016; Yao et al., 2015; Yang et al., 2015). Thus, the replacement of urban permeable surfaces has resulted in the increase of surface runoff volume and the confluence velocity of surface runoff. As shown in Figure 1-3, on the one hand, urbanization has reduced the amount of rainwater infiltration, leading to an increase in the total amount of surface runoff. On the other hand, surface interception has become poorer, which leads to an increase in flood peak discharge and an earlier flood peak. Secondly, the land reclamation of rivers, lakes, or reservoirs has greatly restricted the urban regulation and storage capacity. The continued degradation of the urban river network drainage function means that the majority of surface runoff can only be conveyed through the urban drainage system. This will undoubtedly increase the pressure on the underground drainage system. Thirdly, human activities and urbanization have greatly altered the urban topography. For example, tunnels, underground car parks, and overpasses have been built as a result of urban development. However, the surface elevation of tunnels, culverts, and underground parking is much lower than the surroundings. This means that these urban lowlands are easy to obtain runoff from the surrounding area, which may further exacerbate the risk of urban waterlogging.

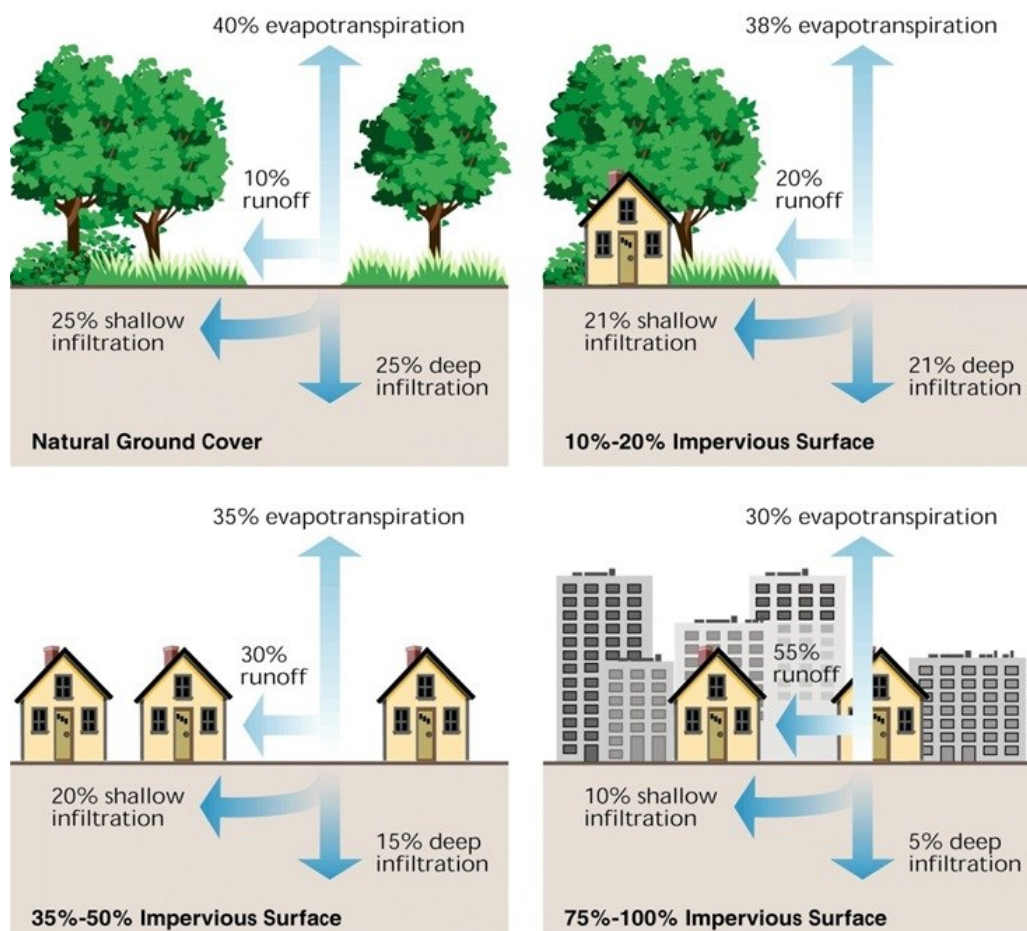


Figure 1-3 Relationship between impervious surfaces and surface runoff. (Source: Stream Corridor Restoration: Principles, Processes, and Practices, 2001.)

Under the dual effect of global climate change and urbanization, the frequency, intensity, and losses caused by urban flood disasters around the world are on the rise, which has attracted the attention of many international organizations, governments, and academia (Hirabayashi et al., 2013; Gallien et al., 2014; Li et al., 2016; Huang et al., 2018; Tang et al., 2018; Yu et al., 2018). The Global Emergency Database (EM-DAT) shows that urban flooding became the most widespread disaster from 2000 to 2019, and the number of disasters increased by 57% compared with 1980-1999 (Global Emergency Database, 2020). According to estimates by the World Bank, the average annual flood loss in 136 mega-large coastal cities in the world in 2005 was approximately US\$6 billion, and this value is expected to increase to US\$52 billion by 2050 (Hallegatte et al., 2013). As shown in Figures 1-4 and 1-5, the number of reported flood events has increased significantly and is mainly concentrated in southern Asia, North America, and South America. Consequently, the number of people affected by floods, financial, economic, and insurance losses are also increasing. For example, in 2010 alone, 178 million people were affected by floods and the total losses in some years such as 1998 and 2010 exceeded US\$40 billion (Jha et al., 2012). The frequent occurrence of global flooding confirms, to some extent, the increasing intensity of climate change and human disturbance. In view of this, it can be inferred that the risk of urban waterlogging may also gradually increase.

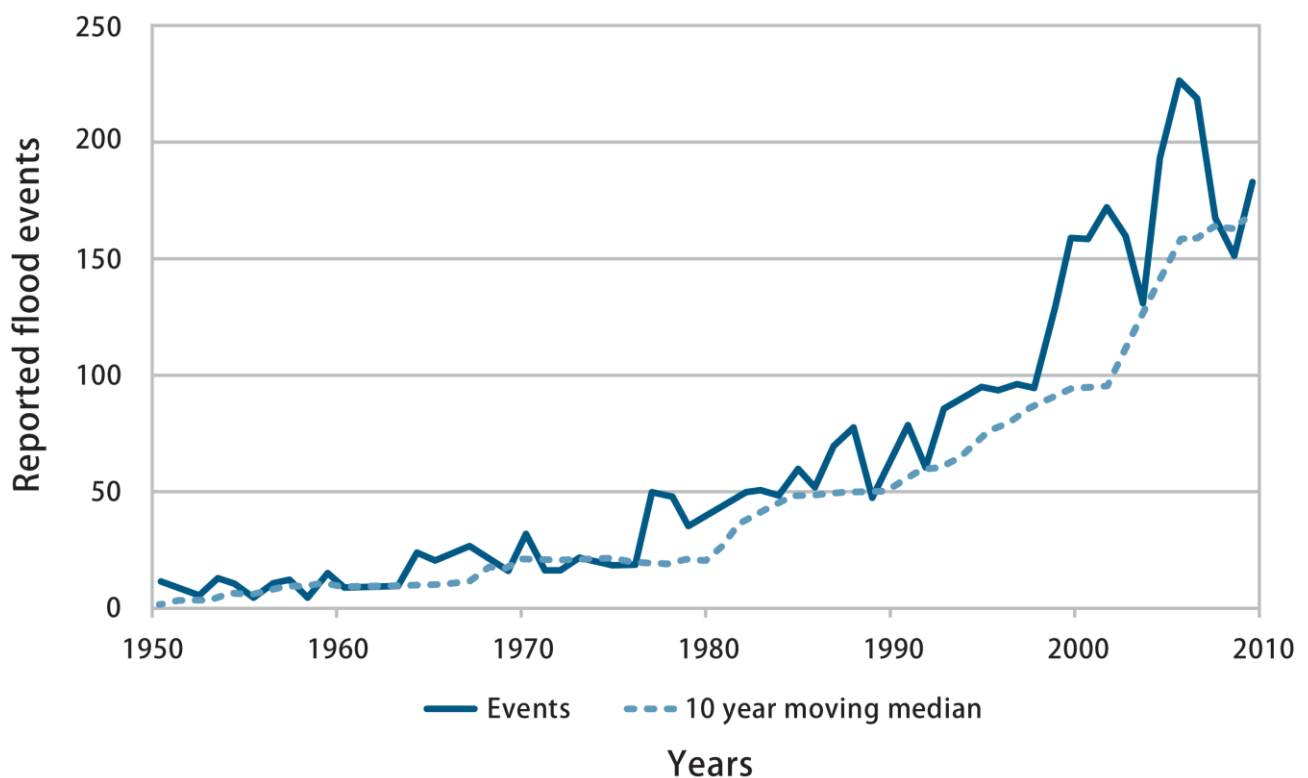




Figure 1-4 Number of reported flood events. (Source: based on EM-DAT/CRED)

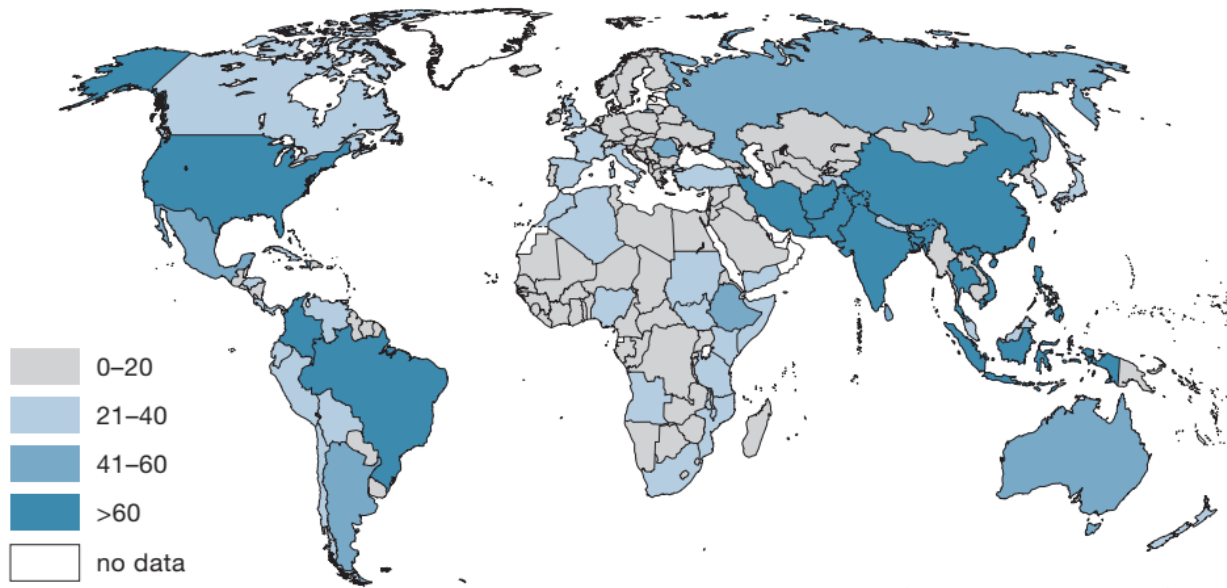


Figure 1-5 The spatial distribution of flood events from 1970 to 2011. (Source: Jha et al., 2012)

In addition, some scholars have used different models to predict future flood hazards under different climate change patterns (Zhang et al., 2017; Few, 2013; Gilroy and McCuen, 2012; Avand et al., 2021). Güneralp et al. (2015) pointed out that from 2000 to 2030, the flood risk of the Danube and Seine river basins in Europe has not increased significantly, but in South Asia and Southeast Asia, nearly three-quarters of the cities are expected to be at risk of high-frequency flooding. Specifically, the cities prone to high-frequency flooding are concentrated in the Chao Phraya River Basin, the Indus River Basin in Thailand, and the Yangtze River and Pearl River Delta basins in China (Figure 1-6). These results all indicate the trend of increasing risk of future flood hazards.

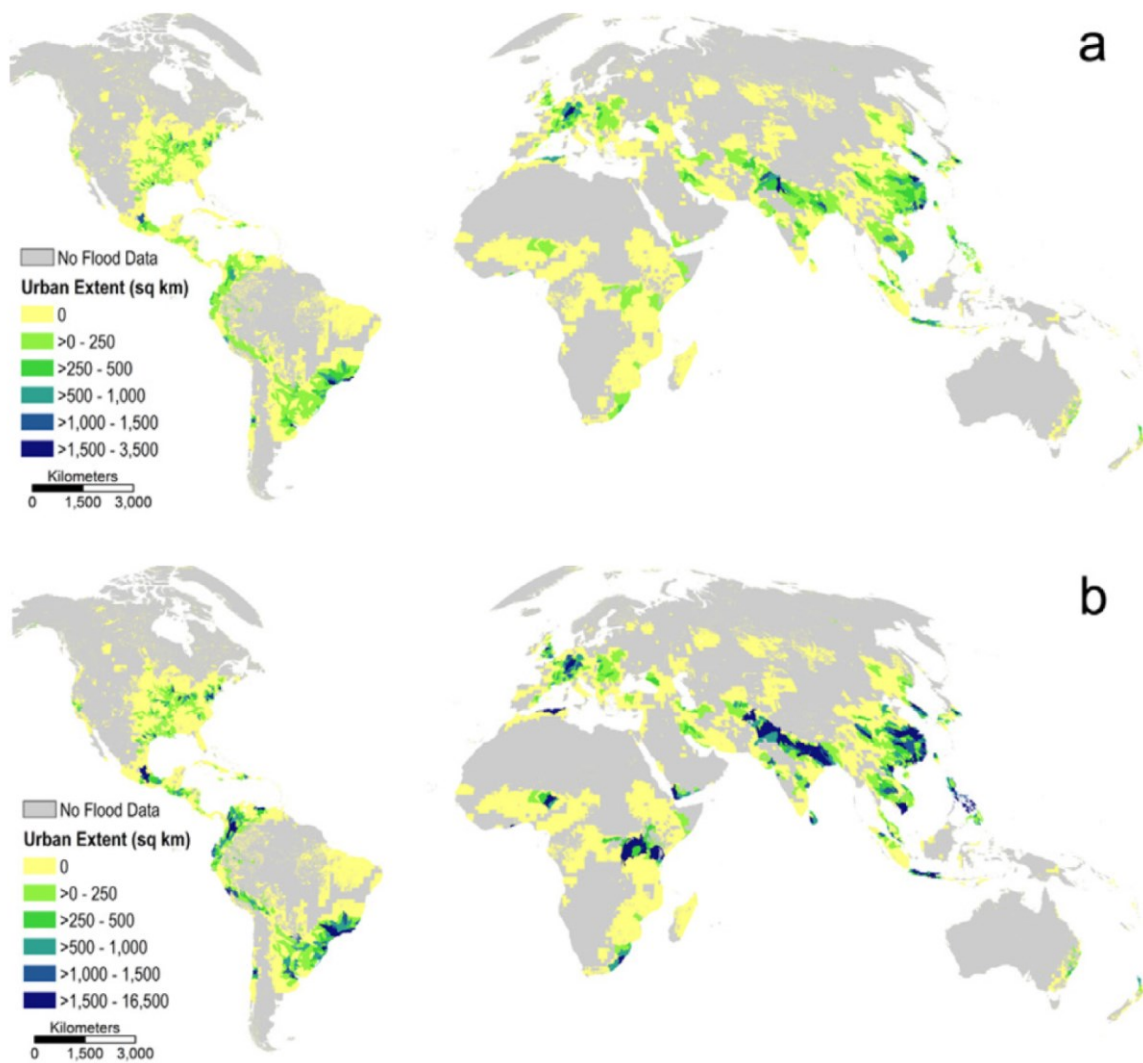


Figure 1-6 Urban land within the high-frequency flood zones in 2000 (a) and 2030 (b). (Source: Güneralp et al., 2015)

Although there are some differences between urban waterlogging and floods (intensity and scope of the disaster), urban waterlogging and floods can transform and influence each other. For example, if the river overflows the embankment (river poured into cities) or the river is silted (rainwater cannot be discharged), these conditions may also lead to urban waterlogging. Moreover, the factors that lead to urban waterlogging and floods have certain similarities (i.e. altered frequency and intensity of rainfall, human activities disrupt urban hydrological cycle). Against the backdrop of a gradual increase in global flood events, it can be predicted that the situation of urban waterlogging is not optimistic either. Therefore, given the increased impact of flooding around the world, we can infer that the number of cities facing the risk of urban waterlogging has also increased significantly.

### 1.1.1 The situation of urban waterlogging in China

At present, China is the largest developing country in the world and also one of the country's most seriously affected by waterlogging in the world. The situation of waterlogging disasters in Chinese cities can be summarized into the following three aspects:

(1) The frequency of urban waterlogging is on the rise. According to data from the Ministry of Housing and Urban-Rural Development, from 2007 to 2015, more than 360 cities across China were waterlogged. Among them, one-sixth of the cities had a single inundation time of more than 12 hours, with a depth of over 0.5 m. As shown in Figure 1-7, from 2010 to 2018, on average, more than 160 cities in China suffered from waterlogging disasters each year, and the direct economic losses were over \$15 billion per year (China Flood and Drought Disaster Bulletin). For example, on July 21, 2012, Beijing was hit by a heavy rainstorm, with an average precipitation of 170 mm and a maximum cumulative precipitation of 318 mm. This torrential rain caused 79 deaths, 63 waterlogging events within the city center, and over 1.9 million people affected. From the evening of March 30 to the morning of March 31, 2014, Shenzhen suffered a rainstorm, with an average rainfall of 125 mm and a maximum cumulative rainfall of 318 mm. This incident caused over 200 waterlogging events and killed 4 people. On May 7, 2017, a torrential rainstorm occurred in Guangzhou, with a maximum precipitation of 184.4 mm/h and the maximum 3-hour precipitation exceeded the historical extreme value of Guangdong Province (382.6 mm). This incident caused 109 houses to collapse, 118 waterlogging events within the city center, about 38.91 km<sup>2</sup> of farmland were flooded, and the economic loss was about \$89 million. Similarly, on May 22, 2020, a heavy rainstorm in Guangzhou (80 mm average hourly precipitation, 167.8 mm maximum hourly precipitation), which killed 4 people and flooded multiple underground parking lots and metro stations. Recently, on July 20, 2021, Zhengzhou city suffered from a severe urban waterlogging disaster (cumulative average precipitation was 449 mm, maximum one-hour precipitation reached 201.9 mm), resulting in 292 deaths and 47 missing persons.

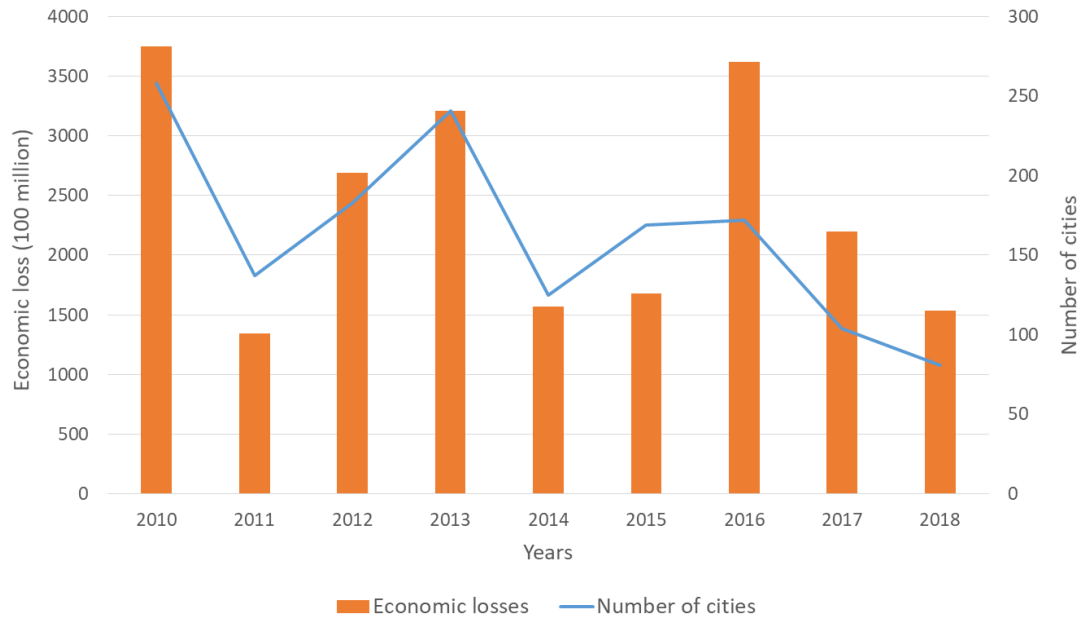


Figure 1-7 Number of cities and economic loss of urban waterlogging in China from 2010 to 2018. (Source: China Flood and Drought Disaster Bulletin, Ministry of Water Resources)

(2) Urban waterlogging is prone to occur. It is worth noting that although the precipitation is within the drainage system design standards (for example 50 mm/h), it has also led to urban waterlogging events in some areas. For example, the rainfall in Beijing on 23 June 2011, the average one-hour rainfall was 50 mm, which was far below the drainage capacity (design standard) of the drainage system. However, with such rainfall conditions, a serious urban waterlogging disaster occurred in Beijing. The city's traffic was paralyzed, causing major economic losses and even resulting in the loss of two lives.

(3) The recurrence rate of urban waterlogging is high. In general, the occurrence of urban waterlogging is mostly distributed in low-lying areas. The risk of urban waterlogging in such regions can be reduced by increasing the implementation of drainage facilities (drainage pipe network, pumping stations, underground storage tanks). However, the construction of these structures is extremely difficult and the workload is enormous. Moreover, the relatively low design standards and inadequate maintenance of drainage networks in developing countries further contribute to the high recurrence rate of waterlogging in such regions. In summary, urban waterlogging in China is becoming more and more serious, causing huge economic losses, and even threatening lives.

### 1.1.2 Urban waterlogging risk in coastal cities

Notably, we find that the frequency of urban waterlogging events, the damage range, the affected population,

and the direct economic losses are all increasing, especially in the densely populated and economically developed coastal regions (Sun et al., 2020; He et al., 2020). Given the unique geographical location and environmental characteristics of coastal regions, urban waterlogging has the characteristics of susceptibility, complexity, and great harm. Firstly, coastal cities are extremely vulnerable to storm surges and rising sea levels. In the context of climate change (notable polar warming), sea levels are gradually rising. Since the 20th century, global sea levels have risen by 10-20 cm (IPCC, 2014). The rise in sea level can inundate some low-lying coastal areas and increase the intensity of storm surges (Shen et al., 2019; Wang et al., 2018; Fang et al., 2021; Fang et al., 2017). This phenomenon highlights the risk of waterlogging in low-lying coastal cities. In addition, in most coastal cities, rainwater is usually collected through drainage systems and then pumped into the sea. However, at extreme high tide, the backflow situation is more serious, resulting in greatly reduced drainage capacity, and surface runoff is difficult to discharge (Wang et al., 2012; Yin et al., 2012). If heavy rainfall and extreme high tide occur simultaneously, the severity of the waterlogging will be greatly aggravated. Secondly, coastal cities have generally experienced rapid development, but it also prompted a series of problems such as unreasonable urban planning, low-capacity (low construction standards) drainage systems, land subsidence, and imperfect urban disaster management systems. Human disturbance, sea-level rise, and storm surges have formed an important disaster chain. Continued human activities have exacerbated sea level rise and storm surge intensity, which in turn magnified the damage of urban waterlogging. At the same time, unreasonable planning will further amplify the hazards of urban waterlogging. As a result, the combined effect of various influencing factors has led to the complexity of urban waterlogging in these regions. Lastly, as coastal cities are the centers of populations, economic activities, these urban waterlogging events may be devastating. The State Oceanic Administration of China estimated that the average annual direct economic loss caused by typhoons/storm surges was US\$1.6 billion from 2012 to 2016 (China Marine Disaster Bulletin, 2016).

As shown in Figure 1-8, the risk of urban waterlogging in the Guangdong-Hong Kong-Macao Greater Bay Area (111°59'42"-115°25'18"E, 21°17'36"-23°55'54"N) is more prominent. The eastern coastal cities, as the most concentrated and economically active regional, are usually highly developed areas with dense population concentration. These regions are a vital part of the national or global economy. Coupled with the impact of rising sea levels, the waterlogging status and its hazards in these coastal cities may be extremely severe (Gallien et al., 2014; Li et al., 2016). Therefore, it is of great practical importance to carry out research on how to effectively alleviate the occurrence of urban waterlogging in metropolitan coastal cities.

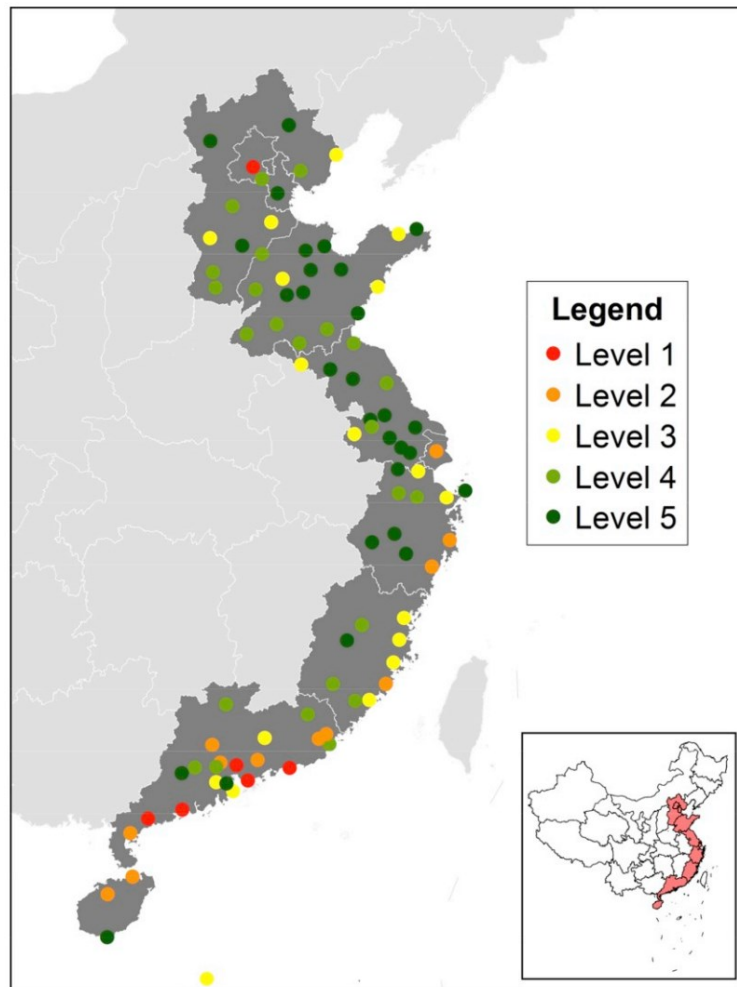


Figure 1-8 The risk of urban waterlogging in Eastern China prefecture-level cities. (Source: Sun et al., 2020)

Facing the increasing risk of urban waterlogging, a proliferation of studies has shown that simple drainage engineering measures are not sufficient to effectively prevent waterlogging disasters (Lin et al., 2021; Lin et al., 2018; Wu et al., 2020; Zhang et al., 2020; Quan, 2014a). Therefore, in the pre-disaster stage, it is necessary to further deepen the understanding of the spatial patterns of urban waterlogging, urban waterlogging driving forces, waterlogging susceptibility assessment and forecasting, and urban waterlogging mitigation strategies to reduce or even prevent related losses. In the context of global climate change, sea-level rise, and urbanization process, two major coastal cities in the Guangdong-Hong Kong-Macao Greater Bay Metropolitan Region, Guangzhou (112°57' to 114°30'E, 22°26' to 23°56'N) and Shenzhen cities (113°45' to 114°37'E, 22°26 to 22°51'N), are selected to test this proposition. This thesis first systematically reveals the spatial distribution pattern of waterlogging events, clarifies the complex mechanism of waterlogging in metropolitan coastal cities by analyzing the effects of natural factors and human activities on urban waterlogging. On this basis, an urban waterlogging simulation model is developed to assess urban


waterlogging susceptibility under different future scenarios (climate change and urbanization). For urban waterlogging mitigation strategies, the role of urban green infrastructure (UGI) on waterlogging and the threshold level of green infrastructure mitigation capacity are also investigated. This thesis is expected to promote urban sustainable development and provides a theoretical and practical reference for the government and urban planners for urban waterlogging prevention and management.


## 1.2 Research significance

### 1.2.1 Sustainable Development Goals

In 2015, all member states of the United Nations adopted the 2030 Agenda for Sustainable Development, which provided a common blueprint for the peace and prosperity of mankind and the planet. It has established 17 Sustainable Development Goals (SDGs) and 169 sub-goals, covering the society, economy, and environment, which point out the direction for the sustainable development and international cooperation of countries in the next 15 years (United Nations: Sustainable Development Goals). This thesis analyzes the complex mechanism of waterlogging and the mitigation effect provided by urban green infrastructure in metropolitan coastal cities. The proposed work connects the SDGs with the Chinese government's strategic development plan, which is in line with the UN 2030 Sustainable Development Goals reported in Table 1-1.

*Table 1-1 The Sustainable Development Goals achieved in this thesis.*

<p>Target 11 Sustainable cities and communities</p> 	<p>11.3 By 2030, enhance inclusive and sustainable urbanization and capacity for participatory, integrated, and sustainable human settlement planning and management in all countries.</p>
	<p>11.5 By 2030, significantly reduce the number of deaths and the number of people affected and substantially decrease the direct economic losses relative to global gross domestic product caused by disasters, including water-related disasters, with a focus on protecting the poor and people in vulnerable situations.</p>
	<p>11.6 By 2030, reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management.</p>

<p>Target 13 Climate action</p> 	<p>13.1 Strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries.</p> <hr/> <p>13.2 Integrate climate change measures into national policies, strategies, and planning.</p>
---	--

### 1.2.2 National level

Increasingly severe urban waterlogging has gradually attracted the attention of the Chinese government. On 25 April 2021, the General Office of the State Council of China issued the "Implementation Opinions on Strengthening Urban Waterlogging Management", proposing to promote urban waterlogging management, striving to significantly improve the waterlogging prevention capacity of cities by 2025 (The State Council of the People's Republic of China). Urban waterlogging is a new type of disaster with urban development. The investigation of the complex mechanism of waterlogging and its mitigation strategies contribute to the implementation of national policies and the realization of sustainable urban development in China.

### 1.2.3 City level

Given the extreme risk of urban waterlogging in coastal cities in China, this thesis selects Guangzhou and Shenzhen in the Guangdong–Hong Kong–Macao Greater Bay Metropolitan Region of Southern China as the study area. These two cities both belong to subtropical maritime monsoon climates, with an annual average rainfall of around 1720 mm and 1935 mm respectively. However, the temporal and spatial distribution of precipitation in the two cities is extremely unbalanced. Temporally, the precipitation is mainly concentrated in April to September, accounting for over 80% of the annual precipitation. Spatially, affected by the prevailing wind direction, the precipitation is decreasing from northeast to southwest (Guangzhou), and southeast to northwest (Shenzhen). Additionally, Guangzhou and Shenzhen are often affected by typhoons and tropical cyclones in summer. From the perspective of topographic conditions, Guangzhou and Shenzhen are both located in delta alluvial plains, with an average elevation of 9.9 m and 46 m, respectively, which is typical of coastal lowlands and extremely vulnerable to sea-level rise or seawater intrusion. The GDP of Guangzhou and Shenzhen in 2020 reached \$384 billion and \$415 billion respectively, which are national economic centers and international cities in China. If urban waterlogging occurs in such regions, it will



undoubtedly have a serious impact on people's production and life and social and economic development. Therefore, this thesis can provide suggestions and guidance for urban waterlogging management in metropolitan coastal cities.

#### *1.2.4 Academic research level*

Given the characteristics of highly urbanized coastal areas, there are various urban landscape elements with different attributes in cities. The distribution of urban landscape elements and their attributes are characterized by spatial heterogeneity. In other words, urban landscape elements (such as DEM, land cover characteristics) vary in different spatial locations. It hints that the driving factors or even the mechanism of urban waterlogging may vary with the change of spatial location, resulting in the high spatial heterogeneity and the non-stationary complex mechanism of urban waterlogging. Additionally, urban waterlogging in coastal areas is affected by various influencing factors such as human activities, sea-level rise, storm surges, rainfall patterns, etc. On the basis of this, the analysis framework of the complex mechanism of urban waterlogging and its mitigation strategies proposed in this thesis establishes a research paradigm of urban waterlogging management and prevention in coastal cities, opens up the research horizon of urban waterlogging, and enriches the theoretical and methodological for urban waterlogging susceptibility assessment.

### **1.3 Literature review**

Based on the research content of this thesis, a literature review is conducted on the current research status of several aspects of urban waterlogging driving factors, urban waterlogging simulation and susceptibility assessment methods, and urban waterlogging mitigation measures.

#### *1.3.1 Causes of urban waterlogging*

Urban waterlogging is a systemic problem, every link of the urban hydrological process may be a factor that affects the formation of urban waterlogging (Zhang et al., 2020; Zhang et al., 2021b). The urban waterlogging influencing factors can be divided into natural factors and anthropogenic factors. In which, natural factors include urban microclimate and geographical conditions (such as meteorological conditions, urban topography, hydrogeology, and soil); while the anthropogenic factors contain land cover features, urban

drainage system, relevant administrative regulations and policy, and other factors related to human activities. In this thesis, we summarize the causes of urban waterlogging into four aspects.

(1) **Meteorological conditions:** due to global climate change, extreme precipitation events occur frequently. The concentration and intensity of precipitation have led to excessive surface runoff in cities, exceeding the urban drainage capacity.

(2) **Topographic conditions:** the low-lying and relatively flat areas within cities are conducive to the accumulation of surface runoff.

(3) **Land cover features:** the rapidly increased impervious surface area blocks the infiltration path of rainwater and alters the original hydrological conditions, thereby disrupting the water balance.

(4) **Drainage facilities:** due to the low design standards and poor management of drainage systems in developing countries, surface runoff is difficult to discharge in a short period.

#### 1.3.1.1 Meteorological conditions

The effect of meteorological conditions on the risk of waterlogging has been extensively documented (Alfieri et al., 2017; Myhre et al., 2019). As the global climate changes, precipitation patterns in various regions are redistributed. The Fifth Assessment Report of the Intergovernmental Panel on Climate Change pointed out that in the context of climate change, the intensity and frequency of extreme precipitation events have increased significantly (IPCC AR5). It can be inferred that the increase and concentration of heavy rainfall events are one of the direct causes of urban waterlogging. Berghuijs et al. (2017) reviewed the variation of floods across multiple continents in recent years and found that climate change may lead to changes in flood magnitude and occurrence rates. Hirabayashi et al. (2013) predict the global flood risk at the end of this century through 11 climate models and used a global river path model with inundation schemes to predict inundation areas under different climate change scenarios. Zhang et al. (2017) assessed the impact of climate change on the probability of inundation in coastal cities based on eight climate models and four CO<sub>2</sub> emission scenarios. Bryndal et al. (2017) analyzed the impact of meteorological conditions on flood risk management, and their findings suggest that extreme precipitation can easily trigger local flooding. Alfieri et al. (2017) proposed a framework to assess the economic loss and population impact caused by river flooding on a global scale in response to global warming. Their results show a significant positive correlation between global warming and future flood risk on a global scale. Additionally, the "urban heat island " effect caused by

urbanization further leads to heat convection over the city. This phenomenon increases the probability and intensity of rainfall in the urban area, forming the "urban rain island" effect (Figure 1-9) (Shepherd, 2002; Han et al., 2014; Shepherd, 2005; Liu et al., 2020a; Mitra and Shepherd, 2015). Simultaneously, as urban air pollutants rise, pollutant particles promote water vapor condensation, which may further increase the frequency of rainfall in urban areas. The Metropolitan Meteorological Experiment (METROMEX) confirms that urbanization has increased the total urban precipitation by 10%-30% in summer, while increasing precipitation by 17% in the autumn and 4% in the spring (Changnon Jr et al., 1971; Changnon S, 2016). Zhao et al. (2021) took Jinan city, China as a case study to analyze the impact of urban expansion on the rain island effect. The results showed that during 1978-2017, the urban expansion in Jinan city significantly increased the frequency and intensity of short-duration precipitation events, and the magnitude and frequency of extreme precipitation events in built-up areas were significantly higher than those in suburbs areas. Therefore, in the context of global climate change and urban microclimate, the changes in the temporal and spatial distribution of precipitation will ultimately affect the entire urban hydrological cycle process, leading to an increased risk of urban waterlogging.

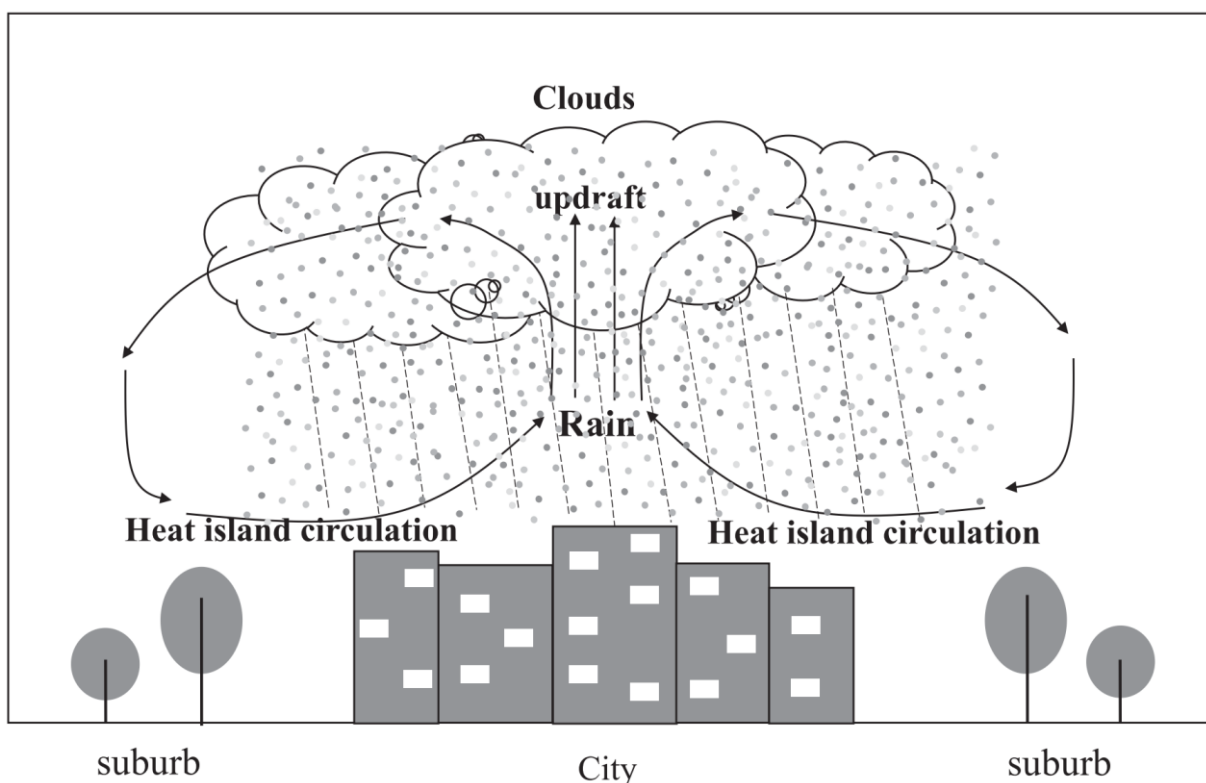


Figure 1-9 Schematic diagram of urban rain island effect. (Source: Yang et al., 2017)

### 1.3.1.2 Topographic conditions

High-intensity human activity has profoundly altered the topographical features of urban areas, forming complex micro-topography. In most cases, urban waterlogging events occur in the region with low elevation and relatively flat terrain, since these areas are easy to obtain runoff from the surrounding area (Figure 1-10). For example: (1) low-lying landforms such as tunnels, culverts, or underground parking lots often become the hot spots of urban waterlogging. (2) Newly built residential communities increase their land surface elevation, which increases the risk of waterlogging in neighboring communities (with lower land surface elevation). (3) The elevation difference of urban arterial roads leads to the formation of local lowlands that become potential locations for waterlogging. Therefore, these urban lowlands often suffer from urban waterlogging problems. The results of Tehrany et al. (2015) indicate that compared with relatively flat areas, high-altitude areas are less likely to experience waterlogging events. In the work of Liu et al. (2021), the impact of surface elevation and slope on urban waterlogging was fully confirmed by using the geographical detect model. Yin et al. (2016) used a hydraulic model (FloodMap-HydroInundation 2D) to assess the impact of land subsidence on the risk of urban flooding in Shanghai. The results show that land subsidence can lead to a non-linear response of flood characteristics. Similarly, the results of Wang et al. (2015) and Wu et al. (2020) both demonstrated that flat and low-lying land surfaces increase the risk of urban waterlogging. In addition, urban built-up areas with low elevations will also be affected by the backflow of rivers and tides, which further threaten water security in low-altitude areas.



*Figure 1-10 The waterlogging in urban lowlands (tunnels and culverts). (Source: China Daily)*

### 1.3.1.3 Land cover features

The rapid development of urbanization has led to drastic changes in the land cover features, which directly affects the surface runoff generation and confluence. Among various land cover types in cities, forests, grasslands, wetlands, and other urban green spaces that can effectively reduce surface runoff are replaced by impervious surfaces (Figure 1-11). These impervious surfaces decrease the infiltration of rainwater, which increases the urban runoff coefficient (Nigussie and Altunkaynak, 2016; Chu et al., 2010; Amaguchi, 2012; Bonneau et al., 2017; Burns et al., 2015). Li et al. (2018) used Boosted Regression Tree (BRT) to assess the relative impact and marginal effects of factors affecting direct runoff. The results indicate that direct runoff is significantly related to urbanization, and high-density residential, commercial and industrial areas have large surface runoff and high runoff coefficients. Hu et al. (2020) investigated the impact of land-use change on surface runoff in the downtown area of Beijing from 1984 to 2019, which indicated that the variation trend of surface runoff is strongly consistent with the proportion of impervious surface. Zeinali et al. (2019) explored the effects of residential areas and rainfall on runoff coefficients using one-way and two-way analysis of variance. On the other hand, urbanization leads to a decrease in surface roughness. Therefore, the speed of surface runoff confluence significantly increases (Zhang et al., 2018a). Nigussie and Altunkaynak

(2016) conducted an analysis on the hydrological effects of urbanization in the river basin, and they believed that the peak discharge is the largest under the scenario of unrestricted urbanization.

Furthermore, urban development also occupies a large amount of urban flood storage space. The continuous degradation of regulation and storage functions increases the risk of urban waterlogging. Lee and Brody (2018) pointed out that urban landscapes with a higher percentage of impervious surfaces may cause more severe flooding than other land-use types. The study of Zhang et al. (2020) confirmed that the impervious surfaces represented by residential areas are the dominant factors for waterlogging at different analysis scales. Yu et al. (2018) analyzed the mechanism of the spatial and temporal pattern of impervious surfaces on urban waterlogging in Guangzhou, which confirmed that with the expansion of cities, the increase of impervious surface leads to the intensification of urban waterlogging. Zhang et al. (2018) took Guangzhou city as a case study to investigate the relative importance of the spatial pattern of impervious surfaces to urban waterlogging and confirm that the percent coverage of building has the most significant impact on the magnitude of urban waterlogging. Sofia et al. (2017) analyzed flood dynamics in northeastern Italy from 1900 - 2010 and found that land-use change was an important factor in increasing the frequency of flood hazards. The research of Huang et al. (2017) shows that the land cover change is the key factor that causes urban waterlogging, which largely affects the degree of inundation at the regional scale.

In contrast to impervious surfaces, urban green infrastructure (UGI) plays a positive role in urban waterlogging. A proliferation of studies have shown that urban green infrastructure, as a permeable surface, can effectively absorb and store rainwater (Zhang et al., 2021a; Kim and Park, 2016; Yao et al., 2015; Yang et al., 2015), and the canopy and rhizome of vegetation can intercept surface runoff (Yang et al., 2013; Liu et al., 2014). As for UGI composition, Armson et al. (2013) found that in a sample plot of 9 m<sup>2</sup> (the land cover includes grassland, trees, and asphalt), the grassland controlled almost all the surface runoff, and the trees reduced 62% of runoff from asphalt. Richards et al. (2015) pointed out that a vegetated area of 7.5% to the catchment area would reduce surface runoff by more than 90%. For the spatial pattern of UGI, a study indicated that the less fragmented urban green spaces are more effective in reducing peak annual average river runoff (Kim and Park, 2016).

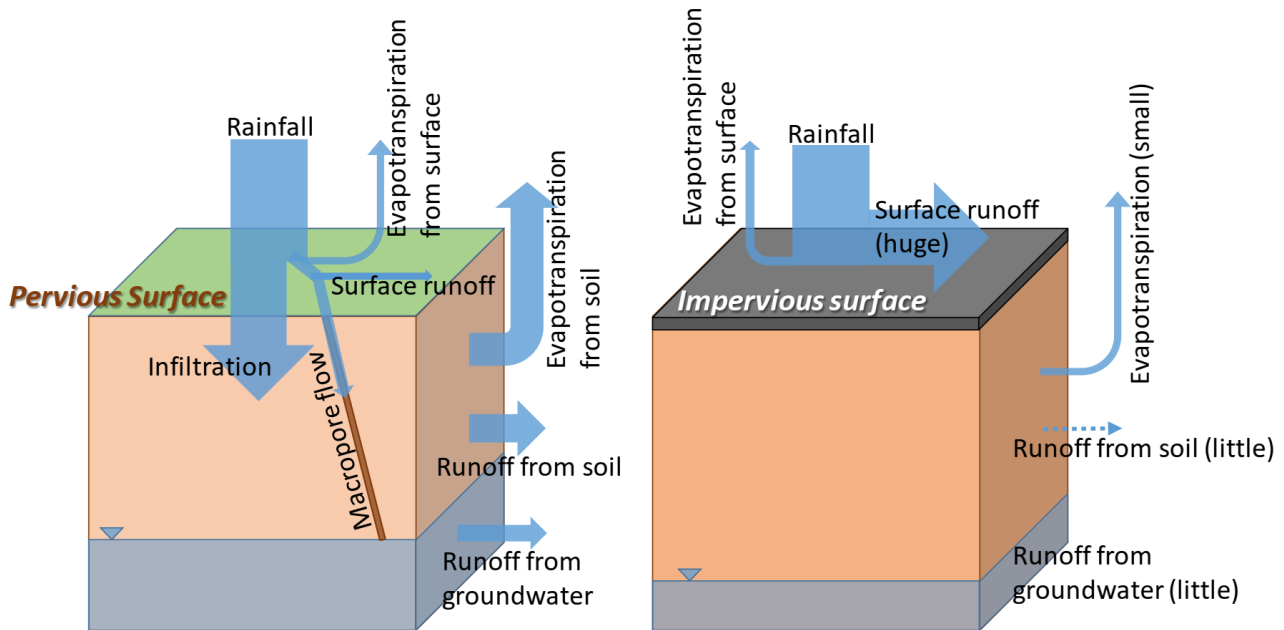


Figure 1-11 The hydrological characteristics of permeable surface and impervious surface. (Source: Tanouchi et al., 2019)

#### 1.3.1.4 Drainage facilities

In the process of rapid urbanization, rivers within cities have been artificially landfilled, which disrupts the urban drainage pattern. Moreover, the area of urban rivers network and lakes continues to shrink, weakening their capacity to regulate and store rainwater runoff. This has led to an increased risk of urban waterlogging. For example, Wang (2017) found that the river network was negatively correlated with Shanghai's rainstorm and waterlogging. Additionally, most of the rainwater drainage networks in developing countries are characterized by low design standards and inadequate maintenance, which makes it difficult to play an active role in the face of heavy rainstorms. In China, the 2019 "Outdoor Drainage System Design Standard" requires that the design standard for drainage facilities in general areas (non-central urban area) is 1-3 years for the recurrence period of precipitation and that in important areas (central urban area) is 3-5 years (Ministry of Housing and Urban-Rural Development, 2019). Compared with design standards of other countries/regions (Table 1-2), it indicates that the design standards of China's urban drainage pipelines are relatively low.

Table 1-2 Comparison of domestic and foreign design standards of drainage facilities. (Source: Zhang et al., 2013)

Country/region	Design standard (rainstorm return period)
China Mainland	1-3 years for general areas (non-central urban area), 3-5 years for important areas (central urban area), 10 years for particularly important areas (medical

---

	facilities, schools, tunnels, transportation hubs, metro system).
China Hong Kong SAR	10 years for rural drainage systems and 50 years for urban drainage systems.
United States	2-15 years for residential areas, 10-100 years for commercial and high-value areas.
European Union	5 years for rural areas, urban centers/industrial areas/commercial areas, 10 years for underground railways/underpasses.
Japan	3-10 years.

---

As a result, the lagging urban drainage system is no longer able to cope with the negative effects of rapid urban development, especially in developing countries where the design standard is relatively low. Thereby, it is easy to accumulate rainwater to form urban waterlogging events when heavy rainfall occurs. Wu and Zhang. (2017) found that the density of the drainage network is positively correlated with the density of urban waterlogging, suggesting that drainage networks are no longer effective in the face of extreme precipitation events. In addition, low-impact development measures such as green roofs, rain gardens, and permeable paving have not been widely used. This means that it is difficult to store and regulate rainwater, which results in rainwater runoff being formed in the shortest time and directly discharged to the pipeline, easily exceeding the maximum load of the drainage pipeline.

### 1.3.2 Urban waterlogging simulation and susceptibility assessment methods

Simulating and predicting the waterlogging variation could provide useful theoretical and practical references for urban waterlogging prevention, sustainable urban development, and urban planning (Zhang, 2015; Zope et al., 2016; Zhao et al., 2018). As many researchers have pointed out, urban waterlogging is influenced by many factors (both natural conditions and human activities). The spatial heterogeneity of the landscape elements leads to the spatial non-stationary and nonlinear characteristics of urban waterlogging. Moreover, buildings, streets, curbstones, and other landscape elements in the urban underlying surface will affect the movement of surface runoff. Therefore, it is a challenging task to construct an accurate and efficient model to simulate and predict the spatial variation of urban waterlogging. In general, methods of characterizing urban waterlogging variation can be summarized into four categories: (1) the multivariate statistical methods, (2) the hydrological and hydrodynamic models, (3) the qualitative model based on expert knowledge, and (4) the machine learning models.



### 1.3.2.1 Multivariate statistical methods

The multivariate statistical methods (such as the stepwise regression model, geographic regression weighted model) are widely used to analyze the impact of various variables on waterlogging (Gaitan and Ten, 2015; Zhang et al., 2020; Wang et al., 2017; Felder et al., 2017; Hong et al., 2017). These methods calculate the impact of each type of condition factor on waterlogging, and thus project the susceptibility of urban waterlogging in the region. Liu et al. (2021) used the geographic regression weighted model (GWR) based on principal component analysis (PCA) to evaluate the risk of urban waterlogging. Tran et al. (2020) based on population density, road density, distance to water bodies, impervious surface percentage, normalized difference vegetation index, and the digital elevation model (30-m resolution), using the GWR method to predict the probability of urban waterlogging risk in Hanoi, Vietnam from 2012 to 2018. Littidej et al. (2019) used GWR to analyze the stability of flood areas in Thailand. Fahy et al. (2019) used spatial models, including spatial error models, spatial lag models, and geographically weighted regression models, to further identify areas that may be severely affected by flooding. However, due to the tremendous landscape heterogeneity, it is difficult to utilize them to simulate the spatial variation of urban waterlogging accurately. Consequently, this method is gradually being replaced by more robust and precise methods.

### 1.3.2.2 Hydrological and hydrodynamic models

Concerning the second group, the hydrological and hydrodynamic models (such as SWMM, MIKE, HEC-RAS, LISFLOOD-FP) are extensively utilized to simulate the urban waterlogging process (Mignot and Dewals, 2019; Zoppou, 2001; Elliott and Trowsdale, 2007; Youssef et al., 2011; Quan 2014b; Bisht et al., 2016; Cheng et al., 2017; Li et al., 2016). Depending on the simulation algorithm, hydrological and hydrodynamic models can be divided into three types: hydrological models, hydrodynamic models, and simplified models. Table 1-3 compares the advantages and disadvantages of these types of models.

*Table 1-3 Comparison of representative urban waterlogging simulation models.*

Model/type	Algorithm	The difficulty of model construction	Computational efficiency	Simulate the process of stagnant water
SWMM/ hydrological model	Nonlinear reservoir method	General	Quick	No

---

FloodMap- HydroInundation2D/ hydrodynamic models	St Venant equations	General	General	Yes
HEC-RAS/ hydrodynamic models	St Venant equations	Difficult	Slow	Yes
LISFLOOD-FP/ hydrodynamic models	Shallow-water equations	General	General	Yes
FCDC/Simplified models	Areas below the elevation of the water surface are submerged	General	Quick	No
CA model/Simplified models	The amount of water exchanged between cells is calculated based on topographic and hydraulic equations	Easy	General	Yes

---

### 1) Hydrological models

In general, the hydrological model divides the urban area into multiple catchments according to the water outlet and regarding each catchment as an independent unit to calculate the processes of surface runoff generation and confluence (Bisht et al., 2016; Kai et al., 2017; Burger et al., 2014; Babaei et al., 2018). At present, representative urban hydrological models mainly include SWMM (Storm Water Management Model) (Rossman, 2010; Gironás et al., 2010), UCURM (University of Cincinnati Urban Runoff Model) (Papadakis and Preul, 1972), ILLUDAS (Illinois Urban Drainage Area Simulator) (Terstriep et al, 1974). Such models are widely used in urban flood modeling due to their low data requirements and high computational efficiency. However, these models typically estimate runoff based on empirical estimation or the curve proposed by the Soil Conservation Service (SCS-CN), which may not be sufficient to describe specific differences in complex urban landscapes (Zhang et al., 2018b; Zope et al., 2016; Seenu et al., 2020). Moreover, due to the complexity of

the underlying surface, the artificial structures (buildings, roads) or trees will change the direction of the surface runoff, resulting in the complicated water flow movement. Therefore, it is difficult to obtain the runoff generation and confluence of various catchments (such as the variation of water depth and flow velocity). This undoubtedly limits the application of hydrological models in urban areas to some extent.

## 2) Hydrodynamic models

Compared with hydrological models, hydrodynamic models with a high spatial resolution are able to simulate surface runoff process under complex urban underlying surfaces, with better accuracy and credibility (Li et al., 2019a; Pradhan et al., 2011; Lin et al., 2006; Pinho et al., 2015). This type of model is based on the partial differential equations of water movement and boundary conditions to calculate surface runoff processes under different topographic conditions. According to the different dimensions of surface runoff processes expressed by the models, the hydrodynamic models can be divided into 1D and 2D hydrodynamic models. Since the 1D hydrodynamic models cannot simulate the lateral diffusion of water waves, the 2D hydrodynamic models based on Saint-Venant's equation are well overcome the deficiencies of the 1D hydrodynamic models (Tsanis and Boyle 2001; Paiva et al., 2011; Felder et al., 2017). In practice, the modeling of these 2D hydrodynamic models based on physical mechanisms relies heavily on a large amount of high-precision local data, such as high-resolution DTM/DEM data and drainage networks. Often, such detailed data is difficult to obtain. Therefore, establishing a hydrodynamic model is a difficult task for many cities. Additionally, these hydrodynamic models require large computational resources, so they can only be applied to a small range of research areas. Although these models can accurately simulate the physical process of waterlogging in relatively small watersheds, the difficulty of modeling and low computational efficiency limit the application and promotion of such models.

## 3) Simplified models

The simplified model is mainly based on high-precision topographic data, which simulates the inundation range of urban waterlogging based on gravity and topography-driven hydraulic balance and hydraulic exchange. This type of model does not consider the complex hydrological or hydrodynamic process and the required data is not as restrictive as the hydrological and hydrodynamic model. Zhang and Pan (2014) proposed a GIS-based urban waterlogging simulation method based on the simplification of the distributed hydrological model. Taking Haining City, China as the study area, a study proposes an improved Simplified Urban Storm Inundation Model (SUSIM) that integrates urban topography, precipitation, surface runoff, and

inundation models, allowing for fast and accurately simulate different inundation conditions (Li et al., 2019b). The cellular automata model was used to simulate the temporal and spatial evolution of floods in Keighley Stockbridge, England (Ghimire et al., 2013). The results were compared with the hydrodynamic model, which showed a good agreement between the two models. Zhang et al. (2014) introduced the flood connected domain calculation (FCDC) method to simulate river inundation. The method considers flow continuity and allows for rapid simulation of the inundation of source floods, such as river floods or levee floods. The simplified model has been widely used because of its advantages such as saving time and simple theory.

#### 1.3.2.3 Qualitative models based on expert knowledge

The qualitative models such as the analytic hierarchy process (AHP) and multi-criteria decision analysis (MCDA) strongly depend on expert knowledge (Nigusse and Adhanom, 2019; Tang et al., 2018b; Samanta et al., 2016; Brito et al., 2019). These qualitative models use the AHP to determine factor weights or integrate explanatory factors into a multicriteria sensitivity map to simulate urban waterlogging events (Zhao et al., 2018; Chowdary et al., 2013). For example, Hong et al. (2018) used the hierarchy process and fuzzy weight evidence to construct a flood susceptibility map. The results point out that this method provides a simple and intuitive explanation of the weights of flood-related variables, which is more appropriate to deal with the fuzziness of expert judgment. Samanta et al. (2016) evaluated the use of MCDA in inland flood risk analysis. Lim and Dong. (2009) presents a combined GIS with spatial MCDA to evaluate flood damage in Suyoung River Basin, Busan, Korea. Paquette et al. (2012) used the MCDA approach combining elevation, catchment, land use, slope, distance to channel, and soil type data to model the spatial extent of flood hazards and assess flood risk in the Nadi River Basin in western Fiji. Roy et al. (2021) used integrated AHP and GIS technology to model and identify urban waterlogging disasters, vulnerabilities, and risk areas in Siliguri, the gateway to northeast India. To some extent, these studies confirm the role of such methods in waterlogging simulation and risk assessment. However, some studies have pointed out that the methods rely on expert knowledge and judgment, which introduces uncertainty (Chowdary et al., 2013).

#### 1.3.2.4 Machine learning models

In the past ten years, machine learning models have been widely used in urban waterlogging simulation, susceptibility modeling, and risk assessment, including Artificial Neural Networks (ANN) (Kia et al. 2012), Support Vector Machines (SVM) (Tehrany et al., 2015), Decision Tree (DT) (Tehrany et al. 2013) and Random Forest (RF) (Zhao et al. 2018). These methods are regarded as a black box to map the relation between input

and output of training samples, which shows advantages in complex data modeling. Most recently, Gupta et al. (2017) identified the urban waterlogging sensitive areas and predicted the severity using an ANN, which indicated that this method could effectively and accurately predict the severity of waterlogging. Zhao et al. (2018) applied a Random Forest model to map the flood susceptibility in mountainous areas of China based on historical flooding records from 1949 to 2000. For the SVM, Tang et al. (2019) applied a particle swarm optimization and an SVM in an integrated approach to evaluate the urban waterlogging susceptibility. Tehrany et al. (2019) evaluate the performance difference between decision tree (DT) and support vector machine (SVM) in assessing the sensitivity of urban flood disasters. Furthermore, Tehrany et al. (2013) also applied a rule-based decision tree to predict the flood susceptible areas in the Kelantan River basin. Ke et al. (2020) considered machine learning (ML) and rainfall thresholds to classify floods and non-flood events, providing a reference for cities to respond to frequent floods. Chen et al., (2021) used six machine learning models (Support Vector Machines, Random Forest, Multilayer Perceptron, Gradient Boosting Decision Tree, eXtreme Gradient Boosting, and Convolutional Neural Network) to assess the flood risk in the Pearl River Delta. This research expands the application of machine learning methods in urban flood risk assessment and deepens our understanding of the underlying mechanisms of flood risk. The results of these studies have confirmed to some extent that the machine learning models are effective in urban waterlogging simulation and risk assessment.

However, these models are sensitive to the quality of the sample data. It is important to select the same number of positive and negative samples to train machine learning. Due to the short duration and spatial dispersion of urban waterlogging, researchers can only obtain relatively few historical waterlogging locations (positive samples), which leads to a limited number of negative samples selected, further affecting the performance of the models. Therefore, some scholars propose to use semi-supervised learning models to make full use of unlabeled data sets to improve model accuracy (Zhao et al., 2019). Similarly, Tang et al. (2019) adopted a random repetition method to select negative samples, which not only theoretically avoids the influence of subjectivity, but also avoids the selection of inaccurate negative samples as much as possible. Furthermore, Tang et al. (2021) proposed an optimized seed propagation algorithm (OSSA), which estimates the potential inundation area based on the altitude and the spatial distribution of natural waters, thereby increasing the number of positive samples. The results point out that this method can significantly improve the accuracy of flooding susceptibility assessment.

### 1.3.3 Mitigation measures for urban waterlogging

#### 1.3.3.1 Traditional drainage facilities

China's current stormwater management concept for flood control and drainage planning is still "drainage oriented", specifically the construction of urban drainage networks and other series of engineering facilities. Local authorities increased underground drainage pipes and pumping stations to speed up rainwater drainage and reduce surface runoff. However, this approach only accelerates the discharge rate of surface runoff but cannot reduce the total surface runoff (Lee et al., 2016). The surface runoff transfer to other places in a short period may bring more pressure on the local drainage systems. Furthermore, drainage facilities block the recharge channel for groundwater, leading to a constant decline of groundwater level, threatening urban geological safety (land subsidence) (Zhang et al., 2012; Zhang et al., 2015). Traditional drainage facilities have the following limitations in alleviating urban waterlogging problems:

##### 1) Difficult to reconstruct drainage facilities

Due to the low standards of China's drainage system, many urban drainage pipelines are seriously blocked, especially in the historical urban district, which directly affects the drainage function. It is a complicated process to upgrade and reform the drainage network. For example, the large-scale reconstruction of the underground drainage system is difficult, expensive, and affects traffic and daily activities. Additionally, intensive urban development leaves insufficient space for underground drainage systems to reconstruct. For most cities, it is almost unrealistic to raise pipeline standards to prevent urban waterlogging in a short period.

##### 2) Poor connection between management departments

In the process of managing urban stormwater problems, there is a lack of coordination and cooperation between relevant departments. For example, as the "Implementation Opinions on Strengthening Urban Waterlogging Management" (General Office of the State Council of China) pointed out that the elevation conflict between urban rivers and drainage systems is prominent. This phenomenon may result in the drainage outlet being submerged by river level or blocked by silt, resulting in poor drainage.

#### 1.3.3.2 Urban river networks and water bodies

It is well known that urban river networks or water bodies can effectively regulate surface runoff. However, along with the process of rapid urbanization, rivers within cities have been artificially landfilled, ditched, and

reduce river discharge sections to obtain more land resources for urban construction (Table 1-4). This has led to serious disruptions to the urban river network, such as a reduction in the complexity of the river network, a reduction in the number of tributaries, and a narrowing of the river channels. In addition, many water bodies, ponds, wetlands and have been encroached by urban construction, reducing the function of stormwater storage and diversion. The superposition of these reasons has caused the urban waterlogging phenomenon to become more and more serious.

*Table 1-4 The impact of urbanization on rivers and water bodies.*

Category	Influence
River	Increase runoff and river flow
	River temperature increases
	Increase soil erosion
	Raise the riverbed
	Increase the pollutant content
	Modify river channel
Water body	Unable to regulate surface runoff
	Reduce groundwater recharge
	Pollution accumulation
	Eutrophication

The urban river networks and water bodies play a key role in stormwater and flood management (Lee et al., 2016). However, due to urban expansion, the natural attributes of rivers and water bodies are often neglected. Taking Guangzhou's downtown area as an example, it is possible to observe that the water area was 131.74 km<sup>2</sup> in 1990, but only 108.48 km<sup>2</sup> in 2010, a decrease of 23.26 km<sup>2</sup> compared to 1990, a reduction of about 17.66%. The river's shoreline decreased from 3833.16 km in 1990 to 2798.48 km in 2010, a reduction of approximately 30% (Chen et al., 2013). With the loss of these water areas, the capacity of rainwater storage in the city is reduced, resulting in an unbalanced urban water cycle, threatening the city's water security. At present, local authorities have realized that urban drainage networks alone cannot adequately resist urban waterlogging disasters. Protecting the natural attributes of river networks is a key measure to address the growing problem of waterlogging in cities.

### 1.3.3.3 Sponge city

Facing the increasingly serious problem of urban waterlogging, scholars have proposed measures such as “sponge city” or “low impact development approach” to mitigate urban waterlogging (Ahammed, 2017; Mai et al., 2018; Dong et al., 2018; Huang et al., 2017; Miao et al., 2019; Shao et al., 2016; Wu et al., 2018; Zimmer et al., 2007). With the concept of "sponge city" and "low impact development" put forward, more and more cities adopt the concept of green development in urban construction, gradually considering the infiltration, accumulation, and retention of rainwater (Figure 1-12). The task of Sponge City is to realize urban rainwater management. It focuses on rainwater storage, retains, and infiltration by using reservoirs, soakaways, storage tanks, artificial wetlands, so as to reduce the pressure of the urban drainage system. Additionally, it is also mindful of the need to maintain public health, prevent water pollution, protect biodiversity and future natural resources. These methods are primarily aimed at increasing the permeable surface of the city (e.g. urban green infrastructure) and thus offsetting the increase in surface runoff.



Figure 1-12 Schematic diagram of the sponge city concept. (Source: Chen and Chen, 2020)

#### 1) Green roof

A green roof is planting a layer of vegetation on the building roof, which has a significant effect on reducing



the total amount of runoff (Figure 1-13) (Peng and Jim, 2015; Berndtsson, 2010; Demuzere et al., 2014; Krebs et al., 2016). The study indicated that the green roof can reduce the runoff by 50% compared with the ungreened roofs (Hall, 2010). Stovin et al. (2012) analyzed the rainwater regulation capability of the experimental green roof by using the rainfall data of 29 months in Sheffield, UK. The results showed that the average interception rate of green roofs was 50.2%. Xu (2007) conducted an experiment, which showed that the drainage rate of the green roof was 5 L/m<sup>2</sup>, while that of the ungreened roof was 16 L/m<sup>2</sup>, within 15min, when the rainfall intensity was 20 L/m. Mentens et al. (2006) indicated that if 10% of the buildings in the area adopt green roof measurement, the total regional surface runoff can be reduced by 2.7%. Although the ability of green roofs to retain and intercept rainwater has been widely recognized by scholars, its retain and intercept capabilities vary greatly in different regions and different environmental conditions. For example, Tang et al. (2011) pointed out that as the thickness of the soil layer increases, the capacity of green roofs to retain and intercept rainwater increases. Teemusk et al. (2007) found that when the rainfall intensity is moderate to light rain, the average interception rate of the green roof is 87%, but in heavy rain, the interception capacity of the green roof will be greatly weakened. Through monitoring in Brussels (Belgium), it is found that green roofs can reduce the surface runoff of individual buildings by 54%, with this reduction being more significant in summer than in winter (Mentens et al., 2006).

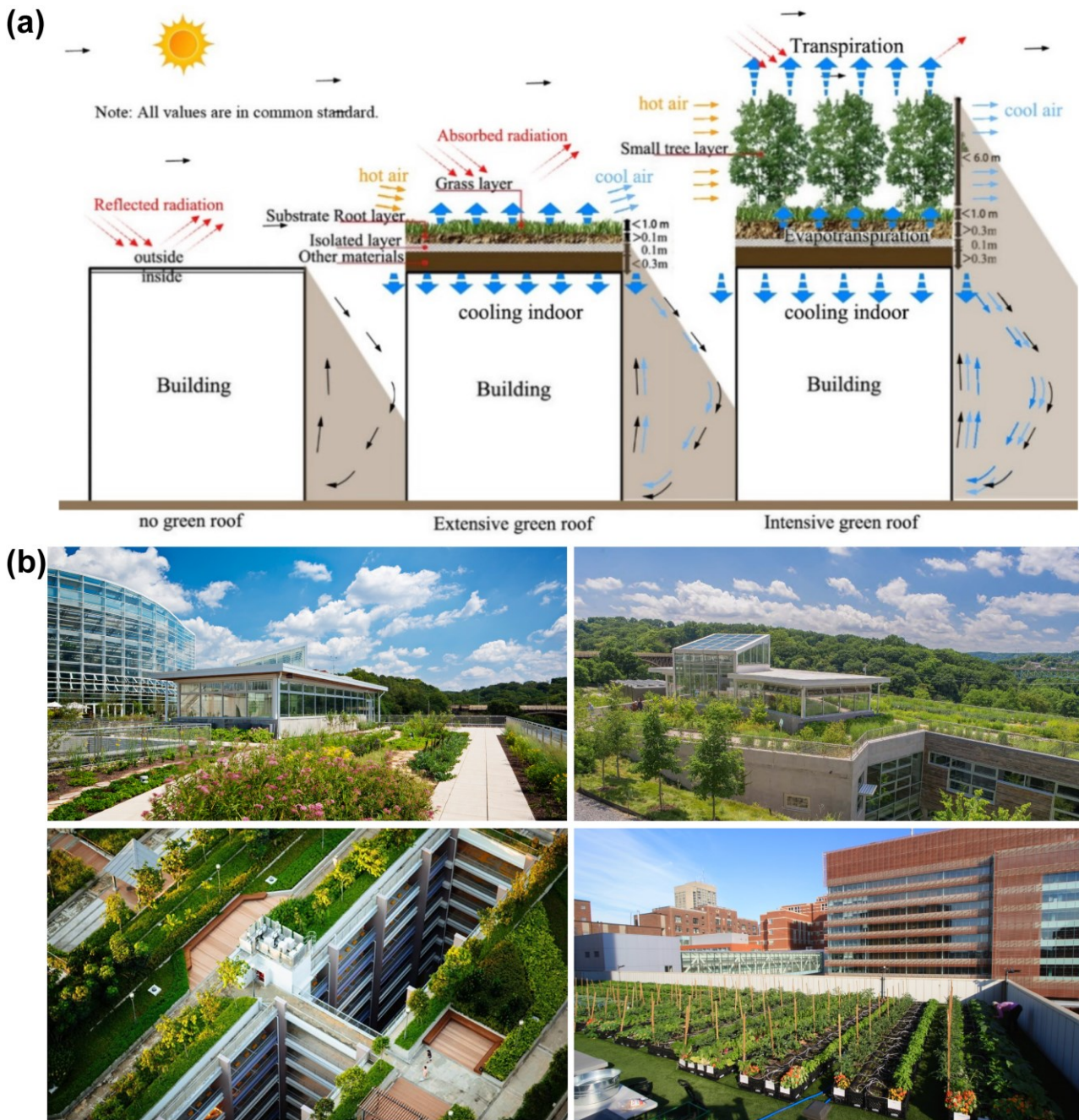


Figure 1-13 Schematic diagram of green roof (a) and the selected Green Roof Awards of Excellence design projects (b).

(Source: Zhang et al., 2019; Green Roofs for Healthy Cities' Awards of Excellence)

## 2) Concave green land

Concave green land similar to rain gardens is a key measure in 'sponge cities'. It refers to vegetation land with a lower elevation than the surrounding environment, which can temporarily retain surface runoff (Figure 1-14) (Ministry of Housing and Urban-Rural Development, 2014). Winston et al. (2016) showed that concave green land can reduce the peak flow by 24-96% for one-year rainfall events. Du et al. (2019) utilized the SCS model to compare the effect of conventional green space, 10 cm concave green land, and 20 cm concave

green land in reducing surface runoff in Shanghai city (China). The results show that the concave green land with a depth of 0.10–0.20 cm can reduce the direct runoff by 23.63–98.35% and the inundation area by 26.09–82.41%, decreasing waterlogging exposed population by 0.40–1.04 million. Liu et al. (2015a) assessed the effectiveness of four types of green infrastructure, including green space enlargement, concave green space, retention ponds, and porous brick pavement, in reducing rainwater runoff. The results noted that when the depth of the concave green space increases from 0 cm (flat) to 10 cm, the ratio of runoff to rainfall decreases from 77.00% to 65.51%, while up to 95% of runoff was reduced when the green space was converted to a 4 cm deep. Wen et al. (2016) showed that under 5-year recurrence rainfall conditions, modifying the green space to a concave pattern with a depth of 0.05 m could reduce runoff and peak flow by 23.20% and 29.11%, respectively. Zhu et al. (2015) used artificial simulation to analyze the flow reduction effect of the concave green space under different rainstorm intensities in Beijing, which indicated that the concave green areas have a significant effect on the reduction of stormwater runoff. Similarly, Wen et al. (2016) and Cai et al. (2011) also demonstrated that a concave-shaped UGI would significantly reduce the surface runoff and peak flood flows.

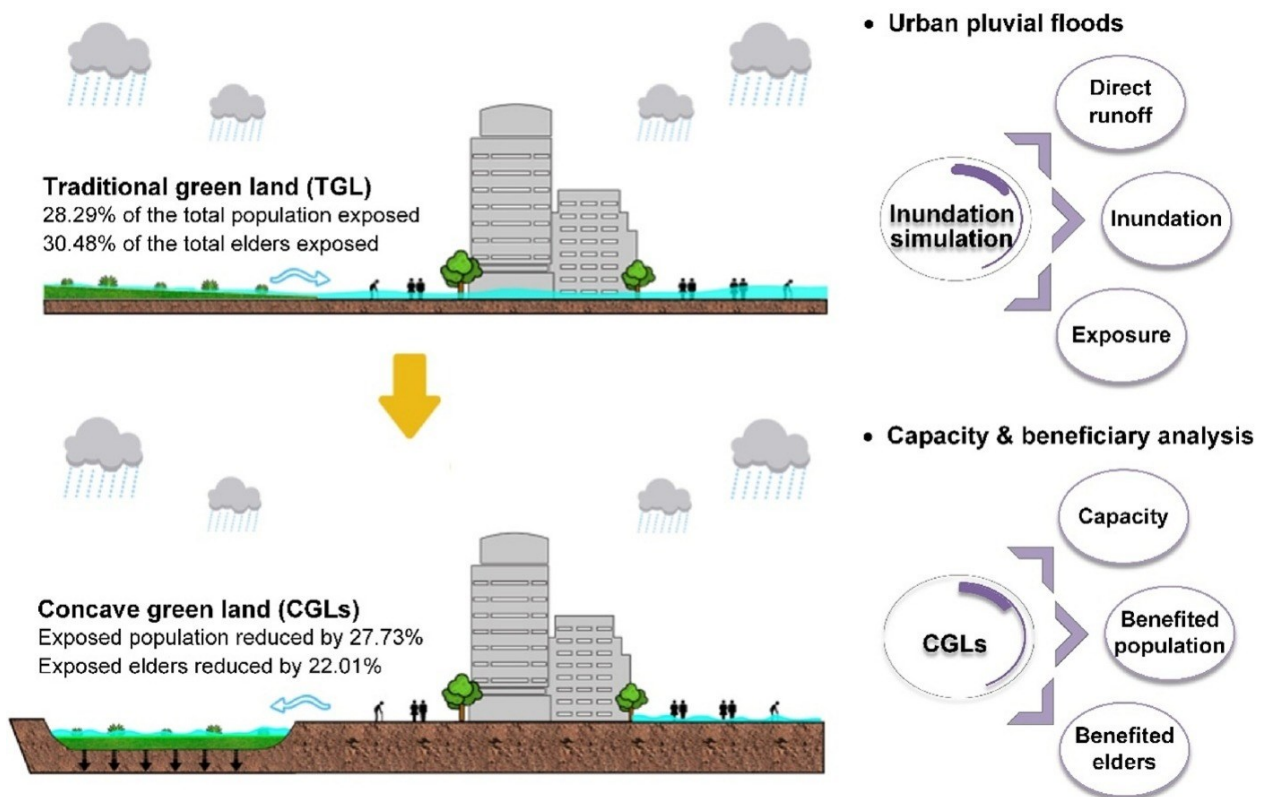


Figure 1-14 Schematic diagram of concave green land. (Source: Du et al., 2019)

### 3) Permeable pavement

The combination of porous brick and permeable pavement can increase the permeable area of the city and effectively alleviate surface runoff (Figure 1-15). Liu et al. (2020c) simulated two rainfall intensities (low intensity is 0.3 mm/min with a depth of 25.4 mm, high intensity is 0.6 mm/min with a depth of 42.0 mm), confirming that permeable surfaces can significantly reduce surface runoff and peak flow. Zhu et al. (2019) used the storm-water management model (SWMM) to quantify the permeable pavement on surface runoff mitigation effect in a two-way six-lane road in Nanjing. The results point out that the permeable pavement has a good effect on reducing the runoff coefficient and flood peaks, which can effectively reduce the pressure of rainwater drainage. Liu et al. (2020b) showed that the average runoff time for permeable pavements was approximately 78.5 minutes for low-intensity rainfall and reduced to only 51.5 minutes for high-intensity rainfall, with an average runoff interception of 52.5% and 42.5% for low intensity and high-intensity rainfall respectively. Liu et al. (2015b) took Beijing as an example, and the SCS model analysis showed that porous pavement could significantly reduce the runoff coefficient by 28.1%. Theoretically, the larger the permeable paving area, the greater the ability to reduce surface runoff. However, some studies point out that the ability of permeable surfaces to reduce surface runoff is influenced by many factors, such as the dryness or wetness of the original surface, the permeable paving material, the amount of rainfall, and the intensity of rainfall (Rodriguez-Hernandez et al., 2015). It has been shown that when surface runoff is excessive (40% of precipitation), even if the percentage of permeable area is increased to 90%, the mitigation effect of permeable pavement on surface runoff is not significant (Liu et al., 2014).

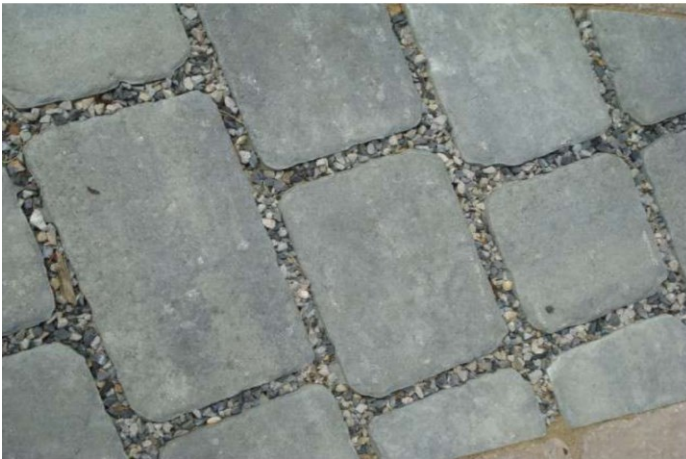
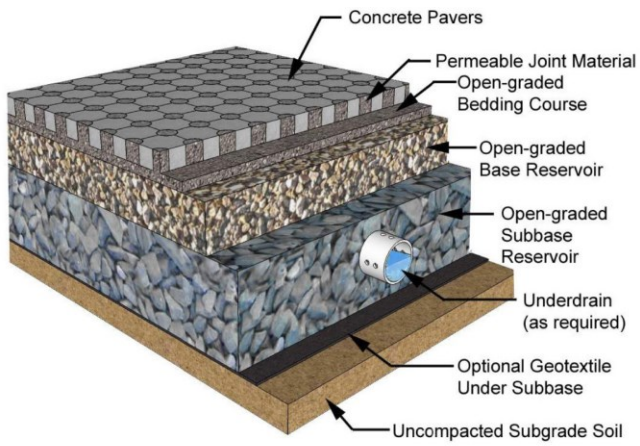


Figure 1-15 Schematic diagram of permeable pavement. (Source: Freeborn et al., 2012)

## 1.4 Research questions and objectives

With the intensification of human activities and climate change, it is foreseeable that in the future, urban waterlogging events in urban centers will undoubtedly increase. Therefore, in order to gain a general knowledge of urban waterlogging patterns at a larger scale (such as in metropolitan areas), it is necessary to identify the spatial-temporal characteristics of urban waterlogging events. However, most of the previous studies have only focused on the waterlogging pattern in relatively small catchments. While studies at a small area may be site-specific, it is difficult to macroscopically investigate the spatial distribution pattern of urban waterlogging. Urban waterlogging is a complex phenomenon, which is the result of the joint action of natural conditions and human activities. However, what is less known, is the relative contribution of each environmental and anthropogenic factor. Most studies just focus on one land cover type - the impervious surface alone, without comprehensively considering topographical factors, overall land cover characteristics (composition and spatial configuration), drainage facilities, precipitation intensity, and other influencing factors. This will undoubtedly bring some uncertainty to the results of the study. Simultaneously, most studies have not carried out a multi-scale analysis, and thus ignoring the scale effects of various influencing factors (Yu et al., 2018; Wang et al., 2015). Due to the great heterogeneity in highly urbanized areas, the analysis of one single scale may not be sufficient. From this perspective, some interesting questions arise: does the impact of environmental and anthropogenic factors on urban waterlogging vary with the scale of analysis? Does the explanatory power of the same influence factor change with the change of scale of analysis? Does the dominant factor determining waterlogging vary across analysis scales? Answering these questions is vital for expanding our scientific understanding of the linkage between urban waterlogging and its environmental and anthropogenic influencing factors.

In highly urbanization areas, the spatial distribution of urban landscape elements and their attributes are characterized by heterogeneous and dynamic. This phenomenon is widely regarded as spatial heterogeneity, describing the complexity of the landscape pattern and its process (Pickett and Cadenasso, 1995; Song et al., 2020; Jiang, 2015). Due to the spatial heterogeneity effect, urban landscape elements (such as topographic features, land cover characteristics, drainage facilities) are different in different spatial locations. Hence, the relationship between landscape elements and urban waterlogging varies with different spatial locations, which is known as spatial non-stationarity. Therefore, this suggests that the waterlogging driving forces in different regions are complex and different. Although there have been considerable studies exploring the

mechanism of urban waterlogging, most of the previous efforts regard the whole study area as spatial homogeneous, without taking spatial heterogeneity into account (Ke et al., 2020; Yu et al., 2018; Sun, 2014; Su et al., 2018; Zhang et al., 2021b; Jian et al., 2021). Analyzing the waterlogging mechanisms at the entire city scale, undoubtedly only the universal mechanisms can be obtained. Thereby ignoring the local driving forces of urban waterlogging. As a result, it is impossible to know how landscape elements in different spatial locations affect urban waterlogging. This shortcoming limited local authorities from developing more target-specific urban waterlogging mitigation strategies for different local conditions.

In general, characterizing the urban waterlogging variation is conducive to revealing the urban waterlogging prone areas, thereby minimizing waterlogging negative effects (Wang et al., 2012; Miao et al., 2019; Tang et al., 2018a). However, as many researchers have pointed out, urban waterlogging is influenced by the natural environment (precipitation and urban topography) and human activities (land-use change and drainage network) (Wu and Zhang 2017; Zhang et al., 2018b; Su et al., 2018). Additionally, the spatial heterogeneity of highly urbanized areas further leads to the non-stationary and non-linear characteristics of urban waterlogging. Thereby, it is difficult to adequately capture the spatial variation of urban waterlogging and identify waterlogging susceptibility areas by traditional statistical methods (i.e. global regression model, stepwise regression). Moreover, we cannot predict the spatial variation of urban waterlogging under different future scenarios (climate change and urban development). Therefore, there is an urgent need for a more robust and computational efficiency model for urban waterlogging simulation and to identify the urban waterlogging susceptibility areas under different future scenarios.

It is widely accepted that urban green infrastructure (UGI), as a permeable surface, can effectively absorb and store rainwater (Kim and Park, 2016; Yao et al., 2015; Yang et al., 2015; Zhang et al., 2021a). In addition, the canopy and rhizome of vegetation can intercept surface runoff, thereby reducing the speed of runoff collection (Yang et al., 2013; Liu et al., 2014). Compared with drainage facilities and pumping stations, urban green infrastructure can be more "environmentally friendly" to discharge surface runoff. However, due to the shortage of land resources in metropolitan areas, it is unrealistic to reduce urban waterlogging by considerably increasing the UGI area. Therefore, it is necessary to understand how to utilize the limited UGI area to maximize the waterlogging mitigation function. However, will the risk of urban waterlogging continue to decrease as the area of UGI increases? Less attention has been paid to investigating the threshold level for the impact of UGI on urban waterlogging. If the mitigation effect of UGI has a threshold level, planning a

larger area of UGI may not provide a more significant mitigation effect. In this context, another research question arises: Is there a threshold level for the effect of UGI on urban waterlogging? Answering these questions can help us improve our understanding of the potential mitigation effect of UGI on urban waterlogging and furnish concrete references for UGI design. Additionally, considerable studies have examined the relationship between UGI's factors and waterlogging. However, previous studies mainly focused on the individual effects of UGI factors on urban waterlogging; instead, the interactive effects of these factors remain unclear. This phenomenon gives rise to some questions: how do the interactions of these UGI factors affect urban waterlogging? Can the interaction between different UGI factors further enhance their mitigation effects on waterlogging?

Given the research gaps mentioned above, this thesis proposes the following research questions:

- 1) What is the spatial-temporal distribution pattern of urban waterlogging events?
- 2) Does the influencing factor of urban waterlogging have different explanatory power under different scales of analysis?
- 3) How can we investigate the spatial heterogeneous driving forces of urban waterlogging?
- 4) How can we obtain the site-specific waterlogging mitigation strategies for different spatial locations?
- 5) How can we capture the urban waterlogging spatial variation and assess urban waterlogging susceptibility under different urbanization and climate change scenarios?
- 6) Is there a threshold level for the mitigation capacity of urban green infrastructure on urban waterlogging?
- 7) How can we enhance the mitigation effect of urban green infrastructure?

Hence, to address these questions, the specific objectives of this study are as follows:

- 1) identify the urban waterlogging hotspots and assess the agglomeration effect of urban waterlogging events.
- 2) clarify the relative contributions of influencing factors and examine whether the dominant factors vary across different analysis scales, thus identifying the scale effect.
- 3) develop an innovative method to investigate the spatial non-stationarity effects of landscape elements on urban waterlogging.



- 4) identify the individual and interaction effect of local driving forces according to different spatial locations.
- 5) propose a new and more robust method - the stepwise cluster analysis model to simulate waterlogging variation and assess urban waterlogging susceptibility under different scenarios.
- 6) explore the quantitative relationship between urban green infrastructure and urban waterlogging to quantify the threshold level of urban green infrastructure mitigation effect.
- 7) assess the effectiveness and stability of urban green infrastructure in mitigating urban waterlogging and reveal the factors that can effectively mitigate waterlogging magnitude.

In light of the research objectives presented, four papers have been written following the logical construction of research questions, objectives, and outcomes:

- i. The first article addresses the first and second research questions and objectives.

Zhang, Q., Wu, Z., Zhang, H., Dalla Fontana, G., & Tarolli, P. (2020). Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China. *Journal of Environmental Management*, 271, 110951. <https://doi.org/10.1016/j.jenvman.2020.110951>

- ii. The second article addresses the third and fourth research questions and objectives.

Zhang, Q., Wu, Z., Guo, G., & Tarolli, P. (2021). A new approach to investigating the spatially heterogeneous driving forces of urban waterlogging. *Under review*

- iii. The third article addresses the fifth research question and objectives.

Zhang, Q., Wu, Z., Guo, G., Zhang, H., & Tarolli, P. (2021). Explicit the urban waterlogging spatial variation and its driving factors: The stepwise cluster analysis model and hierarchical partitioning analysis approach. *Science of The Total Environment*, 763, 143041. <https://doi.org/10.1016/j.scitotenv.2020.143041>

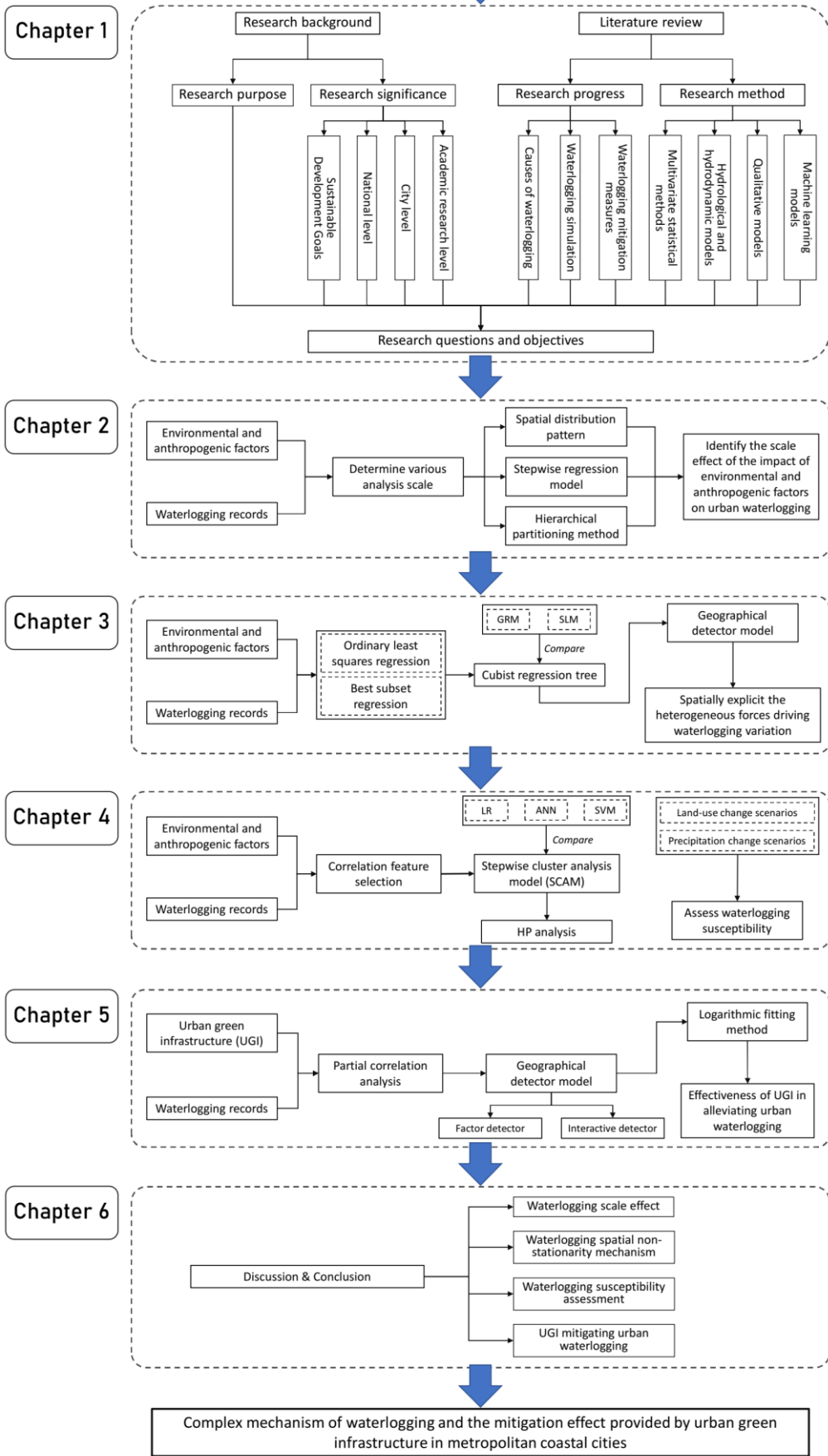
- iv. The fourth article addresses the sixth and seventh research questions and objectives.

Zhang, Q., Wu, Z., & Tarolli, P. (2021). Investigating the Role of Green Infrastructure on Urban WaterLogging: Evidence from Metropolitan Coastal Cities. *Remote Sensing*, 13(12), 2341. <https://doi.org/10.3390/rs13122341>

## 1.5 General organization

This thesis consists of four papers (chapters from 2 to 5) and the purpose of this research is to reach a thorough understanding of the complex mechanism of waterlogging and the mitigation effect provided by urban green infrastructure (UGI) in metropolitan coastal cities. Figure 1-16 illustrates the research framework for this thesis. It started with the reveal of the scale effect of environmental and anthropogenic factors on urban waterlogging, which is presented in chapter 2. Secondly, the mechanism of urban waterlogging and the valuable local driving forces are spatially clarified in chapter 3. Further, the characteristics of waterlogging variation and the identification of waterlogging susceptibility areas under different climate and urbanization scenarios are described in chapter 4. Lastly, in chapter 5, the mitigation effect of urban green infrastructure on urban waterlogging is examined by investigating the threshold level of waterlogging mitigation capacity as well as the UGI factors that can effectively mitigate waterlogging magnitude.

Complex mechanism of waterlogging and the mitigation effect provided by urban green infrastructure in metropolitan coastal cities



*Figure 1-16 The research framework for this thesis.*

The detailed information of each chapter is introduced below:

The first paper (Chapter 2) entitled “Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China” has been published in the *Journal of Environmental Management* in 2020. This paper identifies the relative contribution of each environmental and anthropogenic factor and investigates the stability linking waterlogging to influencing factors at multiple scales of analysis. It confirmed that urban waterlogging events are mainly affected by both land cover characteristics and urban topography. Additionally, this paper stresses the fact that the dominant drivers vary across different analysis scales, presenting a significant scale effect. Due to the appropriate statistical scale may only work for specific influencing factors, a universal “optimal” analysis scale for urban waterlogging studies can not be confirmed. This paper provides additional insights that the best analysis scale for urban waterlogging study should be determined by the characteristics of study areas.

The second paper (Chapter 3) titled “A new approach to investigating the spatially heterogeneous driving forces of urban waterlogging” is under review. It provides a general framework integrating the best subset regression model, cubist regression tree, and geographical detector model to spatially explicit the heterogeneous forces driving waterlogging variation and identify the waterlogging dominant factors with different local conditions. This paper offers the opportunity to objectively select the most representative and meaningful driving factors that effectively describe urban waterlogging variation according to local characteristics. The results indicate that the driving force of urban waterlogging varies with the local site conditions. Understanding the complex site-specific mechanism of urban waterlogging in different watershed units will help us implement more targeted and effective mitigation strategies, rather than a “one-size-fits-all” policy.

The third paper (Chapter 4) entitled “Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China” has been published in the *Science of The Total Environment* in 2020. This study proposes a novel approach that combined a stepwise cluster analysis model (SCAM) and hierarchical partitioning analysis (HPA) within a general framework to fully capture the urban waterlogging spatial variation and identify the waterlogging susceptibility areas under different urbanization and climate change scenarios. This research underlines that the SCAM can provide accurate and detailed simulated results both in urban centers where waterlogging frequently occurs and urban fringe with few

waterlogging events, compared with logistic regression, artificial neural network, and support vector machine. The watershed spatial location and watershed characteristics are relevant aspects to be considered in identifying and assessing waterlogging susceptibility, which provides original insights that urban waterlogging mitigation strategies should be developed according to different local conditions and future scenarios.

The fourth paper (Chapter 5) entitled “Investigating the role of green infrastructure on urban waterlogging: evidence from metropolitan coastal cities” has been published in the *Remote Sensing* in 2021. The purpose of this research is to examine the effectiveness and stability of urban green infrastructure (UGI) in alleviating urban waterlogging and the threshold level of waterlogging mitigation capacity. This paper considered two waterlogging high-risk coastal cities—Guangzhou and Shenzhen, to investigate the individual and interaction effect of UGI’s factors on waterlogging magnitude. It showed that the interaction of UGI factors greatly enhances their individual effects on waterlogging, indicating that the interactive effect of multiple factors can further alleviate the degree of urban waterlogging. Further, it underlined that the impact of UGI on waterlogging presents a threshold phenomenon. The excessive area proportions of UGI within the watershed unit or an oversized UGI patch may lead to a waste of its mitigation effect. This paper provides additional insights that the area proportion of UGI and its mitigation effect should be considered comprehensively when planning UGI.

# CHAPTER 2

## 2. Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China<sup>1</sup>

Qifei Zhang<sup>a,b</sup>, Zhifeng Wu<sup>b,c</sup>, Hui Zhang<sup>d</sup>, Giancarlo Dalla Fontana<sup>a</sup>, Paolo Tarolli<sup>a\*</sup>

- a. Dept. of Land, Environment, Agriculture and Forestry, University of Padova, 35020 Legnaro, PD, Italy.
- b. School of Geographical Sciences, Guangzhou University, 510006 Guangzhou, Guangdong province, China.
- c. Southern Marine Science and Engineering Guangdong Laboratory, 511458 Guangzhou, Guangdong province, China.
- d. College of Surveying and Geo-informatics, North China University of Water Resources and Electric Power, 450046 Zhengzhou, Henan province, China.

---

<sup>1</sup> ***This chapter is a style-edited version of the published article:***

Zhang, Q., Wu, Z., Zhang, H., Dalla Fontana, G., & Tarolli, P. (2020). Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China. *Journal of Environmental Management*, 271, 110951.

***Contributions by the PhD candidate included:*** primary authorship in writing and revision of the entire article; primary responsible for literature review, data collection and quantification, methodology, data processing and analysis, and production of all figures.

## 2.1 Abstract

Urban waterlogging disasters are affected by both environmental conditions and human activities. Previous studies had explored the effect of land-use type on waterlogging in relatively small watersheds. Few, however, have comprehensively revealed the relative contributions of the environmental and anthropogenic factors to urban waterlogging concerning different scales of analysis. Indeed what is less known, are the dominant factors and the appropriate scale of analysis. To overcome this limitation, a novel method that integrates the stepwise regression model with hierarchical partitioning analysis is presented. The purpose is to investigate the complex mechanism of urban waterlogging by identifying the relative contribution of each environmental and anthropogenic factor and the stability linking waterlogging to influencing factors at multiple scales of analysis (i.e. 1 km, 2 km, 3 km, 4 km, and 5 km). We consider waterlogging events in the central urban districts of Guangzhou (PR China) from 2009 to 2015 as a case study. The results show that the spatial distribution of waterlogging events in the central urban area presents a strong agglomeration pattern. The waterlogging hot spots are mainly concentrated in the historical area of Guangzhou. Under all analysis scales, we find that the percent cover of urban green spaces (44.74%), percent cover of residential area (41.03%), and slope.std (36.85%) both have a dominant contribution to urban waterlogging, which suggests the importance of land cover composition in determining urban waterlogging. However, the relative contribution and dominant factors of waterlogging varied across different analysis scales, presenting a strong scale effect. Under a small analysis scale (1 km), the topography factors (slope.std and relative elevation) are confirmed as the dominant variables; however, with the increase of analysis scale, the influence of land cover composition (greenspace, residence area, grassland) and land cover spatial configuration (LPI, AI, Cohesion index) on waterlogging magnitude is greater than other factors. This finding provides additional insights that urban waterlogging can be alleviated by balancing the relative composition of land cover features as well as by optimizing their spatial configuration. Since the optimal statistical scale for urban waterlogging studies only worked for specific influencing factors, the appropriate analysis scale for urban waterlogging study should be determined by the characteristics of study areas. This study has the capability to extend our scientific understanding of the complex mechanisms of waterlogging in highly urbanized coastal cities and provide useful support for the prevention and management of urban waterlogging.

**Keywords:** urban waterlogging; environmental and anthropogenic factors; scale effect; landscape pattern;

land cover features

## 2.2 Introduction

Urban waterlogging is a stagnant water disaster occurring in the urban area, which mainly refers to the phenomenon that short-term heavy rainstorms or continuous precipitation exceed the drainage capacity of a city (Yin et al., 2011; Hammond et al., 2015; Xue et al., 2016). During the rapid urban sprawl and population growth, the proportion of the world's urban population is expected to continue to increase to more than 60% (Nations, 2014). With this rapid urbanization process, the impervious surfaces within the city increased dramatically, which changed the original land cover composition and hydrological conditions, so as consequence increased the risk of urban waterlogging events (Sofia et al., 2014; Du et al., 2015; Sofia et al., 2017; Su et al., 2018; Sofia et al. 2019). Furthermore, with the global climate changes, the frequent occurrence of extreme precipitation events has led to an increment of the severity of waterlogging events worldwide, especially the low-lying coastal cities (IPCC, 2011; Barros et al., 2012; Hallegatte et al., 2013; Woodruff et al., 2013).

China has suffered extensively from the environmental impacts of rapid urbanization and climate change. The record from the Ministry of Housing and Urban-Rural Development of China (MOHURD), shows that around 62% of cities suffer from waterlogging problems, especially in mega coastal cities such as Shanghai, Guangzhou, and Shenzhen (ADB, 2011). For example, as the youngest city in China, on May 11, 2014, Shenzhen suffered the strongest rainstorm since 2008 (226 mm), resulting in 150 roads were seriously waterlogged and the economic loss was more than 11 million dollars. Urban waterlogging events caused huge economic and urban basic functions losses and even pose a serious threat to people's lives (Fahy et al., 2019; Grahm and Nyberg, 2017; Sene, 2013; Yin et al., 2011). Furthermore, coastal cities are usually highly developed areas with dense population concentrations that are a vital part of the national or global economy. The damage of waterlogging events in these cities may be severe (Gallien et al., 2014; Li et al., 2016). Therefore, how to effectively mitigate the occurrence of urban waterlogging is a major problem in the process of sustainable urbanization in China.

In recent years, the frequent occurrence of urban waterlogging events has attracted wide attention from the academic community (Huang et al., 2018; Tang et al., 2018; Yu et al., 2018). Considerable research has shown that besides environmental factors (precipitation and topography), most of the urban waterlogging events



are caused by anthropogenic factors from human activities (Tehrany et al., 2019; Wang et al., 2017; Li et al., 2015; Zhang et al., 2017). It can be summarized in the following three aspects: (1) the urban micro-topography; (2) the disorderly expansion of impervious surfaces; (3) the urban drainage facilities. (1) Urban lowlands and such as underground parking lots and tunnels are conducive to the accumulation of rainwater. Thus, these areas are often suffering from urban waterlogging issues. (2) Land use/land cover change caused by urbanization greatly promotes the formation of urban waterlogging events (Huong and Pathirana, 2013; Shuster et al., 2005; Wu and Zhang, 2017). Among various land cover types within the city, the urban green spaces such as forest, grassland, and wetland that can effectively reduce the surface runoff are replaced by artificial impervious surfaces (cement, asphalt), which greatly changed the hydrological conditions. (3) During the rapid urbanization, most of the river channels or ditches were artificially landfilled, resulting in degradation of the drainage function of the river network and increased vulnerability of the drainage system. In addition, most of the rainwater drainage networks in developing countries are low design standards and have insufficient maintenance, which makes it difficult to play an active role in the face of heavy rainstorms. Therefore, facing the increasingly serious problem of urban waterlogging, scholars have proposed measures such as “sponge city” or “low impact development approach” to mitigate urban waterlogging inundation (Ahammed, 2017; Dong et al., 2018; Huang et al., 2017; Miao et al., 2019; Shao et al., 2016; Wu et al., 2018; Zimmer et al., 2007). The main purpose of these measures is to increase the proportion of permeable surfaces (urban green spaces) within the city so that the surface runoff can be reduced effectively.

Previous studies found that topographic factors, such as elevation and slope, had a strong impact on urban waterlogging (Tehrany et al., 2019; Wang et al., 2012; Zhao et al., 2018). In some cases, however, the results are contradictory. For example, the flat and low-lying surface will increase the risk of urban waterlogging in studies conducted in Huizhou, China (Wang et al., 2015) and Shenzhen, China (Wu and Zhang, 2017), but is associated with reducing the risk of urban waterlogging in Amsterdam, Netherlands (Gaitan et al., 2015). This inconsistency prevents the development of urban waterlogging research. Except for urban topographic factors, considerable studies revealed the impact of land-use change on urban waterlogging (Boudou et al., 2016; Pijl et al., 2018; Schuch et al., 2017; Yao et al., 2017; Zope and Eldho, 2016). Zhang (2018) simulated the impact of urbanization on the urban storm runoff coefficient in Beijing. Chen (2015) took Dongguan city as a case study and analyzed the impact of rapid urbanization on urban waterlogging by the DPSIR hydrological model. These results all confirmed that the impervious surface has a certain impact on urban waterlogging. These studies, however, just only focused on one land cover type - the impervious surface

alone, rather than the overall land cover characteristics (composition and spatial configuration). The quantitative relationship between overall land cover characteristics (i.e. land cover composition and spatial pattern) and urban waterlogging may not be fully understood.

Urban waterlogging is a complex phenomenon, which is the result of the joint action of natural conditions and human activities. However, what is less known, is the relative contribution of each environmental and anthropogenic factor. The influencing factors such as topographic factors, land cover composition, land cover spatial configuration, drainage facilities, and urban morphology are not comprehensively considered, which leads to some biases. Furthermore, the single-factor regression model widely used in previous studies cannot fully reveal the complex mechanisms of urban waterlogging. Simultaneously, most studies have not carried out a multi-scale analysis, and thus ignoring the scale effects of various influencing factors (Yu et al., 2018; Wang et al., 2015). Due to the great heterogeneity in highly urbanized areas, the analysis of one single scale may not be sufficient. Does the influencing factor of urban waterlogging have different effects under different analysis scales? Does the explanatory power of the same influence factor change with the change of analysis scale? Does the dominant factor determining waterlogging vary across analysis scales? Answering these questions is vital for expanding our scientific understanding of the linkage between urban waterlogging and its environmental and anthropogenic influencing factors.

In this work, we overcome this lack of investigation, presenting a case study of the highly urbanized coastal city – the central urban districts of Guangzhou (P.R. China), where the waterlogging events frequently occur. The goal is to explore complex mechanisms of urban waterlogging through multi-scale analysis to clarify the relative contributions of environmental and anthropogenic factors. In which, a novel method that integrates the stepwise regression model with hierarchical partitioning analysis is presented to quantify the complex relationship between urban waterlogging and influencing factors under various analysis scales. We considered urban waterlogging records from 2009 to 2015, and performed the following steps: (1) explore the urban waterlogging spatial pattern and assess the agglomeration effect of urban waterlogging events; (2) reveal the complex relationship between urban waterlogging and environmental and anthropogenic factors in a highly urbanized low-lying coastal city to clarify the relative contributions of influencing factors, and thus identifying the dominant factors; (3) examine whether the dominant factors vary across analysis scales and identify the scale effect. This study provides a theoretical and practical reference for the government and urban planners for urban waterlogging prevention and management.

## 2.3 Study area

Guangzhou City, the selected study area with a total terrestrial area of about 7434.40 km<sup>2</sup>, is located in the south-east part of Guangdong Province, P. R. China, at 112°57'~114°3'E and 22°26'~23°56'N, where is the intersection of the three major tributaries of Pearl River (Fig.2-1). Guangzhou is the political, economic, cultural, and scientific center in southern China, which is known as the “south gate” of China. Guangzhou is located in the subtropical coastal area, which belongs to a marine subtropical monsoon climate with an average annual precipitation of 1623~1899 mm and an average annual rainfall of 189 days (Guangzhou Meteorological Service). The low-lying mountain and hills are located in the northeastern part of Guangzhou, while the southern part is the relatively flat alluvial plain, with an average elevation of 6.6 m asl.

Since the reform and “opening up” policy in the 20th century, Guangzhou has experienced rapid urbanization and urban expansion. The high-density building area blocks the surface water seepage and interrupts the hydrological cycle, thus resulting in the frequency of urban waterlogging events. According to the World Bank's assessment report, Guangzhou ranks first in flood risk among 138 major coastal cities around the world (Hallegatte et al., 2013). In addition, the official statistics provided by Guangzhou Water Authority show that the urban waterlogging event has become a growing problem in the central urban districts of Guangzhou. Therefore, we select the region of the central urban districts of Guangzhou (Liwan, Yuexiu, Tianhe, Haizhu, Baiyun, and Huangpu District) as a study site. In view of the serious risk of urban waterlogging events in this area, it is of certain representativeness and practical significance to investigate the essential conditioning factors of urban waterlogging.

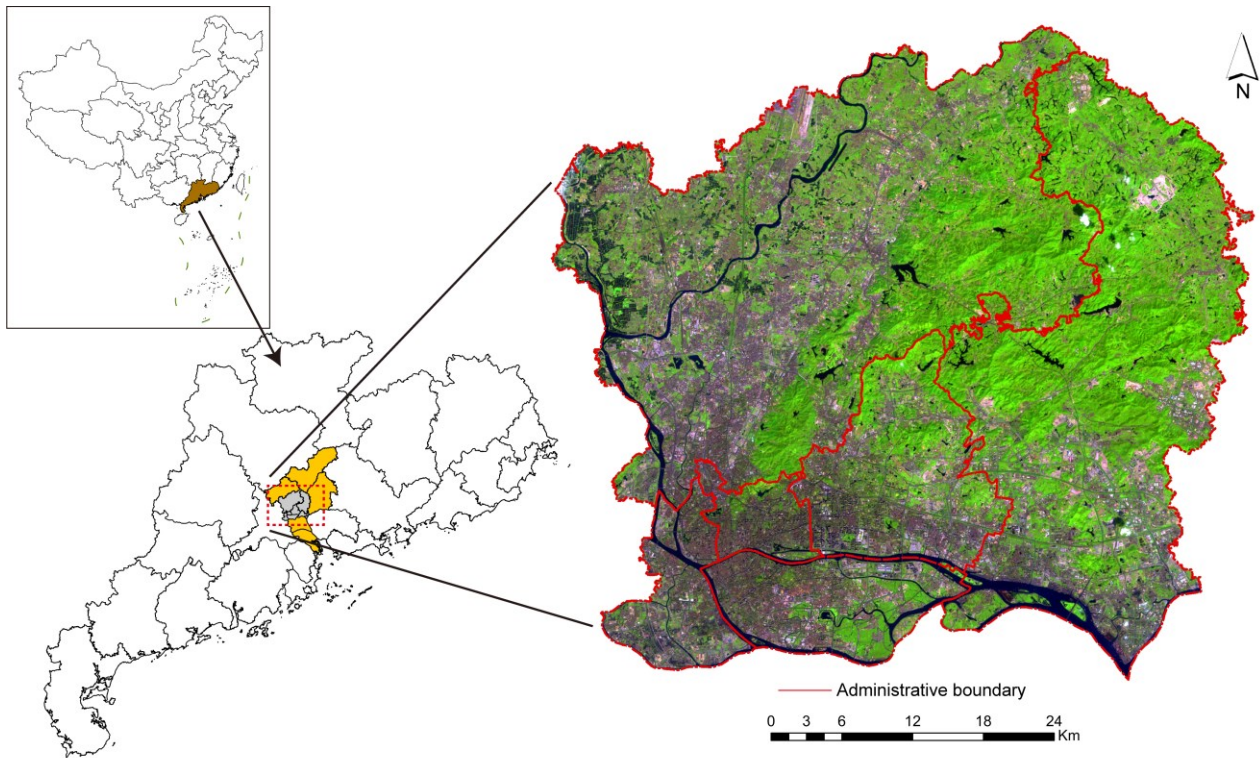


Figure 2-1 The geographic location of Guangzhou Central Urban Districts.

## 2.4 Data and methods

### 2.4.1 Data sources

In this study, we obtained the urban waterlogging records in the Central Urban Districts of Guangzhou from the flood and drought disaster prevention center of Guangzhou Water Authority (Table 2-1). However, the data from the local water authorities only record the address of the waterlogging events, and there is no accurate coordinate. Therefore, we use Google Earth to map the spatial location of these waterlogging events, such as roads intersection, entrance or exit of tunnels, tollgate of highways, and so on. If the location of the waterlogging event is in some part of the road (it is difficult to determine the specific spatial location), we refer to the method recommended by Yu et al. (2018) and Zhang et al. (2018) to regard the geometric center of the road as the spatial location where the urban waterlogging events occur. The Georeferencing tool in ArcGIS Pro is utilized to calibrate the spatial location of waterlogging events with an error of less than 50 meters. Therefore data collection has enough accuracy for assessing a suitable analysis of this work. Finally, we collected for the Central Urban Districts of Guangzhou of 423 urban waterlogging points with a depth of more than 15 cm (Fig.2-2a).

In addition, the digital elevation model (DEM), high-resolution UAV image, urban drainage network, and

building geospatial datasets were utilized as a basis data for exploring the impact of topography, land cover features, drainage facilities, and urban morphology factors on urban waterlogging (Table 2-1).

*Table 2-1 Overview of data used in this study.*

Data	Format	Time	Source
Guangzhou administrative divisions	Esri shapefile	2015	Land Resource Technology Center of Guangdong Province
Digital Elevation Model	Image	2012	Guangzhou Planning and Natural Resources Bureau
1:2000 UAV image	TIFF	2012	
Urban drainage network	Esri shapefile	2012	
Building geospatial datasets	Esri shapefile	2012	
Urban waterlogging records	Text	2009~2015	Guangzhou Water Authority

#### 2.4.2 Determination of analysis scale

Understanding the landscape structure in spatial heterogeneity context requires multi-scale information (Wu, 2004). The concept of scale effect comes from landscape ecology, that is, landscape element has different performances on different spatial and temporal scales. The scale effect is one of the important factors leading to the complication of landscape phenomena (Li et al., 2004; Guo et al., 2012). Single-scale analysis may only provide partial information on the landscape characteristics. Therefore, considering the scale effect can help us gain a comprehensive understanding of how environmental and anthropogenic conditioning factors affect urban waterlogging, which reveals the complex relationship between waterlogging and conditioning factors through appropriate spatial scales.

In this study, we first calculated the average nearest Euclidean distance (626 m) of each waterlogging point in the study area. The minimum analysis scales should be larger than the mean nearest distance of urban waterlogging points (626 m). Thus, in this study, we selected the 1 km grid as the minimum analysis scale. When the analysis scale increased to 5 km, only 37 grid units had waterlogging points, which meant that only 37 data sets could be used to explore the mechanisms of urban waterlogging. If the analysis scale is further

increased, the available data set might be insufficient. Based on these criteria, we set up a series of grid units of different sizes as the analysis scales (i.e. 1km, 2km, 3km, 4km, 5km), respectively (Fig.2-2 b-f). These grid units were used to represent different analysis scales and overlay waterlogging points with different grid sizes. The frequency of waterlogging events is expressed by calculating the number of urban waterlogging points in different grid units.

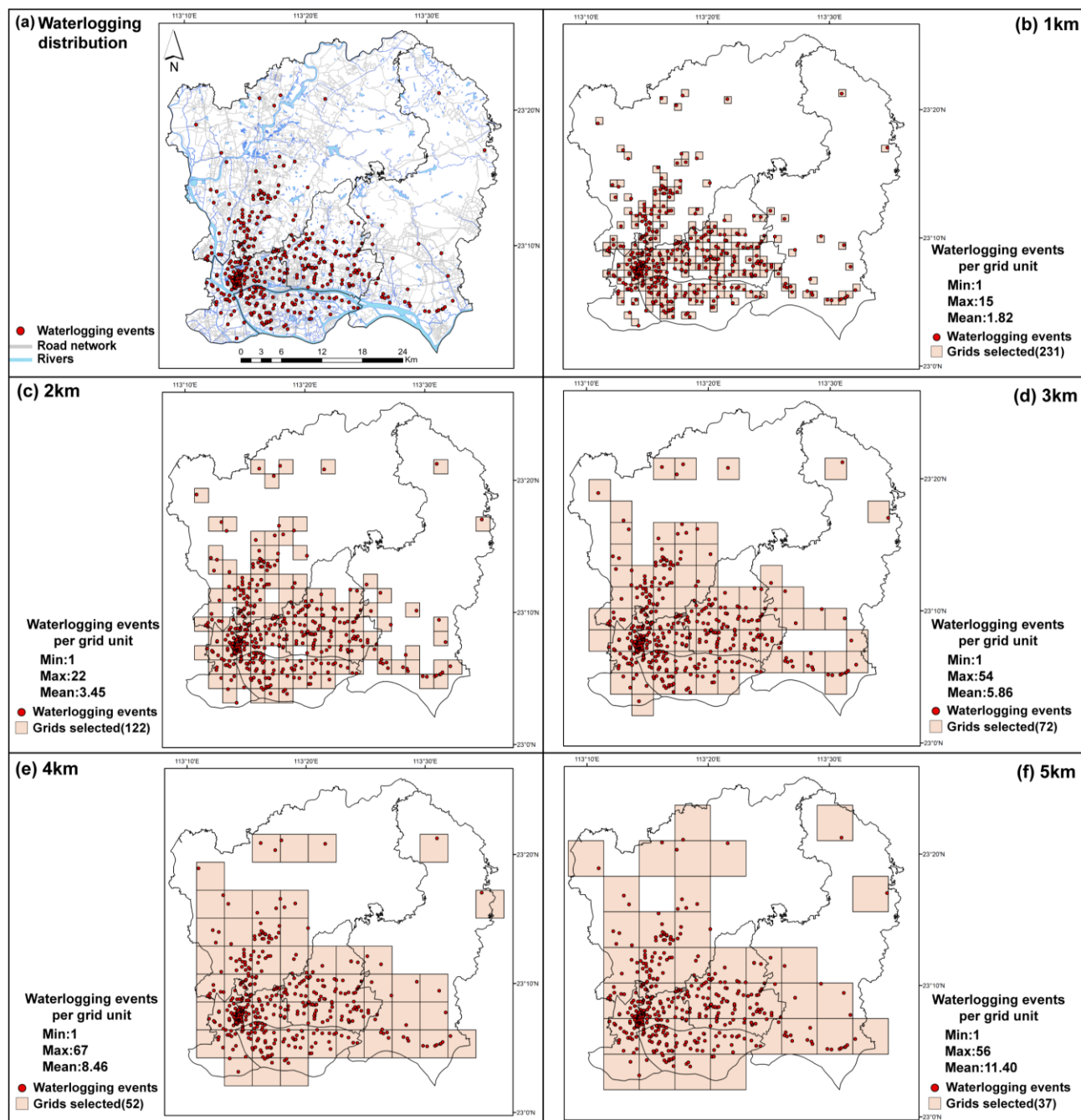


Figure 2-2 The spatial distribution of urban waterlogging events (a) and five different grid scales of 1 km, 2 km, 3 km, 4 km, and 5 km (b-f)

### 2.4.3 The spatial distribution pattern of waterlogging events

In this study, the kernel density and spatial auto-correlation analysis tools were utilized to explore the spatial pattern of urban waterlogging events. The spatial autocorrelation analysis tool measures spatial autocorrelation whether the elements are clustered, discrete, or random according to the element locations and values. Spatial autocorrelation analysis includes two subordinate tools: global spatial correlation and local spatial correlation (Anselin et al., 2006). In this study, the spatial agglomeration patterns of urban waterlogging events in different scales are identified by local spatial correlation (local Moran's I index), which is presented as follows:

$$I = n \times \frac{(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} \times \sum w_{ij} (x_i - \bar{x}) \quad (2.1)$$

where  $w_{ij}$  is the spatial weight matrix;  $x_i$  is the observation  $x$ ;  $\bar{x}$  is the average of all variables and sets the significance level as 0.01. In this study, the spatial agglomeration patterns of urban waterlogging events in different spatial scales were calculated from GeoDa software and shown in Figure 2-6. There were four main cluster types, the specific meanings were as follows:

High-high agglomeration (hot spot): The number of urban waterlogging events in the grid unit and its adjacent grids are significantly higher than the average level, indicating that urban waterlogging events are concentrated in a place.

High-low agglomeration (island): The number of waterlogging events in the cell grid is much higher than the average level, while the waterlogging events in its surrounding grids are lower than the average.

Low-high agglomeration (atoll): The number of waterlogging events in the cell grid is significantly lower than the average level whereas waterlogging events in its adjacent grids are relatively high.

Low-low agglomeration (clod spots): The number of urban waterlogging events and its surrounding grids are relatively lower than the average.

Besides, "non-significance" means that the grid has no significant local spatial correlation association with surrounding surroundings grids.

### 2.4.4 Environmental and anthropogenic factors

In this section, we mainly introduce 27 influencing factors that were used as the independent variables. The influencing factors were categorized as four dimensions: urban topography, land cover characteristics

(composition and spatial configuration), urban drainage facilities, and urban morphology.

#### 2.4.4.1 Topographic variables

The topographic variables of elevation, slope, relative elevation (RE), slope standard deviation (slope.std), curvature were calculated from ArcGIS Pro, while the SAGA GIS software was used to calculate the topographic roughness index (TRI), topographic surface texture index (TSI), and topographic wetness index (TWI) (Fig.2-3).

In most cases, urban waterlogging events occur in the region with low elevation and relatively flat terrain. Since these areas are easy to obtain runoff from the surrounding area and the surface runoff is difficult to discharge in a short time. Therefore, in this study, we selected the elevation and slope indicators to represent the altitude and the flatness of each grid unit. The relative elevation (RE) represents the degree of elevation variation in each grid unit and is calculated as follows:

$$RE = E_{max} - E_{min} \quad (2.2)$$

where  $E_{max}$  represents the maximum elevation in the grid unit  $i$  and  $E_{min}$  represent the minimum elevation in the grid unit  $i$ . When the RE value is large, it means that the elevation in the grid fluctuates greatly. On the contrary, the elevation in the grid remains relatively flat. The standard deviation of slope was calculated to further measure the variation of slope in each grid. If the slope.std value is large, indicating that the terrain in the grid unit fluctuates greatly. Instead, it means that the slope in the grid is relatively gentle. Curvature is another important variable that represents the surface concavity and convexity, which is calculated by the second derivative of the eight adjacent grids (3\*3 pixels).

Topographic indexes such as TRI, TSI, and TWI also have a considerable impact on revealing the formation of urban waterlogging. Topographic Roughness Index (TRI) is one of the morphological factors closely related to flooding, which was developed by Riley, et al. (1999). It is used to quantify the elevation difference between adjacent grids (Werner et al., 2005). The calculation formula is as follows:

$$TRI = Y \left[ \sum (x_{ij} - x_{00})^2 \right]^{1/2} \quad (2.3)$$

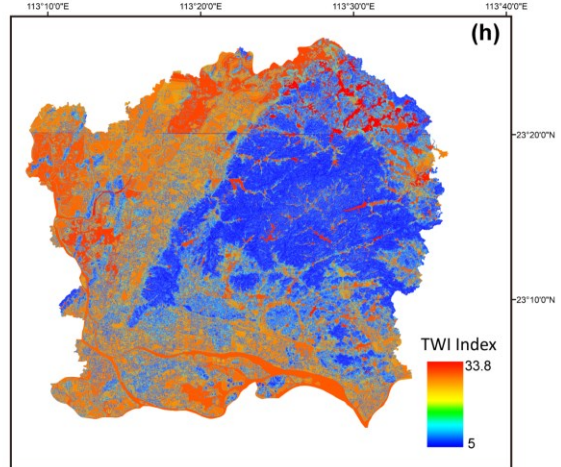
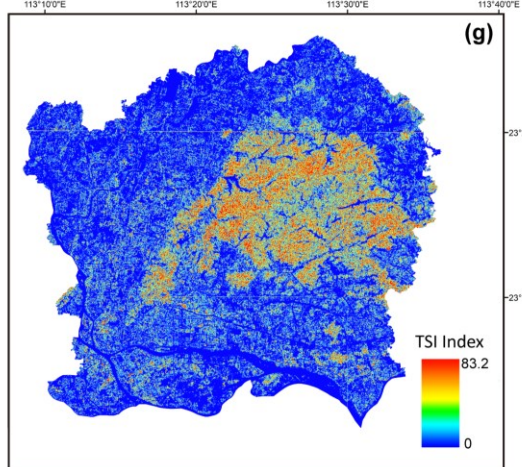
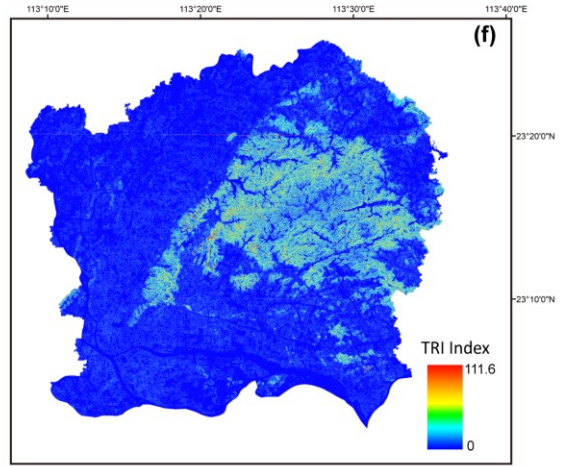
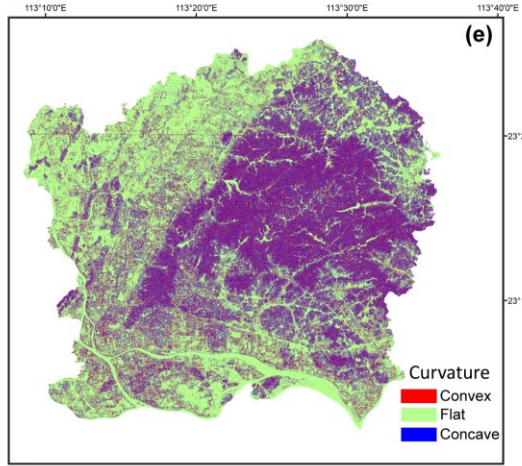
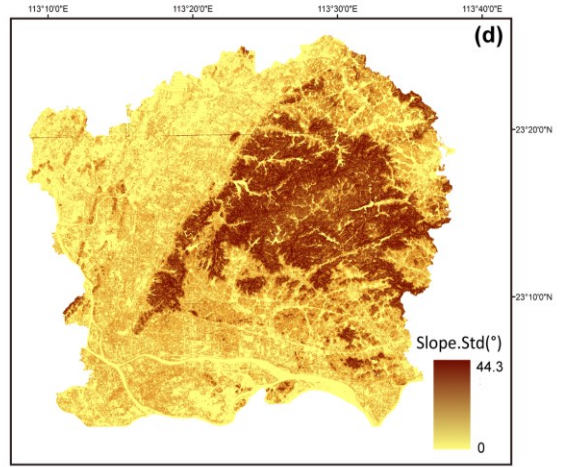
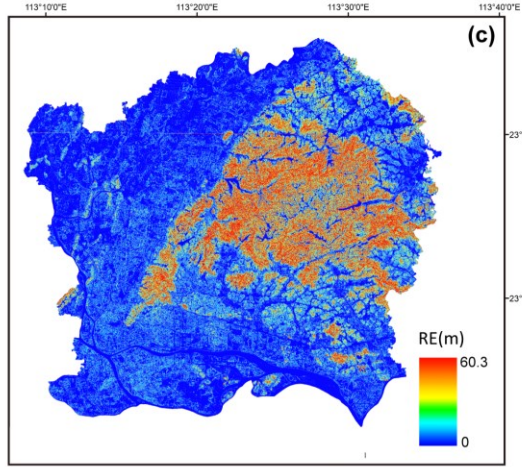
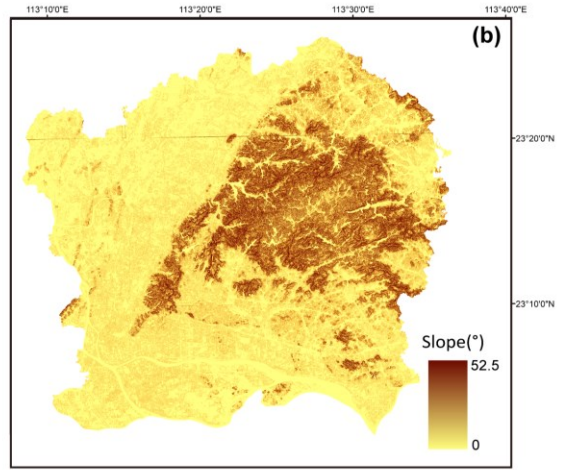
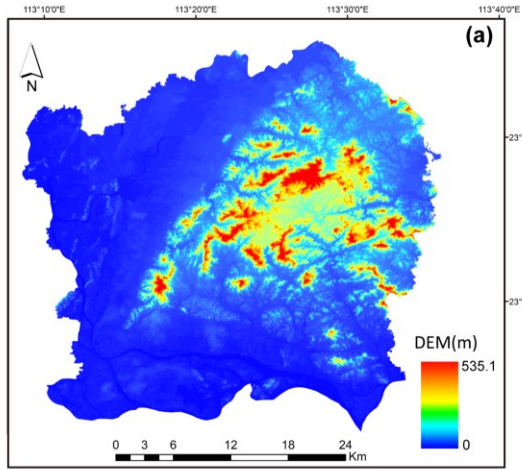
Terrain Surface Texture Index (TSI) is another topographic morphological measurement index, which uses the spatial frequencies of peaks and deep pits to represent the topographic surface texture (Iwahashi and Pike, 2007). The TSI value of each grid represents the frequency of pits and peaks within 10 adjacent grids.



$$TSI = f_{ij} / \sum N_{ij} \quad (2.4)$$

Topographic Wetness Index (TWI) is commonly used to quantify the effect of topography on hydrological processes and it is one of the most important factors for surface runoff generation. The TWI index is widely used to simulate hydrological flow paths and identify flood-prone areas (Lee et al., 2017; Tien Bui et al., 2016). The high value of TWI means the high potential of runoff generation, while the low value means the low potential of runoff generation.

$$TWI = \ln(A_s / \tan \beta) \quad (2.5)$$



*Figure 2-3 The topographic factors of elevation (a), slope (b), relative elevation (c), slope standard deviation (d), curvature (e), TRI (f), TSI (g), TWI (h)*

#### 2.4.4.2 The land cover characteristics

The land cover data is classified from 1:2000 aerial remote sensing images, which divided the land cover types into nine categories: (1) woodland (including forest, plantations, and street trees); (2) grassland (including natural grassland and artificial grassland); (3) garden (including orchard, tea plantation, and other plantation); (4) cultivated land (including vegetable field and paddy field); (5) residential area (including house and building); (6) road network (including street and pavement); (7) construction land (including square, parking lot, and factory); (8) bare land (including soil surface, rock surface, and unused land); and (9) water body (including lake, river, and reservoir) (Fig.2-4). In addition, in order to analyze the impact of different land cover types on urban waterlogging at a macro level, this study establishes another two land cover types: impervious surfaces and urban green spaces. The impervious surfaces are composed of the residential area, road network, and construction land; whereas the woodland, grassland, garden, and cultivated land are classified as urban green spaces.

Landscape pattern metrics can highly condense the information of various landscape features (McGarigal and Marks, 1995). Numerous researchers had developed many landscape pattern metrics to measure the land cover spatial configuration (Gardner et al., 2001; Cushman et al., 2008). Considering the meaning of each landscape pattern metrics and referring to the existing research, six landscape pattern metrics were selected in this study, including: patch density (PD), largest patch index (LPI), landscape shape index (LSI), connectivity index (Cohesion), aggregation index (AI), and landscape richness index (PR). All metrics are specified in Table 2-2, and the calculation process is performed in Fragstats 4.2.

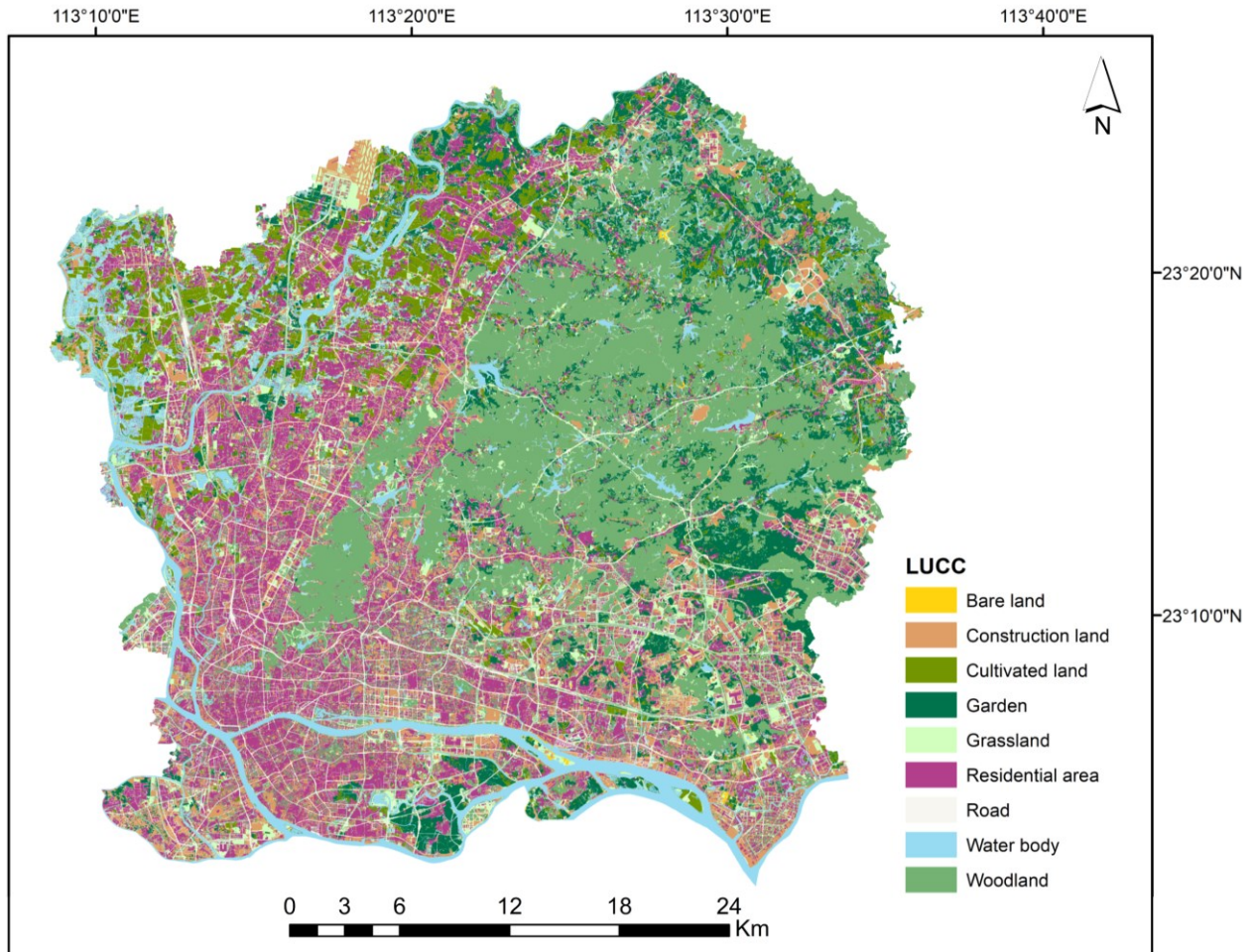


Figure 2-4 The Land-use / land cover map for the study area.

Table 2-2 The selected landscape pattern metrics.

Categories	Landscape pattern metrics	Meaning
Degree of landscape fragmentation	Patch Density (PD)	The ratio of the number of patches to the total landscape area. A larger value means higher fragmentation.
	Largest Patch Index (LPI)	The largest land cover patch within an analysis unit. A simple measure of the dominance of land cover features.
	Landscape Shape Index (LSI)	The shape index is a measure of shape complexity.
	Connectivity Index (Cohesion)	A measure of connectedness.
	Aggregation Index (AI)	Describe the degree of agglomeration or extension trend of different landscape types.

---

Landscape diversity

Landscape Richness Index (PR)

It reflects the uniformity and complexity of the distribution of different landscape types in an area.

---

#### 2.4.4.3 Urban drainage facilities and urban morphology

We calculated the density of the rainwater drainage network to measure the impact of drainage facilities on urban waterlogging (Fig.2-5). Based on the drainage network distribution, the drainage density (DD) was defined as the length of the drainage network in per analysis unit, which can be calculated as follow:

$$DD = \frac{L_i}{S} \quad (2.6)$$

where  $S$  is the total area of each unit ( $\text{km}^2$ ) and  $L_i$  represents the total length of the drainage network (km) in each unit.

Urban morphology refers to the spatial structure and physical characteristics of cities over time, which is the result of various activities of human beings (Chen, 2014). In this study, the urban morphology factor was introduced to verify whether it had a certain impact on urban waterlogging. We used building density (BD) to quantify the urban morphology. The high-resolution building geospatial datasets contain building footprints described by ground-level boundaries. Based on this, we could calculate the building area in each grid unit, and thus, the building density can be calculated as the total area of buildings per unit area (Eq.2.7).

$$BD = \frac{S_B}{S} \quad (2.7)$$

where  $S$  represents the total area of each grid unit ( $\text{km}^2$ ) and  $S_B$  represents the building's area in each grid unit ( $\text{km}^2$ )

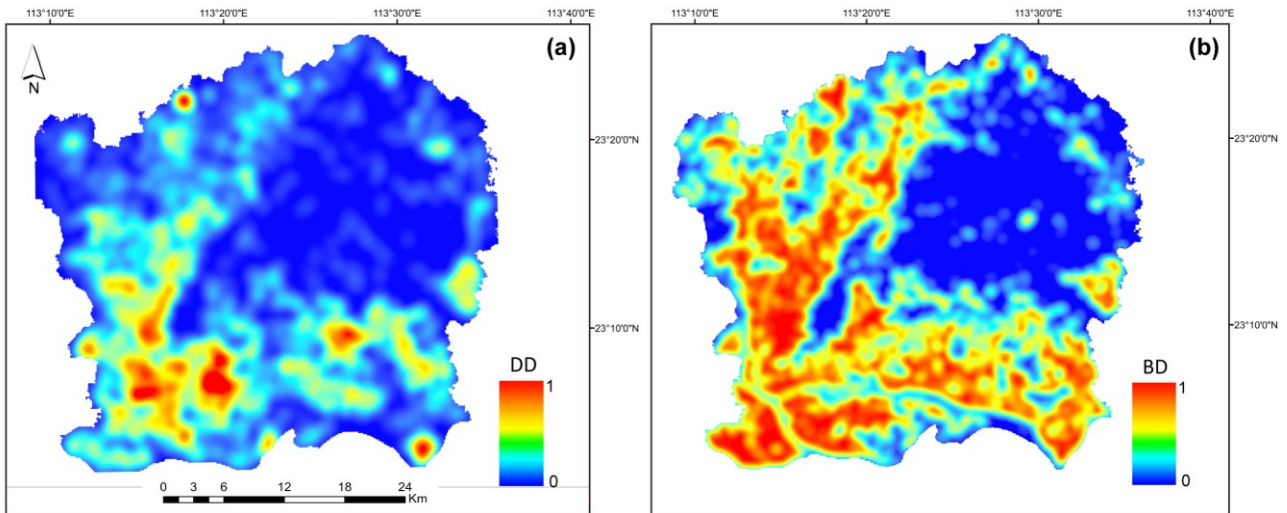


Figure 2-5 The drainage density (a) and building density(b).

#### 2.4.5 Statistical analyses

In this study, we took the number of urban waterlogging points in each grid unit as dependent variables, while considering the topographic variables, the land cover characteristics (composition and spatial configuration), drainage density, and building density as independent variables (Table 2-3). The effects of various influencing factors on urban waterlogging events were quantitatively measured by establishing multiple analysis scales. The Pearson correlation was firstly used to reveal the binary correlation of each factor across all analysis scales. Through this method, the significant explanatory variables related to the waterlogging events (confidence interval  $\geq 95\%$ ) were found out. Furthermore, the robustness of the relationship between the influencing factors and urban waterlogging was further verified through multiscale analysis. Secondly, under the combined effect of multiple factors, the stepwise regression model was used to avoid the multicollinearity problem and confirmed the significant explanatory variables (Chen et al., 2014). The bidirectional elimination stepwise regression can gain a comprehensive explanation rate of the variables entering the model, and thus exclude the insignificant variables. Finally, the significant variables that remained from the stepwise regression model will eventually be used in hierarchical partitioning analysis. The hierarchical partitioning method was utilized to clarify which factor had the dominant effect on waterlogging, which investigated the relative contributions of each factor for the entire hierarchy of models using all combinations of variables that were retained from the stepwise regression model. If  $n$  variables remained from the stepwise regression model, the established all possible combinations  $2^n$  regression models were employed in the hierarchical partitioning method (Table S2-1). The combination of the stepwise regression model and the hierarchical partitioning method is more suitable for multivariate data analysis

than traditional regression analysis methods, considering the heterogeneity of the urban environment. The hierarchical partitioning method was implemented in the “hier.part” package and “gtools” package of R statistical software (R Core Development Team, 2008; Nally, 2000).

*Table 2-3 The classification and description of the independent variables*

Categories of variables		Independent variables
Urban topography	Topography indices	Elevation, relative elevation, slope, slope.std, and curvature
	Hydrology indices	TRI, TSI, and TWI
The land covers composition	Macro perspective	Impervious surface, urban green spaces
	Micro perspective	Woodland, grassland, garden, cultivated land, residential area, road network, construction land, bare land, and water body
The land covers spatial configuration	Degree of landscape fragmentation	Patch density (PD), largest patch index (LPI), landscape shape index (LSI), connectivity index (Cohesion), and aggregation index (AI)
	Landscape diversity	Landscape richness index (PR)
Urban drainage facilities		Drainage density (DD)
Urban morphology		Building density (BD)

## 2.5 Result

### 2.5.1 Spatial pattern of the urban waterlogging events

The kernel density showed that the urban waterlogging events in Guangzhou had a significant spatial agglomeration effect (Figure 2-2). During the period from 2009 to 2015, the high-density area of urban waterlogging events (highlighted as yellow and red) mainly concentrated in the historic urban districts of Guangzhou, such as Liwan, Yuexiu, and Haizhu. In contrast, the low-density areas of urban waterlogging events were mainly concentrated in the northern and eastern mountainous areas, such as Baiyun District and Huangpu District.

Due to the spatial pattern and landscape effect are also affected by the scale effect, so different results may appear at different analysis scales. The local Moran's I index was calculated for determining the multiscale spatial agglomeration pattern of urban waterlogging events (Table 2-4, Fig.2-6). Under all analysis scales, the

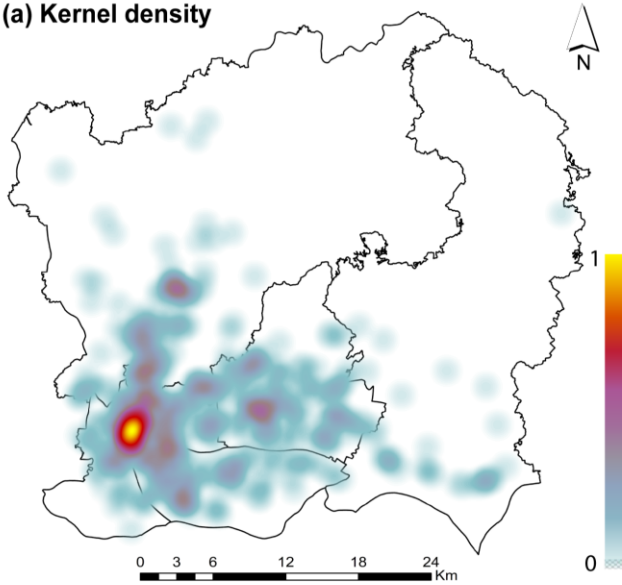
Z-score or P-value indicated a significant clustering effect, which meant the waterlogging hot spots were obvious. The hot spots (high-high concentration) of urban waterlogging events were mainly distributed in the historic urban districts of Guangzhou (Liwan, Yuexiu, and Haizhu district), while the low-high agglomeration grids concentrated around the waterlogging hot spots. The high-low agglomeration grids are sparsely distributed in the low concentration area of urban waterlogging events (Baiyun district and Huangpu district). Although the spatial agglomeration pattern of waterlogging events was similar, the local Moran's I results were slightly different at all analysis scales. The Moran's I index obtained the highest value when the analysis scale is 3 km, which indicated the strongest clustering trend.

*Table 2-4 Moran's I index of waterlogging events under different analysis scales.*

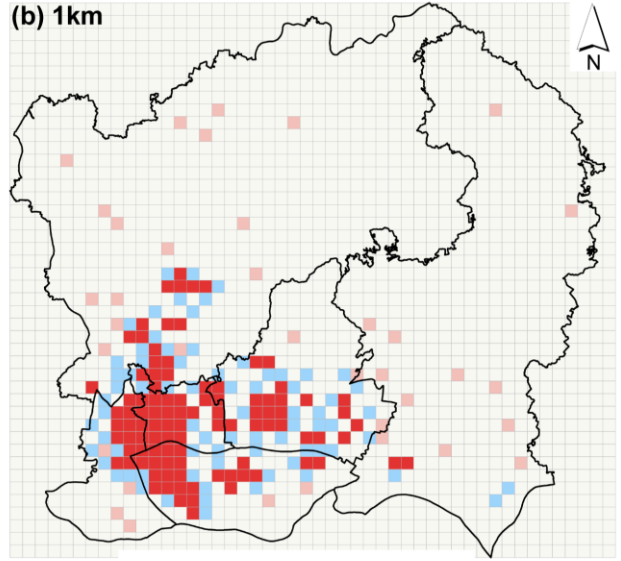
Analysis scales	Moran's I index	Z-test value	P-value
1 km	0.5203	34.3346	0.0000
2 km	0.5952	14.5526	0.0000
3 km	0.7204	10.0144	0.0000
4 km	0.6977	22.9492	0.0000
5 km	0.6058	11.0215	0.0000



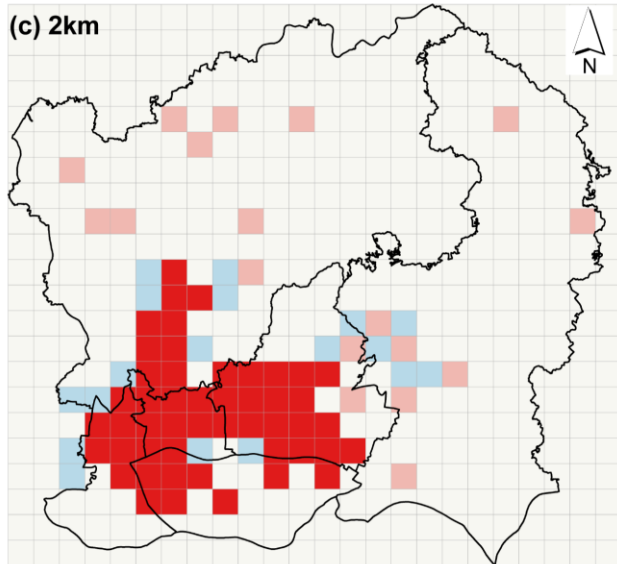
(a) Kernel density



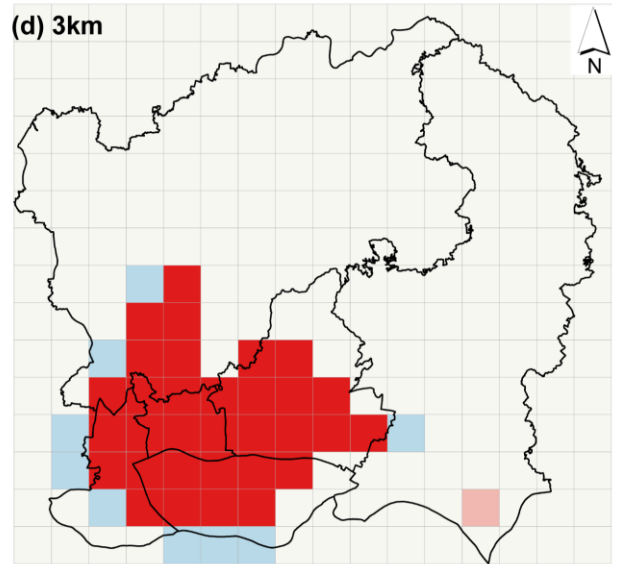
(b) 1km



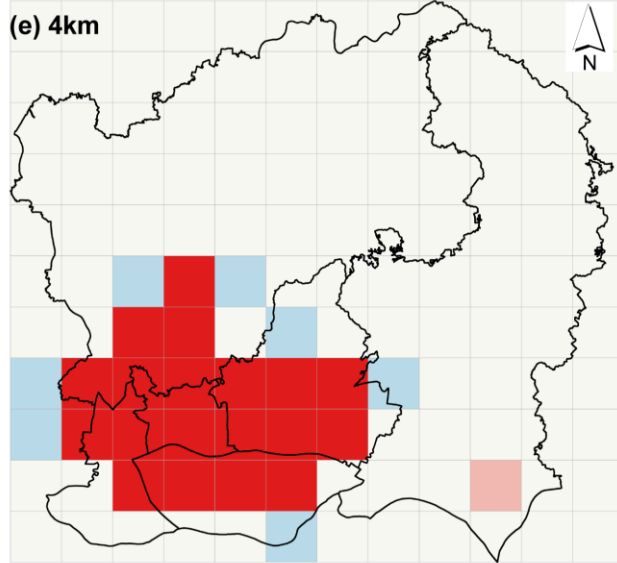
(c) 2km



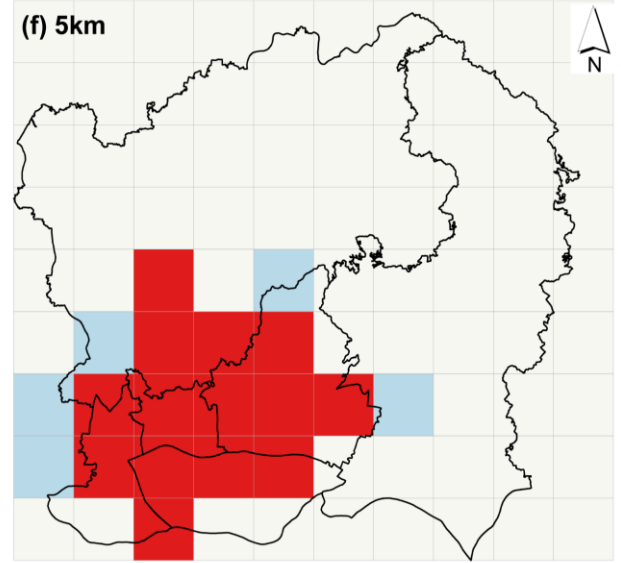
(d) 3km



(e) 4km



(f) 5km



High-high agglomeration      Low-high agglomeration  
High-low agglomeration      Administrative boundary

Figure 2-6 The kernel density of urban waterlogging events (a) and the spatial agglomeration map across different analysis scales (b~f).

For all analysis scales, we noticed that the hot spots of waterlogging events were concentrated in the historic urban area of Guangzhou, presenting a robust spatial pattern (Fig.2-6 b~f). The waterlogging hot-spot mainly covers Xiguan (Zhongshan 7th Road, Hualin Street), Yuexiu, Tianhe Shipai Bridge, South China Normal University, and Jinan University (Fig.2-7). These areas are frequent occurrences of urban waterlogging events, affecting people's normal production and life.

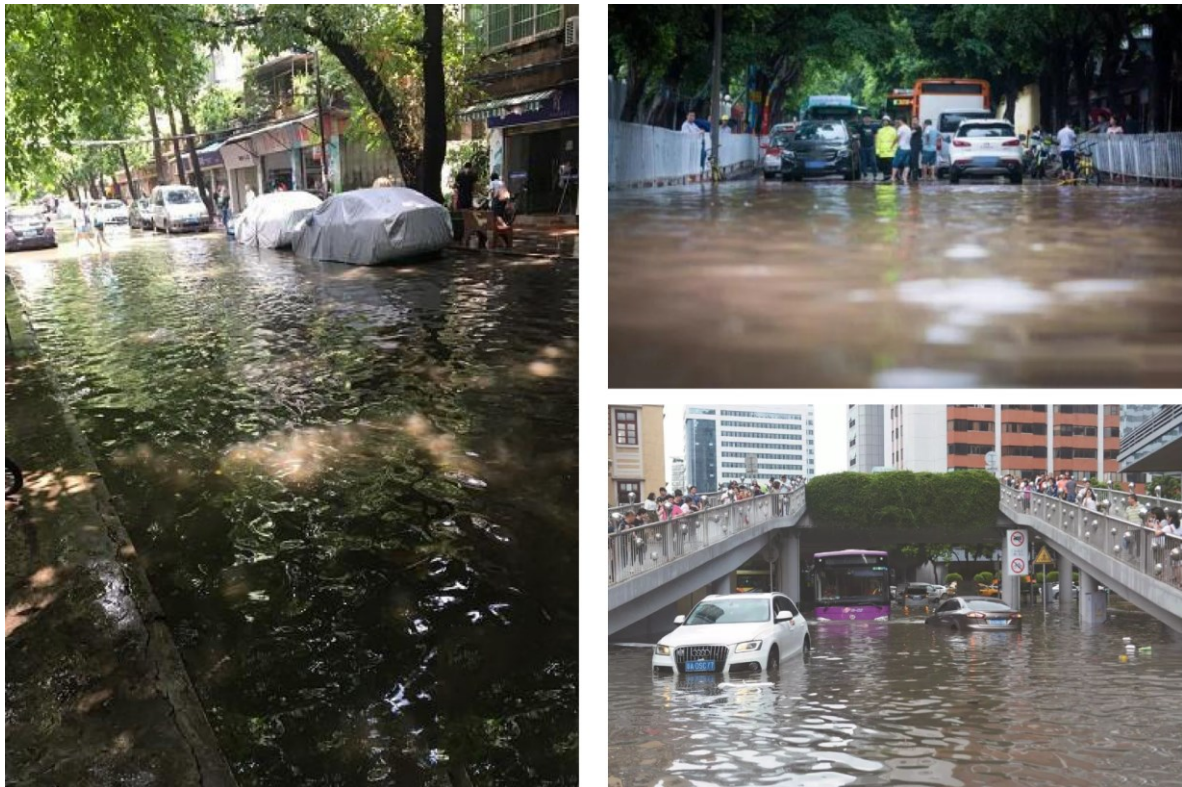


Figure 2-7 The typical urban waterlogging events in Guangzhou (from Guangzhou Daily).

## 2.5.2 Correlations between influencing factors and urban waterlogging at multiple scales

### 2.5.2.1 Effects of the topographic factors on urban waterlogging

The correlation coefficient indicated that the elevation, RE, slope, slope.std, curvature, TRI, and TSI index were all negatively correlated with urban waterlogging events ( $p < 0.05$ ). This indicated that the risk of urban waterlogging was more likely to be reduced where the elevation and slope within the grid unit fluctuate greatly. However, the TWI index had a significant positive correlation with urban waterlogging ( $p < 0.05$ ). The high value of the TWI index meant the greater the potential for accumulated surface runoff, in which the urban waterlogging events were more likely to occur.

When the analysis scale increased from 1 km to 5 km, we found that the elevation factors and TWI presented a contradictory trend. The correlation directions (positive/negative) change with the analysis scale (1 km - 5 km), which fully reflects the scale effect between topographic factors and urban waterlogging. Furthermore, we noticed that in a small analysis scale, the topography variables had a better explanatory power on urban waterlogging, while the research scale increased to 5 km, the correlation coefficient of topographic factors decreased gradually. This may be due to the expansion of the analysis scales, the grid unit covered other influence factors (such as land cover types and landscape pattern) and redundant information, which concealed the interpretability of topographic factors.

*Table 2-5 Pearson correlation coefficients between waterlogging and topographic factors across five analysis scales. (blue color represents a negative correlation, red color represents a positive correlation, and the deeper the color, the greater the correlation coefficient. “\*\*” and “\*\*\*” represent correlation coefficient significance at the 0.05 level and 0.01 level, respectively.)*

Topography factors	Analysis scales				
	1 km	2 km	3 km	4 km	5 km
Elevation	-0.309*	-0.182*	-0.166	-0.097	0.035
RE	-0.587**	-0.503**	-0.213*	-0.193*	-0.184*
Slope	-0.466**	-0.382**	-0.199*	-0.297**	-0.173
Slope.std	-0.628**	-0.321**	-0.356**	-0.209*	-0.153
Curvature	-0.453**	-0.450**	-0.202*	-0.225*	-0.206*
TRI	-0.676**	-0.574**	-0.216*	-0.363**	-0.184*
TSI	-0.429**	-0.349**	-0.308**	-0.248*	-0.267*
TWI	0.435**	0.402**	0.253*	0.273**	-0.187*

### 2.5.2.2 Effects of the land cover composition on urban waterlogging

As can be seen from Table 2-6, the percent cover of urban green spaces (garden, grassland, and woodland) experienced a significant negative correlation with urban waterlogging ( $p < 0.01$ ), while the impervious surfaces (residential area and construction land) presented a significant positive correlation ( $p < 0.01$ ). Results indicated that the larger proportion of impervious surfaces within the grid unit, the more serious is the urban waterlogging. On the contrary, with the increase in the proportion of green spaces, it can effectively reduce the occurrence of urban waterlogging. Moreover, compared with the correlation coefficients of each land cover type, the absolute values of the correlation coefficients of the residential area, garden, grassland, and construction land were larger than other variables under all analysis scales. This means that it was more

effective to reduce the risk of urban waterlogging by properly regulating residential areas or increasing the area of woodland and grassland.

We noted that the correlation directions of construction land and bare land change with the analysis scale. The construction land in a 1 km analysis scale showed a slightly negative correlation with urban waterlogging, while a significant positive correlation was found when the analysis scale was over 1 km. Similarly, the bare land was negatively correlated with urban waterlogging, while presenting a positive correlation at 2 km analysis scale. One of the possible reasons might be the impact of urban topography on urban waterlogging was more obvious in a smaller analysis scale. Furthermore, it was worth noting the absolute values of the correlation coefficients of land cover composition variables in small grid units (1 km up to 2 km) were generally smaller than the large analysis scale (3 km to 5 km). This indicated that there was also a significant scale effect on the relationship between land cover composition and urban waterlogging. The slightly declining trend of correlation coefficients in the 5 km analysis scale may be due to the large analysis scale covering too much redundant information.

*Table 2-6 Pearson correlation coefficients between waterlogging and land cover composition across five analysis scales. (blue color represents a negative correlation, red color represents a positive correlation, and the deeper the color, the greater the correlation coefficient. “\*\*” and “\*\*\*” represent correlation coefficient significance at the 0.05 level and 0.01 level, respectively.)*

Percent cover	Analysis scales				
	1 km	2 km	3 km	4 km	5 km
Impervious surfaces	0.286*	0.322*	0.647**	0.422**	0.366**
Green spaces	-0.282*	-0.452**	-0.544**	-0.385**	-0.424**
Woodland	-0.180*	-0.377**	-0.521**	-0.467**	-0.488**
Grassland	-0.381**	-0.520**	-0.566**	-0.498**	-0.463**
Garden	-0.133*	-0.463**	-0.463**	-0.499**	-0.378**
Cultivate land	-0.057	-0.023	-0.236*	-0.248*	-0.202*
Residential area	0.257*	0.563**	0.663**	0.624**	0.617**
Road network	0.118	0.448**	0.439**	0.498**	0.524**
Construction land	-0.085	0.506**	0.562**	0.511**	0.589**
Bare land	-0.341*	0.123*	-0.093	-0.216*	-0.166*
Water body	-0.093	-0.11	-0.025	-0.11	-0.108

### 2.5.2.3 Effects of the land cover spatial configuration on urban waterlogging

In most instances, the correlation coefficient of spatial configuration variables was smaller than the land

cover composition variables (Table S2-2). However, the correlation coefficients of the LPI of residential area, Cohesion of urban green spaces, PD of woodland, and AI of impervious surface were much higher than other configuration variables. This indicated that these configuration variables also had a certain impact on urban waterlogging. The LPI, Cohesion, and AI indices of green areas such as grassland, garden, and woodland, were negatively correlated with urban waterlogging ( $p < 0.05$ ), while the PD of green spaces presented a positive correlation relationship ( $p < 0.05$ ). As a comparison with urban green spaces, the LPI, Cohesion, AI and PD indices of impervious surface (residential area and construction land) showed an opposite trend. This indicated that the higher the dominance and aggregation of urban green spaces, the lower the probability of urban waterlogging; whereas the concentrated and aggregated impervious surface will increase the probability of urban waterlogging events. Furthermore, the shape of the urban green spaces and impervious surfaces also had a certain impact on urban waterlogging. The landscape richness index (PR) was significantly negatively correlated with the density of the waterlogging ( $p < 0.05$ ). This meant that the increase of landscape abundance or land cover types can effectively reduce the risk of urban waterlogging.

Under different analysis scales, we noticed that the correlation coefficients of configuration variables also have the best performance at the large analysis unit (over 2 km). This suggested that the land cover spatial configuration also experienced an obvious scale effect. However, the relationship between land cover configuration variables and urban waterlogging fluctuated greatly than land cover composition. Compared with topography and land cover composition, the land cover spatial configuration indices affected by the analysis scale were more obvious, which indicated that the land cover spatial pattern was more sensitive to scale effects.

#### 2.5.2.4 Effects of the drainage facilities and urban morphology on urban waterlogging

The impact of drainage density (DD) on waterlogging events was much lower than other conditioning factors (Table 2-7). This means that in the face of a heavy rainstorm, the density of drainage networks may not effectively mitigate the occurrence of urban waterlogging events. Additionally, in some analysis scales (2 km and 5 km), the DD was positively correlated with waterlogging events. Theoretically, increasing the drainage density will undoubtedly accelerate the rainwater drainage, thus reducing the risk of urban waterlogging. However, the positive correlation of drainage density indicated that the existing urban drainage network may no longer function in the face of heavy rainfall events. The building density (BD) also had a certain impact on waterlogging. We found that higher BD significantly leads to higher levels of waterlogging risk. This may be

related to the frequent urban waterlogging events in historical urban areas (Liwan and Haizhu District) where the building density is relatively high. In addition, the higher BD means a higher proportion of impervious surfaces, and thus the risk of urban waterlogging also increases.

*Table 2-7 Pearson correlation coefficients between waterlogging and socioeconomic factors across five analysis scales. (blue color represents a negative correlation, red color represents a positive correlation, and the deeper the color, the greater the correlation coefficient. “\*” and “\*\*” represent correlation coefficient significance at the 0.05 level and 0.01 level, respectively.)*

Analysis scales	Drainage density	Building density
1 km	-0.234*	0.168
2 km	0.062	0.264*
3 km	-0.190*	0.302**
4 km	-0.231*	0.411**
5 km	0.217	0.388**

### 2.5.3 Significant factors affecting urban waterlogging events

The influencing factors that were retained from the stepwise regression model were considered as the significant factors, which would be entered into the hierarchical partitioning analysis (Table 2-8). It was noted that the model yielded the highest explanatory power (77.2%) when the analysis scale is 3 km, indicating that the mechanism of urban waterlogging events could be well investigated. However, when the analysis scale is 1 km, the determination coefficient is relatively low. This may suggest that, under a small analysis scale, more factors should be considered to improve the model explanatory power.

Notably, we also found that the significant variables were different in each analysis scale. Generally, from 1 km to 5 km, the number of variables that remain from the stepwise regression model continues to increase, indicating the mechanisms became more complex. In 1 km unit, there were two topographic factors (slope.std and relative elevation) that remained from the stepwise regression model, while only one land cover characteristics (percent cover of impervious surface) was retained. This meant that the impact of topography factors on urban waterlogging events was greater than other influencing factors at a small analysis scale. When the analysis scale expanded to 2 km, the topography factors (RE, TRI) and the composition of land cover (percent cover of residential area and PD of woodland) survive as significant predictors simultaneously. In 3 km analysis scale, the percent cover of green spaces, the largest patch index of the residential area, the aggregation index of impervious surface, and slope.std finally remained from the

model. When the analysis scale expanded to 4 km, the land surface characteristics (composition and spatial configuration) were confirmed as the significant metrics, while none of the topographic factors were included. In 5 km scale, the retention variables were the cohesion of grassland, aggregation index of impervious surface, percent cover of grassland, and building density. In general, the topographic factors (Slope.std, RE) were the significant factors affecting urban waterlogging when the analysis scales were small. However, when the analysis scale expanded to 3 km, the land cover composition (green space, residential area) exceeded the topographic factors and became important factors in influencing urban waterlogging.

*Table 2-8 Stepwise regression results for influencing factors affecting waterlogging magnitude.*

Analysis Scales	Enter variables	Adj R <sup>2</sup>	F	Sig
1 km	Slope.std, relative elevation (RE) and percent cover of impervious surface	0.319	21.273	0.000
2 km	Relative elevation, TRI, the percent cover of residential area, and PD of woodland	0.560	19.423	0.000
3 km	Slope.std, percent cover of green spaces, LPI of the residential area, and AI of impervious surface	0.772	14.481	0.000
4 km	Percent of residential area, LPI of woodland, PD of construction land, and cohesion of urban green spaces	0.680	16.853	0.000
5 km	Percent cover of grassland, AI of impervious surface, cohesion of grassland, and building density	0.609	24.139	0.000

#### **2.5.4 Relative contributions of significant factors affecting urban waterlogging**

The relative contributions of significant factors on urban waterlogging were revealed through hierarchical partitioning analysis (Fig.2-8). The results of hierarchical partitioning analysis indicated that the slope.std had the largest independent contribution rate (36.85%) to waterlogging in 1 km analysis scale, whereas the percent of impervious surface and relative elevation contributed independently contributed 32.13% and 31.02%, respectively. When the analysis scale expanded to 2 km, the percent of the residential area became the most important driving factor for urban waterlogging, independently contributing 41.03% of waterlogging magnitude variations, whereas the relative elevation, patch density of woodland, and TRI index

contributed 27.72%, 19.65%, and 11.60%, respectively. In 3 km analysis scale, the percent cover of urban green spaces and the largest patch index of the residential area both had the dominant independent contributions to waterlogging variations (44.74% and 35.63%, respectively), while the aggregation index of impervious surface contributed 11.22% with 8.41% for slope.std contribution. When the analysis scale was 4 km, the percent cover of the residential area and largest patch index of woodland had the largest independent contribution rate (40.85% and 31.67%) to waterlogging magnitude, whereas the cohesion of urban green spaces and patch density of construction land contribute 20.30% and 7.18%, respectively. In 5 km scale, the percent cover of grassland was the most important driving factor, independently contributing 34.24% of waterlogging variations, while the aggregation index of impervious surface, cohesion of grassland, and building density contributed 29.28%, 19.36%, and 17.12%, respectively. In general, under all analysis scale, the percent cover of urban green spaces (44.74% in 3 km scale), percent cover of residential area (41.03% in 2 km and 40.85% in 4 km), and slope.std (36.85% in 1 km scale), had the largest contributions. Therefore, we could infer that urban waterlogging events were mainly affected by these three factors.

Under different analysis scales, the dominant factors varied across different analysis scales. In 1 km analysis scale, the slope.std was the dominant driving factor for waterlogging magnitude. However, when the analysis units were up to 2 km, the relative contributions of land surface characteristics (composition and spatial configuration) were much greater than others and became the dominant influencing factor of urban waterlogging. We could conclude that the topography factor driving urban waterlogging events was greater than other significant factors at a small analysis scale; whereas the land cover composition accounted for the dominant contribution rate at the macro-scales (2-5 km).



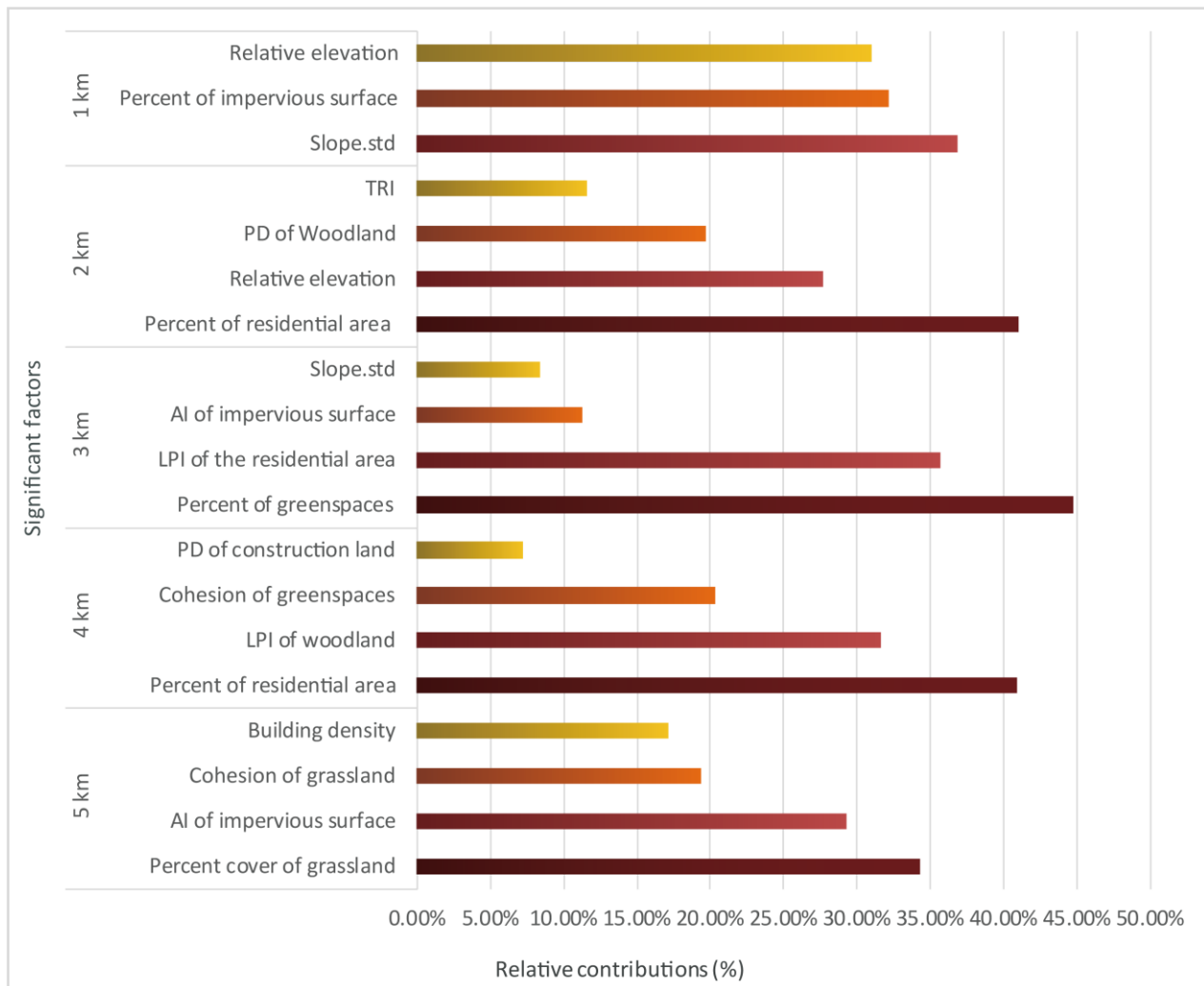


Figure 2-8 The relative contributions of significant factors affecting urban waterlogging across five analysis scales.

## 2.6 Discussion

### 2.6.1 Urban waterlogging agglomeration effect

According to the spatial pattern of urban waterlogging, the waterlogging events are mainly clustered in the historic urban districts of Guangzhou, indicating a significant spatial aggregation effect (Fig.2-6). Consequently, this means that these areas are prone to urban waterlogging events (Fig.2-7). This result is particularly important for urban waterlogging mitigation and risk management. The government and urban planners can gain a comprehensive understanding of the spatial distribution pattern of urban waterlogging events. Thereby, the local authorities can plan or build more drainage facilities for the waterlogging hot spot areas, as well as rational management of existing urban green spaces and impervious surfaces. Furthermore,

the urban waterlogging risk warning can give more priority to the waterlogging hot spot areas when facing heavy rainstorms. With the application of big data and open data (population distribution, urban education, and medical facilities), it is helpful to identify the areas with a high potential risk of urban waterlogging events (Lin et al., 2018). Simultaneously, the population density and urban facilities around urban waterlogging points can reveal the potential impact of urban waterlogging events, which is conducive to early warning and emergency response.

### *2.6.2 Implications for urban waterlogging mitigation*

A record of the relative contributions of each influencing factor is firstly presented and how they affect the urban waterlogging at multiple scales is described. Our results demonstrated that land cover composition plays a more important role in determining urban waterlogging than other influencing factors (Tang et al., 2018; Tehrani et al., 2019; Yu et al., 2018). Increasing impervious surfaces would significantly increase the risk of urban waterlogging, and thus aggravate the urban waterlogging magnitude; whereas the increase of urban green spaces could alleviate urban waterlogging dramatically. The percent cover of residential area (41.03%) and impervious surfaces (32.13%) have a large relative contribution to urban waterlogging under all analysis scales. This finding is also consistent with Zhang et al. (2018) and Quan et al. (2011). As a representative of the impervious surface, the residential area has a great adverse effect on the urban hydrological process, which exacerbates the occurrence of urban waterlogging events. In addition, the increase in impervious surface coverage is also associated with cutting off the hydrological relationship between surface water and groundwater, causing the groundwater level to drop continuously, and thus, threatening urban geological security. In contrast, the percent cover of urban green spaces (44.74%) and grassland (34.24%) also have a large relative contribution, which indicates that these two land cover types can effectively mitigate the urban waterlogging phenomena. This may be due to the fact that urban green spaces are permeable surfaces, which can play a positive role in decreasing surface runoff and regulating rainwater storage, so reducing the risk of waterlogging. These results suggest that in order to alleviate the occurrence of urban waterlogging events to the greatest extent, we should focus on controlling the expansion of the residential areas, and increase the urban green spaces as much as possible, among which the grassland and woodland are the main land cover types.

In the past, urban planning and construction often pursued flat terrain, which is convenient for urban development and building construction. However, the relatively flat terrain (difficult to drainage surface

runoff) will bring greater pressure on drainage systems and eventually increase the risk of urban waterlogging. The result of this study suggests that we should avoid pursuing the flatness of urban surface elevation, which provides important insights into urban planning and waterlogging management. Our results highlighted the necessity of roadbed subsidence or subsided green space. Specifically, the elevation of roadbed subsidence and subsided green space are 5-30 cm lower than the surrounding surface, thereby increasing the slope and RE. Combined with the drainage network or permeable surfaces, the elevation difference of roadbed and green areas can effectively discharge the cumulate surface runoff, which is beneficial to reduce the occurrence of urban waterlogging. These results provide a theoretical and practical reference for the government and urban planners for urban waterlogging prevention and management.

It is interesting to find that the land cover configuration also has a certain impact on the severity of urban waterlogging. Our study explores the relative contributions of these configuration factors, which expand previous studies (Yu et al., 2018; Zhang et al., 2018; Wang et al., 2017). For a fixed relative amount of land cover composition, the occurrence of urban waterlogging events can be significantly increased or decreased by different spatial arrangements. For example, the aggregated distribution pattern of the impervious surface or the fragmentation of urban green spaces would further increase the occurrence of urban waterlogging events, which suggested that we need to give more priority to controlling the dominance and aggregation of each land-use type. While enlarging the area of urban green spaces, the shape complexity of urban green spaces should also be considered. The urban green spaces can be designed as a rectangle, strip with complex boundaries, forming green corridors within the city center, so that the mitigation effect of the green spaces can cover the whole city, thereby reducing the risk of waterlogging events. Furthermore, an increase in cohesion and AI index of green areas can also decrease the magnitude of waterlogging, which suggests that the cluster distributed pattern with high connectivity further reduce the risk of urban waterlogging. This result also reveals the high spatial heterogeneity characteristic of urban landscapes. Forming the green corridors within the city can reduce the occurrence of waterlogging. However, on the other hand, the aggregated distribution of green areas can effectively alleviate the degree of urban waterlogging.

### *2.6.3 Scale effect*

In this study, we notice that the correlation directions of Pearson correlation, the determination coefficients of stepwise regression models, and the relative contributions of the significant factors are different across all analysis scales, thus presenting a strong scale effect. This scale effect phenomenon has led to some

contradictory results. Previous studies based on one single analysis scale may just only provide partial information about the overall landscape features. Furthermore, the analysis scale used by different scholars is different, which will interfere with the explanatory power of various factors and even change the correlation directions. Our study implies that this inconsistency is probably due to the differences in analysis scales. For example, we note that the elevation, TWI, the land cover composition (bare land and construction land), and drainage density are experienced contrary correlation directions. This inconsistency undoubtedly hinders the progress of urban waterlogging research. The result of hierarchical partitioning analysis also indicates that in a small analysis scale (grid unit 1 km) the topographic factors are confirmed as the significant variables and the relative contributions are larger than others. This result highlight that, under a small analysis scale, the influence of topography factors on waterlogging magnitude is greater than other factors. The reason for this may be that at small analysis scales, urban topography has a greater influence on the extent of urban waterlogging. However, with the increasing of the analysis scales, the explanatory power of topography factors gradually declines and even the elevation factors present a contrary relationship, which further confirms why previous studies have produced contradictory results. The dominant drivers vary across different analysis scales, presenting the scale effect of the influencing factors. We finally demonstrate one of the conclusions that the land cover composition contributes the most to urban waterlogging through multi-scale analysis. This result strongly illustrates the importance of the proper analysis scale. Lastly, we also demonstrate that the appropriate statistical scale for urban waterlogging studies only worked for specific influencing factors. Therefore, when we select the appropriate analysis scales, we should according to the characteristics of different study areas. For example, for the low-lying coastal cities with gentle terrain (the differences between the slope and relative elevation are relatively small), the influence of the topography factors can be ignored properly. Therefore, the large grid unit (3 km) can be selected for the appropriate analysis scale to explore the mechanism of urban waterlogging. However, for the mountain area with large terrain fluctuations (the differences between the slope and relative elevation are relatively large), we should select a small analysis scale (1 km).

#### *2.6.4 Limitations*

This study has its limitations. Firstly, only the central urban district of Guangzhou was selected as a case study to explore the relative contribution of environmental and anthropogenic factors on urban waterlogging events. To reveal the influencing factors in one single city may be insufficient. Furthermore, due to the

limitations and accessibility of urban waterlogging records, we collected urban waterlogging events from 2009 to 2015 and the data did not present the specific year of each waterlogging event (just recorded in the period 2009-2015). We could not evaluate the temporal change of the waterlogging events. Also, due to this reason, it is difficult to distinguish the rainfall events corresponding to each waterlogging point. Thus, we assume the rainfall intensity within the central urban districts of Guangzhou is constant. However, this may bring some uncertainty to the results. Consequently, in future research, it is necessary to consider the effect of precipitation in order to reveal the mechanism of urban waterlogging more comprehensively.

## 2.7 Conclusion

Revealing the dominant drivers of urban waterlogging is of great significance to the optimization of urban waterlogging prevention and management. Taking the waterlogging events in the central urban district of Guangzhou (P.R. China) from 2009 to 2015 as an example, we conducted a novel method that combines the stepwise regression model and hierarchical partitioning analysis to investigate the relative contributions of influencing factors to urban waterlogging through multi-scale study. This study gains three conclusions: (1) the urban waterlogging hot spots in Guangzhou are mainly concentrated in the historical urban areas of Guangzhou (Liwang, Yuexiu, Haizhu district), presenting a single-core aggregation pattern. This finding provides important insights to local authorities to identify the high-risk areas of urban waterlogging. (2) Urban waterlogging events are mainly affected by both land cover characteristics and urban topography. Among all conditioning factors, the percent cover of urban green spaces (44.74%), percent cover of residential area (41.03%), and slope.std (36.85%), have a dominant influence on the urban waterlogging magnitude. Furthermore, this result also suggests that holding the proportion of land cover features constant, the urban micro-topography also affects waterlogging magnitude effectively. Furthermore, the effect of land cover spatial configuration on waterlogging cannot be ignored. This result provides additional insights that urban waterlogging can be mitigated by balancing the relative composition of land cover features as well as optimizing their spatial configuration. An even distribution of impervious surfaces or forming green corridors can improve the state of urban waterlogging events. (3) Across different scales of analysis, the correlation directions of Pearson correlation, the determination coefficients ( $R^2$ ) of stepwise regression models, the relative contributions of the significant factors vary with the change of analysis scale, underlining a strong scale effect. Since the dominant drivers vary across different analysis scales, the appropriate statistical scale for urban waterlogging studies only worked for specific influencing factors, and thus the best analysis scale

for urban waterlogging study should be determined by the characteristics of study areas. The results of this study reveal the dominant influencing factors as well as examine the scale effect, which provides important enlightenment for urban waterlogging prevention and management.

### **Author Contributions**

**Qifei Zhang:** conceptualization, methodology, formal analysis, data curation, writing - original draft. **Zhifeng Wu:** resources, writing - review & editing, supervision. **Hui Zhang:** methodology, validation. **Giancarlo Dalla Fontana:** validation, writing - review & editing. **Paolo Tarolli:** conceptualization, methodology, writing - review & editing, overall supervision, project administration.

### **Acknowledgments**

The study was funded by the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0301); Team Project of Guangdong Provincial Natural Science Foundation (2018B030312004); University of Padova research projects DOR1948955/19 "Evaluation of the effectiveness of drainage systems in an agricultural context" and DOR2079232/20 "Soil erosion and waterlogging analysis in agricultural context". The authors sincerely thank the editor and the two anonymous reviewers for their constructive comments and suggestions.

### **Conflicts of Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jenvman.2020.110951>.

Supplementary materials include the following:

Table S2-1 All possible models for hierarchical partitioning at 1 km scale.

Table S2-2 Pearson correlation coefficients between waterlogging magnitude and land cover configuration across five analysis scales.

*Table S2-1 All possible models for hierarchical partitioning at 1 km scale.*

Categories	Models	Enter variables	$R^2$
(i)	(1)	Intercept	0.000
	(2)	Slope.std	0.410
	(3)	RE	0.502
(ii)	(4)	percent cover of impervious surface	0.353
	(5)	Slope.std, RE	0.653
	(6)	Slope.std, percent cover of impervious surface	0.507
	(7)	RE, percent cover of impervious surface	0.565
(iii)	(8)	Slope.std, RE, percent cover of impervious surface	0.685

Table S2-2 Pearson correlation coefficients between waterlogging magnitude and land cover configuration across five analysis scales (blue color represents a negative correlation, red color represents a positive correlation, and the deeper the color, the greater the correlation coefficient. “\*” and “\*\*” represent correlation coefficient significance at the 0.05 level and 0.01 level, respectively).

Analysis Scales	Configuration indicators	Impervious Surface	Urban Greenspace	Woodland	Grassland	Garden	Cultivate	Residence area	Road network	Construction land	Bare land	Water body
1km	PD	-0.276*	0.145*	0.190*	0.262*	0.115	0.038	-0.109	-0.098	-0.138	0.023	0.185*
	LPI	0.225*	-0.272*	-0.102	-0.258*	-0.145*	-0.068	0.153*	0.104	0.185*	-0.007	-0.057
	LSI	0.171*	-0.196*	-0.087	-0.063	-0.027	-0.016	0.177*	0.019	0.190*	-0.201*	-0.110
	Cohesion	0.210*	-0.259*	-0.162*	-0.271*	-0.139*	-0.141*	0.134*	0.101	0.170	-0.127	-0.052
	AI	0.354**	-0.375**	-0.345**	-0.215*	-0.134*	-0.174*	0.347*	0.361**	0.271*	-0.124	-0.134
2km	PD	-0.235*	0.306*	0.457**	0.294*	0.208*	0.208*	-0.214*	-0.272*	-0.392**	0.132	0.243*
	LPI	0.403**	-0.394**	-0.167*	-0.393**	-0.263*	-0.259*	0.253*	0.210*	0.291*	-0.259*	-0.043
	LSI	0.08	-0.278*	-0.390**	-0.304*	-0.092	-0.105	0.129	0.281*	0.303*	-0.184	-0.267
	Cohesion	0.349**	-0.309*	-0.143*	-0.253*	-0.227*	-0.182*	0.134*	0.277*	0.206*	-0.157	-0.142
	AI	0.313*	-0.351**	-0.382**	-0.303*	-0.394**	-0.357**	0.276*	0.225*	0.364**	-0.231*	-0.205*
3km	PD	-0.282**	0.522**	0.558**	0.318*	0.352**	0.241*	-0.395**	-0.339*	-0.510**	0.251*	0.262*
	LPI	0.435**	-0.473**	-0.417**	-0.437**	-0.382**	-0.288*	0.591**	0.395**	0.529**	-0.281*	-0.156*
	LSI	0.122*	-0.293**	-0.264*	-0.247*	-0.217*	-0.243*	0.283*	0.307**	0.237*	-0.132	-0.166*
	Cohesion	0.485**	-0.521**	-0.356**	-0.455**	-0.221*	-0.267*	0.303*	0.304*	0.287*	-0.209*	-0.226*
	AI	0.551**	-0.412**	-0.349**	-0.357**	-0.459**	-0.363**	0.351**	0.359**	0.478**	-0.135	-0.215*
4km	PD	-0.207*	0.485**	0.541**	0.272*	0.334*	0.223*	-0.231*	-0.258*	-0.479**	0.214*	0.219*
	LPI	0.546**	-0.456**	-0.415**	-0.299*	-0.427**	-0.255*	0.214*	0.385**	0.393**	-0.195	-0.122
	LSI	0.222*	-0.201*	-0.149*	-0.138*	-0.284*	-0.138*	0.246*	0.241*	0.241*	-0.044	-0.281*
	Cohesion	0.531**	-0.486**	-0.302*	-0.368**	-0.244*	-0.217*	0.230*	0.336*	0.268*	-0.108	-0.282*
	AI	0.336**	-0.326*	-0.267*	-0.387**	-0.432**	-0.328*	0.315*	0.247*	0.457**	-0.142	-0.271*
5km	PD	-0.213*	0.317*	0.423**	0.262*	0.139*	0.217*	-0.287*	-0.354**	-0.305*	0.170*	0.130
	LPI	0.480**	-0.374**	-0.347**	-0.258*	-0.287*	-0.261*	0.225*	0.350**	0.255*	-0.015	-0.196*
	LSI	0.139*	-0.192*	-0.179*	-0.099	-0.223*	-0.152*	0.316*	0.316*	0.108	-0.125	-0.279*
	Cohesion	0.413**	-0.355**	-0.277*	-0.496**	-0.233*	-0.275*	0.217*	0.218*	0.253*	-0.152*	-0.258*
	AI	0.479**	-0.148	-0.296*	-0.310*	-0.153*	-0.136*	0.272*	0.255*	0.426**	-0.114	-0.256*



## References

1. ADB (2011). People's Republic of China: Urban stormwater management and waterlogging disaster prevention. Technical Assistance Report. Asian Development Bank, Beijing.
2. Ahammed, F. (2017). A review of water-sensitive urban design technologies and practices for sustainable stormwater management. *Sustainable Water Resources Management*, 3. <https://doi.org/10.1007/s40899-017-0093-8>
3. Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, 38, 5–22. <https://doi.org/10.1111/j.0016-7363.2005.00671>
4. Barros, V. , & Stocker, T. F. (2012). Managing the risks of extreme events and disasters to advance climate change adaptation : special report of the intergovernmental panel on climate change. *Journal of Clinical Endocrinology & Metabolism*, 18(6), 586-599. <https://doi.org/10.1017/CBO9781139177245>
5. Boudou, M., Danière, B., & Lang, M. (2016). Assessing changes in urban flood vulnerability through mapping land use from historical information. *Hydrology and Earth System Sciences*, 20(1), 161–173. <https://doi.org/10.5194/hess-20-161-2016>
6. Chen, Y., Zhou, H., Zhang, H., Du, G., & Zhou, J. (2015). Urban flood risk warning under rapid urbanization. *Environmental Research*, 139, 3–10. <https://doi.org/10.1016/j.envres.2015.02.028>
7. Chen F. (2014) Urban Morphology and Citizens' Life. In: Michalos A.C. (eds) *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Dordrecht
8. Cushman, S. A., McGarigal, K., & Neel, M. C. (2008). Parsimony in landscape metrics: Strength, universality, and consistency. *Ecological Indicators*, 8(5), 691–703. <https://doi.org/10.1016/j.ecolind.2007.12.002>
9. Dong, W., Lian, Y., & Zhang, Y. (2018). Sustainable Development of Water Resources and Hydraulic Engineering in China Proceedings for the 2016 International Conference on Water Resource and Hydraulic Engineering. <https://doi.org/10.1007/978-3-319-61630-8>
10. Du, S., Van Rompaey, A., Shi, P., & Wang, J. (2015). A dual effect of urban expansion on flood risk in the Pearl River Delta (China) revealed by land-use scenarios and direct runoff simulation. *Natural Hazards*, 77(1), 111–128. <https://doi.org/10.1007/s11069-014-1583-8>
11. Fahy, B., Brenneman, E., Chang, H., & Shandas, V. (2019). Spatial analysis of urban flooding and

- extreme heat hazard potential in Portland, OR. *International Journal of Disaster Risk Reduction*, 39(September 2018), 101117. <https://doi.org/10.1016/j.ijdr.2019.101117>
12. Gaitan, S., ten Veldhuis, M. Claire, & van de Giesen, N. (2015). Spatial Distribution of Flood Incidents Along Urban Overland Flow-Paths. *Water Resources Management*, 29(9), 3387–3399. <https://doi.org/10.1007/s11269-015-1006-y>
  13. Gallien, T. W., Sanders, B. F., & Flick, R. E. (2014). Urban coastal flood prediction: Integrating wave overtopping, flood defenses and drainage. *Coastal Engineering*, 91, 18–28. <https://doi.org/10.1016/j.coastaleng.2014.04.007>
  14. Gardner, R. H., Gardner, R. H., & Gardner, R. H. (2001). *Landscape ecology in theory and practice*. <https://doi.org/10.1007/b97434>.
  15. Grahn, T., & Nyberg, L. (2017). Assessment of pluvial flood exposure and vulnerability of residential areas. *International Journal of Disaster Risk Reduction*, 21(January), 367–375. <https://doi.org/10.1016/j.ijdr.2017.01.016>
  16. Guo, G., Chen, Y., Wei, J., Wu, Z., & Rong, X. (2012). Impacts of grid sizes on urban heat island pattern analysis. *Shengtai Xuebao/ Acta Ecologica Sinica*, 32(12), 3764–3772. <https://doi.org/10.5846/stxb201107181068>
  17. Hallegatte, S., Green, C., Nicholls, R. J., & Corfee-Morlot, J. (2013). Future flood losses in major coastal cities. *Nature Climate Change*, 3, 802. Retrieved from <https://doi.org/10.1038/nclimate1979>
  18. Hammond, M.J., A.S. Chen, S. Djordjević, D. Butler & O. Mark (2015) Urban flood impact assessment: A state-of-the-art review, *Urban Water Journal*, 12:1, 14-29, DOI: 10.1080/1573062X.2013.857421
  19. Huang, H., Chen, X., Zhu, Z., Xie, Y., Liu, L., Wang, X., ... Liu, K. (2018). The changing pattern of urban flooding in Guangzhou, China. *Science of the Total Environment*, 622–623, 394–404. <https://doi.org/10.1016/j.scitotenv.2017.11.358>
  20. Huang, T., Wang, Y., & Zhang, J. (2017). Simulation and Evaluation of Low Impact Development of Urban Residential District Based on SWMM and GIS. *IOP Conference Series: Earth and Environmental Science*, 74(1). <https://doi.org/10.1088/1755-1315/74/1/012009>
  21. Huong, H. T. L., & Pathirana, A. (2013). Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. *Hydrology and Earth System Sciences*, 17(1), 379–394. <https://doi.org/10.5194/hess-17-379-2013>

22. IPCC. (2011). Intergovernmental Panel on Climate Change Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Cambridge: Cambridge University Press.
23. Iwahashi, J., and Pike, R.J. (2007). Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology* 86, 409-440. <https://doi.org/10.1016/j.geomorph.2006.09.012>
24. Lee, S., Kim, J.-C., Jung, H.-S., Lee, M., & Lee, S. (2017). Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomatics, Natural Hazards and Risk*, 1–19. <https://doi.org/10.1080/19475705.2017.1308971>
25. Li, B., Zhao, Y., & Fu, Y. (2015). Spatio-temporal Characteristics of Urban Storm Waterlogging in Guangzhou and the Impact of Urban Growth. *Journal of Geo-Information Science*, 17(4), 445–450. <https://doi.org/10.3724/SP.J.1047.2015.00445>
26. Li, C., Cheng, X., Li, N., Du, X., Yu, Q., & Kan, G. (2016). A framework for flood risk analysis and benefit assessment of flood control measures in Urban Areas. *International Journal of Environmental Research and Public Health*, 13(8). <https://doi.org/10.3390/ijerph13080787>
27. Li, S., Zeng, H., Xia, J., & Zhang, L. (2004). Current situation and some intending problems of landscape spatial dynamic model. *Ying Yong Sheng Tai Xue Bao = The Journal of Applied Ecology / Zhongguo Sheng Tai Xue Xue Hui, Zhongguo Ke Xue Yuan Shenyang Ying Yong Sheng Tai Yan Jiu Suo Zhu Ban*, 15, 701–706.
28. Lin, Tao, Liu, Xiaofang, Song, & Jinchao, et al. (2018). Urban waterlogging risk assessment based on internet open data: a case study in china. *Habitat International*. <https://doi.org/10.1016/j.habitatint.2017.11.013>
29. McGarigal, K., & Marks, B. (1995). FRAGSTATS—Spatial Pattern Analysis Program for Quantifying Landscape Structure. U. S. Forest Service General Technical Report. <https://doi.org/10.2737/PNW-GTR-351>
30. Miao, Z. T., Han, M., & Hashemi, S. (2019). Correction to: The effect of successive low-impact development rainwater systems on peak flow reduction in residential areas of Shizhuang, China (*Environmental Earth Sciences*, (2019), 78, 2, (51), 10.1007/s12665-018-8016-z). *Environmental Earth Sciences*, 78(3), 1–3. <https://doi.org/10.1007/s12665-019-8088-4>
31. Nally, R. Mac. (2000). Regression and model-building in conservation biology, biogeography and

ecology: The distinction between – and reconciliation of – “predictive” and “explanatory” models. *Biodiversity and Conservation*, 9, 655–671.

32. Nations, U., (2014). *World Urbanization Prospects the 2014 Revision*.
33. Pijl, A., Brauer, C. C., Sofia, G., Teuling, A. J., & Tarolli, P. (2018). Hydrologic impacts of changing land use and climate in the Veneto lowlands of Italy. *Anthropocene*, 22, 20–30. <https://doi.org/10.1016/j.ancene.2018.04.001>
34. R. Quan, L. Zhang, M. Liu, M. Lu, J. Wang and H. Niu, "Risk assessment of rainstorm waterlogging on subway in central urban area of Shanghai, China based on scenario simulation," 2011 19th International Conference on Geoinformatics, Shanghai, 2011, pp. 1-6, doi: 10.1109/Geoinformatics.2011.5981176.
35. R Core Development Team. (2008). *R: A Language and Environment for Statistical Computing*. <http://www.R-project.org/>.
36. Riley, S. J., S. D. DeGloria and R. Elliot. (1999). A terrain ruggedness index that quantifies topographic heterogeneity, *Intermountain Journal of Sciences*, vol. 5, No. 1-4, 1999.
37. Schuch, G., Serrao-Neumann, S., Morgan, E., & Low Choy, D. (2017). Water in the city: Green open spaces, land use planning and flood management – An Australian case study. *Land Use Policy*, 63, 539–550. <https://doi.org/10.1016/j.landusepol.2017.01.042>
38. Sene, K. (2013). Flash Floods, 293–311. <https://doi.org/10.1007/978-94-007-5164-4>
39. Shao, W., Zhang, H., Liu, J., Yang, G., Chen, X., Yang, Z., & Huang, H. (2016). Data Integration and its Application in the Sponge City Construction of CHINA. *Procedia Engineering*, 154, 779–786. <https://doi.org/10.1016/j.proeng.2016.07.583>
40. Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., & Smith, D. R. (2005). Impacts of impervious surface on watershed hydrology: A review. *Urban Water Journal*, 2(4), 263–275. <https://doi.org/10.1080/15730620500386529>
41. Sofia, G., Ragazzi, F., Giandon, P., Dalla Fontana, G., Tarolli, P. (2019). On the linkage between runoff generation, land drainage, soil properties, and temporal patterns of precipitation in agricultural floodplains. *Advances in Water Resources*, 124, 120–138. <https://doi.org/10.1016/j.advwatres.2018.12.003>
42. Sofia, G., Roder, G., Dalla Fontana, G., & Tarolli, P. (2017). Flood dynamics in urbanised landscapes: 100 years of climate and humans' interaction. *Scientific Reports*, 7, 40527.

<https://doi.org/10.1038/srep40527>

43. Sofia, G., Prosdocimi, M., Dalla Fontana, G., Tarolli, P. (2014). Modification of artificial drainage networks during the past half-century: Evidence and effects in a reclamation area in the Veneto floodplain (Italy), *Anthropocene*, 6, 48-62. <https://doi:10.1016/j.ancene.2014.06.005>
44. Su, M., Zheng, Y., Hao, Y., Chen, Q., Chen, S., Chen, Z., & Xie, H. (2018). The influence of landscape pattern on the risk of urban water-logging and flood disaster. *Ecological Indicators*, 92, 133–140. <https://doi.org/10.1016/j.ecolind.2017.03.008>
45. Tang, X., Shu, Y., Lian, Y., Zhao, Y., & Fu, Y. (2018). A spatial assessment of urban waterlogging risk based on a Weighted Naïve Bayes classifier. *Science of the Total Environment*, 630, 264–274. <https://doi.org/10.1016/j.scitotenv.2018.02.172>
46. Tehrany, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena*, 175(December 2018), 174–192. <https://doi.org/10.1016/j.catena.2018.12.011>
47. Tien Bui, D., Pradhan, B., Nampak, H., Bui, Q.-T., Tran, Q.-A., & Phi, Q. (2016). Hybrid Artificial Intelligence Approach Based on Neural Fuzzy Inference Model and Metaheuristic Optimization for Flood Susceptibility Modelling in A High-Frequency Tropical Cyclone Area using GIS. *Journal of Hydrology*, 540, (Accepted, in press). <https://doi.org/10.1016/j.jhydrol.2016.06.027>
48. United Nations, Department of Economic and Social Affairs, Population Division, 2014. *World Urbanization Prospects: The 2014 Revision, CD-ROM Edition*
49. Wang, C., Du, S., Wen, J., Zhang, M., Gu, H., Shi, Y., & Xu, H. (2017). Analyzing explanatory factors of urban pluvial floods in Shanghai using geographically weighted regression. *Stochastic Environmental Research and Risk Assessment*, 31(7), 1777–1790. <https://doi.org/10.1007/s00477-016-1242-6>
50. Wang, J., Gao, W., Xu, S., & Yu, L. (2012). Evaluation of the combined risk of sea level rise, land subsidence, and storm surges on the coastal areas of Shanghai, China. *Climatic Change*, 115(3–4), 537–558. <https://doi.org/10.1007/s10584-012-0468-7>
51. Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., & Bai, X. (2015). Flood hazard risk assessment model based on random forest. *Journal of Hydrology*, 527, 1130–1141. <https://doi.org/10.1016/j.jhydrol.2015.06.008>
52. Werner, M., Hunter, N. M., & Bates, P. D. (2005). Identifiability of Distributed Floodplain Roughness

- Values in Flood Extent Estimation. *Journal of Hydrology*, 314, 139–157.  
<https://doi.org/10.1016/j.jhydrol.2005.03.012>
53. Woodruff, J. D., Irish, J. L., & Camargo, S. J. (2013). Coastal flooding by tropical cyclones and sea-level rise. *Nature*, 504(7478), 44–52. <https://doi.org/10.1038/nature12855>
  54. Wu, J. (2004). Effects of Changing Scale on Landscape Pattern Analysis: Scaling Relations. *Landscape Ecology*, 19, 125–138. <https://doi.org/10.1023/B:LAND.0000021711.40074.ae>
  55. Wu, J., Yang, R., & Song, J. (2018). Effectiveness of low-impact development for urban inundation risk mitigation under different scenarios: A case study in Shenzhen, China. *Natural Hazards and Earth System Sciences*, 18(9), 2525–2536. <https://doi.org/10.5194/nhess-18-2525-2018>
  56. Wu, J., & Zhang, P. (2017). The effect of urban landscape pattern on urban waterlogging. *Dili Xuebao/Acta Geographica Sinica*, 72(3), 444–456. <https://doi.org/10.11821/dlxb201703007>
  57. Xue F., Huang M., Wang W., & Zou L. (2016). Numerical simulation of urban waterlogging based on flood area model. *Advances in Meteorology*, 2016, 1-9. <https://doi.org/10.1155/2016/3940707>
  58. Yao, L., Chen, L., & Wei, W. (2017). Exploring the linkage between urban flood risk and spatial patterns in small urbanized catchments of Beijing, China. *International Journal of Environmental Research and Public Health*, 14(3). <https://doi.org/10.3390/ijerph14030239>
  59. Yin, Z., Yin, J., Xu, S., & Wen, J. (2011). Community-based scenario modelling and disaster risk assessment of urban rainstorm waterlogging. *Journal of Geographical Sciences*, 21(2), 274–284. <https://doi.org/10.1007/s11442-011-0844-7>
  60. Yu, H., Zhao, Y., Fu, Y., & Li, L. (2018). Spatiotemporal variance assessment of urban rainstorm waterlogging affected by impervious surface expansion: A case study of Guangzhou, China. *Sustainability (Switzerland)*, 10(10). <https://doi.org/10.3390/su10103761>
  61. Zhang, H., Cheng, J., Wu, Z., Li, C., Qin, J., & Liu, T. (2018). Effects of impervious surface on the spatial distribution of urban waterlogging risk spots at multiple scales in Guangzhou, South China. *Sustainability (Switzerland)*, 10(5). <https://doi.org/10.3390/su10051589>
  62. Zhang, H., Wu, C., Chen, W., & Huang, G. (2017). Assessing the impact of climate change on the waterlogging risk in coastal cities: A case study of Guangzhou, South China. *Journal of Hydrometeorology*, 18(6), 1549–1562. <https://doi.org/10.1175/JHM-D-16-0157.1>
  63. Zhang, Y., Xia, J., Yu, J., Randall, M., Zhang, Y., Zhao, T., ... Shao, Q. (2018). Simulation and assessment of urbanization impacts on runoff metrics: insights from landuse changes. *Journal of*

Hydrology, 560, 247–258. <https://doi.org/10.1016/j.jhydrol.2018.03.031>

64. Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. *Science of the Total Environment*, 615, 1133–1142. <https://doi.org/10.1016/j.scitotenv.2017.10.037>
65. Zimmer, C. A., Heathcote, I. W., Whiteley, H. R., & Schroter, H. (2007). Low-Impact-Development Practices for Stormwater: Implications for Urban Hydrology. *Canadian Water Resources Journal*, 32(3), 193–212. <https://doi.org/10.4296/cwrj3203193>
66. Zope, Eldho, J. (2016). Impacts of land use-land cover change and urbanization on flooding: A case study of Oshiwara River Basin in Mumbai, India. *Catena*, 145, 142–154. <https://doi.org/10.1016/j.catena.2016.06.009>

# CHAPTER 3

## 3. A new approach to investigating the spatially heterogeneous driving forces of urban waterlogging<sup>2</sup>

Qifei Zhang<sup>a,b</sup>, Zhifeng Wu<sup>b,c,d</sup>, Guanhua Guo<sup>b</sup>, Paolo Tarolli<sup>a\*</sup>

- a. Dept. of Land, Environment, Agriculture and Forestry, University of Padova, 35020 Legnaro, PD, Italy.
- b. School of Geographical Sciences, Guangzhou University, 510006 Guangzhou, Guangdong province, China.
- c. Southern Marine Science and Engineering Guangdong Laboratory, 511458 Guangzhou, Guangdong province, China.
- d. MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, 518000 Shenzhen, China

---

<sup>2</sup> *This chapter is a style-edited version of the published article:*

Zhang, Q., Wu, Z., Guo, G., & Tarolli, P. (2021). A new approach to investigating the spatially heterogeneous driving forces of urban waterlogging. Under review.

*Contributions by the PhD candidate included: primary authorship in writing and revision of the entire article; primary responsible for literature review, data collection and quantification, methodology, data processing and analysis, and production of all figures.*



### 3.1 Abstract

Under the combined effects of climate change and urbanization, urban waterlogging seriously threatens urban sustainable development and human life. It is widely accepted that various landscape elements contribute to the magnitude of urban waterlogging. However, less attention is given to the spatial heterogeneity effects of urban landscape elements on urban waterlogging. Considerable efforts investigate the universal mechanism of urban waterlogging on a global statistical scale by regarding the whole study area as spatially homogeneous while neglecting the spatial heterogeneity issue and the local specific mechanism. Less effort focuses on the spatial heterogeneity driving forces at the local scale, which hinders the implementation of target-specific urban waterlogging mitigation strategies. To shed some light on this topic, an innovative method that integrates the best subset regression model, cubist regression tree, and geographical detector model is presented to spatially explicit the heterogeneous forces driving waterlogging variation and identify the waterlogging dominant factors with different local conditions. The results show that the best subset regression proposed in this study can more objectively select the most representative and meaningful driving factors according to local characteristics since urban waterlogging is largely disturbed by local conditions. By comparing with two other commonly used regression methods (global regression model, spatial lag model), the combination of cubist regression tree and geographical detector model can fully quantify the spatial non-stationarity effect of representative driving factors on waterlogging and spatially explicit the local driving forces. The waterlogging's dominant driving factors and their contribution vary with the local site conditions. The development of site-specific waterlogging mitigation strategies is facilitated by dividing the watershed units as strong dominance and weak dominance watershed units. This denotes that urban waterlogging variation is mainly determined by dominant factors in a strong dominance watershed (the independent contribution of dominant factors has an overwhelming advantage), while waterlogging in a weak dominance watershed is jointly affected by multiple factors (the independent contribution of each factor is similar). The spatial heterogeneous driving forces facilitate the implementation of more targeted and effective mitigation strategies, rather than a "one-size-fits-all" policy. This study extends our scientific understanding of the site-specific mechanism of waterlogging, providing useful information for formulating urban waterlogging mitigation strategies with different local conditions.

**Keywords:** Urban waterlogging; spatially heterogeneous driving forces; spatial non-stationarity relationship; cubist regression tree; geographical detector model

## 3.2 Introduction

With the acceleration of urbanization, the negative effects of population and industrial agglomeration have damaged the urban ecosystem (Wang et al., 2011; Frank et al., 2017; Cepeliauskaite et al., 2020). Urban waterlogging describes the phenomenon that heavy or continuous precipitation far exceeds the retention and drainage capacity of a region, resulting in stagnant water disasters in underground space or low-lying areas (Jiang et al., 2018; Chen et al., 2015; Yu et al., 2018). At present, many cities in China have frequently suffered from waterlogging problems (Ning et al., 2017; Zhang et al., 2020; Huang et al., 2018; Li et al., 2017). Specifically, the increased impervious surface caused by urbanization cuts off the infiltration path of rainwater, thereby disrupting the water balance and increasing the peak flow. This extreme event has brought a series of negative consequences, causing serious casualties and huge socio-economic losses. For example, on July 20, 2021, Zhengzhou city, China suffered from a severe urban waterlogging disaster (cumulative average precipitation was 449 mm, maximum one-hour precipitation reached 201.9 mm), resulting in 292 deaths and 47 missing persons. Aware of these impacts, many local authorities propose waterlogging mitigation strategies to design more sustainable living environments in urban areas.

In fact, the implementation of these waterlogging mitigation strategies depends to a large extent on the understanding of the mechanism of urban waterlogging events. Considerable research has shown that the causes of urban waterlogging can be summarized into four aspects. (1) Meteorological conditions: macroscopically, global climate change has led to an increase in extreme precipitation events, creating background conditions for waterlogging (Tan et al., 2021; Berghuijs et al., 2017; Du et al., 2019; Blöschl et al., 2019). On a regional scale, the change of urban microclimates, such as the urban rain islands effect, further increases the probability and intensity of precipitation in urban areas (Xu et al., 2019; Kim et al., 2017; Liu et al., 2020; Miller and Hutchins, 2017). The effect of meteorological conditions on the risk of waterlogging has been extensively documented. For example, Zhang et al. (2017) assessed the impact of climate change on the probability of inundation in coastal cities based on eight climate models and four CO<sub>2</sub> emission scenarios. Giulia et al. (2017) analyzed flood dynamics in northeastern Italy between 1900 and 2010, which revealed that climate change is a significant factor that increases the frequency of flood hazards. Therefore, the changes in the spatial and temporal distribution of precipitation will ultimately affect the entire urban water cycle process, leading to an increased risk of urban waterlogging. (2) Topographic conditions: first of all, the topographic pattern of high in the west and low in the east determines the flow direction of major rivers in

China. Continuous precipitation in the upper and middle reaches is likely to bring extreme flood control pressure to the downstream cities (usually highly urbanized national cities). Secondly, the effect of urban micro-topography (urban lowland, surface roughness) on urban waterlogging cannot be underestimated. A well-documented consequence of urban lowlands (tunnels, culverts, underground parking lots) within cities are conducive to the accumulation of surface runoff and are often the most severe sites of urban waterlogging (Zhang et al., 2020; Zhang et al., 2021b; Tang et al., 2018; Yalcin et al., 2020). The results of Tehrany et al. (2019) indicate that compared with relatively flat areas, high-altitude areas are less likely to experience waterlogging events. Similarly, Liu et al. (2021) fully confirmed the impact of surface elevation and slope on urban waterlogging by using the geographical detect model. This finding embodies the plains with a gentle slope condition are vulnerable to urban waterlogging. However, the study conduct in Amsterdam, Netherlands (Gaitan et al., 2015), shows a contradictory result, which may be caused by the spatial heterogeneity within the city. (3) Drainage facilities: in the process of rapid urbanization, a large number of rivers within cities have been artificially landfilled, altering the urban drainage pattern. This leads to continuous degradation of the river network and its drainage function, thereby increasing the risk of urban waterlogging. Therefore, local authorities built underground drainage pipes or pumping stations to accelerate the discharge rate of surface runoff. Nevertheless, some studies have pointed out that the drainage facilities in developing countries are characterized by low design standards and poor maintenance, which is difficult to drain surface runoff effectively (Lin et al., 2021; Lin et al., 2018; Wu et al., 2020; Zhang et al., 2020; Quan et al., 2014). (4) Land cover characteristics: urbanization and human activities have led to the continuous encroachment of ecological lands such as woodland, wetland, and farmland. This not only weakens the capacity to regulate surface runoff, but also occupies the spaces for flood storage, which undoubtedly increases the frequency and intensity of urban waterlogging. A proliferation of studies have shown that among various causes of waterlogging, the impact of land cover characteristics is particularly significant, which has gradually become the major reason for the increasing severity of urban waterlogging (Quan et al., 2010; Su et al., 2018; Wu et al., 2012; Yu et al., 2021). Zhang et al. (2020) confirmed that the land cover characteristics are the dominant factors for waterlogging at different analysis scales. Yu et al. (2018) analyzed the mechanism of the spatial and temporal pattern of impervious surfaces on urban waterlogging in Guangzhou. Tran et al. (2020) utilized the OLS regression model to find that the impervious surface and NDVI played a decisive role in the risk of urban waterlogging. In summary, these findings embody the mentioned factors usually coexist and together contribute to the occurrence of waterlogging events.

It is widely accepted that urban waterlogging is the result of both natural conditions and human activities. However, there are various urban landscape elements with different attributes in cities, especially in highly urbanized areas. Moreover, the spatial distribution of urban landscape elements and their attributes are characterized by heterogeneous and dynamic. This phenomenon is widely regarded as spatial heterogeneity, describing the complexity and heterogeneity of the landscape pattern and its process (Pickett and Cadenasso, 1995; Song et al., 2020; Jiang, 2015). Due to spatial heterogeneity, urban landscape elements (such as DEM, land cover characteristics) vary in different spatial locations. As a result, the relationship between landscape elements and urban waterlogging may vary with different spatial locations, which is known as spatial non-stationarity. It hints that the driving factors or even the mechanism of urban waterlogging may vary with the change of spatial location. This indicates that in the urban environment with high spatial heterogeneity, the mechanism of urban waterlogging in different regions is complex and different. Although there have been considerable studies exploring the mechanism of urban waterlogging (Ke et al., 2020; Yu et al., 2018; Sun, 2014; Su et al., 2018; Zhang et al., 2021b; Jian et al., 2021). However, most of the previous efforts regard the whole study area as spatial homogeneous, without taking spatial heterogeneity into account. How landscape elements in different spatial locations affect urban waterlogging in highly heterogeneous urban areas needs further understanding. Therefore, an interesting question arises: how to spatially explicit the mechanism of urban waterlogging?

Most recent studies utilized global statistical methods to explore the relationship between urban waterlogging and landscape elements (Zhang et al., 2020; Lin et al., 2018; Zhang et al., 2018; Wu and Zhang, 2017; Sun, 2014). However, the global statistical methods mainly focus on revealing the general mechanism from a global perspective. These approaches assume that the relationship between urban waterlogging and its drivers is constant across the entire study area, that is, only exploring the universal mechanism of urban waterlogging while ignoring its local specific mechanism. This means that the global statistical methods may not accurately reveal the spatial non-stationarity relationship of urban waterlogging. In view of this, some scholars have proposed to use the geographically weighted regression model (GWR) or spatial lag model (SLM) to analyze this spatial non-stationary relationship (Wang et al., 2017; Liu et al., 2021). However, most efforts neglect the relative contribution of various landscape elements, in which impediments provide valuable site-specific waterlogging driving forces. Additionally, it is still unclear how the interactive effects of landscape elements affect urban waterlogging. In reality, urban waterlogging is affected by multiple factors, rather than by a single factor alone. If only analyze the individual effects of a single landscape element on urban

waterlogging (ignoring the interaction effects), it may lead to certain deviations, especially in highly spatial heterogeneous urban areas (with enormous landscape elements). These shortcomings lead to a lack of understanding of the spatial heterogeneous driving forces of urban waterlogging. Consequently, it cannot provide a theoretical basis for formulating site-specific waterlogging mitigation strategies for different local conditions. From this perspective, some interesting questions emerge: how to assess the spatially heterogeneous effects of landscape elements on urban waterlogging? How to obtain the dominant driving factor of urban waterlogging in different spatial locations? How does the interaction between different landscape elements affect urban waterlogging?

Due to the abovementioned shortcomings, we developed an innovative approach to investigate the spatial non-stationarity mechanism of urban waterlogging by integrating the best subset regression model, cubist regression tree, and geographical detector model. Two highly urbanized coastal cities, Guangzhou and Shenzhen were considered to test this proposition. As urban waterlogging is largely disturbed by local conditions, there is a need in selecting the most representative and meaningful driving factors based on the characteristics of the different study areas. Therefore, the specific objectives of this study were to: (1) determine the representative driving factors of urban waterlogging by different local conditions; (2) investigate the spatial non-stationarity effects of representative driving factors on urban waterlogging; (3) identify the individual and interaction effect of local driving forces, consequently, spatially explicit the driving forces. This method improves our understanding of the spatial non-stationarity mechanism of urban waterlogging and provides additional insights into the spatial variation of urban waterlogging driving forces. These results are helpful for local authorities to develop more target-specific urban waterlogging mitigation strategies for different local conditions.

### **3.3 Study area**

Guangzhou city is located between 112°57' to 114°3'E and 22°26' to 23°56'N, with an area of 7434.40 km<sup>2</sup> (Figure 3-1a). The topography of Guangzhou is high in the northeast and low in the southwest, with a coastal alluvial plain in the south (average elevation 6.6 m). Shenzhen city is located between 113°46' – 114°37'E and 22°27' – 23°52'N, which adjacent to Hong Kong, with an area of 1997.47 km<sup>2</sup> (Figure 3-1b). The GDP of Guangzhou and Shenzhen in 2020 reached \$384 billion and \$415 billion respectively, becoming one of the world's first-tier cities. These two cities both belong to subtropical maritime monsoon climates, with an

annual average rainfall of around 1720 mm and 1935 mm respectively. However, the temporal and spatial distribution of precipitation in the two cities is extremely unbalanced. Temporally, the precipitation is mainly concentrated in April to September, accounting for over 80% of the annual precipitation. Spatially, affected by the prevailing wind direction, the precipitation is decreasing from northeast to southwest (Guangzhou), and southeast to northwest (Shenzhen). Additionally, Guangzhou and Shenzhen are often affected by typhoons and tropical cyclones in summer.

However, due to the topographic features of Guangzhou and Shenzhen (coastal lowlands), under the background of rapid urbanization, coupled with the temporal and spatial differences in precipitation, these two cities have frequently suffered from waterlogging events. On May 7, 2017, Guangzhou was hit by a torrential rainstorm, with a maximum precipitation of 184.4 mm/h and the maximum 3-hour precipitation exceeded the historical extreme value of Guangdong Province (382.6 mm). This incident caused 109 houses to collapse, 118 waterlogging events within the city center, about 38.91 km<sup>2</sup> of farmland were flooded, and the economic loss was about \$89 million. Similarly, on May 22, 2020, a heavy rainstorm in Guangzhou (80 mm average hourly precipitation, 167.8 mm maximum hourly precipitation), which killed 4 people and flooded multiple underground parking lots and metro stations. On May 11, 2014, Shenzhen encountered a heavy rainstorm, resulting in over 170 waterlogging events, more than 2,000 vehicles were flooded, which seriously affected the production and life of citizens. According to the recorded areas of urban waterlogging events, we selected the central urban districts of Guangzhou (1559.82 km<sup>2</sup>) and Shenzhen as the study area. In view of the high spatial heterogeneous and the severe urban waterlogging problems in these two regions, it is appropriate to select these two cities as the study areas.

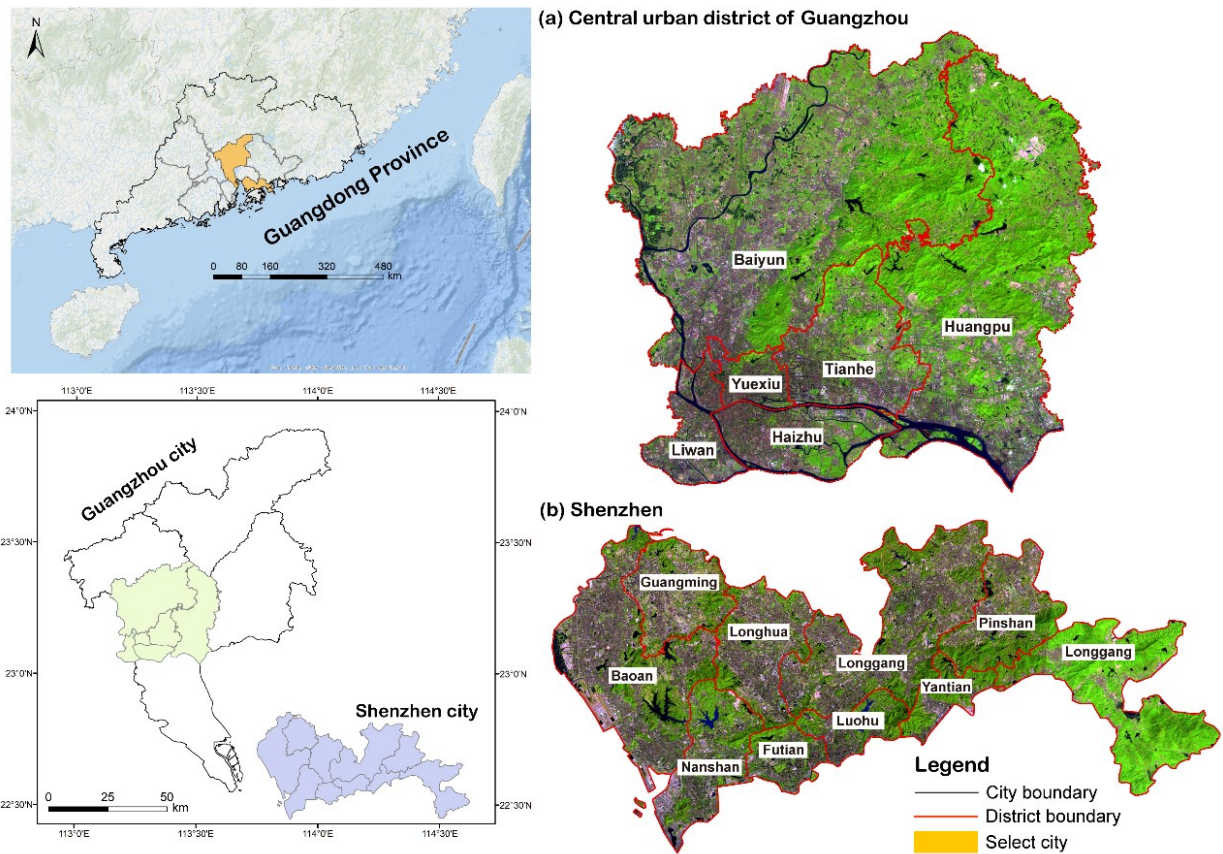


Figure 3-1 The selected study area.

### 3.4 Data and methodology

#### 3.4.1 Spatial data

In this study, we focused on the period from 2009 to 2015. The urban waterlogging data are mainly derived from the address information of urban waterlogging records provided by the Guangzhou and Shenzhen Water Authority. Referring to the study of Yu et al., 2018 and Zhang et al., 2021a, b, we utilized ArcGIS Pro to locate the waterlogging events and established the urban waterlogging inventories data sets in these two cities. Finally, we collected 423 and 353 urban waterlogging events in Guangzhou and Shenzhen, respectively. Then, the waterlogging inventories data sets were superimposed on the watershed division map to calculate the density of waterlogging events in each watershed unit, which was used to characterize the waterlogging magnitude of each watershed (Figure 3-2).

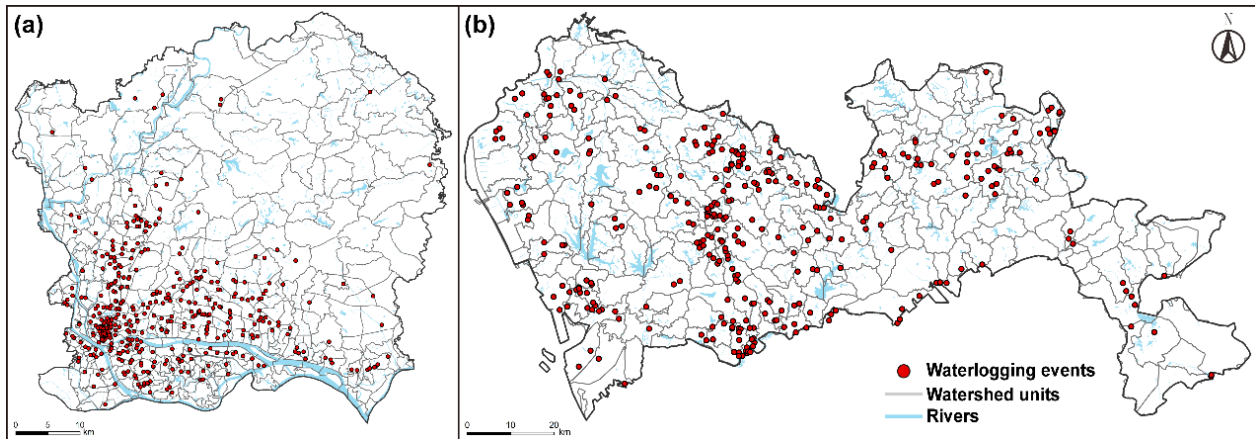


Figure 3-2 The waterlogging events in Guangzhou (a) and Shenzhen (b).

As mentioned previously, waterlogging is influenced by various landscape elements. Therefore, this study also collected the aerial remote sensing images of Guangzhou and Shenzhen acquired in 2013 and the cloud-free Landsat-8 images (path/row: 122-44, 121-44), DEM (spatial resolution 5 m), drainage facilities data, and average precipitation data. As the local water authority only recorded urban waterlogging events for this period (no specific time), we only select one scene remote sensing image during this period. All raw data sources are shown in Table 3-1.

Table 3-1 Metadata information of the spatial data.

Data	Resolution	Time	Source
Waterlogging location	Point	2009 – 2015	Water Resources Authority of Guangzhou and Shenzhen
Landsat-8 OLI imagery	30 m	2013	The United States Geological Survey
Aerial remote sensing images	0.5 m	2013	Land Resources Technology
DEM	5 m	2013	Center of Guangdong Province
Drainage network	Line	2012	
Precipitation	1000 m	2009 – 2015	Meteorological Bureau of Guangzhou and Shenzhen

### 3.4.2 Explanatory factors



Selecting appropriate explanatory factors based on local conditions is crucial to exploring the spatial non-stationarity mechanism of urban waterlogging. Therefore, this study selects as many landscape elements as possible as waterlogging potential explanatory factors, which can be categorized into environmental factors (urban topography, precipitation) and anthropogenic factors (land cover characteristics, urban drainage facilities).

#### 3.4.2.1 Environmental factors

In this study, the elevation (DEM), the standard deviation of elevation (DEM.std), slope, the standard deviation of slope (Slope.std), and the proportion of depression (Dep) were selected to represent terrain surface features (Figure 3-3). In general, the areas with low elevation and gentle slopes have a high likelihood of waterlogging. Therefore, we utilized the zonal statistical tool to calculate the average elevation and average slope of each watershed to represent the topography condition of each watershed unit. The DEM.std and the Slope.std were indicated the fluctuation of elevation and slope over the surface. A large value of the DEM.std and the Slope.std imply the terrain surface is undulating and has higher water flow velocity, so waterlogging is less prone to occur. The depression is the urban lowland surrounded by higher elevation where surface runoff is easy to accumulate (Huang et al., 2019). The proportion of depression (Dep) describes the area proportion of depression in the watershed that is defined as:

$$Dep = \frac{D_i}{A} \quad (3.1)$$

where D represents the area of depression in watershed unit i and A denotes the area of watershed unit i. The identification of depression can refer to the supplementary data. Since the urban waterlogging record does not include a specific year, the average accumulated precipitation was used to illustrate the spatial distribution pattern of precipitation during the period.

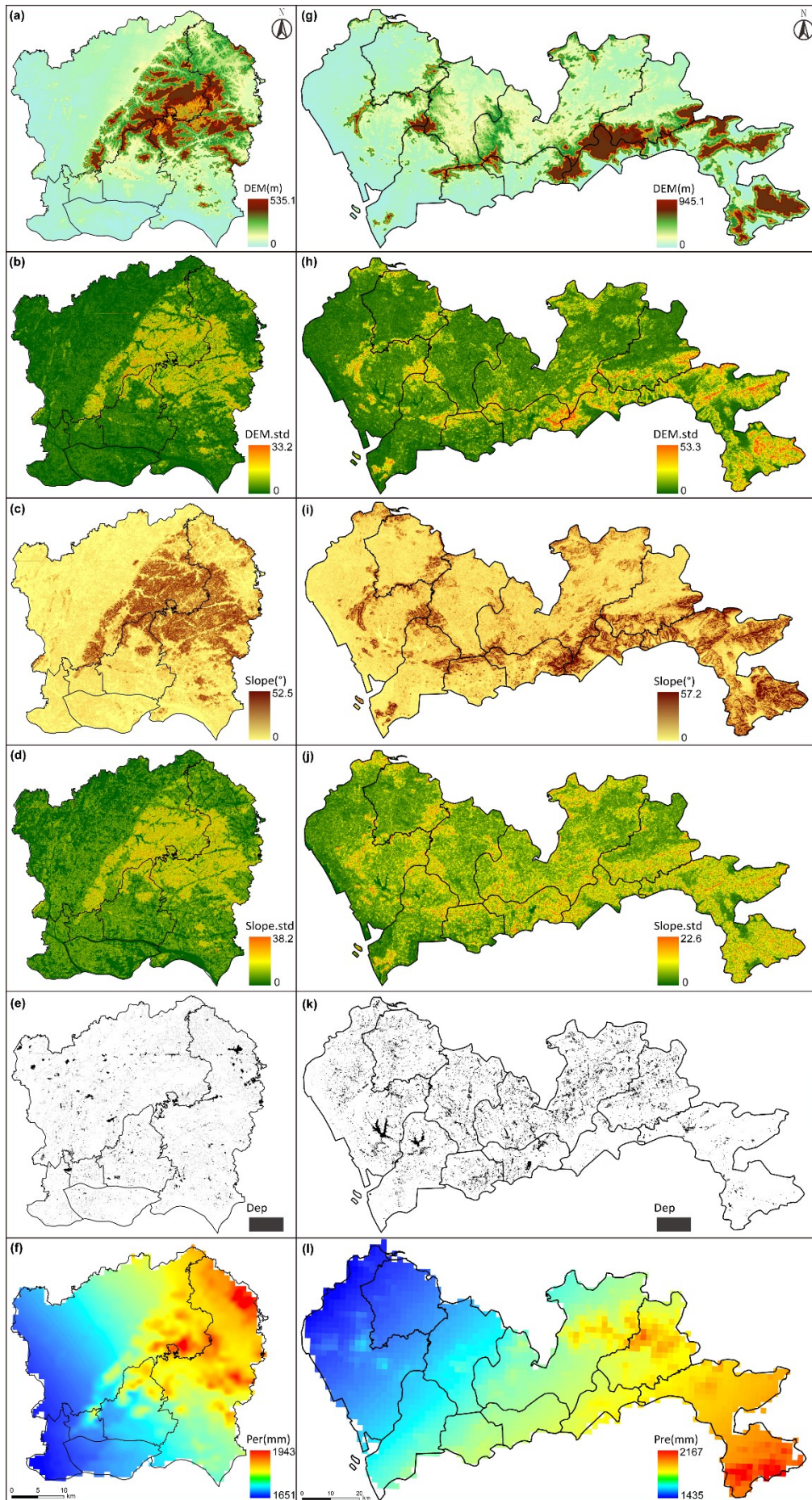


Figure 3-3 The potential environmental factors of waterlogging in Guangzhou (a–f) and Shenzhen (g–l).

### 3.4.2.2 Anthropogenic factors

According to the existing research, it is clear that impervious surfaces cut off the infiltration process of surface rainwater and increase surface runoff, so waterlogging is prone to occur. Urban green spaces have a great impact on the infiltration of surface rainwater and retard runoff velocity, which plays a positive role in regulating rainwater. Therefore, two variables, the area proportion of impervious surface (ISA) and urban green space (UGI) in each watershed unit were extracted as potential driving factors. In this study, we extracted the impervious surfaces and urban green spaces using the aerial remote sensing images, through an object-oriented classification method. The overall accuracy of the classification was 82.5% and 85.8% through the field measurement approach. However, we notice the proportion of green area cannot reflect the biophysical parameters of vegetation. For example, under the same coverage of urban green space, different vegetation growth states or densities have different effects on urban waterlogging. Based on this, this study introduced the normalized differential vegetation index (NDVI) to describe the biophysical parameters of green areas.

The landscape spatial composition and structure can be quantitatively described using landscape pattern metrics (McGarigal, 1995; Cushman et al., 2008). Based on the principle that the landscape pattern metrics are mutually independent, we select three metrics to reflect the spatial configuration characteristics of land cover features, including: (1) landscape fragmentation: mean patch size (MPS) and patch density (PD); (2) landscape aggregation: aggregation index (AI). The equation and description of these UGI metrics were shown in Table 3-2 and were calculated through Fragstats 4.2.

Table 3-2 Selected landscape pattern metrics.

Metric	Calculation	Implication
MPS	$\sum_{i=1}^n \frac{A_i}{n}$	MPS measures the average size of each land cover.
PD	$\frac{A_i}{S}$	PD represents the fragmentation degree of each land cover type.
AI	$\left[ \begin{matrix} g_i \\ \max \rightarrow g_i \end{matrix} \right]$	AI reflects the degree of aggregation of each land cover type.

\* $A_i$ : patch  $i$  area,  $S$ : total area,  $n$ : number of patches,  $g_i$ : number of adjacent patches

Apart from the effect of land cover characteristics, drainage facilities also have a certain effect on waterlogging. Therefore, based on the distribution data of the rainwater pipe network and rainwater outlet, the linear density of the rainwater pipe network (DD) of each watershed unit was extracted as the anthropogenic factors.

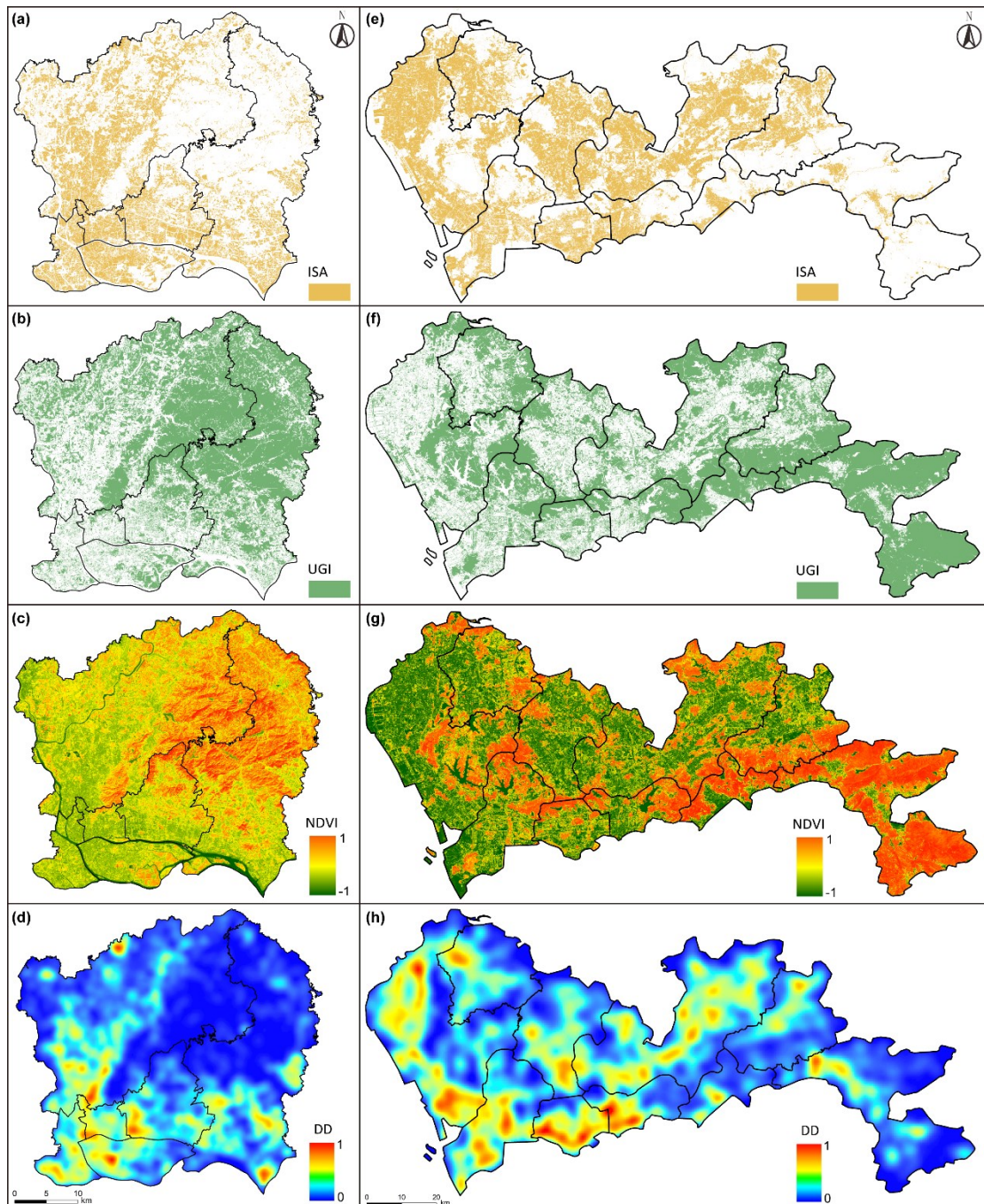
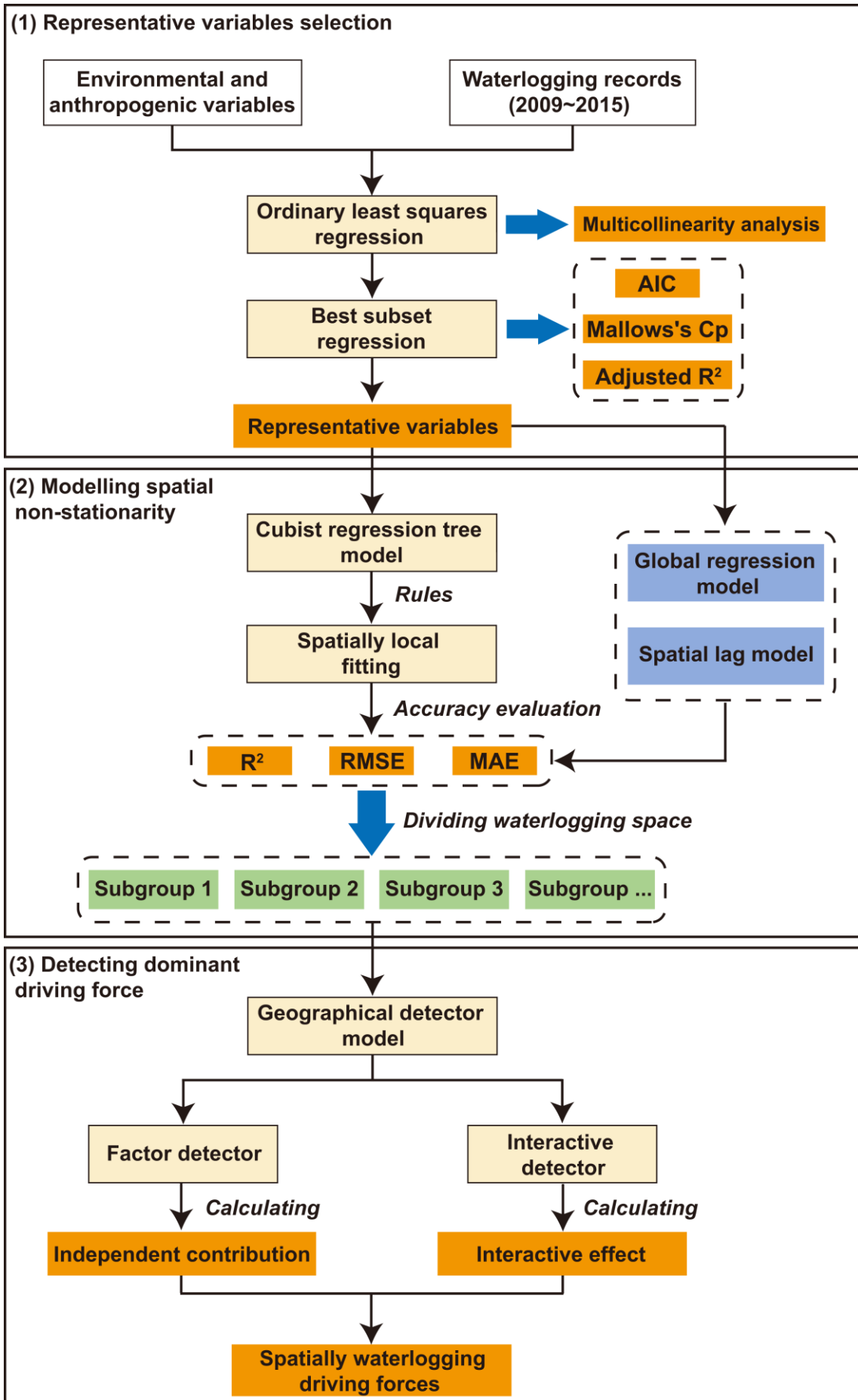


Figure 3-4 The potential anthropogenic factors of waterlogging in Guangzhou (a–d) and Shenzhen (e–h).

### 3.4.3 Methodology

We proposed a novel approach to fully investigate the spatially heterogeneous effects of driving factors on urban waterlogging and to effectively detect the dominant drivers of urban waterlogging in different spatial locations. This approach can be divided into three main steps (Figure 3-5). First, we propose an objective method that utilized ordinary least squares regression and best subset regression to determine the representative driving factors of each study area from potential explanatory factors (Step 1 in Figure 3-5). Secondly, according to the selected representative driving factors, the urban waterlogging spaces were divided into multiple subgroups to reveal the spatial non-stationarity relationship and explicitly identifies the spatial heterogeneity driving forces of urban waterlogging by cubist regression tree (Step 2 in Figure 3-5). Thirdly, we used the geographical detector model to further estimate the individual effect and interaction effect of the driving forces in different regional subgroups (Step 3 in Figure 3-5).



*Figure 3-5 The framework of the proposed approach.*

#### 3.4.3.1 Identify representative driving factors

Since urban waterlogging is a systemic problem, we have introduced various landscape elements as waterlogging potential explanatory factors. However, if all landscape elements are considered as driving factors to enter the regression tree model, it will inevitably lead to multicollinearity problems, increase the load of the model, and reduce the accuracy of the model. Furthermore, it is necessary to select representative driving factors based on the condition of the different study areas.

In this study, the method of selecting representative driving factors for each study area can be decomposed into two steps: the removal of multivariate collinearity and best subsets regression. Firstly, all potential variables were entered into an OLS regression model for collinearity detection. According to the tolerance (Tol) and variance inflation factor (VIF), the potential environmental and anthropogenic variables with strong collinearity (Tol < 0.1 or VIF > 10) are removed, and the significance test is performed at the level of 0.01. Secondly, we input the factors that have been tested for collinearity into the best subset regression model. It considers all the possible combinations of collinearity-tested factors on urban waterlogging. The best subset regression model is a method to select the best subset of independent variables and determine the global optimal regression model (Peng et al., 2018; Wang, 2009; Guo et al., 2021). The core principle of this method is the exhaustive method, which performs regression analysis for all possible combinations of all independent variables to assess the best-fitting models. If there are  $n$  input variables, there are  $2^n - 1$  combined models. Subsequently, we select the best model among all the combined models according to the adjusted  $R^2$ . However, if the adjusted  $R^2$  is similar, the Akaike Information Criteria (AIC) and Mallows's  $C_p$  were also applied to select the best model with the smallest AIC and Mallows's  $C_p$  value. Based on this, we regard the environmental and anthropogenic variables in this best regression model as representative driving factors. Then we plot the relationship between waterlogging density and representative driving factors to further analyze this complex linking.

#### 3.4.3.2 Modeling spatial non-stationarity relationship

In the first step, we respectively determined the representative urban waterlogging driving factors of the two cities. In this step, we develop the cubist regression tree based on these representative driving factors to investigate the complex spatial non-stationarity relationships between urban waterlogging and

representative driving factors. The cubist regression tree is a non-parametric statistical process of data analysis, which can effectively deal with the modeling of discrete and nonlinear relationships (Walton, 2008; Xiao et al., 2007; Guo et al., 2020). The characteristic of the cubist regression tree is the binary tree structure that makes full use of the data in the analysis process. By continuously subdividing the data samples, the binary tree node makes the maximum variation of the variables in the upper node, while with similar homogeneity in the data variables within the same branch nodes. Finally, the variables in the same branch tend to be homogeneous (Guo et al., 2015). Therefore, this study divides the urban waterlogging space into different subgroups through the rules established by the cubist regression tree. The spatial heterogeneity within the subgroups is minimized, while the spatial heterogeneity between groups is maximized. The cubist tree then investigates the spatial non-stationarity relationships between urban waterlogging and representative driving factors in each subgroup. Additionally, this study further compares the cubist regression tree with two commonly used regression methods, the global regression model (GRM) and the spatial lag model (SLM). The accuracy of these models is assessed by correlation coefficients, root mean square error (RMSE), mean absolute error (MAE) to verify the validity and superiority of cubist regression tree in modeling urban waterlogging.

#### 3.4.3.3 Detecting dominant driving forces spatially

The causes of urban waterlogging events are often complex. Spatially detecting the dominant driving factors of waterlogging events and their interactive effect is helpful to provide site-specific suggestions for urban waterlogging prevention. The geographical detector model (GDM) is utilized to calculate the relative contribution of representative driving factors in each subgroup. In this study, the representative factors with the largest contribution in each subgroup were regarded as the dominant factors of urban waterlogging. Thereby revealing which representative factors play a leading role in urban waterlogging at different spatial locations, that is, to determine the dominant driving forces spatially.

The geographical detector model is the new spatial statistical analysis model that can effectively diagnose the spatial heterogeneity of landscape elements and reveal its driving forces (Wang et al. 2010; Zhang et al., 2019; Song et al., 2020). This method has no linear hypothesis, which effectively overcomes the limitations of traditional statistical methods. It has been widely used to study the spatial heterogeneity and driving forces of geographical phenomena (Luo et al., 2016; Wang et al., 2016). The GDM contains four detectors, including factor detector, interaction detector, risk zone detector, and ecological detector (Wang et al., 2016). As the



main purpose of revealing the dominant driving forces spatially, we employ the factor detector and the interaction detector in this study.

The factor detector is used to detect the degree of explanation of each representative driving factor to the spatial differentiation of urban waterlogging, and its formula is as follows:

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} = 1 - \frac{SSW}{SST} \quad (3.2)$$

where SSW is the sum of variance within the stratum, and SST is the total variance of the whole area. The value of q is between 0 and 1, with larger q-values indicating the stronger the explanatory power of representative factors for urban waterlogging, and vice versa. In addition, in order to judge whether the interaction between representative factors will enhance the interpretation of urban waterlogging, this study further uses the interaction detector to quantify the interaction effect of two separate representative factors on the waterlogging. The interactions were classified into five categories: non-linearly weaken, unitary weaken, binary enhancement, independent, and nonlinear enhancement, as shown in Table 3-3.

*Table 3-3 Interaction effect between representative factors.*

Interaction	Equation	Description
Nonlinear weaken	$q(A \cap B) < \min[q(A), q(B)]$	It means that the effect of the interaction of the two factors A and B on the dependent variable Y is less than the independent effects.
Unitary weaken	$\min[q(A), q(B)] < q(A \cap B) < \max[q(A), q(B)]$	It shows that the result of interaction weakens the effect of one of the factors.
Binary enhancement	$q(A \cap B) > \max[q(A), q(B)]$	It indicates that the interaction between factors A and B has a greater impact on Y than a single factor.
Independent	$q(A \cap B) = q(A) + q(B)$	The interaction is independent, indicating that the two factors do not interfere with each other.

Nonlinear  
enhancement

$$q(A \cap B) > q(A) + q(B)$$

It means that the interaction effect has a greater impact on Y than the sum of the independent effects of factors A and B.

### 3.5 Results

#### 3.5.1 Representative driving factors selection

As shown in Table 3-4, after comparing the adjusted  $R^2$ , AIC, and Mallows's Cp of all possible combinations, the model with 6 variables including ISA, DEM, Slope.std, Dep, UGI.MPS, and Per was selected as the best model for Guangzhou, while the model contained 5 variables with ISA, DEM.std, UGI.MPS, ISA.PD, and ISA.AI was selected as the optimal model for Shenzhen. Therefore, the variables in these two models are considered as the representative driving factors of urban waterlogging in Guangzhou and Shenzhen, respectively. It is found that Guangzhou and Shenzhen have similarities and differences in the selection of representative driving factors. For example, ISA and UGI.MPS jointly serve as representative factors for Guangzhou and Shenzhen, while DEM, slope.std, Dep and Per are the representative factors specific to Guangzhou. In contrast, DEM.std, ISA.PD, ISA.AI are the representative factors unique to Shenzhen.

Table 3-4 The best subset regression and representative driving factors selection.

City	Representative driving factors selection	AIC	Mallows's Cp	Adjusted $R^2$
Guangzhou	ISA, DEM, Slope.std, Dep, UGI.MPS, Per	104.93	33.26	0.548
Shenzhen	ISA, DEM.std, UGI.MPS, ISA.PD, ISA.AI	88.77	14.37	0.662

As can be seen in Figure 3-6, the Pearson correlation coefficients for these 11 representative driving factors are all relatively high, indicating a significant correlation between urban waterlogging and the selected representative factors. The DEM, DEM.std, and Slope.std both have a negative correlation with waterlogging; while the Dep experience a strong positive correlation with waterlogging. This means that the fluctuation of the ground surface contributes to the dispersion of surface runoff, thus less prone to waterlogging. In contrast, if the Dep in the watershed unit is too high, the probability of waterlogging will eventually aggravate. In addition, as the frequency and intensity of extreme rainstorms are increasing, this will undoubtedly increase

the impact of precipitation factors on urban waterlogging. Compared with environmental variables, the land cover characteristics have a more significant correlation with waterlogging. The ISA experiences a significant negative impact on the hydrological environment, which will deteriorate the regional waterlogging state. Conversely, the UGI within watershed units has a positive effect on regulating surface runoff and mitigating urban waterlogging. For spatial configuration, there is a positive relationship between ISA.AI and waterlogging, while ISA.PD shows a negative correlation. This suggests that the more aggregated or less fragmented distribution pattern of impervious surfaces, the more prone to waterlogging disasters in this area.

More importantly, from the fitting curve, we found that these correlations are not a simple linear relationship (Figure 3-6). In detail, ISA is exponentially related to the waterlogging density (Figure 3-6a, g). As the increase of ISA, the increasing trend of urban waterlogging is gradually intensified. In Guangzhou city, when the ISA is less than 50%, an increase in the impervious surface area does not significantly increase urban waterlogging density. However, when the area proportion exceeds a threshold (50%), urban waterlogging density increases significantly with the increase of the proportion of impervious surface. We also found that when the UGI.MPS reached 2.0 ha (Guangzhou) or 5.0 ha (Shenzhen), the decreasing trend of urban waterlogging density was no longer obvious (Figure 3-6e, i). This implies that there may be a threshold value for the mitigating effect of urban green spaces on waterlogging. When the MPS of green spaces in a watershed unit is too large, its mitigation effect may be wasted. Therefore, in the process of urban greening, it is recommended to comprehensively consider the MPS of green spaces and their mitigation effect. Furthermore, a similar phenomenon also exists in the spatial configuration of impervious surfaces (PD, AI). As the parameters to measure the terrain variation (DEM, DEM.std, Slope.std), when the terrain variation in the watershed unit is within the threshold, increasing the DEM, DEM.std, or Slope.std can significantly reduce the density of urban waterlogging. However, when the terrain variation of the watershed unit is significant (exceeding the threshold), the downward trend of urban waterlogging density is not obvious.

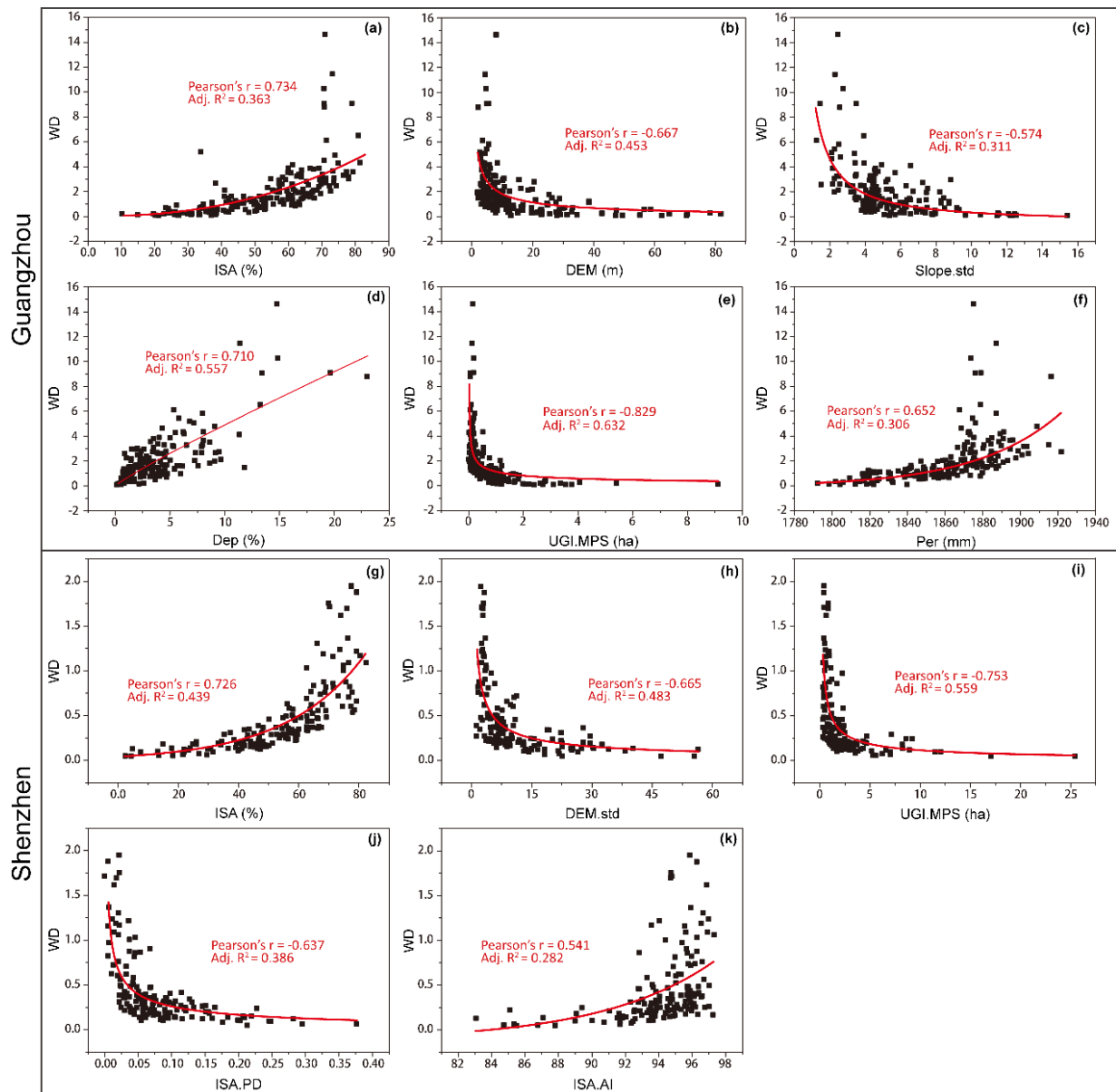


Figure 3-6 The relationship between waterlogging density and representative driving factors of Guangzhou (a–f) and Shenzhen (g–k).

### 3.5.2 The performances of cubist regression tree

As shown in the above section, the fitting curves indicate that the relationship between representative driving factors and urban waterlogging density is not a simple linear relationship. Therefore, this study introduced the cubist regression tree to further reveal this complex mechanism and to compare the accuracy with the GRM and the SLM. The results show that among the three models, the GRM has the lowest accuracy both in Guangzhou and Shenzhen, indicating its poor ability in handling the spatial non-stationarity relationship between urban waterlogging and the representative factors. Compared with the GRM, the SLM has better performance, with higher  $R^2$  and lower error (RMSE, MAE), but not as good as the cubist regression tree. The

adjusted  $R^2$  of the cubist tree model in Guangzhou is 0.79, and the RMSE and MAE are only 1.06 and 0.31, while in Shenzhen the adjusted  $R^2$ , RMSE, and MAE are 0.88, 0.17, and 0.11. These results show that the cubist regression trees have the highest accuracy, both in Guangzhou and Shenzhen, which provides important insight that the cubist regression tree is more effective in detecting the spatial non-stationarity effect.

*Table 3-5 The performances of cubist regression tree, spatial lag model, and global regression model.*

City	Model	Index		
		Adjusted $R^2$	RMSE	MAE
Guangzhou	Cubist tree	0.7964	1.0671	0.3114
	SLM	0.7376	1.2395	0.4327
	GRM	0.5231	1.5532	0.6892
Shenzhen	Cubist tree	0.8824	0.1763	0.1119
	SLM	0.8112	0.2344	0.2207
	GRM	0.6330	0.5975	0.5131

### *3.5.3 The spatial non-stationarity mechanism of urban waterlogging*

The cubist regression tree divides the urban waterlogging space into different heterogeneity subgroups by rules to investigate the spatial non-stationarity mechanism of urban waterlogging. As shown in Figure 3-7, the number of rules in Guangzhou and Shenzhen are 9 and 6, respectively. Based on these rules, the waterlogging spaces of Guangzhou and Shenzhen are divided into 9 and 6 subgroups with the minimum internal spatial heterogeneity and the maximum external spatial heterogeneity. As results revealed, the urban waterlogging space in Guangzhou was divided into more subgroups, implying the variation of waterlogging in Guangzhou is complicated and highly spatially heterogeneous.

The cubist regression tree obtains the equations between urban waterlogging and representative driving factors in different subgroups, which highlights the role of representative factors on urban waterlogging within different subgroups in a heterogeneous environment (Table S3-1). For example, when the ISA in the watershed units (Guangzhou) is less than or equal to 23.14%, the urban waterlogging density variation in these watershed units can be explained by Expression 1. On the contrary, when the ISA in the watershed units

(Guangzhou) is greater than 23.14%, the impact of representative factors on urban waterlogging can be explained by the remaining eight rules. Moreover, we found that some subgroups were influenced by representative factors such as ISA, DEM, and UGI.MPS (Guangzhou Rule 1), while others were influenced by ISA, Dep, and DEM (Guangzhou Rule 7). This finding extends our scientific understanding that although there are consistent representative factors in the same city, different subgroups are affected by different representative factors. It is also important to note that several rules are only available for a very small number of watershed units. For example, rule 3 in Guangzhou and Shenzhen contains only 8 and 10 watershed units. These results show that the driving forces of urban waterlogging are spatially different, which is mainly due to the large spatial heterogeneity of landscape elements in urban areas. In other words, although environmental conditions are common on a regional scale, the driving force of urban waterlogging will vary depending on the local conditions.

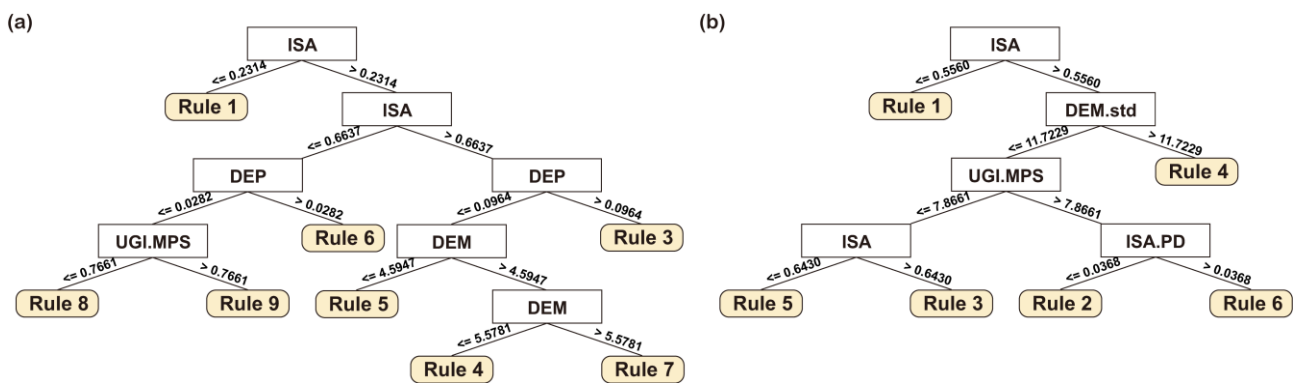


Figure 3-7 The divided rules for Guangzhou (a) and Shenzhen (b).

This study further projected the rules to the remote sensing imagery to identify the landscape features corresponding to the different rules (Figure 3-8). It is interesting to note that the corresponding watershed units of each rule have different landscape characteristics with a strong spatial heterogeneity effect. For example, the spatial pattern of rule 1 and rule 2 of Guangzhou are mainly concentrated in the urban fringe, especially the northeast part of Guangzhou (Huangpu and Baiyun district). The watershed units in these areas have a large topographic relief and the dominant land cover feature is woodland (Figure 3-8A). Additionally, rule 6 tends to correspond to newly built-up areas of the city (Tianhe), where the proportion of impervious surfaces and green areas is more reasonable and the building density is relatively low (Figure 3-8B). On the contrary, the distribution of rules 3 and 4 are mainly distributed in the historic urban district (Liwan and Yuexiu district), where the land cover features are dominated by impervious surfaces with high building density and relatively flat terrain (Figure 3-8C). Similar to Guangzhou, Shenzhen's rule 1 is mainly distributed in the

Dapeng District, which is the location of the National Forest Park. Therefore, the vegetation abundance of watershed units in this region is relatively high (Figure 3-8D). In addition, Shenzhen's rules 3, 4, and 5 are mainly distributed in the city center, and the watershed units corresponding to these rules are occupied by dense built-up areas (Figure 3-8E). In rule 6, the landscape features are characterized by low-rise factories or warehouses (Figure 3-8F).

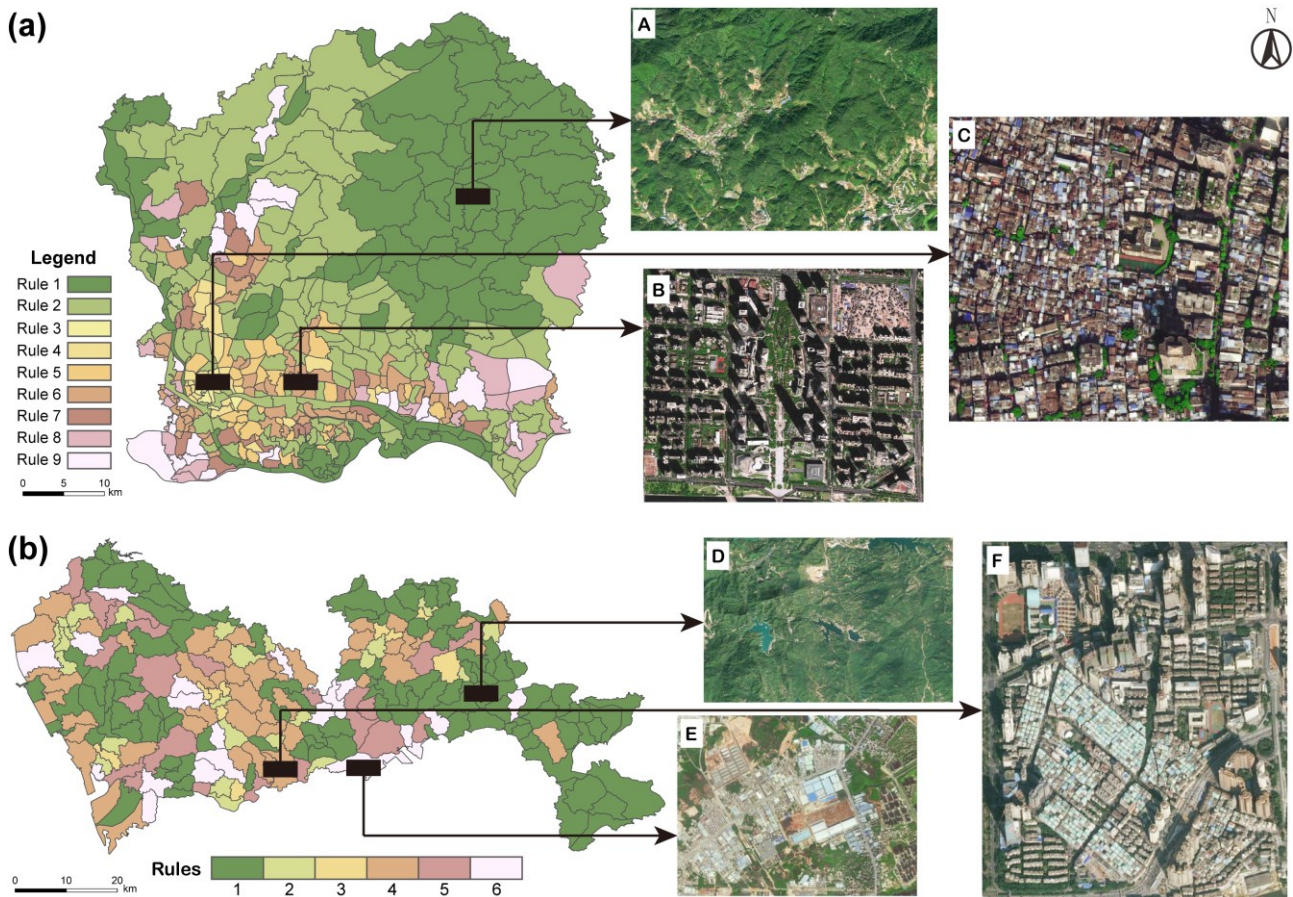


Figure 3-8 The spatial distribution pattern of rules and their corresponding landscape features in Guangzhou (a) and Shenzhen (b).

### 3.5.4 Dominant driving forces spatial assessment

#### 3.5.4.1 Waterlogging dominant factors and their independent contribution

We found that the dominant factors of urban waterlogging have certain differences in watershed units at different spatial locations. In general, the urban waterlogging in Guangzhou is spatially controlled by UGI.MPS, ISA, Dep, DEM, and Slope.std, while in Shenzhen is spatially affected by ISA, DEM.std, ISA.AI, and ISA.PD (Figure 3-9a, b). This implies that the extent of the influence of representative driving factors on urban waterlogging depends on its spatial location. In Guangzhou, the UGI.MPS is the dominant factor for subgroups

1 and 5; the ISA is the dominant driving factor for subgroups 2, 4, 7, and 9. This indicates that in these 6 subgroups, land cover characteristics have a decisive effect on urban waterlogging. However, the watershed units of subgroups 3, 6, and 8 are greatly affected by topographical factors, whose dominant factors are the Dep, DEM, and Slope.std, respectively. In Shenzhen, the ISA is the dominant driving factor of urban waterlogging in subgroups 1 and 4, while ISA.AI and ISA.PD are the dominant factors in subgroups 3 and 5, respectively. The watershed units in subgroups 2 and 6 are mainly influenced by topographic factors, all of which are dominated by Slope.std. Overall, the watershed units in Guangzhou and Shenzhen with ISA as the dominant factor occupied 36.93% and 64.04% of the city's area, respectively. This result indicates that ISA is basically the dominant driver of urban waterlogging in all watershed units in Guangzhou and Shenzhen. The UGI.MPS has become the second dominant factor in Guangzhou, while the second dominant factor in Shenzhen is DEM.std. These two factors occupy 31.27% and 12.36% of the area of Guangzhou and Shenzhen respectively. In contrast, there are relatively few watershed units with Dep and ISA.AI as the dominant factor, which only occupy 4.2% and 6.8% of the area of Guangzhou and Shenzhen.

Furthermore, the relative contribution rates of the dominant factors also have significant spatial heterogeneity (Figure 3-9c, d). The independent contribution of the dominant factor in all watershed units in Guangzhou was 23.28% to 57.82%, with a standard deviation of 0.11, while the independent contribution of Shenzhen ranged from 25.95% to 53.59%, with a standard deviation of 0.09. Generally, the contribution of the dominant factors in Guangzhou fluctuates greatly, which may be due to the large spatial non-stationarity of Guangzhou. It is worth noting that even for the same dominant factor, its contribution varies considerably in different subgroups. For example, in Guangzhou City, the dominant factors in the watershed units corresponding to rules 2, 3, 7, and 9 are all ISA. However, the independent contribution rate of ISA in rule 7 is as high as 57.82%, while the independent contribution of ISA in rule 4 is only 31.07%. Similarly, the dominant factor of watershed units corresponding to rules 2 and 6 of Shenzhen City is DEM.std, but the independent contribution in rule 6 is 40.51%, while it is only 25.95% in rule 2. These results hint that the dominant factors have different explanatory degrees for urban waterlogging in different watershed units.



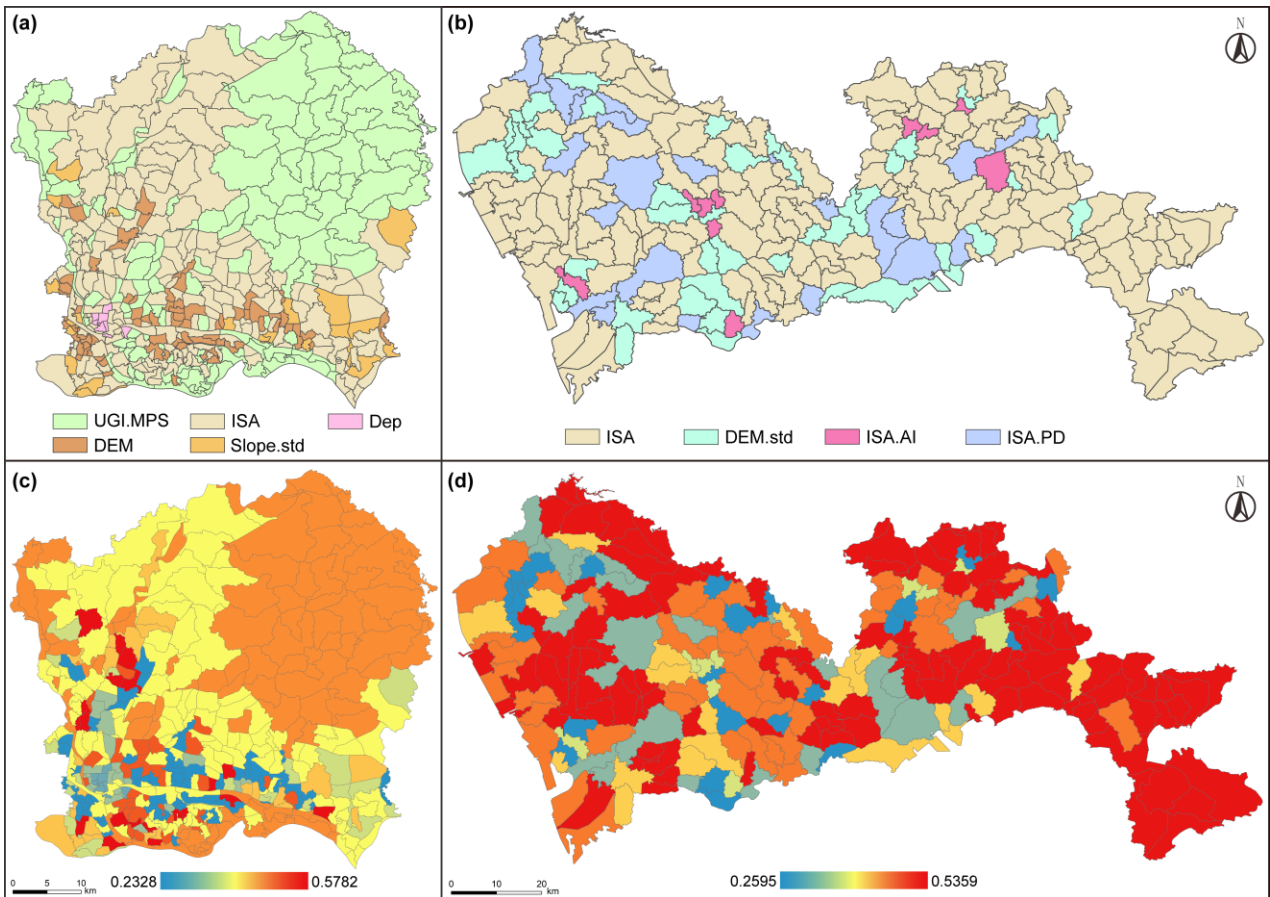


Figure 3-9 The spatial distribution of dominant factors (a, b) and their independent contributions (c, d).

Through further analysis of the independent contributions of representative driving factors to urban waterlogging in each subgroup (Figure 3-10). We found that the independent contributions of the dominant factors in some watershed units have an overwhelming advantage. For example, in Guangzhou subgroup 1, the independent contribution of the UGI.MPS is 51.29%, while the independent contribution of the ISA and DEM (remaining representative factors) are only 11.53% and 8.47%. The contribution of the dominant factors (UGI.MPS) in subgroup 1 exceeds the contribution of the remaining representative factors (ISA, DEM) by over 40%. Similarly, in Shenzhen subgroup 1, the independent contribution of ISA is approximately 37% higher than that of UGI.MPS. On the contrary, the independent contribution of each factor in some watershed units is not much different, without an overwhelming dominant factor. For example, in Guangzhou subgroup 3, the dominant factor Dep (30.41%) has a similar independent contribution to the representative factors such as ISA (27.24%) and DEM (26.51%). This implies that the urban waterlogging magnitude of these watershed units is jointly influenced by multiple factors.

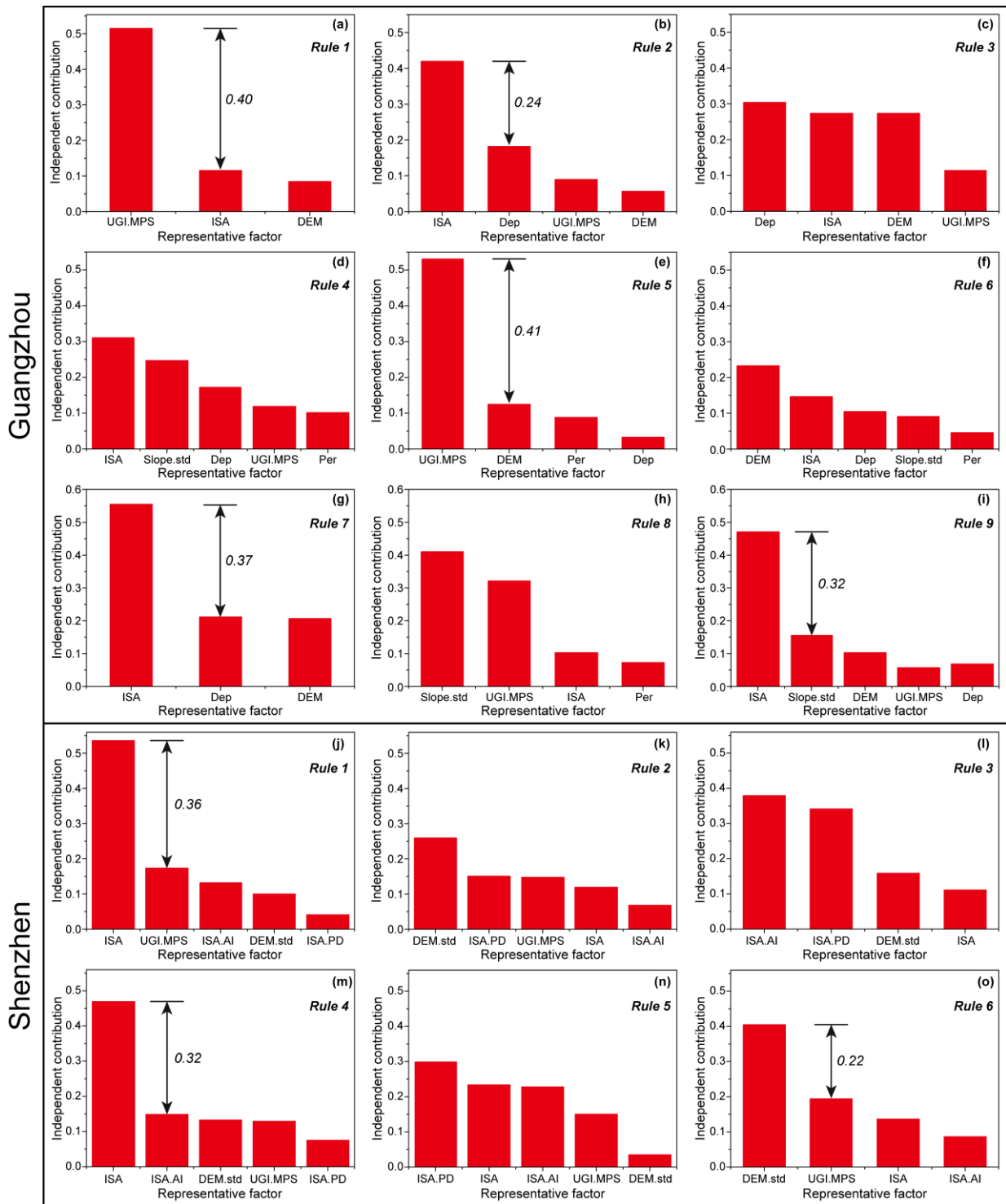


Figure 3-10 The independent contribution of each representative factor in Guangzhou (a–i) and Shenzhen (j–o).

Based on this, if the independent contribution of the dominant factor exceeds the remaining representative factors by more than 20%, we consider that the urban waterlogging in the watershed unit is mainly influenced by its dominant factor and classify it as a strong dominance watershed unit. In contrast, when the independent contribution of the dominant factor is not significantly different from the reminding factors (less

than 20%), we believe that the urban waterlogging of this watershed unit is jointly affected by multiple representative factors, which divide the watershed unit into a weak dominance watershed unit (Figure 3-11). Interestingly, we found that the weak dominance watershed units are predominantly located in the urban core area, which may be due to the complexity of the landscape elements within the urban center.

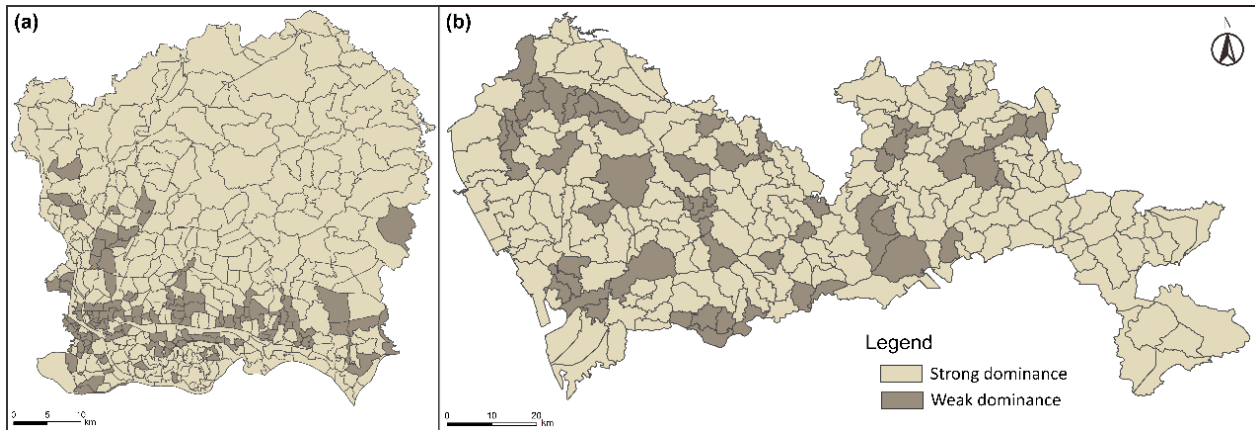


Figure 3-11 The classification of strong dominance and weak dominance watershed unit.

### 3.5.4.2 Interactive effect of representative factors in weak dominance watersheds

After identifying the dominant factors of urban waterlogging in each subgroup, the interaction detector was adopted to further measure the interaction effect of representative factors on urban waterlogging in the weak dominance watershed units (Figure 3-12). It is found that the explanatory power is enhanced by the interaction of the representative driving factors in the weak dominance watershed units, which can be divided into binary enhancement and nonlinear enhancement. This result indicates that the variation of waterlogging in weak dominance watershed units is mainly the result of the interaction of representative factors.

It is worth noting that the intensity of interaction between the representative factors varied significantly among the subgroups. For subgroups 3, 4, 6, and 8 in Guangzhou (weak dominance watershed units), there is a strong interaction between the land cover characteristics (ISA, UGI.MPS) and topographic factors (Dep, DEM, Slope.std), which greatly enhances their independent effects on urban waterlogging. For example, in subgroup 3, the interaction between Dep and UGI.MPS reached 69.98%, which exceeded the independent contribution rate of Dep, the dominant factor of this subgroup, by 39.57%. The interaction between ISA and Slope.std is the strongest in subgroups 4 (72.28%) and 8 (67.21%), which is significantly higher than the independent contribution of the dominant factors in each subgroup. The interactive effect of DEM and ISA in subgroup 6 get 145.46% enhancement compared with the independent effect. These results all indicate the

importance of the interaction between land cover characteristics and topographic factors, which further confirm that waterlogging in these subgroups is mainly affected by the interaction between these two types of factors. Therefore, for these watershed units, on the basis of controlling the ISA, increasing the topographic relief, or reducing the depression will contribute to improving waterlogging states.

For Shenzhen, there is a strong interaction between the land cover composition (ISA) and the landscape configuration (PD, AI). In subgroup 5, the interaction between PD and ISA was as high as 83.72%, which was 180.02% and 258.76% enhancement compared with their independent contributions. Moreover, in subgroup 3, the interaction between AI and PD was 73.02%, which was 35.05% higher than the independent contribution of AI (dominant factors in subgroup 3). This interaction effect emphasizes the importance of the combination of land cover composition and spatial pattern, which is of great significance to metropolises. In addition to the interaction between land cover composition and landscape pattern, the interaction between DEM.std and ISA in Shenzhen subgroup 2 can not be ignored. Compared with the independent contribution of its dominant factor (DEM.std), its interaction effect increased by 82.12%.

It is important to note that the interaction effect of each representative factor in the weak dominance watershed units is significantly stronger than its independent contribution. Therefore, for strong dominance watershed units, the local authorities can develop urban waterlogging mitigation strategies based on its dominant factors. However, for the weak dominance watershed units, it is inappropriate to develop mitigation strategies based on dominant factors as simple as for strong dominance watersheds. We need to comprehensively consider the interaction effect between representative factors, which develop urban waterlogging mitigation strategies that integrate multiple factors according to local conditions.

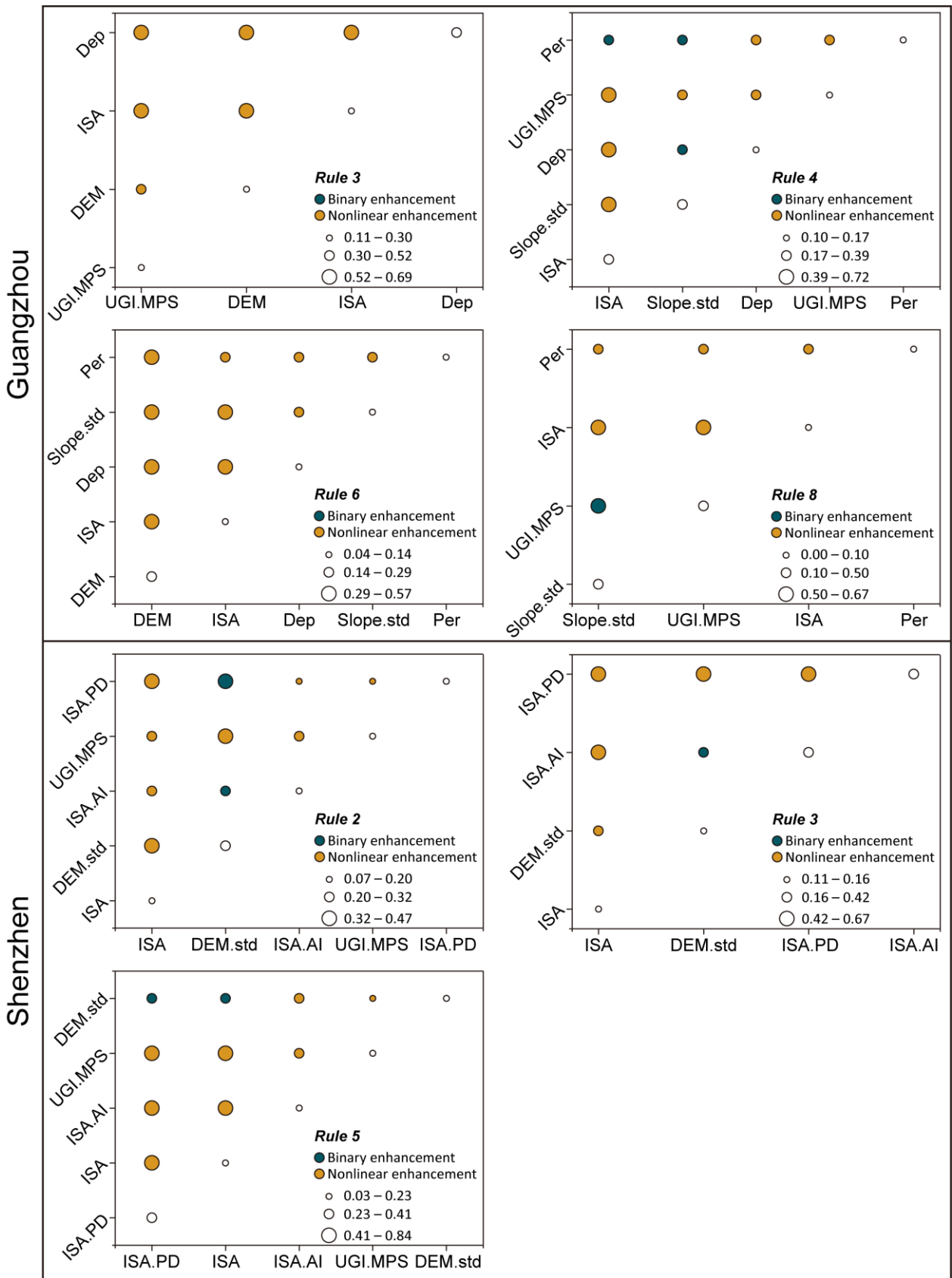


Figure 3-12 The interactive effect of representative factors on urban waterlogging in weak dominance watersheds.

### 3.6 Discussion

### *3.6.1 Select the appropriate driving factors*

As urban waterlogging is largely disturbed by local conditions, which indicates that the driving factors and the mechanisms of waterlogging may vary from place to place. Therefore, the selection of the most representative and meaningful driving factors based on the characteristics of the different study areas becomes the cornerstone of exploring the mechanism of urban waterlogging. Considerable studies select driving factors from both the natural environment and socio-economic aspects respectively (Wu et al., 2012; Liu et al., 2021; Huang et al., 2017; Wu and Zhang, 2017; Tran et al., 2020). For example, an empirical study was conducted on the influence of 27 factors on urban waterlogging in Guangzhou at different spatial scales (Zhang et al., 2020). Wu et al. (2020) selected five landscape pattern metrics to investigate the complex mechanism of urban waterlogging. However, part of the studies relies on subjectively selecting waterlogging driving factors. If these numerous factors are applied directly to the analysis, the multicollinearity problem or overfitting problem will inevitably occur. Moreover, it is difficult to know which driving factor has the most meaningful effect on waterlogging. To overcome this limitation, some scholars have introduced stepwise regression model or correlation analysis to select representative driving factors by eliminating factors with less correlation and multicollinearity. Although the stepwise regression method can "perfectly" meet the requirements of the significance test and multicollinearity, it eliminates most of the driving factors, usually only two or three factors remain. To some extent, the potential explanatory power of various factors is lost or wasted. This suggests that the stepwise regression method only selects the representative driving factors at the statistical aspect. Hence, the selected representative waterlogging driving factors may not be in line with reality.

Therefore, this study proposes an objective method to select representative urban waterlogging driving factors based on local conditions. The combination of the OLS regression and the best subset regression is utilized to explore the impact of all possible combinations of landscape elements on urban waterlogging, which fully considers all the combined models to identify the representative factors. Instead of subjective selection or simply using the stepwise regression model to select representative driving factors for urban waterlogging, the method proposed in this study can more objectively and effectively select representative driving factors. In view of the importance of objectively and quantitatively identifying the representative driving independent variables, the method proposed in this study can be widely applied to other environmental and ecological mechanisms investigation.

### *3.6.2 Spatial non-stationarity effect of driving forces on urban waterlogging*

In reality, urban waterlogging events are characterized by a high degree of spatial heterogeneity, which is usually influenced by many factors. In the past 20 years, a proliferation of studies focused on the driving forces of urban waterlogging (Wu et al., 2020; Zhang et al., 2020; Zhang et al., 2018). However, so far, most studies in exploring this complex mechanism are limited to the use of global single-level analysis methods (Wu and Zhang, 2017; Tran et al., 2021; Sun, 2014). This approach can only analyze the mechanism of urban waterlogging within the entire study area, while inevitably ignoring the influence at a local scale.

Consider the spatial heterogeneity and complexity of the spatial distribution of urban landscape elements. This spatial non-stationarity leads to different effects of landscape elements on urban waterlogging. Based on this, this study successfully revealed the spatial heterogeneity driving forces of urban waterlogging by integrating the cubist regression tree and the geographical detector model. By comparing this method with two other commonly used regression methods, namely GLR and SLM, we found that the proposed method has the highest accuracy that can effectively capture the spatial heterogeneity characteristics of urban waterlogging.

It is worth noting that we find that the driving force of urban waterlogging varies with local conditions. The dominant factors and their independent contribution vary greatly in different subgroups. Although a large number of studies have concluded that the ISA is the dominant factor for urban waterlogging through global statistical methods (Zhang et al., 2018; Wu et al., 2020). This study further spatially reveals the dominant factors of urban waterlogging at the local scale with different background conditions, which highlights the spatial heterogeneity of driving forces. To some extent, these findings indicate that global statistical analysis may not be applicable to highly heterogeneous urban areas, which further prompted us to re-examine the results obtained based on global statistical methods to avoid its shortcomings. The studies based on global statistical methods may only provide a universal and superficial mechanism of waterlogging that can not provide local waterlogging mechanisms. Moreover, the impact of various factors on urban waterlogging is disturbed by local conditions, and even produces contradictory results in different regions. Our study suggests that this inconsistency may be caused by the spatial non-stationarity of landscape elements, which is mainly due to the different ability of landscape elements in different spatial locations to affect urban waterlogging. This finding underscores the importance of revealing the spatial heterogeneity driving forces of urban waterlogging.

### *3.6.3 Site-specific mechanism for urban waterlogging management implication*

A spatial understanding of the mechanism of urban waterlogging will help us to provide theoretical and practical references for waterlogging emergency management, mitigation, and sustainable development. Due to the complexity of urban ecosystems with huge spatial heterogeneous, urban waterlogging is characterized by strong spatial non-stationarity. Indisputably, it indicates that the waterlogging driving forces may vary with local conditions. Therefore, urban waterlogging mitigation strategies developed from the result of global statistical methods may not be applicable in some areas. Strictly speaking, it is unrealistic to develop locally targeted mitigation strategies based on the conclusions drawn from the global statistical method.

In this study, our proposed method successfully characterizes the dominant factors of urban waterlogging in each subgroup at different spatial locations. By understanding the spatially non-stationary effects of landscape elements on urban waterlogging in different watershed units, it will help us to implement more targeted and effective mitigation strategies. Consistent with previous studies, we found that the dominant factor of urban waterlogging in most watershed units is the area proportion of impervious surface (ISA). This is why the previous urban waterlogging mitigation policies mainly focused on reducing the proportion of impervious surfaces. However, our study further found that the ISA was not always the dominant factor in different locations. In some watershed units, topographic factors (DEM, Slope.std, Dep) and land cover characteristics (UGI.MPS, PD, AI) have a much stronger impact on urban waterlogging than ISA (Figure 3-9). Based on these results, we can develop corresponding urban waterlogging mitigation policies according to the local condition of different subgroups in a more targeted way.

(1) In the subgroups where DEM, DEM.std, or Slope.std are regarded as dominant factors (rule 6 and 8 in Guangzhou, rule 2 and 6 in Shenzhen). In these watershed units, we need to pay more attention to the impact of topographical features on urban waterlogging, which highlights the importance of increasing terrain variation. Therefore, local authorities can plan or build more subsidence roadbeds, subsided rain gardens, sunken squares, and ditches in these watershed units, increasing the DEM.std and Slope.std to enhance the function of hydrological regulation and storage.

(2) In the subgroups where Dep are regarded as dominant factors (rule 3 in Guangzhou), this area is the historical urban district of Guangzhou. The elevation of this area is relatively low and gentle. While it is very difficult to alter the topographical conditions, it is recommended that local authorities improve the urban drainage systems design standards in this area by increasing the drainage pipe networks density, pumping



stations, and sluices, so as to improve the efficiency of rainwater drainage. At the same time, the water authorities also need to dredge the rivers to improve their discharge and drainage capacity. Additionally, since surface runoff tends to flow rapidly into the tunnel and cause casualties, transportation in low-lying areas should adopt the viaduct scheme instead of the tunnel scheme.

(3) In the subgroups where ISA are regarded as dominant factors (rule 2, 7, 9 in Guangzhou and rule 1, 4 in Shenzhen). In these watershed units, the ISA is the major cause of urban waterlogging. Therefore, for these watersheds, local authorities need to strictly control the increased rate of impervious surfaces and maintain a reasonable proportion of UGI and ISA.

(4) For the subgroups with ISA.AI and ISA.PD as the dominant factor (rule 3, 5 in Shenzhen). This means that the spatial pattern of ISA also plays a decisive role in urban waterlogging, which embodies the importance of controlling the spatial configuration of ISA. Given a fixed amount of ISA, the aggregated or less fragmented distribution pattern of impervious surfaces will aggravate the occurrence of waterlogging.

(5) For the subgroups with UGI.MPS as the dominant factor (rule 1, 5 in Guangzhou), compared with other measurements, increasing green areas can more effectively alleviate urban waterlogging. Local authorities can introduce green roofs, rain gardens, and ground permeable paving to reduce the peak flow and the pressure of flood drainage for these subgroups. As we all know, UGI has significant effects in intercepting rainfall and reducing runoff, which has a positive effect on alleviating urban waterlogging (Shi, 2020 Li et al., 2020; Kim et al., 2016). Although the benefit of increasing the proportion of UGI is obvious, the land resources for urban greening are often limited. Therefore, it is unrealistic to simply seek a large enough area for urban greening to mitigate urban waterlogging. It is worth noting that Zhang et al. (2021a) found that excessive proportions of UGI within the watershed unit or an oversized UGI patch may lead to a waste of its waterlogging mitigation effect. Therefore, for the watersheds with UGI.MPS is the dominant factor, the local authorities should further weigh the green area and its mitigation effect when formulating policies to increase the area of UGI.

It is worth noting that this study further divided the watersheds into the strong dominant watersheds and the weak dominant watersheds based on the independent contribution of dominant factors (Figure 3-11). In the weak dominant watersheds, the waterlogging magnitude is mainly affected by several representative factors together. For the subgroups 3, 4, 6, and 8 of Guangzhou weak dominant watersheds, the interaction effect between the land cover characteristics (ISA, UGI.MPS) and topographic factors (Dep, DEM, Slope.std)

significantly increases the individual effects of each factor on urban waterlogging. Similarly, for subgroups 2, 3, and 5 as Shenzhen's weak dominant watersheds, the interaction between the composition (ISA) and the landscape configuration (PD, AI) has the dominant effect on waterlogging. It may provide important insights for the mitigation of urban waterlogging. In the strong dominant watersheds, we can develop related mitigation strategies based on the dominant factors in these watershed units. On the other hand, in the weak dominant watersheds, the interaction effects of representative factors must be considered comprehensively when formulating the urban waterlogging mitigation strategy for that watershed unit. This means that urban planners and local authorities should formulate more targeted and effective mitigation strategies based on the driving forces of urban waterlogging in different regions, rather than the "one-size-fits-all" policy.

#### **3.6.4 Limitations**

This research proposes a valuable research framework to explore the spatial heterogeneity driving forces of urban waterlogging. However, this study also has its limitations and should be considered in future work. First, as the obtained data do not record the severity of the hazard (inundation depth, inundation area), the duration, or the specific time of waterlogging events. This inevitably brought some uncertainty to the results. Second, the method proposed in this research has been successfully applied in two highly spatial heterogeneous cities, Guangzhou and Shenzhen. In future studies, we can extend it to more regional comparative studies under different environmental conditions and urbanization backgrounds to further verify the universality and credibility. With the use of sufficient data, introduce the water depth or duration of waterlogging events to more comprehensively investigate the relationship between urban waterlogging and driving forces, so as to provide support for the development of targeted urban waterlogging mitigation policies.

#### **3.7 Conclusion**

Quantitative assessment of the spatial non-stationarity mechanism of waterlogging expands our scientific understanding of the spatial heterogeneity driving forces, which can provide theoretical and practical references for site-specific waterlogging mitigation strategies. In this study, we first utilized the best subset regression model, an objective method to determine the meaningful and representative driving factors of urban waterlogging that are based on the characteristics of local conditions. Then, an innovative approach is proposed to investigate the spatial non-stationarity effects of driving factors on urban waterlogging and map

the dominant drivers in each watershed unit by combining a cubist regression tree model and a geographical detector model. The spatial non-stationarity relationships between urban waterlogging and representative driving factors in each subgroup were spatially explicit as well as the detected dominant driving forces spatially.

It was found that: (1) compared with subjective selection or simply using the stepwise regression method to select representative driving factors for urban waterlogging, the best subset regression proposed in this study can more objectively select the most representative and meaningful driving factors that effectively describe urban waterlogging variation according to local characteristics. (2) By comparing with two other commonly used regression methods (GLR, SLM), the combination of cubist regression tree and geographical detector model can fully quantify the spatial non-stationarity effect of representative driving factors on urban waterlogging and spatially explicit the driving forces with different local conditions. (3) The driving force of urban waterlogging varies with the local site conditions. Understanding the complex site-specific mechanism of urban waterlogging in different watershed units will help us implement more targeted and effective mitigation strategies, rather than a “one-size-fits-all” policy. (4) In the strong dominant watersheds, we can develop related mitigation strategies based on the dominant factors in these watershed units, in contrast to the weak dominant watersheds, the interaction effects of representative factors must be considered comprehensively when formulating the urban waterlogging mitigation strategy for that watershed unit. This research provides an objective and effective tool for urban planners and local authorities to identify the waterlogging spatial heterogeneity driving forces. These new findings facilitate the realization of urban waterlogging mitigation strategies based on different spatial locations.

### **Author Contributions**

**Qifei Zhang:** conceptualization, methodology, formal analysis, data curation, writing - original draft. **Zhifeng Wu:** resources, writing - review & editing, supervision. **Guanhua Guo:** methodology, validation, review & editing. **Paolo Tarolli:** conceptualization, methodology, writing - review & editing, overall supervision, project administration.

## **Acknowledgments**

This study was supported by the NSFC-Guangdong Joint Foundation Key Project (U1901219) and Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (No. GML2019ZD0301). The authors sincerely thank the editor and the two anonymous reviewers for their constructive comments and suggestions.

## **Conflicts of Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Supplementary data

Supplementary materials include the following:

### (1) The identification of depression

The depression is the urban lowland surrounded by higher elevation where surface runoff is easy to accumulate. Therefore, urban waterlogging events frequently occur in depressions, where the stored rainwater easily exceeds the capacity of the urban drainage system. The identification of depression can be delineated using GIS analysis with a high-resolution DEM. The workflow is given in Figure S3-1. Firstly, the original DEM is filled with ArcGIS hydrological analysis module. Secondly, we calculate the difference between the two DEMs (original DEM and filled DEM). The depression can be located by searching for non-zero cells in the derived difference raster. Then, identifying the depressions with the thresholds of depth and area.

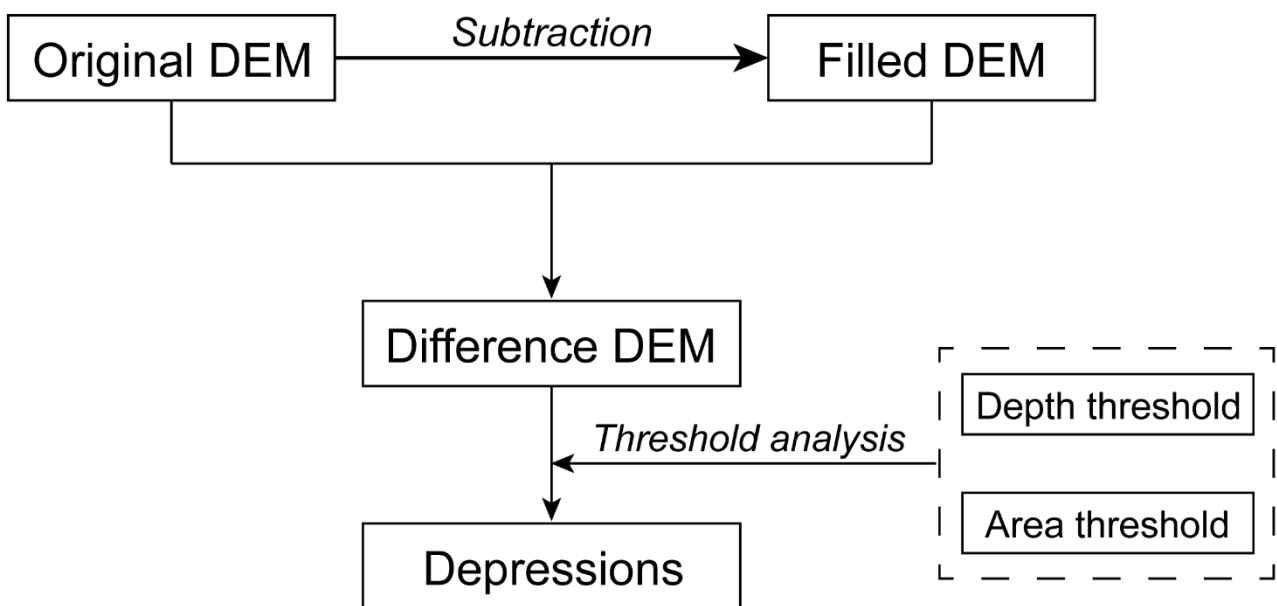


Figure S3-1 The workflow to derive the depression.

The threshold value of depression depth is determined according to the car wading depths. The height of the vulnerable positions such as the air inlet of a car is mostly less than 30 cm. Thus, only depressions with a maximum depth greater than 30 cm were included in the analysis.

The threshold value of the depression area is based on the local conditions of each watershed unit. We test several area thresholds (such as 500 m<sup>2</sup>, 1000 m<sup>2</sup>, 3000 m<sup>2</sup>, 5000 m<sup>2</sup>) in each watershed unit and the number

of depressions corresponding to each area threshold is also counted. We find that as the area threshold increases, the number of depressions gradually decreases. Hence, the turning point on the area-number curve corresponds to the area threshold.

Therefore, the proportion of depression (Dep) describes the area proportion of depression in the watershed that is defined as:

$$Dep = \frac{D_i}{A} \quad (S3.1)$$

where D represents the area of depression in watershed unit i and A denotes the area of watershed unit i.

(2) Divides the urban waterlogging space into different heterogeneity subgroups by rules

Table S3-1 Rules for Guangzhou and Shenzhen.

Guangzhou city
<p><b>Rule 1: [99 cases]</b>            If ISA &lt;= 0.2314            Expression: WD = -0.0974 + 0.63 ISA - 0.03 UGI.MPS + 0.0005 DEM</p>
<p><b>Rule 2: [125 cases]</b>            If Dep &lt;= 0.0282 and ISA &gt; 0.2314 and ISA &lt;= 0.6637            Expression: WD = -0.7701 + 16.84 Dep + 2.32 ISA + 0.0053 DEM - 0.06 UGI.MPS</p>
<p><b>Rule 3: [8 cases]</b>            If Dep &gt; 0.0964 and ISA &gt; 0.6637            Expression: WD = 7.8973 + 0.46 UGI.MPS - 1.07 ISA - 0.35 DEM + 0.91 Dep</p>
<p><b>Rule 4: [21 cases]</b>            If DEM &gt; 4.5947 and DEM &lt;= 5.5781 and Dep &lt;= 0.0964 and ISA &gt; 0.6637            Expression: WD = 28.9743 - 34.37 ISA + 1.17 Slope.std + 0.76 UGI.MPS - 2.42 Dep + 0.0086 Per</p>
<p><b>Rule 5: [34 cases]</b>            If DEM &lt;= 4.5947 and ISA &gt; 0.6637            Expression: WD = 17.0543 - 3.9251 DEM - 0.44 UGI.MPS - 1.375 Dep + 0.0171 Per</p>
<p><b>Rule 6: [74 cases]</b>            If Dep &gt; 0.0282 and ISA &gt; 0.2314 and ISA &lt;= 0.6637            Expression: WD = 70.6914 - 0.0334 Per - 3.3 Dep + 0.15 ISA - 1.25 Slope.std - 1.27 DEM</p>
<p><b>Rule 7: [20 cases]</b>            If DEM &gt; 5.5781 and Dep &lt;= 0.0964 and ISA &gt; 0.6637            Expression: WD = 0.5679 + 9.68 Dep + 0.94 ISA + 0.675 DEM</p>
<p><b>Rule 8: [24 cases]</b>            If ISA &lt;= 0.6637 and Dep &lt;= 0.0282 and UGI.MPS &lt;= 0.7661            Expression: WD = 6.0651 + 2.14 ISA - 0.26 UGI.MPS - 3.32 Slope.std - 0.0028 Per</p>
<p><b>Rule 9: [25 cases]</b>            If ISA &lt;= 0.6637 and Dep &lt;= 0.0282 and UGI.MPS &gt; 0.7661</p>

Expression: $WD = 3.9305 - 5.01 \text{ DEM} + 8.47 \text{ ISA} + 0.28 \text{ UGI.MPS} - 0.41 \text{ Slope.std} + 1.22 \text{ Dep}$
<b>Shenzhen city</b>
<p><b>Rule 1: [110 cases]</b></p> <p>If <math>ISA \leq 0.5560</math></p> <p>Expression: <math>WD = 0.2628 + 0.53 \text{ ISA} + 0.0025 \text{ UGI.MPS} - 0.0035 \text{ ISA.AI} - 0.23 \text{ ISA.PD} + 0.0021 \text{ DEM.std}</math></p>
<p><b>Rule 2: [24 cases]</b></p> <p>If <math>DEM.std \leq 11.7229</math> and <math>ISA &gt; 0.5560</math> and <math>ISA.PD \leq 0.0368</math> and <math>UGI.MPS &gt; 7.8661</math></p> <p>Expression: <math>WD = -2.1749 + 0.1346 \text{ DEM.std} - 27.09 \text{ ISA.PD} + 4.3 \text{ ISA} - 0.0169 \text{ UGI.MPS} - 0.0034 \text{ ISA.AI}</math></p>
<p><b>Rule 3: [10 cases]</b></p> <p>If <math>DEM.std \leq 11.7229</math> and <math>ISA &gt; 0.6430</math> and <math>UGI.MPS \leq 7.8661</math></p> <p>Expression: <math>WD = 0.7646 + 1.37 \text{ ISA} + 0.06767 \text{ DEM.std} - 0.0092 \text{ ISA.AI} - 0.22 \text{ ISA.PD}</math></p>
<p><b>Rule 4: [48 cases]</b></p> <p>If <math>DEM.std &gt; 11.7229</math> and <math>ISA &gt; 0.5560</math></p> <p>Expression: <math>WD = -6.9184 + 2.63 \text{ ISA} - 0.0825 \text{ ISA.AI} - 2.12 \text{ ISA.PD} - 0.0086 \text{ UGI.MPS} + 0.0019 \text{ DEM.std}</math></p>
<p><b>Rule 5: [24 cases]</b></p> <p>If <math>DEM.std \leq 11.7229</math> and <math>ISA &gt; 0.5560</math> and <math>ISA \leq 0.6431</math> and <math>UGI.MPS \leq 7.8661</math></p> <p>Expression: <math>WD = 10.0279 - 10.42 \text{ ISA} - 0.0342 \text{ ISA.AI} - 0.82 \text{ ISA.PD} - 0.0031 \text{ UGI.MPS} + 0.1346 \text{ DEM.std}</math></p>
<p><b>Rule 6: [20 cases]</b></p> <p>If <math>DEM.std \leq 11.7229</math> and <math>ISA.PD &gt; 0.0368</math> and <math>UGI.MPS &gt; 7.8661</math></p> <p>Expression: <math>WD = -0.7601 + 0.0676 \text{ DEM.std} + 0.85 \text{ ISA} + 0.0169 \text{ UGI.MPS} - 0.0825 \text{ ISA.AI}</math></p>

## References

1. Berghuijs, W. R., Aalbers, E. E., Larsen, J. R., Trancoso, R., & Woods, R. A. (2017). Recent changes in extreme floods across multiple continents. *Environmental Research Letters*, 12(11), 114035.
2. Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A., Parajka, J., Merz, B., ... & Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, 573(7772), 108-111.
3. Cepeliauskaite, G., & Stasiskiene, Z. (2020). The framework of the principles of sustainable urban ecosystems development and functioning. *Sustainability*, 12(2), 720.
4. Chen, Y., Zhou, H., Zhang, H., Du, G., & Zhou, J. (2015). Urban flood risk warning under rapid urbanization. *Environmental research*, 139, 3-10.
5. Cushman, S. A., McGarigal, K., & Neel, M. C. (2008). Parsimony in landscape metrics: strength, universality, and consistency. *Ecological indicators*, 8(5), 691-703.
6. Du, S., Van Rompaey, A., & Shi, P. (2015). A dual effect of urban expansion on flood risk in the Pearl River Delta (China) revealed by land-use scenarios and direct runoff simulation. *Natural Hazards*, 77(1), 111-128.
7. Frank, B., Delano, D., & Caniglia, B. S. (2017). Urban systems: A socio-ecological system perspective. *Sociol. Int. J*, 1(1), 1-8.
8. Gaitan, S., & van de Giesen, N. (2015). Spatial distribution of flood incidents along urban overland flow-paths. *Water Resources Management*, 29(9), 3387-3399.
9. Guo, G., Wu, Z., & Chen, Y. (2020). Evaluation of spatially heterogeneous driving forces of the urban heat environment based on a regression tree model. *Sustainable Cities and Society*, 54, 101960.
10. Guo, G., Wu, Z., Cao, Z., Chen, Y., & Zheng, Z. (2021). Location of greenspace matters: a new approach to investigating the effect of the greenspace spatial pattern on urban heat environment. *Landscape Ecology*, 36(5), 1533-1548.
11. Guo, G., Wu, Z., Xiao, R., Chen, Y., Liu, X., & Zhang, X. (2015). Impacts of urban biophysical composition on land surface temperature in urban heat island clusters. *Landscape and Urban Planning*, 135, 1-10.
12. Huang, H., Chen, X., Wang, X., Wang, X., & Liu, L. (2019). A depression-based index to represent topographic control in urban pluvial flooding. *Water*, 11(10), 2115.
13. Huang, H., Chen, X., Zhu, Z., Xie, Y., Liu, L., Wang, X., ... & Liu, K. (2018). The changing pattern of



urban flooding in Guangzhou, China. *Science of The Total Environment*, 622, 394-401.

14. Huang, T., Chen, J., Huang, F., & Su, Z. (2017). A study of urban waterlogging characteristics in Pearl River Delta region based on GIS. *J Guangdong Univ Technol*, 34(1), 24-30.
15. Jian, W., Li, S., Lai, C., Wang, Z., Cheng, X., Lo, E. Y. M., & Pan, T. C. (2021). Evaluating pluvial flood hazard for highly urbanised cities: a case study of the Pearl River Delta Region in China. *Natural Hazards*, 105(2), 1691-1719.
16. Jiang, B. (2015). Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity. *GeoJournal*, 80(1), 1-13.
17. Jiang, Y., Zevenbergen, C., & Ma, Y. (2018). Urban pluvial flooding and stormwater management: A contemporary review of China's challenges and "sponge cities" strategy. *Environmental science & policy*, 80, 132-143.
18. Ke, Q., Tian, X., Bricker, J., Tian, Z., Guan, G., Cai, H., ... & Liu, J. (2020). Urban pluvial flooding prediction by machine learning approaches—a case study of Shenzhen city, China. *Advances in Water Resources*, 145, 103719.
19. Kim, H. W., & Park, Y. (2016). Urban green infrastructure and local flooding: The impact of landscape patterns on peak runoff in four Texas MSAs. *Applied geography*, 77, 72-81.
20. Kim, Y., Eisenberg, D. A., Bondank, E. N., Chester, M. V., Mascaro, G., & Underwood, B. S. (2017). Fail-safe and safe-to-fail adaptation: decision-making for urban flooding under climate change. *Climatic Change*, 145(3), 397-412.
21. Li, H., Ding, L., Ren, M., Li, C., & Wang, H. (2017). Sponge city construction in China: A survey of the challenges and opportunities. *Water*, 9(9), 594.
22. Li, L., Uyttenhove, P., & Van Eetvelde, V. (2020). Planning green infrastructure to mitigate urban surface water flooding risk—A methodology to identify priority areas applied in the city of Ghent. *Landscape and Urban Planning*, 194, 103703.
23. Lin, J., He, X., Lu, S., Liu, D., & He, P. (2021). Investigating the influence of three-dimensional building configuration on urban pluvial flooding using random forest algorithm. *Environmental Research*, 196, 110438.
24. Lin, T., Liu, X., Song, J., Zhang, G., Jia, Y., Tu, Z., ... & Liu, C. (2018). Urban waterlogging risk assessment based on internet open data: A case study in China. *Habitat International*, 71, 88-96.
25. Liu, F., Liu, X., Xu, T., Yang, G., & Zhao, Y. (2021). Driving Factors and Risk Assessment of

Rainstorm Waterlogging in Urban Agglomeration Areas: A Case Study of the Guangdong-Hong Kong-Macao Greater Bay Area, China. *Water*, 13(6), 770.

26. Liu, J., Shao, W., Xiang, C., Mei, C., & Li, Z. (2020). Uncertainties of urban flood modeling: Influence of parameters for different underlying surfaces. *Environmental research*, 182, 108929.
27. Luo, W., Jasiewicz, J., Stepinski, T., Wang, J., Xu, C., & Cang, X. (2016). Spatial association between dissection density and environmental factors over the entire conterminous United States. *Geophysical Research Letters*, 43(2), 692-700.
28. McGarigal, K. (1995). FRAGSTATS: spatial pattern analysis program for quantifying landscape structure (Vol. 351). US Department of Agriculture, Forest Service, Pacific Northwest Research Station.
29. Miller, J. D., & Hutchins, M. (2017). The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *Journal of Hydrology: Regional Studies*, 12, 345-362.
30. Ning, Y. F., Dong, W. Y., Lin, L. S., & Zhang, Q. (2017). Analyzing the causes of urban waterlogging and sponge city technology in China. In *IOP conference series: earth and environmental science* (Vol. 59, No. 1, p. 012047). IOP Publishing.
31. Peng, J., Jia, J., Liu, Y., Li, H., & Wu, J. (2018). Seasonal contrast of the dominant factors for spatial distribution of land surface temperature in urban areas. *Remote sensing of Environment*, 215, 255-267.
32. Pickett, S. T., & Cadenasso, M. L. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science*, 269(5222), 331-334.
33. Quan, R. S., Liu, M., Lu, M., Zhang, L. J., Wang, J. J., & Xu, S. Y. (2010). Waterlogging risk assessment based on land use/cover change: a case study in Pudong New Area, Shanghai. *Environmental Earth Sciences*, 61(6), 1113-1121.
34. Quan, R. S. (2014). Rainstorm waterlogging risk assessment in central urban area of Shanghai based on multiple scenario simulation. *Natural hazards*, 73(3), 1569-1585.
35. Shi, L. (2020). Beyond flood risk reduction: How can green infrastructure advance both social justice and regional impact?. *Socio-Ecological Practice Research*, 2(4), 311-320.
36. Sofia, G., Roder, G., Dalla Fontana, G., & Tarolli, P. (2017). Flood dynamics in urbanised landscapes: 100 years of climate and humans' interaction. *Scientific reports*, 7(1), 1-12.

37. Song, Y., Wang, J., Ge, Y., & Xu, C. (2020). An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data. *GIScience & Remote Sensing*, 57(5), 593-610.
38. Su, M., Zheng, Y., Hao, Y., Chen, Q., Chen, S., Chen, Z., & Xie, H. (2018). The influence of landscape pattern on the risk of urban water-logging and flood disaster. *Ecological Indicators*, 92, 133-140.
39. Sun, Z. (2014). Causal factors of local floods in Beijing central city. *Geo Res*, 33(9), 1668-1679.
40. Tan, X., Wu, X., & Liu, B. (2021). Global changes in the spatial extents of precipitation extremes. *Environmental Research Letters*, 16(5), 054017.
41. Tang, X., Shu, Y., Lian, Y., Zhao, Y., & Fu, Y. (2018). A spatial assessment of urban waterlogging risk based on a Weighted Naïve Bayes classifier. *Science of the total environment*, 630, 264-274.
42. Tehrany, M. S., Jones, S., & Shabani, F. (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena*, 175, 174-192.
43. Tran, D., Xu, D., Dang, V., & Alwah, A. (2020). Predicting urban waterlogging risks by regression models and internet open-data sources. *Water*, 12(3), 879.
44. Walton, J. T. (2008). Subpixel urban land cover estimation. *Photogrammetric Engineering & Remote Sensing*, 74(10), 1213-1222.
45. Wang, C., Du, S., Wen, J., Zhang, M., Gu, H., Shi, Y., & Xu, H. (2017). Analyzing explanatory factors of urban pluvial floods in Shanghai using geographically weighted regression. *Stochastic Environmental Research and Risk Assessment*, 31(7), 1777-1790.
46. Wang, J. F., Li, X. H., Christakos, G., Liao, Y. L., Zhang, T., Gu, X., & Zheng, X. Y. (2010). Geographical detectors - based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science*, 24(1), 107-127.
47. Wang, J. F., Zhang, T. L., & Fu, B. J. (2016). A measure of spatial stratified heterogeneity. *Ecological Indicators*, 67, 250-256.
48. Wang, R., Li, F., Hu, D., & Li, B. L. (2011). Understanding eco-complexity: social-economic-natural complex ecosystem approach. *Ecological complexity*, 8(1), 15-29.
49. Wang, X. (2009). Analysis of problems in urban green space system planning in China. *Journal of forestry Research*, 20(1), 79-82.

50. Wu, J., & Zhang, P. (2017). The effect of urban landscape pattern on urban waterlogging. *Acta Geogr. Sin.*, 72, 444-456.
51. Wu, J., Sha, W., Zhang, P., & Wang, Z. (2020). The spatial non-stationary effect of urban landscape pattern on urban waterlogging: a case study of Shenzhen City. *Scientific reports*, 10(1), 1-15.
52. Wu, X., Yu, D., Chen, Z., & Wilby, R. L. (2012). An evaluation of the impacts of land surface modification, storm sewer development, and rainfall variation on waterlogging risk in Shanghai. *Natural Hazards*, 63(2), 305-323.
53. Xiao, R. B., Ouyang, Z. Y., Zheng, H., Li, W. F., Schienke, E. W., & Wang, X. K. (2007). Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. *Journal of Environmental Sciences*, 19(2), 250-256.
54. Xu, L., Cui, S., Tang, J., Nguyen, M., Liu, J., & Zhao, Y. (2019). Assessing the adaptive capacity of urban form to climate stress: a case study on an urban heat island. *Environmental Research Letters*, 14(4), 044013.
55. Yalcin, E. (2020). Assessing the impact of topography and land cover data resolutions on two-dimensional HEC-RAS hydrodynamic model simulations for urban flood hazard analysis. *Natural Hazards*, 101(3), 995-1017.
56. Yu, H., Zhao, Y., Fu, Y., & Li, L. (2018). Spatiotemporal variance assessment of urban rainstorm waterlogging affected by impervious surface expansion: A case study of Guangzhou, China. *Sustainability*, 10(10), 3761.
57. Yu, H., Zhao, Y., Xu, T., Li, J., Tang, X., Wang, F., & Fu, Y. (2021). A high - efficiency global model of optimization design of impervious surfaces for alleviating urban waterlogging in urban renewal. *Transactions in GIS*.
58. Zhang, H., Cheng, J., Wu, Z., Li, C., Qin, J., & Liu, T. (2018). Effects of impervious surface on the spatial distribution of urban waterlogging risk spots at multiple scales in Guangzhou, South China. *Sustainability*, 10(5), 1589.
59. Zhang, H., Wu, C., Chen, W., & Huang, G. (2017). Assessing the impact of climate change on the waterlogging risk in coastal cities: A case study of Guangzhou, South China. *Journal of Hydrometeorology*, 18(6), 1549-1562.
60. Zhang, L., Liu, W., Hou, K., Lin, J., Song, C., Zhou, C., ... & Wang, Y. (2019). Air pollution exposure associates with increased risk of neonatal jaundice. *Nature communications*, 10(1), 1-9.

61. Zhang, Q., Wu, Z., & Tarolli, P. (2021a). Investigating the Role of Green Infrastructure on Urban WaterLogging: Evidence from Metropolitan Coastal Cities. *Remote Sensing*, 13(12), 2341.
62. Zhang, Q., Wu, Z., Guo, G., Zhang, H., & Tarolli, P. (2021b). Explicit the urban waterlogging spatial variation and its driving factors: The stepwise cluster analysis model and hierarchical partitioning analysis approach. *Science of The Total Environment*, 763, 143041.
63. Zhang, Q., Wu, Z., Zhang, H., Dalla Fontana, G., & Tarolli, P. (2020). Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China. *Journal of Environmental Management*, 271, 110951.

# CHAPTER 4

## 4. Explicit the urban waterlogging spatial variation and its driving factors: the stepwise cluster analysis model and hierarchical partitioning analysis approach<sup>3</sup>

Qifei Zhang<sup>a</sup>, Zhifeng Wu<sup>b,c</sup>, Guanhua Guo<sup>b</sup>, Hui Zhang<sup>d</sup>, Paolo Tarolli<sup>a\*</sup>

- a. Dept. of Land, Environment, Agriculture and Forestry, University of Padova, 35020 Legnaro, PD, Italy.
- b. School of Geographical Sciences, Guangzhou University, 510006 Guangzhou, Guangdong province, China.
- c. Southern Marine Science and Engineering Guangdong Laboratory, 511458 Guangzhou, Guangdong province, China.
- d. College of Surveying and Geo-informatics, North China University of Water Resources and Electric Power, 450046 Zhengzhou, Henan province, China.

---

<sup>3</sup> ***This chapter is a style-edited version of the published article:***

Zhang, Q., Wu, Z., Guo, G., Zhang, H., & Tarolli, P. (2021). Explicit the urban waterlogging spatial variation and its driving factors: The stepwise cluster analysis model and hierarchical partitioning analysis approach. *Science of The Total Environment*, 763, 143041.

***Contributions by the PhD candidate included:*** primary authorship in writing and revision of the entire article; primary responsible for literature review, data collection and quantification, methodology, data processing and analysis, and production of all figures.

## 4.1 Abstract

Urban waterlogging is a hydrological cycle problem that seriously affects human activities and the economy. Characterizing waterlogging variation and its driving factors is conducive to preventing the damage of such disasters. Conventional methods, because of the high spatial heterogeneity and the non-stationary complex mechanism of urban waterlogging, are not able to fully capture the urban waterlogging spatial variation and identify the waterlogging susceptibility areas. A more robust method is recommended to quantify the variation trend of urban waterlogging. Previous studies have simulated the waterlogging variation in relatively small areas. However, the relationship between variables is often ignored, which cannot comprehensively reveal the dominant drivers affecting urban waterlogging. Therefore, a novel approach is proposed that combined the stepwise cluster analysis model (SCAM) and hierarchical partitioning analysis (HPA) within a general framework and verifies the applicability through logistic regression, artificial neural network, and support vector machine. According to the dominant driving factors, different simulation scenarios are established to analyze waterlogging density variation. Results found that the SCAM provides accurate and detailed simulated results both in urban centers where waterlogging frequently occurs and urban fringe with few waterlogging events, which shows an excellent performance with a high classification accuracy and generalization capability. HPA detected that the impervious surface abundance (28.07%), vegetation abundance (20.80%), and cumulate precipitation (16.25%) are the dominant drivers of waterlogging. This result suggests that priority should be given to controlling these three factors to mitigate the risk of waterlogging. It is interesting to note that under different urbanization and rainfall scenarios, the urban waterlogging susceptibility has a considerable variation. The watershed spatial location and watershed characteristics are relevant aspects to be considered in identifying and assessing waterlogging susceptibility, which provides original insights that urban waterlogging mitigation strategies should be developed according to different local conditions and future scenarios.

**Keywords:** urban waterlogging; stepwise cluster analysis model; hierarchical partitioning analysis; natural and anthropogenic factors; scenario simulation

## 4.2 Introduction

Under the background of global climate change and rapid urbanization, waterlogging has become a frequent disaster in Chinese cities (Yu et al., 2018; Huang et al., 2017; Xue et al., 2016). Urbanization has increased the

interaction between human society and the ecological environment, which leads to a series of social-environmental problems (Huang and Shen 2018; Wu et al., 2018). This conversion has changed urban hydrological conditions, thereby reducing water storage capacity and increasing surface runoff. Moreover, the surface roughness of the impervious surface is far lower than vegetation cover (i.e., grassland or forest). Thus the confluence speed of surface runoff will be significantly accelerated, which means that the confluence time of water flow is reduced considerably, as a consequence increasing the pressure of the drainage system (Shuster et al., 2005; Sofia et al., 2017; Chen et al., 2015). These changes led to a direct environmental consequence - increasing the risk of urban waterlogging disasters.

According to the "Statistical Bulletin of Flood and Drought Disasters in China" of the Ministry of Water Resources, an average of 157 cities experienced urban waterlogging from 2006 to 2017 (<http://www.mwr.gov.cn/>). On May 7, 2017, a heavy rainstorm (over 250 mm) occurred in Guangzhou, which affected more than 30,000 people, resulting in a large-scale traffic jam (China Global Television Network). Urban waterlogging prevention has become a prominent shortcoming of China's national flood control, which has severely affected the safety of people's lives and property (Zhang, 2015). Therefore, simulating and predicting the waterlogging variation is conducive to providing useful theoretical and practical references for urban waterlogging prevention, sustainable urban development, and urban planning.

In recent years, the irreversible damage caused by urban waterlogging incidents has highlighted the necessity of implementing waterlogging mitigation measures and management (Ahammed 2017; Sofia et al., 2014; Shao et al., 2016; Zhang et al., 2020; Zhang et al., 2014). In general, characterizing the urban waterlogging variation is conducive to revealing the urban waterlogging prone areas, thereby minimizing waterlogging negative effects (Wang et al., 2012; Miao et al., 2019; Tang et al., 2018a). However, as many researchers have pointed out, urban waterlogging is influenced by the natural environment (precipitation and urban topography) and human activities (land-use change and drainage network) (Wu and Zhang 2017; Zhang et al., 2018b; Su et al., 2018). Furthermore, the landscape heterogeneity leads to non-stationary and non-linear characteristics of urban waterlogging. Thereby, it is difficult to simulate urban waterlogging variation accurately. Generally, methods for characterizing urban waterlogging variation can be summarized into four categories: (1) the multivariate statistical methods, (2) the hydrological and hydrodynamic models, (3) the qualitative model based on expert knowledge, and (4) the machine learning models. For the first method, the multivariate statistical methods (such as stepwise regression model) are widely used to analyze the



impact of various variables on waterlogging (Sofia et al., 2017; Huang et al., 2018; Huang and Shen 2018). However, due to the tremendous landscape heterogeneity, it is difficult to utilize them to simulate the spatial variation of urban waterlogging accurately. Consequently, this method is gradually being replaced by more robust and precise methods. Concerning the second group, the hydrological and hydrodynamic models (such as SWMM, MIKE, HEC-RAS, LISFLOOD-FP) are extensively utilized to simulate the urban waterlogging process (Youssef et al., 2011; Quan 2014; Bisht et al., 2016; Cheng et al., 2017; Li et al., 2016). The storm-water management model (SWMM), as one of the representative hydrological models, is able to analyze various hydrological processes generated by surface runoff (Kai et al., 2017; Burger et al., 2014; Babaei et al., 2018). However, the estimation of runoff from these models are usually based on the empirical estimation or the curve proposed by the Soil Conservation Service (SCS-CN), which may not be sufficient to describe specific differences in complex urban landscapes (Zhang et al., 2018b; Zope et al., 2016). Furthermore, the artificial structures (buildings, roads) or trees will change the direction of the surface runoff, resulting in the complicated water flow movement. This undoubtedly limits the application of hydrological models in urban areas to some extent. To overcome the shortcomings, the two-dimensional hydrodynamic models based on the partial differential equations can better simulate the water flow process under different terrain conditions, which is more suitable for great spatial heterogeneity areas (Tsanis and Boyle 2001; Paiva et al., 2011; Felder et al., 2017). However, these models rely heavily on high-precision local data (such as high-resolution DTM/DEM data and drainage network) and a large amount of computing resources, resulting only suitable for small research areas. Although these models can accurately simulate the physical process of waterlogging in a relatively small catchment, findings in small areas tend to be site-specific, and may not be useful for large-scale studies. Concerning the third category, the qualitative models such as the analytic hierarchy process (AHP) and multi-criteria decision analysis (MCDA) strongly depend on expert knowledge (Tang et al., 2018b; Samanta et al., 2016; Brito et al., 2019). These qualitative models use the AHP to determine factor weights or integrate explanatory factors into a multicriteria sensitivity map to simulate urban waterlogging events (Zhao et al., 2018; Chowdary et al., 2013). For example, Hong et al. (2018) used the hierarchy process and fuzzy weight evidence to construct a flood susceptibility map. However, some studies have pointed out that the methods rely on expert knowledge and judgment, which introduces uncertainty into analysis. For the fourth group, the machine learning models, such as an artificial neural network (ANN), support vector machine (SVM), decision tree (DT), and random forest (RF), have become a common method for urban waterlogging simulation, susceptibility modeling, and risk assessment (Pradhan,

2012). These methods are regarded as a black box to map the relation between input and output of training samples, which shows advantages in complex data modeling. For example, Gupta et al. (2017) identified the urban waterlogging sensitive areas and predicted the severity using an ANN, which indicated that this method could effectively and accurately predict the severity of waterlogging. In Johor River Basin, Malaysia, Kia et al. (2012) integrated the ANN and GIS for flood simulation. The study conducted in Beijing also demonstrates that the ANN is suitable for urban waterlogging risk assessment (Lai et al., 2017). For the SVM, Tang et al. (2019) applied a particle swarm optimization and an SVM in an integrated approach to evaluate the urban waterlogging susceptibility. In Terengganu, Malaysia, different kernel types of SVM models were used to assess the risk of urban floods (Tehrany et al., 2015). Furthermore, Tehrany et al. (2013) also applied a rule-based decision tree to predict the flood susceptible areas in the Kelantan River basin. The results of these studies have confirmed to some extent that the ANN and SVM are effective in simulating urban waterlogging. However, these models are sensitive to the quality of the sample data. If the value of validation samples exceeds the range of training samples, the accuracy of the model will be greatly affected, resulting in poor performance. Furthermore, in the case of a very huge sample size, these models require considerable time consuming and additional modeling parameters.

Compared with the above methods, a more transparent structure and computational efficiency model is needed for urban waterlogging analysis and simulations. The stepwise cluster analysis model (SCAM) is a non-parametric statistical method based on multivariate analysis of variance, which has no statistical assumptions and can process data from different measurement scales. This method is a machine learning paradigm based on a cluster tree, which can be used for multivariate modeling (multiple  $x$  and  $y$ , i.e., supervised learning) and clustering (i.e., unsupervised learning). It has important advantages in investigating the inherent non-linear/discrete relationship between the dependent and independent variables (Fan et al., 2015; Wang et al., 2015; Sun et al., 2019; Wang et al., 2013). Compared to other methods such as ANN and SVM, it performs the more transparent structure of a cluster tree to reflect the complex relationship between the dependent and independent variables. Furthermore, it is also advantageous in the prediction process, which can predict the given independent sample sets or an individual case without given linear/non-linear functions. The studies of Fan et al. (2016), Sun et al. (2019), Zhuang et al. (2016), and Sun et al. (2018) are good examples of using the SCAM for climate prediction, streamflow forecast, and urban ecosystem variation simulation which indicated that SCAM is useful in this respect. However, few studies have applied SCAM for urban waterlogging variation assessment. Therefore, considering the spatial non-stationary and complexity

of urban waterlogging, we attempt to propose the SCAM in this study to capture the spatial variation characteristics of such events. Furthermore, although the machine learning model can obtain accurate analysis results, it cannot provide the relationship between urban waterlogging and various influencing factors. This weakness ignores the relationship between variables, which makes the structure and performance of the model difficult to understand. The quantitative relationship between natural/anthropogenic factors and urban waterlogging is not fully understood. We do not know how much each input factor affects the model. To what extent do these factors contribute to urban waterlogging? Which factors have the dominant effect on waterlogging variation? Answering these questions is essential to provide theoretical references for better waterlogging management and future urban planning. Therefore, while using the SCAM to simulate urban waterlogging spatial variation, we innovatively introduce hierarchical partitioning analysis (HPA) to help us understand the mechanism of urban waterlogging. As a complementary analysis, the HPA reveals the relative contribution of each SCAM input variable, which provides insights into which factor is more critical in determining waterlogging. Furthermore, according to these dominant factors, different simulation scenarios were established to predict the response of waterlogging under the change of waterlogging dominant driving factors, which could provide practical suggestions for urban waterlogging risk identification and control.

The current research aims to explicit the urban waterlogging variation and identify the dominant drivers using the SCAM and ensemble the statistical method of HPA. Then the proposed method will be applied to the example of a highly urbanized coastal metropolis— Guangzhou, P.R. China, selected as a useful case study where waterlogging events frequently occur. In details, the specific objectives are: (1) develop the SCAM and HPA, verify its applicability through accuracy verification indicators (overall accuracy, mean absolute percentage error, modify relative error, correlation coefficient, and Nash-Sutcliffe model efficiency index) and three comparative analysis models: logistic regression (LG), artificial neural network (ANN), and support vector machine (SVM); (2) reveal the spatial variation characteristics of urban waterlogging and clarify the dominant drivers that are deriving waterlogging variations; (3) establish urban waterlogging simulation scenarios to simulate the waterlogging spatial variation under the change of dominant factors, and identify the urban waterlogging susceptibility areas under different development scenarios. This study is expected to provide useful information for further application of the SCAM and HPA in urban waterlogging simulation and assessment as well as bring inspiration to waterlogging risk prevention.

### 4.3 Study area

We selected the central urban districts of Guangzhou (Liwan, Yuexiu, Tianhe, Haizhu, Baiyun, and Huangpu district) as the study area, with a total terrestrial area of 1166.37 km<sup>2</sup>. The Guangzhou city (112°57'~114°3'E and 22°26'~23°56'N) is the central city of the Guangdong-Hong Kong-Macao Greater Bay Metropolitan Region, which is known as one of China's three largest cities with a population of 14.90 million in 2019. This city is also one of the most economically dynamic cities in mainland China. In 2019, the gross domestic product (GDP) of Guangzhou is US\$ 338.23 billion (<http://www.stats.gov.cn/>). It has an annual precipitation of 1623.6~1899.8 mm and 189 precipitation days (<http://www.tqyb.com.cn/>). The rainy season in Guangzhou is spanning from April to September, mainly accounting for about 85% of the annual precipitation, which leading serious urban waterlogging problems in Guangzhou. During the last 40 years, a large number of green areas have been converted into construction land. The rapid urbanization has greatly changed the original hydrological conditions, resulting in a sharp increase in the risk of urban waterlogging. Generally, Guangzhou city is a typical representative of rapid urbanization and industrialization since the “reform and opening up” policy. Therefore, selecting this area to understand the urban waterlogging variation and dominant drivers can not only provide suggestions and guidance to alleviate urban waterlogging, but also provide theoretical bases for other rapidly urbanized regions.

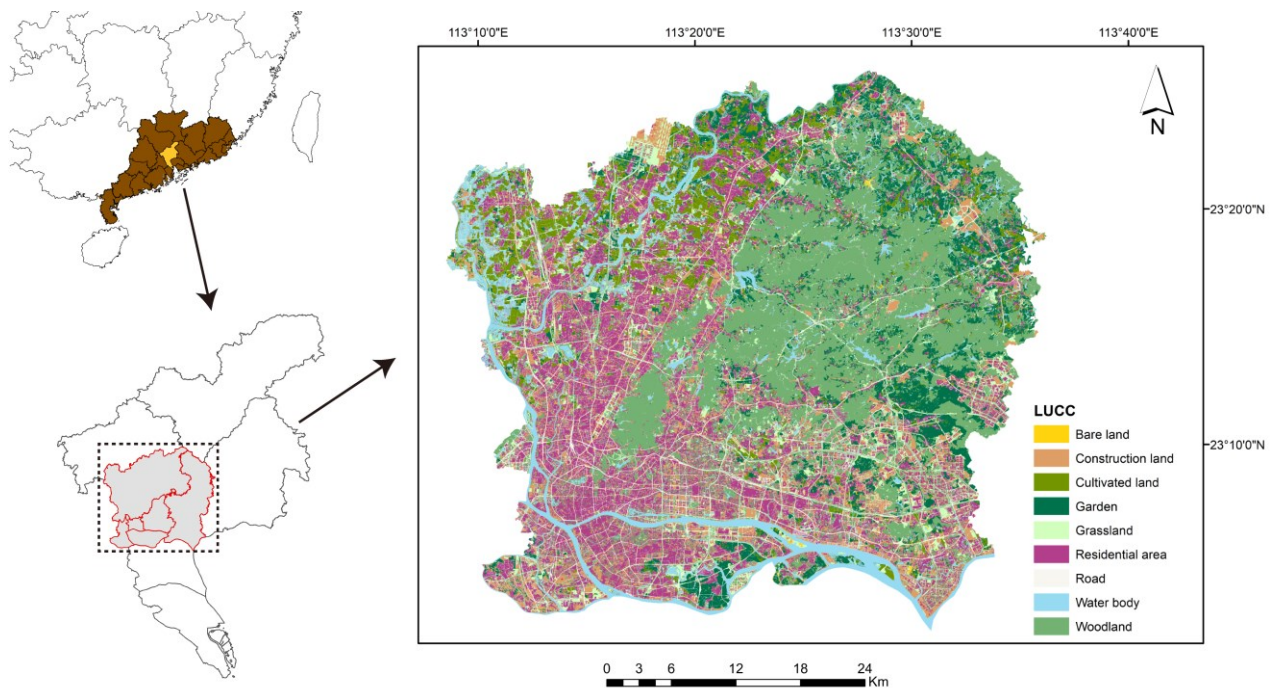


Figure 4-1 The geographic location of Guangzhou Central Urban Districts.

## 4.4 Data and methodologies

### 4.4.1 Data sources

The data in this study included urban waterlogging records, topography (i.e., elevation, slope), precipitation, land cover composition, spatial configuration, gross domestic product (GDP), and drainage facilities. The waterlogging records were used to construct a waterlogging inventory map, while the remaining datasets were utilized as natural and anthropogenic factors of urban waterlogging. We first mapped the spatial location of waterlogging events through Google Earth according to the recorded addresses. Secondly, to have the same projected coordinate system and spatial extent as other spatial datasets, the Google maps with the waterlogging locations were calibrated through the Georeferencing tool in ArcGIS, and the calibrate error is less than 50 m. Finally, from the period of 2009 and 2015, the urban waterlogging events in the study area consisted of 423 urban waterlogging points (Fig.S4-1).

Moreover, the high-resolution DEM (spatial resolution 5 m, vertical accuracy 0.1 m) was acquired from the Geographical Situation Survey Project (GSSP) in 2012. The land cover data (spatial resolution 0.5 m) and urban drainage network were obtained from Guangzhou Planning and Natural Resources Bureau. Subsequently, the spatial distributions of precipitation and gross domestic product were gained from the Resources and Environmental Data Cloud Platform, the Chinese Academy of Sciences.

*Table 4-1 Overview of data used in this study.*

Data	Format	Resolution	Time	Source
Urban waterlogging records	Esri shapefile	Point	2009~2015	Guangzhou Water Authority ( <a href="http://swj.gz.gov.cn/">http://swj.gz.gov.cn/</a> )
Digital Elevation Model	Raster	5 m	2012	Guangzhou Planning and Natural Resources Bureau
Land cover data	Raster	0.5 m	2012	Resources Bureau ( <a href="http://ghzyj.gz.gov.cn/">http://ghzyj.gz.gov.cn/</a> )
Urban drainage network	Esri shapefile	Line	2012	Resources and Environmental Data Cloud Platform ( <a href="http://www.resdc.cn/">http://www.resdc.cn/</a> )
Precipitation	Raster	1 km	2009~2015	Resources and Environmental Data Cloud Platform ( <a href="http://www.resdc.cn/">http://www.resdc.cn/</a> )
Gross domestic product	Raster	1 km	2009~2015	Cloud Platform ( <a href="http://www.resdc.cn/">http://www.resdc.cn/</a> )

### 4.4.2 Watershed unit

The rainwater imbalance is a necessary condition for the occurrence of urban waterlogging. If the rainwater inflow into the watershed exceeds the rainwater discharge capacity, the urban waterlogging event will eventually occur. Using a raster or grid as an analysis unit has the advantages of rapid computation and easy operation (Tang et al., 2019; Zhang et al., 2018a; Zhao et al., 2019; Tehrany et al., 2019; Hong et al., 2018). However, urban waterlogging is a hydrological cycle problem, which is significantly related to the watershed surface characteristics (geomorphology, hydrological conditions, land cover features, surface runoff generation mechanism). It is not suitable to analyze it only using a specific raster unit but instead studied from the perspective of the watershed unit. If grid units are simply used as the analysis units, the effect of watershed geomorphology and hydrodynamic would inevitably be ignored, which undoubtedly brought some biases to the research results. Therefore, in this study, we decided to introduce the watershed as the analysis unit to improve the reliability of the SCAM in characterizing the urban waterlogging variation.

In the field of urban hydro-ecology, scholars have proposed the concept of the "watershed-land" unit, which combines the hydrological watershed unit and urban planning unit (Yu et al., 2018; Zhang et al., 2007). In this study, we generated the watershed unit, which referred to the method recommended by Yu et al. (2018). We first extracted the main drainage channels according to the river distribution and urban drainage network. Then, in the hydrological analysis module of ArcGIS-Pro, the DEM data with a spatial resolution of 5 m was utilized to divide the watershed units. The watershed division was made using the D8 algorithm, which was more efficient in terms of computational analysis at larger scales after comparing it with the D-Infinity algorithm and MFD flow method. Finally, the study area was divided into 468 urban watershed units with an average area of 3.4 km<sup>2</sup> (Fig.S4-1). Based on this, the waterlogging density of each watershed unit could be determined.

#### *4.4.3 Determination of explanatory factors*

In this section, we introduced the fifteen natural and anthropogenic variables (7 natural variables, 8 anthropogenic variables) which were used as the explanatory variables. The natural variables included elevation, elevation standard deviation (elevation.std), relative elevation (RE), slope, slope standard deviation (slope.std), topographic roughness index (TRI), and precipitation (Pre). The anthropogenic variables were composed of vegetation abundance (GV), impervious surface abundance (%ISA), mean patch size (MPS), patch density (PD), largest patch index (LPI), aggregation index (AI), urban drainage density (DD), and gross domestic product (GDP).

#### 4.4.3.1 Topographic and precipitation variables

In this study, the topographic variables were generated from the 5 m DEM data (Fig.S4-2). The elevation, relative elevation (RE), elevation.std, slope, and slope.std were calculated from ArcGIS, while the TRI index was calculated in SAGA GIS software. The RE represents the difference between the maximum elevation and the minimum elevation of each watershed unit.

$$RE = E_{max} - E_{min} \quad (4.1)$$

where  $E_{max}$  and  $E_{min}$  represent the maximum and minimum elevation in watershed unit  $i$ . When RE value is large, it indicates that the elevation fluctuation in the watershed unit is large. In contrast, the elevation fluctuation is relatively gentle. Since the elevation of the terrain always had some sudden changes, the standard deviation of elevation was added to further reflect the fluctuation degree of elevation. Similarly, the slope.std was used to measure the slope fluctuation of each watershed unit. The topographic roughness index (TRI) is a topographic index to quantify the terrain fluctuation between eight adjacent grids (Riley et al., 1999; Werner et al., 2005).

$$TRI = Y \left[ \sum (x_{ij} - x_{00})^2 \right]^{1/2} \quad (4.2)$$

where  $x_{ij}$  is the elevation of each neighbor cell to cell (0,0). Continuous heavy rainfall has a certain effect on urban waterlogging. However, the data obtained from the Guangzhou Water Authority did not list the specific year of each waterlogging event from 2009 to 2015. Therefore, we calculated the average accumulated precipitation (Pre) from 2009 to 2015 to represent precipitation intensity, thereby indicating which regions received more rainfall.

#### 4.4.3.2 Land cover characteristics

In this study, two land cover composition indicators were established, i.e., vegetation abundance (%GV) and impervious surface abundance (%ISA) based on the 0.5 m UAV images, obtained from the Geographical Situation Survey Project (GSSP) in 2012 (Fig.S4-3). The object-oriented classification method was utilized to extract vegetation abundance and impervious surface abundance. The overall accuracy of vegetation abundance and impervious surface abundance exceeds 85%. Finally, the zonal statistics tool was utilized to calculate the proportion of vegetation coverage and impervious surface in each watershed unit. In this study, we investigated the land cover spatial configuration characteristics by calculating the landscape pattern metrics from the perspective of fragmentation and aggregation. Four landscape metrics were contained

including: mean patch size (MPS), patch density (PD), largest patch index (LPI), and aggregation index (AI). The MPS and PD show the average area size and patch density of land cover features respectively, which can indicate the fragmentation of land cover features. The smaller MPS or the greater PD indicates the higher landscape fragmentation. Then, the LPI and PD were utilized to quantify the landscape aggregation. The larger LPI and AI indicate a higher aggregation degree. These metrics can reflect the overall landscape pattern features comprehensively and the redundancy is relatively small. The computing equations and meaning of these landscape pattern metrics were specified in Table 4-2. Fragstats 4.2 was used to calculate these metrics.

Table 4-2 The selected landscape pattern metrics.

Landscape metrics	Equation	Meaning
Mean patch size (MPS)	$\sum_{i=1}^n \frac{s_i}{n}$	The average size of land cover features within a watershed unit
Patch density (PD)	$\frac{s_i}{S}$	The number of patches is divided by the total landscape area. A bigger value means higher fragmentation.
Largest patch index (LPI)	$\max(s_i) \times \frac{100}{S}$	The largest land cover patch within an analysis unit, which can measure the dominance of land cover types
Aggregation Index (AI)	$\left[ \begin{matrix} g_i \\ \max \rightarrow g_i \end{matrix} \right]$	Describe the degree of agglomeration or extension trend of different landscape types.

\* $s_i$  is the area of patch  $i$ ,  $S$  is the total area of all patches,  $n$  is the number of patches,  $g_i$  is the number of adjacent patches

#### 4.4.3.3 Urban drainage network and gross domestic product

We calculated the drainage density (DD) to measure the impact of drainage facilities on urban waterlogging (Fig.S4-4). According to the drainage network data, the line density in the ArcGIS pro was adopted to compute the density of the drainage network in each search radius.

$$DD = \frac{\sum_{i=1}^i (L_i \times V_i)}{A} \quad (4.3)$$

where  $A$  is the area of search radius,  $L_i$  indicates the length of each drainage network within the search radius, and the corresponding weight is  $V_i$ . Finally, the DD is normalized from 0 to 1.

Previous studies had shown that urban waterlogging was affected by environmental conditions and human activities. The GDP could reflect the economic vitality of a region, which could assess the interference of human activities on the natural environment to a certain extent. Therefore, in this study, we utilized the GDP



spatial distribution data obtained from the REDCP to reflect the intensity of human activities. The data was based on national GDP records, land cover features, and nighttime light data, and then a multi-factor weight distribution method is utilized to realize the spatialization of GDP. However, due to the accessibility of GDP spatial distribution data, we only obtained the data in 2010 and 2015. Then we averaged and normalized the GDP data from 2010 to 2015.

#### 4.4.4 Model descriptions and implementations

##### 4.4.4.1 Explanatory variables selection

In modeling urban waterlogging variation, the variables in the selected dataset might be correlated with each other, and such multicollinearity would lead to the instability of the analysis results. It was preferred to select a predictor subset that is highly correlated with the waterlogging and low multicollinearity. Therefore, we used the correlation feature selection (CFS) in the R statistical package of “RWeka” to avoid the multicollinearity problem and to filter out a set of effective predictors for SCAM. The CFS performed a forward or backward search in the predictor subsets and evaluated the value of subsets by considering the individual predictive power and the degree of redundancy. As shown in Table 4-3, we selected these effective variables as the input of prediction factors (i.e., x), and the density of waterlogging event was used as the output factor (i.e., y).

Table 4-3 The effective variables and correlation feature selection result.

Categories	Selected variables	Correlation feature selection
Land cover composition	%ISA	0.6762
	%GV	0.6528
Topographic factors	RE	0.4728
	slope.std	0.4124
	TRI	0.3346
Land cover spatial configuration	GVLPI	0.3523
	GVMPS	0.2491
	ISAMPS	0.2355
Precipitation and drainage facilities	Pre	0.5430
	DD	0.4567

\* GVLPI means the largest patch index for vegetation abundance, the GVMPS and ISAMPS are the mean patch size for vegetation abundance and impervious surface abundance

#### 4.4.4.2 Stepwise cluster analysis modeling approach

The concept of the SCAM is based on the multivariate analysis of variance (Wang et al., 2013). It divides the dependent variables into a cluster tree through a series of cutting or merging actions (Fan et al., 2015). The cutting action is dividing one group into two subgroups, while the merging action is to combine two subgroups. The framework of the SCAM is shown in Fig.S4-5. The establishment of SCAM includes the following parts: (1) construct the effective variables matrix (x) and urban waterlogging density matrix (y); (2) cutting or merging the waterlogging density variables to generate a cluster tree according to the condition of input factors; (3) selecting the best cluster tree bases on the model performance; (4) verifying the applicability of SCAM.

In the SCAM, the criteria for cutting or merging action for cluster trees are based on Wilks' likelihood-ratio  $\Lambda$ , which are defined as follows:

$$\Lambda = \frac{|Q|}{|Q + W|} \quad (4.4)$$

where Q and W represent the sums of squares and cross product (SSCP) matrix within and between groups, respectively. Also, we assume that there are two subclusters c and v, which contain  $n_c$  and  $n_v$  samples. Thus, it can be represented as:  $c_i = (c_{i1}, c_{i2}, c_{i3}, \dots, c_{id})$ ,  $i = 1, 2, 3, \dots, n_c$  and  $v_j = (v_{j1}, v_{j2}, v_{j3}, \dots, v_{jd})$ ,  $j = 1, 2, 3, \dots, n_v$ . Then, the Q and W can be calculated by Eq(4.5-4.8):

$$Q = \sum_{i=1}^{n_c} (c_i - \bar{c})' (c_i - \bar{c}) + \sum_{j=1}^{n_v} (v_j - \bar{v})' (v_j - \bar{v}) \quad (4.5)$$

$$W = \frac{n_c \times n_v}{n_c + n_v} (\bar{c} - \bar{v})' (\bar{c} - \bar{v}) \quad (4.6)$$

$$\bar{c} = \frac{1}{n_c} \sum_{i=1}^{n_c} c_i \quad (4.7)$$

$$\bar{v} = \frac{1}{n_v} \sum_{j=1}^{n_v} v_j \quad (4.8)$$

Previous research results show that when the calculation result of the formula (4.4) is the minimum value, which means that the cutting action of the dependent variable - waterlogging density is the best (Huang, 1992). However, if the formula (4.4) result is huge, which means that the SCAM cannot perform a cutting operation. Thus, the SCAM should perform a merging action to merge the subclusters c and v into the upper-class g. According to Rao's F-statistic, we have:

$$R = \frac{1 - \Lambda^{1/s}}{\Lambda^{1/s}} \times \frac{Z \times S - P \times \frac{(K-1)}{2} + 1}{P \times (K-1)} \quad (4.9)$$

$$Z = n_h - 1 - (P + K)/2 \quad (4.10)$$

$$S = \frac{P^2 \times (K-1)^2 - 4}{P^2 + (K-1)^2 + 5} \quad (4.11)$$

The equation (4.9) obeys the degree of freedom, while K and P are the number of clusters and dependent variables, respectively. In this case, K = 2 (c and v clusters in this study), the Wilks' lambda statistic  $\Lambda$  is performed:

$$F(d, n_c + n_v - d - 1) = \frac{1 - \Lambda}{\Lambda} \times \frac{n_c + n_v - d - 1}{d} \quad (4.12)$$

Therefore, the cutting and merging criteria of the cluster tree become based on the F-test. The first step is the cutting of subclusters. To get the optimal cutting action, the sequence statistical indicator  $\{K^r\}$  is calculated according to the total sample dispersion matrix and inter-group sample dispersion matrix for the dependent variable (cluster g).

$$b_{ij}(k^r, n_g^r) = \frac{n_g^r k^r \times \left\{ \left[ B_i^{(g)}(k^r) - B_i^{(g)}(n_g^r) \right] \times \left[ B_j^{(g)}(k^r) - B_j^{(g)}(n_g^r) \right] \right\}}{n_g^r - k^r} \quad (4.13)$$

$$t_{ij}(n_g^r) = A_{ij}^{(g)}(n_g^r) - n_g^r B_i^{(g)}(n_g^r) B_j^{(g)}(n_g^r) \quad (4.14)$$

$$w_{ij}(k^r, n_g^r) = t_{ij}(n_g^r) - b_{ij}(k^r, n_g^r) \quad (4.15)$$

For each independent variable, the independent variable index  $r^*$  for cutting action judgment is obtained. Thus, we can obtain the optimal cutting point  $K^{*r}$  for cluster g and the corresponding independent variable. If F-test is satisfied:

$$F(P', n_g^{r*} - P' - 1) = \frac{1 - \Lambda(k^{*r}, n_g^{r*})}{\Lambda(k^{*r}, n_g^{r*})} \times \frac{n_g^{r*} - P' - 1}{P'} \geq F_1 \quad (4.16)$$

the cluster g can be divided into two subclusters (c and v) according to the distribution of  $x_{r^*}$ . In contrast, if the equation (4.16) is not satisfied, which means that the cluster g cannot divide into two subclusters c and v. Then the model will test whether the other clusters meet the criteria of equation (4.16). When none of the subclusters can continue to be split, the next step will take the merging action.

The second step is to merge the subclusters. To test whether the subclusters c and v can be combined, we first calculate the following total dispersion matrix and inter-group sample dispersion matrix.

$$t_{ij}(n_c, n_v) = A_{ij}^{(c)}(n_c) + A_{ij}^{(v)}(n_v) - \frac{n_c B_i^{(c)}(n_c) + n_v B_i^{(v)}(n_v)}{n_c + n_v} - \frac{n_c B_j^{(c)}(n_c) + n_v B_j^{(v)}(n_v)}{n_c + n_v} \quad (4.17)$$

$$b_{ij}(n_c, n_v) = \frac{n_c n_v [B_i^{(c)}(n_c) - B_i^{(v)}(n_v)] \times [B_j^{(c)}(n_c) - B_j^{(v)}(n_v)]}{n_c + n_v} \quad (4.18)$$

$$w_{ij}(n_c, n_v) = t_{ij}(n_c, n_v) - b_{ij}(n_c, n_v) \quad (4.19)$$

Then F-test will be carried out if it satisfied the following requirements:

$$F(P', n_c + n_v - P' - 1) = \frac{1 - \Lambda(n_c + n_v - 2, 1)}{\Lambda(n_c + n_v - 2, 1)} \times \frac{(n_c + n_v) - P' - 1}{P'} < F_2 \quad (4.20)$$

then the subclusters c and v can be merged into a cluster g. In contrast, if the equation (4.20) is not satisfied, which means that the subclusters c and v cannot merge into a new cluster. Then the model will consider whether other subclasses can be combined. When cutting and joining action cannot be performed, we have completed all calculations and tests to obtain a cluster tree for each dependent variable. In the process of prediction, it follows the criteria of the cluster tree just established. When predicting a new sample, it is necessary to compare the prediction values with sample value at each intermediate node of the cluster tree. At this time, the predicted value  $y_e$  of the new sample  $e'$  dependent variable can be expressed as:

$$y_i^{(e')} = \frac{1}{n_{e'}} \sum_{k=1}^{n_{e'}} y_{i,k}^{(e')} \quad (4.21)$$

Thus, the urban waterlogging disasters in each watershed simulation can be developed.

#### 4.4.4.3 The hierarchical partitioning analysis

The natural and anthropogenic influencing factors are interdependent and interactive due to the complex mechanisms of urban waterlogging. The HPA was performed to explore the importance of model input factors on urban waterlogging variation. This was effective for clarifying the individual effect of each influencing factor on waterlogging magnitudes. To consider all possible combinations of the factors, the R statistical package of "R-Leaps" was utilized to construct all subsets regression (<https://cran.r-project.org/>). If the input variables of SCAM are n, then n all subsets regression models will be established to employ in the HPA. Then the HPA calculated the relative importance of each input factor by a random transformation of the data matrix of input factors (x matrix) and urban waterlogging density (y matrix). Under the background of great heterogeneity, the HPA is a useful way to determine which variables have a dominant effect in

deriving urban waterlogging variations. This provided theoretical suggestions for constructing urban waterlogging simulation scenarios. Finally, the HPA was implemented in the R statistical package of “hier.part” and “gtools” (R Core Development Team, 2008; Nally, 2000).

#### 4.4.4.4 Model accuracy verification

In this study, we randomly selected 70% watershed units for model calibration and 30% watershed units for validation. We measured the performance of the model from the perspective of accuracy and efficiency. Firstly, the overall accuracy (OA) was introduced to represent the percentage of the watershed units that were correctly identified as waterlogging units and non-waterlogging units. Secondly, the systematic deviation of simulated waterlogging density was calculated through the mean absolute percentage error (MAPE). Thirdly, the modified relative error (MRE) represented the error between simulation value and actual density for each watershed unit. Then, the correlation coefficient ( $R^2$ ) was considered to indicate the difference between the simulated and actual density of each watershed unit to a certain extent. If the simulated density of each watershed was the same as the observed value, the fitting coefficient would be 1. Finally, the Nash-Sutcliffe model efficiency index (NSE) selected as an evaluation index of model efficiency, which represented the model simulation performance, was calculated according to the following equations:

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.22)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^i \left| \frac{WD_{obs,i} - WD_{sim,i}}{WD_{obs,i}} \right| \times 100\% \quad (4.23)$$

$$MRE = \frac{WD_{sim,i} - WD_{obs,i}}{(|WD_{obs,i} + WD_{sim,i}| + 1)} \times 100\% \quad (4.24)$$

$$R^2 = \frac{(i \sum_{i=1}^i WD_{obs,i} \times WD_{sim,i} - \sum_{i=1}^i WD_{obs,i} \times \sum_{i=1}^i WD_{sim,i})^2}{[i \sum_{i=1}^i (WD_{obs,i})^2 - \sum_{i=1}^i (WD_{obs,i})^2] \times [i \sum_{i=1}^i (WD_{sim,i})^2 - \sum_{i=1}^i (WD_{sim,i})^2]} \quad (4.25)$$

$$NSE = 1 - \frac{\sum_{i=1}^i (WD_{obs,i} - WD_{sim,i})^2}{\sum_{i=1}^i (WD_{obs,i} - \overline{WD_{obs,i}})^2} \quad (4.26)$$

where TP (true positive) is the number of watershed units that correctly classify as experienced urban waterlogging disasters units; TN (true negative) is the number of watershed units that correctly classify as non-occurrence urban waterlogging disasters units; FP (false positive) refers to the number of watershed units that are incorrectly classified to experienced urban waterlogging events units; FN (false negative) is the number of watersheds incorrectly classified as non-occurrence urban waterlogging disasters units.  $WD_{obs}$  is the urban waterlogging density observed in watershed unit  $i$ ,  $WD_{sim}$  is the simulated urban waterlogging

density in watershed  $i$ , and  $\overline{WD_{obs,i}}$  is the average urban waterlogging density observed in all watershed units.

Furthermore, to further evaluate the model performance, we compared the SCAM with three widely used urban waterlogging risk assessment and simulation methods (LG, ANN, and SVM). In this study, the sigmoid function was employed for LG. The sigmoid function as an activation function in LG, which was easy to calculate, understand, and implement. It had been adopted by a large number of studies (Tehrany et al., 2014; Neuhold et al., 2011). We selected the radial basis function (RBF) for ANN and SVM based on previous studies, which documented its superior generalization and approximation ability, fast learning convergence, and could approximate any non-linear function to overcome the local minimum problem (Hong et al., 2016; Tehrany et al., 2019; Pradhan, 2012). Subsequently, these three methods were implemented in MATLAB, while the SCAM was performed in R software (R Core Development Team, 2008). For the RBF-ANN model, the number of neurons in the input layer, hidden layer, and output layer was set to 10, 13, and 1, while the maximum number of epochs was set to 1000. The parameters of gamma and c were set to 0.5 and 1 for RBF-SVM.

#### 4.4.4.5 SCAM-HPA framework

The SCAM and HPA framework include four modules: data collection and processing, model development, model evaluation and mapping, waterlogging analysis and scenario simulation module (Fig.4-2).

In the data collection and processing module, all spatial datasets were transformed into the same projected coordinate system, clipped by the Guangzhou Central Urban Districts boundary, and normalized into [0,1]. Then, the spatial datasets corresponding to each watershed unit were extracted. We found that among the 468 watershed units, 227 watershed units had urban waterlogging, while the remaining 241 units had no urban waterlogging records. Base on the number of waterlogging events in each watershed unit, the waterlogging density of each watershed unit could be determined. Besides, the correlation feature selection (CFS) was conducted to avoid the multicollinearity problem and filter out a set of effective predictors. Finally, ten variables (%ISA, %GV, RE, slope.std, TRI, GVLPI, GVMPS, ISAMPS, Pre, DD) remained from the CFS, which would be entered into the model development. Finally, a database containing urban waterlogging density and its corresponding explanatory variables of each watershed unit was obtained.

In the model development module, the LR, RBF-ANN, SVM, and SCAM were established to simulate the spatial variation of urban waterlogging. We used the randomly repeated method to select the training

datasets and testing datasets (Tang et al., 2019). This method could avoid the subjective influence of sample selection and improper training samples. Firstly, we randomly selected 70% of watershed units as a training dataset, while the remaining 30% was used as a testing dataset. When randomly selected samples, we stipulated that 40% of watershed units were positive samples, and the remaining 60% were negative samples, according to the proportion of watersheds that experienced urban waterlogging disasters and those not experienced urban waterlogging disasters. This ensured that both the training and validation datasets contained a certain number of positive and negative samples. Then, to guarantee the model accuracy, the MAPE and OA were selected as evaluation indicators of the training dataset and testing dataset and were assigned values of 0.5 and 0.6, respectively. If the requirements of evaluation indicators are not reached, the model will iterative select positive and negative samples until the evaluation requirements are met.

Concerning the evaluation and mapping module, the urban waterlogging density of each watershed was simulated through these four models (LR, RBF-ANN, SVM, SCAM). To produce an urban waterlogging density simulation map, we transferred each watershed value to ArcGIS pro according to its serial number. Then, the urban waterlogging density was classified according to the natural breaks classification method. Furthermore, the simulation results of these four models were also compared and evaluated by MRE, NSE, and  $R^2$  indicators.

In the waterlogging analysis and scenario simulation module, the dominant factors affecting waterlogging magnitudes were clarified using all the possible combination models of the factors that passed the correlation feature selection. Based on the HPA, we identify the dominant factors that are deriving waterlogging variations. Concerning the scenario simulation module, according to the two most important driving factors (i.e., ISA% and Precipitation), the land-use change scenarios and different rainstorm scenarios were constructed to predict waterlogging density variation.

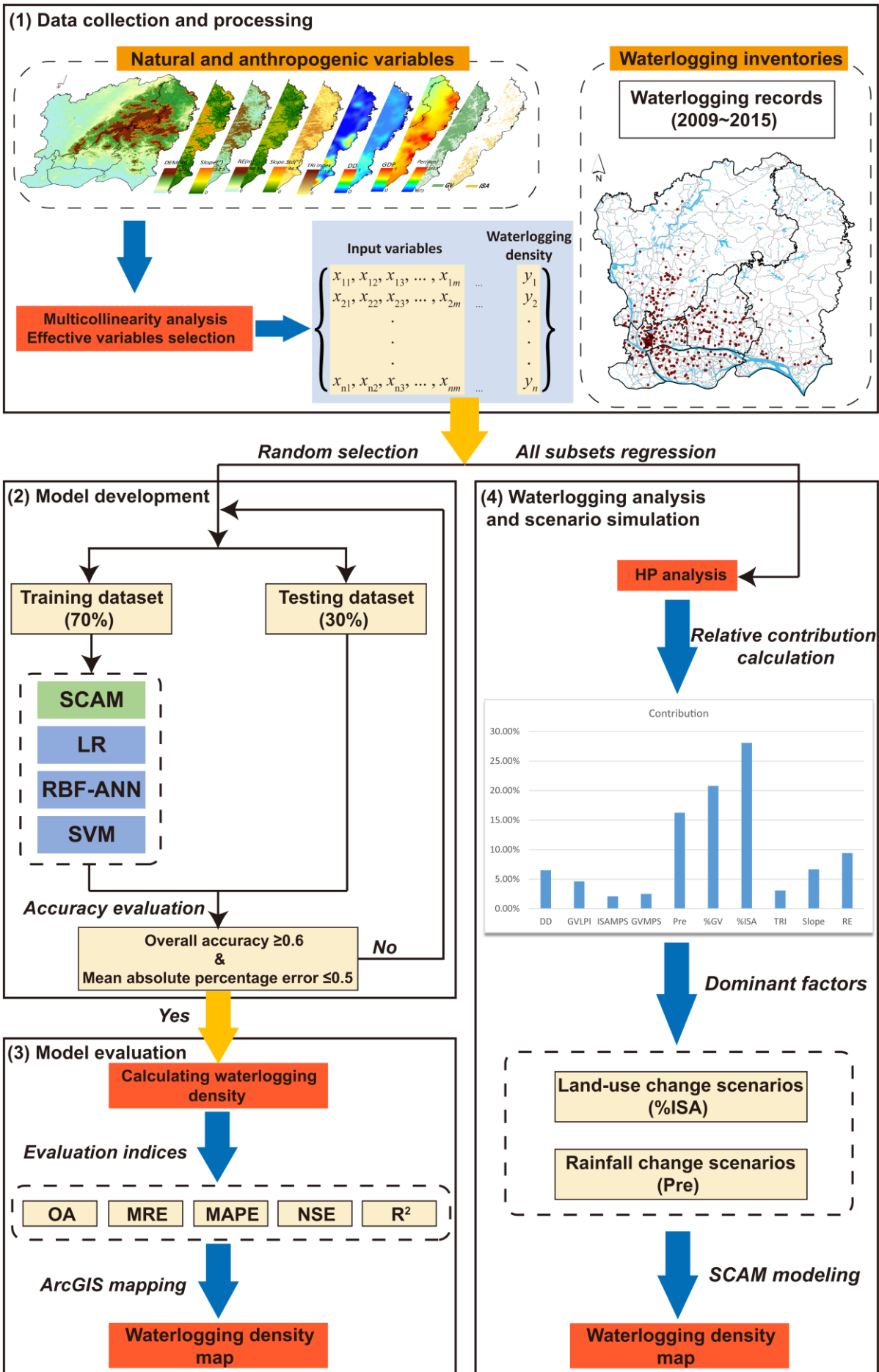




Figure 4-2 The framework of SCAM-HPA.

#### 4.4.5 Scenario simulation

According to the HPA, we found that the impervious surface abundance (%ISA) and precipitation (Pre) has a significant contribution to urban waterlogging variation, which were regarded as the two most important factors affecting urban waterlogging variation. Therefore, in this study, the land-use change scenarios and different rainstorm scenarios were constructed to simulate the variation of urban waterlogging.

Concerning land-use change scenarios, we selected the land-use data of 1995, 2000, 2005, 2010, and 2015 years, elevation, slope, aspect, soil thickness, soil organic matter content, average annual accumulated temperature, average annual precipitation, distance to the lake, river, railway, highway, urban center, economic development zone, rural residential area and town center, population density and urbanization level as spatial drivers. We obtained the data from the Resources and Environmental Data Cloud Platform (<http://www.resdc.cn/>) and calculated it through ArcGIS-Pro. The increasing impervious surfaces, along with the urbanization process, are usually restricted by policies and resources. The growth rate in the future would gradually decrease in the middle and later stages of the urbanization process, which is similar to the regional population development and presented an "s" growth trend (Jiang et al., 2015; Lin et al., 2011; Verburg et al., 2002). Therefore, the LG function of the CLUE-S model was utilized to simulate the change of impervious surface abundance. Finally, the land-use scenarios for the next 20 years from 2020 were calculated. To test the reliability of the CLUE-S simulation results, we compared the 2020 simulation results with the current land-use map and randomly selected 10,000 verification points for the accuracy test. The result indicates that the CLUE-S model met the accuracy requirements of the land-use change scenario simulation.

Compared with the impervious surface abundance in 2020, the impervious surface abundance will increase by 10% in 2027 and nearly 20% in 2036. Therefore, these two land-use scenarios (i.e., impervious surface abundance increased by 10% and 20%) were used to explore the future variation of urban waterlogging. In terms of rainfall scenarios, based on the spatial distribution of precipitation data since 1980, the non-parametric kernel density estimation methods were utilized to calculate the probability density and exceedance probability of cumulative precipitation in each grid unit. The 10-year, 25-year, 50-year, and 100-year precipitation recurrence periods were selected as the rainstorm scenarios, and the corresponding probability levels were 0.1, 0.04, 0.02, and 0.01 respectively. Thus, the annual cumulative precipitation under the four repeating intervals was determined. On this basis, a total of 12 scenarios were constructed to predict

the waterlogging density variation (Table 4-4).

Table 4-4 The land-use change and precipitation scenarios.

Land-use change scenarios	Precipitation scenarios			
	10-year recurrence interval	25-year recurrence interval	50-year recurrence interval	100-year recurrence interval
Impervious surface abundance remains unchanged	B-P10	B-P25	B-P50	B-P100
Impervious surface abundance increases by 10%	S-P10	S-P25	S-P50	S-P100
Impervious surface abundance increases by 20%	L-P10	L-P25	L-P50	L-P100

## 4.5 Results

### 4.5.1 SCAM simulation

In this study, the SCAM trained the calibration and validation data set at eight significance levels to determine the most appropriate cluster tree for urban waterlogging density simulation (Table 4-5). When the significance level was 0.01 ( $\alpha=0.01$ ), the SCAM performed a total of 160 cutting actions and 25 merging actions, and the cluster tree generated a total of 346 total nodes and 136 leaf nodes; while the  $\alpha=0.05$ , the number of cutting and merging actions were significantly increased, and the number of cutting and merging action was 369 and 56, respectively. Compare with  $\alpha=0.01$ , the total number of nodes and leaf nodes increased dramatically. As the  $\alpha$  value grew, the criteria of the cluster tree cutting action became looser. Correspondingly, the cutting actions were more frequent, leading to an increase of intermediate nodes and leaf nodes. This indicated that the significance level had a certain impact on the cluster tree structure.

Table 4-5 The cluster trees of SCAM under eight significance levels.

Significance levels ( $\alpha$ )	Total nodes	Leaf nodes	Cutting actions	Merging actions
0.01	346	136	160	25

0.02	687	257	314	58
0.03	727	277	334	58
0.04	753	272	341	70
0.05	795	314	369	56
0.10	732	326	352	27
0.20	699	338	345	8
0.50	681	341	340	3

Figure 4-3 presented the simulated urban waterlogging density of each watershed unit under all significance levels. We noticed that when the  $\alpha$  value was 0.01 (Fig.4-3b), most of the watershed units had similar simulated values. This might be due to too few cutting and merging actions, resulting in insufficient leaf nodes to reflect the spatial heterogeneity of urban waterlogging. In contrast, with the  $\alpha$  value increased, the structure of cluster trees tended to be more complex, which could adequately reflect the spatial differentiation of waterlogging density in each watershed unit.

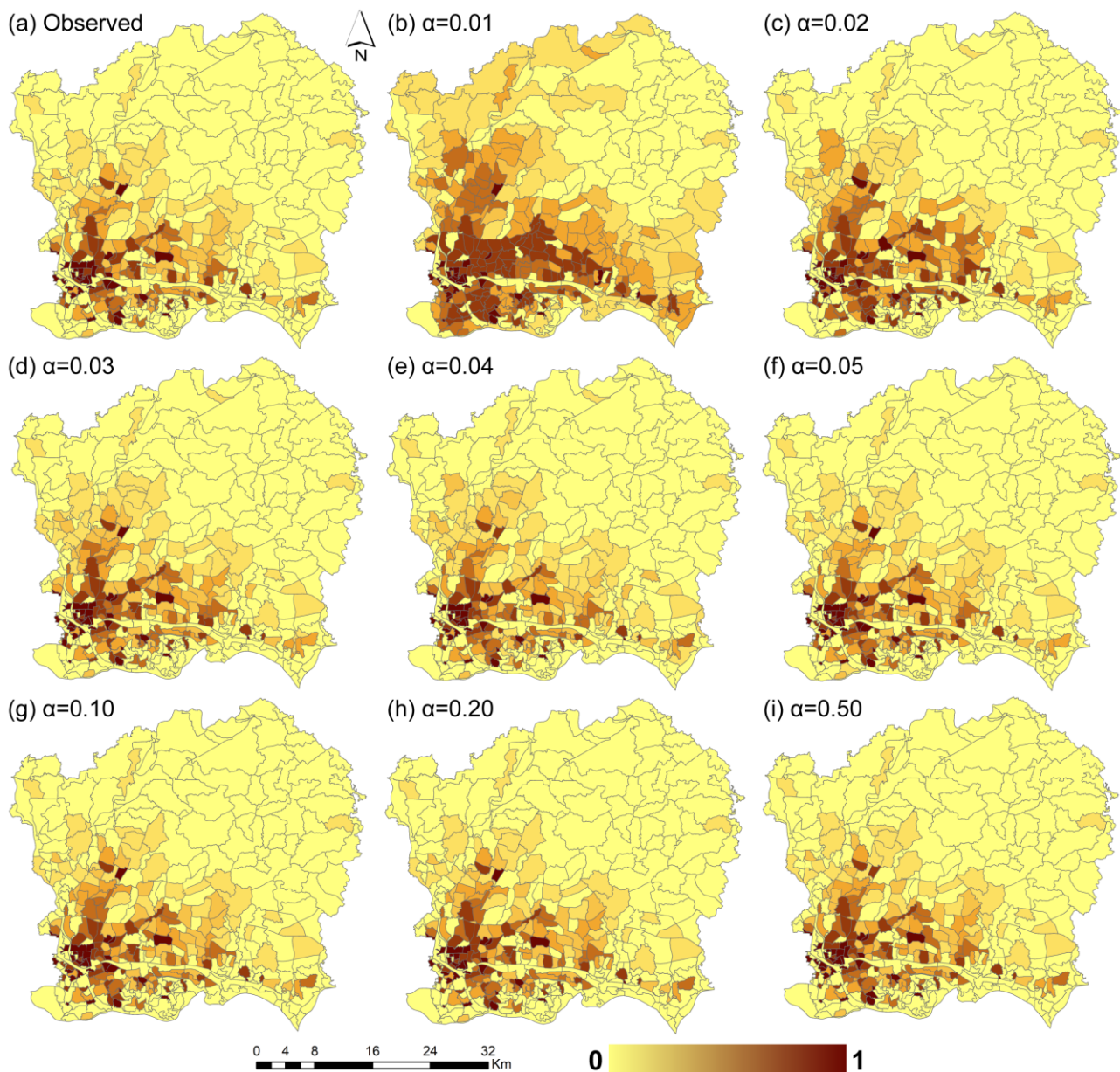


Figure 4-3 The spatial distribution of simulated urban waterlogging density under different significance levels. (a) Observed waterlogging density, (b)~(i) simulated waterlogging density at different significant levels.

The comparison between the simulated and observed waterlogging density under three significance levels is shown in Figure 4-4. The fluctuation trends of simulated waterlogging density were very similar to that of observed waterlogging density, which indicated that the SCAM could effectively capture the complex mechanism linking urban waterlogging to influencing factors. When the  $\alpha$  value is 0.01, the difference between the simulated value and the observed value is the largest in all significance levels, which further confirms that the insufficient leaf node could not accurately reflect the spatial heterogeneity of urban waterlogging density. When the significance level increased to 0.10, the simulation accuracy was significantly

improved. The rising and falling trends of the simulated and observed waterlogging density were well matched. However, when the significant level reached 0.50, the difference between the simulated and observed waterlogging density gradually increased. This might be due to the excessive cutting actions at high significance levels ( $\alpha \geq 0.20$ ), resulting in a large amount of redundant information. As a result, the structure of the cluster tree was too complex, which reduced the model performance and accuracy.

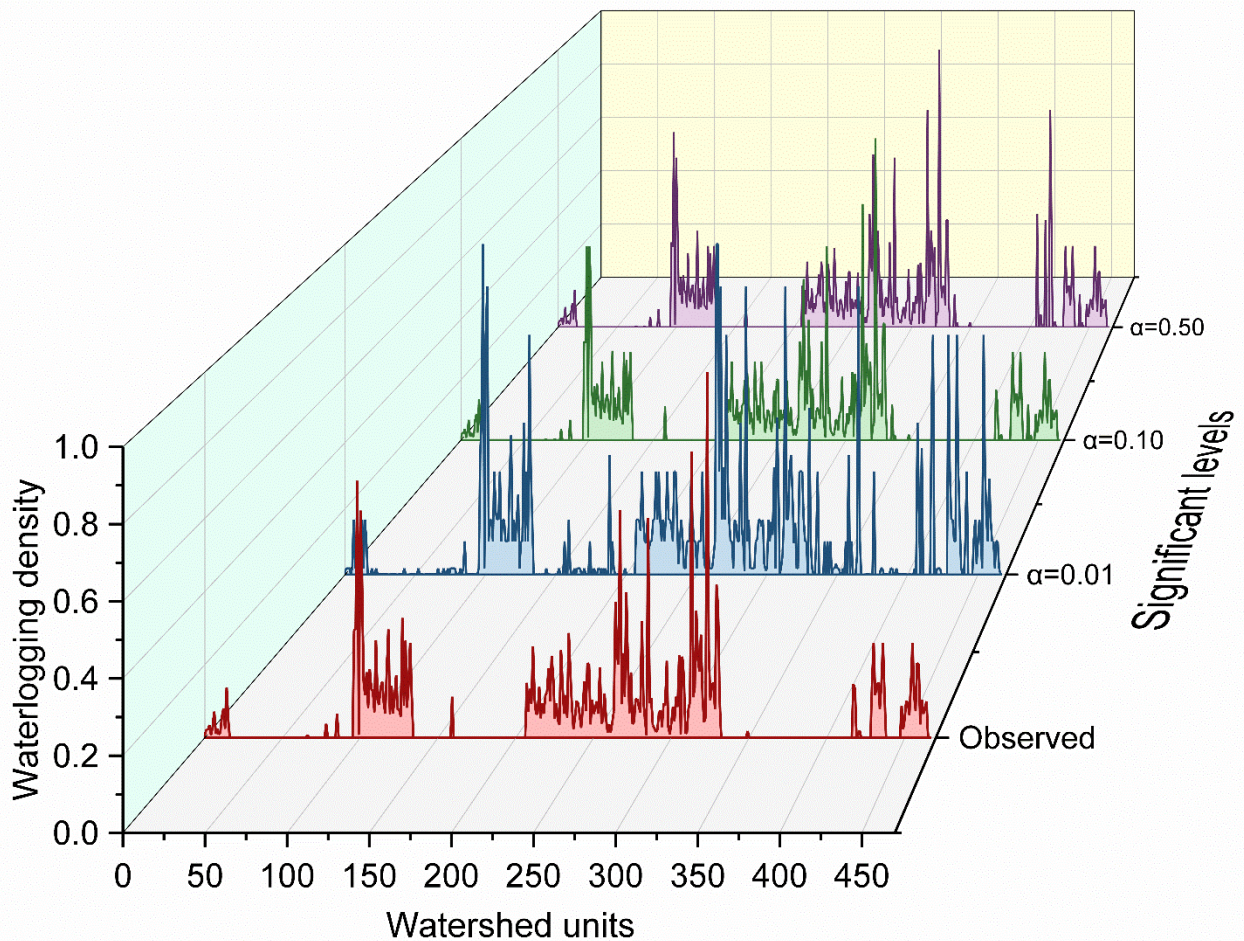


Figure 4-4 The comparison transect of simulated and observed waterlogging density under three significance levels.

#### 4.5.2 Model evaluation and validation

Figure 4-5 depicted the verification results of the overall accuracy (OA), Nash-Sutcliffe model efficiency index (NSE), and the correlation coefficient ( $R^2$ ) of SCAM. The fitting curves (red line) between the simulated and observed waterlogging density were plotted to reflect fitting ability. When the significance level  $\alpha$  was 0.01, the OA, NSE, and  $R^2$  all indicated that the SCAM had poor performance. With the increased significance level, the model accuracy experienced an increasing trend, which showed the model performance gradually improved. When the significance level was 0.10, the SCAM obtained the best performance. The OA, NSE, and

$R^2$  values increased to 0.93, 0.86, and 0.85 in the training dataset, and 0.81, 0.70, and 0.74 in the testing dataset. However, when the significant level exceeded 0.20, the effectiveness of SCAM experienced a downward trend. The value of OA, NSE, and  $R^2$  decreased slightly further confirming that an overly complex cluster tree would reduce the accuracy of the model.

The modified relative error (MRE) and mean absolute percentage error (MAPE) between the simulated and observed waterlogging density were shown in Figure 4-6. Analogously, we noticed that the SCAM had the minimum MRE and MAPE value in the training and testing dataset when the significance level was 0.10. When  $\alpha=0.01$ , the MRE indicator fluctuates significantly, the proportion of samples with MRE less than 10% was 32.5%. The MAPE was 23.42% and 34.46% in the training and testing dataset. However, when the significance levels increased to 0.10, the fluctuation of MRE became more gentle, the proportion of samples with MRE less than 10% was increased to 85.7%. Simultaneously, the MAPE value was reduced to 4.87% and 11.26% in the training and testing dataset. The results of OA, NSE,  $R^2$ , MRE, and MAPE both indicated that the model was optimal at the significance level of 0.10. Finally, according to the verification result, we selected the significant level  $\alpha=0.10$  as the most appropriate level for SCAM to simulate urban waterlogging variation.

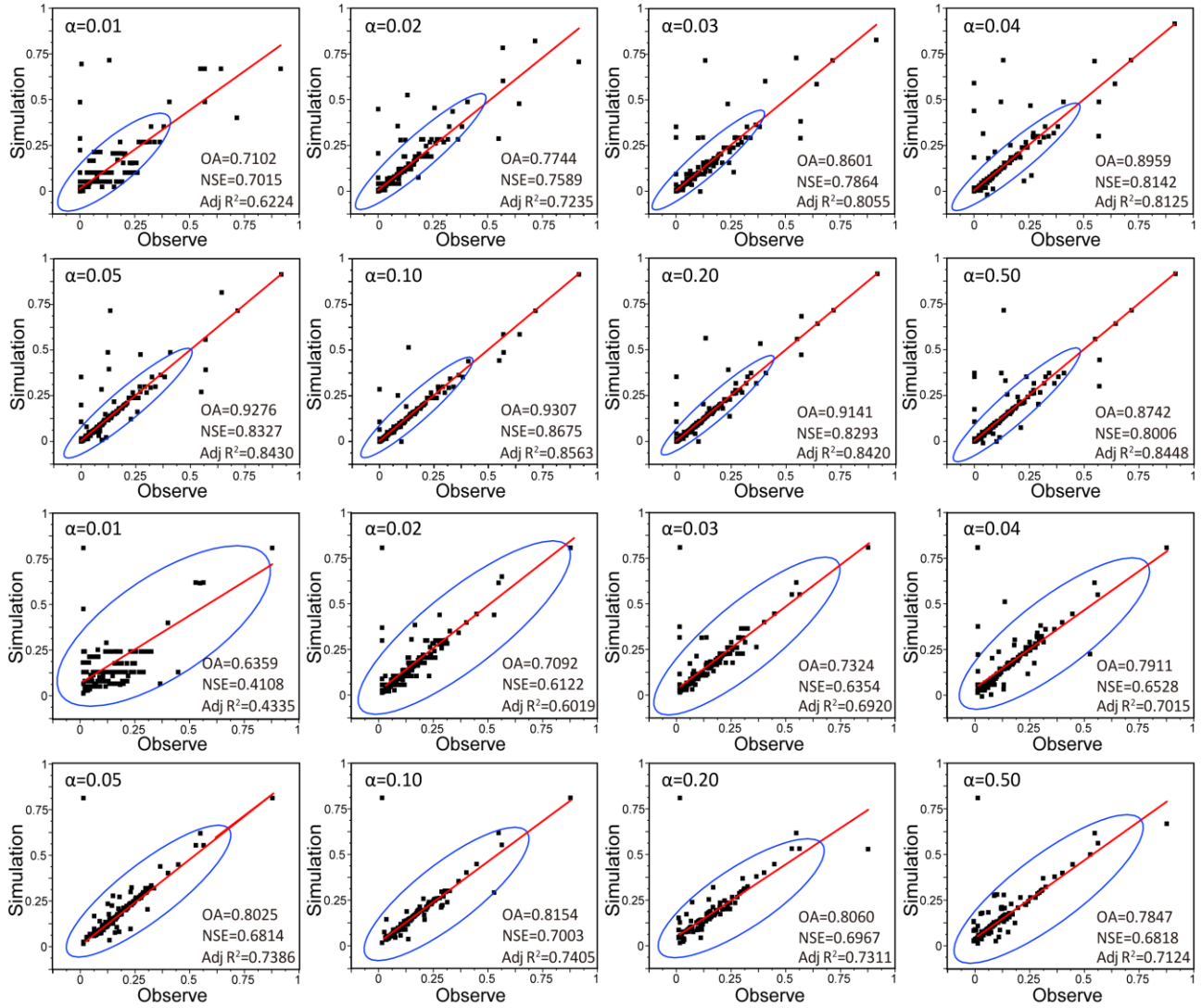


Figure 4-5 The OA, NSE, and scatter plots of simulated and observed waterlogging density. (the red line indicated the fitting curve and the ellipse indicated the 95% confidence ellipse)

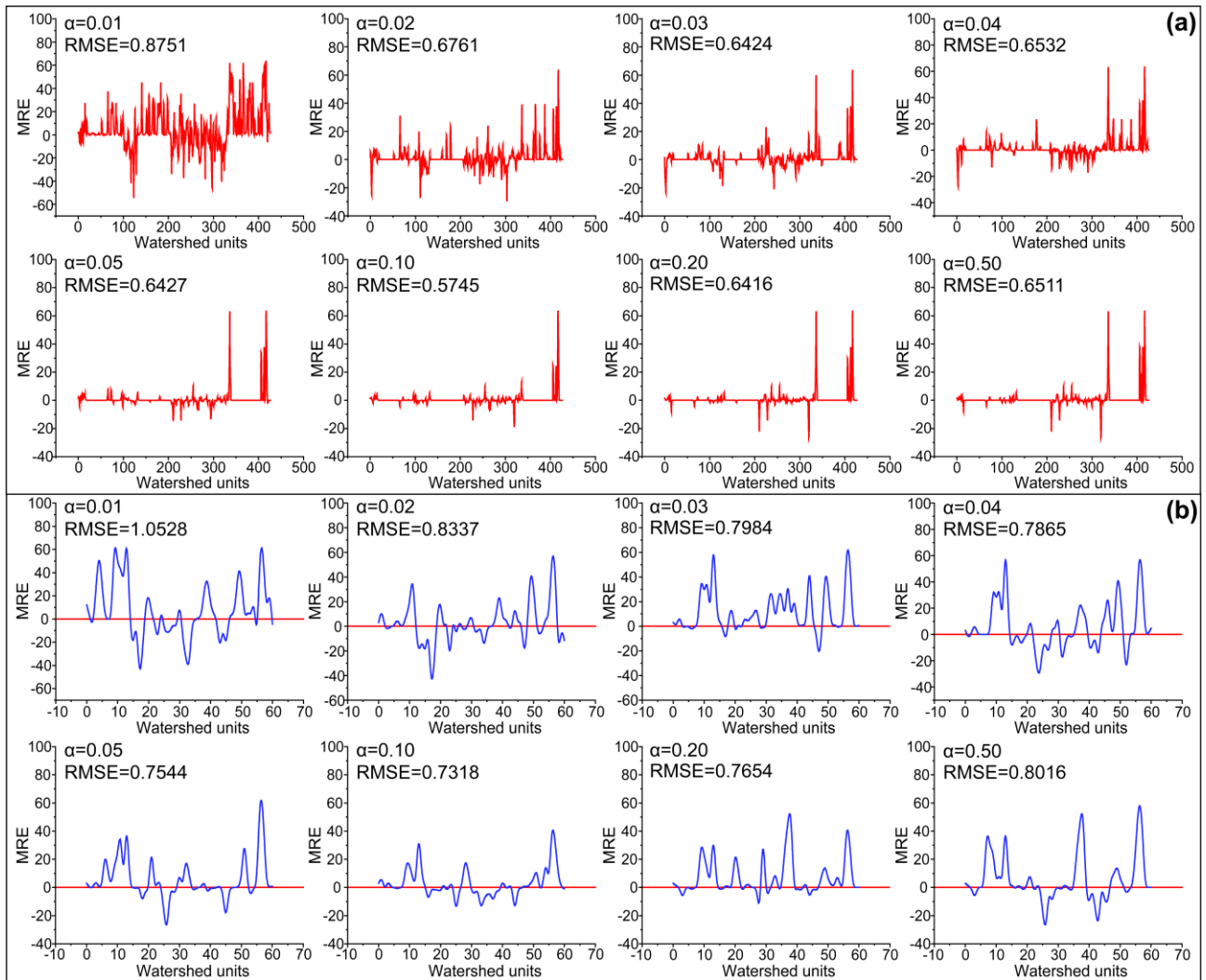


Figure 4-6 The MRE and MAPE between the simulated and actual waterlogging density.

In this study, we further compared the SCAM with three analysis models (LR, RBF-ANN, SVM) to examine the effectiveness and reliability of the SCAM (Table 4-6). We noticed that the SCAM obtained the highest OA, NSE,  $R^2$ , and the lowest MAPE, which indicated that the SCAM had the best performance than other models. The highest OA was from SCAM (0.93 and 0.82 in calibration and validation data sets), which indicated that 93% and 82% of watershed units in calibration and validation data sets were correctly identified, followed by SVM, RBF-ANN, and LR. The NSE index of SCAM in calibration and validation data sets (0.87 and 0.70) illustrated that the extreme values of waterlogging density in the Guangzhou central urban districts could be effectively simulated. The NSE index of the SVM model was 0.81 and 0.68 in training and testing data sets, while the NSE of the RBF-ANN was 0.67 and 0.59, respectively. These indicated that the RBF-ANN and SVM model had an acceptable performance in extreme value simulation. However, the LR model had the lowest NSE index, only 0.55 and 0.42 in calibration and validation data sets, which indicated that the LR model could not effectively



identify the complex mechanisms of urban waterlogging. Furthermore, the MAPE value showed that the SCAM and SVM model had a relatively small error between simulated and recorded waterlogging density. In contrast, the RBF-ANN and LR had relatively large MAPE, with a maximum error of 33.41%. Therefore, we could infer that the LR model was not suitable for capturing the non-stationary and non-linear interaction. The  $R^2$  of SCAM (calibration: 0.86 and validation: 0.74) was significantly higher than other models, especially the LR model (calibration: 0.60 and validation: 0.52). This finding revealed that the simulated value of SCAM was very close to the observed waterlogging density. In the face of great spatial heterogeneity in highly urbanized areas, a more robust method was needed to quantify the variation of urban waterlogging events. Although, the nonlinear mapping ability of RBF-ANN and SVM could improve the accuracy to some extent. However, the SCAM based on Wilks' lambda statistic could handle nonlinear urban waterlogging variation more precisely and reliably.

*Table 4-6 The performances of SCAM and compared models.*

Index	Calibration dataset				Validation dataset			
	LR	RBF-ANN	SVM	SCAM	LR	RBF-ANN	SVM	SCAM
OA	0.71	0.82	0.89	0.93	0.62	0.71	0.74	0.82
NSE	0.55	0.67	0.81	0.87	0.42	0.59	0.68	0.70
MAPE	26.45%	15.13%	9.82%	4.87%	33.41%	25.05%	18.94%	11.26%
$R^2$	0.60	0.76	0.82	0.86	0.52	0.61	0.69	0.74

#### 4.5.3 Comparison of the simulation results of each model

The waterlogging density of Guangzhou central urban district was classified into five levels (namely the lowest, low, medium, high, and highest level) by using the natural breaks classification method (Table 4-7). In general, the urban waterlogging density presents tremendous spatial heterogeneity. The high level and highest level waterlogging density watersheds accounted for 8.33% of the study area and mainly concentrated in the historic urban districts of Guangzhou, such as Liwan, Yuexiu, and Haizhu (Figure 4-7a). These areas have experienced an intense urbanization process, with a large proportion of impervious surfaces. Furthermore, the design standards of drainage networks in historic urban districts are relatively low and inefficient, making it difficult to drain rainwater in time. Therefore, the waterlogging density level in these areas is relatively high. On the contrary, the low-level watersheds accounted for 69.07% area, and are mainly concentrated in the Baiyun and Huangpu districts. The dominant land-use types in these two districts are forest and farmland,

with a record low frequency of urban waterlogging events.

*Table 4-7 Waterlogging density level.*

Density level	Lowest	Low	Medium	High	Highest
Waterlogging density	< 0.043	0.043-0.127	0.127–0.445	0.445–0.624	> 0.624
Proportion (%)	69.07%	14.16%	8.44%	6.25%	2.08%

As shown in figure 4-7 (c-f), we presented the comparative result of the SCAM and LR, RBF-ANN, and SVM models. Region *i* belongs to a rural area with good ecological conditions. According to historical records, there are no urban waterlogging events in this area. Therefore, we could assume the waterlogging density of each watershed in region *i* was the lowest level (Fig.4-7a). The SCAM identified this region correctly as the lowest level of waterlogging density; while the SVM, RBF-ANN, and LR model classified a large part of watershed units as low level or even medium level. This implies that the LR, RBF-ANN, and SVM models may have an overestimation problem in low-density areas. Unlike region *i*, region *ii* is located in the city center, where the building density and impervious surface abundance are relatively high (Fig.4-7b). According to the waterlogging database, this region has a large amount of waterlogging records, indicating it as a typical waterlogging prone area. As shown in Fig.4-7c-f, we notice that all models accurately identify this region as the high level or highest level of waterlogging density, which indicates that these models have better performance in urban waterlogging high-density regions than low-density regions. We also find that some watershed units in region *ii* present the lowest urban waterlogging density level. However, the LR and SVM models almost classify all the watershed units as high or highest level, whereas the simulated result of SCAM is closer to the actual values. Our results demonstrate that the SCAM can provide more accurate and detailed results in both high-density and low-density regions.

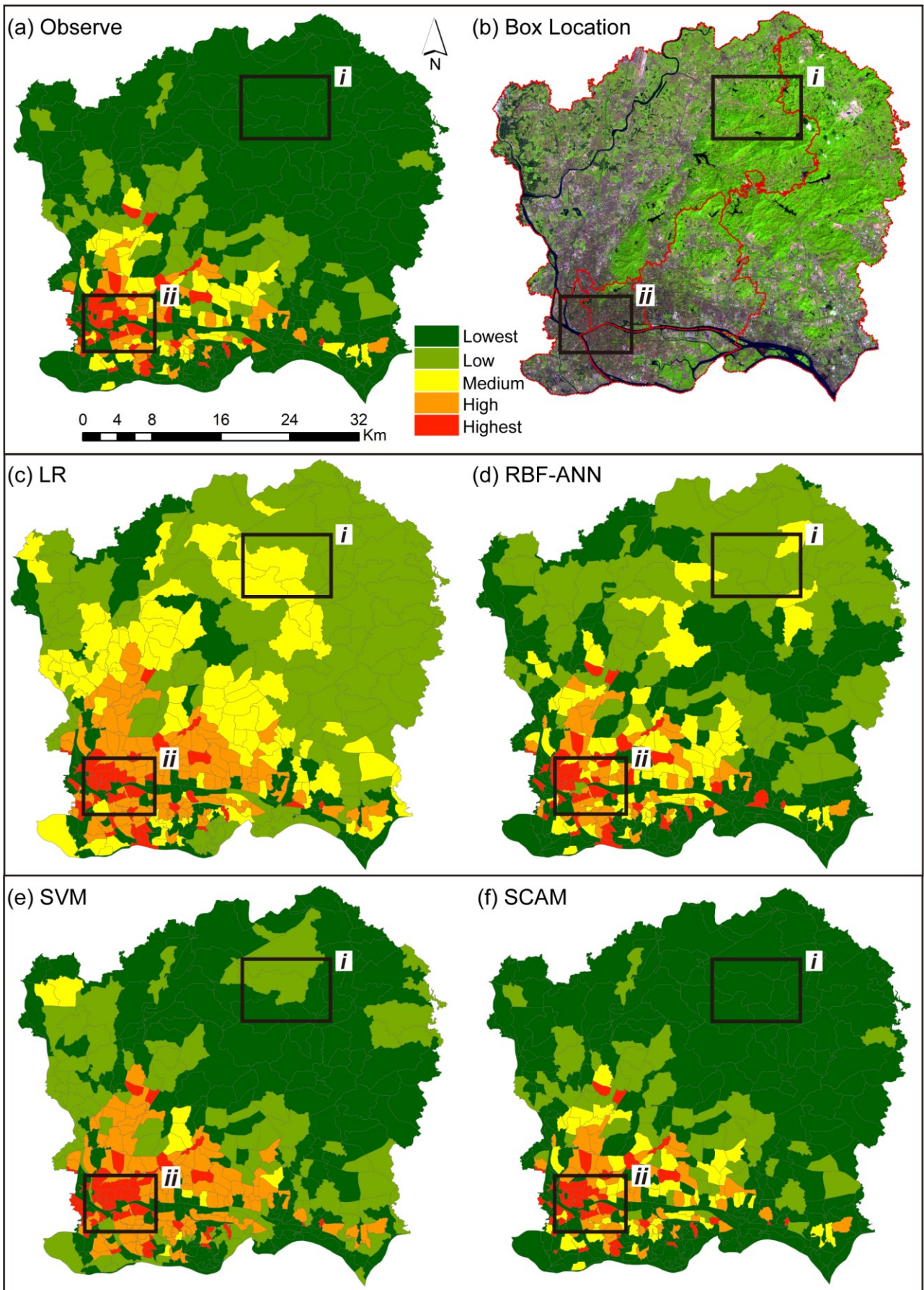


Figure 4-7 Comparison of urban waterlogging density maps. (a) Observed waterlogging density, (b) location of the region

*i and ii, (c) waterlogging density simulated by LR, (d) waterlogging density simulated by RBF-ANN, (e) waterlogging density simulated by SVM, (f) waterlogging density simulated by SCAM.*

#### 4.5.4 The dominant drivers of urban waterlogging variation

The HPA method, which integrated with the SCAM, was utilized to quantify the independent effect of each input variable on waterlogging and specified their relative contributions (Table 4-8). The impervious surface abundance (28.07%), vegetation abundance (20.80%), and precipitation (16.25%) present the dominant contribution rate. Compared with land cover composition, the cumulative precipitation (16.25%) is also a significant driver for urban waterlogging. If rainfall exceeds the rainwater discharge capacity, the risk of urban waterlogging is significantly increased. Therefore, it is necessary to give more priority to controlling these three factors. Moreover, it is interesting to find that the topography factors (RE, slope.std, TRI) and land cover configuration (MPS and LPI) also have an appreciable effect on waterlogging magnitude. The RE, slope.std, and TRI index contribute 9.40%, 6.66%, and 3.10%, respectively, while the three landscape pattern metrics contribute nearly 10% to waterlogging variations.

*Table 4-8 The relative contributions of model input variables affecting urban waterlogging.*

Model input variables	RE	Slope.s td	TRI	%ISA	%GV	Pre	ISA MPS	GV MPS	GVLPI	DD
Contribution	9.40%	6.66%	3.10%	28.07%	20.80%	16.25%	2.10%	2.50%	4.63%	6.50%

#### 4.5.5 Waterlogging simulation under different scenario

According to the dominant factors of urban waterlogging, 12 scenarios were forcibly inputted into the SCAM to simulate waterlogging density variation (Fig.4-8). Under the same precipitation scenario, when the proportion of impervious surface increase by 20% (L-P10 scenario), the average urban waterlogging density in the central urban district of Guangzhou is 0.297, while the average urban waterlogging density is 0.058 under the B-P10 scenario (impervious surface abundance remains unchanged). Compare with urban expansion scenarios, when the cumulative precipitation is the B-P100 scenario, the average waterlogging density is 0.092. This value is higher than the current reality situation (0.051) but far less than the land-use change scenario (L-P100). This result highlights that urban expansion (land-use change) significantly increases urban waterlogging density, which is the main cause of urban waterlogging deterioration.

Table 4-9 The average urban waterlogging density under different scenarios.

Scenarios	Waterlogging density	Scenarios	Waterlogging density	Scenarios	Waterlogging density
B-P10	0.058	S-P10	0.112	L-P10	0.297
B-P25	0.062	S-P25	0.125	L-P25	0.312
B-P50	0.077	S-P50	0.141	L-P50	0.350
B-P100	0.092	S-P100	0.177	L-P100	0.414

Furthermore, we also noticed that watershed units in different locations have different sensitivity to land-use change and rainfall change scenarios. The watershed units located in the northern and eastern parts of the central urban district (Baiyun and Huangpu district) are more sensitive to land-use change scenarios than those in the southwest urban districts (Liwan, Haizhu, Yuexiu district). In the land-use change scenarios, the watersheds with the highest increase rates of waterlogging density are mainly located in the northern and eastern regions. The urban waterlogging-prone area gradually spread from the urban historic districts (Liwan, Haizhu, Yuexiu district) to the urban fringe (Baiyun and Huangpu district). This might be because a large amount of urban green spaces (i.e., forest, wetland, and grassland) in these regions were converted into impervious surfaces during the urbanization scenarios, resulting in a significant increase in the risk of urban waterlogging.

On the contrary, under the same land-use change scenario, with the increase of cumulative precipitation, the waterlogging density in the urban historic districts (Liwan, Haizhu, Yuexiu district) increases more significantly than that in the suburban district (Baiyun and Huangpu district). This is mainly due to the inadequate drainage system and relatively high impervious surface abundance in the historic urban district. Hence, the waterlogging density in these regions increases dramatically. At the same time, the vegetation abundance in the urban fringe is relatively high, which could effectively absorb the increased rainfall, so the waterlogging density does not increase significantly.

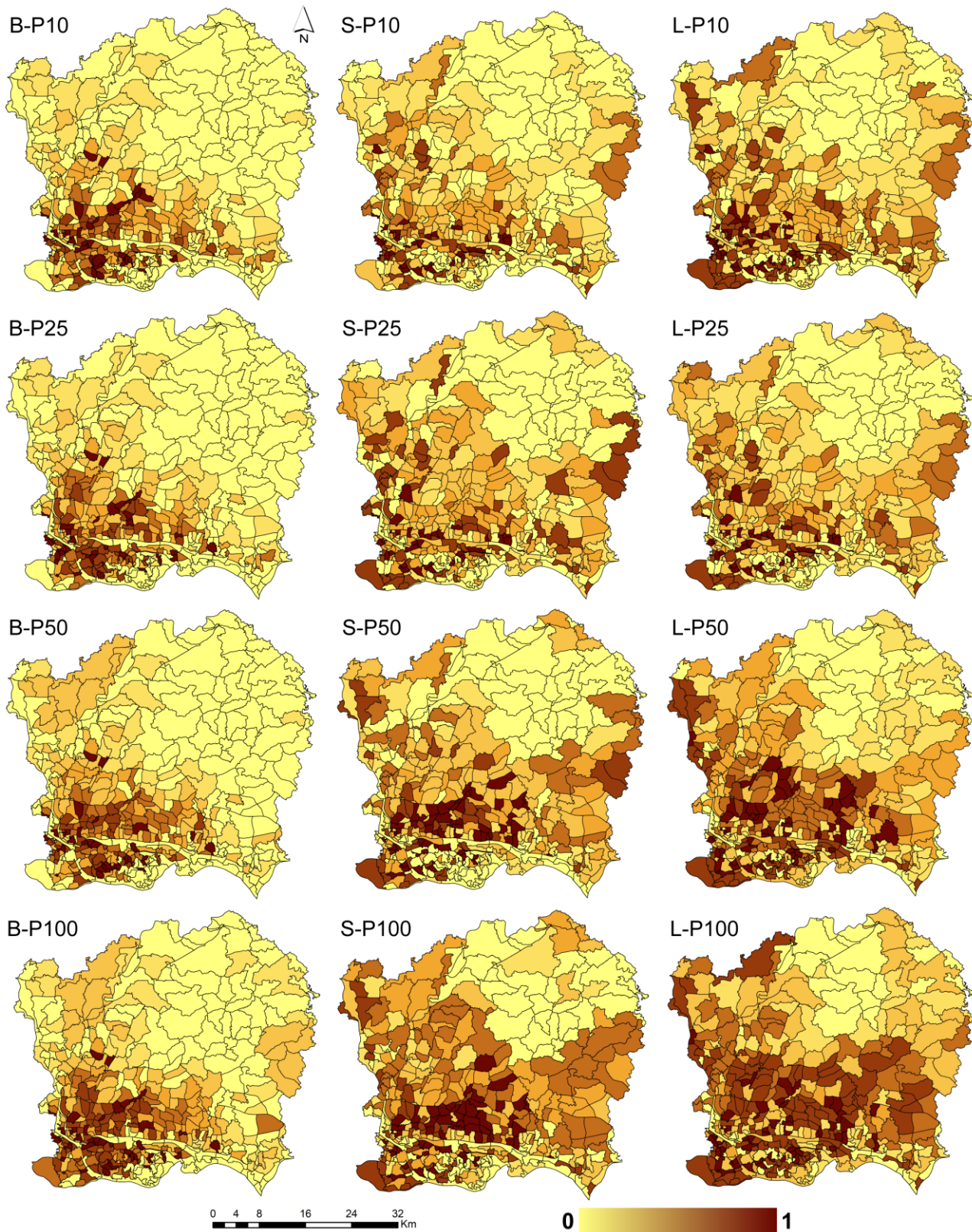


Figure 4-8 The spatial distribution of waterlogging density in different scenarios.

## 4.6 Discussion

Our results found that the high-density urban waterlogging watersheds are mainly clusters in the Liwan, Yuexiu, and Haizhu districts, indicating these regions are highly susceptible to urban waterlogging events (Fig.4-7a). Thereby, the local authorities can pay more attention to the risk warning for these regions. Simultaneously, with the application of spatio-temporal big data (distribution of population, public institutions), we can predict and assess the urban waterlogging vulnerability to minimize waterlogging negative effects. Through HPA, our result demonstrates that both land cover composition and precipitation are the dominant drivers of urban waterlogging, which is consistent with previous studies (Quan et al., 2014; Zhang et al., 2020). Results from this study provide additional insights that the land cover composition – impervious surface and vegetation, had a dominant effect in deriving waterlogging variations, which independently accounted for 48.87% of the contribution. Therefore, it is necessary to control the expansion of impervious surfaces and increase urban green space as much as possible to reduce the risk of urban waterlogging. The cumulative precipitation is also indicated that as rainfall increases, the state of urban waterlogging gradually deteriorates (IPCC, 2011; Barros et al., 2012; Hallegatte et al., 2013).

In general, our results reflect that land-use change affects the spatial variation of urban waterlogging more significantly than precipitation, which further verifies the land cover feature is the critical driver of urban waterlogging. However, it is interesting to note that watersheds in different spatial locations have different sensitivity to land-use change and rainfall change scenarios. This indicates that the watershed spatial location and watershed characteristics (i.e., land cover characteristics, drainage network density, microtopography) should be considered in the waterlogging sensitivity identification. This result is of considerable significance to urban waterlogging prevention. Due to the rapid urbanization, a large number of urban green spaces, rivers, wetlands in the city center have been converted into various development areas (such as commercial, industrial, and residential land). However, in the current urban eco-environmental management policies, whenever the green spaces or wetlands in the city center are occupied by urbanization, the green areas in the urban fringe (less pressure for urban development) are used to fill the area loss in the city center. Although this approach can maintain the relative balance of urban green areas at the statistical level. However, the spatial replacement of green spaces has resulted in a continuous net loss of vegetation abundance in the central urban districts. The consequence is the lack of ecological benefits (such as waterlogging mitigation effect) required for such urban areas. This may lead to the further deterioration of urban waterlogging in the city center (urban green spaces continue to decrease), even though the urban green areas within the whole city remain relatively balanced. Our study indicates that a spatial understanding

of each mitigation measure is required before adopting urban waterlogging prevention strategies.

The urban waterlogging prone areas with large impervious surface abundance, are more sensitive to rainfall change scenarios. Thereby, the government should rationally manage the existing land cover features for urban waterlogging mitigation, even though the pressure of land-use development is relatively high. It is effective to increase the vegetation abundance to form green corridors, low elevation greenbelts, or green roofs to increase the accumulation, evaporation, and infiltration of rainwater. Furthermore, if the land-use pressure in the central districts is relatively high (difficult to increase the urban green areas), it is necessary to increase drainage density as well as optimize the spatial pattern of land cover. In contrast, for the urban waterlogging low sensitive areas, where the vegetation coverage is relatively high, the watershed units are more sensitive to land-use change scenarios. Thus the local authorities should be strictly controlled the speed of urban expansion to avoid exacerbating the frequency of urban waterlogging disasters in newly developed areas. This result implies that urban waterlogging mitigation strategies should be formulated according to different local conditions and future scenarios.

This study novel attempts to introduce the SCAM to capture urban waterlogging spatial variation characteristics at the watershed level. The results confirmed that the SCAM-HPA framework has many advantages over other methods, as it can effectively investigate the inherent non-linear/discrete relationship, performs fast, no statistical assumptions, and can process data from different measurement scales. Another advantage of this approach is that it can clarify the waterlogging dominant drivers, which can provide practical suggestions for urban waterlogging risk identification and control. It can be applied as a general framework to other cities for urban waterlogging risk identification and urban waterlogging control. However, its limitations can be concluded as follows. Firstly, as a drawback of SCAM, this method has high requirements for the predictor set (screening effective predictor set is required) and cannot capture the new extremum. Secondly, although we have obtained average precipitation data, precipitation intensity and rainfall duration are also important factors affecting urban waterlogging. Due to data limitations and accessibility, we cannot obtain these data for this study. Consequently, in future research, the introduction of these data will help to improve the accuracy of urban waterlogging risk identification. Thirdly, the urban waterlogging records may be insufficient, which only records from the period of 2009-2015. There is no specific time for each waterlogging event. In future research, the water depth and frequency of urban waterlogging should be further considered.



## 4.7 Conclusion

This study proposed a novel approach to explicit the urban waterlogging spatial variation and determine its dominant drivers by implementing the SCAM and HPA. Specifically, the SCAM was applied to simulate the urban waterlogging variation and identified the urban waterlogging susceptibility areas under different scenarios. This study gains three conclusions: (1) the proposed SCAM can successfully capture the non-stationary and non-linear interaction between urban waterlogging and explanatory factors. Compared with LG, RBF-ANN, and SVM, the SCAM provides more accurate and detailed simulated results both in highly urbanized urban centers where waterlogging frequently occurs and suburban areas with few waterlogging events. This result indicates that the SCAM has high classification accuracy and generalization capability, which is suitable for the simulation and prediction of urban waterlogging in highly heterogeneous urbanized areas. (2) The HPA reveals that the impervious surface abundance (28.07%), vegetation abundance (20.80%), and cumulative precipitation (16.25%) are the dominant drivers of urban waterlogging variation. This result indicates that urban waterlogging magnitude is mainly affected by both land cover composition and precipitation. Therefore, it is necessary to give more priority to controlling these three factors. Furthermore, the urban micro-topography and land cover configuration also have a certain influence on waterlogging magnitude, which indicates that changing land cover spatial arrangements will also affect the occurrence of waterlogging while keeping the land cover composition constant. (3) In general, with the increase of cumulative precipitation and impervious surface abundance, the urban waterlogging density increases significantly. It is interesting to note that under different urbanization and precipitation scenarios, the urban waterlogging susceptibility areas have a considerable variation. This result provides valuable insights that urban waterlogging mitigation strategies should be formulated according to different future conditions. It suggests that for the waterlogging prone regions (central urban district), the impervious surface and urban green area should be properly adjusted, even though the land-use pressure is very high. In contrast, for the waterlogging low sensitive areas (urban fringe), it is necessary to restrict the expansion of impervious surfaces to avoid deteriorating the state of urban waterlogging. The results of this study demonstrate that the combination of SCAM and HPA can provide an effective and feasible solution for urban waterlogging variation simulation. They also highlight the dominant drivers of urban waterlogging, which provide theoretical and practical enlightenment for urban waterlogging prevention and management.

## **Author Contributions**

**Qifei Zhang:** Conceptualization, Methodology, Formal analysis, Data curation, Writing - original draft. **Zhifeng Wu:** Resources, Writing - review & editing, Supervision. **Guanhua Guo:** Validation, Writing – review & editing. **Hui Zhang:** Methodology, Validation. **Paolo Tarolli:** Conceptualization, Methodology, Writing - review & editing, Overall supervision, Project administration.

## **Acknowledgments**

The study was partly funded by the University of Padova research projects (DOR1948955/19; DOR2079232/20), the NSFC-Guangdong Joint Foundation Key Project (U1901219), and the Team Project of Guangdong Provincial Natural Science Foundation (grant number 2018B030312004).

## **Conflicts of Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2020.143041>.

Supplementary materials include the following:

Figure S4-1 The spatial distribution of urban waterlogging events and watershed units.

Figure S4-2 The topographic factors of elevation (a), relative elevation (b), elevation standard deviation (c), slope (d), slope standard deviation (e), TRI (f), precipitation (g).

Figure S4-3 The vegetation abundance (GV) and impervious surface abundance (ISA) for the study area.

Figure S4-4 The drainage density (a) and GDP (b).

Figure S4-5 The framework of SCAM.

Figure S4-6 Different land-use scenarios. (a) Impervious surface abundance remains unchanged, (b) impervious surface abundance increases 10%, (c) impervious surface abundance increases 20%.

Figure S4-7 Different rainfall scenarios. (a) 10-year recurrence interval, (b) 25-year recurrence interval, (c) 50-year recurrence interval, (d) 100-year recurrence interval.

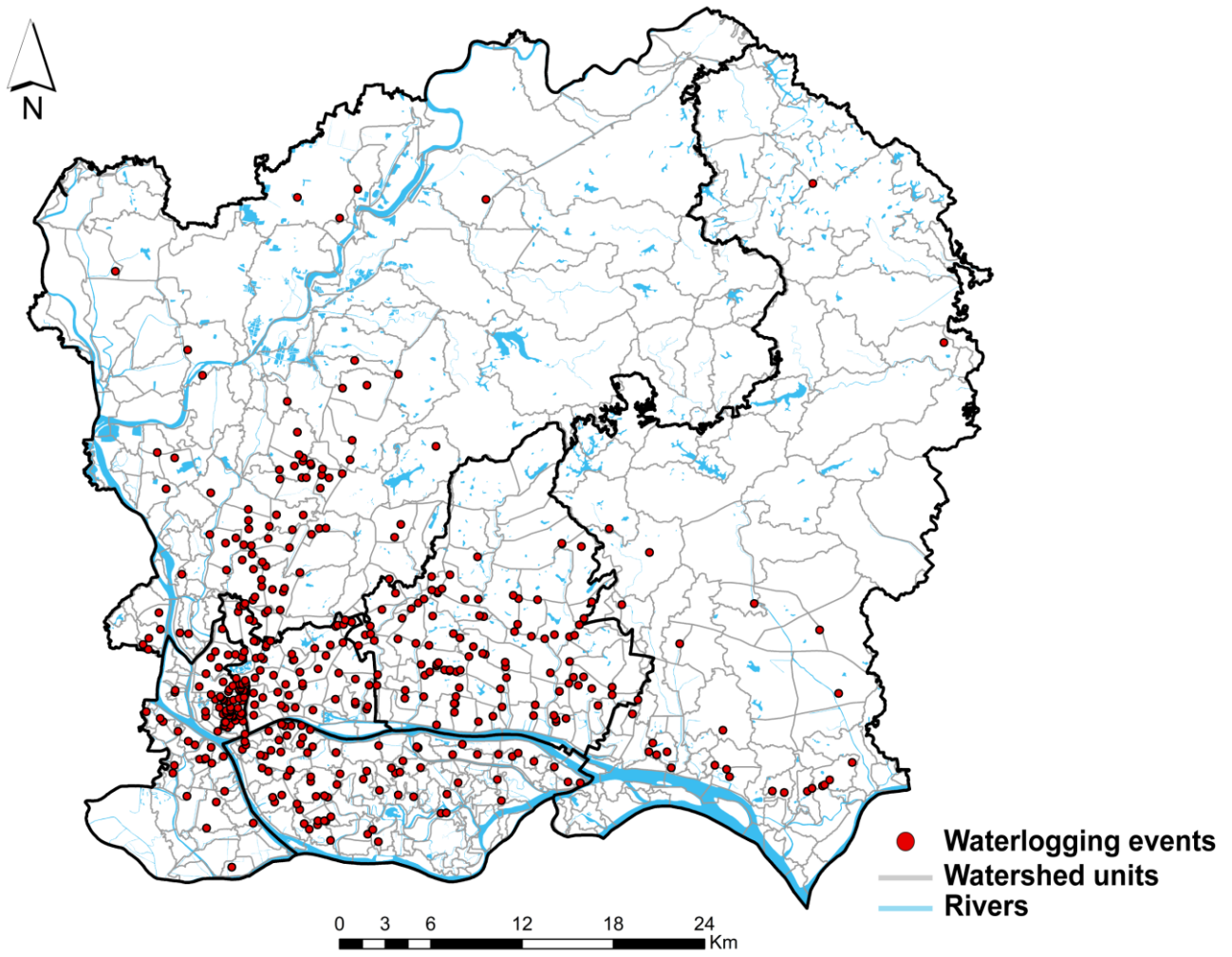


Figure S4-1 The spatial distribution of urban waterlogging events and watershed units.

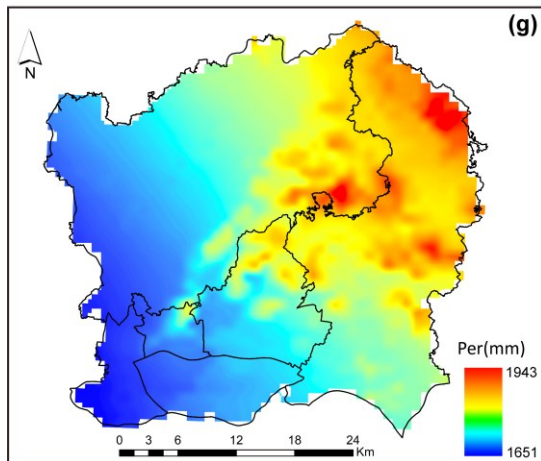
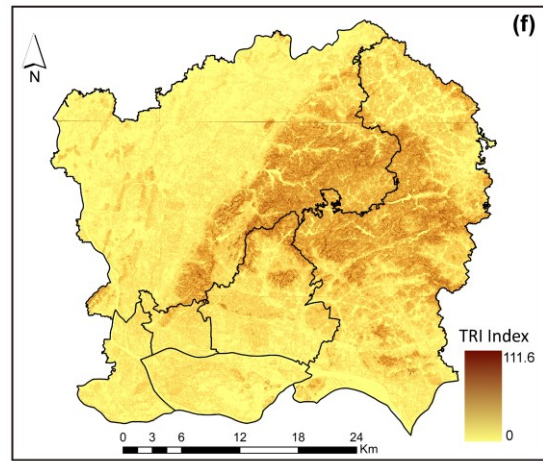
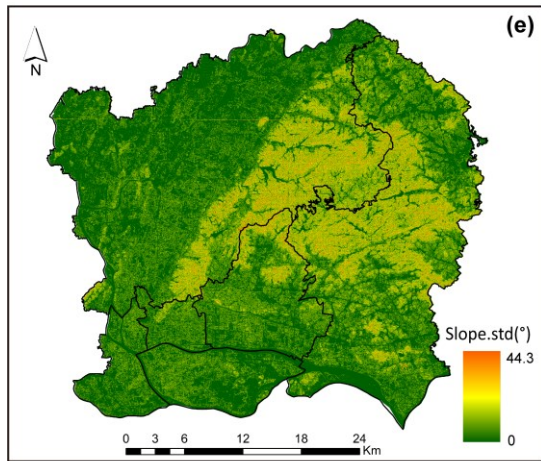
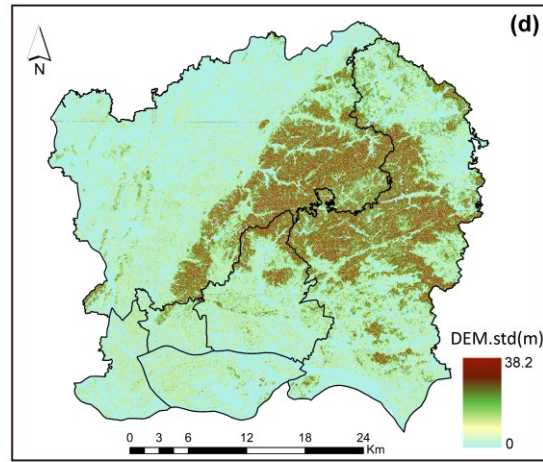
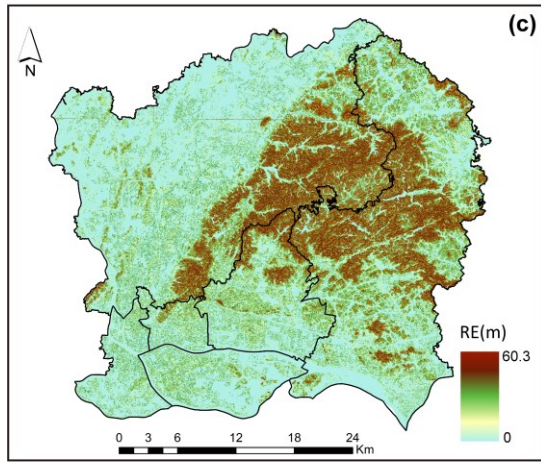
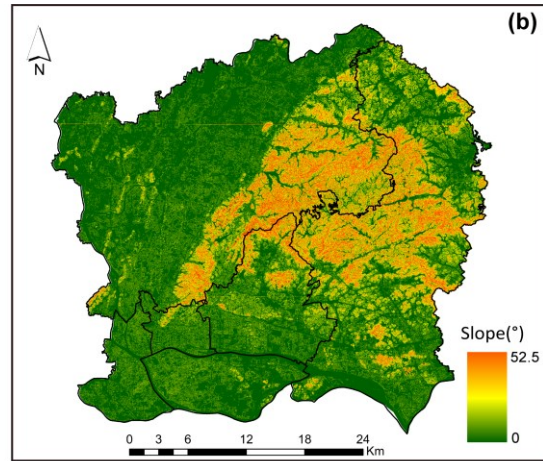
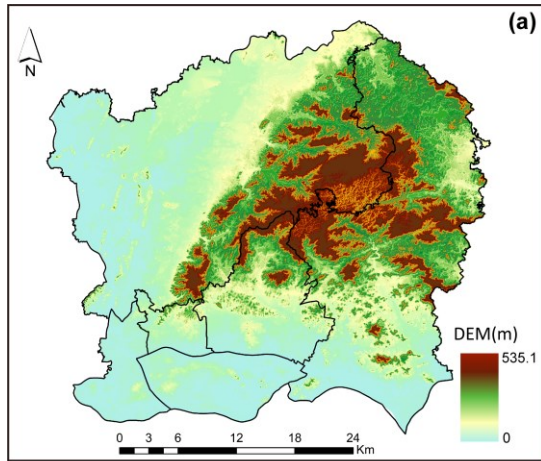


Figure S4-2 The topographic factors of elevation (a), slope (b), relative elevation (c), standard deviation of elevation (d), standard deviation of slope (e), TRI (f), precipitation (g).

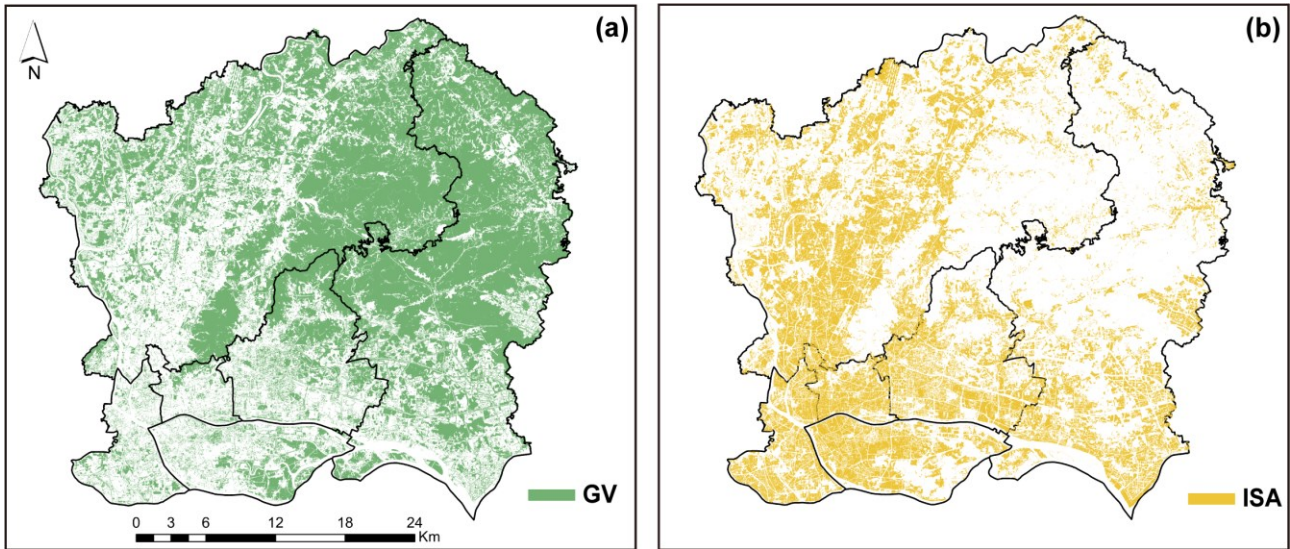


Figure S4-3 The vegetation abundance (GV) and impervious surface abundance (ISA) for the study area.

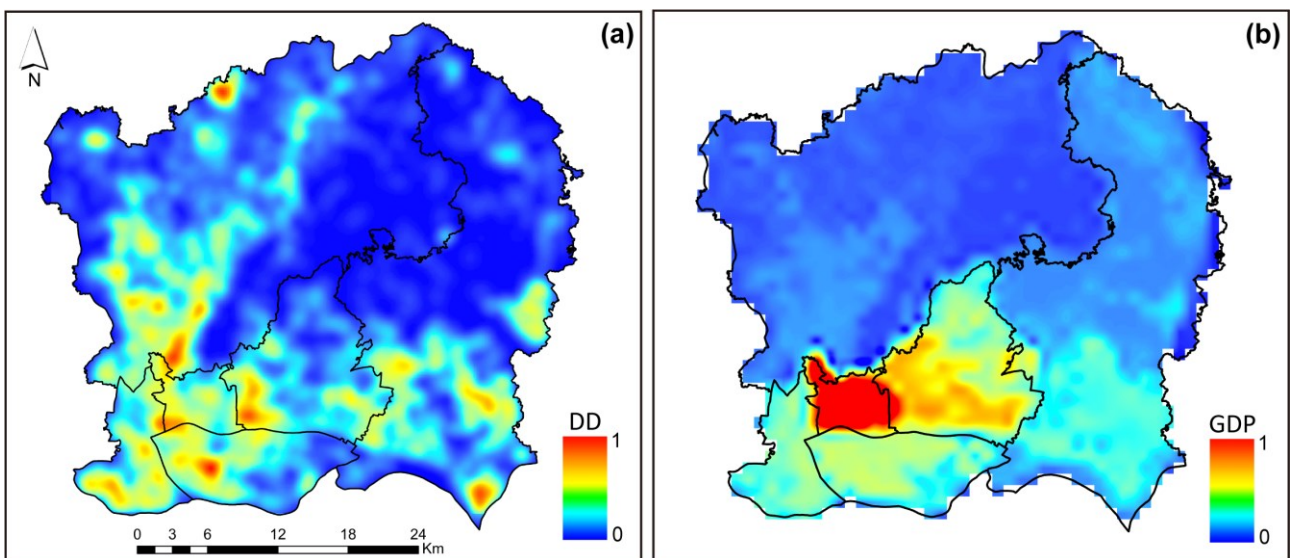


Figure S4-4 The drainage density (a) and GDP (b).

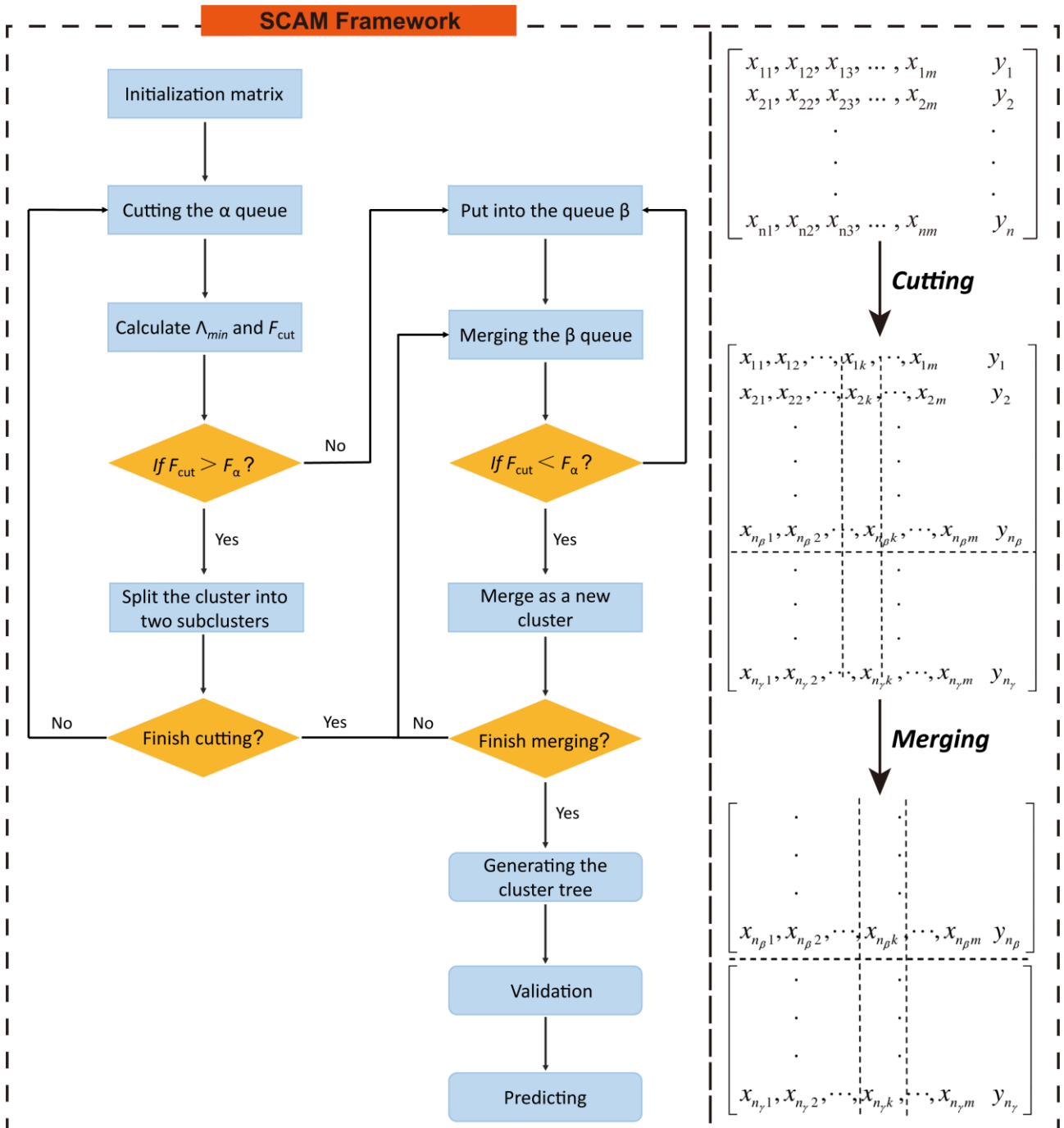


Figure S4-5 The framework of SCAM.

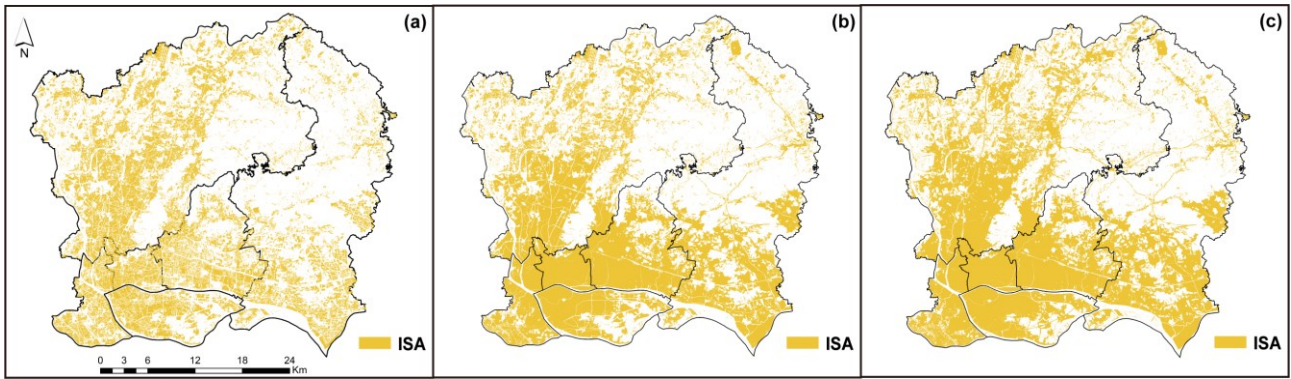


Figure S4-6 Different land-use scenarios. (a) Impervious surface abundance remains unchanged, (b) impervious surface abundance increases 10%, (c) impervious surface abundance increases 20%.

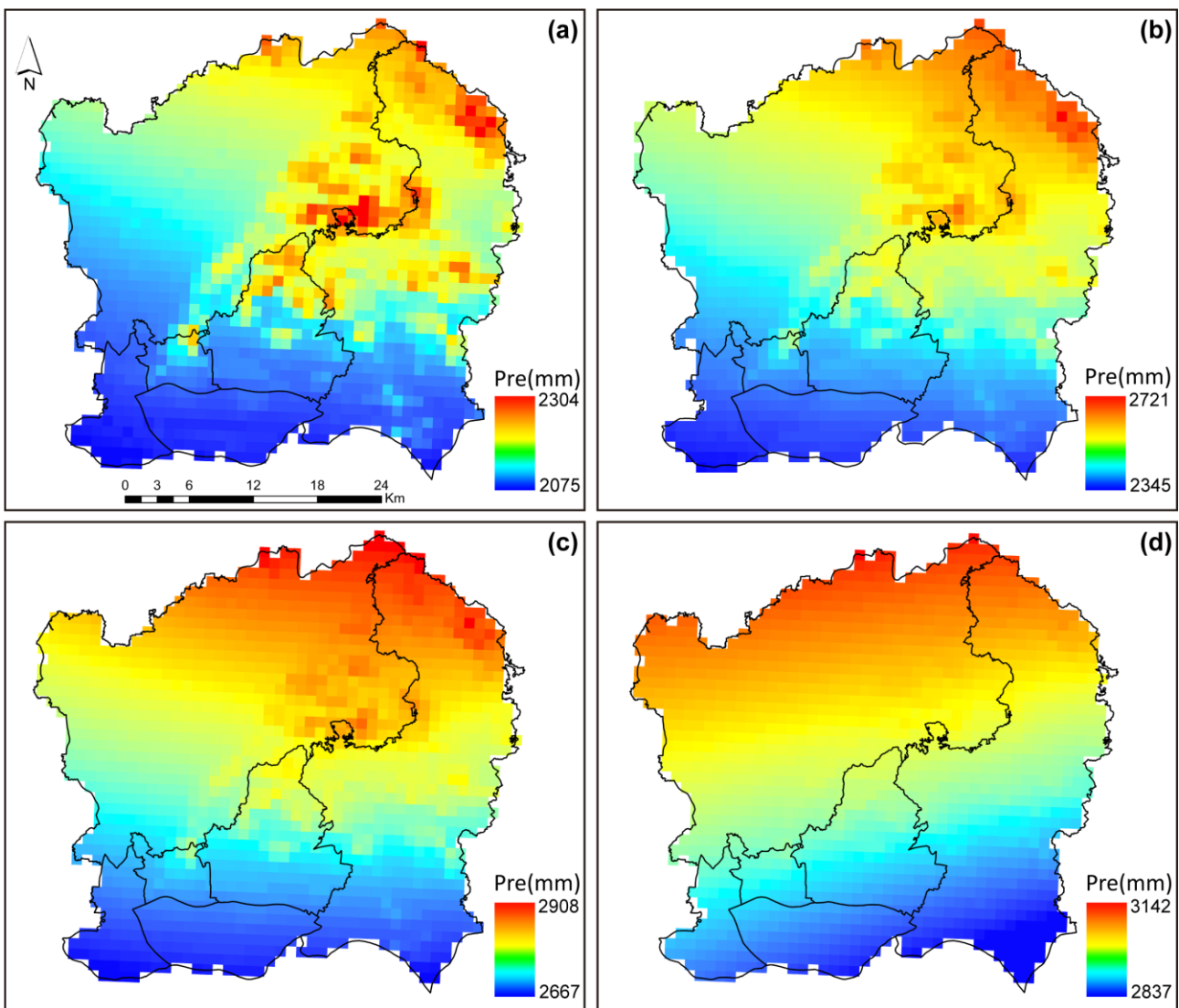


Figure S4-7 Different rainfall scenarios. (a) 10-year recurrence interval, (b) 25-year recurrence interval, (c) 50-year recurrence interval, (d) 100-year recurrence interval.



## References

1. Ahammed, F., 2017. A review of water-sensitive urban design technologies and practices for sustainable stormwater management. *Sustainable Water Resources Management*, 3. <https://doi.org/10.1007/s40899-017-0093-8>
2. Babaei, S., Ghazavi, R., & Erfanian, M., 2018. Urban flood simulation and prioritization of critical urban sub-catchments using SWMM model and PROMETHEE II approach. *Physics and Chemistry of the Earth*, 105 (July 2017), 3–11. <https://doi.org/10.1016/j.pce.2018.02.002>
3. Barros, V., & Stocker, T. F., 2012. Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change. *Journal of Clinical Endocrinology & Metabolism*, 18(6), 586-599. <https://doi.org/10.1017/CBO9781139177245>
4. Bisht, D. S., Chatterjee, C., Kalakoti, S., Upadhyay, P., Sahoo, M., & Panda, A., 2016. Modeling urban floods and drainage using SWMM and MIKE URBAN: a case study. *Natural Hazards*, 84(2), 749–776. <https://doi.org/10.1007/s11069-016-2455-1>
5. Brito, M. M., Almoradie, A., & Evers, M., 2019. Spatially-explicit sensitivity and uncertainty analysis in a MCDA-based flood vulnerability model. *International Journal of Geographical Information Science*, 33(9), 1788–1806. <https://doi.org/10.1080/13658816.2019.1599125>
6. Burger, G., Sitzenfrei, R., Kleidorfer, M., & Rauch, W., 2014. Parallel flow routing in SWMM 5. *Environmental Modelling and Software*, 53, 27–34. <https://doi.org/10.1016/j.envsoft.2013.11.002>
7. Chen, Y., Zhou, H., Zhang, H., Du, G., & Zhou, J., 2015. Urban flood risk warning under rapid urbanization. *Environmental Research*, 139, 3–10. <https://doi.org/10.1016/j.envres.2015.02.028>
8. Cheng, T., Xu, Z., Hong, S., & Song, S., 2017. Flood Risk Zoning by Using 2D Hydrodynamic Modeling: A Case Study in Jinan City. *Mathematical Problems in Engineering*, 2017. <https://doi.org/10.1155/2017/5659197>
9. China Global Television Network. [https://news.cgtn.com/news/3d597a4d34637a4d/share\\_p.html](https://news.cgtn.com/news/3d597a4d34637a4d/share_p.html)
10. Chowdary, V. M., Chakraborty, D., Jeyaram, A., Murthy, Y. V. N. K., Sharma, J. R., & Dadhwal, V. K., 2013. Multi-Criteria Decision Making Approach for Watershed Prioritization Using Analytic Hierarchy Process Technique and GIS. *Water Resources Management*, 27(10), 3555–3571. <https://doi.org/10.1007/s11269-013-0364-6>

11. Fan, Y. R., Huang, G. H., Li, Y. P., Wang, X. Q. , & Li, Z., 2016. Probabilistic prediction for monthly streamflow through coupling stepwise cluster analysis and quantile regression methods. *Water Resources Management*, 30(14), 5313-5331. <https://doi.org/10.1007/s11269-016-1489-1>
12. Fan, Y. R., Huang, W., Huang, G. H., Li, Z., Li, Y. P., & Wang, X. Q. , et al., 2015. A stepwise-cluster forecasting approach for monthly streamflows based on climate teleconnections. *Stochastic Environmental Research and Risk Assessment*, 29(6), 1557-1569. <https://doi.org/10.1007/s00477-015-1048-y>
13. Felder, G., Zischg, A., & Weingartner, R., 2017. The effect of coupling hydrologic and hydrodynamic models on probable maximum flood estimation. *Journal of Hydrology*, 550, 157–165. <https://doi.org/10.1016/j.jhydrol.2017.04.052>
14. Guangzhou Meteorological Service. <http://www.tqyb.com.cn/>
15. Guangzhou Planning and Natural Resources Bureau. <http://ghzyj.gz.gov.cn/>
16. Guangzhou Water Authority. <http://swj.gz.gov.cn/>
17. Gupta, A., Bansal, A., Gupta, R., Naryani, D., & Sood, A., 2017. Urban Waterlogging Detection and Severity Prediction Using Artificial Neural Networks. 2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS). IEEE. <https://doi.org/10.1109/HPCC-SmartCity-DSS.2017.6>
18. Hallegatte, S., Green, C., Nicholls, R. J., & Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. *Nature Climate Change*, 3, 802. <https://doi.org/10.1038/nclimate1979>
19. Hong, H., Pradhan, B., Bui, D.T., Xu, C., Youssef, A.M., Chen, W., 2016. Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Sichuan area (China). *Geomat. Nat. Haz. Risk* 1–26. <https://doi.org/10.1080/19475705.2016.1250112>
20. Hong, H., Tsangaratos, P., Iliu, I., Liu, J., Zhu, A. X., & Chen, W., 2018. Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. *Science of the Total Environment*, 625, 575–588. <https://doi.org/10.1016/j.scitotenv.2017.12.256>
21. Huang G., Shen Z., Mardin R. (2019) Overview of Urban Planning and Water-Related Disaster Management. In: Huang G., Shen Z. (eds) *Urban Planning and Water-related Disaster Management*.

Strategies for Sustainability. Springer, Cham. [https://doi.org/10.1007/978-3-319-90173-2\\_1](https://doi.org/10.1007/978-3-319-90173-2_1)

22. Huang, H., Chen, X., Zhu, Z., Xie, Y., Liu, L., Wang, X., ... Liu, K., 2018. The changing pattern of urban flooding in Guangzhou, China. *Science of the Total Environment*, 622–623, 394–404. <https://doi.org/10.1016/j.scitotenv.2017.11.358>
23. Huang, T., Wang, Y., & Zhang, J., 2017. Simulation and Evaluation of Low Impact Development of Urban Residential District Based on SWMM and GIS. *IOP Conference Series: Earth and Environmental Science*, 74(1). <https://doi.org/10.1088/1755-1315/74/1/012009>
24. IPCC., 2011. Intergovernmental Panel on Climate Change Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Cambridge: Cambridge University Press.
25. Jiang, W., Chen, Z., Lei, X., Jia, K., & Wu, Y., 2015. Simulating urban land use change by incorporating an autologistic regression model into a CLUE-S model. *Journal of Geographical Sciences*, 25(7), 836–850. <https://doi.org/10.1007/s11442-015-1205-8>
26. Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A., 2012. An artificial neural network model for flood simulation using gis: Johor river basin, Malaysia. *Environmental Earth Sciences*, 67(1), 251-264. <https://doi.org/10.1007/s12665-011-1504-z>
27. Lai, W. L., Wang, H. R., Wang, C., Zhang, J., & Zhao, Y., 2017. Waterlogging risk assessment based on self-organizing map (som) artificial neural networks: a case study of an urban storm in Beijing. *Journal of Mountain Science*. <https://doi.org/10.1007/s11629-016-4035-y>
28. Li, C., Cheng, X., Li, N., Du, X., Yu, Q., & Kan, G., 2016. A framework for flood risk analysis and benefit assessment of flood control measures in Urban Areas. *International Journal of Environmental Research and Public Health*, 13(8). <https://doi.org/10.3390/ijerph13080787>
29. Lin, Y., Chu, H., Wu, C., & Verburg, P. H., 2011. Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling – a case study. *International Journal of Geographical Information Science*, 25(1), 65–87. <https://doi.org/10.1080/13658811003752332>
30. Miao, Z. T., Han, M., & Hashemi, S., 2019. Correction to: The effect of successive low-impact development rainwater systems on peak flow reduction in residential areas of Shizhuang, China (*Environmental Earth Sciences*, (2019), 78, 2, (51), 10.1007/s12665-018-8016-z). *Environmental Earth Sciences*, 78(3), 1–3. <https://doi.org/10.1007/s12665-019-8088-4>

31. Nally, R. Mac., 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between – and reconciliation of – “predictive” and “explanatory” models. *Biodiversity and Conservation*, 9, 655–671. <https://doi.org/10.1023/A:1008985925162>
32. National Bureau of Statistics. <http://www.stats.gov.cn/>
33. Neuhold, C., Stanzel, P., & Nachtnebel, H. P., 2011. Integrating River Bed Dynamics to Flood Risk Assessment. *Sediment Transport in Aquatic Environments*. InTech. <https://doi.org/10.5772/20307>
34. Paiva, R. C. D., Collischonn, W., & Tucci, C. E. M., 2011. Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *Journal of Hydrology*, 406(3–4), 170–181. <https://doi.org/10.1016/j.jhydrol.2011.06.007>
35. Pradhan, B., 2012. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using gis. *Computers & Geosciences*, 51(2), 350-365. <http://dx.doi.org/10.1016/j.cageo.2012.08.023>
36. Quan, R., 2014. Risk assessment of flood disaster in Shanghai based on spatial–temporal characteristics analysis from 251 to 2000. *Environmental Earth Sciences*, 72(11), 4627–4638. <https://doi.org/10.1007/s12665-014-3360-0>
37. R Core Development Team., 2008. R: A Language and Environment for Statistical Computing. <http://www.R-project.org/>
38. Rai, P. K., Chahar, B. R., & Dhanya, C. T., 2017. GIS-based SWMM model for simulating the catchment response to flood events. *Hydrology Research*, 48(2), 384–394. <https://doi.org/10.2166/nh.2016.260>
39. Resources and Environmental Data Cloud Platform. <http://www.resdc.cn/>
40. Riley, S. J., S. D. DeGloria and R. Elliot., 1999. A terrain ruggedness index that quantifies topographic heterogeneity, *Intermountain Journal of Sciences*, vol. 5, No. 1-4,1999.
41. Samanta, S., Koloa, C., Pal, D. K., & Palsamanta, B., 2016. Flood risk analysis in lower part of Markham river based on multi-criteria decision approach (MCDA). *Hydrology*, 3(3), 1–13. <https://doi.org/10.3390/hydrology3030029>
42. Shao, W., Zhang, H., Liu, J., Yang, G., Chen, X., Yang, Z., & Huang, H., 2016. Data Integration and its Application in the Sponge City Construction of CHINA. *Procedia Engineering*, 154, 779–786. <https://doi.org/10.1016/j.proeng.2016.07.583>
43. Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., & Smith, D. R., 2005. Impacts of

impervious surface on watershed hydrology: A review. *Urban Water Journal*, 2(4), 263–275.  
<https://doi.org/10.1080/15730620500386529>

44. Sofia, G., Prosdocimi, M., Dalla Fontana, G., Tarolli, P., 2014. Modification of artificial drainage networks during the past half-century: Evidence and effects in a reclamation area in the Veneto floodplain (Italy), *Anthropocene*, 6, 48-62. <https://doi.org/10.1016/j.ancene.2014.06.005>
45. Sofia, G., Roder, G., Dalla Fontana, G., & Tarolli, P., 2017. Flood dynamics in urbanised landscapes: 100 years of climate and humans' interaction. *Scientific Reports*, 7(July 2016), 1–12. <https://doi.org/10.1038/srep40527>
46. Su, M., Zheng, Y., Hao, Y., Chen, Q., Chen, S., Chen, Z., & Xie, H., 2018. The influence of landscape pattern on the risk of urban water-logging and flood disaster. *Ecological Indicators*, 92, 133–140. <https://doi.org/10.1016/j.ecolind.2017.03.008>
47. Sun, J., Li, Y. P., Gao, P. P., Suo, C., & Xia, B. C., 2018. Analyzing urban ecosystem variation in the city of dongguan: a stepwise cluster modeling approach. *Environmental Research*, 166(OCT.), 276-289. <https://doi.org/10.1016/j.envres.2018.06.009>
48. Sun, J., Li, Y. P., Suo, C., & Huang, G. H., 2019. Identifying changes and critical drivers of future temperature and precipitation with a hybrid stepwise-cluster variance analysis method. *Theoretical and Applied Climatology*. <https://doi.org/10.1007/s00704-018-02758-9>
49. Tang, X., Hong, H., Shu, Y., Tang, H., Li, J., & Liu, W., 2019. Urban waterlogging susceptibility assessment based on a PSO-SVM method using a novel repeatedly random sampling idea to select negative samples. *Journal of Hydrology*, 576(June), 583–595. <https://doi.org/10.1016/j.jhydrol.2019.06.058>
50. Tang, X., Shu, Y., Lian, Y., Zhao, Y., & Fu, Y., 2018a. A spatial assessment of urban waterlogging risk based on a Weighted Naïve Bayes classifier. *Science of the Total Environment*, 630, 264–274. <https://doi.org/10.1016/j.scitotenv.2018.02.172>
51. Tang, Z., Yi, S., Wang, C., & Xiao, Y., 2018b. Incorporating probabilistic approach into local multi-criteria decision analysis for flood susceptibility assessment. *Stochastic Environmental Research and Risk Assessment*, 32(3), 701–714. <https://doi.org/10.1007/s00477-017-1431-y>
52. Tehrany, M. S., Jones, S., & Shabani, F., 2019. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena*, 175(December 2018), 174–192. <https://doi.org/10.1016/j.catena.2018.12.011>

53. Tehrany, M. S., Pradhan, B., & Jebur, M. N., 2013. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology*, 504, 69–79. <https://doi.org/10.1016/j.jhydrol.2013.09.034>
54. Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N., 2015. Flood susceptibility assessment using gis-based support vector machine model with different kernel types. *Catena*, 125, 91-101. <http://dx.doi.org/10.1016/j.catena.2014.10.017>
55. Tehrany, M.S., Pradhan, B., Jebur, M.N., 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J. Hydrol.* 512, 332–343. <https://doi.org/10.1016/j.jhydrol.2014.03.008>
56. Tsanis, I. K., & Boyle, S., 2001. A 2D hydrodynamic/pollutant transport GIS model. *Advances in Engineering Software*, 32(5), 353–361. [https://doi.org/10.1016/S0965-9978\(00\)00098-3](https://doi.org/10.1016/S0965-9978(00)00098-3)
57. Verburg, P. H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., & Mastura, S. S. A., 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. *Environmental Management*, 30(3), 391–405. <https://doi.org/10.1007/s00267-002-2630-x>
58. Wang, J., Gao, W., Xu, S., & Yu, L., 2012. Evaluation of the combined risk of sea level rise, land subsidence, and storm surges on the coastal areas of Shanghai, China. *Climatic Change*, 115(3–4), 537–558. <https://doi.org/10.1007/s10584-012-0468-7>
59. Wang, X., Huang, G., Lin, Q., Nie, X., Cheng, G., Fan, Y., ... Suo, M., 2013. A stepwise cluster analysis approach for downscaled climate projection - A Canadian case study. *Environmental Modelling and Software*, 49, 141–151. <https://doi.org/10.1016/j.envsoft.2013.08.006>
60. Wang, X., Huang, G., Zhao, S., & Guo, J., 2015. An open-source software package for multivariate modeling and clustering: applications to air quality management. *Environmental Science and Pollution Research*, 22(18), 14220–14233. <https://doi.org/10.1007/s11356-015-4664-7>
61. Werner, M., Hunter, N. M., & Bates, P. D., 2005. Identifiability of Distributed Floodplain Roughness Values in Flood Extent Estimation. *Journal of Hydrology*, 314, 139–157. <https://doi.org/10.1016/j.jhydrol.2005.03.012>
62. Wu, J., & Zhang, P., 2017. The effect of urban landscape pattern on urban waterlogging. *Dili Xuebao/Acta Geographica Sinica*, 72(3), 444–456. <https://doi.org/10.11821/dlxb201703007>
63. Wu, J., Yang, R., & Song, J., 2018. Effectiveness of low-impact development for urban inundation risk mitigation under different scenarios: A case study in Shenzhen, China. *Natural Hazards and*

Earth System Sciences, 18(9), 2525–2536. <https://doi.org/10.5194/nhess-18-2525-2018>

64. Xue, F., Huang, M., Wang, W., & Zou, L., 2016. Numerical Simulation of Urban Waterlogging Based on FloodArea Model. *Advances in Meteorology*, 2016. <https://doi.org/10.1155/2016/3940707>
65. Youssef, A. M., Pradhan, B., & Hassan, A. M., 2011. Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environmental Earth Sciences*, 62(3), 611–623. <https://doi.org/10.1007/s12665-010-0551-1>
66. Yu, H., Zhao, Y., Fu, Y., & Li, L., 2018. Spatiotemporal variance assessment of urban rainstorm waterlogging affected by impervious surface expansion: A case study of Guangzhou, China. *Sustainability (Switzerland)*, 10(10). <https://doi.org/10.3390/su10103761>
67. Zhang, B., 2015. Flood Disaster of China 2014. *China Flood and Drought Management*, (1), 59–65.
68. Zhang, H., Cheng, J., Wu, Z., Li, C., Qin, J., & Liu, T., 2018a. Effects of impervious surface on the spatial distribution of urban waterlogging risk spots at multiple scales in Guangzhou, South China. *Sustainability (Switzerland)*, 10(5). <https://doi.org/10.3390/su10051589>
69. Zhang, Q., Wu, Z., Zhang, H., Dalla Fontana, G., & Tarolli, P., 2020. Identifying dominant factors of waterlogging events in metropolitan coastal cities: the case study of Guangzhou, China. *Journal of Environmental Management*, 271, 110951. <https://doi.org/10.1016/j.jenvman.2020.110951>
70. Zhang, S. L., Gan, J. Y., Zeng, Q. L., & Lu, G. N., 2007. Automatic compartmentalization of urban rainwater catchments on water outlet supported by gis technology. *Journal of Hydraulic Engineering*, 38(3), 325-294.
71. Zhang, S., & Pan, B., 2014. An urban storm-inundation simulation method based on GIS. *Journal of Hydrology*, 517, 260–268. <https://doi.org/10.1016/j.jhydrol.2014.05.044>
72. Zhang, Y., Xia, J., Yu, J., Randall, M., Zhang, Y., Zhao, T., Shao, Q., 2018b. Simulation and assessment of urbanization impacts on runoff metrics: insights from landuse changes. *Journal of Hydrology*, 560, 247–258. <https://doi.org/10.1016/j.jhydrol.2018.03.031>
73. Zhao, G., Pang, B., Xu, Z., Peng, D., & Xu, L., 2019. Assessment of urban flood susceptibility using semi-supervised machine learning model. *Science of the Total Environment*, 659, 940–949. <https://doi.org/10.1016/j.scitotenv.2018.12.217>
74. Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T., 2018. Mapping flood susceptibility in mountainous areas on a national scale in China. *Science of the Total Environment*, 615, 1133–1142. <https://doi.org/10.1016/j.scitotenv.2017.10.037>

75. Zhuang, X.W., Li, Y.P., Huang, G.H. et al., 2016. Assessment of climate change impacts on watershed in cold-arid region: an integrated multi-GCM-based stochastic weather generator and stepwise cluster analysis method. *Clim Dyn* 47, 191–209. <https://doi.org/10.1007/s00382-015-2831-7>
76. Zope, Eldho, J., 2016. Impacts of land use-land cover change and urbanization on flooding: A case study of Oshiwara River Basin in Mumbai, India. *Catena*, 145, 142–154. <https://doi.org/10.1016/j.catena.2016.06.009>



# CHAPTER 5

## 5. Investigating the Role of Green Infrastructure on Urban Waterlogging: Evidence from Metropolitan Coastal Cities<sup>4</sup>

Qifei Zhang<sup>a,b</sup>, Zhifeng Wu<sup>b,c,d</sup>, Paolo Tarolli<sup>a\*</sup>

- a. Department of Land, Environment, Agriculture and Forestry, University of Padova, Legnaro 35020, Italy.
- b. School of Geographical Sciences, Guangzhou University, Guangzhou 510006, China.
- c. Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou 511458, China.
- d. MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, Shenzhen 518000, China.

---

<sup>4</sup> *This chapter is a style-edited version of the published article:*

Zhang, Q., Wu, Z., & Tarolli, P. (2021). Investigating the Role of Green Infrastructure on Urban WaterLogging: Evidence from Metropolitan Coastal Cities. *Remote Sensing*, 13(12), 2341.

**Contributions by the PhD candidate included:** primary authorship in writing and revision of the entire article; primary responsible for literature review, data collection and quantification, methodology, data processing and analysis, and production of all figures.

## 5.1 Abstract

Urban green infrastructures (UGI) can effectively reduce surface runoff, thereby alleviating the pressure of urban waterlogging. Due to the shortage of land resources in metropolitan areas, it is necessary to understand how to utilize the limited UGI area to maximize the waterlogging mitigation function. Less attention, however, has been paid to investigating the threshold level of waterlogging mitigation capacity. Additionally, various studies mainly focused on the individual effects of UGI factors on waterlogging but neglected the interactive effects between these factors. To overcome this limitation, two waterlogging high-risk coastal cities—Guangzhou and Shenzhen, are selected to examine the effectiveness and stability of UGI in alleviating urban waterlogging. The results indicate that the impact of green infrastructure on urban waterlogging largely depends on its area and biophysical parameter. Healthier or denser vegetation (superior ecological environment) can more effectively intercept and store rainwater runoff. This suggests that while increasing the area of UGI, more attention should be paid to the biophysical parameter of vegetation. Hence, the mitigation effect of green infrastructure would be improved from the “size” and “health”. The interaction of composition and spatial configuration greatly enhances their individual effects on waterlogging. This result underscores the importance of the interactive enhancement effect between UGI composition and spatial configuration. Therefore, it is particularly important to optimize the UGI composition and spatial pattern under limited land resource conditions. Lastly, the effect of green infrastructure on waterlogging presents a threshold phenomenon. The excessive area proportions of UGI within the watershed unit or an oversized UGI patch may lead to a waste of its mitigation effect. Therefore, the area proportion of UGI and its mitigation effect should be considered comprehensively when planning UGI. It is recommended to control the proportion of green infrastructure at the watershed scale (24.4% and 72.1% for Guangzhou and Shenzhen) as well as the area of green infrastructure patches (1.9 ha and 2.8 ha for Guangzhou and Shenzhen) within the threshold level to maximize its mitigation effect. Given the growing concerns of global warming and continued rapid urbanization, these findings provide practical urban waterlogging prevention strategies toward practical implementations.

**Keywords:** urban waterlogging; green infrastructure; composition and spatial configuration; geographical detector model; nonlinear relationship

## 5.2 Introduction

Urban waterlogging is caused by surface runoff exceeding the local drainage capacity of a city due to short-term heavy rainfall [1,2,3]. With the acceleration of globalization, the natural surface within the city has undergone drastic changes [4,5]. This phenomenon leads to numerous social-environmental-ecological problems [6,7,8,9]. The driving factors and spatial variability of waterlogging have been extensively studied [10,11]. Specifically, the man-made land covers destroy the original urban hydrological cycle, which impedes the natural infiltration of rainwater and reduces the storage capacity of the underlying urban surface. These phenomena have led to the frequent occurrence of urban waterlogging events [12,13,14]. The Intergovernmental Panel on Climate Change (IPCC) fifth assessment report states that the intensity and frequency of extreme precipitation events have increased significantly [15]. Therefore, in the context of climate change and intense human activities, it will undoubtedly lead to the frequent occurrence of urban waterlogging disasters, posing a growing threat to human well-being. For example, in 2017, 104 cities in China suffered from urban waterlogging disasters, affecting 2.18 million people and causing direct economic losses of \$2.47 billion [16]. From 21 to 22 July 2012, Beijing suffered the strongest rainstorm and waterlogging disaster in 61 years (460 mm maximum precipitation). The torrential rain triggered flash floods, resulting in 79 deaths, 10,660 houses collapsing, and \$1875 billion in losses. Coincidentally, as the youngest city in China, in April 2019, a sudden, instantaneous heavy precipitation (maximum half an hour rainfall, 73.4 mm) caused 11 deaths in Shenzhen. Consequently, strengthening the ability to prevent waterlogging disasters has become an important issue of sustainable urban development and the UN's 2030 Sustainable Development Goals (SDGs).

The importance of mitigating the risk of urban waterlogging has been widely recognized by society. This requires understanding the mechanisms of urban waterlogging first. Considerable studies have shown that urban waterlogging events are caused by environmental factors and human activities [17,18,19,20]. In terms of environmental factors, urban waterlogging is mainly affected by meteorological conditions and urban microtopography. In the context of global warming, the frequency and intensity of extreme rainstorms have increased [20,21,22]. Moreover, the phenomenon that the precipitation in many cities is significantly higher than that in the surrounding suburbs has become more prominent in recent years, known as the "urban rain island" effect. If the "rain island effect" occurs concentratedly in the rainy season, which is more likely to cause waterlogging disasters. In urban microtopography, the area with higher elevation is less prone to

waterlogging. On the contrary, low-lying areas tend to accumulate surface runoff, which is why tunnels and underpasses are prone to waterlogging [10,21,23]. Among anthropogenic factors, drainage facilities and land cover composition have a significant impact on urban waterlogging. However, some studies have pointed out that the drainage facilities in developing countries generally suffer low design standards and inadequate management, which makes it difficult to play an active role in the face of heavy rainstorms [10,21]. In terms of land cover composition, numerous studies have shown that the impact of land use is particularly significant compared to other factors [24,25,26], which has gradually become the major cause of the increasing severity of urban waterlogging disasters.

At present, the Municipal Administration builds underground drainage pipelines and pumping stations to speed up rainwater drainage, thereby reducing the flow of surface runoff. However, this approach only accelerates the discharge rate of surface runoff but cannot reduce the total surface runoff. Surface runoff transfer to other regions in a short period may bring more pressure on the local drainage systems. Furthermore, drainage facilities block the recharge channel for groundwater, leading to a constant decline of groundwater level, threatening urban geological safety [27,28]. Compare with drainage pipelines to accelerate rainwater drainage, the concept of “sponge cities” or “low-impact development methods” has been proposed to reduce the total amount of surface runoff. These methods aim to increase the permeable surface (such as green infrastructure) in cities, thereby counteracting the increase of surface runoff [29,30]. Therefore, understanding how urban green infrastructure can alleviate urban waterlogging is of great significance for urban sustainable living environment planning and management.

Urban green infrastructures (UGI) mainly refer to natural vegetation or artificial vegetation, such as urban parks, grassland, wetland, forest, and unmanaged green areas [31,32,33,34]. A considerable number of studies have pointed out the extensive ecological services of UGI, such as reducing stormwater runoff [35,36], regulating local microclimates [37,38], and purifying rainwater [39,40]. In particular, UGI increased the permeability to regulate surface runoff and peak flows. A proliferation of studies have shown that UGI, as a permeable surface, can effectively absorb and store rainwater [41,42,43], and the canopy and rhizome of vegetation can intercept surface runoff, thereby reducing the speed of runoff collection [44,45]. For example, Yang et al. [43] used the improved soil and water conservation service mode to evaluate the average accumulation of urban green space in Yixing city, and their result indicated that the average water storage of urban green space accounted for more than 88% of the annual rainfall. Liu et al. [45] show the effectiveness

of UGI in urban flooding reduction at a community scale. These studies all demonstrated that UGI has a positive influence on waterlogging. Furthermore, some studies further examined the impact of UGI composition and spatial pattern on waterlogging [46,47]. As for UGI composition, Armson et al. [48] found that in a sample plot of 9 m<sup>2</sup> (the land cover includes grassland, trees, and asphalt), the grassland controlled almost all the surface runoff, and the trees reduced 62% of runoff from asphalt. Richards et al. [49] pointed out that a vegetated area of 7.5% to the catchment area would reduce surface runoff by more than 90%. For the spatial pattern of UGI, a study indicated that the less fragmented urban green spaces are more effective in reducing peak annual average river runoff [41]. In addition to the UGI composition and spatial configuration, the morphology of UGI also had a substantial influence on urban waterlogging mitigation. The study in Shanghai (China) confirmed that the concave green space could effectively mitigate pluvial floods [36]. Similarly, Wen et al. [50] demonstrated that a concave-shaped UGI would significantly reduce the surface runoff and peak flood flows.

The above studies have demonstrated that the impact of UGI on urban waterlogging is associated with various factors, such as UGI composition, spatial configuration, and morphology. Considerable studies have examined the relationship between UGI's factors and waterlogging through regression coefficients or hydrologic models [41,45,51,52]. However, previous studies mainly focused on the individual effects of UGI factors (composition or spatial configuration) on urban waterlogging; instead, the interactive effects of these factors remain unclear. The influence of UGI on urban waterlogging is not only affected by one factor alone. Only analyzing the individual effect of a UGI factor on urban waterlogging while ignoring the interactive effect may lead to biases, especially for the great heterogeneity urbanized area. From this perspective, some interesting questions emerge: How do the interactions of these UGI factors affect urban waterlogging? Can the interaction between different UGI factors further enhance their effects on waterlogging?

It is widely accepted that increasing the area of UGI may further increase its impact on urban waterlogging, thereby reducing the risk of urban waterlogging. However, will the risk of urban waterlogging continue to decrease as the area of UGI increases? In this context, another research question arises: Is there a threshold level for the effect of UGI on urban waterlogging? Less attention has been paid to investigating the threshold level for the impact of UGI on urban waterlogging. Moreover, given the shortage of urban land resources, it is unrealistic to reduce urban waterlogging by considerably increasing the UGI area. If the effect of UGI has a threshold level, planning a larger area of UGI may not provide a more significant mitigation effect. Therefore,

it is necessary to understand the threshold level of UGI affecting urban waterlogging so that the limited UGI resource can be used to minimize the negative influence of urban waterlogging. Additionally, it is worth noting that many studies just involved a single city or region, which present inconsistent results among the studies [41,44]. These inconsistent results are not sufficient to fully examine the effect of UGI on waterlogging, which makes it difficult to apply in UGI planning and urban management. This highlights the urgency of conducting cross-regional comparative studies to further verify the universal effect of UGI on urban waterlogging.

Therefore, this study aims to shed some light on the above two research gaps by taking two waterlogging high-risk Chinese cities for a comparative study to address the following questions: (1) How does the interaction effect of UGI's factors affect urban waterlogging? Which UGI factors are the dominant factors affecting urban waterlogging? (2) Is there a threshold level for the impact of UGI on urban waterlogging? Answering these questions can help us improve our understanding of the potential mitigation effect of UGI on urban waterlogging and furnish concrete references for UGI design.

## 5.3 Materials and methods

### 5.3.1 Study area and data

#### 5.3.1.1 Study area

Two major cities in the Guangdong–Hong Kong–Macao Greater Bay Metropolitan Region, Guangzhou and Shenzhen cities, are selected for this study (Figure 5-1). Guangzhou City (112°57' to 114°30'E, 22°26' to 23°56'N) is located in the downstream of the Pearl River Basin, the central and southern part of Guangdong Province, with an area of 7434.40 km<sup>2</sup>. Shenzhen City (113°45' to 114°37'E, 22°26' to 22°51'N), with an area of 1997.47 km<sup>2</sup>, is located on the eastern bank of the Pearl River Estuary. The average annual precipitation of Guangzhou and Shenzhen are 1720.6 mm and 1933.3 mm, respectively, belonging to subtropical monsoon climate [53,54]. The two cities are among the four national cities in mainland China, with a permanent resident population of 15.31 million (Guangzhou) and 13.44 million (Shenzhen), respectively, which together account for 47% of Guangdong's GDP in 2019 (\$756 billion; <http://www.stats.gov.cn/>; access on 8 January 2021).

With the substantial increase of extreme rainfall events, urban waterlogging events frequently occur in these two low-lying coastal cities [2]. For example, on 22 May 2020, four people were killed in extraordinarily heavy rainfall in Guangzhou, with an average hourly rain intensity exceeding 80 mm, and the maximum precipitation

in 3 hours at 288.5 mm. From 29–30 August 2018, a heavy rainstorm occurred in Shenzhen for two consecutive days (269 mm average cumulative precipitation, 97 mm maximum hourly precipitation), the first-ever recorded in local meteorological history. This event resulted in approximately 150 waterlogging events, 10 local riverbank collapses, and 37 landslides. Given the densely populated area and the serious risk of urban waterlogging disaster in this region, selecting these two cities to investigate the effect of UGI on urban waterlogging has a certain practical significance. Guangzhou Water Authority has only recorded the urban waterlogging events in the central urban districts (Liwan, Yuexiu, Tianhe, Haizhu, Baiyun, and Huangpu district). Hence, we select these central urban districts of Guangzhou (1559.82 km<sup>2</sup>) and Shenzhen city as our study sites.

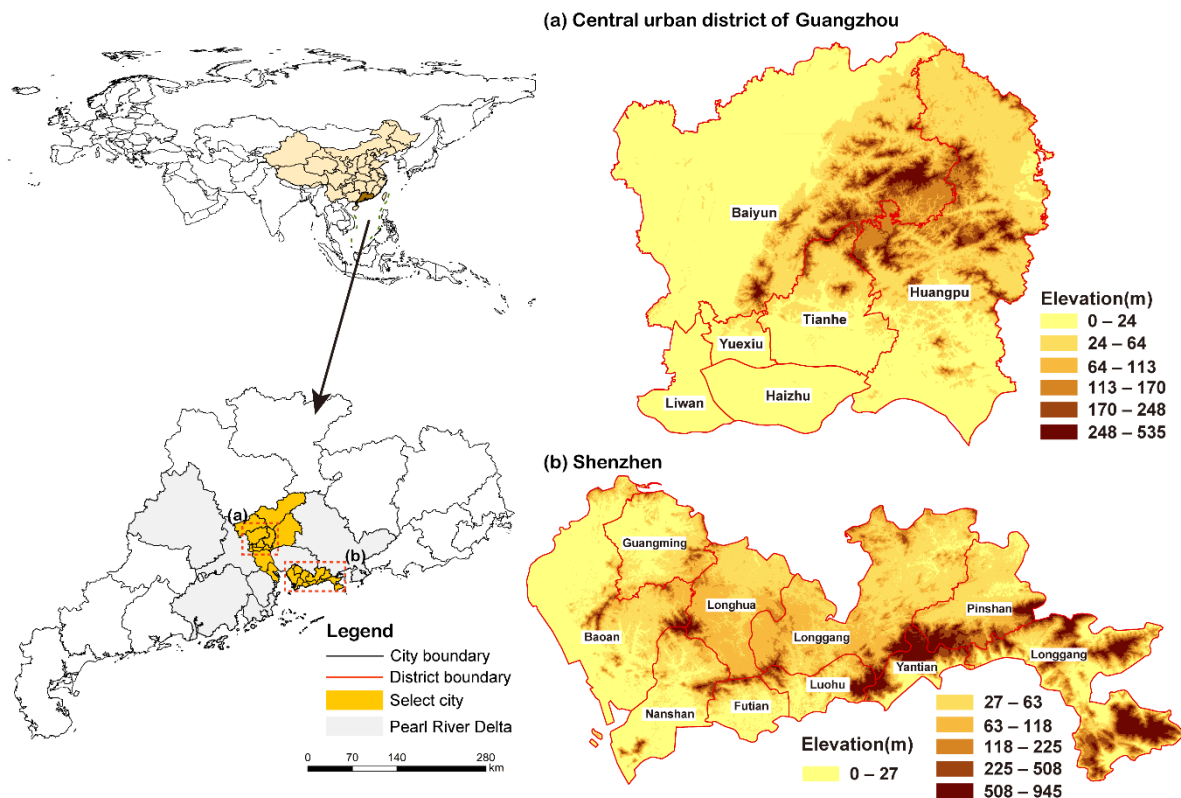


Figure 5-1 The location of the Guangzhou central urban districts and Shenzhen city.

### 5.3.1.2 Dataset

In this study, we concentrate on the period from 2009 to 2015. The data include urban waterlogging records, UAV images, Landsat-8 Operational Land Imager imagery, DEM, precipitation, and drainage facilities (Table 5-1). As mentioned in previous studies, the waterlogging records obtained from Guangzhou and Shenzhen Water Authority only contain location information [10,55]. Therefore, we utilized ArcGIS Pro to locate the

spatial location of urban waterlogging events. Finally, we collected 423 and 353 records in Guangzhou and Shenzhen from 2009 and 2015. The composition and spatial configuration of UGI were obtained from UAV aerial images (spatial resolution 0.5 m). The cloud-free Landsat-8 OLI imageries (path/row: 122-44, 121-44) were utilized in this study to calculate the biophysical parameter of UGI. Subsequently, we utilized DEM (spatial resolution 5 m, vertical accuracy 0.1 m) to generate auxiliary variables, including elevation and slope. Lastly, other auxiliary variables, such as precipitation and drainage density, were also collected. Local water authorities only recorded urban waterlogging events in this period (without a specific year). Therefore, we only selected remote sensing images, DEM data, drainage network, river network, and precipitation data in this period.

*Table 5-1 List of the data sources.*

Data	Format	Time	Detail	Source
Waterlogging locations	Shapefile	2009~2015	Point	Guangzhou Water Authority ( <a href="http://swj.gz.gov.cn/">http://swj.gz.gov.cn/</a> ) Shenzhen Water Authority ( <a href="http://swj.sz.gov.cn/">http://swj.sz.gov.cn/</a> )
Landsat-8 OLI imagery	GeoTIFF	2013	30 m path/row: 122-44, 121-44	The United States Geological Survey ( <a href="http://www.usgs.gov/">http://www.usgs.gov/</a> )
UAV images	Raster	2012	0.5 m	Guangzhou Planning and Natural Resources Bureau
Digital Elevation Model	Raster	2012	5 m	( <a href="http://ghzyj.gz.gov.cn/">http://ghzyj.gz.gov.cn/</a> )
Drainage network	Shapefile	2012	Line	Shenzhen Planning and Natural Resources Bureau
River network	Shapefile	2012	Line	( <a href="http://pnr.sz.gov.cn/">http://pnr.sz.gov.cn/</a> )
Precipitation	Raster	2009~2015	1 km	Geographical Information Monitoring Cloud Platform ( <a href="http://www.dsac.cn/">http://www.dsac.cn/</a> )

### 5.3.2 Integrated framework

The integrated framework was developed to analyze the interactive effects of UGI factors on waterlogging and quantify the threshold level (Figure 5-2). Urban waterlogging is a systemic problem. The occurrence of waterlogging is related to the destruction of the hydrological cycle in the watershed unit [55,56,57]. When the rainwater is unbalanced, rainfall or rainwater inflow exceeds the drainage capacity, urban waterlogging



events will eventually occur. The watershed unit reflects the hydrological characteristics of an area, which has more natural and ecological significance. It is not appropriate to analyze urban waterlogging from the perspective of a point or a raster grid (buffer zone), as it ignores the hydrodynamics of the surface. Therefore, we investigated the effect of UGI at the watershed level.

First, the density of waterlogging per unit area within each watershed unit was calculated based on the waterlogging record. Several metrics were utilized to measure green infrastructure composition and spatial configuration (area proportion, biophysical parameter, and spatial configuration). Second, other auxiliary variables, including elevation, slope, precipitation, and drainage density, were adopted as control variables. Then, the urban waterlogging density was regarded as a dependent variable, while the UGI composition and spatial configuration were considered explanatory variables. Fourth, the correlation between urban waterlogging density and explanatory variables was examined through partial correlation analysis. Fifth, the interaction effect of UGI factors on waterlogging we examined through the geographical detector model. Lastly, we quantified the threshold level of UGI affecting urban waterlogging using the logarithmic fitting method.

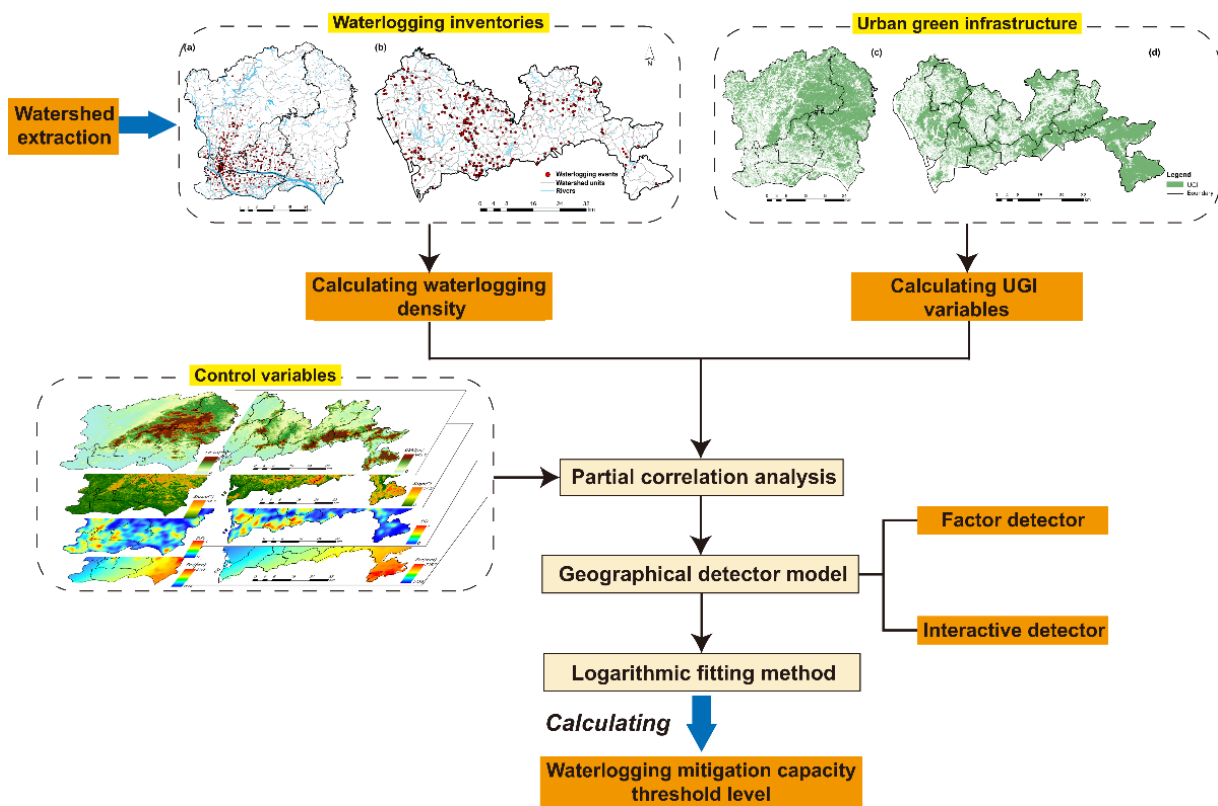


Figure 5-2 The integrated framework of this study.

### 5.3.3 Watershed unit

According to the method proposed by Yu et al. [55] and successfully applied in Zhang et al. [10], the DEM, urban river and drainage network, and the hydrological analysis module of ArcGIS pro were utilized to divide the watershed units through the D8 algorithm. Although the D8 algorithm is more efficient at the urban scale than other algorithms, including D-Infinity and MFD. Due to the flat topography of the Pearl River Basin, the extracted watershed boundaries need to be modified using urban rivers and drainage networks [57,58]. Finally, we divided the Guangzhou central urban district and Shenzhen into 351 and 276 watershed units (Figure 5-3).

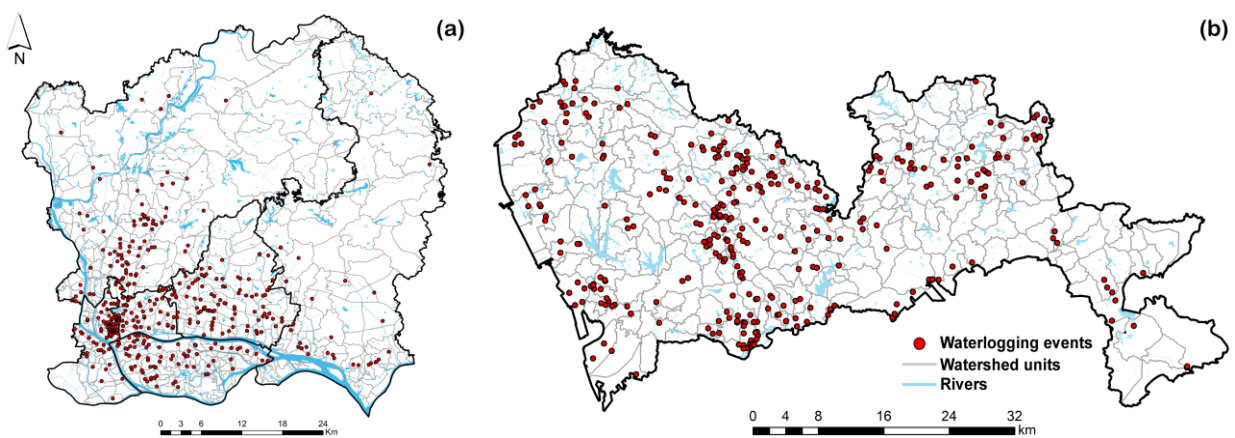


Figure 5-3 The urban waterlogging events and watershed units for (a) Guangzhou central urban district and (b) Shenzhen.

### 5.3.4 Measuring green infrastructure composition and spatial configuration

In this study, we mapped the green infrastructure of Guangzhou central urban district and Shenzhen using 0.5 m aerial images. These images were obtained from the Geographical Situation Survey Project (GSSP) in 2012. The flight missions were conducted with DB-2S and IFAUAV-3 platform, and the image forward overlap and image-side overlap were set as 83% and 57%. The images and POS information were then imported into Pix4D Mapper to create the stitching project. After performing automatic aerial triangulation, the images were corrected by acquiring ground control point (GCP) data and a third-order polynomial model. Finally, the horizontal RMSE values were 0.763 m and 0.871 m in Guangzhou and Shenzhen, sufficient for green infrastructure extraction.

According to the field research, the woodland, grassland, garden, and cultivated land were defined as UGI in this study. We extracted the green infrastructure through an object-oriented classification method using the

eCognition Developer software. The classification accuracy assessment was computed from ground-truthing analysis by randomly selecting over 100 points for each city. Ultimately, the overall accuracy of classification was 85.8%, 81.2%, and the kappa coefficient was 0.78, 0.75, for Guangzhou and Shenzhen, respectively (Figure 5-4). Subsequently, the area proportion of UGI within different watershed units was calculated.

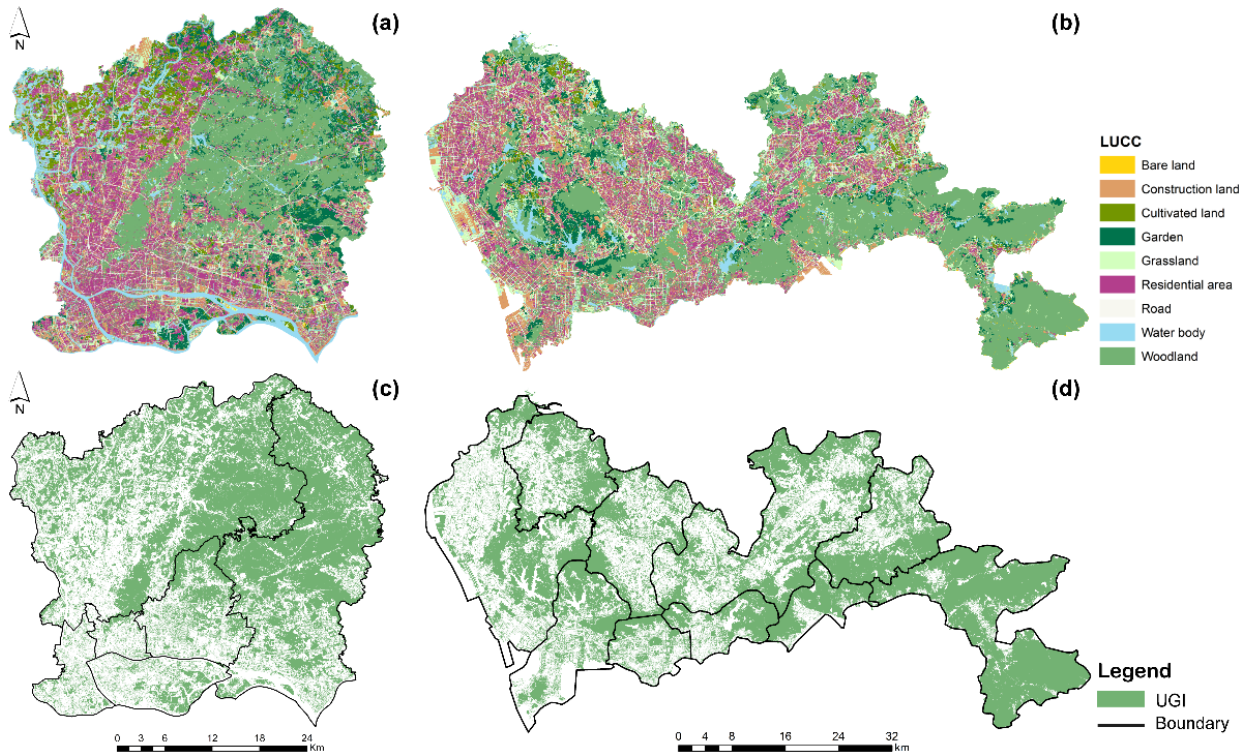


Figure 5-4 The land cover maps (a),(b) and UGI maps (c),(d) for Guangzhou central urban district (a),(c) and Shenzhen (b),(d).

The UGI area proportion refers to the area ratio of green infrastructure in a watershed unit, however, it could not reflect the biophysical parameter of UGI. Under the same green infrastructure coverage ratio, different vegetation growth statuses or densities have different effects on urban waterlogging. For example, dense vegetation may be more conducive to reducing surface runoff, while sparse vegetation may be less effective. Biophysical parameters should represent these vegetation gradients. Therefore, we used the enhanced vegetation index (EVI) to describe the biophysical parameter of UGI (Figure 5-5), which derived from the multispectral optical band (blue and red) and near-infrared (NIR) band of the Landsat-8 OLI images (Equation 5.1).

$$EVI = 2.5 \times \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6.0 \times \rho_{RED} - 7.5 \times \rho_{BLUE} + 1} \quad (5.1)$$

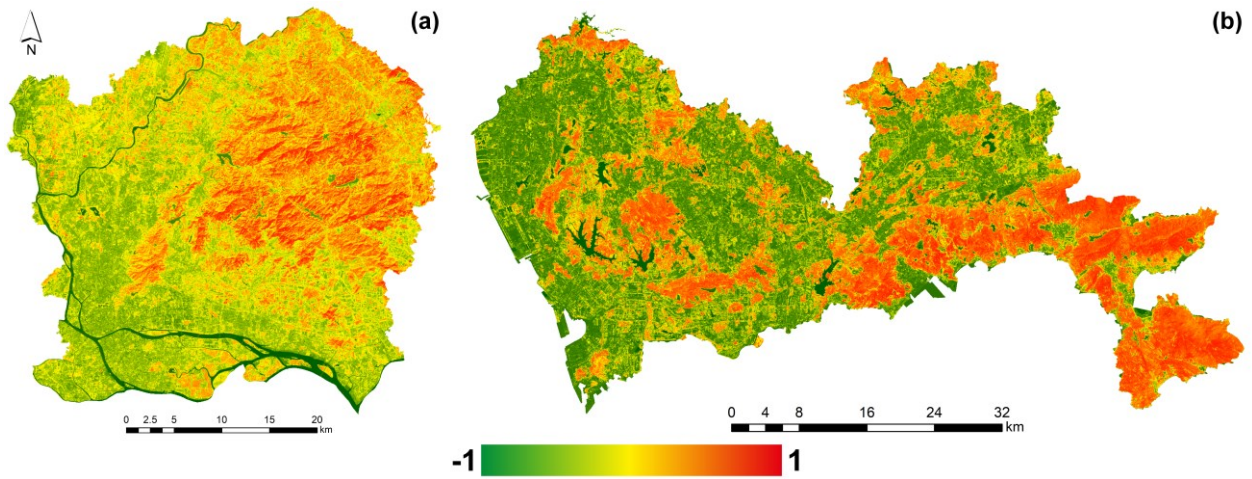


Figure 5-5 The EVI for (a) Guangzhou central urban district and (b) Shenzhen.

The spatial pattern of land cover features can be described by landscape pattern metrics [59]. In recent decades, the landscape pattern metrics have achieved unprecedented development, plenty of indicators have been developed to reveal the characteristics of landscape spatial patterns [60,61]. In this study, three landscape metrics were selected to reflect the spatial configuration characteristics of UGI, including (1) landscape fragmentation: mean patch size (MPS) and landscape division index (LDI); (2) landscape aggregation: aggregation index (AI). The equation and description of these UGI metrics were shown in Table 5-2 and were calculated through Fragstats 4.2.

Table 5-2 List of landscape pattern metrics.

Landscape metrics	Equation*	Description
MPS	$\sum_{i=1}^n \frac{A_i}{n}$	Reflects the average patch size of a certain land cover type.
LDI	$1 - \sum_{i=1}^n \left(\frac{A_i}{S}\right)^2$	Reflects the degree of fragmentation.
AI	$\left[ \frac{g_i}{\max \rightarrow g_i} \right]$	Measures the spatial distribution pattern.

\* $A_i$ : patch i area, S: total area, n: number of patches,  $g_i$ : number of adjacent patches.

### 5.3.5 Control variables

Urban waterlogging is the result of the combination of environmental conditions and human activities. In addition to the significant influence of land cover composition (UGI) on urban waterlogging, the urban microtopography, rainfall, and drainage facilities also have a non-negligible impact on urban waterlogging

[10,21]. In order to accurately quantify the effect of UGI on urban waterlogging, it is essential to exclude the influence of other relevant variables on waterlogging. Therefore, the topography (elevation and slope), average precipitation, and drainage density are adopted as control variables to avoid these distractions (Figure 5-6). Firstly, the topographic variables of elevation and slope were calculated from the DEM data through ArcGIS Pro. Then, as the urban waterlogging record does not include the specific years, this study used the average cumulative precipitation (Pre) reflecting the spatial distribution difference of rainfall during this period. Lastly, we calculated the drainage network density (DD) by line density module in ArcGIS pro to reflect the drainage capacity.

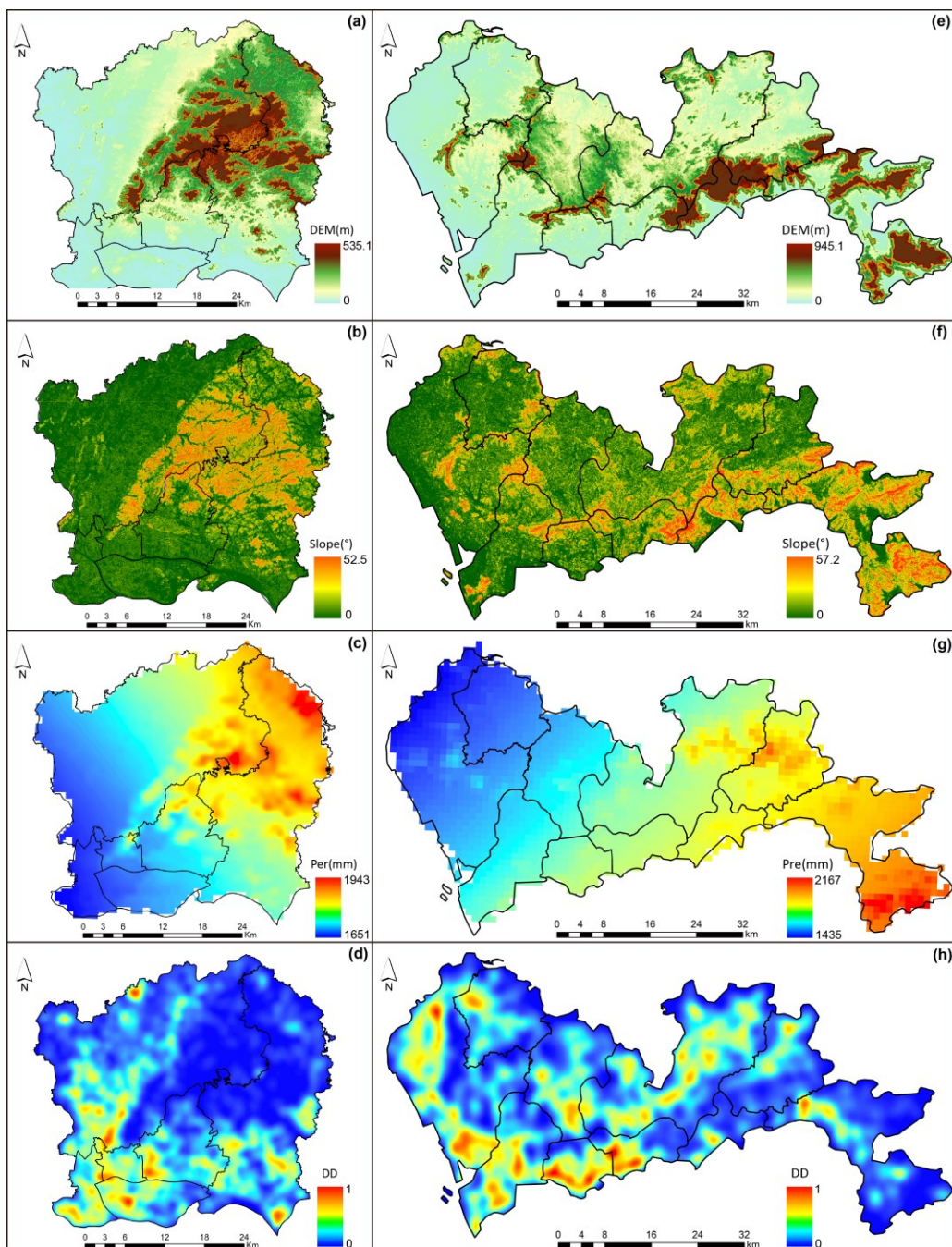


Figure 5-6 The auxiliary variables for (a)–(d) Guangzhou central urban district and (e)–(h) Shenzhen.

### 5.3.6 Statistical analyses

#### 5.3.6.1 UGI and waterlogging clusters extraction

In this study, the spatial autocorrelation analysis Getis-G statistic was adopted to investigate the spatial distribution pattern of UGI and urban waterlogging events. The Getis-G statistic allows us to detect whether the elements (green infrastructure and water-logging event) are clustered, discrete, or randomly distributed, which has been widely applied in geography and economy [38,62]. This allows the identification of spatial agglomeration effect with statistical significance (99%, 95%, 90% confidence level). In this study, the proportion of UGI and waterlogging density within each watershed unit was used as input attributes to distinguish the spatial agglomeration effect (hot spots or cold spots) of UGI and waterlogging. The calculation formula is as follows:

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n \omega_{i,j} x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n x_i x_j}, \forall j \neq i \quad (5.2)$$

where  $\omega_{i,j}$  is the spatial weight matrix;  $x_i$  and  $x_j$  are the attribute values of the  $i$  and  $j$  variables respectively. The Getis-G statistic will return five values: General G observed value, General G expected value, Z-test, and p-value. A positive Z-test indicates a high value of the attribute (UGI proportion and waterlogging density) spatial clustering (hot spots), which means that the density of urban waterlogging or the area proportion of UGI in the region is relatively large. Conversely, a negative Z-test indicates a low value (UGI proportion and waterlogging density) of spatial clustering (cold spots), which implies that the density of urban waterlogging or the area proportion of UGI in the region is relatively low. There were six main cluster types, and the specific meanings were as follows:

The waterlogging or UGI hot spots at 99%, 95%, 90% confidence level: The density of urban waterlogging events or the area proportion of UGI in the watershed unit and its adjacent watersheds are significantly higher than the average level, indicating that urban waterlogging events or UGI distribution are concentrated in a place.

The waterlogging or UGI cold spots at 99%, 95%, 90% confidence level: The density of urban waterlogging events or the area proportion of UGI in the watershed unit and its surrounding units are relatively lower than the average level, which implies urban waterlogging events and UGI distribution are much fewer in the region.

### 5.3.6.2 Partial correlation analysis

The partial correlation analysis was first used to reveal the binary correlation of waterlogging and UGI. As urban waterlogging is a systemic problem, the relationship between UGI and urban waterlogging is affected by multiple variables [10,17,21]. For example, improving the condition of green infrastructure or increasing drainage facilities can both reduce the risk of urban waterlogging. Therefore, to accurately measure green infrastructure's effect on urban waterlogging, the partial correlation analysis with control variables was utilized to examine the correlation and stability between UGI factors and urban waterlogging density. The partial correlation can effectively prevent the correlation between two variables from being contaminated by other correlations. In this case, any influencing factors that have potential effects on urban waterlogging were regarded as control variables for partial correlation analysis. Hence, the elevation, slope, precipitation, and drainage density were adopted as control variables to examine UGI composition and spatial configuration for its partial correlation with urban waterlogging.

### 5.3.6.3 Geographical detector model

Spatial heterogeneity is a major characteristic of spatial data. The geographical detector model is a spatial statistic tool based on stratified spatial heterogeneity, which has been widely used to investigate spatial heterogeneity and driving forces of geographical phenomena [63,64,65]. The geographical detector model can be divided into four parts according to their specific analytical functions: factor detector, risk detector, ecological detector, and interactive detector [66]. As the main purpose of this study, the factor detector and interactive detector were employed to reveal which green infrastructure factor has a more important impact on urban waterlogging and how these factors interact with each other.

The explanatory power of different factors to the dependent variable can be expressed by the PD value (power determinant) calculated by the factor detector. It compares the total variance of the factor in different subregions with the total variance of the factor in the whole study area to assess the impact of the green infrastructure factor on waterlogging:

$$PD = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} = 1 - \frac{SSW}{SST} \quad (5.3)$$

where  $N$  is the number of units across the region;  $N_h$  indicate the number of units in stratum  $h$ ,  $h = 1, 2, \dots, L$ ;  $\sigma^2$  and  $\sigma_h^2$  represent the global variance in the entire study area and variance in stratum  $h$ ;  $SSW$  and  $SST$  refer to the total variance within stratum and across the region. The value of  $PD$  ranges from 0 to 1. When  $PD$

= 1, the UGI factor fully explains the spatial distribution of urban waterlogging; when PD = 0, the UGI factor has no relationship with the variation of waterlogging.

The interaction detector is used to detect whether the combined effect of two individual factors on a dependent variable is significantly greater or less than the individual effect of a single factor [63]. It is determined by comparing the sum of the PD values of the two factors with the PD values of the two-factor interaction. The interaction detector divides the interactions between two factors into seven categories, as shown in Table 5-3.

Table 5-3 Types of interaction between two factors.

Interaction	Description
Nonlinear weaken	$PD(X1 \cap X2) < \text{Min}[PD(X1), PD(X2)]$
Unitary weaken	$\text{Min}[PD(X1), PD(X2)] < PD(X1 \cap X2) < \text{Max}[PD(X1), PD(X2)]$
Binary enhancement	$PD(X1 \cap X2) > \text{Max}[PD(X1), PD(X2)]$
Independent	$PD(X1 \cap X2) = PD(X1) + PD(X2)$
Nonlinear enhancement	$PD(X1 \cap X2) > PD(X1) + PD(X2)$

#### 5.3.6.4 Thresholds level of UGI affecting waterlogging

First, we plot the relationship between waterlogging density and UGI factors to further analyze this complex linking. As shown in Figure 5-7, we notice that waterlogging density varies with the UGI factors but gradually approaches a stable level. This indicates that urban waterlogging density no longer decreases or increases with the UGI factor when the UGI factor reaches a certain range. For example, the UGI area proportion exceeds a certain range, the decreasing trend of urban waterlogging is gradually gentle (Figure 5-7a). Therefore, we can consider this value range reached by the UGI factor as the threshold value. Second, we notice that the relationship between waterlogging density and the UGI factor is similar to a logarithmic function. The logarithmic fitting is adapted to reflect the nonlinearity, which can be expressed as:

$$y = a \ln(x) + b \quad (5.4)$$

where y represents the waterlogging density, x is the UGI indicator, a and b are coefficients. Third, we calculate the derivative of the logarithmic fitting expressions to obtain the variation rate of waterlogging density (Figure 5-7b). According to the variation rate of waterlogging density, we find that at the beginning, the decreasing



rate of waterlogging density is very large, but with the increase of UGI indicators, the decline rate gradually tends to be flat. Therefore, we further calculated the unit decline rate of urban waterlogging density for each UGI indicator. When the unit decline rate is less than 0.01, we consider that the urban waterlogging density remains relatively stable, which no longer decreases significantly with the driver. On this basis, we regard the inflection point when the unit decline rate reaches 0.01. Accordingly, the value of the UGI indicators corresponding to the inflection point is considered the threshold value, which is defined as the limits of the impact of each UGI factor on urban waterlogging. For example, the area proportion of UGI corresponding to the inflection point is 24.4% (Figure 5-7b). This means that when the proportion of UGI exceeds 24.4%, the waterlogging density will not decrease significantly with the increase of the green area. Therefore, the threshold level of the UGI area proportion affecting urban waterlogging is 24.4%. Finally, the logarithmic fitting and derivation are implemented in the R package of “basicTrendline” [67].

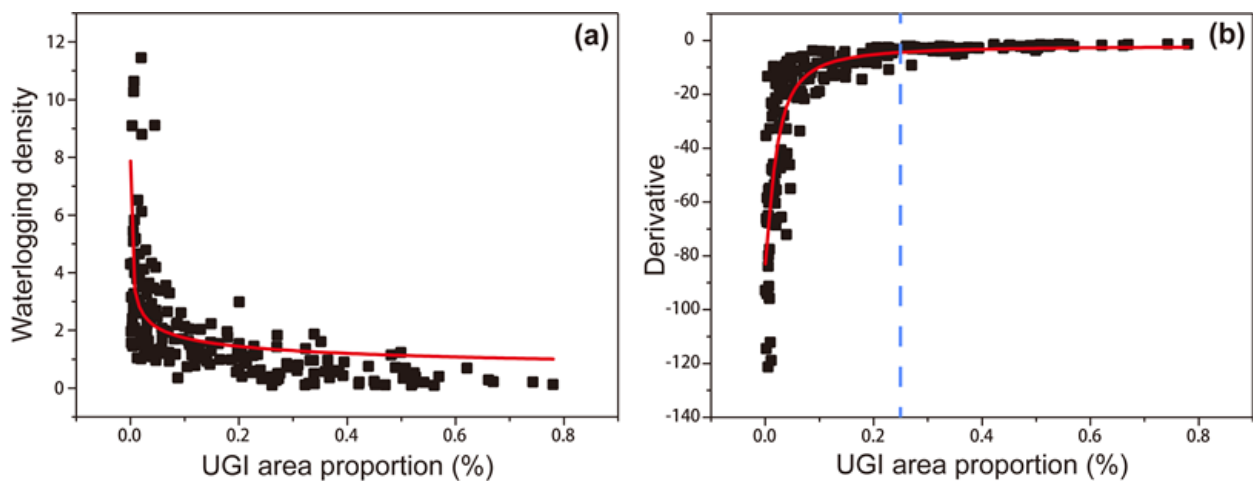


Figure 5-7 The logarithmic fitting (a) and the derivative (b) between UGI area proportion and waterlogging. The red line indicates the nonlinear fitting curve and the derivative curve, and the blue dashed line represents the threshold value.

## 5.4 Results

### 5.4.1 Spatial patterns of UGI and urban waterlogging

#### 5.4.1.1 Spatial pattern of UGI between two cities

As shown in Table 5-4, within Guangzhou central urban districts, the UGI accounts for 33.92% of the total area, 21.41%, 4.73%, 2.26%, and 5.52%, woodland, grassland, cultivated land, and garden, respectively. For Shenzhen city, the proportion of UGI is much greater than that of Guangzhou, accounting for 45.86% of the city. The woodland, grassland, cultivated land, and garden account for 27.78%, 7.82%, 2.75%, and 7.51%,

respectively. Regarding spatial configuration, the MPS and AI values are greater in Shenzhen city, indicating the clustered and continuous distribution of UGI. In contrast, Guangzhou's MPS and AI values are relatively small, while the LDI value is large, indicating a more scattered and fragmented distribution of UGI in Guangzhou.

*Table 5-4 Spatial pattern of UGI in Guangzhou and Shenzhen.*

Composition	City							
	Guangzhou				Shenzhen			
Woodland	21.41%				27.78%			
Grassland	4.73%				7.82%			
Cultivated land	2.26%				2.75%			
Garden	5.52%				7.51%			
Spatial configuration	Range	Mean	Median	S.D.	Range	Mean	Median	S.D.
	MPS	0.01 - 3.72	0.77	0.34	1.4	0.29 - 4.32	1.52	1.21
LDI	0.04 - 0.87	0.51	0.59	0.23	0.006 - 0.63	0.21	0.17	0.15
AI	67.94 - 99.53	91.47	88.52	2.81	65.17 - 97.98	95.77	94.33	4.26

The cluster effect of UGI between the two cities is shown in Figure 5-8. The hot spots of UGI (highlighted as red) in Guangzhou are mainly concentrated in the northwest part where the land cover features are dominated by woodland. In contrast, the cold spots of UGI were concentrated in the Liwan, Yuexiu, and Haizhu districts (southwest part of Guangzhou's central urban districts). These three districts are part of the historical urban area of Guangzhou and possess a high abundance of impervious surfaces. In Shenzhen, the UGI hot spots were mainly distributed in the Dapeng district, where the Dapeng Peninsula National Geopark is located. Compare with Guangzhou, Shenzhen's UGI cold spots presents a multicentric distribution pattern, mainly distributed in each urban district (Futian, Luohu, Baoan, Longgang District).

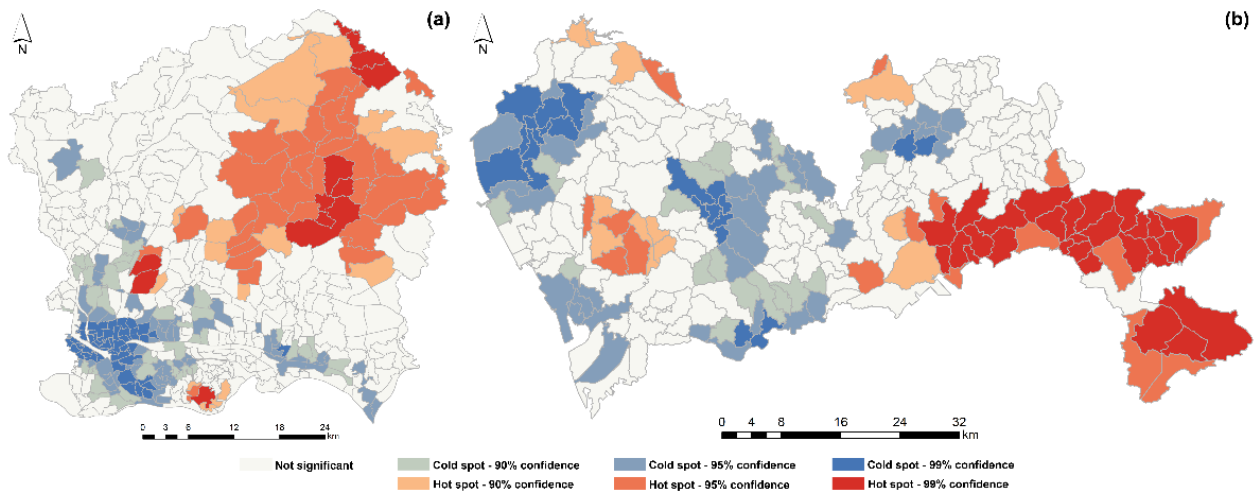


Figure 5-8 The UGI spatial agglomeration map for (a) Guangzhou and (b) Shenzhen.

#### 5.4.1.2 Urban waterlogging spatial agglomeration effect between two cities

The Getis-G statistic shows that the area covered by waterlogging hot spots in Guangzhou is approximately 17.35%, while the area covered by waterlogging cold spots is about 22.19% (Table 5-5). In Shenzhen, around 23.68% of Shenzhen was covered by urban waterlogging hot spots, and 29.52% by cold spots. Both indicate that urban waterlogging in Guangzhou and Shenzhen has a significant clustering effect.

Table 5-5 Descriptive statistics of urban waterlogging hot spots and cold spots.

City	Area percentage (%)		General G	z-score	p-value
	Hot spots	Cold spots			
Guangzhou	17.35	22.19	0.00008	18.33	0.00
Shenzhen	23.68	29.52	0.00002	12.94	0.00

In Guangzhou, the waterlogging hot spots (highlighted as red) are concentrated in the southwestern part, which has a relatively high proportion of impervious surfaces (Figure 5-9). In contrast, the waterlogging cold spots (highlighted as blue) are mainly distributed in the northwestern part with a relatively high UGI abundance. As for Shenzhen, the waterlogging hot spots are sparsely distributed in the urban sub-centers (Futian, Luohu, Longhua, Longgang District), while the urban waterlogging cold spots are mainly clustered in the eastern Dapeng district with better natural conditions. Although both cities show the agglomeration effect of urban waterlogging, the spatial clustering effect is more prominent in Guangzhou; instead, the hot spots and cold spots in Shenzhen present a more dispersed distribution pattern. This phenomenon may be due to the differences in the spatial distribution patterns of UGI between the two cities. Furthermore, a

comparison between Figures 5-8 and 5-9 reveals that the aggregation effects of UGI and urban waterlogging exhibit a coupling trend. For example, the hot spots of the UGI correspond to the cold spots of urban waterlogging and vice versa.

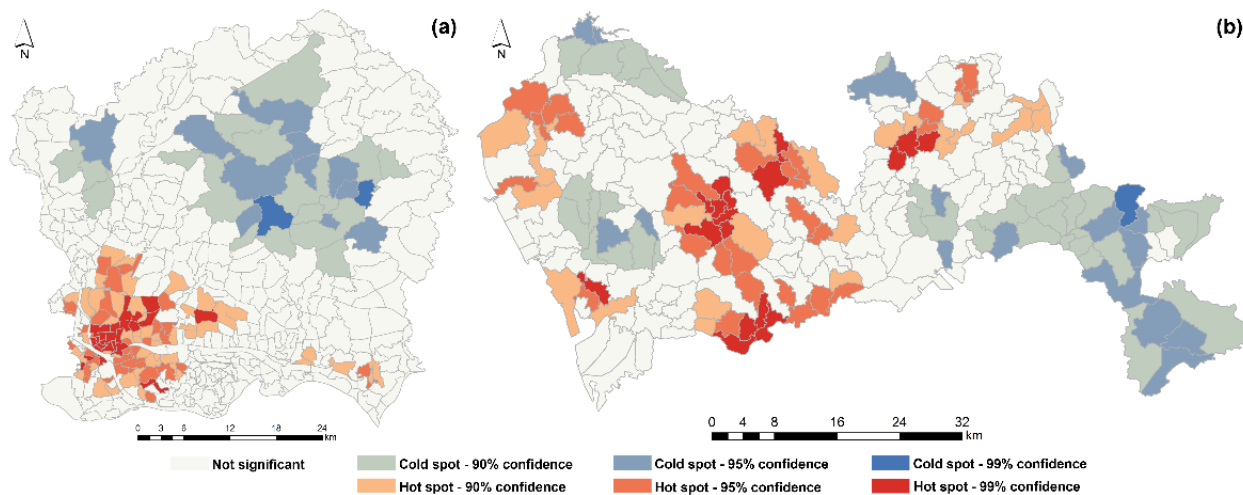


Figure 5-9 The waterlogging spatial agglomeration map for (a) Guangzhou and (b) Shenzhen.

#### 5.4.2 Impacts of UGI on urban waterlogging

##### 5.4.2.1 Partial correlations between UGI and waterlogging

As shown in Table 5-6, we found that the UGI area proportion and EVI both presented a significant negative correlation with waterlogging in two cities ( $p < 0.01$ ). This suggests that UGI such as woodland and grassland play a crucial role in regulating rainfall and reducing surface runoff or that the waterlogging density decreases with the increase of the UGI area proportion or EVI. From the perspective of landscape fragmentation, the MPS shows a negative correlation, while LDI experiences a positive correlation. This result indicates that a UGI with a large mean area or low fragmentation is less prone to waterlogging. As for landscape aggregation, the AI of green infrastructure has a negative effect on waterlogging, which implies that the clustered distribution of UGI is also conducive to the mitigation of urban waterlogging. The correlation results suggest that optimizing the spatial arrangement of green infrastructure also matters, which positively alleviates urban waterlogging.

Table 5-6 Partial correlation coefficients between UGI and waterlogging.

	City	Guangzhou	Shenzhen
Composition	EVI	-0.338**	-0.445**
	UGI	-0.471**	-0.657**

	Woodland	-0.428**	-0.556**
	Grassland	-0.354**	-0.498**
	Garden	-0.272**	-0.379**
	Cultivate land	-0.133	-0.237*
<hr/>			
Spatial configuration	MPS	-0.382**	-0.542**
	LDI	0.347**	0.422**
	AI	-0.278	-0.394**

#### 5.4.2.2 Individual and interactive effects of UGI factors on urban waterlogging

The factor detector examined the relative importance (individual effect) of UGI factors on urban waterlogging. As shown in Figure 5-10, the PD values for all influencing factors ranged from 0.05 to 0.42. First, the UGI compositions (area proportion and EVI) in both cities show the strongest impact on waterlogging, which both have an explanatory power of over 30%. The result indicates that the proportion of UGI and EVI has an almost equally important effect in alleviating urban waterlogging. However, the individual effect of UGI spatial configuration on waterlogging is relatively small. The PD values for MPS in Guangzhou and Shenzhen are 0.142 and 0.171, respectively, while the other spatial conformation indices (LDI and AI) are even smaller. It hints that urban waterlogging is mainly affected by UGI area proportion and EVI, rather than the spatial configuration. Under the background of rapid urbanization and continuous expansion of impervious surfaces, the importance of properly regulating the UGI area proportion and EVI to alleviate the risk of urban waterlogging is highlighted. Second, we notice that the PD value of UGI factors in Shenzhen is generally higher than in Guangzhou. This indicates that the single effect of UGI factors on waterlogging density in Shenzhen is greater than that in Guangzhou. Although there are slight differences in PD values between cities, all confirm that UGI area proportion has the greatest impact on urban waterlogging.

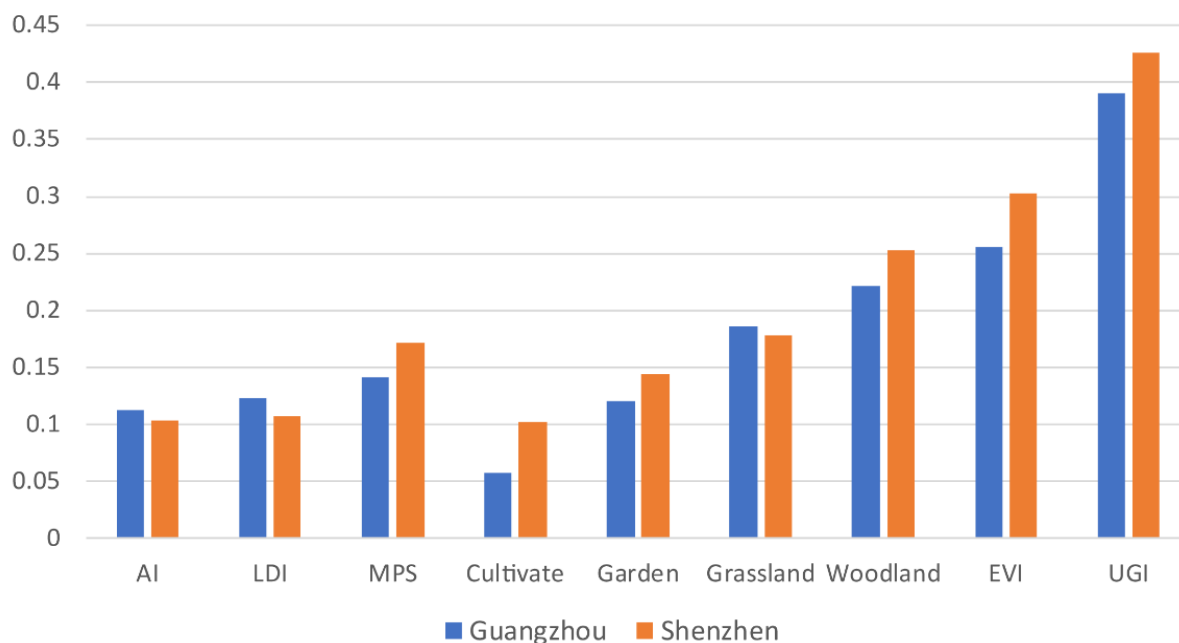


Figure 5-10 The PD values of UGI factors in different cities.

However, it is difficult to further reveal the mechanism of UGI on urban waterlogging through the individual effect. Therefore, the interaction detector was utilized to quantify the interactive effects of UGI factors on waterlogging. As shown in Table 5-7, the interaction detector calculated the interaction between five factors on urban waterlogging. The results indicate that the interaction of UGI factors greatly enhances their individual effects on waterlogging. In these 10 pairs of interactions, all the factors have strong binary enhancement, and some even show a non-linear enhancement. The largest interaction in Guangzhou and Shenzhen is the UGI area proportion interacting with EVI, followed by EVI interacting with MPS. This illustrates the importance of the interaction between the UGI area and its biophysical parameter (reflect vegetation healthy and density). Neither the size of the UGI nor its biophysical parameter can be ignored. Appropriate UGI area combined with good vegetation cover (EVI) will further improve the mitigation of urban waterlogging. Regarding the spatial configuration of Guangzhou, although AI has a relatively lower PD value than UGI area proportion from the single factor detector results, the interactive effect of UGI area proportion and AI factor get 261.61% enhancement compared with the single effect. Additionally, the results of the single factor detector results in Shenzhen show that the impact of LDI is not very important for urban waterlogging. However, the interactive effect of UGI area proportion and LDI accounts for around 274.58% enhancement. Furthermore, the interactive effect of EVI and MPS obtains over 233% and 195% enhancement for Guangzhou and Shenzhen, respectively. These results underscore the importance of the combination of UGI composition

(area proportion and EVI) and spatial configuration. Based on a certain percentage of UGI, the interaction of UGI composition and configuration can further enhance its impact on urban waterlogging, which has important implications for the metropolis with a shortage of urban land resources. Lastly, similar to individual effects, most PD values of the Shenzhen interactive effect are higher than those of Guangzhou. This means that there is some variation in the ability of UGI to influence urban waterlogging under different urban backgrounds. Considering the vegetation conditions in Shenzhen (high cover), it can be inferred that UGI can affect urban waterlogging to a greater extent in cities with better vegetation conditions.

*Table 5-7 Partial correlation coefficients between UGI and waterlogging.*

Factor	Guangzhou city		Factor	Shenzhen city	
	Interactive PD	Enhancement		Interactive PD	Enhancement
PD (UGI $\cap$ EVI)	0.525	Binary	PD (UGI $\cap$ EVI)	0.557	Binary
PD (UGI $\cap$ MPS)	0.446	Binary	PD (UGI $\cap$ MPS)	0.483	Binary
PD (UGI $\cap$ LDI)	0.424	Binary	PD (UGI $\cap$ LDI)	0.442	Binary
PD (UGI $\cap$ AI)	0.405	Binary	PD (UGI $\cap$ AI)	0.439	Binary
PD (EVI $\cap$ MPS)	0.474	Nonlinear	PD (EVI $\cap$ MPS)	0.505	Nonlinear
PD (EVI $\cap$ LDI)	0.362	Binary	PD (EVI $\cap$ LDI)	0.407	Binary
PD (EVI $\cap$ AI)	0.311	Binary	PD (EVI $\cap$ AI)	0.374	Binary
PD (MPS $\cap$ LDI)	0.252	Binary	PD (MPS $\cap$ LDI)	0.308	Nonlinear
PD (MPS $\cap$ AI)	0.267	Nonlinear	PD (MPS $\cap$ AI)	0.289	Nonlinear
PD (LDI $\cap$ AI)	0.238	Nonlinear	PD (LDI $\cap$ AI)	0.231	Nonlinear

#### 5.4.3 Threshold level of UGI affecting waterlogging

Next, we aimed to better demonstrate the impact of UGI on the urban waterlogging magnitude. According to the relative importance of UGI factors on urban waterlogging, the relationship between UGI factors (UGI area proportion, EVI, MPS) and waterlogging density was plotted (Figure 5-11). As for UGI area proportion (Figure 5-11a,g), the nonlinear fitting curve (red line) indicates that as the proportion of green infrastructure increases, the decreasing trend of urban waterlogging gradually becomes more gentle. In Guangzhou, the downward trend of urban waterlogging density is significant when the area proportion is below 15%. However, when the area proportion exceeds 20%, the decreasing rate gradually slows down, and the waterlogging

density remains relatively stable. Similarly, we also found that when the proportion of UGI in Shenzhen exceeds 60%, the decreasing trend of urban waterlogging density is also not obvious. This means that if the area of green infrastructure in the watershed exceeds the threshold, continuing to increase the proportion of green infrastructure may not significantly improve its mitigation effect. By deriving the function (Figure 5-11d,j), we find that when the UGI proportion exceeds 24.4% and 72.1%, the decline rate of urban waterlogging gradually approaches zero. This suggests that urban waterlogging barely declines as the proportion of green space increases. Correspondingly, we can consider that 24.4% and 72.1% of UGI area proportion are the threshold values for Guangzhou and Shenzhen. It hints that the area proportion of UGI within a watershed unit needs to be maintained at a certain level to effectively exert its waterlogging mitigation effect. If the green infrastructure proportion exceeds the threshold, the mitigation effect is no longer enhanced, which indicates the saturation effect of urban waterlogging mitigation. Therefore, the area proportion of UGI within watersheds should be weighed comprehensively regarding the benefits of the urban waterlogging mitigation effect.

As for the biophysical parameter, the EVI also presents a strong logarithmic correlation with waterlogging both in Guangzhou and Shenzhen (Figure 5-11b,h). The EVI thresholds in Guangzhou and Shenzhen are 0.36 and 0.43 when calculating the derivatives. When the EVI is less than the threshold value, the waterlogging density and the derivative values in Guangzhou and Shenzhen decrease significantly. At the same time, the EVI exceeds the threshold, the decline rate decreases gradually and insignificantly. This result also suggests that vegetation also has a saturation effect in the interception and infiltration of surface runoff.

Regarding the effect of spatial configuration, it can be found that a similar phenomenon also exists in the relationship between spatial configuration (MPS) and waterlogging (Figure 5-11c,i). In detail, the threshold value of MPS is 1.9 ha (Guangzhou) and 2.8 ha (Shenzhen). The threshold of MPS indicates that each green patch needs to maintain a certain area to achieve the optimal mitigation effect. This means that in addition to controlling the area proportion of UGI within a watershed unit, the size of green infrastructure patches also needs to be weighed.



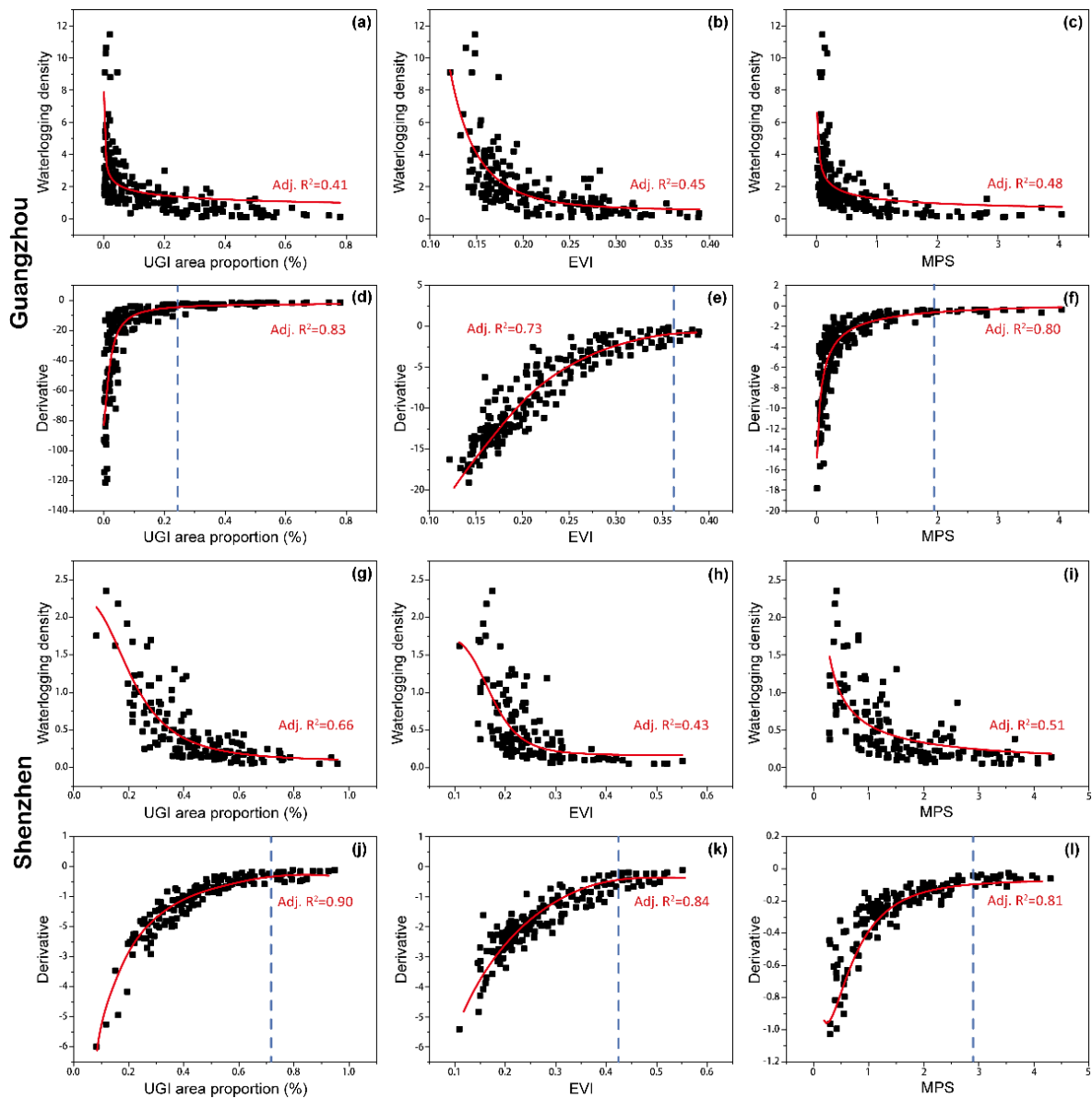


Figure 5-11 The relationship between UGI variables and waterlogging for (a–f) Guangzhou and (g–l) Shenzhen. The red line indicates the nonlinear fitting curve and the derivative curve, and the blue dashed line represents the threshold value.

## 5.5 Discussion

### 5.5.1 Spatial variations of urban waterlogging

As shown in Figure 5-9, the urban waterlogging events in Guangzhou and Shenzhen have an obvious cluster effect. The waterlogging hotspots are mainly located in areas with a high proportion of impervious surfaces and low vegetation abundance, which couples with the spatial distribution of UGI hotspots (Figure 5-8). It hints that urban waterlogging hot spots tend to correspond to cold spots for green infrastructure. Since

implementing the "open-door" policy and economic reforms in 1978, many cities in China have experienced rapid urbanization. The selected study areas, Guangzhou and Shenzhen, are representative of rapid urbanization in China. During this process, the underlying surface inside cities has changed dramatically, particularly the southwestern part of Guangzhou and the central and western parts of Shenzhen (with a high proportion of impervious surfaces). Accordingly, increase the risk of urban waterlogging in these regions. However, it is worth noting that the northeastern part of Guangzhou and the eastern part of Shenzhen are less affected by urban expansion, mainly because the urban expansion is strictly restricted in these areas as there are national forest parks. Therefore, the vegetation abundance is relatively high in these areas, which has become the hotspot for green infrastructure distribution. Despite the relatively high average annual rainfall and the low density of the drainage network in these areas (Figure 5-6), the occurrence of urban waterlogging is much lower than that of the city center, becoming the waterlogging cold spots. This phenomenon indirectly confirms the positive effect of green infrastructure on urban waterlogging. Moreover, some studies have pointed out that the influence of drainage facilities on urban waterlogging is not as great as commonly believed [10,23,25,68]. The more important impact is the land cover composition, as our study further confirms. This further explains why urban waterlogging events are less frequent in the suburbs (with ample green areas).

Moreover, the spatial agglomeration effect of urban waterlogging and green infrastructure also provides corresponding enlightenment for waterlogging prevention and reduction. Local authorities can develop local mitigation strategies for waterlogging hotspot regions, such as increasing the green area or optimizing the spatial configuration of UGI. Simultaneously, urban development in these areas needs to pay more attention to the proportion of impervious surface area, which avoids further encroachment of impervious surfaces on scarce green areas. In the event of heavy rainfall, these measures can help reduce or even prevent urban waterlogging. Furthermore, with the advent of multisource big data (population mobility data, traffic travel data, population density), the emergency management department can provide early warning to these hotspot regions to minimize the negative impact of waterlogging, such as the evacuation of the elderly or closing the underground parking lot. This will help accurately assess the risk of urban waterlogging and improve early warning and emergency response.

### *5.5.2 The mitigation effect of UGI on waterlogging*

UGI is often recommended for regulating surface runoff, purifying rainwater, and reducing negative

environmental impacts [68,69,70,71]. Our results demonstrated that UGI has a considerable effect on urban waterlogging, even after controlling the impact of urban topography, precipitation, and drainage facilities. This result is consistent with previous studies, which confirmed the role of UGI in mitigating urban waterlogging risk [43,45,47]. However, our results further expand our understanding of the mechanism of green infrastructure alleviating urban waterlogging.

Firstly, we found that choosing the most representative and important UGI metric is crucial in this study. Focusing on the different behavior of various UGI metrics to influence the urban waterlogging magnitude, the results of cross-site evaluation suggest that the area proportion of UGI to be the most dominant factor influencing urban waterlogging. The larger the area of a UGI, the more considerable effect it has in regulating urban waterlogging magnitude. This finding is also consistent with Yao et al. [42] and Yang et al. [43]. However, this study further reveals the relative contribution of different UGI compositions. For the area proportion of UGI, woodland and grassland have the greatest impact on urban waterlogging, which provides implications to urban planners on the importance of preserving woodlands and grasslands in cities.

Additionally, it is interesting to note that the impact of green infrastructure on urban waterlogging also depends on its vegetation status (biophysical parameter). Our result demonstrated that the influence of biophysical parameters (EVI) could not be ignored or simply equated with area proportion. However, most previous studies ignore the influence of biophysical parameters (EVI), which only analyze the effects of different sizes of green infrastructure on urban waterlogging [41,42,43,44,45]. For example, Armson's study confirmed that when the green patch area reached 9 m<sup>2</sup>, it could effectively reduce the surface runoff [48]. Coincidentally, Kim et al. [41] pointed out that the larger the green space area, the greater its effect on average runoff. However, ignoring the biophysical parameters of UGI will inevitably lead to some deviation. This is mainly because densely vegetated plots and sparsely vegetated plots have completely different effects on urban waterlogging for the same area. Our discoveries further deepen our understanding of the mechanism of UGI alleviating urban waterlogging, which helps implement more effective UGI planning strategies to mitigate urban waterlogging. For the same UGI area, healthier or denser vegetation (superior ecological environment with high EVI value) can more effectively intercept and store rainwater runoff, thereby contributing to the mitigation of urban waterlogging. Therefore, while increasing the area of UGI, it is also necessary to improve the vegetation conditions (biophysical parameter) of UGI.

Traditionally, urban planners increased the area of UGI to create a more pleasant human settlement

[32,45,49]. However, with the rapid urbanization, the land-use pressure within cities has increased significantly, showing that impervious surfaces continue to encroach on green infrastructure. Therefore, it is particularly important to optimize the spatial configuration of UGI under limited land resource conditions. Previous studies have also noted the impact of the spatial configuration of UGI on urban waterlogging [41,44,48]. Our study also confirmed the effect of spatial configuration on waterlogging. The mitigation effect of UGI on waterlogging can be increased or decreased through different spatial arrangements while keeping the green infrastructure area constant, which has important implications for the metropolis that lack land resources for UGI construction. For example, for most city centers, after experiencing rapid urbanization, there are not enough land resources for UGI construction. Therefore, it is necessary to carry out a spatial reorganization of existing green infrastructure to make it clustered distributed with less fragmentation.

Considerable studies have mainly focused on the individual effect of UGI factors on urban waterlogging [41,44], while neglecting the interactive effect of UGI factors on urban waterlogging. However, our results show that the interaction of UGI factors greatly enhances its impact on urban waterlogging. This will undoubtedly further enhance our scientific knowledge in mitigating waterlogging. For example, the UGI area combined with EVI or spatial configuration will further improve the mitigation of urban waterlogging. These results underscore the importance of the combination of UGI composition and spatial configuration, as the individual effect is not sufficient. This suggests that our proposed method can reveal in more detail how the interaction of UGI composition and spatial configuration affects urban waterlogging. Additionally, this finding refreshes our perception of the importance of the interaction of landscape patterns. In previous studies, the importance of the UGI spatial configuration is often overlooked [41,44,48], mainly due to the spatial configuration having a relatively small individual effect. The interaction effect has led to a renewed awareness of the importance of the UGI landscape patterns for urban waterlogging mitigation. The interaction between spatial configuration and composition can more significantly improve the mitigation capacity of UGI. Therefore, we cannot ignore the role of the spatial configuration due to its relatively small individual effect. This also provides valuable and practical references for the urban planner to optimize the spatial configuration of UGI in urban centers.

### *5.5.3 Threshold level of waterlogging mitigation effect*

It is well known that increasing the area proportion of UGI helps alleviate urban waterlogging. Most previous studies have only confirmed that increasing the area of green infrastructure can more effectively alleviate

urban waterlogging [41,42,43,44,45,50]. However, this result cannot be practically applied to guide UGI planning, as the land use pressure within the urban centers is too large to increase the green area greatly. Compared with previous studies, our results find that there is a threshold level for the waterlogging mitigation effect (Figure 5-11). The impact of green infrastructure on urban waterlogging is not a simple linear relationship. As the area proportion of UGI within the watershed exceeds the threshold, the waterlogging density will not continue to decline as the UGI area increases. This provides a new perspective on urban waterlogging mitigation strategies—blindly increasing the area of green infrastructure may not bring much improvement to urban waterlogging mitigation. The excessive proportions of UGI within the watershed unit may lead to a waste of its mitigation effect. Therefore, the area proportion of UGI and its mitigation effect should be considered comprehensively when planning UGI. Moreover, we can infer that under the same area of UGI, replacing a single large-area UGI (exceeding the threshold) with several small green infrastructure patches may provide a more significant mitigation effect for urban waterlogging. At present, some local governments have built a single enormous green area in urban new districts. However, these actions may lead to a great waste of their mitigation effects. It is recommended to control the area proportion of UGI within the threshold value to mitigate urban waterlogging more effectively.

Furthermore, it is necessary to point out that the threshold values of MPS indicate that green infrastructure patches need to be maintained in a certain area. Green infrastructure patches that are too small or too large may not be effective in alleviating urban waterlogging. This means that when replacing a single large-area UGI, it is not advisable to use too small and fragmented patches. At the same time, the biophysical parameters also need to be weighed. In general, this provides practical implementations for urban green infrastructure planning: the proportion of green infrastructure at the watershed scale and the green infrastructure area at the patch scale are recommended not to exceed the threshold.

These results provide considerable implications for UGI management and planning. At present, if UGI in an urban center is occupied by artificial land cover, the loss of UGI area is generally filled by the land resources at the urban fringe (with low land-use pressure). However, the UGI area remains relatively balanced at the city level. It is a net loss of the UGI area in the urban center, which undoubtedly negatively affects the management of urban waterlogging. Additionally, the vegetation abundance, biophysical parameter, and spatial configuration of UGI are not consistent during spatial replacement, which further leads to the relative loss of the urban waterlogging mitigation effect. Even if the area of UGI in the entire city remains relatively

balanced, this may further worsen the urban waterlogging status. Therefore, despite the great pressure on land use and development in the urban center, it is still necessary to retain an appropriate area of UGI.

#### *5.5.4 Limitations and uncertainties*

This study provides a potentially valuable idea for investigating the interaction effect and threshold level of UGI on urban waterlogging. However, this study has its limitations and they should be considered in future work. Firstly, the complex relationship between UGI and waterlogging was analyzed in two cities using only the historical record of urban waterlogging. Although we explicitly demonstrate the effect of UGI composition and spatial configuration, we may not apply this across all regions. The mitigation effect of various UGI factors may largely depend on the urban background. Revealing the role of UGI in urban waterlogging in other regions may help us confirm the universality of our findings. Secondly, the urban waterlogging data did not record the size (water depth, area), duration, and the specific year of each event. We only analyzed the waterlogging mitigation effect from the whole period, which inevitably brings some uncertainty to the results. Thirdly, only several commonly used UGI metrics were used in this study. Other three-dimensional metrics, such as green volume, were not taken into account. Consequently, in future research, the mitigation effect of UGI can reveal further insights from the perspective of the mitigation intensity and mitigation scale. It is suggested to explore the differences in the mitigation intensity of different compositions and spatial patterns of green infrastructure on urban waterlogging and the scale of this mitigation effect. Moreover, with sufficient data, we can introduce three-dimensional indicators of green infrastructure, as well as the water depth and duration of waterlogging events to investigate the effect of UGI on urban waterlogging more comprehensively.

## **5.6 Conclusion**

In the context of the UN's 2030 Sustainable Development Goals, two highly urbanized coastal cities were selected for a cross-regional study to investigate the interaction effect and threshold level of UGI on urban waterlogging. The results support three conclusions.

Firstly, the area proportion and EVI of UGI both have a non-negligible effect in alleviating urban waterlogging. The impact of green infrastructure on urban waterlogging largely depends on its area and vegetation status. Healthier or denser vegetation (superior ecological environment) can more effectively intercept and store rainwater runoff. This finding provides practical insights into UGI planning, i.e., while increasing the area of UGI, more attention should also be paid to the biophysical parameter of vegetation, thereby improving the

mitigation effect of green infrastructure from the “size” and “health”. Secondly, the interaction of UGI factors greatly enhances their individual effects on waterlogging. The UGI composition (area percentage and biophysical parameter) and the spatial configuration can effectively alleviate urban waterlogging. This result offers insights into the importance of the interactive enhancement effect between UGI composition and spatial configuration. Under limited area for green infrastructure, it is more necessary to optimize the UGI composition and spatial configuration. Lastly, the impact of UGI on waterlogging presents a threshold phenomenon. Blindly increasing the area of green infrastructure may not greatly improve the alleviation of urban waterlogging. Excessive proportions of UGI within the watershed unit or an oversized UGI patch may lead to a waste of mitigation effects. Therefore, it is necessary to control the UGI area (both in the watershed unit and patch size) within a certain range to play a corresponding role in mitigating urban waterlogging. Since the threshold values of some UGI indicators are different among cities, the thresholds are disturbed by regional characteristics. Therefore, the UGI-based waterlogging prevention strategies should be adapted to local conditions. Given the growing concerns of global warming and continued rapid urbanization, we believe that our findings provide useful enlightenment for local authorities in urban waterlogging prevention, green infrastructure management, and sustainable development.

### **Author Contributions**

Conceptualization, Q.Z. and P.T.; methodology, Q.Z. and P.T.; software, Q.Z.; validation, P.T., Z.W. and Q.Z.; formal analysis, Q.Z.; investigation, Q.Z.; resources, Z.W. and P.T.; data curation, Z.W. and Q.Z.; writing—original draft preparation, Q.Z.; writing—review and editing, P.T. and Z.W.; visualization, Q.Z.; supervision, P.T. and Z.W.; project administration, P.T. and Z.W.; funding acquisition, P.T. and Z.W. All authors have read and agreed to the published version of the manuscript.

### **Funding**

This research was funded by the University of Padova research project, grant number DOR1948955/19; the National Key R & D Program of China, grant number 2018YFB2100702; the NSFC-Guangdong Joint Foundation Key Project, grant number U1901219; and the Team Project of Guangdong Provincial Natural Science Foundation,

grant number 2018B030312004.

### **Data Availability Statement**

The remote sensing data and code used in this study are openly available on the website. The high-resolution topographic data are not publicly available due to the confidentiality rules.

### **Acknowledgments**

The authors are grateful for the waterlogging records provided by the Guangzhou and Shenzhen Water Resources Bureau. The authors also would like to thank the Land Resources Technology Center of Guangdong Province and Shenzhen Planning and Natural Resources Bureau for providing basic urban data. We thank the reviewers and editors for their comments.

### **Conflicts of Interest**

The authors declare no conflict of interest.



## References

1. Lin, J.; He, X.; Lu, S.; Liu, D.; He, P. Investigating the influence of three-dimensional building configuration on urban pluvial flooding using random forest algorithm. *Environ. Res.* 2020, 196, 110438.
2. Zhang, Q.; Wu, Z.; Guo, G.; Zhang, H.; Tarolli, P. Explicit the urban waterlogging spatial variation and its driving factors: The stepwise cluster analysis model and hierarchical partitioning analysis approach. *Sci. Total Environ.* 2020, 763, 143041.
3. Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Xu, L. Assessment of urban flood susceptibility using semi-supervised machine learning model. *Sci. Total Environ.* 2019, 659, 940–949.
4. United Nations. World Urbanization Prospects: The 2018 Revision. Available online: <https://www.un.org/development/desa/pd/content/world-urbanization-prospects-2018-revision> (accessed on 15 January 2021).
5. Qiu, J.; Cao, B.; Park, E.; Yang, X.; Zhang, W.; Tarolli, P. Flood Monitoring in Rural Areas of the Pearl River Basin (China) Using Sentinel-1 SAR. *Remote Sens.* 2021, 13, 1384.
6. Liqueste, C.; Kleeschulte, S.; Dige, G.; Maes, J.; Grizzetti, B.; Olah, B.; Zulian, G. Mapping green infrastructure based on eco-system services and ecological networks: A Pan-European case study. *Environ. Sci. Policy* 2015, 54, 268–280.
7. Le, M.T.; Cao, T.A.T.; Tran, N.A.Q. The role of green space in the urbanization of Hanoi city. In *E3S Web of Conferences*; EDP Sciences: Ulis, France, 2019; Volume 97.
8. Zhong, Y.D.; Jia, Y.W.; Li, Z.W. Spatial and temporal changes of maximum 1 h precipitation intensity in Beijing region in last 53 years. *J. China Hydrol.* 2013, 33, 32–37.
9. Lyu, H.M.; Shen, S.L.; Zhou, A.N.; Zhou, W.H. Flood risk assessment of metro systems in a subsiding environment using the interval FAHP-FCA approach. *Sustain. Cities Soc.* 2019, 50, 101682.
10. Zhang, Q.; Wu, Z.; Zhang, H.; Dalla Fontana, G.; Tarolli, P. Identifying dominant factors of waterlogging events in metro-politan coastal cities: The case study of Guangzhou, China. *J. Environ. Manag.* 2020, 271, 110951.
11. Sofia, G.; Roder, G.; Dalla Fontana, G.; Tarolli, P. Flood dynamics in urbanised landscapes: 100 years of climate and humans' interaction. *Sci. Rep.* 2017, 7, 1–12.
12. Huang, H.; Chen, X.; Zhu, Z.; Xie, Y.; Liu, L.; Wang, X.; Wang, X.; Liu, K. The changing pattern of

urban flooding in Guang-zhou, China. *Sci. Total Environ.* 2018, 622, 394–401.

13. Yang, H.; Dick, W.A.; McCoy, E.L.; Phelan, P.L.; Grewal, P.S. Field evaluation of a new biphasic rain garden for stormwater flow management and pollutant removal. *Ecol. Eng.* 2013, 54, 22–31.
14. Amado, L.; Albuquerque, A.; Santo, A.E. Influence of stormwater infiltration on the treatment capacity of a LECA-based horizontal subsurface flow constructed wetland. *Ecol. Eng.* 2012, 39, 16–23.
15. IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2014.
16. Ministry of Water Resources of the People's Republic of China. Available online: <http://www.mwr.gov.cn> (accessed on 22 January 2021).
17. Tehrany, M.S.; Jones, S.; Shabani, F. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena* 2019, 175, 174–192.
18. Wang, C.; Du, S.; Wen, J.; Zhang, M.; Gu, H.; Shi, Y.; Xu, H. Analyzing explanatory factors of urban pluvial floods in Shang-hai using geographically weighted regression. *Stochastic Environ. Res. Risk Assess.* 2017, 31, 1777–1790.
19. Li, B.; Zhao, Y.; Fu, Y. Spatio-temporal characteristics of urban stormwaterlogging in Guangzhou and the impact of urban growth. *Geo Inf. Sci.* 2015, 17, 445–450.
20. Zhang, H.; Wu, C.; Chen, W.; Huang, G. Assessing the impact of climate change on the waterlogging risk in coastal cities: A case study of Guangzhou, South China. *J. Hydrometeorol.* 2017, 18, 1549–1562.
21. Wu, J.; Sha, W.; Zhang, P.; Wang, Z. The spatial non-stationary effect of urban landscape pattern on urban waterlogging: A case study of Shenzhen City. *Sci. Rep.* 2020, 10, 1–15.
22. Huong, H.T.L.; Pathirana, A. Urbanization and climate change impacts on future urban flooding in Can Tho City, Vietnam. *Hydrol. Earth Syst. Sci.* 2013, 17, 379–394.
23. Wang, J.; Gao, W.; Xu, S.; Yu, L. Evaluation of the combined risk of sea level rise, land subsidence, and storm surges on the coastal areas of Shanghai, China. *Clim. Chang.* 2012, 115, 537–558.
24. Zhao, G.; Pang, B.; Xu, Z.; Yue, J.; Tu, T. Mapping flood susceptibility in mountainous areas on a national scale in China. *Sci. Total Environ.* 2018, 615, 1133–1142.
25. Wu, J.S.; Zhang, P.H. The effect of urban landscape pattern on urban waterlogging. *Acta Geogr. Sin.* 2017, 3, 26–36.

26. Liu, F.; Liu, X.; Xu, T.; Yang, G.; Zhao, Y. Driving Factors and Risk Assessment of Rainstorm Waterlogging in Urban Agglomeration Areas: A Case Study of the Guangdong-Hong Kong-Macao Greater Bay Area, China. *Water* 2021, 13, 770.
27. Zhang, B.; Li, N.; Wang, S. Effect of urban green space changes on the role of rainwater runoff reduction in Beijing, China. *Landsc. Urban Plan.* 2015, 140, 8–16.
28. Zhang, B.; Xie, G.; Zhang, C.; Zhang, J. The economic benefits of rainwater-runoff reduction by urban green spaces: A case study in Beijing, China. *J. Environ. Manag.* 2012, 100, 65–71.
29. Dong, W.; Lian, Y.; Zhang, Y. Sustainable Development of Water Resources and Hydraulic Engineering in China Proceedings for the 2016 International Conference on Water Resource and Hydraulic Engineering; Springer: Cham, Switzerland, 2018.
30. Pour, S.H.; Abd Wahab, A.K.; Shahid, S.; Dewan, A. Low impact development techniques to mitigate the impacts of climate-change-induced urban floods: Current trends, issues and challenges. *Sustain. Cities Soc.* 2020, 62, 102373.
31. Palliwoda, J.; Banzhaf, E.; Priess, J.A. How do the green components of urban green infrastructure influence the use of eco-system services? Examples from Leipzig, Germany. *Landsc. Ecol.* 2020, 35, 1127–1142.
32. Yu, Z.; Fryd, O.; Sun, R.; Jørgensen, G.; Yang, G.; Özdil, N.C.; Vejre, H. Where and how to cool? An idealized urban thermal security pattern model. *Landsc. Ecol.* 2020, 1–10.
33. Amundsen, O.M.; Allen, W.; Hoellen, K. Green infrastructure planning: Recent advances and applications. In *Planners Advisory Service Memo*; American Planning Association: Chicago, IL, USA, 2009.
34. Schuch, G.; Serrao-Neumann, S.; Morgan, E.; Choy, D.L. Water in the city: Green open spaces, land use planning and flood management—An Australian case study. *Land Use Policy* 2017, 63, 539–550.
35. Lovell, S.T.; Taylor, J.R. Supplying urban ecosystem services through multifunctional green infrastructure in the United States. *Landsc. Ecol.* 2013, 28, 1447–1463.
36. Du, S.; Wang, C.; Shen, J.; Wen, J.; Gao, J.; Wu, J.; Lin, W.; Xu, H. Mapping the capacity of concave green land in mitigating urban pluvial floods and its beneficiaries. *Sustain. Cities Soc.* 2019, 44, 774–782.
37. Guo, G.; Wu, Z.; Chen, Y. Complex mechanisms linking land surface temperature to greenspace spatial patterns: Evidence from four southeastern Chinese cities. *Sci. Total Environ.* 2019, 674, 77–

87.

38. Guo, G.; Wu, Z.; Xiao, R.; Chen, Y.; Liu, X.; Zhang, X. Impacts of urban biophysical composition on land surface temperature in urban heat island clusters. *Landsc. Urban Plan.* 2015, 135, 1–10.
39. Pickett, S.T.; Cadenasso, M.L.; Grove, J.M.; Boone, C.G.; Groffman, P.M.; Irwin, E.; Warren, P. Urban ecological systems: Scientific foundations and a decade of progress. *J. Environ. Manag.* 2011, 92, 331–362.
40. Luan, X.; Yu, Z.; Zhang, Y.; Wei, S.; Miao, X.; Huang, Z.Y.X.; Teng, S.N.; Xu, C. Remote sensing and social sensing data reveal scale-dependent and system-specific strengths of urban heat island determinants. *Remote Sens.* 2020, 12, 391.
41. Kim, H.W.; Park, Y. Urban green infrastructure and local flooding: The impact of landscape patterns on peak runoff in four Texas MSAs. *Appl. Geogr.* 2016, 77, 72–81.
42. Yao, L.; Chen, L.; Wei, W.; Sun, R. Potential reduction in urban runoff by green spaces in Beijing: A scenario analysis. *Urban For. Urban Green.* 2015, 14, 300–308.
43. Yang, L.; Zhang, L.; Li, Y.; Wu, S. Water-related ecosystem services provided by urban green space: A case study in Yixing City (China). *Landsc. Urban Plan.* 2015, 136, 40–51.
44. Yang, X.; You, X.Y.; Ji, M.; Nima, C. Influence factors and prediction of stormwater runoff of urban green space in Tianjin, China: Laboratory experiment and quantitative theory model. *Water Sci. Technol.* 2013, 67, 869–876.
45. Liu, W.; Chen, W.; Peng, C. Assessing the effectiveness of green infrastructures on urban flooding reduction: A community scale study. *Ecol. Model.* 2014, 291, 6–14.
46. Roy, S.; Byrne, J.; Pickering, C. A systematic quantitative review of urban tree benefits, costs, and assessment methods across cities in different climatic zones. *Urban For. Urban Green.* 2012, 11, 351–363.
47. Silva, M.M.; Costa, J.P. Urban floods and climate change adaptation: The potential of public space design when accommodating natural processes. *Water* 2018, 10, 180.
48. Armson, D.; Stringer, P.; Ennos, A.R. The effect of street trees and amenity grass on urban surface water runoff in Manchester, UK. *Urban For. Urban Green.* 2013, 12, 282–286.
49. Richards, P.J.; Farrell, C.; Tom, M.; Williams, N.S.G.; Fletcher, T.D. Vegetable rain gardens can produce food and reduce stormwater runoff. *Urban For. Urban Green.* 2015, 14, 646–654.
50. Wen, L.; Weiping, C.; Chi, P. Modeling the effects of green infrastructure on storm water runoff

reduction on community scale. *Acta Ecol. Sin.* 2016, 36, 1686–1697.

51. Du, S.; He, C.; Huang, Q.; Shi, P. How did the urban land in floodplains distribute and expand in China from 1992–2015? *Environ. Res. Lett.* 2018, 13, 034018.
52. Pijl, A.; Brauer, C.C.; Sofia, G.; Teuling, A.J.; Tarolli, P. Hydrologic impacts of changing land use and climate in the Veneto lowlands of Italy. *Anthropocene* 2018, 22, 20–30.
53. Guangzhou Meteorological Service. Available online: <http://www.tqyb.com.cn> (accessed on 22 February 2021).
54. Shenzhen Meteorological Bureau. Available online: <http://weather.sz.gov.cn> (accessed on 22 February 2021).
55. Yu, H.; Zhao, Y.; Fu, Y.; Li, L. Spatiotemporal variance assessment of urban rainstorm waterlogging affected by impervious surface expansion: A case study of Guangzhou, China. *Sustainability* 2018, 10, 3761.
56. Xue, F.; Huang, M.; Wang, W.; Zou, L. Numerical Simulation of Urban Waterlogging Based on FloodArea Model. *Adv. Meteorol.* 2016, 2016, 1–9.
57. Zhang, S.L.; Gan, J.Y.; Zeng, Q.L.; Lu, G.N. Automatic compartmentalization of urban rainwater catchments on water outlet supported by GIS technology. *J. Hydraul. Eng.* 2007, 38, 325–329.
58. Zuo, D.; Xu, Z.; Yao, W.; Jin, S.; Xiao, P.; Ran, D. Assessing the effects of changes in land use and climate on runoff and sediment yields from a watershed in the Loess Plateau of China. *Sci. Total Environ.* 2016, 544, 238–250.
59. McGarigal, K. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure; US Department of Agriculture, Forest Service, Pacific Northwest Research Station: Corvallis, OR, USA, 1995; Volume 351.
60. Turner, M.G.; Gardner, R.H.; O'Neill, R.V.; O'Neill, R.V. *Landscape Ecology in Theory and Practice*; Springer: New York, NY, USA, 2001; Volume 401.
61. Cushman, S.A.; McGarigal, K.; Neel, M.C. Parsimony in landscape metrics: Strength, universality, and consistency. *Ecol. Indic.* 2008, 8, 691–703.
62. Yang, Z.; Sliuzas, R.; Cai, J.; Ottens, H.F. Exploring spatial evolution of economic clusters: A case study of Beijing. *Int. J. Appl. Earth Obs. Geoinf.* 2012, 19, 252–265.
63. Wang, J.F.; Li, X.H.; Christakos, G.; Liao, Y.L.; Zhang, T.; Gu, X.; Zheng, X.Y. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the

Heshun Region, China. *Int. J. Geogr. Inf. Sci.* 2010, 24, 107–127.

64. Zhang, L.; Liu, W.; Hou, K.; Lin, J.; Song, C.; Zhou, C.; Huang, B.; Tong, X.; Wang, J.; Rhine, W.; et al. Air pollution exposure associates with increased risk of neonatal jaundice. *Nat. Commun.* 2019, 10, 1–9.
65. Yin, Q.; Wang, J.; Ren, Z.; Li, J.; Guo, Y. Mapping the increased minimum mortality temperatures in the context of global climate change. *Nat. Commun.* 2019, 10, 1–8.
66. Wang, J.F.; Zhang, T.L.; Fu, B.J. A measure of spatial stratified heterogeneity. *Ecol. Indic.* 2016, 67, 250–256.
67. R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
68. Sofia, G., Prosdocimi, M., Dalla Fontana, G., & Tarolli, P. Modification of artificial drainage networks during the past half-century: Evidence and effects in a reclamation area in the Veneto floodplain (Italy). *Anthropocene*. 2014, 6, 48-62
69. Hansen, R.; Pauleit, S. From multifunctionality to multiple ecosystem services? A conceptual framework for multifunctionality in green infrastructure planning for urban areas. *Ambio*. 2014, 43, 516–529.
70. Szota, C.; Coutts, A.M.; Thom, J.K.; Virahsawmy, H.K.; Fletcher, T.D.; Livesley, S.J. Street tree stormwater control measures can reduce runoff but may not benefit established trees. *Landsc. Urban Plan.* 2019, 182, 144–155.
71. Zhang, H.; Cheng, J.; Wu, Z.; Li, C.; Qin, J.; Liu, T. Effects of impervious surface on the spatial distribution of urban water-logging risk spots at multiple scales in Guangzhou, South China. *Sustainability* 2018, 10, 1589.

# CHAPTER 6

## 6. Discussion

This thesis aims to reach a thorough understanding of the complex mechanism of waterlogging and the mitigation effect provided by urban green infrastructure in metropolitan coastal cities. Through four articles, the scale effects of urban waterlogging influencing factors, the spatial heterogeneous driving forces of waterlogging, the waterlogging susceptibility areas under different development scenarios, and the mitigation effect of urban green infrastructure on waterlogging are respectively investigated.

Firstly, for the first research question, our results successfully identified the urban waterlogging hot spots by using local spatial autocorrelation analysis. The identification of the urban waterlogging hot spots can help the government to gain a comprehensive understanding of the spatial distribution pattern of urban waterlogging events, which is particularly important for urban waterlogging mitigation and risk management. In detail, based on this result, local authorities can plan or build more drainage facilities as well as rational management of existing urban green spaces and impervious surfaces for the waterlogging hot spot areas. In addition, more priority can be given to urban waterlogging hot spot areas during the early warning of urban waterlogging risks when facing extreme weather. With the rapid development of spatial big data, the spatial distribution of population, medical and educational facilities can help to more accurately assess the vulnerability of urban waterlogging hotspots.

For the second research question, we first record the relative contributions of each influencing factor and describe their impact on urban waterlogging at multiple scales. Our results note that under different analysis scales, the dominant factors vary across different analysis scales, which present a strong scale effect. We infer that the inconsistent results of previous studies may be due to this scale effect. For example, the flat and low-lying surface will increase the risk of urban waterlogging in studies conducted in Huizhou, China (Wang et al., 2015) and Shenzhen, China (Wu and Zhang, 2017), but is associated with reducing the risk of urban waterlogging in Amsterdam, Netherlands (Gaitan et al., 2015). This is because different scholars use different

scales of analysis, which will interfere with the explanatory power of various factors and even change the correlation directions. These results further confirm why previous studies have produced different or even contradictory results. From the other side, due to the variation of waterlogging dominant factors at different analysis scales, we cannot simply determine a universal “optimal” analysis scale for urban waterlogging studies. This means that the appropriate analysis scale should be chosen according to the specific influencing factors and the characteristics of study areas.

For the third research question, we successfully reveal the spatial heterogeneity driving forces of urban waterlogging by integrating the cubist regression tree and the geographical detector model. In the past 20 years, a proliferation of studies focused on the driving forces of urban waterlogging (Wu et al., 2020; Zhang et al., 2020; Zhang et al., 2018). However, so far, most efforts in exploring this complex mechanism are limited to the use of global single-level analysis methods (Wu and Zhang, 2017; Tran et al., 2020; Sun, 2014). The global statistical methods mainly focus on revealing the general mechanism within the entire study area, while inevitably ignoring the local specific mechanism (Yu et al., 2018). Considering the spatial heterogeneity and complexity of the spatial distribution of urban landscape elements, the relationship between landscape elements and urban waterlogging may vary with different spatial locations. Based on our proposed method, this research confirms the driving force of urban waterlogging varies with local conditions. To some extent, these findings indicate that global statistical analysis may not be applicable to highly heterogeneous urban areas, which further promotes us to re-examine the results obtained based on global statistical methods to avoid its shortcomings.

For the fourth research question, we further characterize the individual and interaction effect of local driving forces at different spatial locations to develop the site-specific waterlogging mechanism. Beyond the relative contribution as revealed in previous studies (Zhang et al., 2020; Yu et al., 2018), we further detect the interaction effect locally, and thus, improve our understanding of the site-specific waterlogging mitigation strategies. Some watershed units have an obvious dominant factor with an overwhelming independent contribution that far exceeds the remaining factors. Therefore, for these watershed units, the local authorities can develop urban waterlogging mitigation strategies based on their dominant factor. In contrast, some watershed units are jointly influenced by multiple factors, with the independent contribution of each factor is not much different, but a strong interaction enhancement of each factor. For these watershed units, it is inappropriate to develop mitigation strategies based on one single factor; instead, the interaction effect



between each factor needs to be comprehensively considered. This allows us to develop more targeted and effective waterlogging mitigation strategies, rather than the “one-size-fits-all” policy. In other words, urban waterlogging mitigation strategies developed from the result of global statistical methods may not be applicable in some areas. Strictly speaking, it is unrealistic to develop locally targeted mitigation strategies drawn from the general conclusions.

For the fifth question, we successfully simulate and predict urban waterlogging in highly heterogeneous urbanized areas through the combination of the stepwise cluster analysis model (SCAM) and hierarchical partitioning analysis (HPA). Compared with logistic regression, artificial neural networks, support vector machines, this proposed method (SCAM-HPA) has high classification accuracy and generalization capability, which can be applied as a general framework to other cities for urban waterlogging risk identification and control. Under different urbanization and precipitation scenarios, we notice that watershed units in different locations have different sensitivity to land-use change and rainfall change scenarios. For example, areas with large impervious surface abundance are more sensitive to rainfall change scenarios. Thereby, the government should rationally manage the existing land cover features (vegetation abundance) or increase drainage capacity to cope with future urban waterlogging, even though the pressure of land-use development is relatively high. In contrast, urban fringe regions are more sensitive to land-use change scenarios. Thus the local authorities should strictly control the speed of urban expansion to avoid exacerbating the frequency of urban waterlogging disasters in these areas.

For the sixth research question, we demonstrate that the mitigation effect of UGI on waterlogging presents a threshold phenomenon. This implies that the excessive proportions of UGI may lead to a waste of its mitigation effect. In contrast to previous studies, most efforts have only confirmed that increasing the area of green infrastructure can more effectively alleviate urban waterlogging (Yang et al., 2013; Wen et al., 2016). This is why most urban waterlogging mitigation strategies are based on increasing the amount of green space. Traditionally, urban planners increase the area of UGI to reduce surface runoff (Yu et al., 2020; Richards et al., 2015). However, as the land use pressure within the urban centers is too large to increase the green area greatly, which cannot be practically applied to guide UGI planning. Therefore, our findings provide a new perspective on urban waterlogging mitigation strategies—blindly increasing the area of green infrastructure may not bring much improvement to urban waterlogging mitigation. It is recommended to consider comprehensively the area proportion of UGI and its mitigation effect when planning UGI.

Lastly, for the seventh research question, our results indicate that the area proportion and biophysical parameters (EVI) of UGI both have a non-negligible effect in alleviating urban waterlogging. However, most previous studies ignore the influence of EVI, which only analyzes the effects of different sizes of green infrastructure on urban waterlogging (Kim and Park, 2016; Yao et al., 2015; Yang et al., 2015). Ignoring the biophysical parameters of UGI will inevitably lead to some deviation. For the same UGI area, healthier or denser vegetation (better vegetation conditions with high EVI value) can more effectively intercept and store surface runoff. Therefore, while increasing the area of UGI, it is also necessary to improve the vegetation conditions, thereby enhancing the mitigation effect of urban waterlogging. Moreover, our study further confirmed the interaction effect of spatial configuration on waterlogging. This finding refreshes our perception of the importance of the interaction of landscape patterns. The interaction between spatial configuration and composition can more significantly improve the mitigation capacity of UGI. In previous studies, the importance of the UGI spatial configuration is often overlooked (Kim and Park, 2016; Yang et al., 2013; Armson et al., 2013), which is mainly due to its relatively small individual effects. However, the interaction effect has led to a renewed awareness of the importance of the UGI landscape patterns for urban waterlogging mitigation. The results of this study provide a theoretical reference for urban planners to optimize the spatial configuration of UGI in urban centers.

These results contribute to extending our understanding of the complex mechanism of waterlogging in highly urbanized coastal cities and provide a theoretical and practical reference for the prevention and control of urban waterlogging and the design of UGI. However, this thesis has its limitations that should be considered in future work. Firstly, the urban waterlogging data were obtained from local water authority historical records. However, due to the limitations and accessibility of urban waterlogging records, the urban waterlogging data did not record the size (water depth, area), duration, and the specific year of each event (just recorded in the period 2009-2015). We only analyze the waterlogging mechanism from the whole period, which inevitably brings some uncertainty to the results. Also, due to this reason, we could not evaluate the temporal change of the waterlogging events. Secondly, two major cities in Guangdong–Hong Kong–Macao Greater Bay Metropolitan Region, Guangzhou, and Shenzhen cities, were selected for this research. To reveal the mechanism of waterlogging and its mitigation strategy in two cities may be limited. Revealing the waterlogging driving forces and the role of UGI in urban waterlogging in other regions may help us confirm the universality of our findings. Additionally, the proposed methods that integrated the best subset regression model, cubist regression tree, and geographical detector model, and the SCAM-HPA framework

have not been tested in multiple regions. Therefore, in future studies, we can extend it to more regional comparative studies under different environmental conditions and urbanization backgrounds to further verify the universality and credibility.

# CHAPTER 7

## 7. Conclusions

Under the combined effects of climate change and urbanization, urban waterlogging seriously threatens urban sustainable development. Given the severe urban waterlogging problems in metropolitan coastal cities, two highly urbanized coastal cities, Guangzhou and Shenzhen are selected as the study area. To have a comprehensive understanding of the urban waterlogging mechanism, this thesis first offers the interpretation of the scale effect and identifies the dominant factors of urban waterlogging at different analysis scales. Secondly, considering the scale effect and the spatial heterogeneity effects of urban landscape elements on urban waterlogging, an innovative method that integrates the cubist regression tree, and geographical detector model is proposed to map the spatial heterogeneous driving forces of urban waterlogging. Further, the thesis highlights the fact that conventional methods have difficulty in adequately capturing the spatial variation of urban waterlogging as well as identifying the waterlogging susceptibility areas. A more robust method that combines stepwise cluster analysis model (SCAM) and hierarchical partitioning analysis (HPA) is proposed to simulate urban waterlogging variation and assess waterlogging susceptibility under different land cover change and climate change scenarios. Lastly, the effectiveness and stability of urban green infrastructure (UGI) in mitigating urban waterlogging are assessed from two aspects: quantifying the threshold level of UGI alleviating urban waterlogging and identifying UGI factors that can effectively alleviate urban waterlogging (independent contribution and interaction effect). The main conclusions of this thesis are as follows:

### 1) Spatial distribution of urban waterlogging

The local Moran's I and Getis-G statistic both indicate the spatial distribution of waterlogging events in Guangzhou and Shenzhen has a significant spatial agglomeration effect. The urban waterlogging hot spots in Guangzhou are mainly concentrated in the historical urban areas of Guangzhou (Liwan, Yuexiu, Haizhu district), presenting a single-core aggregation pattern. As for Shenzhen, the waterlogging hot spots are

sparingly distributed in the urban sub-centers (Futian, Luohu, Longhua, Longgang District). Although both cities show the agglomeration effect of urban waterlogging, the spatial clustering effect is more prominent in Guangzhou; instead, the agglomeration effect in Shenzhen presents a more dispersed distribution pattern.

## 2) Scale effects of various influencing factors

The correlation directions of Pearson, the determination coefficients of stepwise regression models, the relative contributions of influencing factors, and the dominant factors to waterlogging are different across all analysis scales. Under a small analysis scale (i.e. 1 km), the influence of topography factors on waterlogging magnitude is greater than other factors. However, with the increasing of the analysis scales (i.e. 3 km), the explanatory power of topography factors gradually declines and the land cover composition is gradually becoming the dominant factor of waterlogging. Since the dominant drivers vary across different analysis scales, the appropriate analysis scale for urban waterlogging studies only works for specific influencing factors, and thus the appropriate analysis scale for urban waterlogging study should be determined by the characteristics of study areas.

## 3) Urban waterlogging spatially heterogeneous mechanism

The combination of cubist regression tree and geographical detector model can fully quantify the spatial non-stationarity effect of representative driving factors on urban waterlogging and spatially explicit the driving forces with different local conditions. To some extent, these findings indicate that global statistical methods that only provide a general mechanism may not be applicable to highly heterogeneous urban areas. Indisputably, the local waterlogging mechanisms provide a theoretical and practical reference for the development of site-specific waterlogging mitigation strategies. In strong dominant watersheds, we can develop related mitigation strategies based on the dominant factors in these watershed units. On the other hand, in weak dominant watersheds, the interaction effects of representative factors must be considered comprehensively when formulating the urban waterlogging mitigation strategy. Understanding the complex site-specific mechanism of urban waterlogging in different spatial locations will help us implement more targeted and effective mitigation strategies, rather than a “one-size-fits-all” policy.

## 4) Urban waterlogging susceptibility assessment

The proposed SCAM-HPA framework is an effective and feasible solution for urban waterlogging variation simulation, which provides accurate and detailed simulated results both in urban centers where waterlogging

frequently occurs and urban fringe with few waterlogging events. By comparing the performance with other models, the SCAM-HPA can be applied as a general framework to other cities for urban waterlogging simulation and risk assessment. Under different urbanization and precipitation scenarios, the urban waterlogging susceptibility areas have a considerable variation. For the areas with large impervious surface abundance, with the increase of cumulative precipitation, the increase in waterlogging density is significantly higher than that of areas with a lower abundance of impervious surface. This indicates that these regions are more sensitive to rainfall change scenarios. Thereby, the impervious surface and urban green area in these areas should be properly adjusted, even though the land-use pressure is very high. In contrast, for the areas with relatively high vegetation coverage, in the land-use change scenarios, a large number of green spaces are converted into impervious surfaces, resulting in a significant increase in the risk of urban waterlogging. These watershed units are more sensitive to land-use change scenarios. Therefore, the local authorities should be strictly controlled the speed of urban expansion to avoid exacerbating the frequency of urban waterlogging disasters.

#### 5) The threshold level of waterlogging mitigation capacity

This thesis reveals that the impact of green infrastructure on urban waterlogging is not a simple linear relationship. As the area proportion of UGI within the watershed exceeds the threshold level, the waterlogging density will not continue to decline as the UGI area increases. The excessive proportions of UGI within the watershed unit may lead to a waste of its mitigation effect. Additionally, the threshold values of the mean patch size (MPS) indicate that green infrastructure sizes need to be maintained in a certain area. Green infrastructure sizes that are too small or too large may not be effective in alleviating urban waterlogging. Therefore, the area proportion of UGI and its mitigation effect should be considered comprehensively when planning UGI. Since the threshold values of some UGI indicators are different among cities, the thresholds values are disturbed by regional characteristics. Therefore, the UGI-based waterlogging prevention strategies should be adapted to local conditions.

#### 6) Enhance the mitigation capacity of green infrastructure on waterlogging

The impact of green infrastructure on urban waterlogging largely depends on its area and vegetation status. This finding provides practical insights into UGI planning, i.e., while increasing the area of UGI, more attention should also be paid to the biophysical parameter of vegetation, thereby improving the mitigation effect of green infrastructure from the “size” and “health”. Moreover, the interaction of UGI factors greatly enhances

its impact on urban waterlogging. The interaction between UGI area proportion and EVI has the largest contribution to the mitigation of urban waterlogging, indicating that neither the size of the UGI nor its biophysical parameter (vegetation status) can be ignored. The combination of UGI area proportion and good vegetation status further improves the mitigation of urban waterlogging. Furthermore, this thesis highlights the importance of the combination of UGI composition and spatial configuration. Based on a certain percentage of UGI, the interaction of UGI composition and configuration can further enhance its impact on urban waterlogging, which has important implications for the metropolis with a shortage of urban land resources for green infrastructure.

In the context of rapid urbanization in recent years, urban waterlogging has gradually become one of the most critical issues threatening human activities and the economy. Based on the conclusions drawn from this thesis, in future urban development, on the one hand, it is necessary to reduce the widespread of impervious surfaces as much as possible to protect urban green infrastructure that plays an important role in the urban hydrological cycle. On the other hand, there is also a need to implement site-specific urban waterlogging mitigation strategies according to local conditions as well as introduce effective engineering solutions, such as sponge cities, and low-impact development to reduce urban waterlogging. Given the growing concerns of global warming and continued rapid urbanization, we believe that our findings provide useful enlightenment for local authorities in urban waterlogging management and UN 2030 Sustainable Development Goals.

## 8. References

1. Ahammed, F. (2017). A review of water-sensitive urban design technologies and practices for sustainable stormwater management. *Sustainable Water Resources Management*, 3.
2. Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., ... & Feyen, L. (2017). Global projections of river flood risk in a warmer world. *Earth's Future*, 5(2), 171-182.
3. Amaguchi, H., Kawamura, A., Olsson, J., & Takasaki, T. (2012). Development and testing of a distributed urban storm runoff event model with a vector-based catchment delineation. *Journal of Hydrology*, 420, 205-215.
4. Armson, D., Stringer, P., & Ennos, A. R. (2013). The effect of street trees and amenity grass on urban surface water runoff in Manchester, UK. *Urban Forestry & Urban Greening*, 12(3), 282-286.
5. Avand, M., Moradi, H. R., & Ramazanzadeh Lasboyee, M. (2021). Spatial prediction of future flood risk: an approach to the effects of climate change. *Geosciences*, 11(1), 25.
6. Babaei, S., Ghazavi, R., & Erfanian, M. (2018). Urban flood simulation and prioritization of critical urban sub-catchments using SWMM model and PROMETHEE II approach. *Physics and Chemistry of the Earth, Parts A/B/C*, 105, 3-11.
7. Berghuijs, W. R., Aalbers, E. E., Larsen, J. R., Trancoso, R., & Woods, R. A. (2017). Recent changes in extreme floods across multiple continents. *Environmental Research Letters*, 12(11), 114035.
8. Berndtsson, J. C. (2010). Green roof performance towards management of runoff water quantity and quality: A review. *Ecological engineering*, 36(4), 351-360.
9. Bisht, D. S., Chatterjee, C., Kalakoti, S., Upadhyay, P., Sahoo, M., & Panda, A. (2016). Modeling urban floods and drainage using SWMM and MIKE URBAN: a case study. *Natural Hazards*, 84(2), 749-776.
10. Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A., Parajka, J., Merz, B., ... & Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, 573(7772), 108-111.
11. Bonneau, J., Fletcher, T. D., Costelloe, J. F., & Burns, M. J. (2017). Stormwater infiltration and the 'urban karst'—A review. *Journal of hydrology*, 552, 141-150.
12. Brito, M. M., Almoradie, A., & Evers, M., (2019). Spatially-explicit sensitivity and uncertainty analysis in a MCDA-based flood vulnerability model. *International Journal of Geographical Information Science*, 33(9), 1788–1806.
13. Bryndal, T., Franczak, P., Krocak, R., Cabaj, W., & Kołodziej, A. (2017). The impact of extreme



rainfall and flash floods on the flood risk management process and geomorphological changes in small Carpathian catchments: a case study of the Kasiniczanka river (Outer Carpathians, Poland). *Natural Hazards*, 88(1), 95-120.

14. Burger, G., Sitzenfrei, R., Kleidorfer, M., & Rauch, W., (2014). Parallel flow routing in SWMM 5. *Environmental Modelling and Software*, 53, 27–34.
15. Burns, M. J., Fletcher, T. D., Duncan, H. P., Hatt, B. E., Ladson, A. R., & Walsh, C. J. (2015). The performance of rainwater tanks for stormwater retention and water supply at the household scale: an empirical study. *Hydrological Processes*, 29(1), 152-160.
16. Cai, J. B., Lin, N., Du X. S., & Feng, Y. G. (2011). Effect Analysis of Low-lying Greenbelt to Reduce Urban Runoff Depth and Runoff Coefficient [J]. *Urban Roads Bridges & Flood Control*, 6, 119-122.
17. Changnon Jr, S. A., Huff, F. A., & Semonin, R. G. (1971). METROMEX: An investigation of inadvertent weather modification. *Bulletin of the American Meteorological Society*, 52(10), 958-968.
18. Changnon, S. (Ed.). (2016). METROMEX: a review and summary.
19. Chen, J., Huang, G., & Chen, W. (2021). Towards better flood risk management: assessing flood risk and investigating the potential mechanism based on machine learning models. *Journal of environmental management*, 293, 112810.
20. Chen, K., Wang, X., Li, D., & Xue, D. (2013). The morphological evolution of river and water body in urban area of Guangzhou City in 1990–2010. *Scientia Geographica Sinica*, 33(2), 223-230.
21. Chen, Y., & Chen, H. (2020). The Collective Strategies of Key Stakeholders in Sponge City Construction: A Tripartite Game Analysis of Governments, Developers, and Consumers. *Water*, 12(4), 1087.
22. Chen, Y., Zhou, H., Zhang, H., Du, G., & Zhou, J. (2015). Urban flood risk warning under rapid urbanization. *Environmental Research*, 139, 3–10.
23. Cheng, T., Xu, Z., Hong, S., & Song, S., (2017). Flood Risk Zoning by Using 2D Hydrodynamic Modeling: A Case Study in Jinan City. *Mathematical Problems in Engineering*, 2017.
24. China Flood and Drought Disaster Bulletin. Available at: <http://www.mwr.gov.cn/sj/tjgb/zgshzhgb/>
25. China Urban Development Report No. 12 (2019). Available at: <https://www.ssap.com.cn/c/2019-10-30/1081675.shtml>
26. Chowdary, V. M., Chakraborty, D., Jeyaram, A., Murthy, Y. V. N. K., Sharma, J. R., & Dadhwal, V. K., (2013). Multi-Criteria Decision Making Approach for Watershed Prioritization Using Analytic

Hierarchy Process Technique and GIS. *Water Resources Management*, 27(10), 3555–3571.

27. Chu, H. J., Lin, Y. P., Huang, C. W., Hsu, C. Y., & Chen, H. Y. (2010). Modelling the hydrologic effects of dynamic land - use change using a distributed hydrologic model and a spatial land - use allocation model. *Hydrological Processes*, 24(18), 2538-2554.
28. Demuzere, M., Orru, K., Heidrich, O., Olazabal, E., Geneletti, D., Orru, H., ... & Faehnle, M. (2014). Mitigating and adapting to climate change: Multi-functional and multi-scale assessment of green urban infrastructure. *Journal of environmental management*, 146, 107-115.
29. Dong, W., Lian, Y., & Zhang, Y. (2018). Sustainable Development of Water Resources and Hydraulic Engineering in China Proceedings for the 2016 International Conference on Water Resource and Hydraulic Engineering.
30. Du, S., Van Rompaey, A., & Shi, P. (2015). A dual effect of urban expansion on flood risk in the Pearl River Delta (China) revealed by land-use scenarios and direct runoff simulation. *Natural Hazards*, 77(1), 111-128.
31. Du, S., Wang, C., Shen, J., Wen, J., Gao, J., Wu, J., ... & Xu, H. (2019). Mapping the capacity of concave green land in mitigating urban pluvial floods and its beneficiaries. *Sustainable Cities and Society*, 44, 774-782.
32. Elliott, A. H., & Trowsdale, S. A. (2007). A review of models for low impact urban stormwater drainage. *Environmental modelling & software*, 22(3), 394-405.
33. Fahy, B., Brenneman, E., Chang, H., & Shandas, V. (2019). Spatial analysis of urban flooding and extreme heat hazard potential in Portland, OR. *International journal of disaster risk reduction*, 39, 101117.
34. Fang, C., & Wang, D. (2011). Comprehensive development measuring and improving roadmap of China's urbanization quality. *Geography Research*, 30(11), 1931-1945.
35. Fang, J., Liu, W., Yang, S., Brown, S., Nicholls, R. J., Hinkel, J., ... & Shi, P. (2017). Spatial-temporal changes of coastal and marine disasters risks and impacts in Mainland China. *Ocean & Coastal Management*, 139, 125-140.
36. Fang, J., Wahl, T., Fang, J., Sun, X., Kong, F., & Liu, M. (2021). Compound flood potential from storm surge and heavy precipitation in coastal China: dependence, drivers, and impacts. *Hydrology and Earth System Sciences*, 25(8), 4403-4416.
37. Felder, G., Zischg, A., & Weingartner, R. (2017). The effect of coupling hydrologic and hydrodynamic

models on probable maximum flood estimation. *Journal of hydrology*, 550, 157-165.

38. Few, R. (2013). *Flood hazards, vulnerability and risk reduction* (pp. 20-39). Routledge.
39. Freeborn, J. R., Sample, D. J., & Fox, L. J. (2012). Residential stormwater: methods for decreasing runoff and increasing stormwater infiltration. *Journal of Green Building*, 7(2), 15-30.
40. Gaitan, S., & Ten Veldhuis, J. A. E. (2015). Opportunities for multivariate analysis of open spatial datasets to characterize urban flooding risks. *Proceedings of the International Association of Hydrological Sciences*, 370, 9-14.
41. Gallien, T. W., Sanders, B. F., & Flick, R. E. (2014). Urban coastal flood prediction: Integrating wave overtopping, flood defenses and drainage. *Coastal Engineering*, 91, 18–28.
42. Ghimire, B., Chen, A. S., Guidolin, M., Keedwell, E. C., Djordjević, S., & Savić, D. A. (2013). Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. *Journal of Hydroinformatics*, 15(3), 676-686.
43. Gilroy, K. L., & McCuen, R. H. (2012). A nonstationary flood frequency analysis method to adjust for future climate change and urbanization. *Journal of hydrology*, 414, 40-48.
44. Gironás, J., Roesner, L. A., Rossman, L. A., & Davis, J. (2010). A new applications manual for the Storm Water Management Model(SWMM). *Environmental Modelling & Software*, 25(6), 813-814.
45. Global Emergency Database (2020). CRED—Human cost of disasters. An overview of the last 20 years 2000-2019. <https://cred.be/sites/default/files/CRED-Disaster-ReportHuman-Cost2000-2019.pdf>
46. Green Roofs for Healthy Cities' Awards of Excellence. Available at: <https://greenroofs.org/awards>
47. Güneralp, B., Güneralp, İ., & Liu, Y. (2015). Changing global patterns of urban exposure to flood and drought hazards. *Global environmental change*, 31, 217-225.
48. Gupta, A., Bansal, A., Gupta, R., Naryani, D., & Sood, A. (2017, December). Urban waterlogging detection and severity prediction using artificial neural networks. In *2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)* (pp. 42-49). IEEE.
49. Hall, A. (2010). *Green infrastructure case studies: municipal policies for managing stormwater with Green Infrastructure*. EPA Office of Wetlands, Oceans and Watersheds, 2.
50. Hallegatte, S., Green, C., Nicholls, R. J., & Corfee-Morlot, J. (2013). Future flood losses in major

coastal cities. *Nature climate change*, 3(9), 802-806.

51. Han, J. Y., Baik, J. J., & Lee, H. (2014). Urban impacts on precipitation. *Asia-Pacific Journal of Atmospheric Sciences*, 50(1), 17-30.
52. He, T., Xiao, W., Zhao, Y., Deng, X., & Hu, Z. (2020). Identification of waterlogging in Eastern China induced by mining subsidence: A case study of Google Earth Engine time-series analysis applied to the Huainan coal field. *Remote Sensing of Environment*, 242, 111742.
53. Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., ... & Kanae, S. (2013). Global flood risk under climate change. *Nature climate change*, 3(9), 816-821.
54. Hong, H., Pradhan, B., Bui, D. T., Xu, C., Youssef, A. M., & Chen, W. (2017). Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Suichuan area (China). *Geomatics, Natural Hazards and Risk*, 8(2), 544-569.
55. Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A. X., & Chen, W. (2018). Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. *Science of the total environment*, 625, 575-588.
56. Hu, S., Fan, Y., & Zhang, T. (2020). Assessing the effect of land use change on surface runoff in a rapidly urbanized city: A case study of the central area of Beijing. *Land*, 9(1), 17.
57. Huang, H., Chen, X., Zhu, Z., Xie, Y., Liu, L., Wang, X., ... Liu, K. (2018). The changing pattern of urban flooding in Guangzhou, China. *Science of the Total Environment*, 622–623, 394–404.
58. Huang, Q., Wang, J., Li, M., Fei, M., & Dong, J. (2017). Modeling the influence of urbanization on urban pluvial flooding: a scenario-based case study in Shanghai, China. *Natural Hazards*, 87(2), 1035-1055.
59. Huang, T., Wang, Y., & Zhang, J. (2017). Simulation and Evaluation of Low Impact Development of Urban Residential District Based on SWMM and GIS. *IOP Conference Series: Earth and Environmental Science*, 74(1).
60. IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2014.
61. Jha, A. K., Bloch, R., & Lamond, J. (2012). *Cities and flooding: a guide to integrated urban flood risk management for the 21st century*. World Bank Publications.
62. Jian, W., Li, S., Lai, C., Wang, Z., Cheng, X., Lo, E. Y. M., & Pan, T. C. (2021). Evaluating pluvial

flood hazard for highly urbanised cities: a case study of the Pearl River Delta Region in China. *Natural Hazards*, 105(2), 1691-1719.

63. Jiang, B. (2015). Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity. *GeoJournal*, 80(1), 1-13.
64. Ke, Q., Tian, X., Bricker, J., Tian, Z., Guan, G., Cai, H., ... & Liu, J. (2020). Urban pluvial flooding prediction by machine learning approaches—a case study of Shenzhen city, China. *Advances in Water Resources*, 145, 103719.
65. Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A., (2012). An artificial neural network model for flood simulation using gis: Johor river basin, Malaysia. *Environmental Earth sciences*, 67(1), 251-264.
66. Kim, H. W., & Park, Y. (2016). Urban green infrastructure and local flooding: The impact of landscape patterns on peak runoff in four Texas MSAs. *Applied geography*, 77, 72-81.
67. Kim, Y., Eisenberg, D. A., Bondank, E. N., Chester, M. V., Mascaro, G., & Underwood, B. S. (2017). Fail-safe and safe-to-fail adaptation: decision-making for urban flooding under climate change. *Climatic Change*, 145(3), 397-412.
68. Krebs, G., Kuoppamäki, K., Kokkonen, T., & Koivusalo, H. (2016). Simulation of green roof test bed runoff. *Hydrological processes*, 30(2), 250-262.
69. Lee, E. H., Lee, Y. S., Joo, J. G., Jung, D., & Kim, J. H. (2016). Flood reduction in urban drainage systems: Cooperative operation of centralized and decentralized reservoirs. *Water*, 8(10), 469.
70. Lee, Y., & Brody, S. D. (2018). Examining the impact of land use on flood losses in Seoul, Korea. *Land use policy*, 70, 500-509.
71. Li, C., Cheng, X., Li, N., Du, X., Yu, Q., & Kan, G. (2016). A framework for flood risk analysis and benefit assessment of flood control measures in Urban Areas. *International Journal of Environmental Research and Public Health*, 13(8).
72. Li, C., Liu, M., Hu, Y., Shi, T., Qu, X., & Walter, M. T. (2018). Effects of urbanization on direct runoff characteristics in urban functional zones. *Science of the Total Environment*, 643, 301-311.
73. Li, W., Lin, K., Zhao, T., Lan, T., Chen, X., Du, H., & Chen, H. (2019a). Risk assessment and sensitivity analysis of flash floods in ungauged basins using coupled hydrologic and hydrodynamic models. *Journal of hydrology*, 572, 108-120.
74. Li, Y., Hu, T., Zheng, G., Shen, L., Fan, J., & Zhang, D. (2019b). An Improved Simplified Urban Storm

Inundation Model Based on Urban Terrain and Catchment Modification. *Water*, 11(11), 2335.

75. Lim, K. S., & Lee, D. R. (2009). The spatial MCDA approach for evaluating flood damage reduction alternatives. *KSCE Journal of Civil Engineering*, 13(5), 359-369.
76. Lin, B., Wicks, J. M., Falconer, R. A., & Adams, K. (2006, March). Integrating 1D and 2D hydrodynamic models for flood simulation. In *Proceedings of the Institution of Civil Engineers-Water Management* (Vol. 159, No. 1, pp. 19-25). Thomas Telford Ltd.
77. Lin, J., He, X., Lu, S., Liu, D., & He, P. (2021). Investigating the influence of three-dimensional building configuration on urban pluvial flooding using random forest algorithm. *Environmental Research*, 196, 110438.
78. Lin, T., Liu, X., Song, J., Zhang, G., Jia, Y., Tu, Z., ... & Liu, C. (2018). Urban waterlogging risk assessment based on internet open data: A case study in China. *Habitat International*, 71, 88-96.
79. Littidej, P., & Buasri, N. (2019). Built-Up Growth Impacts on Digital Elevation Model and Flood Risk Susceptibility Prediction in Muaeng District, Nakhon Ratchasima (Thailand). *Water*, 11(7), 1496.
80. Liu, F., Liu, X., Xu, T., Yang, G., & Zhao, Y. (2021). Driving Factors and Risk Assessment of Rainstorm Waterlogging in Urban Agglomeration Areas: A Case Study of the Guangdong-Hong Kong-Macao Greater Bay Area, China. *Water*, 13(6), 770.
81. Liu, J., Shao, W., Xiang, C., Mei, C., & Li, Z. (2020). Uncertainties of urban flood modeling: Influence of parameters for different underlying surfaces. *Environmental research*, 182, 108929.
82. Liu, J., Shao, W., Xiang, C., Mei, C., & Li, Z. (2020a). Uncertainties of urban flood modeling: Influence of parameters for different underlying surfaces. *Environmental research*, 182, 108929.
83. Liu, W., Chen, W., & Feng, Q. (2018). Field simulation of urban surfaces runoff and estimation of runoff with experimental curve numbers. *Urban Water Journal*, 15(5), 418-426.
84. Liu, W., Chen, W., & Peng, C. (2014). Assessing the effectiveness of green infrastructures on urban flooding reduction: A community scale study. *Ecological Modelling*, 291, 6-14.
85. Liu, W., Chen, W., & Peng, C. (2015a). Influences of setting sizes and combination of green infrastructures on community's stormwater runoff reduction. *Ecological Modelling*, 318, 236-244.
86. Liu, W., Chen, W., Peng, C., Wu, L., & Qian, Y. (2015b). A water balance approach to assess rainwater availability potential in urban areas: the case of Beijing, China. *Water Science and Technology: Water Supply*, 15(3), 490-498.
87. Liu, W., Feng, Q., Chen, W., & Deo, R. C. (2020b). Stormwater runoff and pollution retention

performances of permeable pavements and the effects of structural factors. *Environmental Science and Pollution Research*, 27, 30831-30843.

88. Liu, W., Feng, Q., Deo, R. C., Yao, L., & Wei, W. (2020c). Experimental study on the rainfall-runoff responses of typical urban surfaces and two green infrastructures using scale-based models. *Environmental management*, 66(4), 683-693.
89. Mai, Y., Zhang, M., Chen, W., Chen, X., Huang, G., & Li, D. (2018). Experimental study on the effects of LID measures on the control of rainfall runoff. *Urban Water Journal*, 15(9), 827-836.
90. Mentens, J., Raes, D., & Hermy, M. (2006). Green roofs as a tool for solving the rainwater runoff problem in the urbanized 21st century?. *Landscape and urban planning*, 77(3), 217-226.
91. Miao, Z. T., Han, M., & Hashemi, S. (2019). The effect of successive low-impact development rainwater systems on peak flow reduction in residential areas of Shizhuang, China. *Environmental earth sciences*, 78(2), 51.
92. Mignot, E., Li, X., & Dewals, B. (2019). Experimental modelling of urban flooding: A review. *Journal of Hydrology*, 568, 334-342.
93. Miller, J. D., & Hutchins, M. (2017). The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *Journal of Hydrology: Regional Studies*, 12, 345-362.
94. Min, S. K., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. *Nature*, 470(7334), 378-381.
95. Ministry of Housing and Urban-Rural Development of the People's Republic of China (2019). "Outdoor Drainage System Design Standard".
96. Ministry of Housing and Urban-Rural Development of the People's Republic of China (2014). *Technical Guidance for Sponge City: Constructing Low Impact Development Rainwater System* (MHURD, 2014-275).
97. Mitra, C., & Shepherd, J. M. (2015). Urban precipitation: A global perspective. In *The Routledge Handbook of Urbanization and Global Environmental Change* (pp. 176-192). Routledge.
98. Myhre, G., Alterskjær, K., Stjern, C. W., Hodnebrog, Ø., Marelle, L., Samset, B. H., ... & Stohl, A. (2019). Frequency of extreme precipitation increases extensively with event rareness under global warming. *Scientific reports*, 9(1), 1-10.
99. Nigusse, A. G., & Adhanom, O. G. (2019). Flood hazard and flood risk vulnerability mapping using

- geo-spatial and MCDA around Adigrat, Tigray region, Northern Ethiopia. *Momona Ethiopian Journal of Science*, 11(1), 90-107.
100. Nigussie, T. A., & Altunkaynak, A. (2016). Assessing the hydrological response of Ayamama watershed from urbanization predicted under various landuse policy scenarios. *Water Resources Management*, 30(10), 3427-3441.
  101. Paiva, R. C. D., Collischonn, W., & Tucci, C. E. M., (2011). Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *Journal of Hydrology*, 406(3–4), 170–181.
  102. Papadakis, C., & Preul, H. C. (1972). University of Cincinnati urban runoff model. *Journal of the Hydraulics Division*, 98(10), 1789-1804.
  103. Papalexiou, S. M., & Montanari, A. (2019). Global and regional increase of precipitation extremes under global warming. *Water Resources Research*, 55(6), 4901-4914.
  104. Paquette, J., & Lowry, J. (2012). Flood hazard modelling and risk assessment in the Nadi River Basin, Fiji, using GIS and MCDA. *The South Pacific Journal of Natural and Applied Sciences*, 30(1), 33-43.
  105. Peng, L. L., & Jim, C. Y. (2015). Economic evaluation of green-roof environmental benefits in the context of climate change: The case of Hong Kong. *Urban Forestry & Urban Greening*, 14(3), 554-561.
  106. Pickett, S. T., & Cadenasso, M. L. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science*, 269(5222), 331-334.
  107. Pinho, J., Ferreira, R., Vieira, L., & Schwanenberg, D. (2015). Comparison between two hydrodynamic models for flooding simulations at river Lima basin. *Water resources management*, 29(2), 431-444.
  108. Pradhan, B., & Youssef, A. M. (2011). A 100 - year maximum flood susceptibility mapping using integrated hydrological and hydrodynamic models: Kelantan River Corridor, Malaysia. *Journal of Flood Risk Management*, 4(3), 189-202.
  109. Quan, R. S. (2014a). Rainstorm waterlogging risk assessment in central urban area of Shanghai based on multiple scenario simulation. *Natural hazards*, 73(3), 1569-1585.
  110. Quan, R. S. (2014b). Risk assessment of flood disaster in Shanghai based on spatial–temporal characteristics analysis from 251 to 2000. *Environmental earth sciences*, 72(11), 4627-4638.
  111. Quan, R., Zhang, L., Liu, M., Lu, M., Wang, J., & Niu, H. (2011, June). Risk assessment of rainstorm waterlogging on subway in central urban area of Shanghai, China based on scenario simulation. In



- 2011 19th International Conference on Geoinformatics (pp. 1-6). IEEE.
112. Rai, P. K., Chahar, B. R., & Dhanya, C. T. (2017). GIS-based SWMM model for simulating the catchment response to flood events. *Hydrology Research*, 48(2), 384–394.
  113. Richards, P. J., Farrell, C., Tom, M., Williams, N. S., & Fletcher, T. D. (2015). Vegetable raingardens can produce food and reduce stormwater runoff. *Urban Forestry & Urban Greening*, 14(3), 646-654.
  114. Rodriguez-Hernandez, J., Andrés-Valeri, V. C., Ascorbe-Salcedo, A., & Castro-Fresno, D. (2016). Laboratory study on the stormwater retention and runoff attenuation capacity of four permeable pavements. *Journal of Environmental Engineering*, 142(2), 04015068.
  115. Rossman, L. A. (2010). Modeling low impact development alternatives with SWMM. *Journal of Water Management Modeling*.
  116. Roy, S., Bose, A., Singha, N., Basak, D., & Chowdhury, I. R. (2021). Urban waterlogging risk as an undervalued environmental challenge: An Integrated MCDA-GIS based modeling approach. *Environmental Challenges*, 4, 100194.
  117. Samanta, S., Koloa, C., Pal, D. K., & Palsamanta, B. (2016). Flood risk analysis in lower part of Markham river based on multi-criteria decision approach (MCDA). *Hydrology*, 3(3), 1–13.
  118. Seenu, P. Z., Rathnam, E. V., & Jayakumar, K. V. (2020). Visualisation of urban flood inundation using SWMM and 4D GIS. *Spatial Information Research*, 28(4), 459-467.
  119. Shao, W., Zhang, H., Liu, J., Yang, G., Chen, X., Yang, Z., & Huang, H. (2016). Data Integration and its Application in the Sponge City Construction of CHINA. *Procedia Engineering*, 154, 779–786.
  120. Shen, Y., Morsy, M. M., Huxley, C., Tahvildari, N., & Goodall, J. L. (2019). Flood risk assessment and increased resilience for coastal urban watersheds under the combined impact of storm tide and heavy rainfall. *Journal of Hydrology*, 579, 124159.
  121. Shepherd, J. M. (2005). A review of current investigations of urban-induced rainfall and recommendations for the future. *Earth Interactions*, 9(12), 1-27.
  122. Shepherd, J. M., Pierce, H., & Negri, A. J. (2002). Rainfall modification by major urban areas: Observations from spaceborne rain radar on the TRMM satellite. *Journal of applied meteorology*, 41(7), 689-701.
  123. Sofia, G., Roder, G., Dalla Fontana, G., & Tarolli, P. (2017). Flood dynamics in urbanised landscapes: 100 years of climate and humans' interaction. *Scientific reports*, 7(1), 1-12.
  124. Song, Y., Wang, J., Ge, Y., & Xu, C. (2020). An optimal parameters-based geographical detector

- model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data. *GIScience & Remote Sensing*, 57(5), 593-610.
125. Stovin, V., Vesuviano, G., & Kasmin, H. (2012). The hydrological performance of a green roof test bed under UK climatic conditions. *Journal of hydrology*, 414, 148-161.
  126. *Stream Corridor Restoration: Principles, Processes, and Practices* (2001). Available at: [http://www.nrcs.usda.gov/Internet/FSE\\_DOCUMENTS/stelprdb1044574.pdf](http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1044574.pdf)
  127. Su, M., Zheng, Y., Hao, Y., Chen, Q., Chen, S., Chen, Z., & Xie, H. (2018). The influence of landscape pattern on the risk of urban water-logging and flood disaster. *Ecological Indicators*, 92, 133-140.
  128. Sun, S., Zhai, J., Li, Y., Huang, D., & Wang, G. (2020). Urban waterlogging risk assessment in well-developed region of Eastern China. *Physics and Chemistry of the Earth, Parts A/B/C*, 115, 102824.
  129. Sun, Z. (2014). Causal factors of local floods in Beijing central city. *Geo Res*, 33(9), 1668-1679.
  130. Tan, X., Wu, X., & Liu, B. (2021). Global changes in the spatial extents of precipitation extremes. *Environmental Research Letters*, 16(5), 054017.
  131. Tang, L., Ni, G., Liu, M., & Sun, T. (2011). Simulation on the effect of green roofs on the stormwater runoff and waterlogging release. *J China Hydrol*, 4, 20-4.
  132. Tang, X., Hong, H., Shu, Y., Tang, H., Li, J., & Liu, W. (2019). Urban waterlogging susceptibility assessment based on a PSO-SVM method using a novel repeatedly random sampling idea to select negative samples. *Journal of Hydrology*, 576, 583-595.
  133. Tang, X., Li, J., Liu, W., Yu, H., & Wang, F. (2021). A method to increase the number of positive samples for machine learning-based urban waterlogging susceptibility assessments. *Stochastic Environmental Research and Risk Assessment*, 1-18.
  134. Tang, X., Shu, Y., Lian, Y., Zhao, Y., & Fu, Y. (2018a). A spatial assessment of urban waterlogging risk based on a Weighted Naïve Bayes classifier. *Science of the Total Environment*, 630, 264–274.
  135. Tang, Z., Yi, S., Wang, C., & Xiao, Y., (2018b). Incorporating probabilistic approach into local multi-criteria decision analysis for flood susceptibility assessment. *Stochastic Environmental Research and Risk Assessment*, 32(3), 701–714.
  136. Tanouchi, H., Olsson, J., Lindström, G., Kawamura, A., & Amaguchi, H. (2019). Improving urban runoff in multi-basin hydrological simulation by the HYPE model using EEA urban atlas: A case study in the Sege River Basin, Sweden. *Hydrology*, 6(1), 28.

137. Teemusk, A., & Mander, Ü. (2007). Rainwater runoff quantity and quality performance from a greenroof: The effects of short-term events. *Ecological engineering*, 30(3), 271-277.
138. Tehrany, M. S., Kumar, L., & Shabani, F. (2019). A novel GIS-based ensemble technique for flood susceptibility mapping using evidential belief function and support vector machine: Brisbane, Australia. *PeerJ*, 7, e7653.
139. Tehrany, M. S., Pradhan, B., & Jebur, M. N. (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *Journal of Hydrology*, 504, 69–79.
140. Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N. (2015). Flood susceptibility assessment using gis-based support vector machine model with different kernel types. *Catena*, 125, 91-101.
141. Terstriep, M. L., & Stall, J. B. (1974). The Illinois urban drainage area simulator, ILLUDAS. Bulletin (Illinois State Water Survey) no. 58.
142. The State Council of the People's Republic of China (2021). Implementation Opinions on Strengthening Urban Waterlogging Control. Available at: [http://www.gov.cn/zhengce/2021-05/27/content\\_5612869.htm](http://www.gov.cn/zhengce/2021-05/27/content_5612869.htm)
143. Tran, D., Xu, D., Dang, V., & Alwah, A. (2020). Predicting urban waterlogging risks by regression models and internet open-data sources. *Water*, 12(3), 879.
144. Tsanis, I. K., & Boyle, S. (2001). A 2D hydrodynamic/pollutant transport GIS model. *Advances in Engineering Software*, 32(5), 353–361.
145. United Nations Sustainable Development Goals (2015). Available at: <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>
146. United Nations The World's Cities in 2016. United Nations Population Division (2017). Available at: [http://www.un.org/en/development/desa/population/publications/pdf/urbanization/the\\_worlds\\_cities\\_in\\_2016\\_data\\_booklet.pdf](http://www.un.org/en/development/desa/population/publications/pdf/urbanization/the_worlds_cities_in_2016_data_booklet.pdf).
147. United Nations World Urbanization Prospects (2018). Available at: <https://esa.un.org/unpd/wup/Download/>.
148. United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition. Available at: <https://esa.un.org/unpd/wup/>.
149. Wang, C., Du, S., Wen, J., Zhang, M., Gu, H., Shi, Y., & Xu, H. (2017). Analyzing explanatory factors

- of urban pluvial floods in Shanghai using geographically weighted regression. *Stochastic Environmental Research and Risk Assessment*, 31(7), 1777–1790.
150. Wang, D., Fang, C., Gao, B., Huang, J., & Yang, Q. (2011). Measurement and spatio-temporal distribution of urbanization development quality of urban agglomeration in China. *Chinese Geographical Science*, 21(6), 695-707.
  151. Wang, J., Gao, W., Xu, S., & Yu, L. (2012). Evaluation of the combined risk of sea level rise, land subsidence, and storm surges on the coastal areas of Shanghai, China. *Climatic Change*, 115(3–4), 537–558.
  152. Wang, J., Yi, S., Li, M., Wang, L., & Song, C. (2018). Effects of sea level rise, land subsidence, bathymetric change and typhoon tracks on storm flooding in the coastal areas of Shanghai. *Science of the total environment*, 621, 228-234.
  153. Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., & Bai, X. (2015). Flood hazard risk assessment model based on random forest. *Journal of Hydrology*, 527, 1130–1141.
  154. Wen, L., Weiping, C., & Chi, P. (2016). Modeling the effects of green infrastructure on storm water runoff reduction on community scale. *Acta Ecol. Sin*, 36(6).
  155. Winston, R. J., Dorsey, J. D., & Hunt, W. F. (2016). Quantifying volume reduction and peak flow mitigation for three bioretention cells in clay soils in northeast Ohio. *Science of the Total Environment*, 553, 83-95.
  156. Wu, J., & Zhang, P. (2017). The effect of urban landscape pattern on urban waterlogging. *Acta Geogr. Sin*, 72, 444-456.
  157. Wu, J., Sha, W., Zhang, P., & Wang, Z. (2020). The spatial non-stationary effect of urban landscape pattern on urban waterlogging: a case study of Shenzhen City. *Scientific reports*, 10(1), 1-15.
  158. Wu, J., Yang, R., & Song, J. (2018). Effectiveness of low-impact development for urban inundation risk mitigation under different scenarios: A case study in Shenzhen, China. *Natural Hazards and Earth System Sciences*, 18(9), 2525–2536.
  159. Xu H. Investigation on Green Roof Structure. Beijing: Beijing University of Forestry Press; 2007.
  160. Xu, L., Cui, S., Tang, J., Nguyen, M., Liu, J., & Zhao, Y. (2019). Assessing the adaptive capacity of urban form to climate stress: a case study on an urban heat island. *Environmental Research Letters*, 14(4), 044013.
  161. Yang, L., Zhang, L., Li, Y., & Wu, S. (2015). Water-related ecosystem services provided by urban

- green space: A case study in Yixing City (China). *Landscape and urban planning*, 136, 40-51.
162. Yang, P., Ren, G., & Yan, P. (2017). Evidence for a strong association of short-duration intense rainfall with urbanization in the Beijing urban area. *Journal of Climate*, 30(15), 5851-5870.
  163. Yang, X., You, X. Y., Ji, M., & Nima, C. (2013). Influence factors and prediction of stormwater runoff of urban green space in Tianjin, China: laboratory experiment and quantitative theory model. *Water science and technology*, 67(4), 869-876.
  164. Yao, L., Chen, L., Wei, W., & Sun, R. (2015). Potential reduction in urban runoff by green spaces in Beijing: A scenario analysis. *Urban Forestry & Urban Greening*, 14(2), 300-308.
  165. Yin, J., Yin, Z., Wang, J., & Xu, S. (2012). National assessment of coastal vulnerability to sea-level rise for the Chinese coast. *Journal of Coastal Conservation*, 16(1), 123-133.
  166. Yin, J., Yu, D., & Wilby, R. (2016). Modelling the impact of land subsidence on urban pluvial flooding: A case study of downtown Shanghai, China. *Science of the Total Environment*, 544, 744-753.
  167. Youssef, A. M., Pradhan, B., & Hassan, A. M. (2011). Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. *Environmental Earth Sciences*, 62(3), 611–623.
  168. Yu, H., Zhao, Y., Fu, Y., & Li, L. (2018). Spatiotemporal variance assessment of urban rainstorm waterlogging affected by impervious surface expansion: A case study of Guangzhou, China. *Sustainability (Switzerland)*, 10(10).
  169. Yu, Z., Fryd, O., Sun, R., Jørgensen, G., Yang, G., Özdil, N. C., & Vejre, H. (2021). Where and how to cool? An idealized urban thermal security pattern model. *Landscape Ecology*, 36(7), 2165-2174.
  170. Zeinali, V., Vafakhah, M., & Sadeghi, S. H. (2019). 'Impact of Urbanization on Temporal Distribution Pattern of Storm Runoff Coefficient. *Environmental monitoring and assessment*, 191(9), 1-16.
  171. Zhang, B. (2015). Flood Disaster of China 2014. *China Flood and Drought Management*, (1), 59–65.
  172. Zhang, B., Li, N., & Wang, S. (2015). Effect of urban green space changes on the role of rainwater runoff reduction in Beijing, China. *Landscape and Urban Planning*, 140, 8-16.
  173. Zhang, B., Xie, G., Zhang, C., & Zhang, J. (2012). The economic benefits of rainwater-runoff reduction by urban green spaces: a case study in Beijing, China. *Journal of environmental management*, 100, 65-71.
  174. Zhang, C., Lv, Y. P., Chen, Y., He, X. H., & Zhi, X. H. (2013). Inspirations of international technology and standard systems of drainage systems to China.

175. Zhang, G., He, B. J., Zhu, Z., & Dewancker, B. J. (2019). Impact of morphological characteristics of green roofs on pedestrian cooling in subtropical climates. *International journal of environmental research and public health*, 16(2), 179.
176. Zhang, H., Cheng, J., Wu, Z., Li, C., Qin, J., & Liu, T. (2018a). Effects of impervious surface on the spatial distribution of urban waterlogging risk spots at multiple scales in Guangzhou, South China. *Sustainability (Switzerland)*, 10(5).
177. Zhang, H., Wu, C., Chen, W., & Huang, G. (2017). Assessing the impact of climate change on the waterlogging risk in coastal cities: A case study of Guangzhou, South China. *Journal of Hydrometeorology*, 18(6), 1549-1562.
178. Zhang, Q., Wu, Z., & Tarolli, P. (2021a). Investigating the Role of Green Infrastructure on Urban WaterLogging: Evidence from Metropolitan Coastal Cities. *Remote Sensing*, 13(12), 2341.
179. Zhang, Q., Wu, Z., Guo, G., Zhang, H., & Tarolli, P. (2021b). Explicit the urban waterlogging spatial variation and its driving factors: The stepwise cluster analysis model and hierarchical partitioning analysis approach. *Science of The Total Environment*, 763, 143041.
180. Zhang, Q., Wu, Z., Zhang, H., Dalla Fontana, G., & Tarolli, P. (2020). Identifying dominant factors of waterlogging events in metropolitan coastal cities: The case study of Guangzhou, China. *Journal of Environmental Management*, 271, 110951.
181. Zhang, S., & Pan, B. (2014). An urban storm-inundation simulation method based on GIS. *Journal of Hydrology*, 517, 260-268.
182. Zhang, S., Wang, T., & Zhao, B. (2014). Calculation and visualization of flood inundation based on a topographic triangle network. *Journal of hydrology*, 509, 406-415.
183. Zhang, Y., Xia, J., Yu, J., Randall, M., Zhang, Y., Zhao, T., ... & Shao, Q. (2018b). Simulation and assessment of urbanization impacts on runoff metrics: insights from landuse changes. *Journal of hydrology*, 560, 247-258.
184. Zhao, G., Pang, B., Xu, Z., Yue, J., & Tu, T. (2018). Mapping flood susceptibility in mountainous areas on a national scale in China. *Science of the Total Environment*, 615, 1133-1142.
185. Zhao, Y., Xia, J., Xu, Z., Zou, L., Qiao, Y., & Li, P. (2021). Impact of Urban Expansion on Rain Island Effect in Jinan City, North China. *Remote Sensing*, 13(15), 2989.
186. Zhu, H., Yu, M., Zhu, J., Lu, H., & Cao, R. (2019). Simulation study on effect of permeable pavement on reducing flood risk of urban runoff. *International journal of transportation science and technology*,

8(4), 373-382.

187. Zhu, Y., Bi, H., Chang, Y., & Hai, X. (2015). Runoff Reduction Effect of Sunken Lawn under Different Designed Rainstorm Intensity in Beijing Area. *Bulletin of Soils and Water Conservation*, 35(2), 121-124.
188. Zimmer, C. A., Heathcote, I. W., Whiteley, H. R., & Schroter, H. (2007). Low-Impact-Development Practices for Stormwater: Implications for Urban Hydrology. *Canadian Water Resources Journal*, 32(3), 193–212.
189. Zope, P. E., Eldho, T. I., & Jothiprakash, V. (2016). Impacts of land use–land cover change and urbanization on flooding: A case study of Oshiwara River Basin in Mumbai, India. *Catena*, 145, 142-154.
190. Zoppou, C. (2001). Review of urban storm water models. *Environmental Modelling & Software*, 16(3), 195-231.