A Study of Land Development Patterns of Shinkansen Station Areas Based on Image Matching Algorithms

陳，思テイ

https://hdl.handle.net/2324/4784383

出版情報：九州大学, 2021, 博士（工学）, 課程博士
バージョン：
権利関係：

A Study of Land Development Patterns of Shinkansen Station Areas Based on Image Matching Algorithms

CHEN Siting

2021

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Foremost, I would like to acknowledge and give my warmest thanks to my supervisor, Prof. ZHAO Shichen, for his invaluable advice, continuous support, and patience during my study at Kyushu University. His immense knowledge and plentiful experience have encouraged me in my academic research and daily life.

Besides, I would like to thank the rest of my thesis committee: Prof. SAKAI Takeru, Prof. KUROSE Takefumi, and Prof. YOSHITAKE Tetsunobu, for their encouragement, insightful comments and suggestions to improve this dissertation.

I would like to thank the Institute of Decision Science for a Sustainable Society (IDS3), Kyushu University, who provided me with the scholarship during my 5-year graduate study (for the master's degree and doctorate). Their support has allowed me to focus on my study and research without worrying about the living expenses.

My sincere thanks also go to my fellow lab mates in Urban Planning Lab, Kyushu University: Dr. CHEN Qi, Dr. XU Feifan, Dr. DANG Sheng, REN Yujie, DU Mengge, NI Molin, SONG Jingying, ZHANG Yichuan, TSAI Chiayu, CHI Jiemeng, HIGUCHI Go, YAMAMOTO Ryuta, ONIZUKA Yuki, KANEFUJI Satoya, for the discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the past few years.

Specially, I would like to thank my family, my parents, CHEN Shiqiang and YANG Tinglian, without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my study.

At last, I want to express my gratefulness and a deep sense of appreciation to all those individuals and institutions whose invaluable contributions and support have helped me in writing up and finishing this humble work.

ABSTRACT

Faced with the unprecedented nationwide depopulation in Japan, a major concern with the recently extended Shinkansen has been the trade-off between the benefits of expanding the network to the remote regions against the huge investment. An appropriate land development planning around a Shinkansen station has been proved to have a significant positive impact on the economic development in the station area and the station city. This study aimed to systematically summarize the land development patterns around existing Shinkansen Stations, explicitly reveal the trend of the current land use situation and dynamic transformation situation of Shinkansen station areas, which will help better understand the current issues thus providing reference to urban planners and policymakers of cities whose Shinkansen station is under construction or currently planned.

To identify the land development patterns of the existing Shinkansen station areas, this study started with developing a land-use pattern comparison method based on image matching algorithms. In contrast to previous land-use pattern comparison approaches which are on the basis of certain aspects of quantifiable land-use pattern feature, the image matching approach treats land-use maps as images in which the feature value of each pixel represents the land-use category of the corresponding land-use cell, and then applies image smoothing as a feature extractor to integratively characterize the compositional and configurational features of a land-use map with a land-use pattern feature descriptor. The similarity between the land-use patterns can be calculated by the similarity of the land-use feature descriptors. In this way, the image matching approach allows us to identify both the compositional and configurational features of the land-use patterns, and consider not only the overall trend throughout the study area but also reflect the relationship between each individual land-use cell and its neighbors. Besides, it provides a quantitative measure for the similarity level between land-use patterns that can be converted to distances for further processing in the clustering. The algorithm of the proposed land-use pattern comparison method was explained and validated in Chapter 2, then it was applied to analyze the land development patterns of Shinkansen station areas.

Specifically, the Shinkansen station area was defined as the 1000-meter radius range around a Shinkansen station, and the 400-meter radius range around a Shinkansen station was identified as the core station area. For the 1000-meter radius range area, we employed the land-use pattern feature-based comparison method proposed in Chapter 2 to classify the land-use patterns of 91 existing Shinkansen station areas into 6 types, and identified the trend of the land-use patterns in station areas of cities with different population scales. Then we applied the proposed method to identify the changes of land-use patterns of Shinkansen station areas between 2009 and 2016, thereby, the land use transformation situation of Shinkansen station areas and the location trend of the changes were clarified.

As to the 400-meter core station area, we considered it as a spatial indifference zone and focused on the urban functional composition of land uses in these areas. In Chapter 5, based on 16 variables which reflect the development situation of residential, commercial, business and amenity facilities in the core station areas, we classified the urban functional composition situation of the core station areas around existing Shinkansen stations into 6 groups, and discussed the relationship between the functional composition features of a core station area and the location condition of the station. Besides, Chapter 5 also analyzed the functional transformation situation in the existing core Shinkansen station areas.

Accordingly, this study provided a comprehensive and systematic insight into the trend of land development situation and the dynamic land use transformation situation Shinkansen station areas.

Key words:

Shinkansen, HSR, Land Use, Land-use Transformation, Spatial Pattern Comparison, Image Similarity

Chapter 1

INTRODUCTION

1.1 Background

1.1.1 The Significant Impacts of The Shinkansen High-Speed Rail (HSR) Network on The Regional and Urban Development

The development of the Japanese High-Speed Rail (HSR) system, which is considered as one of the most significant technological breakthroughs in the second half of the 20th century, has brought tremendous influence on Japanese society and national economies. [1] The first Shinkansen HSR line between Tokyo and Osaka, the Tokaido Shinkansen, was officially opened for commercial operation in 1964. It has shortened the travel time between these two cities from 8 hours on a conventional train to less than 4 hours. The operation of the Tokaido Shinkansen has not only successfully increased the capacity of the conventional Tokaido Line, but has also increased the business communication between Tokyo and Osaka, encouraged tourism along the line, and promoted regional development. Since then, the broad socio-economic impact of the Shinkansen HSR system on regional and urban development has attracted consistent worldwide attention.[2]

As a high-speed, mass-transit railway system, the Shinkansen has not only played an important role in intercity transportation in Japan but has also been considered as a key to correct the regional imbalance. Researchers suggested that the Tokaido Shinkansen has significantly accelerated the growth of the major cities on the line and has also speeded up the transformation of the corridor into a huge megalopolis. [3] However, simultaneously with the development of the Tokaido Shinkansen, the rapid economic growth from the latter half of the 1950s was accompanied by the excessive concentration of population and industries in the major cities, and population loss and economic activity decreasing in rural and outlying areas. The success of the Tokaido Shinkansen has made the Japanese government believe that a high-speed link such as the Shinkansen HSR system can strengthen interregional communication and can effectively disperse the economic activity or population from the developed regions to intermediate regions along the line, thus releasing the social strain and promote the regional balancing. [3,4] Therefore, the Japanese government approved the *Nationwide Shinkansen Railway Construction Act* in 1970 to expand the Shinkansen network nationwide. Since then, the Shinkansen HSR network has been continuously expanding. Up to now, the Shinkansen network has expanded to 92 stations[*1)], connecting the southernmost city of the Kyushu island - Kagoshima to the Northernmost island - Hokkaido. Besides, the Hokkaido Shinkansen, Hokuriku Shinkansen, Nagasaki section of Kyushu Shinkansen, and the Linear Chuo Shinkansen are under construction, the Shinkansen network is going to connect more cities. (Fig. 1-1)

However, the expansion of the Shinkansen network has not always been progressing smoothly. Fig.1-2 summarized the number of newly opened Shinkansen stations at different times. As the post-war economic recovery has subsided, as the national economy and personal incomes have grown, and as economic concerns have been replaced by the pursuit of quality of life, the extension of the Shinkansen network encountered resistance in some regions and complaints about noise on existing lines. [3,5] In some areas, the planning of new stations was frustrated by protracted land acquisition issues, the expansion of the Shinkansen network did not make progress smoothly as planned. Besides, with the expansion of the Shinkansen network, the sparsely populated areas did not have enough travel demands, the unprofitable railways showed up in rural areas aggravated the financial burden of the national railways (JNR) company and finally led to the privatization of JNR at the end of the 1980s. The national planning of the nationwide Shinkansen network was set aside for a time. [6]

In addition, with the Shinkansen network extending to rural and outlying regions, researchers observed that the influences of the Shinkansen construction on regional development have not always been consistently positive. Much empirical research implicated that the construction of a high-speed rail system does not naturally cause the development of a region. It was pointed out that the mythical belief that the construction of Shinkansen can solve both transport and regional development problems is derived from its routing. The miraculous success of the incipient Shinkansen system is closely related to a non-negligible fact that it was apparently designed to serve the most important region in the country, the Tokaido region, which is a region that was experiencing or expected to experience growth. [7] In other words, the construction of Shinkansen did not trigger the growth in its serving region, the growth was dispersed from existing centers to sub-centers, and indigenous growth in the region was attracted to the sub-centers around locations with access to the Shinkansen. [6] However, the growth did not always occur that way, the impact of an HSR system on different cities varies dramatically across different cities. Researchers specified that the redistribution effect generally leads to growth in the existing centers at the expense of the periphery regions. [8–10] It has been observed in the worldwide practices of HSR development that the HSR may strengthen the locational advantages of large central cities, and extend the accessibility gap between large central cities and intermediate cities, thus accelerating the agglomeration of population and economic factors around large central cities. [11–13] Researchers highlighted the risk of spatial polarization imposed by the HSR [14], argued that the introduction of an HSR system does not naturally bring positive benefits to the less-developed regions, and emphasized the importance of careful planning and policy intervention to effect necessary ancillary investment. [15,16]

Nevertheless, the Shinkansen HSR system is still expected to play a catalyzing role in driving the spatial and urban transformation process. In the "Grand Design of National Spatial towards 2050" released by The Ministry of Land, Infrastructure, Transport, and Tourism (MLIT) in 2014, the development of regional high-speed connections such as the Shinkansen network is considered to be essential to the formation of a compact regional network which is the key to achieve agglomeration economies and support the effective services delivery and regional innovation in a population decreasing society. [17] The expansion of the Shinkansen network has been slowly but continuously progressed by the privatized regional Japan Railway (JR) companies. Although the future development of the Shinkansen network may not progress exactly as the 1970's basic plan depicted, undoubtedly the Shinkansen network is going to continuously play an important role in Japan's intercity transportation and regional development. Therefore, it is necessary to consistently pay attention to the impact of such widespread implementation of a development program, so that policymakers can make an appropriate adjustment to the future development plan.

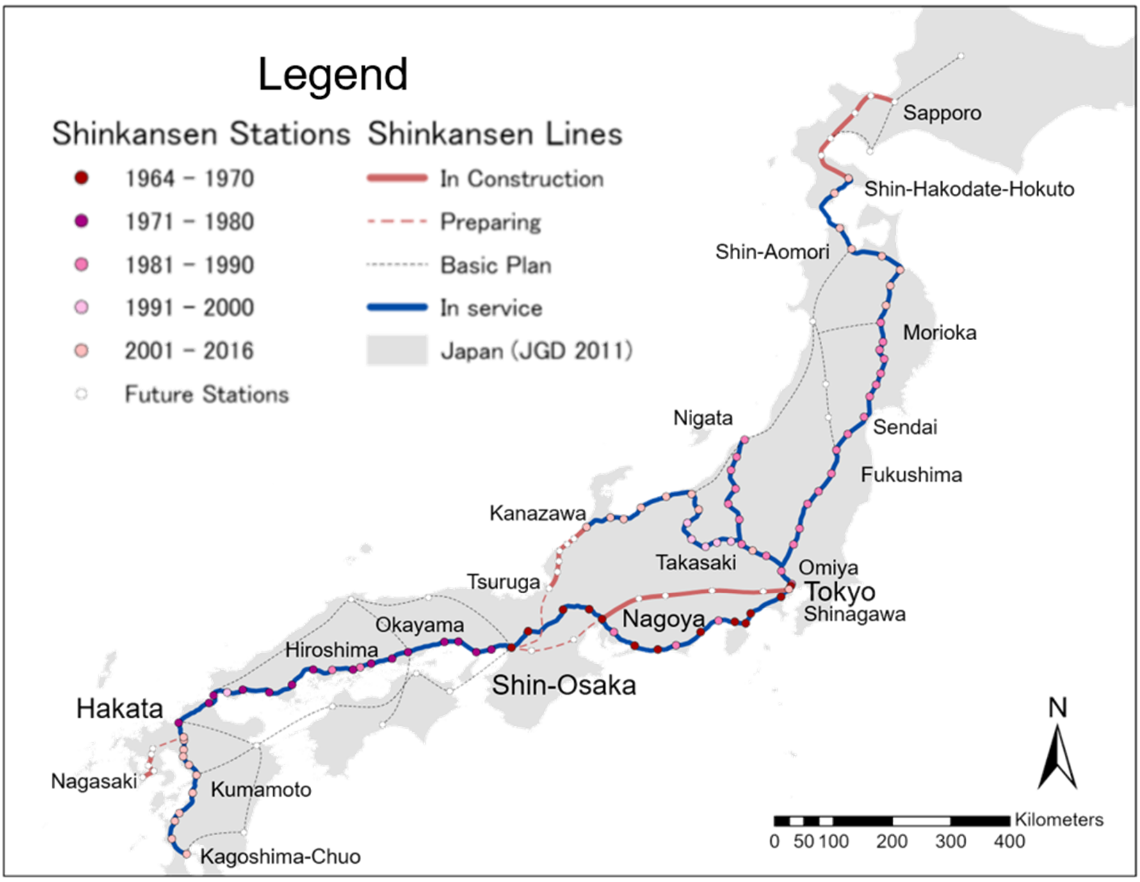


Fig. 1-1 The Progress of the Construction of Nationwide Shinkansen Network System According to the Nationwide Shinkansen Railway Construction Act

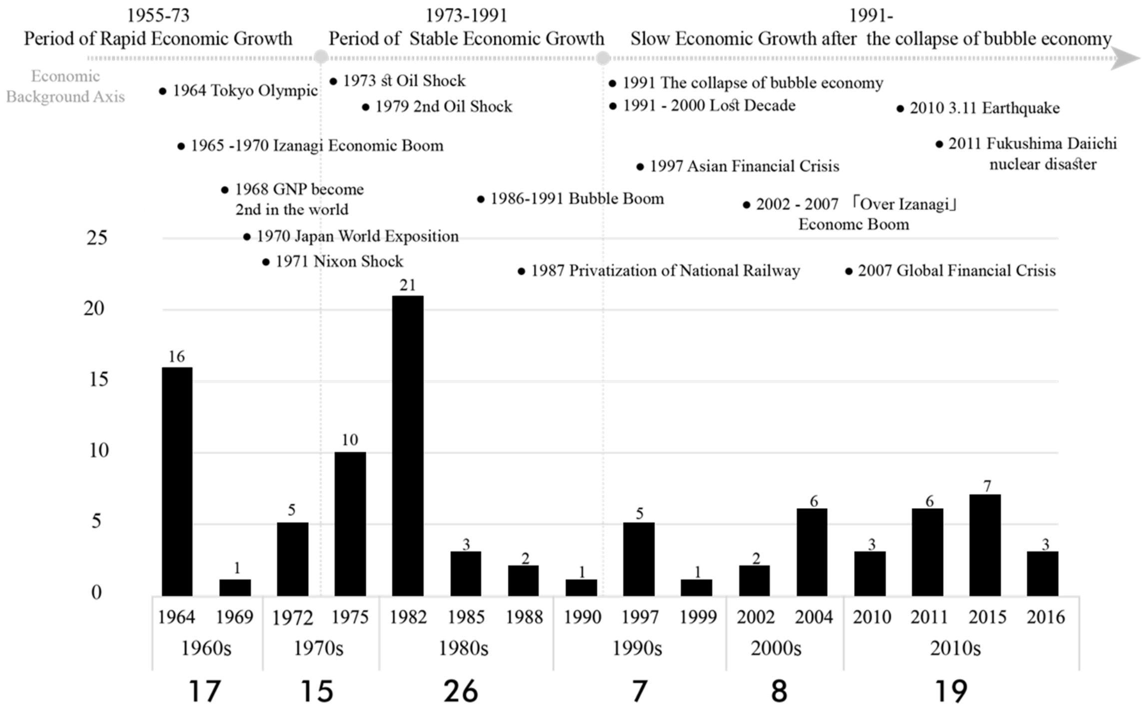


Fig. 1-2 The Development of The Shinkansen System and the National Economic Climate in Different Times

1.1.2 The Concern of Urban Development around The Shinkansen in The Context of Nationwide Depopulation

The report of the 2015 National Population and Household Census declared that Japan's total population has been decreasing since 2010. Although aging and depopulation have been a widespread concern in Japan for many years, the total population has kept increasing since 1920. However, according to the 2015 National Census, among all 1965 administrative regions, 1490 had population declined between the year 2010 and 2015, only 21.9% of the total 1965 administrative regions were able to keep positive population growth, which signified that depopulation is no longer a problem for several regions in Japan but an inevitable new reality for urban planners and decision-makers in nationwide. [18]

Besides, the majority of the Shinkansen development experience was accumulated by the studies of the early Shinkansen lines constructed in the post-war economic growth period. Researchers have specified that the success of the early Shinkansen was considered to be inextricably linked with the increasing transport demands and the increasing population. [2] It is noteworthy that the initial motivation for constructing the Shinkansen was no more than increasing the carrying capacity of the Tokaido Line. The region along the conventional Tokaido Line was the most important region in Japan. Although it only takes 16% of Japan's total area, it contains three of the four great industrial zones of Japan and seven of ten large cities with a population of one million or more. This area concentrated about 34% of the total population and 60% of Japan's industrial products, making the Tokaido Line the most important conventional railway line in the country. With the continuous annual increment of traffic volume on the Tokaido Line, it was approaching its upper limit in 1952. [19] That was when the plan of the Shinkansen construction came into being. The extension of the Shinkansen from Osaka to Fukuoka was under a similar condition. This region contains about 20% of Japan's population and 33% of the national industrial output, and the determination of construction of Sanyo Shinkansen was when traffic volume on the conventional trunk was about to reach saturation point. Namely, the construction of the early Shinkansen has been to meet the traffic demand, qualitatively and quantitatively. In contrast, the planning of later Shinkansen lines did not follow the traffic demand, instead, they were planned to drive the development of the less-developed areas and reduce the regional imbalance, although later studies have demonstrated the implementation of Shinkansen may have failed to meet this expectation. [20]

Indubitably, improvement in accessibility improves competitiveness and brings development opportunities to less-developed regions. However, the Shinkansen HSR connection also created a convenient express for the central megapolis to siphon resources from these regions. The past decades have witnessed the development of the Shinkansen HSR system accelerated the urbanization and regional development in the densely populated areas with the cities along the Tokaido Line and Sanyo Line as representatives, however, in the rural and outlying regions, few successful experiences have been reported insofar. Nevertheless, people did not lose faith in the positive regional impact of Shinkansen and have never stopped seeking a more appropriate development mode to make full use of the development opportunities brought by the Shinkansen construction. Faced with the unprecedented nationwide depopulation and the limited government financial budget, one major concern with the recent extended Shinkansen has been the trade-off between the benefits of expanding the network to the remote regions against the huge investment. [21] The urban development around Shinkansen in the current depopulation society aroused much attention and reconsideration.

1.1.3 The Land Development around Shinkansen Stations is Crucial for The Eventual Successful Operation of An HSR Station

The Shinkansen stations usually serve as the transportation hubs in the cities, and the construction of Shinkansen HSR is considered as a catalyst for urban restructuring of the station areas. Accompanied with the construction of Shinkansen, there are always urban development plans carried out to improve the quality of the living environment, attract investments, led population aggregation and commercial prosperity. Empirical evidence through academic work emphasizes that good station-area planning is a very important prerequisite for the eventual successful operation of an HSR station. [13,22]

In the field of urban planning, researchers have been focused on the land use around the Shinkansen stations. Urban land carries all sorts of human activities, the study of land use helps us integratively understand the characteristics of the population, industrial, and economic activities distributed around the stations. Researchers use the land-use pattern to describe the characteristic of the layout or arrangement of the land uses in a typical urban area, and devise land-use models to abstracts the complex city system thus explaining the underlying development rules of land use arrangement in the cities. Early land-use models, such as Burgess and Park’s concentric zone model and Park’s concentric zone model, were based on a large number of empirical observations and described the general pattern of a typical city area. Although these models were descriptive and did not directly reveal the causes for a given geometrical form, they provided an intuitive summary of the urban structure of observed areas through visual assessment of the land use pattern of the observed areas.[23] On the contrary, modern land-use models provide more accurate and quantitative approaches to explore the underlying causes for any particular land-use patterns and make conditional predictions of land-use transformation [24,25], nevertheless, these models are often localized and customized for one specific city region, thus they might not be adaptable to other regions.

The study of the land use around Shinkansen station is essentially a sub-part of the study of urban land use. Influenced by the Shinkansen, the urban areas around the Shinkansen stations have similar properties of transportation hubs and city gateways. However, the majority of previous studies of land use surround Shinkansen stations focused on one or some specific stations[26–31], and most were successful cases along the early constructed Shinkansen lines (Tokaido and Sanyo). As a result, our understanding of the land-use development around Shinkansen stations has been limited in the situation of several case studies, it remains unclear how much could these case studies represent the general land development situation of Shinkansen station areas. Besides, as the socio-economic background has been different from when the Tokaido and Sanyo Shinkansen were constructed, there exists an urgent necessity for studies that summarize the land development pattern around Shinkansen stations in the current context of nationwide depopulation.

1.2 Statement of the Research Problem

Appropriate land development planning around a Shinkansen station will significantly positively impact the economic development in the station area and the station city. However, our current understanding of the land development of Shinkansen station areas is insufficient. The successful experience at several regional center cities can hardly be adapted to the less developed rural and outlying regions, there is an extreme lack of appropriate reference for the small cities. In concern of the unprecedented nationwide depopulation, it will be necessary to conduct studies that explicitly reveal the land-use situation around Shinkansen stations in both the developed and less developed cities, thus helping make clear the current issues, providing references to urban planners and policymakers of cities whose Shinkansen station is under construction or currently planned.

This study aims to clarify the land development patterns of Shinkansen station areas in the context of nationwide depopulation. In this research, *land development* has two aspects of meaning: the static land-use composition and configuration **situation** and the dynamic **process** of land-use transformation. Thus, the *land development pattern* in this research refers to the regularity of the land-use composition and distribution situation and the land-use transformation process. Although the land development around each Shinkansen station has unique characteristics, as the transportation hubs and city gateway areas, our hypothesis is that station areas with similar development conditions must share some common land development features. The goal is to review the land use development around all 92 existing Shinkansen stations in the context of nationwide depopulation, detect the common features and classify them as different land development patterns. In other words, this research is going to specify the following questions:

1. **What** and **where** are different land use categories located in the Shinkansen station areas?
2. **How** have land uses around Shinkansen stations transformed?

1.3 Literature Review

The research problem of this study mainly relates to two topics: the study of the Shinkansen stations area development and the study of land use patterns. This part develops a theoretical framework of this study by reviewing the academic works in the field of relevant research topics, thus addressing the research gaps in the previous literature, and specifying the research goal of this study.

1.3.1 The Shinkansen Station Area Development

The study of land development of Shinkansen station areas is a sub-branch of the research field of land development around railway stations. The station areas' superiority in accessibility distinct them from ordinary urban areas, researchers have put consistent effort to reveal the characteristics of the land use distribution and transformation in the station areas, and considerable achievements have been accumulated. Previous literature relevant to land development around stations mainly concentrated on the following 3 perspectives: ①. the function arrangement and spatial structure of land use, which aimed to clarify the characteristic of development intensity and industrial distribution around stations; ②. The economic effects around the stations, including the change of land price, rental, employment, and other economic activities agglomerated in the station areas; ③. The quality of the living environment, which discusses the transport organization, community development, urban vitality of the station areas.

Moreover, previous literature can be divided into theoretical studies and empirical studies. Theoretical studies explained the underlying mechanism of the development activities around stations based on the socio-economic theories and built up a model of appropriate land development. Namely, the theoretical studies aimed to illustrate what the land development around stations **should** be. One of the most far-reaching and profound theories is the transit-oriented development (TOD) concept, which advocates maximizing the amount of residential, business, and leisure space within walking distance of public transit [32] and promotes a symbiotic relationship between dense, compact urban form and public transport use. However, TOD is a comprehensive idea for the urban development around transit stations, there is no single, all-encompassing definition that represents the TOD concept in its many forms. Besides, there are many similar concepts such as the walkable city, the compact city and so on, they all share the following common elements: ①. Higher intensity of development around a public transit station; ②. Mixed-land use and compactness; ③. Pedestrian- and cycle-friendly environments; ④. Development that is conducive to transit riding; ⑤. Stations as community hubs. [33,34] These concepts usually serve as the guiding ideology for the urban planning and design strategies of a station area. However, there is no single standard for a unified range for a station area. Another important concept is the “three-development-zone” theoretical model for the development of surrounding areas of transportation facilities proposed by Schütz E. (1998) [35], which distinguished the 5-10 minutes accessible area, the 10-15 minutes accessible area, and the peripheral areas around an HSR station as the primary, secondary, and tertiary development zones respectively, and clarified the characteristics of urban development in these three different zones. [16] Following this framework, many researchers defined the station area as a circle structure centered around a transit station and usually range from 400 - 1000 meters. Besides, in Japan, along with the long history of HSR development, there is also an extensive history and ample experience of station area development. Abundant previous studies discussed the station-adjacent developments (駅前開発), and usually, they use station catchment area (駅勢圏) to name the area whose development might be influenced by a railway station. Also, there is no unified standard for the range of a station catchment area, it was usually defined as a range of 600 - 2000 meters around a station and varied with specific research problems. Generally, the major purpose of the theoretical is to give an illustration of the principles to develop a good station area, however, there is always a distance between theory and practices, and this is where the empirical studies shed light on.

Empirical studies focused on the real development situation around one or several specific stations and usually were usually to compare the real situation with some station area development theories, thus testifying the adequacy of the development strategies under a particular development theory. [26–31] However, their findings can hardly be generalized and adapted to other station areas. Several studies gave the conclusions after summarizing the situation around multiple stations. Moon C., Sato S., and Tonuma K. (1997) analyzed urban improvement projects around 56 Shinkansen stations and clarified the time serial and spatial characteristics of the development process around Shinkansen stations. [36] This study was a qualitative analysis, and it summarized the urban development patterns around Shinkansen stations based on the planning documents from municipalities, and it was conducted in the 1990s, which was about 30 years ago, the current situation might be different. Matsumoto E., Ubaura M. (2013) analyzed the land development situation around 17 Shinkansen stations in the suburbs and clarified the problems in the process of urbanization through case studies. [37] However, the current land development situation of other non-suburb stations remains unclear.

Therefore, this study contributes to the studies on the land development around Shinkansen station by summarizing the current land development situation around existing Shinkansen stations and provide a comprehensive insight into the land development situation of Shinkansen station areas in the context of current nationwide depopulation.

1.3.2 Land Use Pattern Comparison

The study of land development of Shinkansen station areas is also a sub-branch of the study field of land use patterns. *Land use pattern* is the layout or arrangement of the uses of the land. Clarifying the land-use patterns of Shinkansen station areas is essentially a land-use pattern comparison task. It involves identifying the land use pattern features and clustering the station areas with similar land-use patterns together thus finding some land-use patten types.

Land use pattern comparison is a fundamental approach to understand urban development in urban studies: comparing the land uses across regions helps to identify the areal variation of land development in the observed regions, whereas comparing land uses over time helps to detect changes brought by underlying urban development processes. A large body of urban studies has involved land use pattern comparison, among which the most traditional and commonly used method is visual comparison, which is to present maps side-by-side as a tool for visual comparison and identify the similarities and differences based on the human visual system’s ability to recognize shapes and patterns.[38] The visual comparison of maps has been productive of urban models and theories, for instance, Burgess and Park's (1925) concentric zone model was built on large observations and comparisons of American cities. However, limited by human visual perception, the visual comparison approach is incapable of detecting the fragmental or imperceptible differences between maps, and cannot provide an accurate measure of the degree of similarity.

Quantitative methods of land use pattern comparison are generally based on quantifiable measurements of land use patterns. Quantitative measures of land use patterns fall into two categories: compositional indices and configurational indices.[39] The composition refers to the non-spatial properties of a land-use map (e.g. the type, number, and statistical properties of land uses); whereas, the configuration refers to the spatial distribution characteristics of land-use units, whose quantification is complicated and problematic due to its ambiguous definition. [40,41] Considerable efforts have been put into developing methods to quantify the configurational characteristic of a land-use pattern, and there are now literally hundreds of quantitative measures of spatial pattern that have been proposed to quantify the configuration from the aspects of shape, form, density, intensity, clustering, centrality, and dispersion.[42–44] However, none of them can individually summarize the overall compositional and configurational characteristics of land uses.

Besides, the quantitative measures of a spatial pattern can be broken down into *global level* measures and *local level* measures. Global spatial pattern indices measure the spatial variation throughout a study area, while local indices focus on the spatial variation surround each individual.[45] At the global level, only the overall trends of the spatial variation in the study area will be interpreted, for example, the center, shape, and orientation of the elements' distribution, or whether the elements are randomly distributed or spatially clustered in the study area. At the local level, the relation between the occurrence and attribute of an element with its neighbors will be clarified. Many indices have been proposed to measure the spatial relationships between land-use units at different levels, such as aggregation index, contagion index, clumpiness index, and so on [39,46], although there is no standard for an appropriate observation scale when applying these indices to implement the spatial analysis. Researchers have referred this uncertainty problem given rise to different geographical division schemes to the modifiable areal unit problem (MAUP). [47–49] As yet there still lacks sound solution to this problem and the study area for current spatial analysis is usually determined based on experience.

In addition, limited by the inability of finding an adequate descriptor that can summarize the overall configurational characteristic of a spatial pattern, the quantitative comparison of spatial patterns insofar was mostly based on a set of spatial pattern indices, as a result, the comparison result can be largely influenced by the selection of indices. Besides, many of the spatial pattern indices themselves have compound meanings, a similarity measure built on multifarious indices can lead to opacity and even misinterpretation. [50–52] Moreover, many researchers argued that both local vs global and compositional vs configurational characteristics of a spatial pattern are independent and interactive, and they highlighted the importance of simultaneously considering and distinguishing them. [41,51–54]

To address the existing problems mentioned above, researchers have tried to seek solutions from other fields. Particularly, in the field of computer vision, computer scientists have also been intensively developing methods for image matching and comparison, which is analogous to land-use map comparison.[38] With the rapid development of computer technology, significant progress has been made in image matching techniques. Computer scientists have developed approaches to extract image features and represent them with an image feature descriptor for further processing such as matching or clustering. [55] In recent years, applications of these methods in urban and geographic contexts have been increasing, however, most were image semantic retrieve tasks, such as land use discrimination and object detection from satellite imagery [56–59], or spatial qualities analysis based on landscape elements identification from streetscape imagery [60–62]. However, few attempts have been made to extract the geospatial pattern information in the land-use maps with image processing techniques.

Therefore, to clarify the land development patterns of Shinkansen station areas, firstly, it is necessary to build an effective land-use pattern comparison approach that allowing us to ①. distinguish both the compositional and configurational characteristics between spatial patterns, ②. consider not only the overall trend throughout the study area but also reflect the relationship between each individual and its neighbors and ③. Provide a quantitative similarity measure that can be converted to distances for further processing in the clustering.

1.4 Aims and Objects of the Study

On the basis of the literature review, to answer the research problems of this study, firstly we specify the definition of a Shinkansen station area in this research. As Fig.1-3 suggests, it basically follows the “three-development-zone” concept, with integrative consideration of the average walking speed of normal adults and aged people [63], the core station area is defined as a 5-minute walk area around a Shinkansen station, and the station area is defined as a 1-kilometer radius range area around a Shinkansen station.

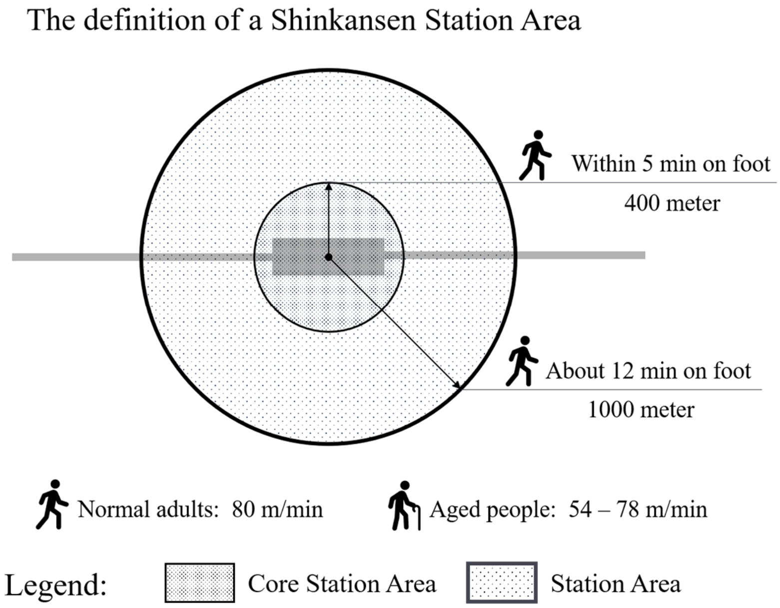


Fig. 1-3 The Definition of a Station Area

Besides, when analyzing the land development patterns, this study considers the 400-meter zone as a spatial indifference zone. In other words, when discussing the land-use pattern of the 1-km zone, this study will analyze the characteristics of the composition and spatial relationships between land uses, while for the 400-meter zone, this study will only focus on the functional composition.

Specifically, this study aimed to review the land development situation around 92 existing Shinkansen station areas and summarize their development patterns. The focus is the current land development state and dynamic transformation situation during 2009 and 2016. *2) The research goal is to identify the land development patterns around Shinkansen stations.

The achievement of the main purpose is based on the following specific objectives:

1). To develop a land-use pattern comparison method that integratively quantifies the similarity level of both land use composition and the spatial relationships based on image matching algorithms.

2). To study the land use situation of the existing 91*3) Shinkansen station areas, classify them into different types and identify their land-use pattern features based on the proposed land-use pattern comparison method.

3). To analyze the land use transformation situation around Shinkansen stations during 2009 and 2016, and to identify the location trend of the transformations.

4). To analyze the urban functional composition of land development in 92 station core areas, identify functional composition features and make clear the transformation relationships between urban functions.

5) To discuss the relation between the land development pattern of a station area and the location condition of the station city.

1.5 Dissertation Structure

The dissertation outline can be divided into six chapters as Fig. 1-4 suggests. The first chapter is the introduction which primarily gives the aim of the present research work. The rest of the dissertation is arranged as follows:

Chapter 2 proposed and validated the land-use pattern comparison method based on the image matching algorithms.

Chapter 3 applied the method developed in chapter 2 to analyze the land-use patterns in the 1km radius range around the existing 91*3) Shinkansen stations.

Chapter 4 applied the method developed in chapter 2 to compare the land-use patterns of 91 Shinkansen station areas between 2009 and 2016 and clarified the land use transformation situation of Shinkansen station areas.

Chapter 5 analyzed the functional composition and transformation situation of land uses in the 400-meter radius range of the 92 Shinkansen station core areas.

Chapter 6 summarized the whole dissertation.

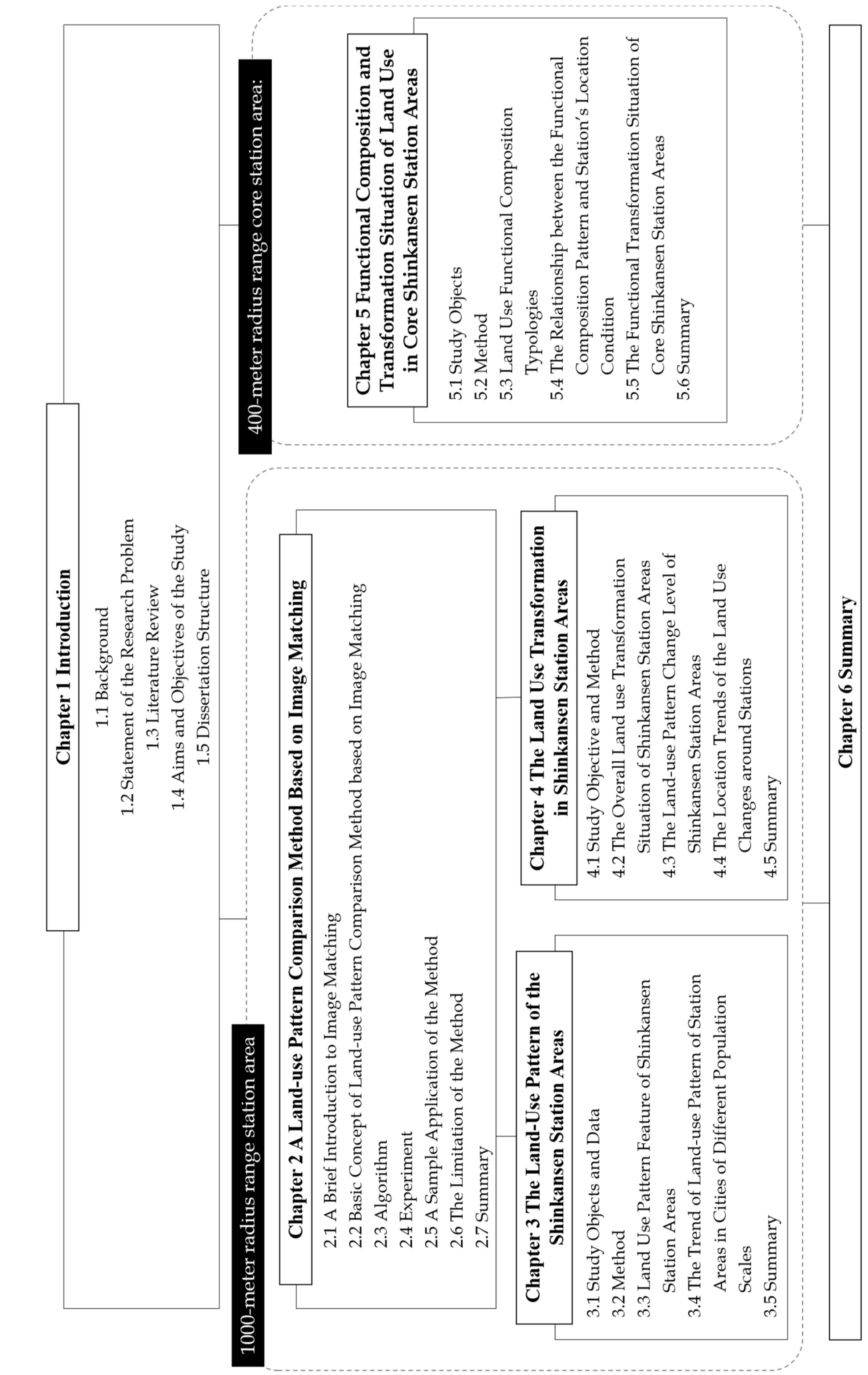


Fig. 1-4 Dissertation Structure

NOTES

*1). It is generally accepted that the maximum running speed of a high-speed rail should be in excess of 200Km/h, thus the Akita Shinkansen and Yamagata Shinkansen were not discussed in this research for their maximum running speed is 130Km/h.

*2). The latest urban land-use subdivision mesh data published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism was the land-use data of 2016.

*3) Okutsugaru-Imabetsu station locates in rural area, thus it is not included in the urban land-use subdivision mesh data published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism.

Chapter 2

A Land-use Pattern Comparison Method Based on Image Matching

Following the research framework of the whole dissertation, this chapter aimed to develop a land-use pattern comparison method based on image matching. According to the review of the previous literatures, we have highlighted the shortage of the existing land-use pattern comparison approaches, and clarified the goal is to develop a land-use pattern similarity measure that quantifies the similarity level of both composition and spatial relationships between land use units simultaneously, thus preparing for further processing of clustering.

This begins with a brief introduction to image matching (**Section 2.1**), to give a glimpse of the key techniques that can be employed to develop the land-use pattern comparison method, and help develop the framework to convert the land-use pattern comparison problem to an image matching task. (**Section 2.2**) Then we explain the specific algorithm (**Section 2.3**). To show the effectiveness of the developed method, we conducted experiments on the simulated random maps generated by a landscape generator. (**Section 2.4**) Then, we show the applicability of the proposed approach through a sample application to compare the land-use patterns of 6 sample Shinkansen station areas. (**Section 2.5**) Finally, we discuss the parameters in the algorithm, specify the principles to set the parameters, and clarify the limitations of the proposed approach. (**Section 2.6**)

2.1 A Brief Introduction to Image Matching

Image matching is an important concept in computer vision and object recognition. Fundamentally, the image matching tasks are to identify then correspond the same or similar structure/content from two or more images. [64] For the human visual system, it is not difficult to integrate visual information including color, distribution, orientation, style, and objects contained in the images thus identifying the similarity and dissimilarity between images, which is however not as easy for computers. Computers see an image as arrays of feature values (such as brightness or color intensity value), and each value provides visual information depicted in one pixel that constitutes the image. [65] Image comparison in computer vision is not an easy task because the definition of similar content is ambiguous.

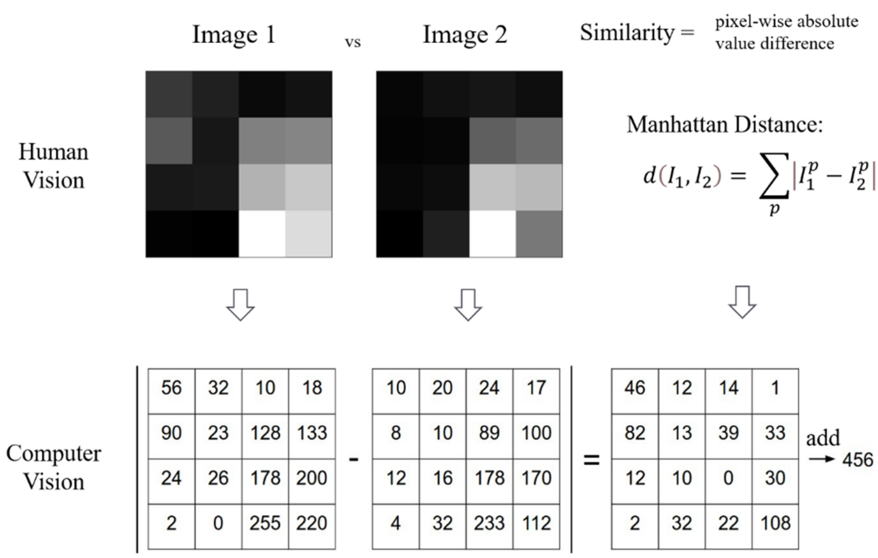


Fig. 2-1 Early-Stage Image Matching: Image Similarity Based on Pixel-wise Absolute Value Difference

In the early stage, the image matching is based on the pixel-by-pixel comparison. As Fig 2-1 suggests, computers view two images as grids of numbers between [0,255]. In Fig.2-1, the (dis)similarity is measured by the Manhattan Distance between 2 sample images, whose result is 456. Besides, there exist tons of (dis)similarity measures, such as the equally widely used Euclidean distance and correlation, they share the same strategy of measuring the similarity by the sum of the pixel-wise differences. This is the simplest and most straightforward strategy for image matching, also known as the *area-based method* and has been employed in early object recognition works, which matches an object in an image with a template and a sliding window. [66] However, this method treats each pixel independently, thus it is incapable of interpreting the semantic information represented by the combination of pixels. Therefore, this approach is not widely used because it is basically sensitive to all the image transformations like scaling, transition, rotation, reflection, shear, partial occlusion, and color difference.

In recent decades, with the rapid development of computer technology, particularly the progress in machine learning, computer scientists have developed many accurate, robust, and efficient approaches for image matching. The mainstream for image matching has been the *feature-based method*. [64] The fundamental idea of feature-based approaches is that instead of directly measuring the pixel-wise value differences, they first characterize and numerically represent the image features with some feature descriptors (or feature vector) and then match images based on the extracted features. Namely, feature-based image matching can generally be broken down into two intrinsic stages: *feature extraction* and *feature matching*.

Feature extraction has been a major study part of computer vision. The main goal of feature extraction is to extract the most relevant information from an image to obtain an appropriate and robust descriptor. The definition of the appropriate image features is application-dependent, and in most tasks that involve image matching, the image features are learned through a deep learning process. [55] However, feature extraction in deep learning is essentially a data-driven process, although it is now a very active area in computer vision [67–69], people's understanding of it is still limited. Moreover, the task of land-use pattern comparison in this research is not a learning task. Therefore, here we only focus on the traditional feature extraction approaches, which derive image features through the devised algorithms using the information present in the image itself. In contrast to the learned features in deep learning, the manually designed image features are also known as *handcrafted features*.

The handcrafted features are devised to characterize an image and make it more invariant to issues brought by some kinds of image transformations such as occlusions and variations in scale and illuminations. [70] The image features can be divided into *global features* and *local features*. Global features describe the visual content of the entire image by a single vector, whereas local features aim to detect the interest points (IPs) in an image and describe them by a set of vectors. [71] Commonly used global features include color, texture, shape, which are also known as three major low-level visual features [72], while widely used local features include SIFT, SURF, BRIEF, and ORB. [73] Besides, tons of techniques have recently been proposed to unify global and local features to integrate low-level features like color, texture, and shape into more robust higher-level features for better performance in the image matching tasks. [74] Seeking better-handcrafted features has been another hot topic in computer vision because handcrafted features are manually designed, they are more controllable, and they can usually help optimize the performance of the learned features in deep learning.

However, it is beyond the scope of this research to introduce the whole universe of image feature extraction. The purpose of this part is only to give a glimpse of the major challenges and key techniques of image matching tasks, thus helping convert the land-use pattern comparison task to an image matching task. According to the introduction above, it is not difficult to find analogical connections between land-use pattern comparison and image matching, particularly in the part of unifying the global and local features. In association with the compositional and configurational properties of land-use patterns, here we introduce *histogram* as the representation of global feature and *image smoothing* as the method to extract spatial relationships.

1). Histogram: Color histogram is one of the most commonly used image representations. An image is essentially a two-dimensional matrix whose elements are natural numbers corresponding to the brightness or color intensity values. The histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. As Fig.2-2 suggest, (a) is a grayscale image with the size of 273×187, (b) is the histogram of image (a). The height of each histogram bin represents the pixel count in the original image with the corresponding color intensity value.

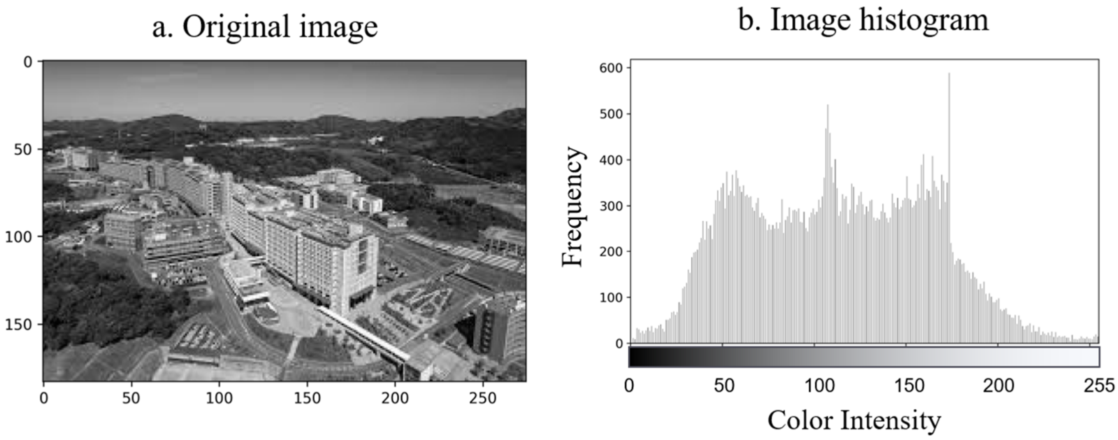


Fig. 2-2 Image Representation with Color Histograms

For color images, an RGB image for example, is composed of three two-dimensional arrays, with the elements in each array representing the color intensity value of red, green, and blue, respectively. Generally, each array of the intensity values is called a color channel. An RGB image has three color channels, therefore, it has three histograms with each individually representing the frequency of the color intensity in each channel.

The histogram provides the global information about the color intensity composition in an image; however, it discards the information of the spatial relationship between pixels. Fig.2-2 for example, from the histogram (b) we can read the composition proportions of different color intensity in the original image, but we cannot trace back the locations of different intensity values in the original image (a). Therefore, we can say the histogram is to an image what areal proportions are to a land-use pattern. We still need a method to extract the spatial relationships between pixels or land-use units, and this is where the image smoothing comes into play.

2). Image Smoothing: Image smoothing is a local pre-processing method to suppress noise in image data by some form of averaging of brightness values in some neighborhood. It is achieved by convolution operations. Fig.2-3 illustrates the smoothing process for a single pixel, the value in the output image pixel $f(i,j)$ is calculated as a weighted sum of the feature values (usually color intensity or brightness) in a local neighborhood K of the pixel value $g(i,j)$ in the input image. The contribution of the pixels in the neighborhood K is weighted by coefficients w, and the array of the coefficients can be called as a *convolution kernel*. In Fig.2-3, the source cell $g(i,j) = 0$ is smoothed by a 3×3 mean kernel, namely, the contributions of neighbor pixels are equally weighted, and as a result, the output value becomes $f(i,j) = 226.7$.

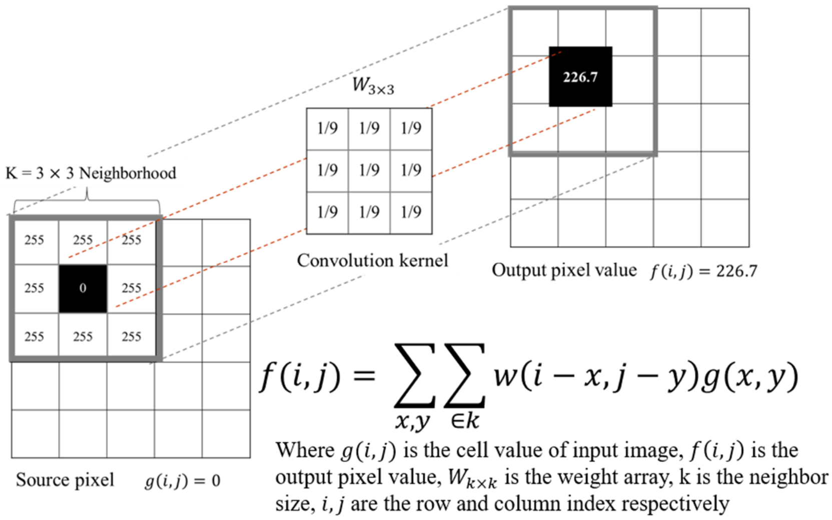


Fig. 2-3 Smoothing Operation of a Single Pixel

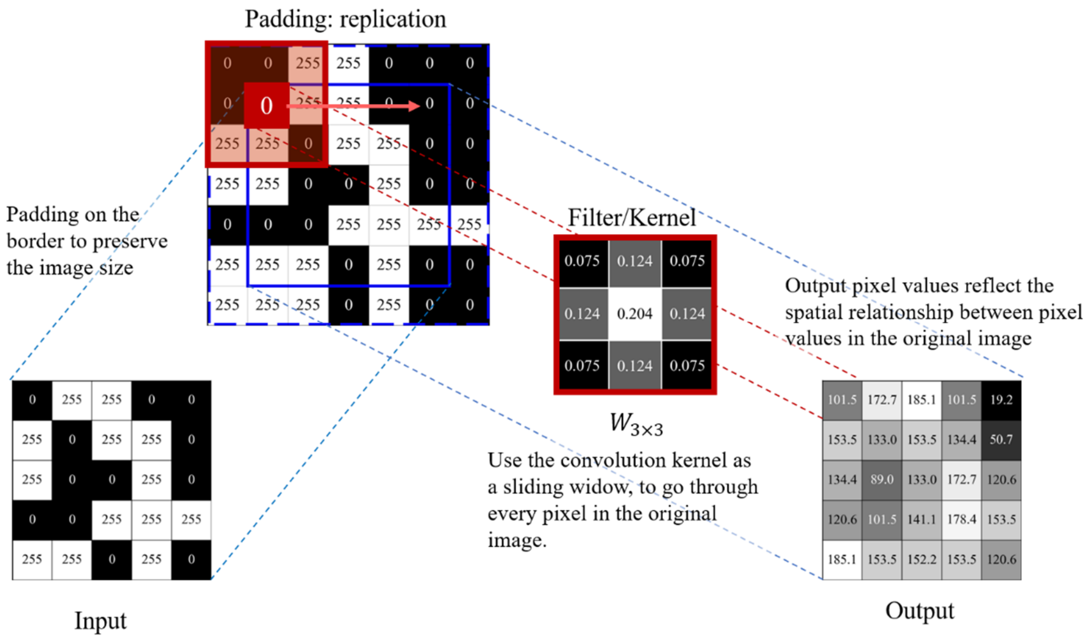


Fig. 2-4 The Image Smoothing Process for an Image

For the whole image, as shown in Fig.2-4, we can consider using the convolutional kernel as a sliding window to go through every pixel in the original image. Every pixel and its neighbor pixels are multiplied by the corresponding coefficient in the kernel, and the output pixel value equals a linear combination of itself and its neighbor pixel values. Besides, when it comes to the edge pixels, we copied the pixel values on the border to preserve the shape of the image. *1) In this way, the output pixel value no longer represents its own brightness, it reflects the information of its neighbor pixel values in the original image. Besides, as Fig.2-4 suggests, the coefficients in the convolution kernel are not necessarily equal. The coefficients control how much the information of a pixel will be reflected in the output values, it can be customized to extract the needed information and reduce the undesired information in image data. Therefore, the convolution kernel is also called a *filter*. In some sense, the devise of a feature extractor is in essence to tune the coefficients in the filter.

Accordingly, it becomes clear how to use the histogram to represent the global feature of the image pixel values composition, and also how to use image smoothing to obtain the information in each pixel’s neighborhood and represent it in the output feature value. This sheds some light on how to extract both the compositional and configurational features and integrate both local and global features of a land-use pattern. The next section of this chapter will introduce the image matching approach to the land-use pattern comparison. Based on the image smoothing and histogram comparison algorithm, we will devise an apt land-use pattern extractor to numerically represent the land-use pattern features with a pattern feature descriptor, then matching the land-use patterns based on the extracted features

2.2 Basic Concept of Land-use Pattern Comparison Method based on Image Matching

A land-use map is composed of units that represent the land-use attribute; this research only discusses the categorical land-use map in which each square cell represents one land-use category. For a categorical land-use map, the land-use pattern includes two aspects of features: the composition proportions of land use categories and the spatial relationships between land uses. Furthermore, the spatial relationship refers to the aggregation level of each land-use category and the location relation between different land-use categories. In this research, two maps are considered to have similar land-use patterns only when the land uses constitute them have both similar composition proportion and spatial relationships. Therefore, the goal is to extract land-use pattern features that appropriately reflect both the compositional and spatial information in a land-use map.

Inspired from the image matching algorithms, a land-use map can be viewed as an image composed of feature values representing the land-use categories. Here we call the value representing the land-use categories as *land-use feature values*. The main idea of our approach is to replace the land-use feature value of every land-use cell with a feature value that represents the land-use composition and distribution situation of its surrounding area through image smoothing operation, thereby the overall land-use pattern features can be represented by the distribution of the output feature values. To be clear, we call the land-use composition and distribution situation in a land-use cell's neighborhood as the *neighborhood pattern* and the distribution of the output feature values as the *land-use pattern descriptor* (pattern descriptor for short). In essence, the pattern descriptor is a list of numbers, however, in contrast to the ordinary image descriptors which are usually composed of every pixel's feature value, the land-use pattern descriptor consists of the counts of different output feature values after smoothing, which can also be represented as the distribution or histogram of the output feature values. Thus, the land-use pattern similarity can be computed based on the pattern descriptors.

Fig. 2-5 is a simple example that illustrates the basic concept of our approach. As (a) suggests, map-a and map-b are binary maps consisting of 10×10 cells. Assuming that black (B) and white (W) cells represent two categories of land-use, both map-(a) and map-(b) have composition proportions of B: W = 7: 3. If representing the black cells and white cells with the color intensity values of 0 and 255, respectively*2), their composition proportions can be represented by the histograms of the cell values. According to the histogram correlation, the similarity between map-(a) and map-(b) is 1, indicating they are the same land-use pattern. However, the spatial relationships of black and white cells in the two maps are clearly different: the white cells are aggregated in map-(a) while randomly distributed in map-(b). Following our approach, we replaced each cell's value with the mean of cells in its 3 × 3 neighborhood. Accordingly, the pattern descriptors of the two maps are represented as the histograms in (b). Many cell values of map-(b) were averaged by their neighbors to a value between 0 and 255 since they are dispersed, whereas many cell values of map-(a) remained or towards the initial value of 0 or 255 because they are more aggregated. As a result, the similarity became 0.38, suggesting two maps are different patterns, which is intuitively more acceptable.

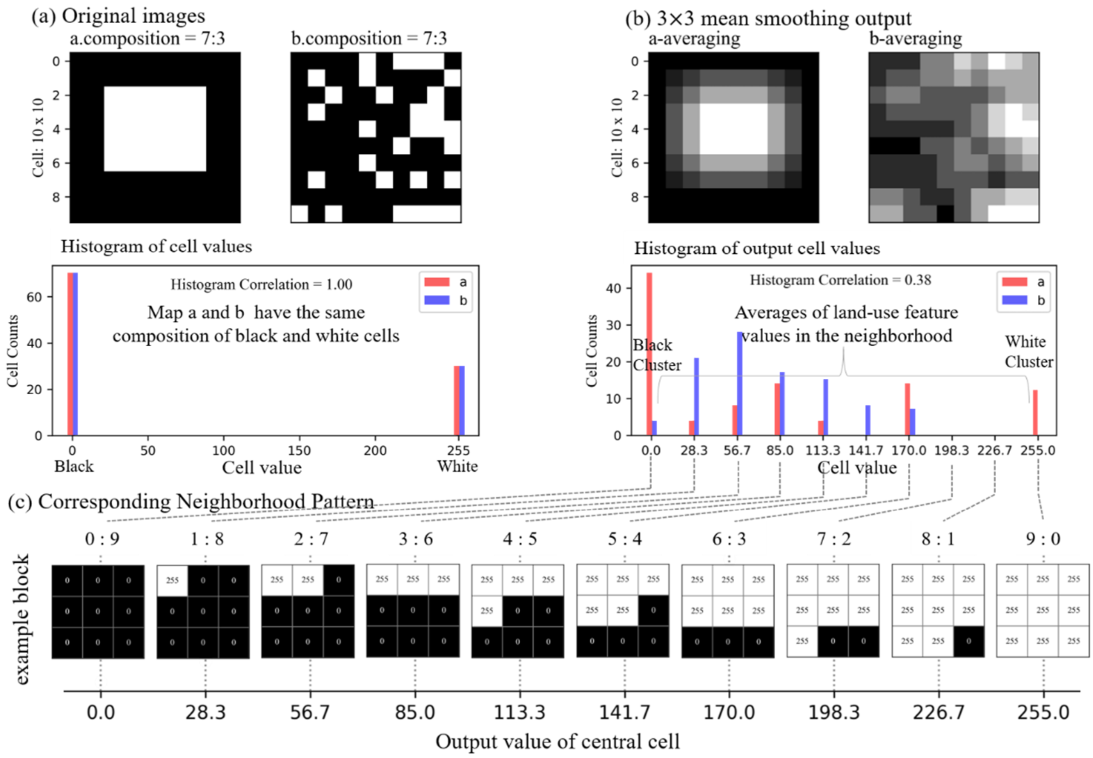


Fig. 2-5 Illustration for the Basic Concepts of the Proposed Method

Fig.2-5 also illustrates how the distribution of output feature values represents the land-use pattern features. When employing a 3×3 mean matrix in the image smoothing process, the corresponding relationship between the output feature value and the land-use combination in a cell's 3×3 neighborhood is shown in Fig.2-5 (c). Namely, if a black cell is surrounded by 8 black neighbors, its output feature value remains 0. Similarly, if a white cell is surrounded by 8 white neighbors, its output feature value remains 255. While the composition proportion of the black and white cells is B: W = 8: 1 in a cell's 3×3 neighborhood, the output feature value of the central cell is 28.3, and when B: W = 7: 2, the output feature value of the central cell will become 56.7, and so on. In this way, each output feature value represents one combination situation of land uses in the neighborhood.

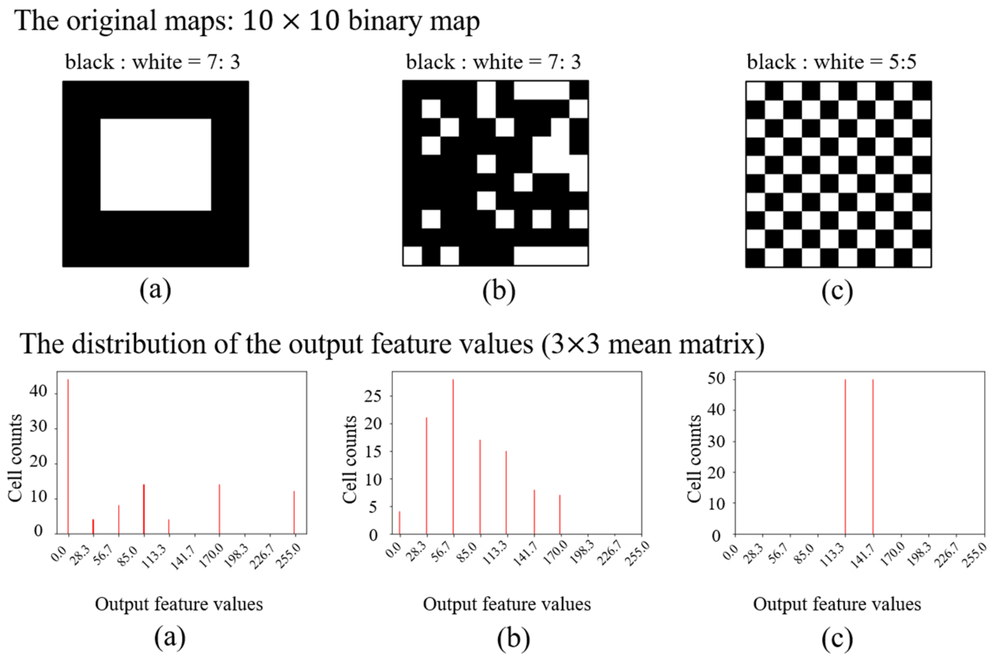


Fig. 2-6 Illustration for the Land-use Pattern Feature Descriptors

To illustrate the land-use pattern features represented by the histogram of the output feature values after smoothing, Fig.2-6 displays three sample binary maps-(a), (b), (c) and the distribution of their output feature values after smoothing with a 3×3 mean matrix. After smoothing, the output feature value of each cell corresponds to a land-use combination situation in Fig.2-5 (c). The height of each histogram bar represents the cell counts of the corresponding output feature values. Therefore, in the histograms of output feature values, the cell counts of the initial land-use feature values (0,255) represent the number of land-use cells surrounded by cells of the same categories as themselves. Hence, a land-use category is more aggregated, it will have more cells remain the initial land-use feature values. In Fig.2-6 for example, comparing (a) and (b), although they have the same composition proportions, the histogram of (a) has more cells remaining 0 and 255 for the black and white cells in the original image (a) are more aggregated. Moreover, in histogram (a), the cell count for 0 is greater than 255, reflecting that the black cells take larger areal occupancy than the white cells in the original image.

Besides, each output feature value corresponds to one land-use combination situation in the defined neighborhood (as Fig.2-5 (c) suggests), thus the number of different output feature values reflected the spatial variations of location relation between land uses. The image (c) in Fig.2-6 for example, the black and white are uniformly distributed, in the 3×3 neighborhood, the number of the black and white cells are always B: W = 4: 5 or 5: 4, as a result, after smoothing, the output cell values only have 2 different values. In contrast, in the original images (a) and (b), the location relation between black and white cells varies in different areas, the output feature values spread to more different values.

Accordingly, it becomes clear how the land-use pattern descriptor composed of the cell counts of different output feature values represent the composition and spatial relationship information and integrate the global and local features of the original land-use map. The key is to devise a filter letting every different output feature value corresponding to a unique neighborhood pattern, in this way, the land-use pattern features can be represented by the counts of different output feature values, i.e., the land-use pattern descriptor. Then we can compare the land-use patterns based on the pattern descriptors and quantify the similarity level by computing the distance or correlation between the pattern descriptors.

2.3 General Framework

2.3.1 Algorithms

The proposed method is composed of 3 major steps, as shown in Fig.2-7, the first step is to represent a categorical land-use map with a numeric array of certain initial feature values X; then to replace the feature value of each cell $x(i,j)$ with a weighted averaging of its neighborhood $F(i,j)$ through image smoothing operation $G_w(X)$; finally is to compute the similarity based on the distributions of the output feature values $d(H1, H2)$.

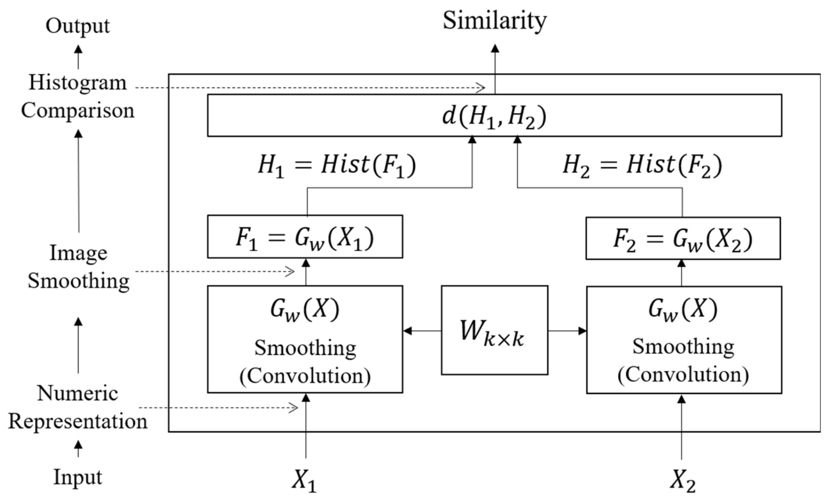


Fig. 2-7 General Framework of the Proposed Method

1) Numeric Representation of Land Use Categories: The crux of our method is that the output feature value of a land-use cell after smoothing corresponds to its neighborhood pattern: the feature value of a land-use cell will remain unchanged if it is the same land-use category with all its neighbor cells; otherwise, its feature value will be shifted to another value to label its neighborhood pattern. Therefore, the absolute value of the initial numeric representation of a land-use category is meaningless. However, to ensure output feature value varies when the neighborhood pattern varies, the principle to initialize the land-use feature values is to avoid that any land-use feature values being a linear combination of other land-use feature values. In this study, we introduced the Euler's Constant and set an exponential relationship between land-use feature values, i.e., for a map with n categories of land uses ($n \geq 2$), the land-use feature values a_i are set as:

$$a_i = \begin{cases} \frac{z}{e^{\frac{i}{\varphi}}} & (i = 0,1,2,\dots,n-2) \\ 0 & (i = n-1) \end{cases} \qquad \text{Equation (2.1)}$$

where z is the range to set all land-use feature values, φ is a parameter to control the interval between land-use feature values, i is the index for each land-use category. The order of the land-uses when setting the feature values is unimportant.

In this chapter, the land-use feature values were set to values between [0,255], i.e., $z = 255$, so that the feature values of land-use cells correspond to the color intensity value in the gray color space, which is easier for visualization. Moreover, φ is the parameter to ensure the balance of the intervals between feature values, thus avoiding the situation where there are huge intervals between the minority of feature values while most feature values squeeze in a small interval. In this paper, we set $\varphi = 2$. The determination of z and φ will be further discussed in section 2.6.

2) Image Smoothing Operation: The image smoothing operation is used to convert the land-use feature value to a linear combination of its neighborhood, i.e., the neighborhood pattern, thus representing the land use combination situation in the neighborhood with the output feature values. As introduced in section 1, a value in the output image pixel $F(i,j)$ is calculated as a weighted averaging of feature values (usually color intensity of brightness) in a local neighborhood K of the pixel value $g(i,j)$ in the input image X. The contribution of the pixels in the neighborhood K is weighted by coefficients w. Thus, the image smoothing operation $G_w(X)$ can be represented as:

$$F(i,j) = G_w\big(X(i,j)\big) = \sum\sum_{(x,y)\in K} w(i-x, j-y)g(x,y) \qquad \text{Equation (2.2)}$$

In concern of the geospatial interpretation of land-use maps, in this study, we call the kernel a *weight matrix* and the kernel size as *neighborhood size,* respectively.

The values of weight matrix w and the neighborhood size k define the neighborhood patterns, thus determining the composition elements of the land-use pattern descriptor, which will be further explained in sections 2.3.2 and 2.3.3.

3) Histogram Similarity: After the image smoothing, the land-use pattern descriptor can be represented by the distribution of output feature values. Therefore, the similarity of land-use patterns can be measured by the similarity of the distribution of output feature values i.e., histogram similarity. In this study, the histogram similarity is defined as the correlation between two histograms, which can be represented as:

$$d(H_1, H_2) = \frac{\sum_I (H_1(I) - \overline{H}_1)(H_2(I) - \overline{H}_2)}{\sqrt{\sum_I (H_1(I) - \overline{H}_1)^2 \sum_I (H_2(I) - \overline{H}_2)^2}} \qquad \text{Equation (2.3)}$$

where $\overline{H}_k = \frac{1}{N}\sum_J H_k(J)$. N is the total number of histogram bins. The value of N will be further discussed in section 2.6. In the experiments of section 2.4, $N = 256$.

With histogram correlation, the similarity of two land-use patterns is evaluated by a value between $[-1,1]$. The correlation value closer to 1 suggests greater similarity, whereas the correlation value closer to 0 indicates a greater difference between the land-use patterns, and a value closer to -1 indicates an inverse relationship between land-use patterns.

2.3.2 Define Location Relation with Weight Matrix

In the image smoothing process, every neighborhood pattern is represented by a linear combination of the land-use feature values in the defined neighborhood. The values in the weight matrix determine the contribution rates of neighbor cells. By assigning different weights to different positions, a land-use feature value will have a different contribution rate when its position differs. In this way, even when the land-use cells in the neighborhood have the same composition, the output feature value will vary if their distribution situation varies, thus the image smoothing will distinguish them as different neighborhood patterns. Notably, the numeric value of weights does not represent the real spatial influences between land-use cells, it only labels the location relations between land-use cells in the neighborhood. Accordingly, the weight matrix defines discrimination standards for the spatial variation in the neighborhood. In practice, the devise of the weight matrix depends on how we want to distinguish the location relations of land-use cells in the neighborhood.

Generally, there are two ways to consider the location relationships between the land-use cells in the neighborhood: one is to consider the neighborhood as a spatial indifference zone, in this case, neighbors within the specified distance weight equally (Fig.2-8-a); another is to distinguish neighbors in different positions, i.e., the weights of neighbors vary with different location relation. In this paper, we discuss two situations: distinguishing neighbors' positions by their distances to the central land-use cell (Fig.2-8-b) and distinguishing neighbors' positions with different corner and edge contiguity relations (Fig.2-8-c).

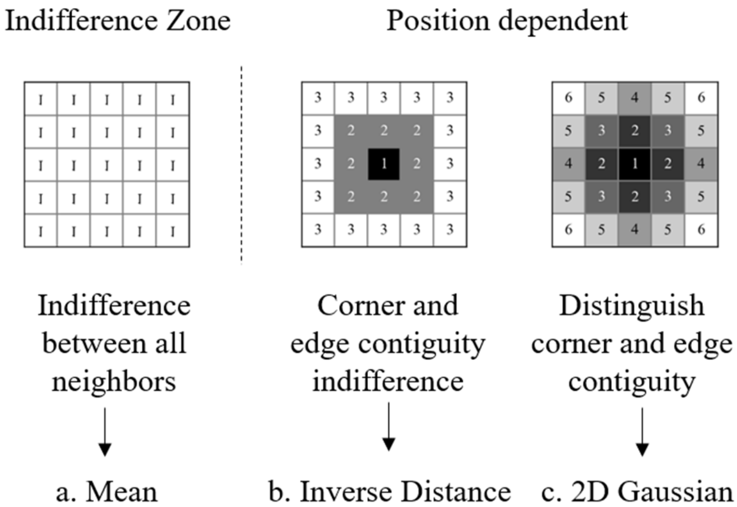


Fig. 2-8 Location Relations of Neighbor Cells Defined by Different Wight Matrix

1) Mean Weight Matrix: In the case of considering the neighborhood as a spatial indifference zone (Fig.2-8-a), the neighborhood pattern is determined by only the land-use composition of neighbor cells, and the distribution situation of land uses in the neighborhood is considered indifferent. In this case, all the land-use cells in the neighborhood are given the same weight in the smoothing operation, thus the weight matrix for a $k \times k$ neighborhood is:

$$w(i,j) = \frac{1}{k \times k} \qquad \text{Equation (2.4)}$$

where k is the size of the neighborhood, and k should be odd, (i,j) is the location coordinates of the cell in the neighborhood.

2) Inverse Distance Weight Matrix: When the locations of neighbor cells are distinguished by their distances to the central land-use cell (Fig.2.8-b), the weights for land-use cells vary with their distance to the central cell. In this paper, we use a negative exponential function of the distance to the center cell to determine the weight for each land-use cell. Be noted that we introduced the Euler number as a component of the exponent to avoid the situation where a larger weight is an integral multiple of a certain smaller weight, thus ensuring different land-use distribution situation corresponds to a different output value. The relation between the weight of a land-use cell and its distance to the central land-use cell is represented as:

$$w(i,j) = \alpha \times d^{-\frac{e}{\mu}} \qquad \text{Equation (2.5)}$$

where μ controls the gradient between weights. To ensure that even the land-use cells in the periphery of the neighborhood have a significant contribution rate to the output feature values, in the experiments of section 2.4, we set $\mu = 2$. d can be represented as the larger number between the row distance or column distance between the neighbor cell (i,j) and the center cell coordinate $(\frac{k-1}{2}, \frac{k-1}{2})$, therefore, the inverse distance weight matrix can be represented as:

$$w(i,j) = \alpha \times (\max(\{\left|i - \frac{k-1}{2}\right|, \left|j - \frac{k-1}{2}\right|\}))^{-\frac{e}{2}} \qquad \text{Equation (2.6)}$$

where k is the size of the neighborhood, and k should be odd, (i,j) is the location coordinates of a cell in the neighborhood, α is the scale factor to ensure $\sum_i \sum_j w(i,j) = 1$.

3) 2D Gaussian Weight Matrix: In more complex cases where the locations of land-use cells in the neighborhood are to be distinguished by the different corner and edge contiguity relations as Fig.3-c suggests, the 2D Gaussian equation can be employed to determine weights for the neighbor cells, i.e.:

$$w(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{(2\sigma^2)}} \qquad \text{Equation (2.7)}$$

where is σ the standard deviation of the distribution, (x,y) is the location coordinates of a point in the 2d-plane.

In an image, the 2d plane is composed of square cells, thus the location coordinate (x,y) is represented by a cell's row and column index (i,j), specifically: $x = i - \frac{k-1}{2}$, $y = j - \frac{k-1}{2}$, thus the weight for a land-use cell in the neighborhood can be computed by the following equation:

$$w(i,j) = \alpha \times e^{\frac{-((i-\frac{k-1}{2})^2+(j-\frac{k-1}{2})^2)}{(2\sigma^2)}} \qquad \text{Equation (2.8)}$$

Where k is the size of the neighborhood, and k should be odd, (i, j) is the location coordinates of a cell, σ is the Gaussian standard deviation (in experiments of section 2.4, $\sigma = 1$), and α is the scale factor chosen so that $\sum_i \sum_j w_{ij} = 1$.

2.3.3 Define Land-use Clusters with Neighborhood Size

In the image smoothing operation, the initial land-use feature value of a cell is replaced with a weighted averaging of its neighborhood. In this process, a land-use cell's feature value remains the original only if it has the same land-use property with all its neighbor cells. Therefore, if the size of a land-use patch in the map is equal to or greater than the defined neighborhood size, the central cell's feature value will remain the original, and the greater the land-use patch's size is, the more land-use cells remain the original feature value after smoothing, indicating the land use is more aggregated; on the contrary, when the size of a land-use patch is smaller than the neighborhood size, the feature value of the central cell will be shifted, suggesting the land use is more fragmented or dispersed. Therefore, the neighborhood size defines the minimum land-use cluster size, which influences the discrimination of the aggregation level of land uses.

As a whole, the neighborhood pattern is defined by the arrangement of n categories of land use in the neighborhood, therefore, the number of land-use categories, the location relationship between neighbor cells (the weight matrix), together with the neighborhood size determines the pattern descriptor thus influencing the similarity evaluation for land-use patterns. However, there is no standard weight matrix and neighborhood size, we need to customize according to specific applications. In the next section, the experiments intuitively illustrate how different combination of weight matrix and neighborhood size define the land-use pattern descriptor.

2.4 Experiments

To examine the influence of parameters and the effectiveness of the proposed method, we conducted two experiments with simulated land-use maps: one for binary land-use maps, another for multi-categorical land-use maps. The test maps were created by the neutral landscape models (NLM), which can generate maps with varied compositional and configurational characteristics without considering any specific spatial process. [75] Specifically, we employed the NLMpy neutral landscape generator [76] to create categorical maps with specified compositional and configurational parameters.

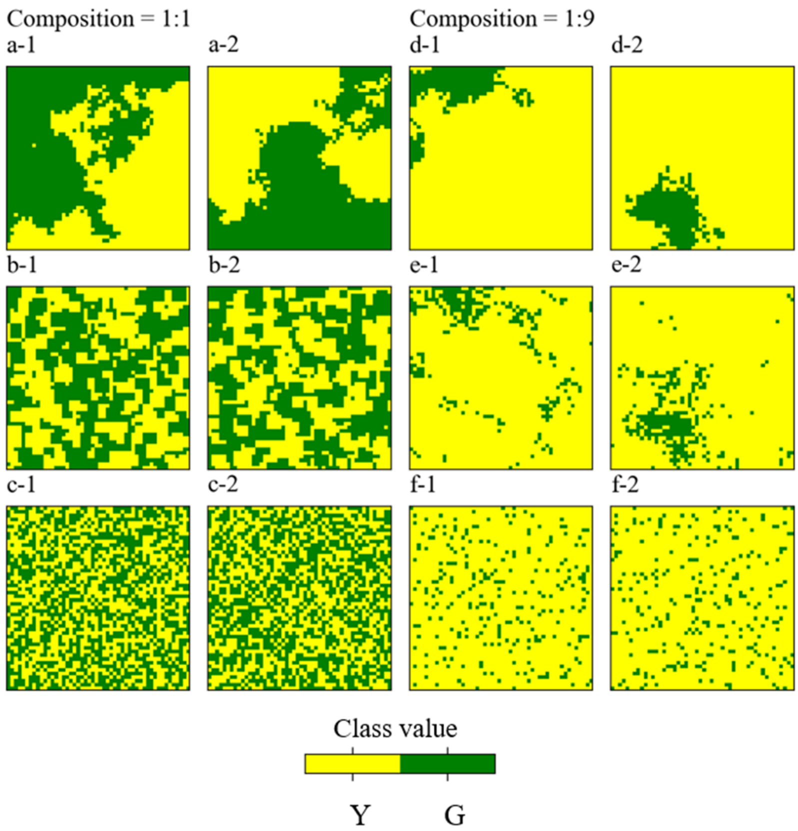
For each experiment, we generated 6 groups of random binary maps with different land-use pattern features, and each group contains 2 maps that were generated with exactly the same compositional and configuration parameters. The expectation is that the proposed method identifies the land-use pattern of maps in the same group and distinguishes the land-use pattern of maps in different groups.

Besides, in order to show the influence of weight matrix and neighborhood size on the land-use feature pattern extraction, the test maps were smoothed with different combinations of weight matrices and neighborhood sizes. The extracted land-use pattern features will be represented by the histograms of the output feature values. Then the land-use pattern similarities were computed based on the histograms, to examine whether the results are consistent with the experimental presuppositions.

2.4.1 Experiment of Binary Land-use Map Comparison

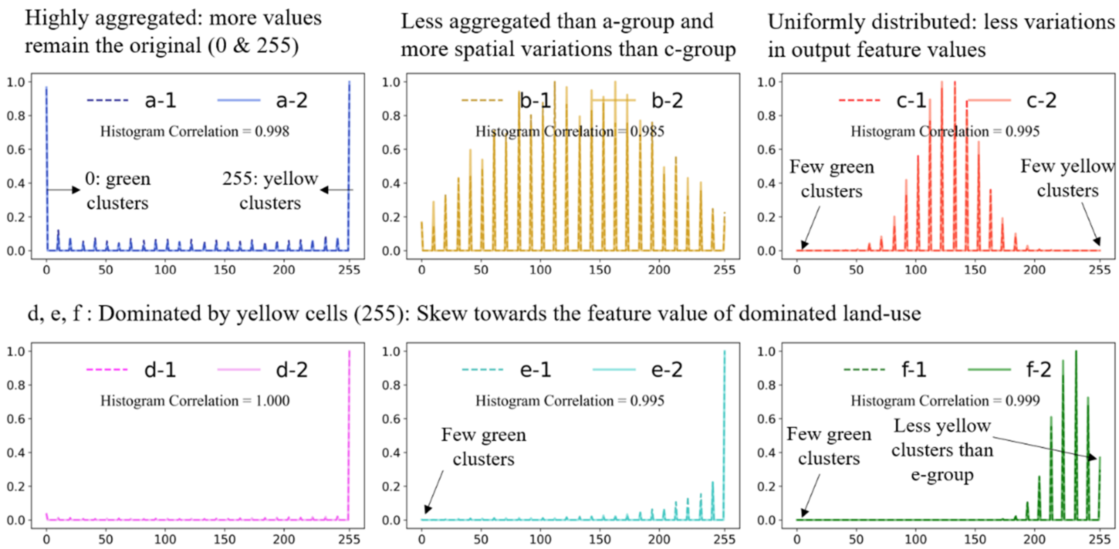
1) Test Data: The random binary maps generated for the binary land-use map comparison are shown in Fig. 2-9. With the NLMpy python module, 6 groups of binary maps were generated with different land-use pattern features, each map is composed of 50×50 land use cells. The yellow (Y) and green (G) cells in maps of a, b, c groups have the same composition proportions of Y: G = 1:1, while d, e, f groups have the same composition proportions of Y: G = 9: 1. As to the spatial relationships, maps of a, d groups have the same aggregation level, while maps in the e group are less aggregated. Besides, maps in the b group are composed of land-use patches with the size in the range of [3,5] units, while c, f groups are both completely random and uniformly distributed.

2) Experiment and Result: According to the framework of the proposed method, for the binary land-use maps, firstly, the land-use feature values were initialized as (Y = 255*e^0, G = 0) according to Equation (2.1). Then the test maps were smoothed with different combinations of weight matrices (mean, inverse distance, and 2D Gaussian) and neighborhood sizes (k = 3,5,7,9).



With the NLMpy python module, the aggregation level of maps in a, d groups were controlled by **mdp()** function (parameter h = 0.75), b group was controlled by **randomRectangularCluster()** function (parameter $minL$ = 3, $maxL$ = 5), e group was based on **mpd()** function (parameter h = 0.1), c, f groups was based on **random()** function. As to the composition, with **classifyArray()** function, land uses in a, b, c groups were classified into 2 categories of equal proportion (parameter *weights* = [0.5,0.5]), while in d, e, f groups, the proportions are Y:G = 9:1 (paramter *weights* = [0.9, 0.1]).

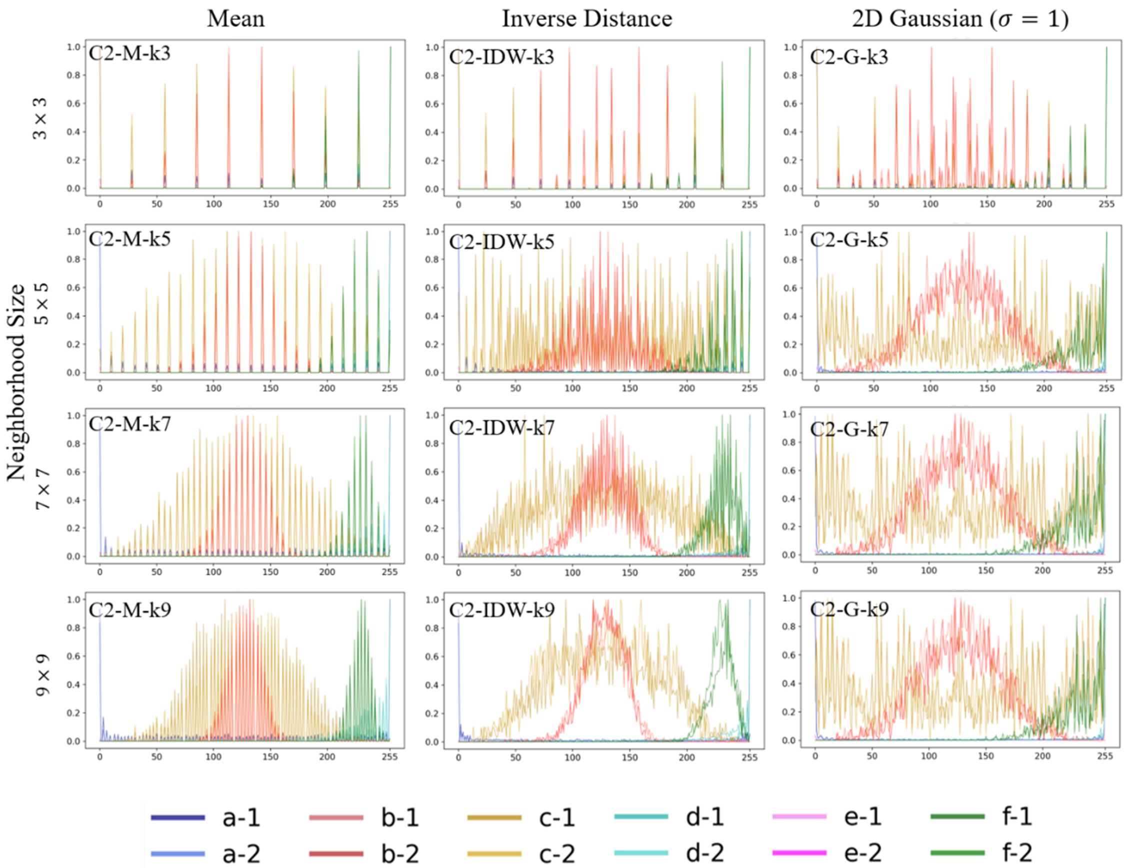
Fig. 2-9 The Simulated Binary Land-use Maps



* Each graph consists of two histograms with similar colors (an opaque dashed line and a translucent solid line) that are almost with each other.

Fig. 2-10 The Land-use Pattern Feature Descriptors of Each Group of Test Maps Extract by 5×5 Mean Matrix

Fig.2-10 displays the histograms of output feature values of each group's maps after smoothing with a 5×5 mean matrix, illustrating the land-use pattern descriptors extracted from the original maps. Specifically, the counts of the initialized land-use feature values indicate the areal occupancy and aggregation level: Land-use cells in maps of group-a, d are more aggregated than other groups, thus they have more land-use cells remain the original value (0, 255); in contrast, land use in maps of group-b, c, d, e are relatively dispersed, they have more land-use cells shifted to other output values. Besides, comparing the histogram graph of group-a, b, c with group-d, e, f, it can be found when there is a dominant land-use, more output feature values remain the original or towards the feature value of the dominant land-use category, the histogram skews to the feature value of the dominant land-use, thus reflecting the composition proportions of the land uses. Furthermore, the number of different output feature values reflects the spatial variations. For instance, comparing the histogram graph of group-b, c, the output feature values of maps in group-b are spread to more different values while the output feature values of maps in group-c concentrate in a small range, corresponding to that the land uses in the original maps of group-c distributed more uniformly than group-b. On the whole, the land-use descriptors represented by the histogram of the output feature values after smoothing captured both the compositional and configurational features of the original maps and unified the local and global features of the land-use patterns, and successfully identified the land-use pattern of maps in the same group. When put the histograms in Fig.2-10 together, it becomes the histogram graph-C2-M-k5* in Fig.2-11, the histograms distinguished the land-use patterns of different groups.



The notation at the top-left of each histogram graph denotes the number of land use categories contains in the original map and the combination of weight matrix and neighborhood size employed in the smoothing operation. Specifically, the notation takes the form of "C (the number of land-use categories)- (the type of weight matrix)-k (the neighborhood size)". As to the type of weight matrix, "M" denotes the mean matrix, "IDW" represents the inverse distance matrix, and "G" represents the 2D Gaussian matrix. For example, "C2-M-k3" means the histograms are the distribution of output feature value of binary maps after smoothing by a 3 × 3 mean matrix. This notation rule is consistent in other parts of this dissertation.

Fig. 2-11 The Land-use Pattern Feature Descriptors Extracted with Different Combination of Weight Matrices and Neighborhood Sizes

Fig.2-11 shows the histograms of all tests maps after smoothing with different combinations of weight matrices and neighborhood sizes, illustrating the influence of weight matrices and neighborhood sizes on the land-use pattern descriptor. The histograms of maps in the same groups were given similar line colors. It can be observed that in either situation, the histograms of maps in the same group almost overlap with each other, and the histograms of maps in different groups distinguish from each other. The similarity is calculated based on the land-use pattern descriptors according to equation (3). Therefore, although the histogram correlations vary when the combination of weight matrix or neighborhood size in the smoothing operation varies, they all suggest consistent results with the histogram graphs in Fig.2-11. As an example, Table.2-1 displays the results of land-use pattern similarity when the weight matrix = Mean and the neighborhood size $k = 3$, 5, and 7, respectively.

In Table 2-1, the correlations between maps of the same group are close to 1, suggesting their land-use patterns are almost the same, whereas the correlations between maps of different groups indicate their land-use patterns distinct from each other in different degrees. The fact that similarity evaluation varies with different neighborhood sizes illustrated how the definition of the minimum land-use cluster size affects the land-use pattern descriptor. Particularly, the similarities of a-b[*4)] for instance, when the neighborhood size $k = 3$, their similarity is about 0.7. However, while $k = 5$, their similarity largely decreased to 0.2. The reason is that when creating maps in the b-group, the aggregation level of land uses was controlled by setting the land-use patch sizes to [3,5]. Therefore, when the neighborhood size is defined as $k = 3$, many of the land-use cells will be able to remain their initial value after smoothing, which is similar with a-group; whereas, when $k = 5$, only the land-use cells in the center of the patches larger than 5×5 remain unchanged after smoothing, in this case, much fewer land-use cells' feature value remained the original, thus its similarity with a-group becomes lower. Accordingly, it becomes clear that the neighborhood size determines the aggregation level discrimination standard, thus influencing the pattern descriptor.

Besides, in the situation of C2-M-K3, the similarities between a-b group and a-d group are both about 0.7, however, the similarity between b-d group are about 0.4, which demonstrated that the proposed method are able to evaluate the similarity between land-use pattern with integrative consideration of composition and distribution properties.

Consequently, this experiment exemplified how the pattern descriptors represent the land-use pattern features in the corresponding map, and the results verified the effectiveness of our method in comparing the land-use pattern of binary land-use maps.

Table 2-1 The Similarity of Land-use patterns (n=2)

Histogram Correlation Result of C2-M-k3

id		a-1	a-2	b-1	b-2	c-1	c-2	d-1	d-2	e-1	e-2	f-1	f-2
Composition		Y:G =	1:1					9:1					
a-1	1:1	1.000											
a-2	1:1	0.999	1.000										
b-1	1:1	0.703	0.673	1.000									
b-2	1:1	0.712	0.681	0.997	1.000								
c-1	1:1	0.130	0.100	0.714	0.694	1.000							
c-2	1:1	0.138	0.106	0.721	0.702	0.999	1.000						
d-1	9:1	0.749	0.751	0.461	0.454	0.022	0.026	1.000					
d-2	9:1	0.746	0.748	0.464	0.457	0.028	0.031	1.000	1.000				
e-1	9:1	0.716	0.712	0.519	0.515	0.092	0.097	0.982	0.983	1.000			
e-2	9:1	0.721	0.717	0.510	0.506	0.081	0.085	0.987	0.988	1.000	1.000		
f-1	9:1	0.542	0.524	0.525	0.540	0.094	0.107	0.691	0.693	0.802	0.790	1.000	
f-2	9:1	0.573	0.556	0.535	0.549	0.098	0.110	0.739	0.741	0.843	0.832	0.996	1.000

Histogram Correlation Result of C2-M-k5

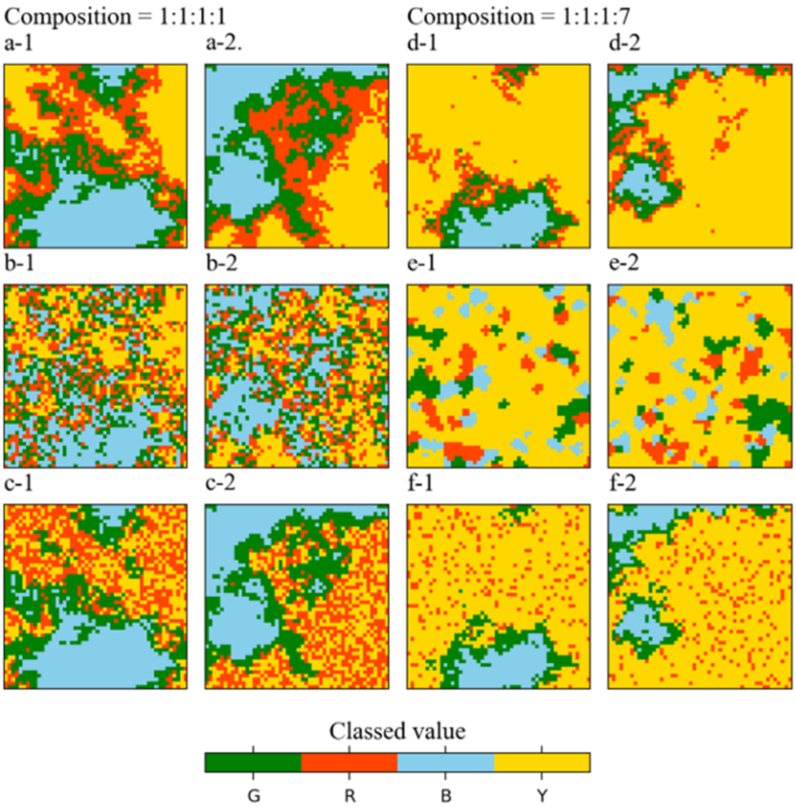
id		a-1	a-2	b-1	b-2	c-1	c-2	d-1	d-2	e-1	e-2	f-1	f-2
Composition		Y:G =	1:1					9:1					
a-1	1:1	1.000											
a-2	1:1	0.998	1.000										
b-1	1:1	0.246	0.198	1.000									
b-2	1:1	0.234	0.183	0.985	1.000								
c-1	1:1	0.093	0.064	0.727	0.720	1.000							
c-2	1:1	0.095	0.065	0.747	0.733	0.995	1.000						
d-1	9:1	0.738	0.740	0.088	0.067	0.003	0.003	1.000					
d-2	9:1	0.734	0.736	0.091	0.070	0.005	0.006	1.000	1.000				
e-1	9:1	0.705	0.701	0.176	0.157	0.025	0.025	0.958	0.959	1.000			
e-2	9:1	0.711	0.708	0.158	0.138	0.025	0.026	0.973	0.973	0.995	1.000		
f-1	9:1	0.218	0.198	0.231	0.237	-0.024	-0.024	0.208	0.209	0.456	0.376	1.000	
f-2	9:1	0.239	0.220	0.228	0.233	-0.025	-0.025	0.239	0.239	0.484	0.404	0.999	1.000

Histogram Correlation Result of C2-M-k7

id		a-1	a-2	b-1	b-2	c-1	c-2	d-1	d-2	e-1	e-2	f-1	f-2
Composition		Y:G =	1:1					9:1					
a-1	1:1	1.000											
a-2	1:1	0.996	1.000										
b-1	1:1	0.125	0.080	1.000									
b-2	1:1	0.129	0.083	0.975	1.000								
c-1	1:1	0.074	0.038	0.730	0.762	1.000							
c-2	1:1	0.076	0.039	0.745	0.770	0.995	1.000						
d-1	9:1	0.743	0.736	-0.004	-0.004	-0.007	-0.007	1.000					
d-2	9:1	0.735	0.727	-0.001	-0.001	-0.004	-0.004	1.000	1.000				
e-1	9:1	0.681	0.671	0.040	0.031	-0.008	-0.009	0.902	0.902	1.000			
e-2	9:1	0.704	0.694	0.034	0.031	0.004	0.004	0.943	0.944	0.976	1.000		
f-1	9:1	0.077	0.070	0.033	0.016	-0.038	-0.039	0.046	0.044	0.374	0.185	1.000	
f-2	9:1	0.088	0.082	0.024	0.014	-0.038	-0.038	0.063	0.061	0.378	0.204	0.987	1.000

2.4.2 Experiment of 4-category Land-use maps Comparison

1) Test Data: With the NLMpy python module, we generated 6 groups of random maps composed of 50×50 land-use cells and contained 4 categories (yellow, blue, red, green = Y, B, R, G) land uses, each group have different land-use pattern features. As Fig.2-12 suggests, maps of a, b, c groups have equal composition proportions of Y:B:R:G = 1:1:1:1, while d, e, f groups are Y:B:R:G = 7:1:1:1. As to spatial relationships, land uses in group-a, d have the same aggregation level; group-b are fragmented and dispersed, while e-group are composed of small patches; e, f-group have a hierarchical structure where the red cells randomly dispersed in the yellow area.



For the 4-category maps, the aggregation level of maps in a, d groups were both controlled by **mdp()** function (parameter h = 0.75), and b group was also based on **mdp()** function (parameter h = 0.1), e group was controlled by **randomClusterNN()** function (parameter p = 0.4),c, f groups were hierachical combinations of **mdp()** model (parameter h = 0.75 for yellow, green and blue cells) and **random()** function (for red cells). As to the composition, land uses in a, b, c groups have the equal proportion of Y:B:R:G = 1:1:1:1, while in d, e, f groups are Y:B:R:G = 7:1:1:1.(**classifyArray()** function)

Fig. 2-12 Simulated Land-use Maps of 4 categories

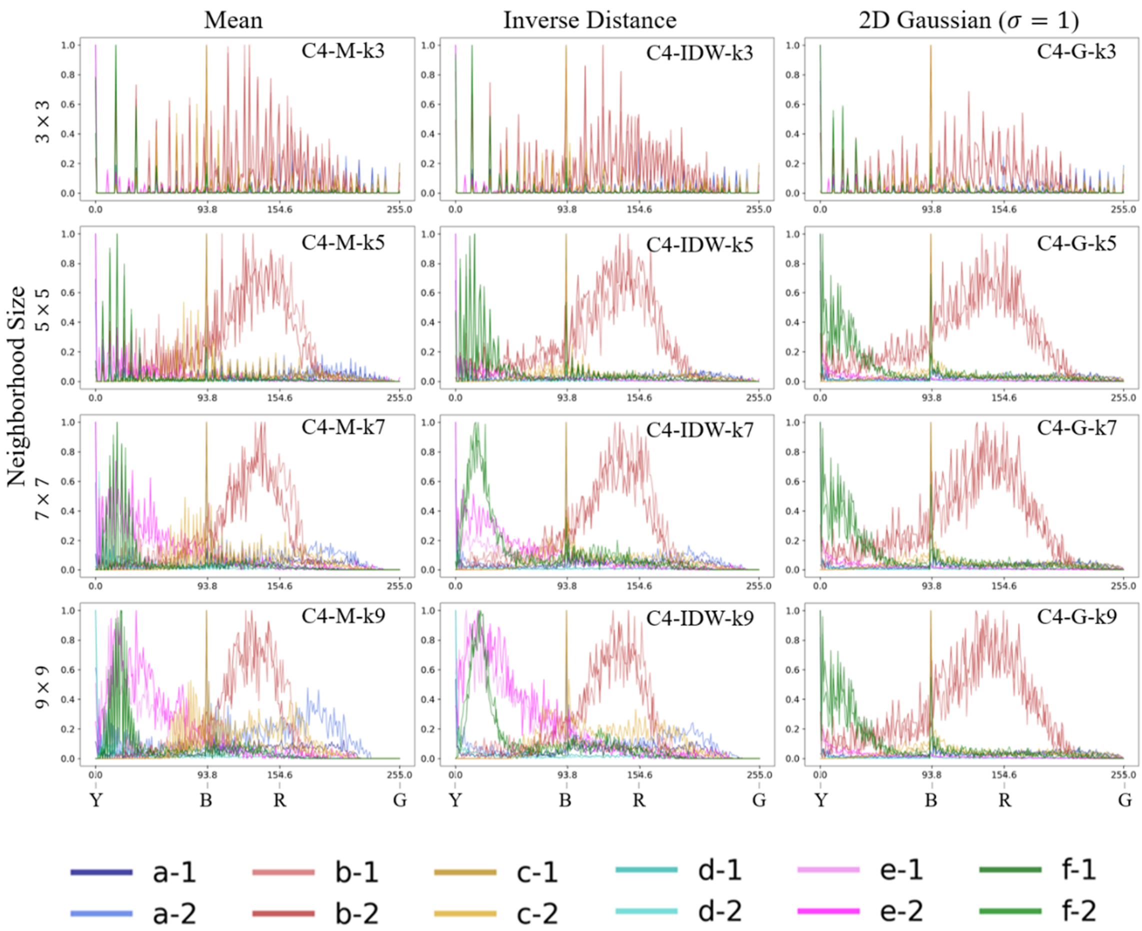


Fig. 2-13 Land-use Feature Descriptors of 4-category Maps Extracted with Different Combination of Weight Matrix and Neighborhood Size

2) Experiment and Result: For land-use maps (Fig.2-12) composed of 4 categories of land uses (G, R, B, Y), the land-use feature values were initialized according to Equation (1), i.e. the four feature values for 4 categories are ($G = 255 * e^{0}, R = 255 * e^{\frac{-1}{2}}, B = 255 * e^{-1}, Y = 0$). Then the test maps were smoothed with different combinations of weight matrices (mean, inverse distance, and 2D Gaussian) and neighborhood sizes (k = 3,5,7,9), the pattern descriptors in each smoothing situation are represented as histogram graphs in Fig.2-13. In each histogram graph of Fig.2-13, the histograms of maps in the same group were given similar colors, thus the land-use pattern descriptors of each group can be observed more clearly. As a result, it can be observed in each graph of Fig.2-13, the land-use pattern descriptors defined by different combinations of weight matrix and neighborhood sizes successfully distinguished the land-use pattern features of maps in different groups.

As Fig.2-13 suggests, the counts at the initial land-use feature values corresponding to the areal occupancy and the aggregation level of each land-use category: specifically, the more aggregative a land-use category is, the more land-use cells remain the original feature values (e.g., yellow cells in a, d, e group, blue cells in all groups), on the contrary, the more dispersed, fewer cells remain their initial feature values. As to the areal occupancy, when there are dominant land-use cells in a map, the histograms skew to the feature value of the dominant land-use (e.g., d, e, f groups). Besides, the number of different output feature values corresponding to the spatial variation of location relations between land-use cells, the more fragmented and heterogeneous, the output feature values will be scattered to more different values, the histograms have a higher frequency of fluctuations (e.g., b group). Consistent with the binary maps, the land-use descriptors well represented the desired land-use pattern features.

Also, similarity evaluation based on the histogram correlations suggested consistent results with the histograms in Fig.2-13. Table 2-2 displays the similarity assessment results when the maps were smoothed with different 5×5 weight matrices (Mean, IDW, 2D-Gaussian), to illustrate how different weight matrices affect the land-use pattern feature descriptors. When applying a complex matrix (such as 2D-Gaussian), the location relations of land-use cells are distinguished more delicately, thus more different numbers will be needed to label different neighborhood patterns, and the pattern descriptor will contain more elements. Therefore, when computing the similarity based on the pattern descriptors, the results will vary when employ a different weight matrix as a feature extractor. Nevertheless, as Table 2-2 suggests, in spite of the different weight matrices employed in the image smoothing process, the histogram correlations between maps of the same group are close to 1, suggesting they are almost the same; the histogram correlations between different groups successfully distinguished the maps in different groups according to the similarity level of their land-use pattern descriptors(histogram). Take group-a for an example, the similarities of a-c are about 0.7, which is the highest similarity between all different groups. This is understandable since they have the same composition proportion, and the blue and green cells were generated with the

Table 2-2 The Similarity of Land-use Patterns (n = 4)

Histogram Correlation Result of C4-M-k5

id		a-1	a-2	b-1	b-2	c-1	c-2	d-1	d-2	e-1	e-2	f-1	f-2
Composition		1:1:1:1						7:1:1:1					
a-1		1.000											
a-2		0.936	1.000										
b-1	1:1:1:1	0.231	0.197	1.000									
b-2		0.122	0.087	0.912	1.000								
c-1		0.742	0.574	0.306	0.243	1.000							
c-2		0.627	0.515	0.292	0.236	0.925	1.000						
d-1		0.604	0.755	0.028	-0.020	0.043	0.036	1.000					
d-2		0.559	0.720	0.031	-0.010	-0.004	-0.004	0.997	1.000				
e-1	7:1:1:1	0.478	0.635	-0.043	-0.083	-0.050	-0.040	0.925	0.925	1.000			
e-2		0.449	0.597	-0.066	-0.103	-0.050	-0.039	0.881	0.877	0.980	1.000		
f-1		0.268	0.307	0.108	0.057	0.072	0.056	0.236	0.202	0.365	0.401	1.000	
f-2		0.191	0.250	0.098	0.062	-0.008	-0.012	0.234	0.206	0.373	0.406	0.989	1.000

Histogram Correlation Result of C4-IDW-k5

id		a-1	a-2	b-1	b-2	c-1	c-2	d-1	d-2	e-1	e-2	f-1	f-2
Composition		1:1:1:1						7:1:1:1					
a-1		1.000											
a-2		0.949	1.000										
b-1	1:1:1:1	0.203	0.163	1.000									
b-2		0.044	0.021	0.877	1.000								
c-1		0.774	0.587	0.299	0.176	1.000							
c-2		0.718	0.562	0.294	0.180	0.952	1.000						
d-1		0.604	0.780	0.011	-0.051	0.034	0.023	1.000					
d-2		0.561	0.746	0.015	-0.036	-0.015	-0.024	0.997	1.000				
e-1	7:1:1:1	0.499	0.676	-0.064	-0.124	-0.064	-0.077	0.947	0.944	1.000			
e-2		0.469	0.636	-0.090	-0.142	-0.067	-0.083	0.905	0.900	0.979	1.000		
f-1		0.281	0.304	0.005	-0.102	0.096	0.059	0.272	0.238	0.383	0.442	1.000	
f-2		0.164	0.217	-0.012	-0.091	-0.044	-0.073	0.282	0.258	0.415	0.466	0.969	1.000

Histogram Correlation Result of C4-G-k5

id		a-1	a-2	b-1	b-2	c-1	c-2	d-1	d-2	e-1	e-2	f-1	f-2
Composition		1:1:1:1						7:1:1:1					
a-1		1.000											
a-2		0.961	1.000										
b-1	1:1:1:1	0.202	0.180	1.000									
b-2		0.010	0.000	0.836	1.000								
c-1		0.769	0.608	0.288	0.127	1.000							
c-2		0.743	0.591	0.258	0.109	0.971	1.000						
d-1		0.627	0.785	0.028	-0.060	0.039	0.031	1.000					
d-2		0.587	0.753	0.028	-0.048	-0.011	-0.018	0.998	1.000				
e-1	7:1:1:1	0.535	0.698	-0.037	-0.121	-0.060	-0.068	0.965	0.964	1.000			
e-2		0.518	0.673	-0.052	-0.141	-0.057	-0.067	0.934	0.931	0.987	1.000		
f-1		0.424	0.443	-0.100	-0.227	0.180	0.141	0.418	0.392	0.553	0.608	1.000	
f-2		0.282	0.350	-0.119	-0.211	-0.035	-0.068	0.478	0.466	0.626	0.667	0.930	1.000

same configurational parameters. Also, the similarities between a-d are greater than 0.5, suggesting a relatively high similarity between maps in a-group and d-group. This is also rational because although their land uses composition proportion differs, they have the same spatial relationships. Moreover, although both c-group and d-group have relatively high similarities with a-group, the similarities between c-d are close to 0 since they differ in both composition proportions and spatial distribution. This demonstrated that the proposed method assesses the similarity of land-use patterns according to both the land-use composition and their spatial relationships.

Besides, the high similarities between group d-e in Table 2-2 illustrated the neighborhood size's influence on the discrimination of land-use pattern: because the land-use patches are generally larger than 5 × 5 in maps of d and e groups, the image smoothing with a 5 × 5 matrix will not be able to differentiate them. To distinguish the land-use patterns in d and e groups, a larger neighborhood size should be defined when executing image smoothing.

On the whole, the experiment results illustrated how the pattern descriptors defined by the proposed method represent the land-use pattern features in corresponding maps, and verified the effectiveness of this method in comparing the land-use patterns between the multi-categorical land-use maps. Also, this part intuitively illustrated the influence on the land-use pattern descriptor when employing different combinations of weight matrix and neighborhood size as a feature extractor. In addition, due to the computation efficiency, this part computed the histogram similarity based on 256 bins, which reduced the accuracy of the similarity computation. This relates to the limitations of the proposed method, which will be further discussed in section 2.6.

2.5 A Sample Application of the Method

To illustrate the applicability of our method, we applied it to the comparative analysis of the land-use pattern in a 1km × 1km square area surrounding Shinkansen stations. The sample Shinkansen stations include stations of 4 regional center cities (Takasaki, Toyama, Nagano, and Kanazawa) and 2 major cities (Sendai and Hakata). The land-use data is the 100-meter land use subdivision mesh data of urban areas published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport and Tourism, 2016. The land-use composition and distribution situations of the 6 sample station areas are shown in Fig. 2-14, and the land-use composition proportions[*5)] of the station areas are shown as shown in Table 2-3.

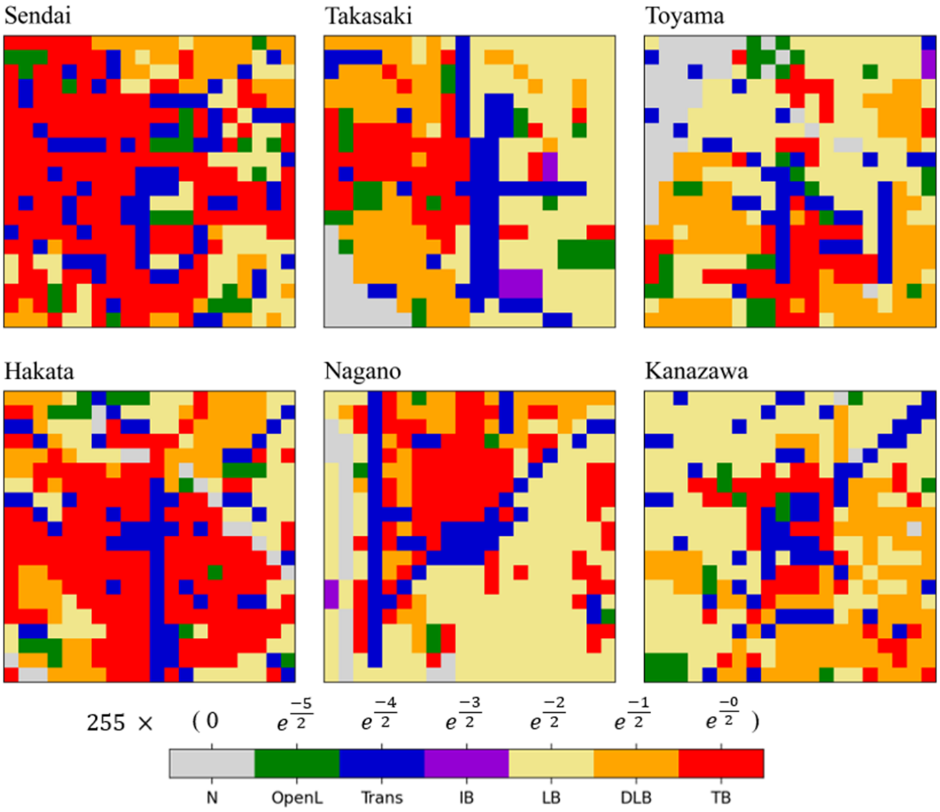


Fig. 2-14 Land-use Maps of Shinkansen Station Areas

Table 2-3 Land-use Composition Proportions of the Shinkansen Station Areas

ID		Station	TB	DLB	LB	IB	Trans	OpenL	N
A	1	Sendai	49.00%	12.25%	15.50%	0.00%	17.25%	6.00%	0.00%
	2	Hakata	44.75%	15.00%	17.50%	0.00%	14.50%	4.50%	3.75%
B	3	Takasaki	16.00%	21.75%	33.50%	2.00%	15.00%	6.75%	5.00%
	4	Nagano	25.00%	10.50%	43.50%	0.50%	13.75%	1.50%	5.25%
	5	Toyama	14.25%	23.50%	33.00%	0.50%	10.50%	6.75%	11.50%
	6	Kanazawa	12.25%	26.25%	42.25%	0.00%	14.50%	4.00%	0.75%

①TB: Tall Building Land; ② DLB: Dense Low-rise Building Land; ③ LB: Low-rise Building Land; ④ IB: Industrial Building Land; ⑤ Trans: Transportation land;⑥ OpenL: Open Land; ⑦ N: Non-construction Land

Table 2-4 Similarity Based on Land-use Composition Proportions

Station	Sendai	Hakata	Takasaki	Nagano	Toyama	Kanazawa
Sendai	1.000	0.993	0.721	0.762	0.666	0.620
Hakata	0.993	1.000	0.775	0.805	0.737	0.679
Takasaki	0.721	0.775	1.000	0.941	0.983	0.982
Nagano	0.762	0.805	0.941	1.000	0.915	0.923
Toyama	0.666	0.737	0.983	0.915	1.000	0.963
Kanazawa	0.620	0.679	0.982	0.923	0.963	1.000

Table 2-5 Similarities Based on Land-use Pattern Feature Descriptors Extracted with the 3×3 Mean Matrix

Station	Sendai	Hakata	Takasaki	Nagano	Toyama	Kanazawa
Sendai	1.000	0.718	0.162	0.157	0.174	0.097
Hakata	0.718	1.000	0.233	0.266	0.196	0.153
Takasaki	0.162	0.233	1.000	0.688	0.562	0.582
Nagano	0.157	0.266	0.688	1.000	0.449	0.569
Toyama	0.174	0.196	0.562	0.449	1.000	0.587
Kanazawa	0.097	0.153	0.582	0.569	0.587	1.000

M-K3: In the image smoothing operation, wight matrix = Mean, neighborhood size $k = 3$
The number of bins $N = 10^4$

The similarities based on the land-use composition is shown in Table 2-4, accordingly, the 6 sample stations can be distinguished into 2 groups: A-group is 2 stations of the major cities where the Tall Building land occupies about half of the area, whereas B-group includes 4 stations of the regional center cities whose Low-rise Building land takes over 1/3 of the total station area and the Tall Building land together with the dense low-building land takes about 40% of the total areas. However, it is hard to further distinguish the spatial characteristics of land-use distribution between the station areas according to land-use composition proportion.

Therefore, we applied the proposed method to compare the land-use patterns in the sample station areas. Firstly, the feature values of 7 categories of land uses ([TB, DLB, LB, IB, Trans, OpenL, N]) were initialized according to Equation (2.1) where the parameter $z = 255$ and $\varphi = 2$. As a result, the feature values for the land-use categories are initialized as 255 × [TB $=e^{\frac{-0}{2}}$, DLB = $e^{\frac{-1}{2}}$, LB $=e^{\frac{-2}{2}}$, IB = $e^{\frac{-3}{2}}$, Trans = $e^{\frac{-4}{2}}$, OpenL = $e^{\frac{-5}{2}}$, N = 0], as shown in Fig.2-14. Then land-use maps were smoothed with a 3×3 mean matrix. Since the mesh data is composed of 100 meter × 100 meter cells, the 3×3 neighborhood on the map corresponds to a 300 meter × 300 meter area in the city, we consider it as a reasonable range for the identification of a land-use cluster, and it is a range that can be considered to be spatial indifferent. The similarity is computed based on Equation (3), and here we set the number of bins $N = 10^4$ to ensure the accuracy when computing the similarities between pattern descriptors. The results are as shown in Table 2-5. This result not only distinguished the compositional difference between stations of A-group (Sendai, Hakata) and B-group (Takasaki, Nagano, Toyama, Kanazawa), but also furtherly distinguished their differences in spatial distribution: among 4 stations of the regional center cities, Takasaki and Nagano are apparently more similar than other two stations.

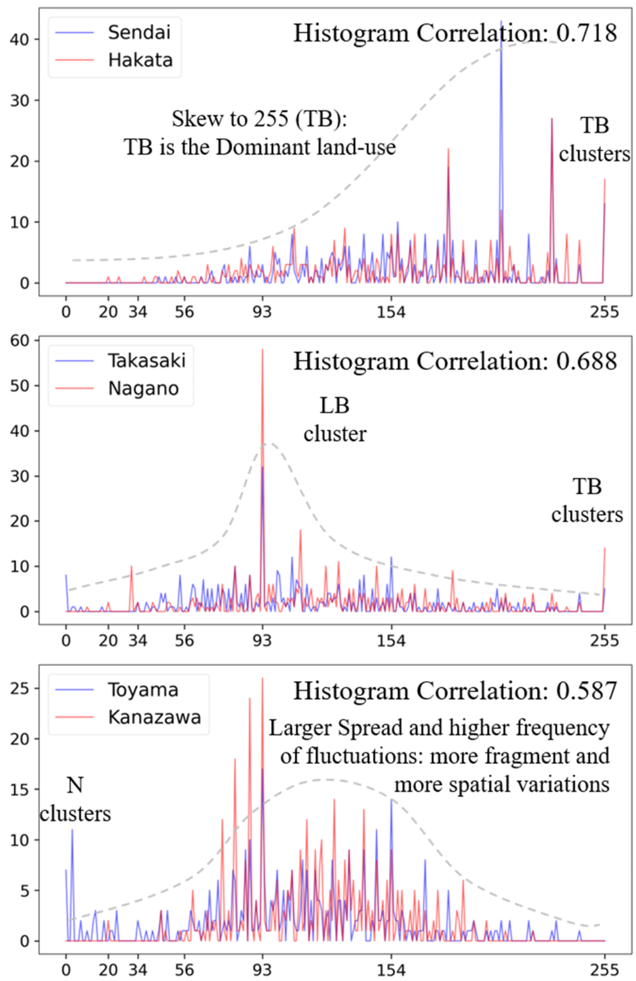


Fig. 2-15 Histogram of Smoothed Land-use Maps of Shinkansen Station Areas

Fig.2-15 displays the pattern descriptors of maps as histograms, providing intuitive reference to analysis the land-use pattern features of the station areas. As shown in Fig.2-15, the histogram of Sendai and Hakata skew to 255, indicates the tall-building land cells are aggregative and dominant in the station area; whereas the histograms of station areas of 4 regional cities concentrate around $255 * e^{-1}$ (LB$\approx$ 93), indicates the low building land takes most of the total station area. Moreover, the counts at $255 * e^{-1}$ in the histograms of Takasaki and Nagano are greater than in Toyama and Kanazawa, indicating the low-building land cells are more aggregated in Takasaki and Nagano. Besides, histograms of Toyama and Kanazawa have a larger spread and higher frequency of fluctuations, suggesting the distribution of land-use cells in Toyama and Kanazawa are more fragmented and have more spatial variations.

In summary, this application demonstrated the proposed method provided an effective method to quantify the similarity level of the land-use patterns in urban areas with integrative consideration of the compositional and configurational characteristics. Moreover, the histogram of output feature values provided an intuitive reference for analyzing the characteristics of land-use patterns.

2.6 The Limitation of the Method

The histogram graphs in Fig.2-11 and Fig.2-13 in section 2.4 have illustrated that, the land-use pattern descriptor varies when employing different combinations of weight matrix and neighborhood size in the image smoothing, thus influencing the land-use pattern similarity evaluation. This section discusses the parameters that influence the land-use pattern descriptor, thus clarifying the limitations and the applicability of the method.

2.6.1 The Trade-off between Computation Efficiency and Accuracy

According to the algorithm in section 2.3, the initial land-use feature values are defined by Equation (2.1), while the elements of the land-use pattern descriptor are the counts for different output feature values which are determined by the number of different neighborhood patterns. Generally, for a land-use map composed of n land-use categories, the number of elements constituting the pattern descriptor equals the number of ways to arrange n-category of land uses in $k \times k$ land-use cells according to the location relations defined by the weight matrix. Their mathematical relation is explained in Note*6). Table 2-6 gives the number of elements contained in the land-use pattern descriptor when employing different combination of neighbor cells' location relation definition (weight matrix) and neighborhood size (k) in different situations of land-use composition (n categories). With the number of land-use categories, the complexity of land-use cells' location relation and the neighborhood size increasing, more numbers will be needed to label all different neighborhood patterns, the pattern descriptor will consist of more elements. According to equation (1), the elements of pattern descriptor will be distributed in the range of $[0, z]$. Therefore, the number of descriptor elements become greater, the intervals between the elements become smaller. However, when computing the similarity according to Equation (2.3), the number of the histogram bins N influences the shape of the histogram. When the number of histogram bins is less than the number of descriptor elements, the counts of multiple output feature values will be merged into one histogram bin, the counts at each histogram bin no longer correspond to a unique neighborhood pattern, which changes the shape of histogram thus influencing the accuracy of the land-use pattern similarity evaluation.

Table 2-6 the Number of Elements Composed of the Land-use Pattern Descriptor in Different Situations

n	k=3			k=5			k=7		
	Mean	IDW	Gaussian	Mean	IDW	Gaussian	Mean	IDW	Gaussian
2	10	18	50	26	306	11250	50	7650	22781250
3	55	135	675	351	20655	6834375	1275	6712875	3.E+12
4	220	660	4900	3276	639540	1.E+09	22100	2.E+09	3.E+16
5	715	2475	24500	23751	11991375	6.E+10	292825	2.E+11	7.E+19
6	2002	7722	95256	142506	2.E+08	2.E+12	3162510	2.E+13	5.E+22
7	5005	21021	308700	736281	2.E+09	4.E+13	28989675	9.E+14	2.E+25
8	11440	51480	871200	3365856	1.E+10	6.E+14	2.E+08	3.E+16	3.E+27
9	24310	115830	2205225	13884156	9.E+10	7.E+15	2.E+09	9.E+17	3.E+29

* n: categories of land-uses contains in the maps; k: neighborhood size
* The scientific notation is used when the value is greater than 10^8

Accordingly, to ensure accuracy when computing similarity between two land-use pattern descriptors, the bin width of the histogram should be narrower than the intervals between the descriptor elements to avoid the merge of the histogram bins. Therefore, assuming the intervals between the output feature values are uniform, the number of bins should at least equal the number of elements constituting the pattern descriptor (i.e., numbers in Table 2-6). However, according to Equation (2.3), the computational complexity of histogram similarity grows with the number of histogram bins, making this method computationally inefficient. In experiments of section 2.4, we set the number of bins equals 256, as a result, in the cases when the number of descriptor elements is greater than 256 (such as C2-IDW-k5, C4-M-k5), the counts at each bin correspond to multiple neighborhood patterns, making the correspondence between the shape of the histogram and the land-use pattern feature less discernible. Therefore, although the results in Table-2-2 are acceptable, they are inaccurate.

2.6.2 Parameter Determination and Optimization

Indubitably, the merge of histogram bins will change the shape of the histogram. However, we found when the intervals between output feature values are relatively even, the merge of histogram bins will cause less change to the shape feature of the histogram. Therefore, when generating the land-use feature values, we kept the intervals relatively even by controlling the parameters of z and φ in equation (1). Besides, when devising the weight matrix, the contribution rate of each land-use cell should be significant by adjusting the parameters of μ in equation (2.5) and σ in equation (2.8). [*7] However, as to the optimal value of these parameters and the range of a rational number of histogram bins N, further studies are needed. In practice, we can avoid the merge of histogram bins by setting the width of bins smaller than the intervals between the output feature values (e.g., section 2.5 set $N = 10^4$), although this will sacrifice the computational efficiency.

An alternative is to reduce the number of pattern descriptor elements by defining the neighborhood with a simple location relation (e.g., Mean) and small neighborhood size (e.g., 3×3,5×5). For a fixed number of land-use categories, the number of descriptor elements is determined by how the neighborhood patterns are defined. The more complicated the neighborhood patterns are defined, the pattern descriptor will reflect more detailed features on the neighborhood level, making the characteristics of the overall land-use pattern less discernible, which might contradict the purpose of the land-use pattern identification. In practice, it is rarely necessary for the urban land-use pattern comparison to define the location relation with a complex matrix such as IDW or 2D Gaussian. In urban studies, the common land-use maps are middle scales such as 1:100-meter, 1:200-meter, and so on. In such cases, to distinguish the land-use clusters in the urban area, usually a 3×3 or 5×5 neighborhood will be sufficient, and it is usually unnecessary to distinguish the locations in the corresponding city area of such neighborhood range. As the sample application of section 2.5 shows, the 3×3 mean matrix has helped to identify the land-use pattern features well. Therefore, when dealing with the land-use patterns in urban studies, a 3×3 or 5×5 mean matrix is recommended. Although the 3×3 or 5×5 matrix indeed has limited ability to distinguish land-use patterns when the land-use clusters in the maps are generally greater than 3×3 or 5×5 units, in such cases, optimizing approaches such as the map resolution transformation or multi-layer feature extraction can be helpful. Nevertheless, such optimizing approaches still need further studies. Considering the scales of commonly used land-use maps in urban studies, it is a good start with a 3×3 or 5×5 mean matrix.

2.7 Summary

In this chapter, we introduced an image matching approach to the land-use pattern comparison and developed a novel land-use pattern comparison that integratively measures the similarity level of both compositional and configurational features of land-use patterns. Specifically, the proposed approach applies image smoothing as a land-use pattern feature extractor to characterize and numerically quantify the integrative characteristics of composition and spatial relationships as a land-use pattern descriptor. The land-use pattern descriptor we defined can be represented by the distribution of smoothed output feature values (histogram), thereby the land-use pattern similarity can be computed based on the land-use pattern descriptors. This chapter explained the algorithms of the proposed method, and verified its effectiveness through experiments with simulated random maps, and demonstrated its applicability in urban studies through a sample comparative analysis of land-use patterns in Shinkansen station areas. Besides, this paper discussed the principles to choose the parameters in the algorithm and clarified the limitation and applicability of the proposed method.

In contrast to many existing land-use pattern comparison approaches which are based on some specific indices that measure the spatial structure of a phenomenon and reveal the underlying spatial process, the proposed method in essence is an image matching approach that compares maps based on the land-use pattern features extracted by the devised feature extractor. Therefore, the proposed method focuses on the visual similarity of land-use patterns without considering the underlying spatial process. Instead, this method provides a fine similarity measurement between land-use patterns with integrative consideration of the composition and spatial relationships between land uses, which provides a good source for further classification.

However, further study is needed to optimize the parameters in the algorithm and improve the applicability of this method. Nevertheless, this study is a meaningful attempt to apply the image matching techniques to the land-use pattern comparison. Compared with the traditional land-use analysis method, the image matching approach is more efficient in handling a large amount of map data. Furthermore, this study presented a method to extract geospatial pattern information from maps through the image smoothing process and provided references to further applications of image processing and machine learning to the study of urban development regularity.

Notes

*1). For a gray scale $(n \times n)$ image and $(k \times k)$ filter/kernel, the dimension of the image resulting from a convolution operation is $(n - k + 1) \times (n - k + 1)$, this is known as the boundary effect in image processing. For example, for an (10×10) image and 3×3 kernel, the output resulting after convolution operation would be of size (8×8). To preserve the shape of the image, a common solution is to make the input image large by simply adding some values to the border of the image, which is called padding. Padding can be implemented by adding a certain constant, using a replicated copy of the boundary pixel, and other methods. [65] In the proposed method, we simply replicate a copy of the boundary pixel, based on the spatial assumption of near things are usually more similar than distant things.

*2). For all the land-use maps in this study, the colors are only for visualizing different land-use categories on the maps, the land-use feature values do not necessarily correspond to the color intensity values.

*3) The histograms are normalized to [0,1] to make the fluctuations trends of histograms easier to observe.

*4) “The similarities of a-b” is short for “the similarity between maps of a-group and b-group”. Other parts of this paper also take this form when discussing the similarities between maps of two groups.

*5) The original land-use data has 17 subdivisions of land uses, which are too many for the analysis of the land-use pattern of the urban area, thus we merged some subdivisions and reclassified them into 7 categories. The original subdivision of land-use can refer to:

https://nlftp.mlit.go.jp/ksj/gml/codelist/LandUseCd-09-u.html

*6) For a land-use map composed of n land-use categories, the number of different neighborhood patterns equals the number of ways to arrange n-category of land uses in $k \times k$ land-use cells according to the location relations defined by the weight matrix.

When the locations of neighbors are defined by mean matrix, all the $k \times k$ land-use cells are the same (Fig.2-8-a), the number of elements composed of the land-use pattern descriptor is:

$$c = C_{k^2+n-1}^{n-1} = \frac{(k^2+n-1)!}{(n-1)!\,(k^2)!}\ (k = 3, 5, 7 \dots)$$

when employ the inverse distance weight matrix, the $k \times k$ land-use cells are distinguished into $\frac{k+1}{2}$ groups according to the land-use cells distance to the center cell, and except the center cell, each group has $4d-4$ ($d = 3,5,7 \ldots, k$) cells (Fig.2-8-b), the number of different neighborhood patterns equals the number of combination of ways of to arrange n categories of land uses in each group of land-use cells in the neighborhood, therefore, the number of elements composed of land-use pattern descriptor can be represented as:

$$c = c_n^1 \prod_{d=3}^{k} c_{4d-4+n-1}^{n-1}$$
$$= n \prod_{d=3}^{k} \left(\frac{(4d-4+n-1)!}{(n-1)!\,(4d-4)!}\right) \quad (k = 3,5,7 \ldots)$$

Similarly, in the case of the 2D-Gaussian weight matrix, the land-use cells in a $k \times k$ neighborhood are distinguished into $\frac{(k+3)(k+1)}{8}$ groups. Among them, besides the center cell, there are $k-1$ groups which contain 4 land-use cells with the same weights, and $\frac{(k-3)(k-1)}{8}$ groups which contain 8 land-use cells with the same weights (Fig.2-8-c), thus the number of pattern descriptor elements is:

$$C = c_n^1 (c_{4+n-1}^{n-1})^{k-1} (c_{8+n-1}^{n-1})^{\sum_{d=1}^{\frac{k-3}{2}} d}$$
$$= n\left(\frac{(4+n-1)!}{(n-1)!\,4!}\right)^{k-1} \left(\frac{(8+n-1)!}{(n-1)!\,8!}\right)^{\frac{(k-3)(k-1)}{8}} \quad (k = 3,5,7 \ldots)$$

*7) When the Gaussian standard deviation $\sigma = 1$, the weight decreases drastically with the distance to the central cell increasing. In this case, when the neighborhood size $k > 5$, the weights of land-use cells outside the 5×5 neighborhood are close to 0, thus these land-use cells' feature values barely have a contribution to the output value. This explains the reason that the histograms land-use maps smoothed with 2D-Gaussian in Fig.4-3 and Fig.5-3 barely changed when neighborhood size $k > 5$.

Chapter 3

The Land-Use Patterns of the Shinkansen Station Areas

This chapter applies the land-use pattern comparison method developed in Chapter 2 to compare the land-use patterns between 91 existing Shinkansen station areas. The image-matching-based method provides a similarity measurement between the land-use patterns of different Shinkansen station areas, based on which we computed the pairwise distances and classified the land-use patterns of station areas into different groups with Hierarchical clustering. In this way, the station areas of the same groups will have not only similar land-use composition proportions but also similar spatial characteristics. Specifically, **Section 3.1** specifies the research objects and the map data source. Then, **Section 3.2** introduces the method to calculate the land-use pattern similarity and the method to convert the similarities to the pairwise distance for clustering. **Section 3.3** clarifies the land use pattern features based on the clustering result. **Section 3.4** discusses the trend of land-use patterns of station areas in cities of different population scales. **Section 3.5** summarizes this chapter.

3.1 Study Objects and Data

This part focused on the land-use situation in the 1-kilometer radius range around each of 91*1) existing Shinkansen stations, as Fig. 3-1 suggests. The land-use data is the 100-meter land use subdivision mesh data of urban areas of 2016 published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism. Namely, as Fig.3-1 shown, every square land-use unit in the map corresponds to a 100×100 m^2 area in the city. Besides, the original land-use data has 17 subdivisions, which contains some subdivisions we are not interested in. Therefore, we merged the subdivision into 7 categories, as Table 3-1 suggests.

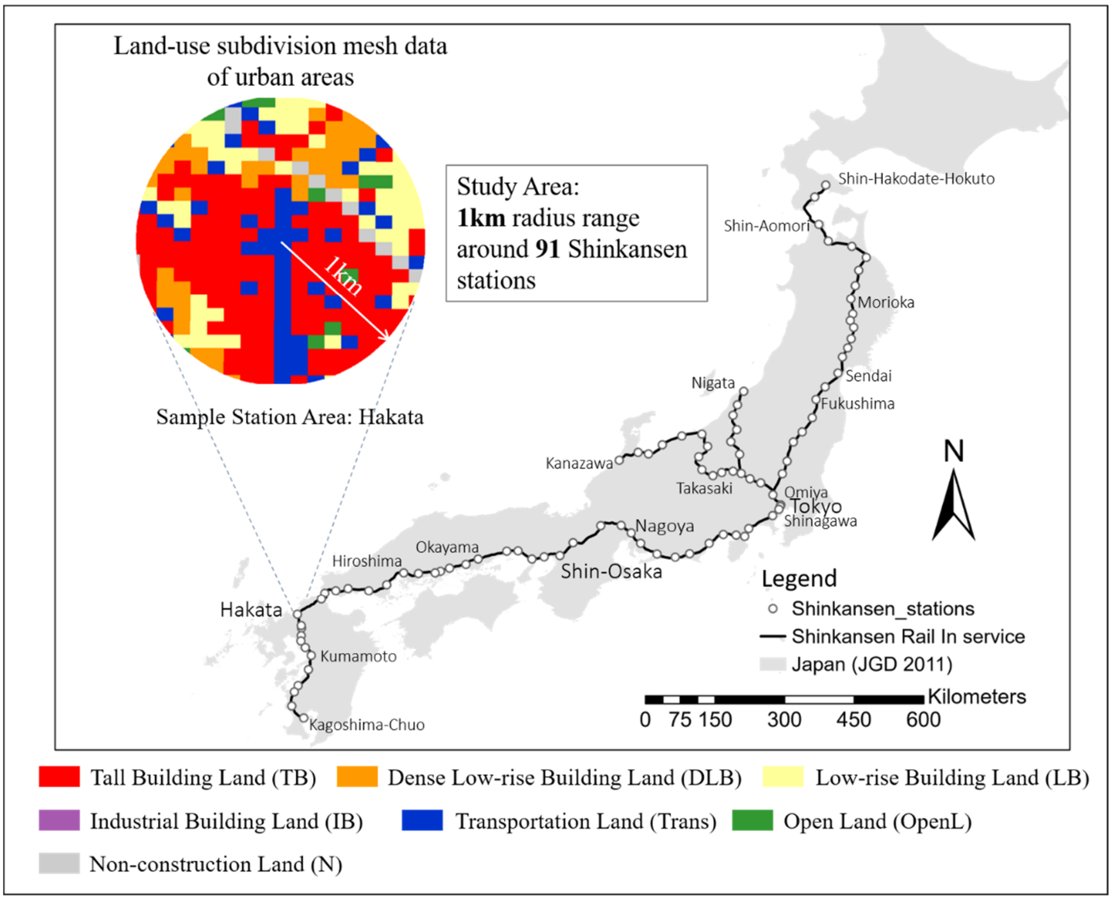


Fig. 3-1 Study Objects: 91 Shinkansen Stations and The Range of a Station Area

Table 3-1 The Subdivisions of Land Use Data

ID	Land-use categories	Abbrev.	Description
1	Tall Building Land (高層建物)	TB	Urban area dominated by buildings with 4 floors or more
2	Dense Low-rise Building Land (低層建物密集地)	DLB	Urban area where buildings of 3 stories or less are densily concentrated
3	Low-rise Building Land(低層建物)	LB	Urban area where buildings of 3 stories or less are distributed collectively and in a relatively low-density form
4	Industrial Building Land　(工場)	IB	Urban area dominated by industrial buildings
5	Transportation Land	Trans	Urban area dominated by land for transportation uses, includes the road (道路) and rail (鉄道) in the original subdivision data.
6	Open Land	OpenL	Urban area dominated by open land including the public facility land (公共施設等用地), open space (空地), park and greenland (公園・緑地), golf course (ゴルフ場) in the original subdivision.
7	Non-construction Land	N	Corresponds to other land-use categories in the original subdivision, including farmland(田), other agricultural land (その他の農用地),forest(森林), wasteland (荒地), rivers and lakes (河川地及び湖沼), seaside (海浜),and sea(海水域).

3.2 Method

To make clear the land-use patterns around 91 existing Shinkansen stations, we extract the land-use pattern features of the station areas then classify them into different groups. Specifically, as shown in Fig. 3-2, the analysis can be divided into 2 steps: firstly, we compute the similarities between land use maps based on the proposed land-use pattern comparison method developed in chapter 2; then, we compute the pairwise distance based on the similarity and implement hierarchical clustering to classify the station areas into different groups. The land-use patterns are clarified by summarizing the land-use pattern features of each group.

3.2.1 Image Similarity

To calculate the similarity, firstly, we need to convert the map data to a numeric array. As Fig. 3-3 suggests, the map was resampled to a 20×20 image, so that the image has the same resolution as the original map. *2) Then, we initialize the land-use feature values based on Equation (2.1). Specifically, there are 7 categories of land-uses in the source maps, however, we need one more feature value to label

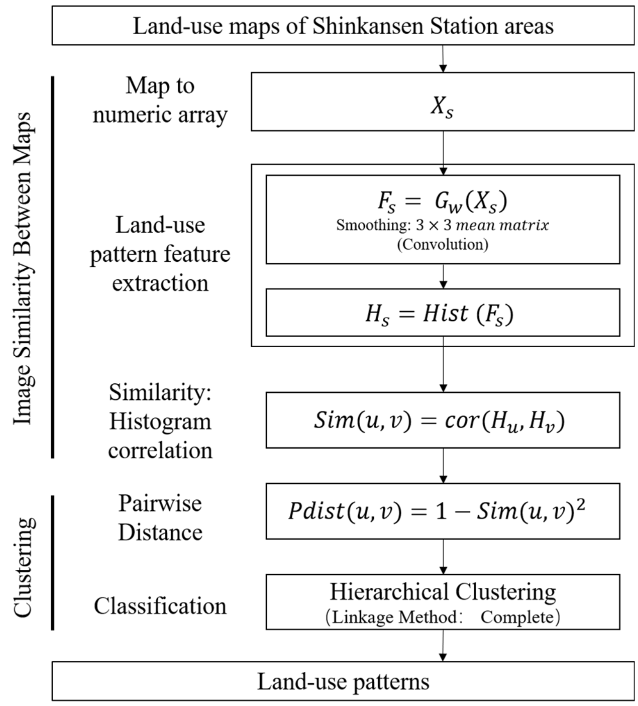


Fig. 3-3 Study Flow and Method

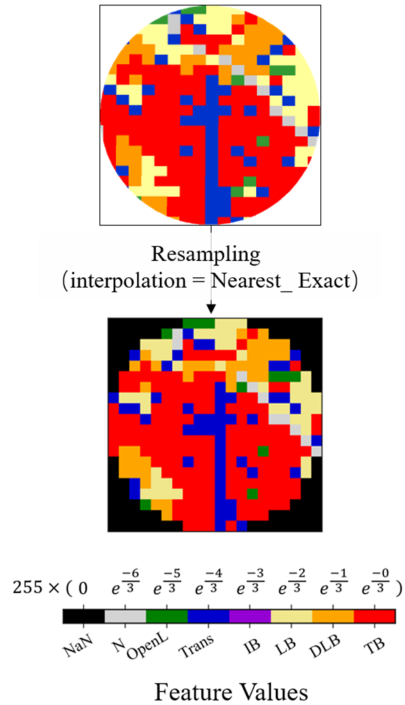


Fig. 3-2 The Initialization of Land-use Feature Values

the background (NaN) of the image. Namely, we need $n = 8$ initial feature values to label all the units in a map. Therefore, considering the intervals between the initial values, we set the parameters in Equation (2.1) as $z = 255$ and $\varphi = 3$. Namely, the numeric representations of land use maps are based on following Equation 3.1:

$$a_i = \begin{cases} \frac{255}{e^{\frac{i}{3}}} & (i = 0,1,2,\ldots,n-2) \\ 0 & (i = n-1) \end{cases} \qquad \text{Equation (3.1)}$$

Where i is the index for each land-use category.

As a result, the units in the map were labelled according to their land-use properties, as 255 × [TB $=e^{\frac{-0}{3}}$, DLB = $e^{\frac{-1}{3}}$, LB $=e^{\frac{-2}{3}}$, IB = $e^{\frac{-3}{3}}$, Trans = $e^{\frac{-4}{3}}$, OpenL = $e^{\frac{-5}{3}}$, N = $e^{\frac{-6}{3}}$,NaN = 0], as shown in Fig. 3-3.

Then, considering the map scale and the study area, the land-use pattern features were extracted with a 3 × 3 mean matrix and represented as the histograms of the output feature values. The similarities were computed based on the histogram correlation between land-use pattern feature descriptors according to Equation (2.3). According to Table 2-6, when calculating similarities, the number of histograms $N = 12000$.

3.2.2 Hierarchical Clustering

To classify the land-use patterns based on their land-use pattern features, firstly, we need to compute the pairwise distances between station areas. However, the similarities were based on the histogram correlations, for which we cannot operate dynamic clustering such as K-means. Instead, we used hierarchical clustering. For hierarchical clustering, the pairwise distances of the land-use pattern features of two station areas $Pdist(u, v)$ can be computed according to Equation (3.2):

$$Pdist(u, v) = 1 - Sim(u, v)^2 \qquad \text{Equation (3.2)}$$

Where $Sim(u, v)$ denotes the histogram correlation between the land-use patterns of two maps.

We take the squared form of the similarity for it avoids dealing with minus correlations and it corresponds to the form of Euclidean distance for standardized data, therefore, it preserves the original resemblances better. Besides, for the process of iteratively merging clusters to clusters, we use the complete-linkage method [77]as Equation (3.3) suggests:

$$d(u, v) = Max(pdist(u[i], v[j])) \qquad \text{Equation (3.3)}$$

Namely, merge the clusters based on the maximum distance between members of clusters. The complete-linkage method is more robust to outliers and allows us to find compact clusters. [78]

As a result, the dendrogram is shown in Fig.3-5. The cophenetic correlation coefficient [79] evaluates how well a dendrogram preserves the pairwise distances between the elements in the original dataset. The cophenetic correlation coefficient when calculating the pairwise distances with a squared form is 0.72, suggesting the clustering result matches the original resemblances well.

The final step is to cut the dendrogram to get the final clustering result. Any unsupervised clustering task has to face the problem of the optimal number of clusters, which is still a problem that has not been solved well. For hierarchical clustering, it has the advantage of not having to pre-define the number of clusters. Instead, it records the sequences of merges or splits with a dendrogram, which provides an intuitive reference for choosing the number of clusters. Besides, we introduced a variant of the *elbow method* to help determine the number of clusters. Since the desired clustering result is when reaching the maximum inter-cluster variation and the minimum total intra-cluster variation, the elbow method is based on the concept that the optimal number of clusters is most efficient for minimizing the total intra-cluster variation. [80] In other words, the optimal number of clusters is when adding another cluster does not improve much better the total intra-cluster variation. To correspond to the distance in the dendrogram, here instead of the within-cluster sum of squares, Fig.3-4 shows the inter-cluster distance growth, and the optimal cluster number is the clustering step where the acceleration of distance growth is the biggest. Accordingly, the optimal cluster number is 6, therefore the dendrogram is cut-off to get 6 clusters as shown in Fig. 3-5.

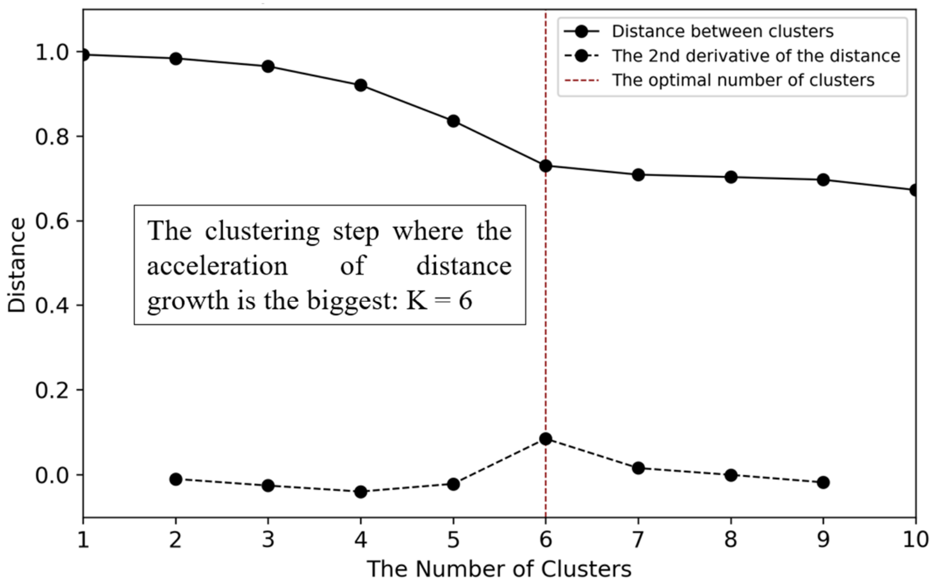


Fig. 3-4 The Optimal Number of Clusters (Elbow Method)

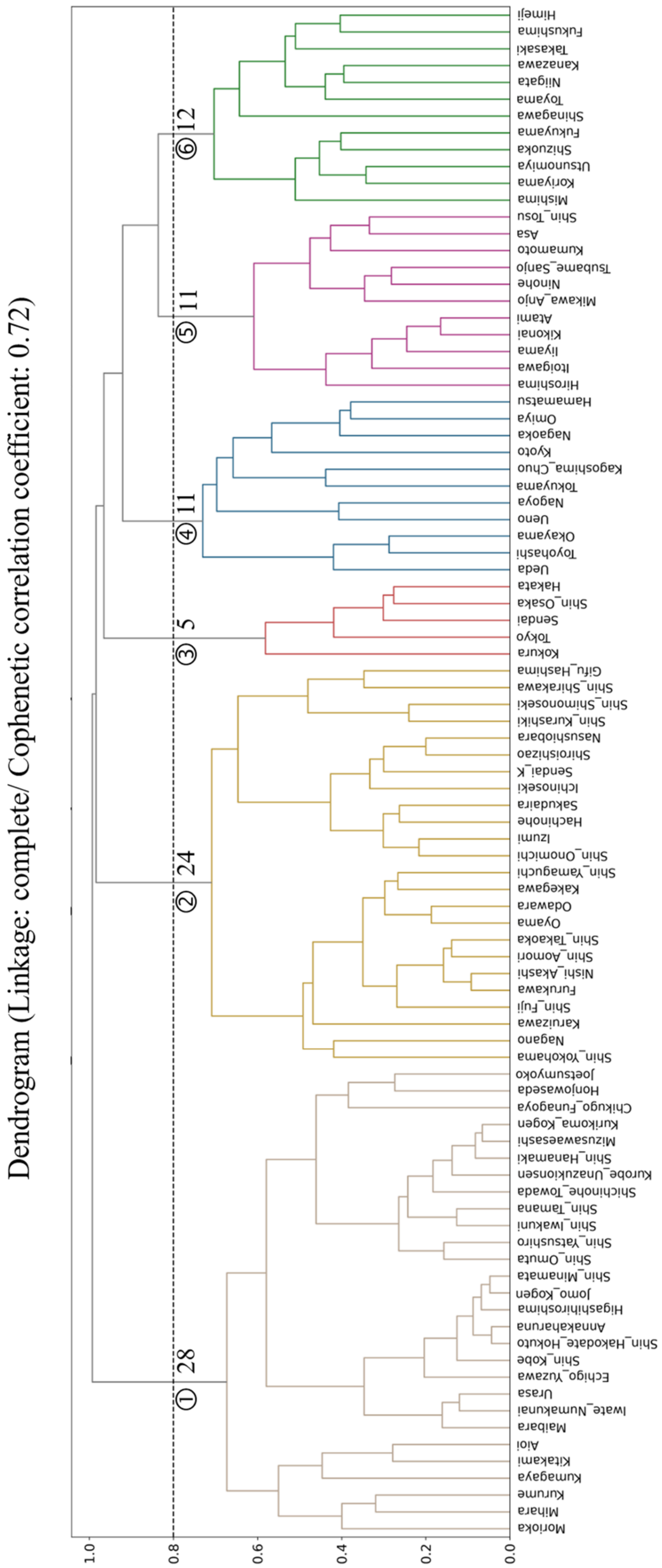
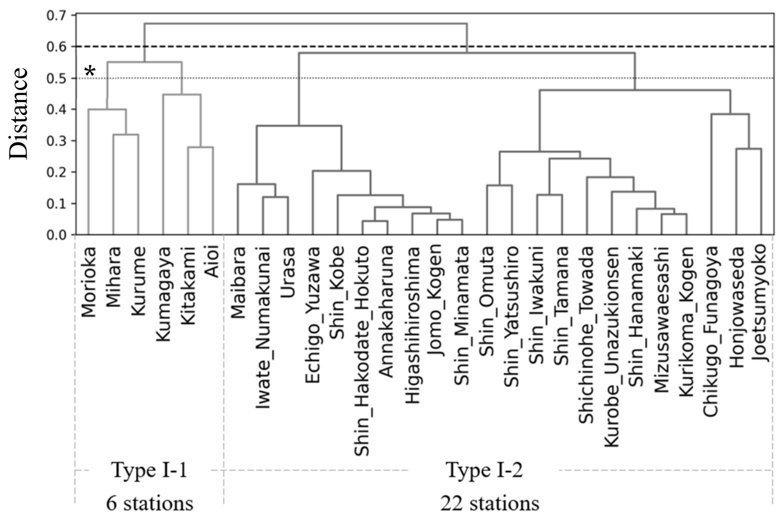


Fig. 3-5 Dendrogram for Land-use Pattern Clustering

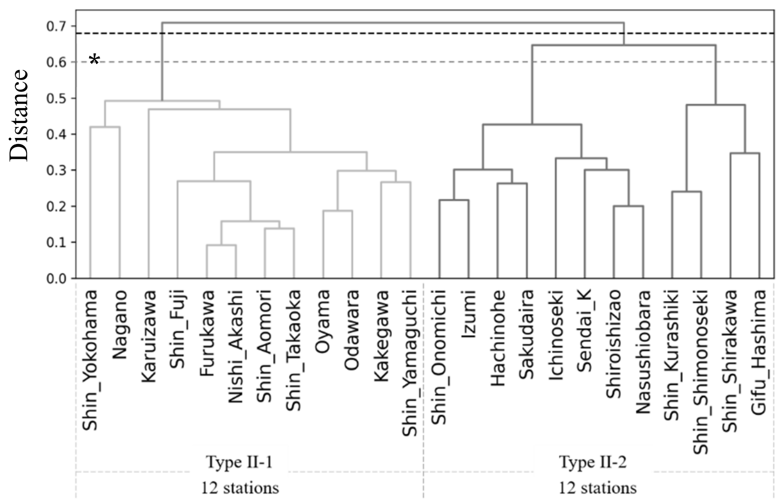
According to the classification results, the land-use patterns of 91 Shinkansen station areas were classified into 6 groups. However, it can be found that Type-I and Type-II are both composed of more than 20 stations, while other groups only contain about 10 members or fewer. Moreover, by observing the dendrogram of group 1 and group 2, clearly there exists intra-clusters in these two groups. Therefore, with integratively consideration of the number of members in every group and the inter-cluster distances, we furtherly classified Type-I and Type-II into 2 sub-types, as Fig.3-5 suggests.

Sub-clusters for Type-I



*When cut-off at the distance of 0.5, there will be 2 groups composed of only 3 members, it will be insufficient for their representativeness.

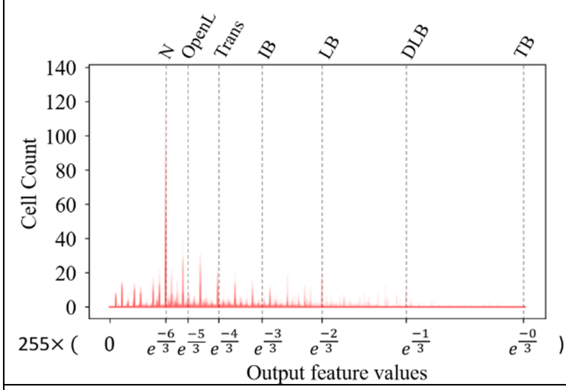
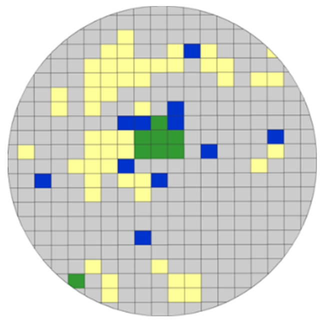
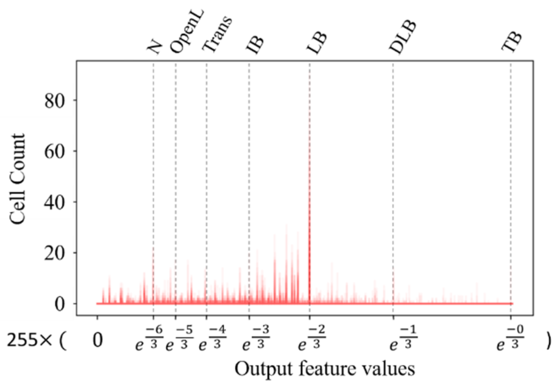
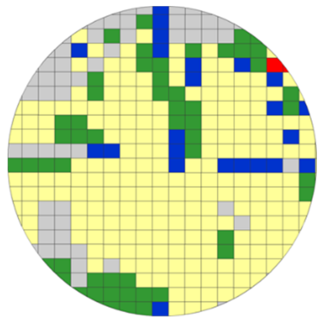
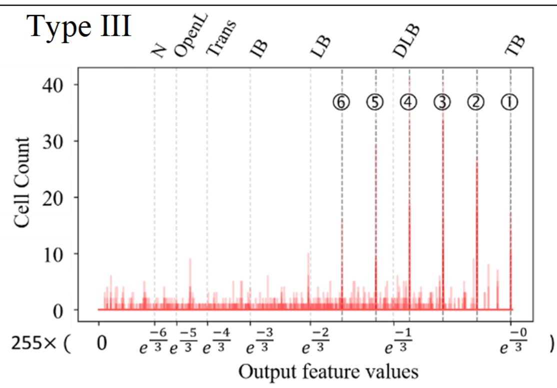
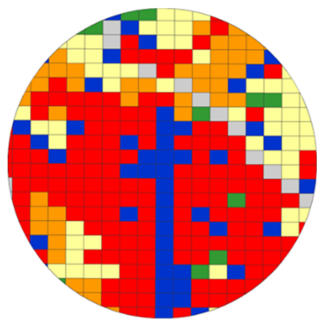
Sub-clusters for Type-II

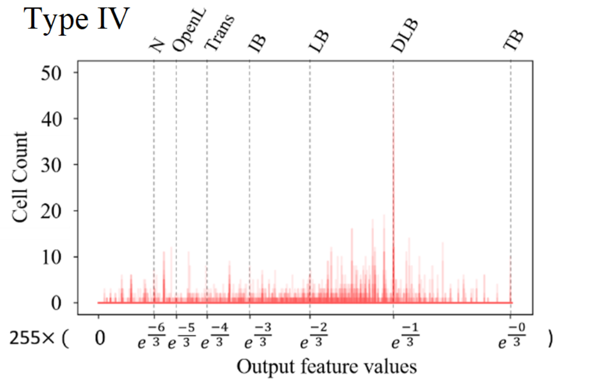
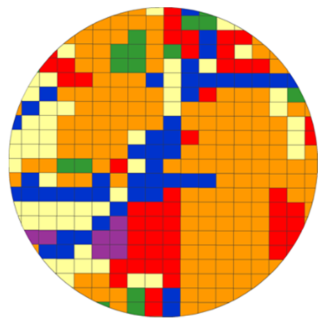
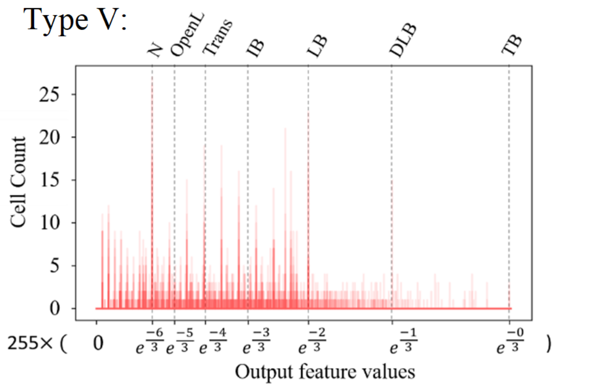
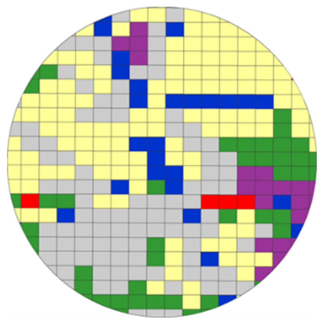
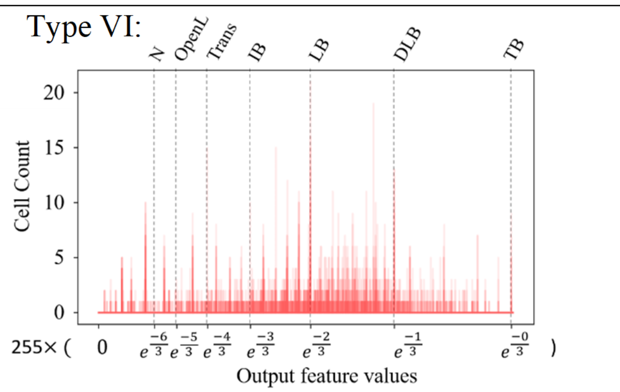
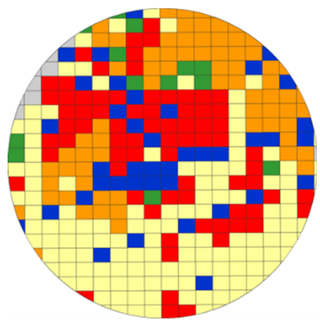


*When cut-off at the distance of 0.6, there will be a group composed of only 4 members, it might be an insufficient sample number.

Fig. 3-6 Identifying the Subtypes

Table 3-2 The Land-use Patterns Clustering Result

<table>
<tr><th>ID</th><th>Land use Pattern Features</th><th>Sample Stations</th></tr>
<tr><td rowspan="2">Type-I N-dominant Type</td><td>
</td><td>
Shin-Hanamaki</td></tr>
<tr><td colspan="2">Type I-1 (6 stations): Morioka, Mihara, Kurume, Kumagaya, Kitakami, Aioi
Type I-2 (22 stations): Maibara, Iwate-Numakunai, Urasa, Echigo-Yuzawa, Shin-Kobe, Shin-Hakodate-Hokuto, Annakaharuna, Higashihiroshima, Jomo-Kogen, Shin-Minamata, Shin-Omuta, Shin-Yatsushiro, Shin-Iwakuni, Shin-Tamana, Shichinohe-Towada, Kurobe-Unazukionsen, Shin-Hanamaki, Mizusawaesashi, Kurikoma-Kogen, Chikugo-Funagoya, Honjowaseda, Joetsumyoko</td></tr>
<tr><td rowspan="2">Type-II LB-dominant Type</td><td>
</td><td>
Shin-Aomori</td></tr>
<tr><td colspan="2">Type II-1 (12 stations): Oyama, Sakudaira, Shin-Takaoka, Shin-fujim Hachinohe, Nishi-Akashi, Nasushiobara, Odawara, Shin-Shirakawa, Shin-Yokohama, Kakegawa, Shin-Yamaguchi
Type II-2 (12 Stations): Shin-Onomichi, Izumi, Karuizawa, Gifu-Hashima, Shin-Aomori, Sendai(K), Ichinoseki, Shiroishizao, Shin-Kurashiki, Shin-Shimonoseki, Furukawa, Nagano</td></tr>
<tr><td rowspan="2">Type-III TB-dominant Type</td><td>
</td><td>
Hakata</td></tr>
<tr><td colspan="2">Type III (5 stations): Kokura, Tokyo, Sendai, Shin-Osaka, Hakata</td></tr>
</table>

ID	Land use Pattern Feature	Sample Stations
Type-IV DLB-dominant Type		 Okayama
	Type IV (11 stations): Ueda, Toyohashi, Okayama, Ueno, Nagoya, Tokuyama, Kagoshima-Chuo, Kyoto, Nagaoka, Omiya, Hamamatsu	
Type-V Low-intensity Intermixed Type		 Shin-Tosu
	Type V (11 Stations): Hiroshima, Itoigawa, Iiyama, Kikonai, Atami, Mikawa-Anjo, Ninohe, Tsubame-Sanjo, Kumamoto, Asa, Shin-Tosu	
Type-VI High-intensity Intermixed Type		 Niigata
	Type VI (12 stations): Mishima, Koriyama, Utsunomiya, Shizuoka, Fukuyama, Shinagawa, Toyama, Niigata, Kanazawa, Takasaki, Fukushima, Himeji	

Tall Building Land (TB) · Dense Low-rise Building Land (DLB) · Low-rise Building Land (LB)
Industrial Building Land (IB) · Transportation Land (Trans) · Open Land (OpenL)
Non-construction Land (N)

Table 3-3 The Land-use Pattern Features of Sub-type I-1 and I-2

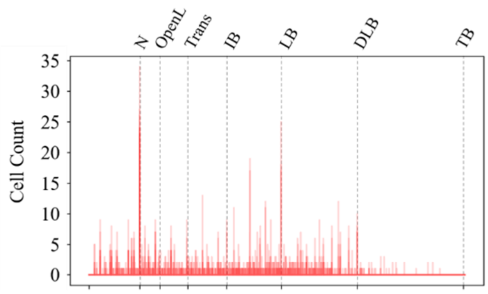
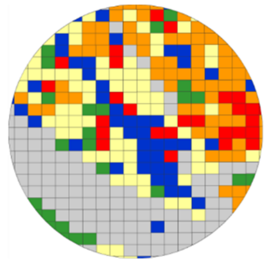
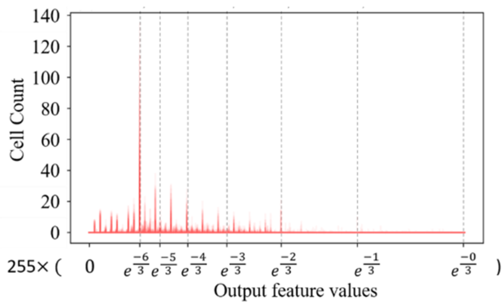
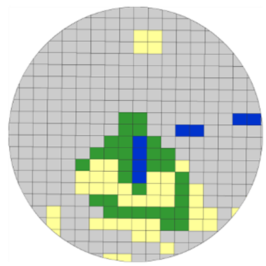
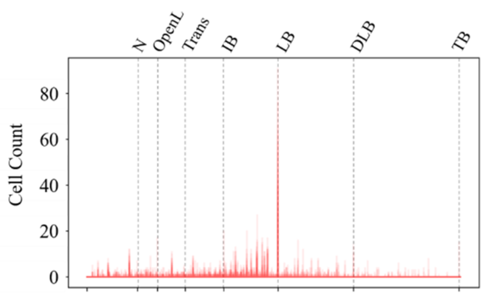
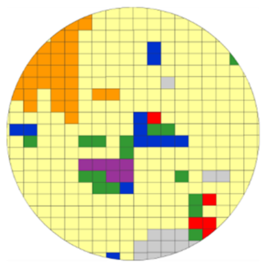
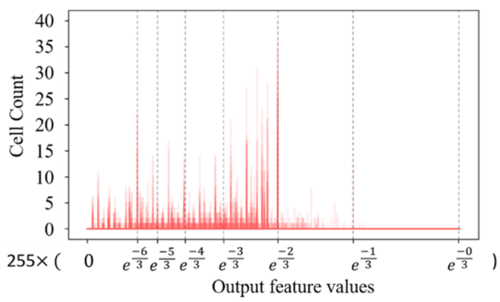
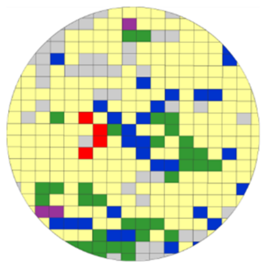
ID	Land-use Pattern Features	Sample Station Area
Type I-1		 Morioka
Type I-2		 Annakaharuna

Table 3-4 The Land-use Pattern Features of Sub-type II-1 and II-2

ID	Land-use Pattern Features	Sample Station Area
Type II-1		 Furukawa
Type II-2		 Gifu-Hashima

Tall Building Land (TB)   Dense Low-rise Building Land (DLB)   Low-rise Building Land (LB)

Industrial Building Land (IB)   Transportation Land (Trans)   Open Land (OpenL)

Non-construction Land (N)

3.3 Land Use Pattern Features of Shinkansen Station Areas

According to the hierarchical clustering, the land-use patterns of 91 existing Shinkansen station areas have been classified into 6 major types and 2 of which were furtherly classified into 2 subtypes, respectively. To clarify the land-use pattern features of each group, Table 3-2 displays the histograms of the output feature values to help visualize the land-use pattern feature descriptors of each group. Table-3-3 and Table 3-4 separated the histograms in Type-I and Type-II according to the subtypes. The histogram of every station area was given an opacity $alpha = \frac{1}{(the\ number\ of\ group\ members)}$. Take Type-I for example, there are 28 stations in total classified as Type-I, therefore, the histogram for every station in Type-I had an opacity of $1/_{28}$. In this way, when put all the histograms of the same group together, only the common parts of all group members will have 100% opacity. As a result, the land-use pattern features can be interpreted from the histograms of every group in Table 3-2. Specifically, we will interpret the land-use pattern features at the global level and local level respectively, and from the aspects of both land-use composition and spatial relationships.

3.3.1 Global Land-use Pattern Features

At the global level, the land-use pattern features include the overall composition and the distribution situation throughout the station area. Specifically, *composition* refers to the areal proportions of land uses, and distribution situation can be concluded by the *aggregation level* of each land-use category and the *spatial variations* of the location relationships between land uses.

In the histograms, the composition proportions are reflected by the skewness, the aggregation level of each land-use category is reflected by the count of the initialized land-use feature value, and the spatial variation is reflected by the number of unique output feature values, which can be observed from the fluctuations of the histogram. Accordingly, the global land-use pattern features of the 6 types of station areas can be interpreted as follows:

Type-I N-dominant type (28 stations): The histograms of Type-I have the single peak at the cell count of $255 \times e\hat{}\,(-2)$, which corresponds to the land-use feature value of N, and the output values concentrated around N. Moreover, there are barely any output feature values located between the initial feature values of LB and TB, indicating station areas of Type-I generally do not contain the TB and DLB. Therefore, Type-I can be concluded as the N-dominant type. In these station areas, the Non-construction land not only takes the most areal proportion but is also highly aggregated, forming large continuous unconstructed areas, whereas other land-use categories have much less occupancy and are dispersed in the station areas, as the sample station Shin-Hanamaki suggests.

Type-II LB-dominant type (24 stations): The histograms of Type-II have the single peak at the cell count of $255\times e\verb|^|((-2)/3)$, which is the initial land-use feature value of the Low-rise Building land. There are few output feature values located between the initial feature value of LB and TB. Therefore, Type-II can be named the LB-dominant type. Station areas in this group are characterized by a large continuous area of buildings of 3 stories or less, which are distributed collectively in a low-density form, whereas other land-use categories take much less areal proportions and are dispersed in the station area. Shin-Aomori station is a typical example of this group.

Type-III TB-dominant type (5 stations): The histograms of Type-III are distinct from other groups for it has significantly more cell counts at 255, indicating the land-use clusters of TB. Besides, if we trace the land-use combinations represented by the feature values of the peaks at ①~⑥, it can be found they represent land uses cells whose neighborhood is composed of TB: Trans = ①9:0, ②8:1, ③7:2, ④6:3, ⑤5:4, and ⑥4:5, respectively. Besides, comprehensively the histograms skew towards the initial land-use feature value of TB. Thus, the land-use pattern features of Type-III can be concluded as the TB-dominant type. In these station areas, Tall-building land takes the most areal occupancy. However, instead of forming a large continuous area, the TB cells intermixed with Trans cells. A typical example of Type-III is the Hakata station.

Type-IV DLB-dominant type (11 stations): The histograms of Type-IV differ from other groups for the significant single peak at the output feature value represent the cluster of DLB. Meanwhile, the majority of the output feature values distribute around the feature value of the DLB cluster. Therefore, Type-IV was named the DLB-dominant type. In station areas of this group, DLB cells take the most areal occupancy. They are highly aggregated, forming large continuous areas of land where buildings of 3 stories or less densely concentrated. In contrast, other land use categories take minor occupancy and distribute dispersedly in the station area. Okayama station is a typical example of Type-IV.

Type-V Low-intensity intermixed type (11 stations): the histograms of Type-V are distinct from type-I ~ IV for they have no significant single peak. Instead, the cell counts at the output feature values for N and LB are equally both slightly higher than other feature values. Besides, the majority of feature values distribute between [0, 255× e^((−2)/3)], and the histograms have more fluctuations than Type-I ~ IV. Accordingly, the pattern features of Type-V are concluded as the Low-intensity intermixed type. These station areas are mainly formed by the low-intensity land-use categories, i.e., N and LB cells. They clustered as small patches and intermixed, forming a fragmented pattern. A typical example of this group is the Shin-Tosu station.

Type-VI High-intensity intermixed type (12 stations): the histograms of Type-VI do not have a significant single peak either. However, in contrast to Type-V, the histograms of this group have small peaks at the output feature values represent the clusters of TB (255), DLB (255 × e^((−1)/3)), Trans (255 × e^((−2)/3)), indicating these station areas have relatively large number of TB cells, DLB cells, and LB cells. Besides, most output feature values distribute around the initial land-use feature values for DLB and LB, indicating these two land-use categories take relatively greater areal proportion than other land-use categories, suggesting a relatively high construction intensity. Also, the fluctuations of the histograms of this group indicate that the land-use cells in these station areas are intermixed, and the land-use pattern is fragmented. Therefore, Type-VI can be named as the High-intensity intermixed type. A typical representation of this group is the Niigata station.

3.3.2 The Differences between Subtypes

Within Type-I, the N-dominant type, subtype-I-1 and subtype-I-2 distinguishes with each other in the land-use composition. As Table 3-3 suggests, the peak value in the histograms of the output maps appears at e^{-2}, indicating both of them are characterized by large continuous unconstructed land. However, the cell count of the peak differs a lot, and the feature values for subtype I-1 spread to more values, indicating the composition proportion of Non-construction land in I-1 is lower than in I-2, and I-1 has higher spatial variation level than I-2.

As to the subtypes II-1 and II-2, according to Table 3-4, they are both characterized by the peak value at the feature values representing the Low-rise Building land. However, the specific peak value indicates that the Low-rise Building in II-2 has less occupancy and aggregation level. In other words, although both station areas of type II-1 and II-2 are dominated by Low-rise building land, in II-1 other land uses are also relatively aggregated, thus the Low-rise building land in II-1 forms a large continuous area. In contrast, in II-2, other land uses are scattered in the Low-rise building land area, therefore the Low-rise building land in II-2 looks more fragmented than in II-1.

3.3.3 Local Land-use Pattern Features

The local land-use pattern feature refers to the composition and distribution situation in the defined neighborhood. However, in this study, we have defined the $300\times300m^2$ neighborhood as a spatial indifference zone by employing the mean matrix in the smoothing operation. The local land-use pattern features here will only be interpreted as the land-use compositions in the neighborhood.

Since every different land-use combination in the 3×3 neighborhood corresponds to a unique output feature value, the neighborhood land-use composition situation can be traced back from the output feature values. Table 3-5 shows the 5 most frequent feature values in the output map and their corresponding land use combinations in the 3×3 neighborhood. Fig. 3-7 is the diagram for the 5 most frequent neighborhood land-use composition situations. As Table 3-5 suggests, the total percentage of the 5 most frequent feature values' cell count takes less than 30% of the total cell counts. However, considering the number of all possible land-use combinations is 11440, namely the theoretical probability of observing each feature value is only 1/11440, the first 5 feature values can still give an intuition of the land-use location relation in the station area.

Table 3-5 The First 5 Frequent Land Use Combinations in Land-use Cells' Neighborhood in Station Areas of Different Land-use Patterns

Type ID	Rank	Feature Values*	Percentage	Corresponding Land-use Combination							
				TB	DLB	LB	IB	Trans	OpenL	N	NaN
I	1	34.51049805	15.70%	0	0	0	0	0	0	9	0
	2	45.22281647	4.09%	0	0	1	0	0	0	8	0
	3	55.93513489	3.87%	0	0	2	0	0	0	7	0
	4	66.64745331	2.78%	0	0	3	0	0	0	6	0
	5	77.35977173	2.03%	0	0	4	0	0	0	5	0
Total Percentage			**28.48%**								
II	1	130.9213715	10.80%	0	0	9	0	0	0	0	0
	2	123.8431396	2.92%	0	0	8	0	1	0	0	0
	3	116.7649078	2.55%	0	0	7	0	2	0	0	0
	4	120.209053	2.55%	0	0	8	0	0	0	1	0
	5	109.4967346	2.22%	0	0	7	0	0	0	2	0
Total Percentage			**21.04%**								
III	1	234.1352539	6.16%	8	0	0	0	1	0	0	0
	2	213.2705078	6.06%	7	0	0	0	2	0	0	0
	3	192.4057617	5.35%	6	0	0	0	3	0	0	0
	4	171.5410156	3.38%	5	0	0	0	4	0	0	0
	5	255.0	2.84%	9	0	0	0	0	0	0	0
Total Percentage			**23.79%**								
IV	1	182.7154846	6.11%	0	9	0	0	0	0	0	0
	2	176.9605865	2.08%	0	8	1	0	0	0	0	0
	3	169.8823547	1.61%	0	8	0	0	1	0	0	0
	4	157.0492096	1.53%	0	7	0	0	2	0	0	0
	5	171.2056885	1.44%	0	7	2	0	0	0	0	0
Total Percentage			**12.77%**								
V	1	34.51049805	4.90%	0	0	0	0	0	0	9	0
	2	77.35977173	3.15%	0	0	4	0	0	0	5	0
	3	130.9213715	3.05%	0	0	9	0	0	0	0	0
	4	88.07209778	2.98%	0	0	5	0	0	0	4	0
	5	66.64745331	2.80%	0	0	3	0	0	0	6	0
Total Percentage			**16.88%**								
VI	1	130.9213715	2.38%	0	0	9	0	0	0	0	0
	2	182.7154846	1.74%	0	9	0	0	0	0	0	0
	3	123.8431396	1.54%	0	0	8	0	1	0	0	0
	4	116.7649078	1.15%	0	0	7	0	2	0	0	0
	5	169.8823547	0.91%	0	8	0	0	1	0	0	0
Total Percentage			**7.72%**								

* When computing the percentage, the cell count for the feature value of 0 is not included, because it represents background area. Besides, when ordering the cell counts for the feature values, the feature values whose corresponding land use combination contains 0 are excluded, because they reflect much information of the image background (blank area) and limited information about the land uses.

*Except 0 and 255, the output feature values are all irrational numbers, for the convenience of expression, here displays the approximate values.

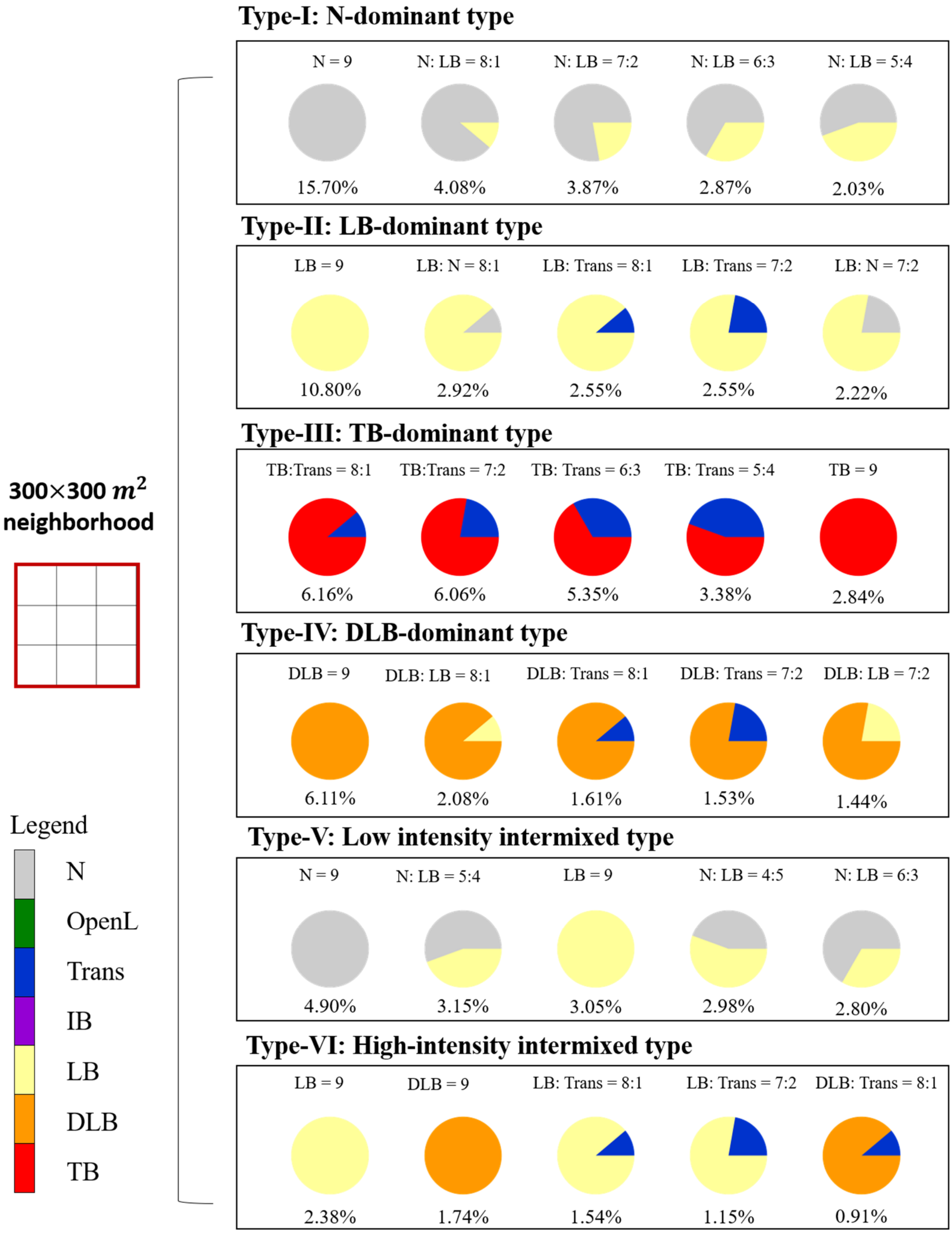


Fig. 3-7 Diagram of the First 5 Frequent Land Use Contiguity Relation at Defined Neighborhood Level in Different Land-use Patterns

Accordingly, it can be found in station areas of Type-I, the N-dominant type, the most frequent feature value corresponds to the situation where the 3×3 neighborhood is composed of only the N cells, and its percentage of 15.7% suggests a significant pattern feature. Besides, the first five frequent neighborhood land-use combinations suggested that Type-I's major spatial contiguity relation at the neighborhood level is between N cells and LB cells, where N cells usually take more than 50% of the areal occupancy. For station areas of Type-II, the LB-dominant type, the 5 most frequent land-use combinations in the 3×3 neighborhood are composed of LB cells, N cells, and Trans cells. Nevertheless, in either case, N cells and Trans cells take less than 2/9 of the neighborhood. For station areas of Type-III, the TB-dominant type, there is no single feature value concentrated significantly more cells than others. Instead, the first 5 frequent feature values correspond to the land-use combinations of TB cells and Trans cells in the 3×3 neighborhood, indicating the major contiguity relationship in these station areas is between TB and Trans. In the station areas of Type IV, the DLB-dominant type, the 5 most frequent feature values correspond to the neighborhood land-use combinations of DLB, LB, and Trans, where the DLB cells are always dominant in the neighborhood. In contrast, in station areas of Type-V and Type-VI, none feature value has a significantly higher percentage than others, which corresponds to the global features of higher spatial variation and a fragmented pattern in these station areas. However, we can still observe that the first 5 frequent neighborhood land-use combinations in Type-V are composed of the N and LB cells, while in Type-VI are the LB and DLB cells. These reflected the difference in the construction intensity between Type-V and Type-VI.

We can find consistency between the local pattern features and the global pattern features. The local features give an insight into the location relation of land uses in the defined neighborhood (300×300m^2), specifically, the contiguity relationship between different land use categories. Takes Type-I for example, with 15.7% of land use cells' neighborhood is composed of Non-construction land, and almost 30% of land use cells' neighborhood is formed by N and LB cells, among which the occupancy of LB cells is always less than 50%, when assembling all this neighborhood pattern together, we will get a station area where the Non-construction land covers most of the station area while small patches of Low-rise

Building land scatters in the non-construction area. Similarly, for other types, we can depict the land-use pattern feature of the whole station area by assembling the neighborhood land-use combinations together. Therefore, the local pattern features help us understand how different land-use categories are intermixed and how land uses are arranged to form the land-use pattern of the whole station area.

3.4 The Trend of land-use Pattern of Station Areas in Cities of Different Population Scales

So far, we have classified the land-use patterns of 91 existing Shinkansen station areas into 6 types and clarified the land-use pattern features of each type. The land-use pattern features of each group were summarized in Table 3-6. And to give a better intuition of the construction situation in each type of station area, Table 3-6 displays the aerial photography of the example stations. Then, this section is going to detect the relation between land-use patterns and the property of the station cities.

Fig. 3-8 shows the locations of the land-use pattern types. To further look into the trend of the station areas' land-use patterns in different cities, we classified the Shinkansen station cities into 4 groups according to the city population because the population scale usually corresponds to the integrative city development potential and dynamics. Specifically, small cities are cities whose population is under 100 thousand, mid-sized cities are cities whose population is between 100 and 300 thousand, core cities are cities whose population is over 300 thousand, and the government ordinance designated cities (政令指定都市) are cities that have a population greater than 500 thousand and have been designated as such by order of the Cabinet of Japan according to the Local Autonomy Law (地方自治法). Figure 3-9 summarized the composition of the city's population scale in each land-use pattern type group. The trend of land-use patterns of station areas in cities with different population scales can be concluded as follow:

Type-I N-dominant type (28 stations) is characterized by large continuous unconstructed land. Among all station areas of land-use pattern Type-I, 92.8%

are located in small cities or mid-sized cities, with the small cities takes 64.2%and mid-sized cities takes 28.6%. Namely, the N-dominant type is mainly observed in cities whose population is less than 300 thousand. Particularly the intermediate stations whose Shinkansen were recently constructed, such as intermediate stations along Kyushu Shinkansen Line. In these station areas, the formation of the large unconstructed area has various causes. Some are related to the development stages, such as station areas in the initial development period, especially the newly constructed suburb stations along recently extended Shinkansen lines like Shin-Omuta stations and Shin-Hakodate-Hokuto stations, etc. Also, there are suburb station areas whose development has stagnated for years, such as Shin-Hanamaki station, Higashihiroshima station, etc. Besides, there are also major city stations in this group, such as Morioka station and Shin-Kobe station. These stations are located near great rivers or mountains. In such station areas, unconstructed land is formed due to the topographic limitation.

Type-II LB-dominant type (24 stations) is mainly identified in station areas in cities with a population of less than 300 thousand, particularly in cities with a population of between 100-300 thousand. According to Fig. 3-9, 87.5% of the station areas in the Type-II group locate in small cities or mid-size cities, among which the small city takes 29.1% and the mid-sized city takes 58.3%. Among station areas of Type-II, 12 station areas where the Low-rise Building land aggregated together and formed a large continuous area of buildings that are 3 stories or less were identified as Subtype II-1. In contrast, the other 12 station areas where the small patches of other land use scattered in the aggregated low-rise building area, breaking the low-rise building areas into fragmented clusters were identified as Subtype II-2. Fig. 3-9 suggests that both the station areas of Subtype II-1 and II-2 are intermediate cities and mostly locates along the Tokaido Line, Sanyo Line, and Tohoku Line.

Type-III TB-dominant type (5 stations): are all located in government ordinance designated cities and typically the central cities of major metropolitan areas. Such station areas are characterized by high-intensity land development and remarkable transportation property, usually serving as centers for aggregating city functions and the national-level transportation hubs.

Type-IV DLB-dominant type (11 stations): is marked for the aggregated Dense Low-rise Building land whose areal occupancy dominates the whole station area. Fig.3-9 suggests that all the station areas of Type-IV locate in cities whose population is over 100 thousand, and more than half are government ordinance-designated cities. Moreover, it can be observed from Fig.3-8 that the majority of station areas in the Type-IV group locates in cities along Tokaido Line and Sanyo Line, which had relatively long development history. Besides, many of the station areas in Type-IV locates in important region center cities such as Okayama and Hamamatsu.

Type-V Low-intensity intermixed type (11 stations): is a fragmented pattern formed by the Non-constructed land and Low-rise Building land mixing with other land use categories. In these areas, the Non-constructed land, Low-rise Building land, and other land have composition proportions approximate to 1: 1: 1, there is no single dominant land-use category. 72% of these station areas locate in cities whose population is less than 100 thousand. These station areas are either recently constructed stations in suburb areas like Shin-Tosu station, Tsubame-Sanjo station, or station areas limited by the topography or constructions in the built-up areas, such as Atami station and Kumamoto station.

Type-VI High-intensity intermixed type (12 stations): is mainly observed in station areas of cities that have a population of over 100 thousand, particularly cities whose population is over 300 thousand. In these station areas, Tall-building land, Dense Low-rise Building land, Low-rise Building land cells aggregate as small patches, intermixing together, forming some discernable land-use zones. These station areas are mainly located in the built-up areas of regional center cities, with a relatively long development history. However, the urban redevelopment of these areas is generally less dynamic than the station areas of Type-III and usually have minor scales. The different land use zones in the station areas reflect different stages of the redevelopment process in these station areas.

3.5 Summary

In conclusion, this chapter classified the land-use patterns of 91 existing Shinkansen station areas into 6 types based on the land-use pattern comparison method we have developed in Chapter 2 and then clarified the land-use pattern features including the land use composition proportions, aggregation level, and spatial variation properties at the global level, and the major contiguity relationships on the local level. Besides, this chapter discussed the relation between the land-use pattern and the population scales of the station cities. As a result, it became clear that the land-use pattern of a station area generally correlates to the population scale of the station city.

According to the clustering result, the construction situation around the existing Shinkansen station becomes clear. In contrast to the previous existing clustering approaches, which tend to classify station areas based on the land-use composition proportions, development intensity, station scales, or station cities properties, the image matching approach identifies the land-use pattern based on the location situation of land-use cells in the station areas. Therefore, the clustering result reveals the feature of land-use patterns regardless of the specific spatial process or the underlying influential factors. For example, we can observe that the station areas of Type-I, the N-dominant type, include station areas that are newly developed, station areas whose development has stagnated for years due to their isolated locations, and also major city stations whose land development is limited by the topography. According to the clustering result and the background information of the station areas, we can find due to various influential factors, these station areas with different properties formed the same land-use pattern in the time slice we observed.

However, the land-use pattern also influenced by the development stage of the station area due to various construction time of Shinkansen. The next chapter will discuss the dynamic land use transformation situation in the station areas to help deeper understand to land development situation around Shinkansen stations.

Table 3-6 Summary of the Land-use Pattern Features of Shinkansen Station Areas

Typical Sample Station Area								
			Shin-Hanamaki	Shin-Aomori	Hakata	Okayama	Shin-Tosu	Niigata
Features / Type			I	II	III	IV	V	VI
Global Level	Composition Situation	Dominated land use (near or over 50%)	Non-Construction Land	Low-rise Building Land	Tall Building Land	Dense Low-rise Building Land	-	-
		Composition Proportion Relation (Average)	N: LB: Others≈ 3:1:1	LB: N: Others≈ 5:2:3	TB: Trans: Others ≈ 5:2:3	DLB + TB + LB + Trans ≈ 90%	LB: N: Others≈ 1:1:1	LB + DLB ≈ 50% TB ≈ Trans ≈ 15%
	Aggregation Situation		N: aggregated Others: dispersed	LB: aggregated Others: dispersed	TB: aggregated but not continuously	DLB: aggregated but not continuously	LB, N: aggregated as small patches Others: dispersed	TB,DLB,LB: small patches Others dispersed
	Spatial Variation Level (Relative)		homogeneous	homogeneous	heterogenous	heterogenous	heterogenous	heterogenous
Local Level	Major Contiguity Relation		N + N N + LB (LB $< \frac{1}{2}$)	LB+ LB LB + N (N $< \frac{1}{4}$) LB + Trans (Trans $< \frac{1}{4}$)	TB + Trans TB + TB	DLB +DLB DLB + LB (LB $\leq \frac{1}{4}$) DLB + Trans (Trans $\leq \frac{1}{4}$)	N + LB N + N LB + LB	-

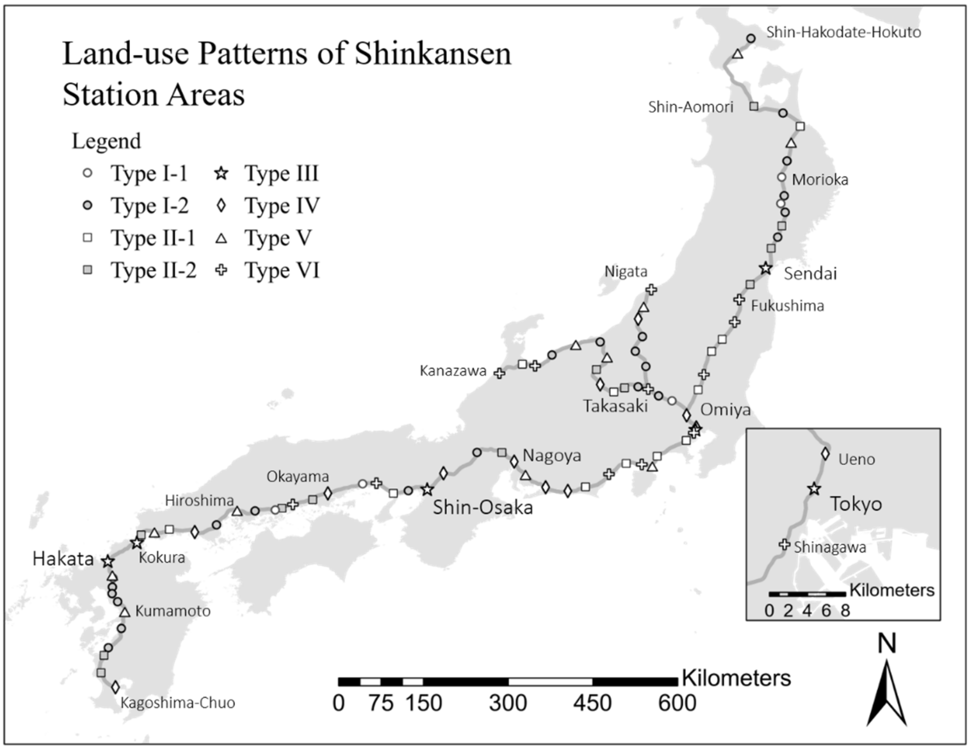
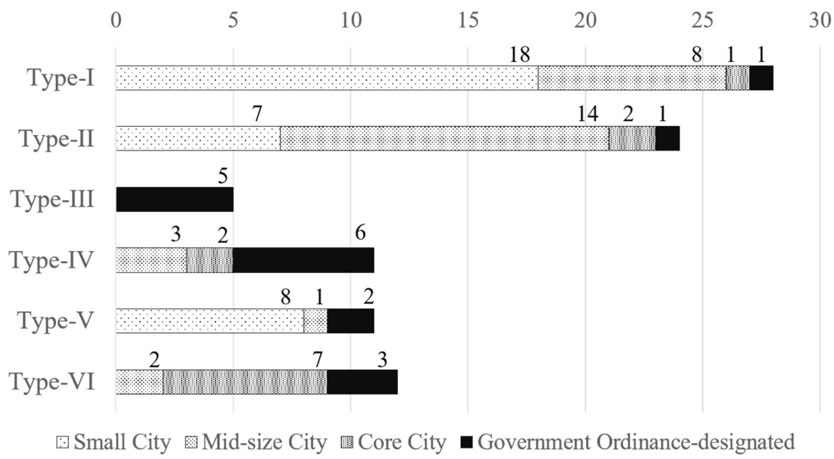


Fig. 3-8 Locations of Different Land-use Pattern Types



* Small city: population less than 100 thousand * Mid-size city: population between 100-300 thousand * Core city: population more than 300 thousand * Government ordinance-designated city: population greater than 500 thousand and has been designated as such by order of the Cabinet of Japan under the Local Autonomy Law

Fig. 3-9 The Trend of Land-use Pattern of Shinkansen Station Areas in Cities of Different Population Scales

Notes

*1) Okutsugaru-Imabetsu station is located in a rural area, thus it is not included in the urban land-use subdivision mesh data published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism.

*2) According to the standard digital map 100-meter grid defined by the Geographic Institution (GSI), the 1000-meter range around a station is not exactly 20×20 land-use cells. The land-use cells on the boundary will be cut-off. Therefore, when converting the land-use maps to images, the image needs to be resampled to make the station area composed of 20×20 cells. The resampling is based on the nearest neighbor interpolation (CV2: INTER_NEAREST_EXACT).

Chapter 4

Land Use Transformation in Shinkansen Station Areas

This chapter focuses on the land use transformation situation of Shinkansen station areas during 2009 and 2016. **Section 4.1** introduces the study flow and method. **Section 4.2** summarizes the overall land-use change situations in 91 existing Shinkansen station areas. Then, with the land-use pattern comparison method developed in chapter 2, **Section 4.3** compares the land-use patterns of Shinkansen station areas over time, identifying the station areas with significant land use transformation. **Section 4.4** analyzes the location trend of the land use transformation around stations, clarifies the transformation situation in the Shinkansen station areas of the cities with different population scales. **Section 4.5** summarizes this chapter.

4.1 The Study Objective and Method

This chapter aimed to make clear the dynamic land use transformation situation around Shinkansen stations during 2009 and 2016 *1). The source data is the 100-meter land use subdivision mesh data of urban areas of 2009 and 2016 published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism. Consistent to the land use pattern analysis in Chapter 3, this chapter also focused on the 1000-meter radius range around 91 existing Shinkansen stations, as Fig.3-1 suggests. The land-use information of 2009 and 2016 are unified as Table 3-1 shows.

Instead of focusing on one or some specific station areas, this study aimed to provide a comprehensive perspective to the dynamic land development situation around Shinkansen stations. This chapter discusses the land use transformation situation from 3 levels of perspectives: the national level, the regional level, and the local level. The study structure and method are shown in Fig 4-1.

At the national level, we focused on the overall land-use changes of 91 existing Shinkansen station areas, based on the statistical summary of the land-use composition of 2009 and 2016 and statistical hypothesis test, we testified whether and how the land uses of Shinkansen station areas has significantly changed during 2009 and 2016.

At the regional level, the goal is to clarify where is the land use transformation, i.e., to identify which station areas developed more dynamically, and which station areas' land development remains stagnated. Specifically, we will employ the land-use pattern comparison method developed in Chapter 2 to evaluate the level of land-use changes in the station area. According to the proposed land-use comparison method, we can get the similarity level $sim(X_{09}, X_{16})$ of a station area's land-use patterns at different times. The similarity level ranges in [-1, 1], and 1 means the land-use patterns are exactly the same. Therefore, the level of land-use pattern changes can be computed by $1 - sim(X_{09}, X_{16})$. Accordingly, we discussed the relation between the land-use pattern change level and the hierarchy of station cities.

At the local level, we concentrated on the dynamically developed areas, aimed to clarify the location trends of the land development around stations.

Specifically, we focused on the central tendency, circumferential distribution trend, and radial distribution trend of the land development relative to the station. With the geographic distribution analysis, we examined whether the land development of each land-use category differs in the different orientations of the station and whether the land development of each land-use category varies with the distance to the station. Besides, we also analyzed location trends of land development in station areas of different city groups, thus discussing the spatial pattern of the dynamic land development in different cities.

This part aimed to provide a comprehensive insight into the dynamic land development situation of the Shinkansen station areas. According to analysis from a multilevel perspective, the relation between the land development of Shinkansen station areas and the regional development will become clearer.

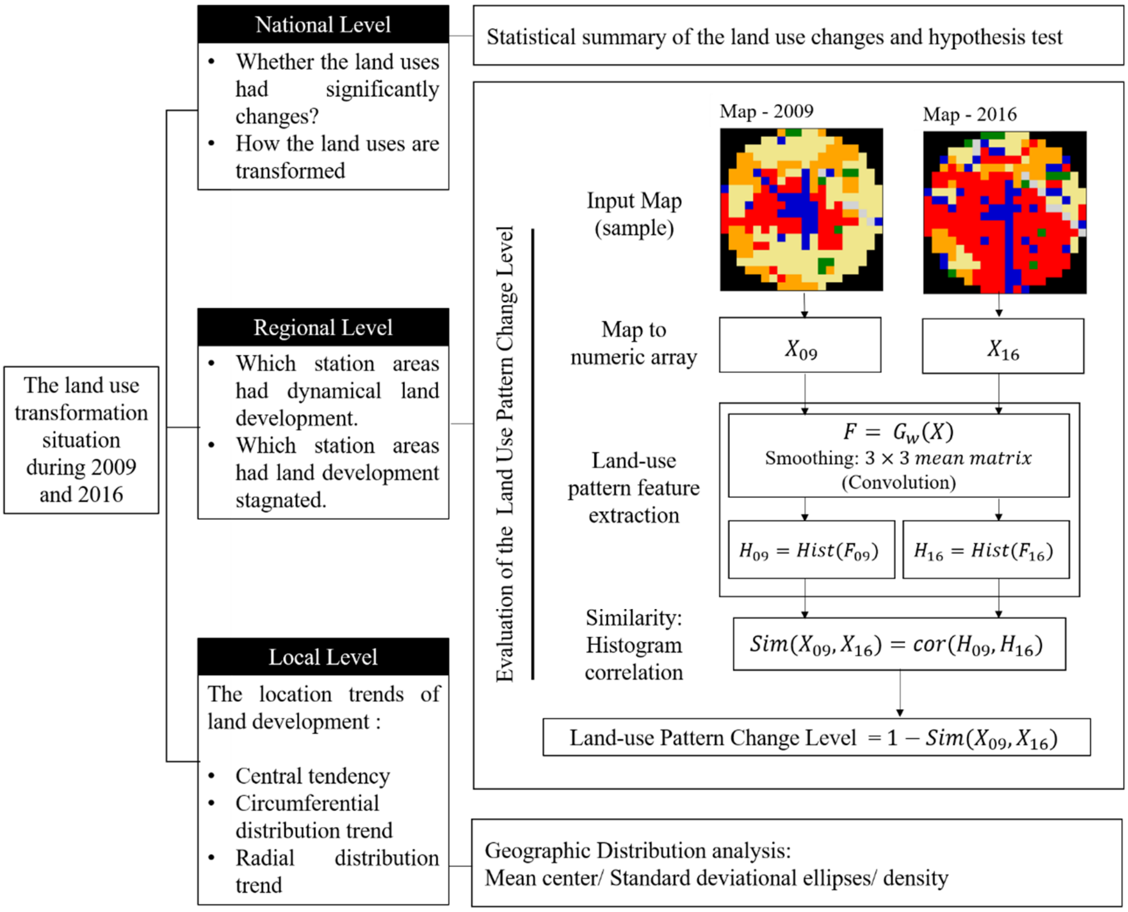


Fig. 4-1 Study Flow and Method

4.2 The Overall Land Use Transformation Situation of Shinkansen Station Areas

Table 4-1 summarized the overall land-use composition change of 91 existing Shinkansen station areas. To test whether the changes are significant, we need to conduct a statistical hypothesis test. Before the statistical hypothesis test, we conducted the Shapiro-Wilk normality test [81] to the composition proportion of each land use category at 91 station areas, to help determine the appropriate statistical hypothesis test method. As a result, the p-values of the Shapiro-Wilk normality test for all the land use categories are less than 0.05, indicating the distribution of the land-use composition proportions is significantly different from the normal distribution. Therefore, we use the Wilcoxon signed-rank test [82] to examine whether the differences in the land-use composition proportions are statistically significant.

In consequence, as Table 4-1 suggests, except for the Industrial Building land and the Open land, the p-values of the Wilcoxon test for the composition proportions of other land use categories are all less than 0.001, suggesting the mean of the composition proportions in 2016 are statistically significantly different with the composition proportions in 2009. It can be found that the average composition proportions of the Tall Building land and the Transportation land increased 3.66% and 5.41%, respectively. In contrast, the Dense Low-rise Building land, Low-rise Building land, and Non-construction land decreased by 2.87%, 4.34%, and 1.85%, respectively, indicating that comprehensively the land development in 91 existing Shinkansen station areas were becoming more intensive.

Specifically, Table 4-2 summarized the overall land-use transformation situation in 91 existing Shinkansen station areas with a land-use transition matrix. In the land-use transition matrix, the row index r is the land-use category of 2009, the column index c is the land-use category of 2016. The cell value of actual area (r, c) represents the amount of land uses of the corresponding transformation situation from r to c, and the probability suggests the proportion of the land-use area of such transformation situation occupies the total area of corresponding land-use category. For example, 1031.67 in (TB, TB) represents that 1031.67 ha

Tall Building land remained unchanged during 2009 and 2016, which is 72.56% of the total area of Tall Building land in 2009, also 486.53 in (DB, TB) represents that 485.53 ha Dense Low-rise Building land of 2009 has become Tall Building land in 2016, which is about 11.01% of the total Dense Low-rise Building land in 2009. Accordingly, the overall land-use transformation situation of 91 existing Shinkansen station areas can be interpreted from Table 4-2.

In Table 4-2, we are particularly interested in the land use that has changed. As to Tall Building land, most transition is to Transportation land, indicating that the development in Tall Building land area was mostly for increasing road density or traffic squares thus improving the quality of traffic system in the station area. Similarly, for the Dense Low-rise Building land and Low-rise Building land, the transition to Transportation land takes 8.41% and 7.45% of the total respectively, implying the land readjustment undergoing in the station areas. Also, the major transition from Dense Low-rise Building land and Low-rise Building land to the Tall Building land indicates the increase of the overall land development intensity in station areas. Besides, it can be observed the transition from the Non-construction land to Low-rise Building land and transportation land, this mostly indicates the land development at the newly opened station areas in the suburb areas.

Comprehensively, it can be concluded that during the year of 2009 and 2016, the land development in the Shinkansen station areas has progressed significantly. Concretely, the land development in the Shinkansen station areas is generally going through the process from Non-construction land to construction land, from low-rise to high rise, and the land development has the characteristic of improving the intensity and traffic quality. However, the overall situation cannot represent the land development at every individual stations. In next section, we are going to clarify where did the transformations happened.

Table 4-1 Overall Land Use Change in 91 Existing Shinkansen Station Areas during 2009 and 2016

Land-use	2009		2016		Δmean.	P((Wilcoxon test))	
	mean.	sd.	mean.	sd.			
TB	5.01%	10.83%	8.67%	13.38%	3.66%	<0.001	****
DLB	15.58%	19.25%	12.71%	16.26%	-2.87%	<0.001	****
LB	38.07%	19.07%	33.72%	17.68%	-4.34%	<0.001	****
IB	1.97%	4.22%	1.62%	3.43%	-0.34%	0.042	*
Trans	3.93%	3.32%	9.34%	5.35%	5.41%	<0.001	****
OpenL	5.75%	5.19%	6.09%	4.57%	0.33%	0.027	*
N	29.69%	26.83%	27.85%	24.97%	-1.85%	<0.001	****

Table 4-2 Overall Land Use Transformation Situation in 91 Existing Shinkansen Station Areas

Land-use Transition Matrix ($\frac{\text{Actual Area: ha}}{\text{Probability}}$)

		2016							Total	Reduce
	Category	TB	DLB	LB	IB	Trans	OpenL	N		
2009	TB	1031.67	30.50	56.18	3.16	244.19	35.22	20.82	1421.74	390.07
		0.7256	0.0215	0.0395	0.0022	0.1718	0.0248	0.0146	1.0000	0.2744
	DLB	486.53	3196.93	254.08	1.06	371.95	64.79	44.85	4420.19	1223.27
		0.1101	0.7233	0.0575	0.0002	0.0841	0.0147	0.0101	1.0000	0.2767
	LB	757.50	255.43	7969.73	33.15	800.42	333.70	597.42	10747.34	2777.62
		0.0705	0.0238	0.7416	0.0031	0.0745	0.0310	0.0556	1.0000	0.2584
	IB	6.42	9.07	42.82	381.03	41.79	53.68	17.72	552.53	171.50
		0.0116	0.0164	0.0775	0.6896	0.0756	0.0972	0.0321	1.0000	0.3104
	Trans	78.41	51.23	113.67	6.30	782.80	35.79	40.03	1108.24	325.44
		0.0708	0.0462	0.1026	0.0057	0.7063	0.0323	0.0361	1.0000	0.2937
	OpenL	69.25	34.53	248.51	18.40	120.65	966.61	167.93	1625.88	659.27
		0.0426	0.0212	0.1528	0.0113	0.0742	0.5945	0.1033	1.0000	0.4055
	N	26.86	29.06	833.89	13.48	281.88	229.20	6955.70	8370.07	1414.37
		0.0032	0.0035	0.0996	0.0016	0.0337	0.0274	0.8310	1.0000	0.1690
	Total	2456.65	3606.75	9518.88	456.57	2643.68	1719.00	7844.47		
		1.0344	0.8558	1.2711	0.7138	1.2202	0.8218	1.0829		
	Increase	1424.98	409.82	1549.15	75.55	1860.88	752.39	888.77		
		0.3087	0.1326	0.5295	0.0242	0.5139	0.2273	0.2519		

4.3 The Land-use Pattern Change Level of Shinkansen Station Areas

To identify where did the land use transformation happened, we computed the similarity level between each station area's land-use patterns of 2009 and 2016 based on the image matching based method we have developed in Chapter 2. And the level of the land-use pattern changes can be evaluated by $1 - similarity$. Consequently, the average of the land-use pattern change level is about 0.16, however, 54 station areas' land-use pattern change level is under the average. Indicating the majority of the land development concentrated in a small part of stations. Fig 4-2 illustrated the land-use pattern change levels of 91 existing stations areas. Accordingly, it can be found that the dynamic transformation is generally concentrated around stations located in the major cities, most are the regional core cities or the government ordinance-designated cities.

To examine whether the dynamic land development situation differs in station areas of different city hierarchy, Fig.4-3 summarized the land-use pattern change level of station areas according to the station city scale. As Fig. 4-3 suggests, apart from the government ordinance-designated cities (政令指定都市), the cities are classified into 4 ranks according to the city population: small cities refer to cities whose population is under 100 thousand, mid-sized cities refer to cities whose population is between 100 and 300 thousand, and core cities are cities whose population is over 300 thousand but not designated cities. As a result, the mean land-use pattern changes level of the small cities and the mid-sized cities are 0.097 and 0.121 respectively, and the mean land-use pattern change level of core cities and designated cities are 0.258 and 0.267. To test whether the differences between different city groups are statistically significant, statistical hypothesis test is needed. Firstly, the Shapiro-Wilk normality test was implemented to check whether the land-use pattern level is normally distributed, which helps determine the statistical hypothesis test method. Because the p-value of the Shapiro-Wilk test is much smaller than 0.001, indicating the land-use pattern levels' distribution is significantly different from the normal distribution, we implemented the Kruskal-Wallis test [82] to the differences of the mean change levels between station areas of different cities groups. The result suggests the mean change level between the station areas of small cities and the mid-sized cities are not statistically different, and the difference between the core cities and designated cities are not statistically different, while the differences between other cities groups are significant.

Therefore, it can be concluded that on the regional level, the dynamic land development situation around Shinkansen stations generally corresponds to the development trends of the station city: the major cities in the national city system often have more developing dynamics, the station areas also have more significant changes, while the land development of station areas in small cities and mid-sized cities generally have much less progress. Particularly, it can be observed in Fig.4-2, the stations located in or near the Tokyo metropolitans and the Kitakyushu-Fukuoka metropolis had significantly drastic changes, which is consistent with the city development in these regions. For the stations in the Tokyo metropolitans, located in the most important city region of the whole country, their development was driven by the inexhaustible development dynamics of the city. Especially, influenced by the coming 2020 Tokyo Olympics, a large number of urban redevelopment projects were implemented in the Metropolitan of Tokyo, and the Shinkansen station areas were of course included. As to the Kitakyushu-Fukuoka metropolitan area, although it is far away from the national economic center, Kanto area, as the most important metropolitan area in the Kyushu island, it has the definite attraction and influence on the whole Kyushu area. The fact that Fukuoka Prefecture is the prefecture in Kyushu region which had a positive population growth according to the report of national census of 2015 and the population growth ratio of the Fukuoka cities ranks the first among all government ordinance-designated cities demonstrated the development dynamics of the station city. [83] In related to the recently opened Kyushu Shinkansen (Hakata-Kagoshima section) and currently constructing Nagasaki section, Fukuoka city’s role as the regional center has been largely strengthened. The development around Hakata station as the transportation hub of a regional center city seems natural. However, as to the Tohoku region, it can be found cities northern than Sendai has little changes. Similarly, to the stations along the Hokuriku Shinkansen (Takasaki-Kanazawa section). Although it has been reported that the opening of Shinkansen has positively benefited the tourism of these region, nevertheless, it is a short-term influence. [21] Study and trials are needed to find appropriate development plans for these regions.

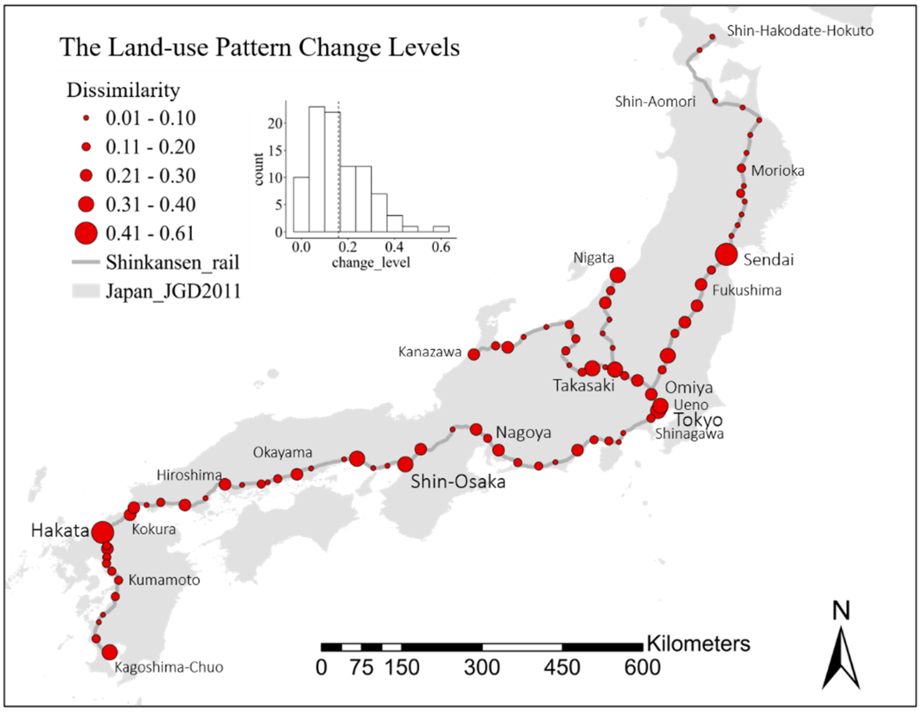


Fig. 4-2 The Level of Land-use Pattern Change of Shinkansen Station Areas

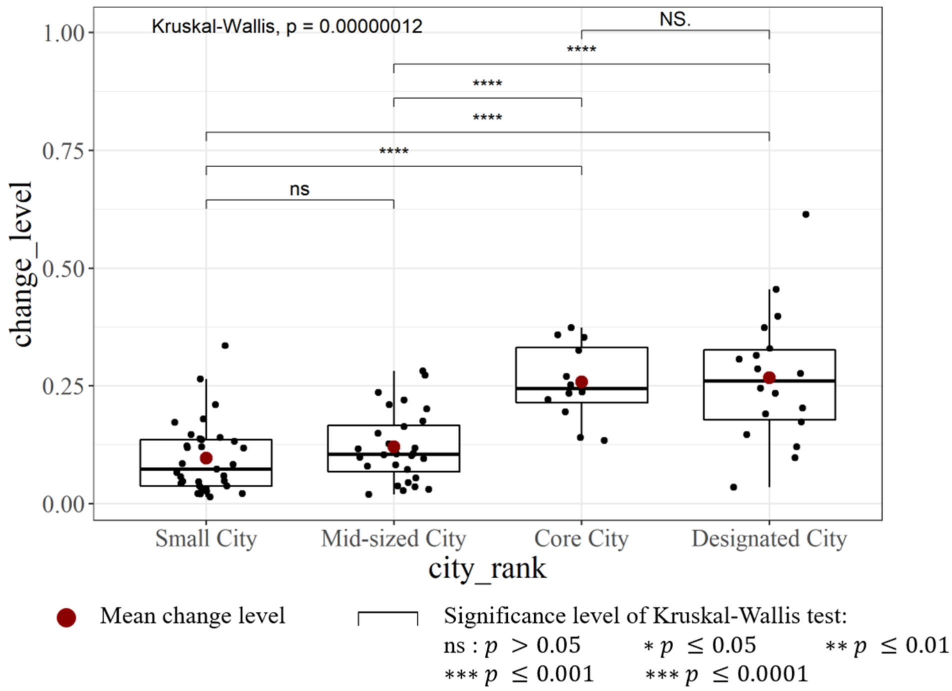


Fig. 4-3 The Boxplot of the Land-use Pattern Change Level of Station Areas in Cities of Different Scales

4.4 The Spatial Pattern of the Land Use Changes around Stations

This section focused on the trend for the locations of the land use changes around the station. The purpose is to identify the spatial pattern of the land use changes' distribution around the Shinkansen stations. Specifically, this section will clarify the spatial pattern of the land use changes from 2 perspectives: the *circumferential distribution* trend and the *radial distribution* trend. The circumferential distribution trend is to examine whether the land use changes differ on different orientation of the station. The radial distribution trend is to clarify whether the land use changes vary with the distance to the station.

To reduce the influence of the inherent differences of the land use data deriving from the land use discrimination biases [*2)], we selected the station areas whose land-use pattern change level is over 0.10. Thereby, 58 stations were selected, including 14 station areas in small cities, 16 station aeras in mid-sized cities, 12 station areas in core cities, and 16 station areas of the government ordinance designated cities.

To make clear the location trend of the land use changes, we summarized the relative locations of the land use changes to the Shinkansen station. As Fig 4-4 suggests, the station areas were rotated to match the form of the standard station area, in which the horizontal axis coincide to the extension direction of Shinkansen railway and the positive direction of vertical axis indicates the orientation of the central city (city hall). Then the frequency of the land use change at each location (100×100-meter cell) in the standardized station area will be summarized according to city hierarchy of the station city, to observe whether the location trends of the land use changes and to testify whether the land use changes had different spatial pattern in different hierarchy of cities.

In addition, section 4.2 has demonstrated that the change of Industrial land and Open land is not statistically significant, hence, this section will either not discuss their location trends. Besides, this section is only interested in the newly increased land use, regardless of the original land use categories. Therefore, this section will only discuss the location trend of the newly increased Tall Building land, Dense Low-rise land, Low-rise land and Transportation land.

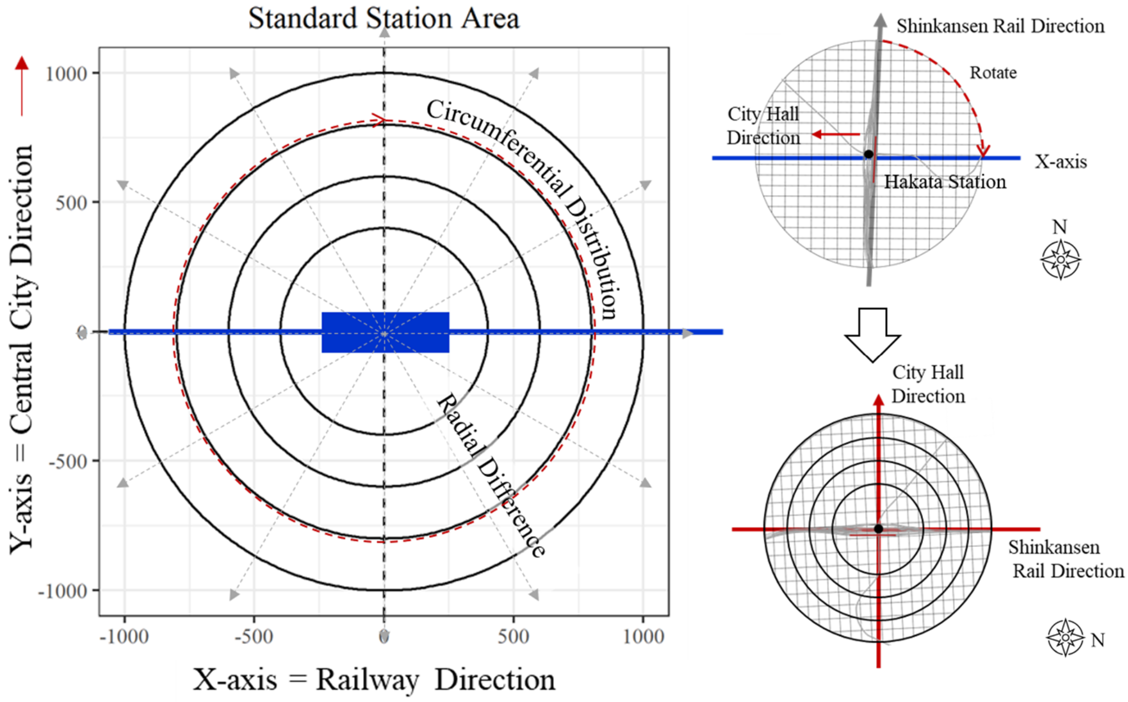


Fig. 4-4 The Diagram for the Standardization of a Station Area

As a result, the relative locations of the newly increased Tall Building land, Dense Low-rise Building land, Low-rise Building land and Transportation land in different hierarchy of cities are summarized in Fig.4-5. It can be observed that the newly increased Tall Building land mostly concentrates in station areas of the core cities and designated cities, while more of the newly increased Low-rise Building land were at the station areas of small cities and mid-sized cities. The amount of the newly increased Dense Low-rise Building land did not differ significantly in station areas of mid-size cities, core cities and the designated cities, but station areas of small cities clearly had less newly increased Dense Low-rise Building than station areas of other 3 city groups. Besides, the amount of the newly increased Transportation land corresponds to the hierarchy of the station city. Accordingly, the composition characteristics of the land use changes in different cities groups can be concluded from Fig. 4-5: the majority land use changes in station areas of small cities and mid-sized cities are land-use transition to the Low-rise Building land and Transportation land, while the land use changes in core cities and designated cities are characterized by large amount of the newly increased Tall Building land and Transportation land.

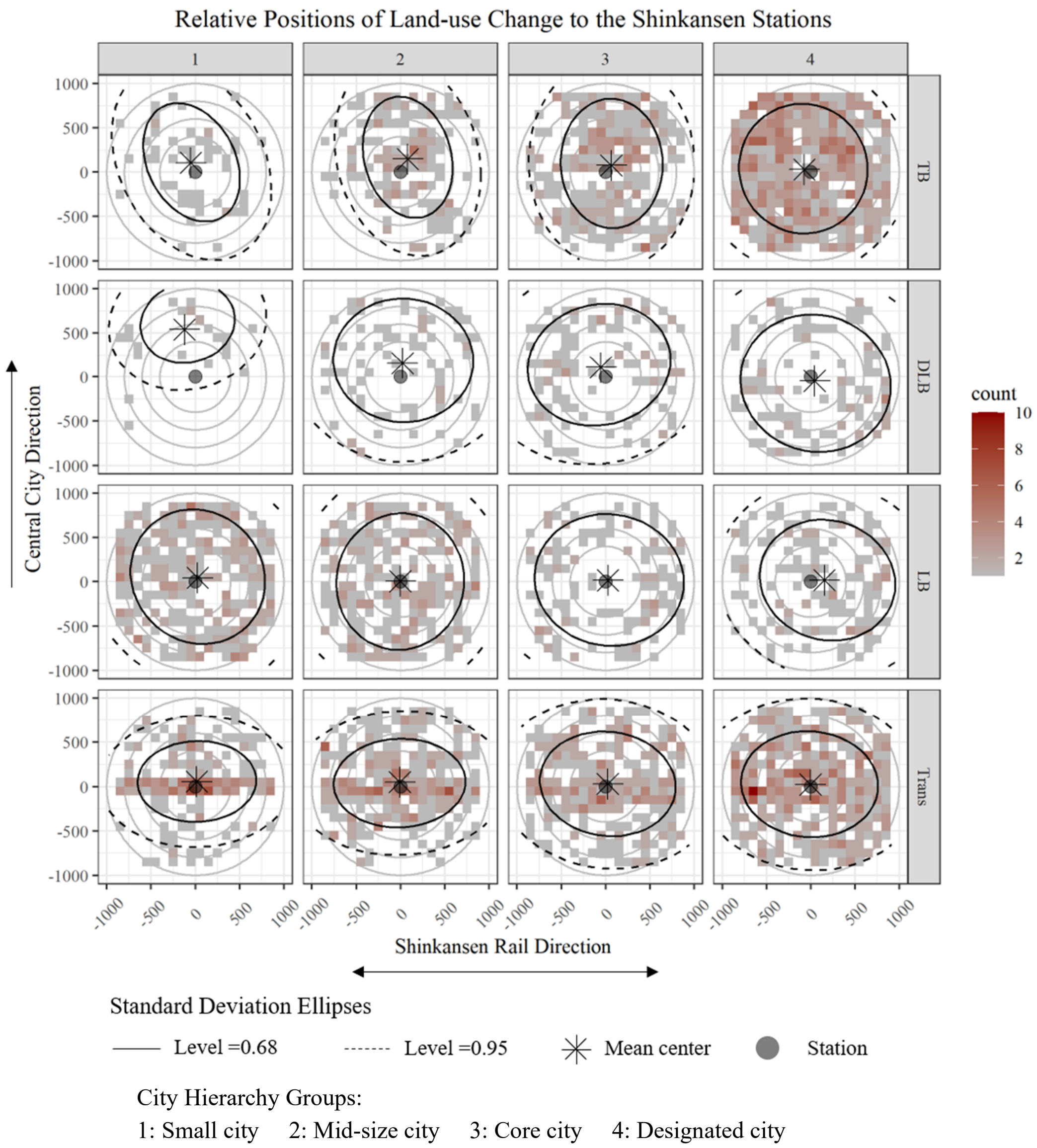


Fig. 4-5 The Circumferential Distribution of Land Development around Shinkansen Stations

4.4.1 The Circumferential Distribution Trends of Land Development

To examine the circumferential distribution characteristics of the land use changes, we added the *mean center* and the *standard deviational ellipse* to each plot of Fig.4-5. Mean center is calculated by the average of all the newly increased land use cells' geometry center coordinates. The standard deviational ellipse defined by the standard deviation of the newly increased land use cells' x-coordinates and y-coordinates from the mean center. The mean center reflects the central tendency of the newly increased land uses. The ellipse allows us to see if the distributions of the newly increased land uses are elongated and hence has a particular orientation. Also, the length of the major and minor axes of the ellipse reflects the dispersion trends of the newly increased land uses.

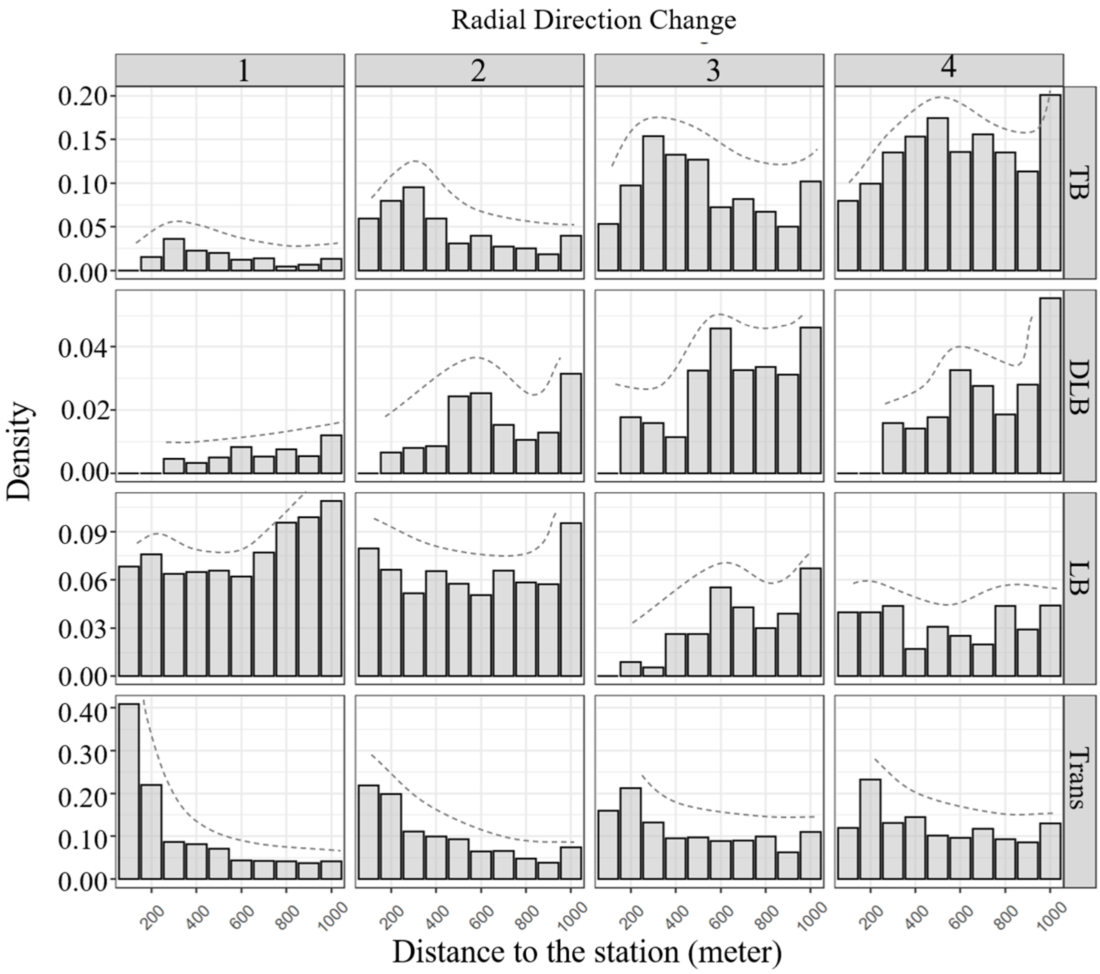
Accordingly, it can be observed that the newly increased land uses indeed have different location trends around stations, the circumferential distribution trend of each land use categories can be summarized as follow:

1) Tall Building land: In station areas of the designated cities, the newly developed TB cells are uniformly distributed around stations, there is no clear directional trend. However, in small cities, mid-sized cities, and core cities, the major axes of the ellipse are almost perpendicular to the direction of the railway extension, indicating in station areas of these cities, the development of TB has the directional trend in the perpendicular direction of the railway, and the development extends on both sides of the railway. According to the mean center, the development of TB in station areas of small and mid-sized cities is located slightly more on the side towards the city center.

2) Dense Low-rise Building land: The circumferential distribution trend of the newly developed DLB cells has no remarkable differences in station areas of mid-size cities, core cities, and designated cities. The length of the major and minor axes of the standard deviational ellipse is almost equal, and the mean centers are close to the station, indicating that the development of DLB has no significant directional trend in station areas of these cities. In contrast, the mean center of the newly developed DLB cells is located in the center city side, and it is distant from the station, indicating the development of DLB in small cities concentrates on the central city side of the stations.

3) Low-rise Building land: as the standard deviational ellipse suggests, in station areas of either city group, the length of the major axes and the minor axes are almost equal, and the mean center is close to the station, indicating the development of the LB locates uniformly in either orientation around the station.

4) Transportation land: it can be observed the major axes of standard deviational ellipses of the newly developed Trans cells in station areas of 4 city groups are parallel with the extension of railways and concentrated along the railway. However, in station areas of core cities and designated cities, the minor axes of the ellipse are longer than the ellipse of the station areas in small and mid-sized cities. Therefore, apart from the construction of the railway, in the station areas of the core cities and designated cities, there might also be a large number of land development for the improvement of inner traffic quality of the station area.



City Hierarchy Groups:
1: Small city 2: Mid-size city 3: Core city 4: Designated city

Fig. 4-6 The Radial Distribution of Land Use Changes to The Central Station

4.4.2 The Radial Distribution Trend of Land Use Changes

To analyze the radial distribution trends of the land development with the distance to the station, Fig 4-6 summarized the density of the land use development in every 100-meter circular ring region around station. Because with the distance to the station increasing, the area of the circular ring increases, it is naturally more likely to observe more land use development. Therefore, instead of summarizing the area of the newly increased land use cells, we analyzed the density which is calculated by the total area of the newly increased land use cell dividing by the area of the circular ring. Accordingly, the radial distribution trends of the land development of Tall Building land, Dense Low-rise Building Land, Low-rise Building land and Transportation land in station areas of 4 different hierarchy of station cities are illustrated in Fig. 4-6.

1) Tall Building land: In station areas of small and mid-sized cities, the newly developed TB cells concentrated more within the 400-meter radius area. In contrast, in station areas of core cities, the newly developed TB concentrated more between the 200~500-meter radius area. As to station areas of designated cities, the density of newly developed TB increased with the distance to the station within the 500-meter radius range of that station area, however, on average there is no significant increase or decrease trend in the range of the 200~1000-meter radius area.

2) Dense Low-rise Building land: The newly developed DLB has the location trend to distant from the station. Generally, there is barely any newly developed DLB within the 200-meter radius range station area. In the station areas of small cities, the newly developed DLB slightly increases with the distance to the station. In station areas of mid-sized cities and core cities, the newly developed DLB is more concentrated in the area that is 400-meter or more distant from the station. In the designated cities, the newly developed DLB was outside the 200-meter radius range area and was more likely to be located outside the 500-meter radius station area.

3) Low-rise Building land: In small cities, the newly developed LB has no significant difference within the 600-meter radius range area, while it was more likely to concentrate in the area outside the 600-meter range. As to station areas of mid-size cities, the newly developed LB has no significant radial distribution trend. Namely, the newly developed LB is distributed uniformly in the 1000-meter radius range station area. In station areas of core cities, it can be observed that the newly developed LB is generally outside the 300-meter radius range of the station area and is more concentrated outside the 600-meter radius station areas. As to the station areas of designated cities, there is no remarkable radial distribution trend.

4) Transportation land: The newly developed Trans is generally more concentrated near stations. However, in station areas of core cities and designated cities, this concentration trend is much less significant than in small and mid-sized cities. This can be related to the fact that there are more newly opened Shinkansen stations in small and mid-sized cities, therefore, it is more likely to observe the land development related to the railway construction in these areas.

4.5 Conclusion and Discussion

In conclusion, this chapter analyzed the dynamic land development situation around Shinkansen stations from the perspectives of the national level, regional level and the local level. On national level, this chapter summarized the land use transformation of the 91 existing Shinkansen station areas, testified that the Tall Building land and Transportation land has significantly increased, while the Dense Low-rise Building land, Low-rise Building land and Non-construction Building land has significantly decreased, indicating the comprehensive land development during 2009 and 2016 has the trend of improving the intensity and the traffic quality. Then on regional level, we computed the dissimilarity level between the land use pattern of 2009 and 2016 of each station area, thereby, we identified the station areas that has dynamically developed and the station areas that remain stagnated. In consequence, the major cities tend to have more dynamic changes while the small cities have much less land development, the dynamic development situation generally correlated to the station cities' hierarchy in the national city system. Finally, this chapter analyzed the location trends of newly increased Tall Building land, Dense Low-rise Building land, Low-rise Building land and Transportation land in station areas of different hierarchy of station cities. As a result, it can be concluded that the land development in station area of small cities and mid-sized is characterized by the land use transition to the Low-rise Building land, while the station areas of core cities and designated cities are characterized by the increasingly developed Tall Building land and Transportation land, and the location trends for the development of each land use category varies in different cities groups.

However, this study only focused on the phenomenon, as to the reason for the land use development, it can be complicated, and it always varies in different cities. The purposed of this part is to give an insight to the dynamic land development situation of Shinkansen station area, to provide reference to reconsider the current the land development strategies of the Shinkansen station area and the relation between the Shinkansen construction and the station area development.

Notes

*1). The latest urban land-use subdivision mesh data published by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism was the land-use data of 2016.

*2). The land-use subdivision mesh data provided by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism is based on the discrimination of satellite maps. Due to the quality of the satellite map and algorithm of the land-use discrimination, there exist inherent biases in the results. Therefore, the land-use data between different years have inherent differences caused in the data collection process.

Chapter 5

Functional Composition and Transformation Situation of Land Use in Core Shinkansen Station Areas

Insofar, we have clarified the land-development situation of Shinkansen station areas. However, the land-use data provided by the Policy Bureau of Land, Infrastructure, Transport and Tourism is about the construction situation of the land, it cannot specify the composition situation of urban functions like residential, commercial, business and amenity facilities. Therefore, this chapter looks deeper into the functional composition of the core Shinkansen station area, namely, the 400-meter radius range around Shinkansen stations. We collected the data reflecting the development situation of urban functions including residential, commercial, business and amenity facilities, then we classified the core Shinkansen station areas into different groups based on their function composition and clarified the characteristics of each group (**section 5.3**). Then we examine the location conditions of each group, discussing the relationship between the location conditions and the land development situation of Core Shinkansen station areas. (**Section 5.4**) Finally, we use population to reflect the development of residential and the employment to reflect the economic activities, thus analyzing the functional transformation situation in the core Shinkansen station areas. (**Section 5.5**)

5.1 Study Objects

This chapter focuses on the urban functions of land uses in the core Shinkansen station areas. Specifically, the *urban functional composition* of land uses refers to the condition of urban functions including commercial, business, residential and amenity facilities aggregating in a core station area. In this part, we focused on the 400-meter radius range around 92 exiting Shinkansen stations, as shown in Fig.5-1. We will analyze the composition situation of the urban functions in the core station areas and classify them into different typologies, then discuss the relationship between land development situation and the stations' location condition. Besides, we also examined the functional transformations in the core station area, thus giving a deep insight into the land development status and the dynamic development trends of the core station areas.

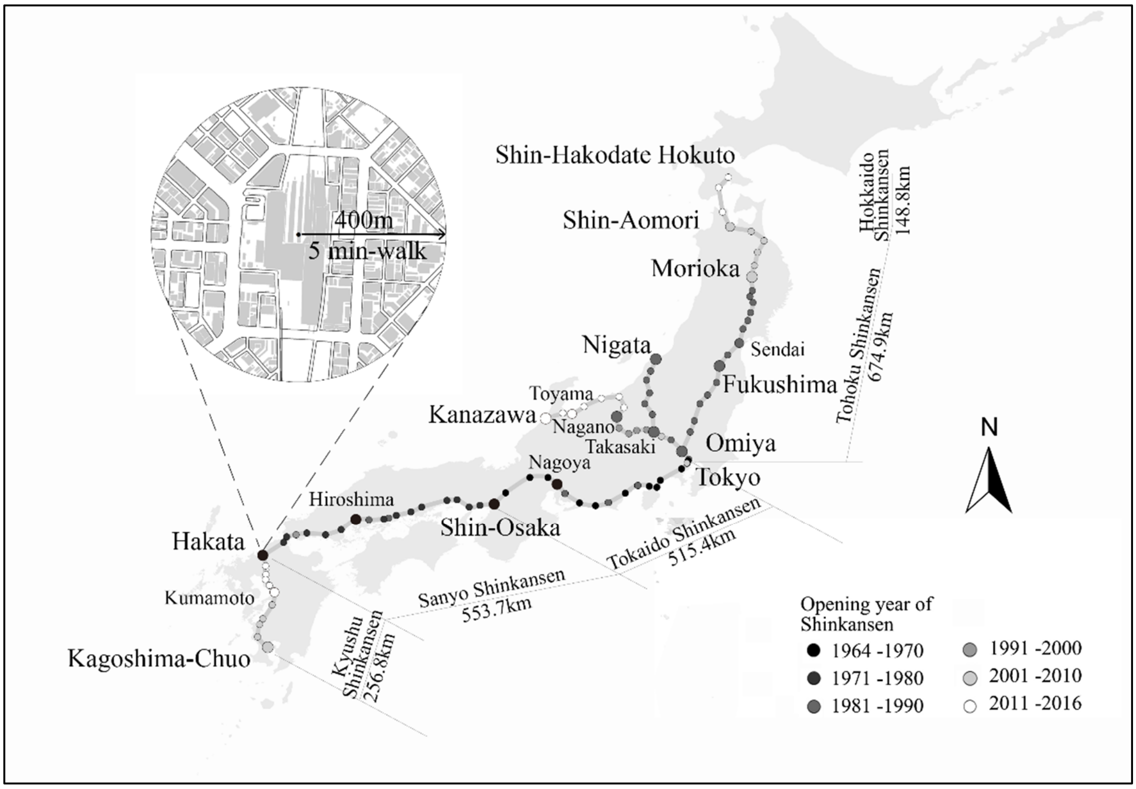


Fig. 5-1 Study Object

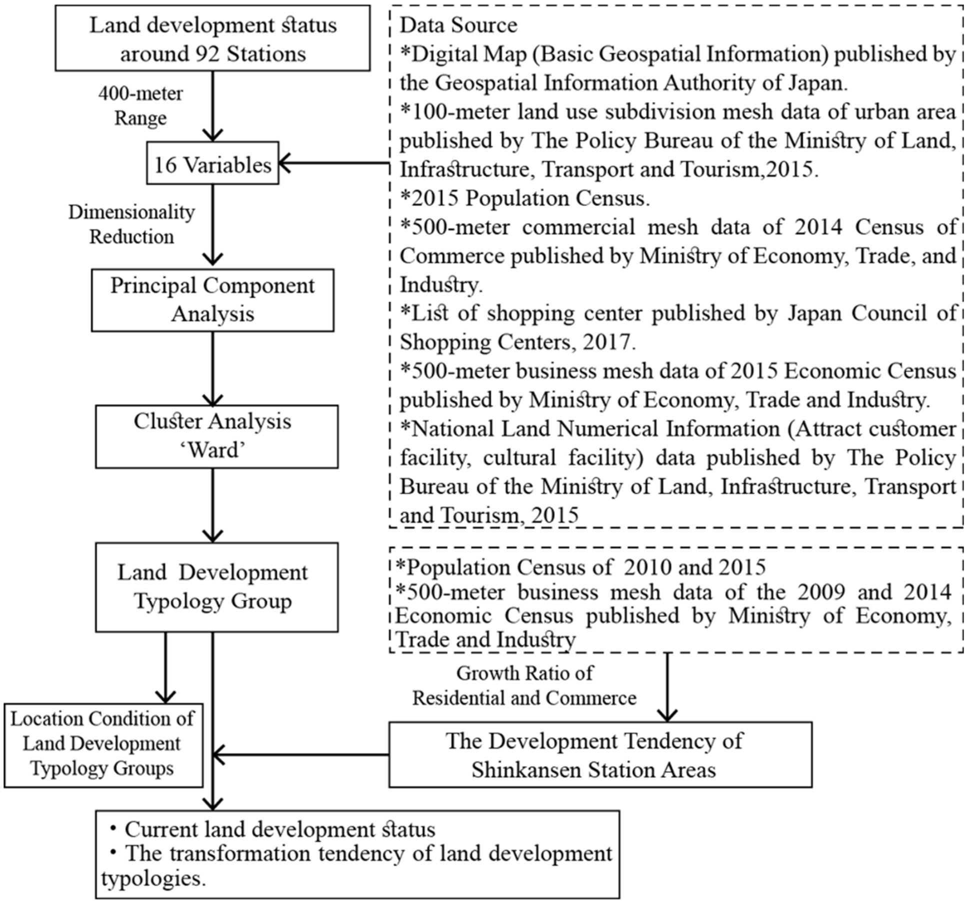


Fig. 5-2 Study Flow and Data Source

5.2 Method

To elucidate the functional composition and transformation situation of land uses in core station areas, the study flow is as shown in Fig.5-2.

Firstly, for the functional composition of the land uses, this study used 16 variables related to commercial, business, residential and other amenity functions, to reflect land development status around stations, which including building coverage ratio and road density from *Digital Map (Basic Geospatial Information)*, 100-meter land use subdivision mesh data of 2015, population and household from the *Population Census of 2015*, the number of retail shops and commercial employees from the *Census of Commerce of 2014*, the shopping center area from the *List of Shopping Center of 2017*, the number of companies and employees from the *Economic Census of 2015*, and attract customer facility and cultural facility data from *National Land Numerical Information* published in 2015. Through the Principal Component Analysis and the Hierarchical Cluster Analysis (Ward's method), we classified the current land development around 92 Shinkansen stations into different typologies. (Section 5.3)

Then, we analyzed the location condition of stations in each typology group, thus discussed the association between the land development typology and the location condition. In this study, location condition of a Shinkansen station refers to the characteristic of the regional environment that affects the location of a Shinkansen station, including 3 aspects: ① Region: the social and economic status of the city to which the station belongs; ② Transportation network: station's status in the transportation network which including the Shinkansen system itself and the access network to the station; ③Surrounding development condition: original development situation of the surrounding area. (Section 5.4)

Considering the growth of the population reflects the development of residential, and the growth of the employments to some extent indicates the commerce vitality, this study extracted the population in station area from the mesh-data of *Population Census of 2010 and 2015*[*2)], and the number of employments in the station area from *the business mesh-data of 2009 and 2014*. Through analyzing the growth of population and employments of Shinkansen station areas between 2010 and 2015, we clarified the dynamic development tendency of residential and commerce vitality of station areas. (Section 5.5)

5.3 Land Use Functional Composition

5.3.1 Variables to Reflect Urban Functional Composition Situation of Core Shinkansen Station Area

To clarify the current land development status around Shinkansen stations, 16 variables were selected from open data sources to reflect the development situation of commercial, business, residential and other amenity functions. The variable definitions and descriptive statistics are shown in Table 5-1, due to different measurement scales, the standard deviation of 16 variables go from 0.005 to more than 100,000. Since principal component analysis seeks to maximize the variance of each component, the result will be influenced a lot by the variables with large variances when variables are measured in different scales. Therefore, the standardization of data is necessary when conducting principal component analysis.

Before conducting principal component analysis, the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity were conducted to test whether the selected 16 variables were adequate for principal component analysis. The KMO measure verified the sampling adequacy for the analysis, with KMO = 0.63, which is "mediocre" according to Kaiser. [84] Bartlett's test of sphericity, χ^2 (92) =2624.24 ， $p<0.001$, indicating that correlation between variables were sufficiently large for PCA.

Table 5-1 Variable Definitions and Descriptive Statistics

Indexes	ID	Variables	Description	Mean	Std.dev.
Construction Condition (CC)	1	Building Coverage Ratio	The ratio of the building area divided by the land(site) area.	0.11	0.053
	2	Road Density	The ratio of the length of the total road network*3) to the station area. (m/m^2)	0.03	0.005
Land use*4) (L)	3	Tall Building Area	Land area of urban areas where commercial, business buildings and condominiums with 4 stories or more are densely concentrated. (m^2)	64716.42	87328.64
	4	Low-rise Building Area	Land area of urban areas where housings of 3 stories or less are distributed collectively and in a relatively low-density form. (m^2)	169596.23	115948.11
	5	Dense Low-rise Building Area	Land area of urban areas where housings of 3 stories or less are densely concentrated. (m^2)	46730.09	65515.26
	6	Other Urban Construction Land Area	Land area of other urban construction land including park, green land, road, railway, golf, and empty lot. (m^2)	147184.57	59263.79
	7	Urban Non-construction Land Area	Land area of non-construction land including farmland, forest, river and lake, sea and other non-construction land in urban area. (m^2)	74159.23	107328.60
Residential (R)	8	Population	Population living in station area.	1684.29	1314.18
	9	Household	The number of households in station area.	883.56	768.43
Commercial (C)	10	Number of Retail Shops	Total number of retail shops in station area.	78.75	98.70
	11	Commercial Employees	Total number of employees working at the commercial facilities in station area.	679.81	968.46
	12	Area of Shopping Center *5)	Total shop area of shopping centers in station area. (m^2)	17729.34	34317.63
Business (B)	13	Number of Companies	Total number of office and company in station area.	427.90	526.74
	14	Number of Business Employees	Total number of employees working at the office and company in station area	7506.49	14533.25
Amenity (A)	15	Number of Amenity Facility*6)	The number of amenity facility including public facilities and customer attract facilities.	1.44	2.06
	16	Number of Amenity Facility Types	The number of types of amenity facilities.	1.20	1.58

5.3.2 Principal Components of Land Development

According to the result of the Principal Component Analysis in Table 5-2, the first 4 four components whose eigenvalues were greater than 1 were chosen as the principal components, and the cumulative proportion suggests that 79.8% of the information contained in the original data can be explained with the first 4 components, which is sufficient enough to explain the original data set. [85] As to the interpretation of each component, the absolute value of 0.5 (weight: $0.5^2 = 0.25$) is commonly considered as a criterion that indicates the variable has a correlation with the component, and the larger absolute value suggests stronger

Table 5-2 Component Loadings of Variables in PCA Matrix

Indexes	Id	Variables	Comp 1	Comp 2	Comp 3	Comp 4
CC	1	Building Coverage Raito	0.891	0.279	0.034	0.023
	2	Road Density	0.052	0.757	-0.088	0.169
L	3	Tall Building Area	0.868	-0.143	-0.213	0.045
	4	Low-rise Building Area	-0.555	0.420	0.007	0.618
	5	Dense Low-rise Building Area	0.338	0.417	0.344	-0.532
	6	Other Urban Construction Land	0.504	-0.173	0.190	0.185
	7	Non-Urban Construction Land	-0.591	-0.497	-0.150	-0.480
R	8	Population	0.442	0.779	-0.146	-0.184
	9	Household	0.531	0.691	-0.204	-0.213
C	10	Number of Retail Shops	0.931	-0.131	-0.101	-0.033
	11	Commercial Employees	0.906	-0.204	-0.222	0.039
	12	Area of Shopping Center	0.704	-0.236	-0.358	0.044
B	13	Number of Companies	0.947	-0.140	-0.179	0.040
	14	Number of Business Employees	0.773	-0.326	-0.174	0.153
A	15	Number of Amenity Facilities	0.666	-0.101	0.668	0.048
	16	Number of Amenity Facility Types	0.669	-0.097	0.660	0.072
		Eigen value	7.620	2.643	1.427	1.078
		Standard deviation	2.761	1.626	1.195	1.038
		Proportion of Variance	0.476	0.165	0.089	0.067
		Cumulative Proportion	0.476	0.641	0.731	0.798

Comp1: Commerce; Comp2: Residential;
Comp3: Amenity; Comp4: Density (of low-rise building area)

correlation. [86] Since the component loadings of component 1 were generally large, we focused on the variables whose absolute value was greater than 0.8 (wight: 0.8^2=0.64) (grey cells). However, the component loading of component 2, 3, and 4 were all smaller than 0.8. Based on the principle that each of the variables explains no more than one component, we focused on the variables whose absolute value was about or greater than 0.7 at component 2 (grey cell), about or greater than 0.6 at component 3 (grey cell), and for component 4 we focused on the variables whose absolute value is greater than 0.5 (grey cell).

According to the correlation between variables and principal components, the four principal components were interpreted as follows:

Component-1 was interpreted as commerce axis, for it highly positively correlated to the number of companies, employees, retail shops, building coverage ratio and Tall Building area, which reflects the characteristics of central business and commercial areas. Besides, all the variables reflect the development are positively correlated while the undeveloped-related variables are negatively correlated, thus this axis can to some extent also reflect the comprehensive development intensity of an area.

Component-2 was considered as a residential axis for it highly positively correlated with population, household, and road density, while negatively correlated to all the variables related to commerce, which reflected the characteristics of a typical residential area.

Component-3 was defined as the amenity for it is highly positively correlated to the number and types of amenity facilities, while negatively related to the variables that related to commerce, namely, this axis reflected the amount and diversity of non-commercial amenity functions.

Component-4 reflects the density of Low-rise Buildings, for it has a high positive loading on Low-rise Building area variable, and negative loading on Dense Low-rise Building area variable. Namely, component-4 is an indicator of the density of housings with 3 or less stories in the station area.

5.3.3 The Patterns of Functional Composition of Core Station Areas

The hierarchical cluster analysis (Ward's method) was conducted based on the principal component scores of 92 station areas, and according to the gap statistic, this study classified the land development status of 92 station areas into 6 typologies[*7)]. According to the average component scores of each typology group (Table 5-3), the 6 typologies can be defined as: A. Primitive state (15 stations), B. Low-density residential-oriented typology (35 stations), C. Intensive mixed residential-oriented typology (21 stations), D. Low-density commerce-oriented typology (8 stations), E. Dense commerce-oriented typology (7 stations), F. Integrated core typology (6 stations). Fig.5-3 shows the distribution of projection in the two-dimensional coordinate system composited by component 1 and 2 (Commerce - Residential), and component 3 and 4 (Amenity – Density of Low-rise Building). Fig.5-4 shows the cluster dendrogram and the result of land development typology groups. Based on the projection values at each component axis, the characteristic of each typology group was elucidated and summarized as follows:

1). Primitive state (15 stations): refers to station areas who have extremely low values on both commerce and residential components, which indicates that these station areas were in the initial stage of development with almost no development of commercial, business or residential.

2). Low-density residential-oriented typology (35 stations): refers to station areas with relatively high values on residential and amenity axes, whereas extremely low value on commerce axis, which indicates station areas in this typology group were residential areas without barely any commercial facilities. Besides, the value on the Low-rise Building density axis suggests that the buildings in these station areas were not in a high-density form. Therefore, these station areas are considered as low-density residential-oriented.

3). Intensive mixed residential-oriented typology (21 stations): refers to station areas with high values on all functional axes of commerce, residential, and amenity, among which the value on the residential axis was largest, which indicates the land development of these station areas were not only residential, but also intensive and multifunctional. Thus, these station areas are defined as intensively developed residential-oriented areas.

4). Low-density commerce-oriented typology (8 stations): refers to station areas with low values on all functional axes, among which values on the commerce axis was relatively higher, for which the functional composition of this group can be considered as commerce-oriented. However, since component 1 also reflects the comprehensive development intensity, compared with the largest average value of 6.368 among six groups, the average value of 0.292 indicates the commerce of this group is relatively low density. Therefore, this group is defined as low-density commerce- oriented typology.

5). Dense commerce-oriented typology (7 stations): refers to station areas with extremely high values on commerce axis, while the values on residential and amenity axes are extremely low, which indicates the land development in these station areas were dense commercial and business with barely any residential and amenity facilities.

6). Integrated core typology (6 stations): refers to station areas where both commerce and amenity were densely concentrated, while only the development of residential does not match up with the development orientation of these areas, which can be considered as characteristics of station areas located in the city central area, where commercial, business and other multi-functional amenity facilities were intensively gathered.

Table 5-3 The Average Component Scores of Each Typology Group

ID	Land Development Typology	Comp1	Comp2	Comp3	Comp4
A	Primitive State	-2.828	-2.028	0.182	-0.949
B	Low-density Residential-oriented	-1.748	0.513	0.169	0.696
C	Intensive Mixed Residential-oriented	1.620	1.384	0.309	-0.753
D	Low-density Commerce-oriented	0.292	-0.135	-1.586	0.488
E	Dense Commerce oriented	4.148	-0.128	-1.723	-0.332
F	Integrate core	6.368	-2.435	1.604	0.684

Comp1: Commerce; Comp2: Residential;
Comp3: Amenity; Comp4: Density (of low-rise building area)

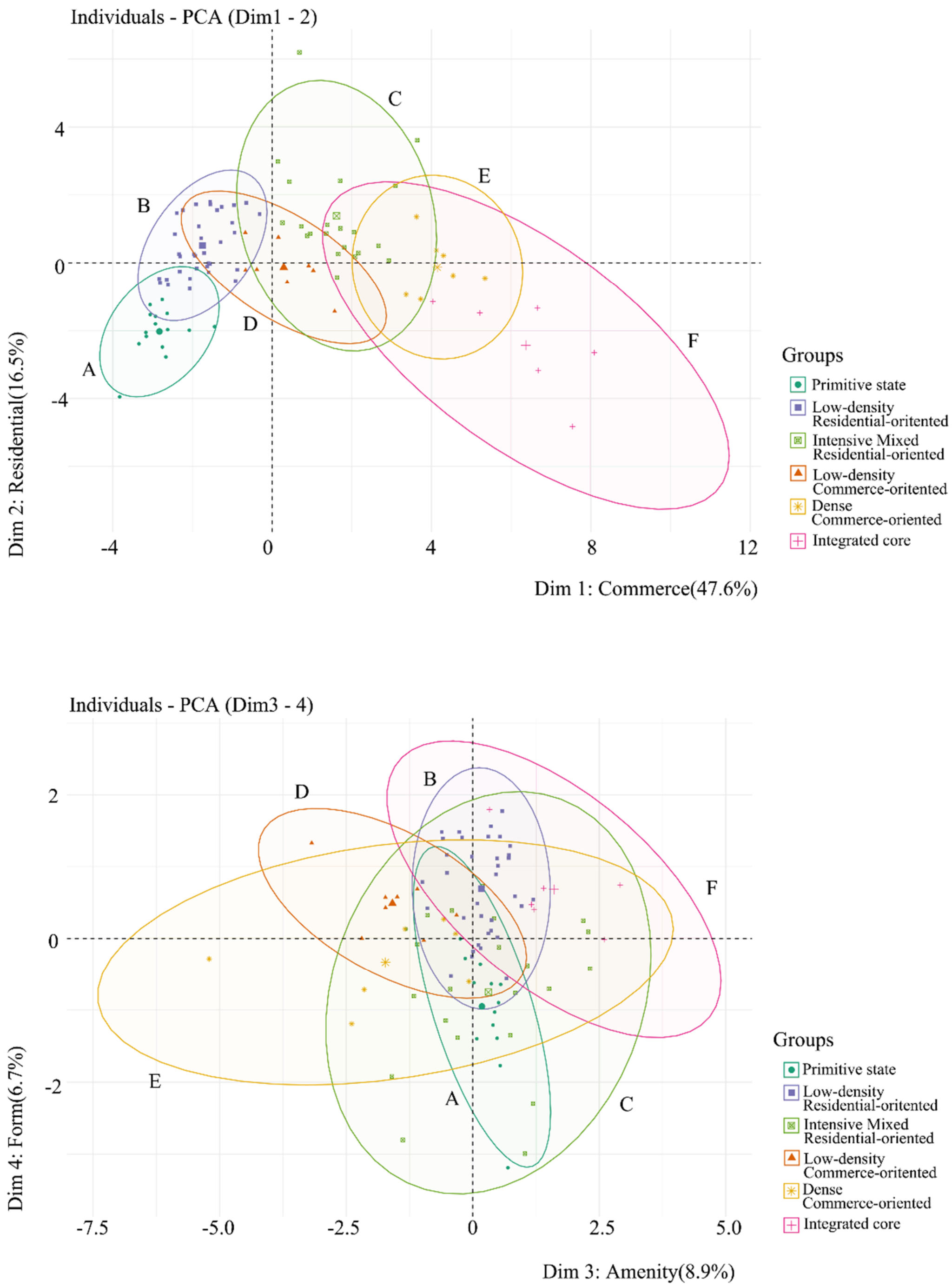


Fig. 5-3 Distribution of Stations in the Space of Principal Components (dim 1-2, dim 3-4)

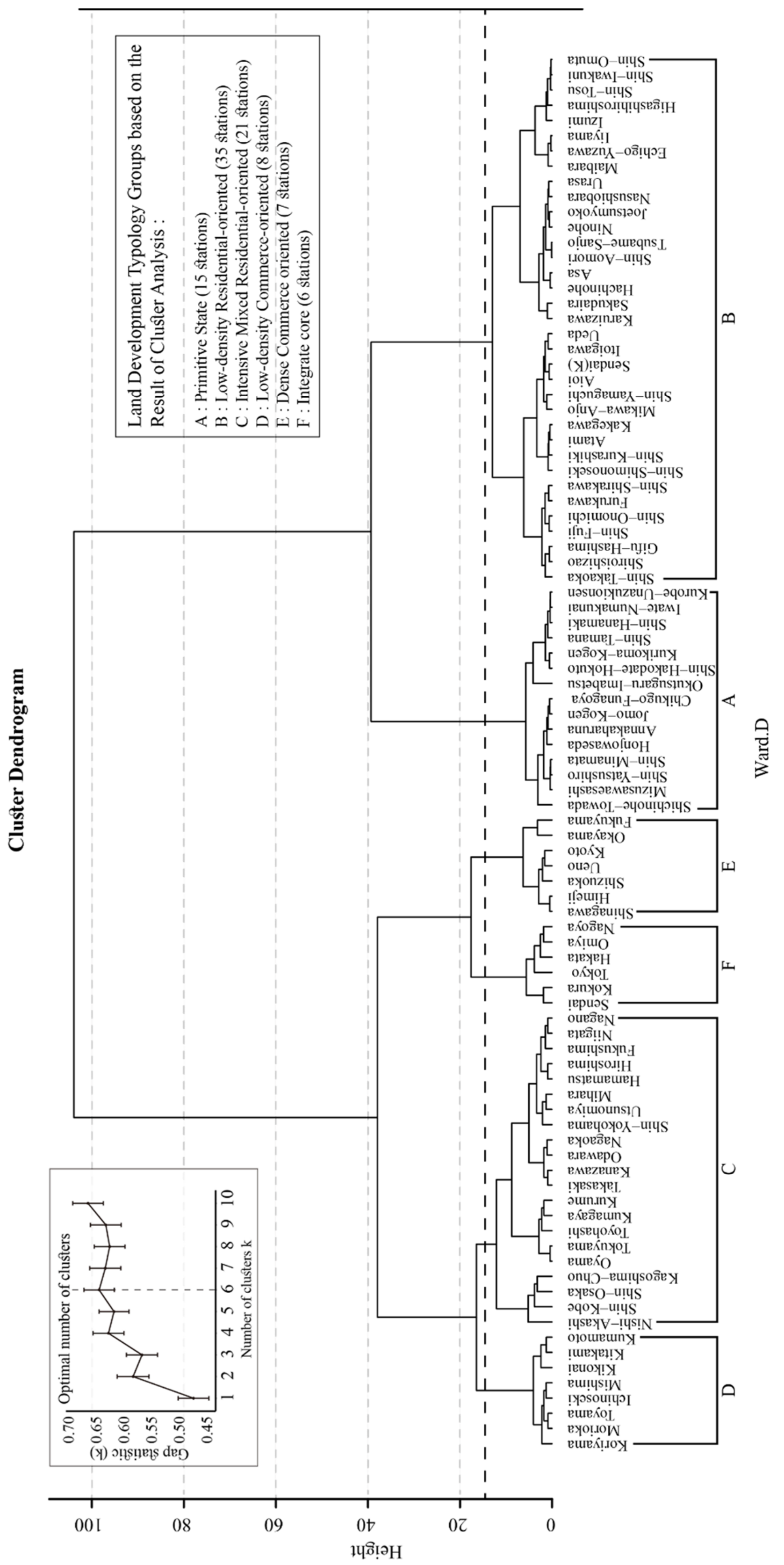


Fig. 5-4 Cluster Dendrogram and the Result of the Land Development Typology Groups

5.4 The Relationship between The Functional Composition Pattern and Station's Location Condition

The locational advantages will largely influence the development of station areas; therefore, this part discusses the relationship between functional composition pattern and the station's location condition. Specifically, we analyze the location condition of a Shinkansen station from 3 perspective: regional, transportation network, and the local level. For regional level, we focused on the station city's hierarchy in the national city system; for transportation network, we analyze the accessibility of the station area; for local level, we examine the initial construction situation of station area when Shinkansen started operation. As a whole, we expect to make clear the relationship between the functional composition pattern of station areas and the station's location condition.

5.4.1 Regional Perspective

The distribution of land development typology groups is shown in Fig.5-5. It can be found that dense commerce-oriented, and integrated core typology of stations are mostly located in the major cities along Tokaido Line and Sanyo Line, and many intermediate stations along these two lines are intensive mixed residential-oriented. However, primitive state stations are mostly located at the remote area along the Kyushu Shinkansen, Hokuriku Shinkansen, and the north section of Tohoku Shinkansen. These characteristics of each typology's spatial distribution can be related to regional status and the economic background when Shinkansen were constructed.

As to the regional status, it can be found that all stations of dense commerce-oriented typology and integrated core typology were located in the central city of Japan's main metropolitan areas. All these cities have a population*9) over 200,000, and concentrate major administrative organs, important airports, important ports, universities, large companies, and many other facilities to support the intensive economic activities and social welfare in these regions. Therefore, for these cities, the intensive development around Shinkansen stations seems like a natural result. In contrast, stations along Joetsu Shinkansen, Hokuriku Shinkansen, Kyushu Shinkansen and the north section of Tohoku

Shinkansen were mostly located in the small local cities with much fewer population, less social resources, therefore it is much difficult to attract the investment compared with major cities.

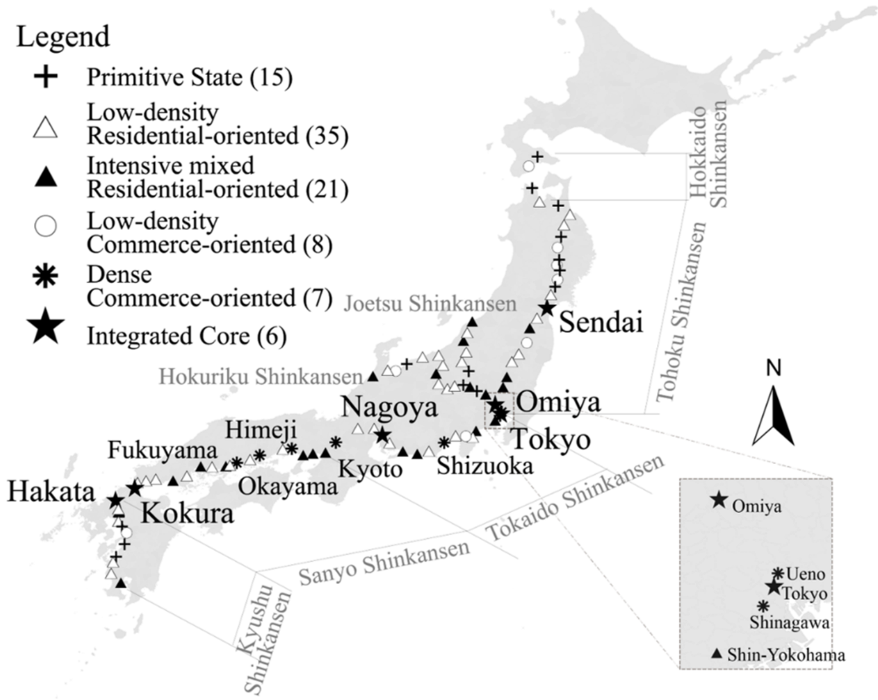


Fig. 5-5 The Distribution of Land Development Typologies

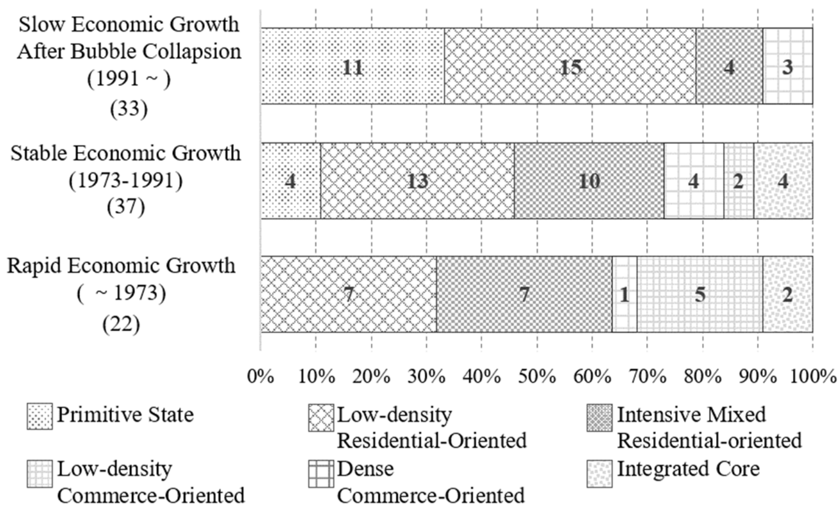


Fig. 5-6 The Functional Composition Pattern of Station Area and Economic Background of the Opening Year of Shinkansen

Besides, according to the constructed year of Shinkansen at stations in each typology group (Fig.5‒6), it can be found that the dense commerce-oriented station areas and integrated core station areas where commerce intensively developed, constructed Shinkansen in the post-war economic rapid growth period, while primitive state stations areas all constructed Shinkansen after the year of 1973. Notably, among 33 stations where Shinkansen were constructed after the Bubble Collapsing, only 4 stations had intensive development at the surrounding areas, namely, other 87% of the stations had barely any development at the surrounding areas. It is reasonable to attribute the stagnation of development to the short development time, however, the difference of land development typology in different period also implied that the economic environment such as the land market had impact on the land development mode of station area, although further study is needed to clarify the mechanism of the influence.

5.4.2 Transportation Network Perspective

A station's status in the transportation network refers to the *efficiency* to access the station areas and *connectivity* to other points in the network. The volume of ridership directly reflects the utilization situation of a station, thus comprehensively represents the station status in the transportation network. More specifically, since Shinkansen usually has a longer services distance than conventional rails, it plays an important role in strengthening the interregional linkage, while the conventional rails and buses usually serve the intra-regional linkage in a smaller area. Thus, the location condition of a station in the transportation network can be considered from two levels: interregional network and the intra-regional network, namely, the Shinkansen network itself and the access network to the station.

According to the daily passenger volume shown in Fig.5‒7, comprehensively, the development intensity corresponded to the ridership. Besides, it can be observed that the average daily passengers of integrated core and dense commerce-oriented stations were mostly larger than 100000, while the average daily passengers of stations with lower intensity of development were less than 100000, especially, primitive state stations were less than 5000. Therefore, the average daily passenger volume of 100000 and 5000 can be considered as threshold for evaluating the development potential of a station area.

According to the route information and departure schedule of each station (including Shinkansen, conventional railway and bus) *10), as shown in Fig.5-8, dense commerce-oriented and integrated core station areas were not only served by a higher efficient interregional network (Fig.5-8: a), but also better access network with more alternative routes (Fig.5-8: d, e) and higher daily departure frequency (Fig.5-8: b, c). The superiority in both interregional network and intra-regional network makes these station areas a transportation hub in the network, therefore, economic activities tend to agglomerate near these stations. In contrast, primitive station area and low-density residential-oriented developed station areas were usually served with very limited alternative routes with much lower departure frequency, therefore they were much less competitive to get investments from large companies.

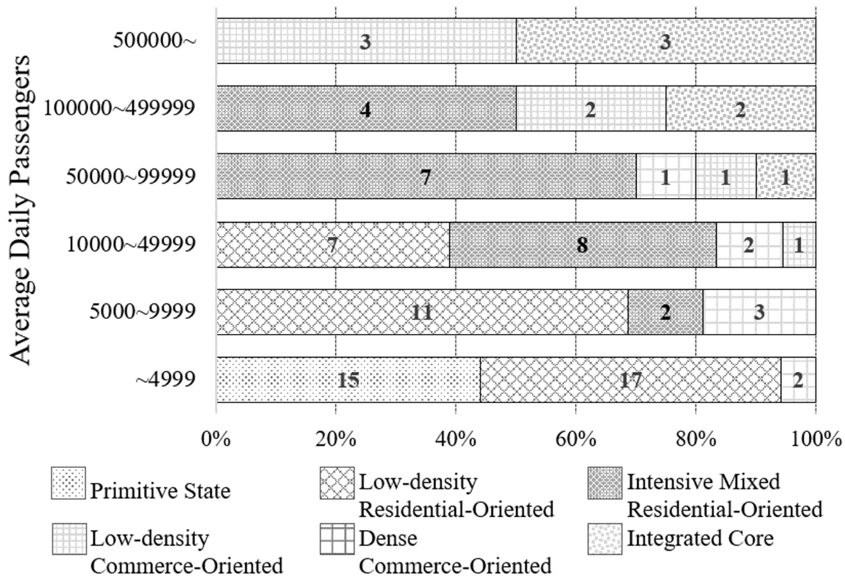


Fig. 5-7 The Functional Composition Pattern of Station Areas and Average Daily Passengers

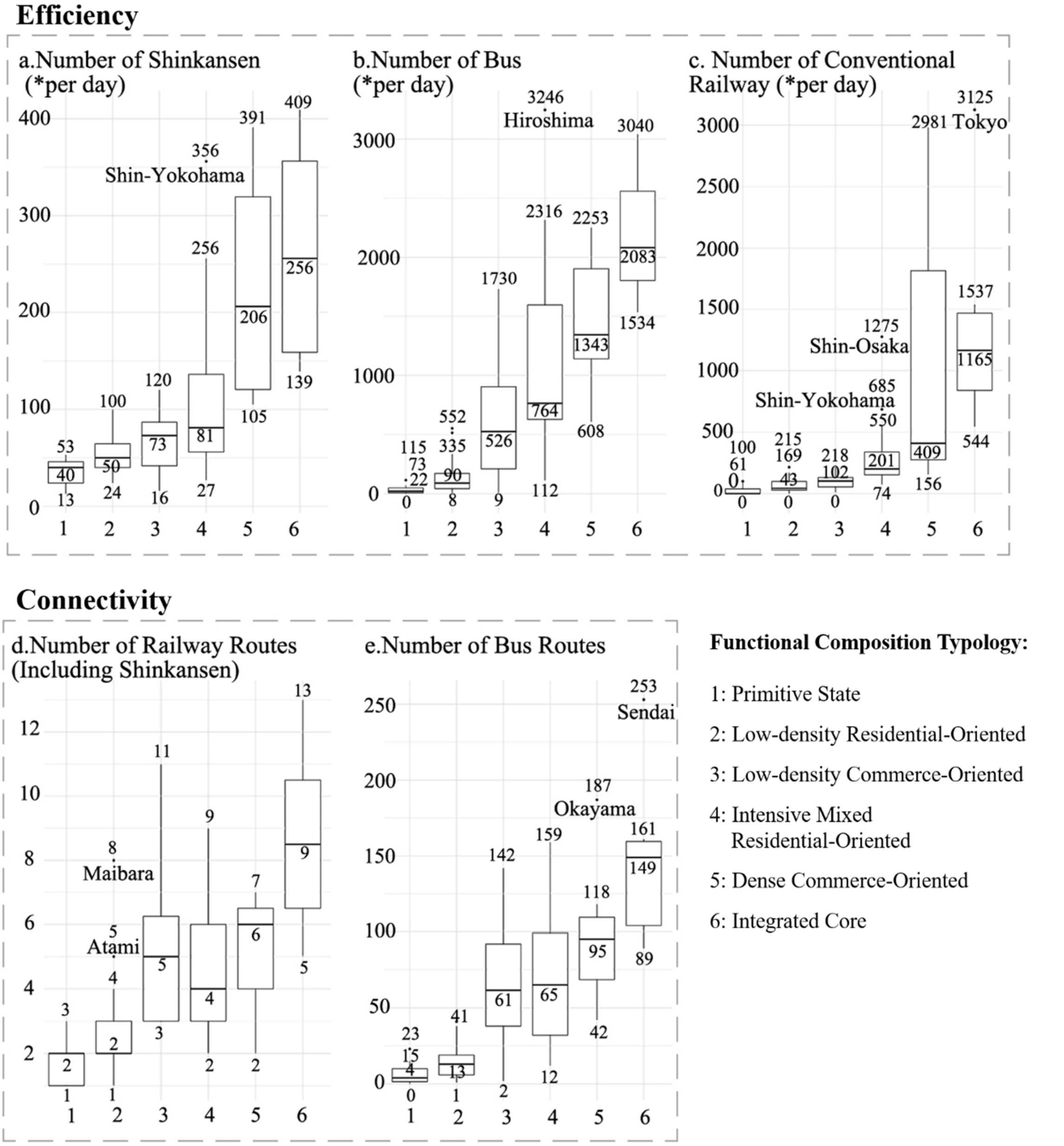


Fig. 5-8 Station Status in Transportation Network

5.4.3 Local-level Perspective

Most Shinkansen stations were reconstructions of existing railway stations where the surrounding areas were already to some extent developed. Fig.5-9 shows the proportion of newly built stations and reconstructed stations in each land development typology groups. Among 22 newly built stations, however, only 2 stations had the surrounding area intensively developed, 11 remained undeveloped and 9 had residential developed in low-density status, which reflected that Shinkansen's impact of stimulating the urban development of surrounding area was limited in those newly built station areas.

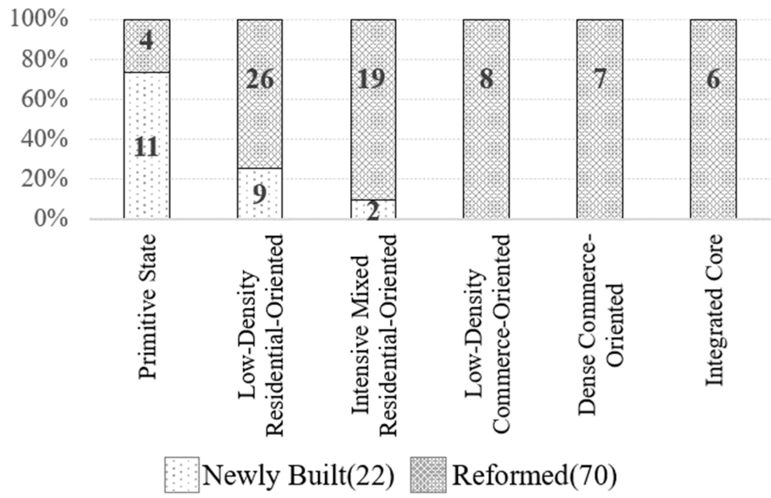


Fig. 5-9 The Functional Composition Pattern and The Initial Local Construction Situation

Consequently, although it is widely accepted that Shinkansen can promote the urban development of the surrounding area, the fact that only 34 stations had their surrounding area intensively developed and 58 station areas remained undeveloped or developed with low intensity, suggested that the development-promoting effect was limited. Undoubtedly, the introduction of Shinkansen services to major cities in the rapid economic development period has been seen to have been a major success, however, for those intermediate stops located at remote area, the success in those major cities were not necessarily transferable. It has to be recognized the success of the Shinkansen station areas in major cities were under very specific circumstances, they occupy all the superiority in the region and the transportation network, it is natural for these station areas to concentrate economic activities. However, for intermediate station in the local cities, the development of a positive benefit affected by Shinkansen developments requires careful planning and policy intervention to create development opportunities.

5.5 The Functional Transformation Situation of Core Shinkansen Station Areas

5.5.1 The Change of Functional Composition Situation

According to the comprehensive situation of the growth of population and employment in 92 Shinkansen station area (Table 5-4), it was found that during 2010 and 2015, the core surrounding area of 92 Shinkansen stations as a whole did concentrate more population and economic activities, nevertheless, the net population growth ratio of 3.68% and net employment growth ratio of 1.66% indicates the growth was not remarkable, the increment was almost even with the decrement. Besides, it is notable that although the total number of employments increased, the decrease happened at more than half of the Shinkansen stations.

Table 5-4 The Statistical Summary of Population Growth and Employment Growth around Shinkansen Stations

	Population Growth	Population Growth Ratio	Employment Growth	Employment Growth Ratio
Mean	66	0.0315	133	0.0797
Median	28.5	0.0254	-7	-0.0097
Net Growth	6080	0.0368	12273	0.0166
Number of Stations Positive Growth	55	55	41	41
Number of Stations Negative Growth	37	37	51	51

According to the growth of population and employment at each station, the development tendency of 92 Shinkansen station areas were classified into following 4 groups, the grouping result is as shown in Table 5-5.

1). Recession (24): refers to station areas where both the population and employments decreased, which indicated the decline of both residential and commerce vitality.

2). Residential Growth (26): refers to station areas where population increased while employments decreased, which suggests these station areas were transforming to residential-oriented areas.

3). Commerce Growth (13): refers to station areas where population decreased while employments increased, which suggests that these areas were transforming to commerce-oriented areas by replacing residential to commerce.

4). Integrated Developing (29): refers to the station areas where both population and employment increased, namely, both residential and commerce were developed.

Table 5-5 Functional Transformation Type of Core Shinkansen Station Area

Type	P*	E*	Stations
Recession (24)	—	—	Okutsugaru-Imabetsu, Shin-Hakodate-Hokuto, Kikonai, Tokuyama, Jōetsumyōkō, Karuizawa, Itoigawa, Kakegawa, Shin-Ōsaka, Iiyama, Echigo-Yuzawa, Iwate-Numakunai, Hachinohe, Kurobe-Unazukionsen, Shin-Fuji, Fukushima, Mihara, Odawara, Nasushiobara, Mizusawaesashi, Asa, Ichinoseki, Kyoto, Shin-Iwakuni.
Residential Growth (26)	+	—	Utsunomiya, Shin-Kurashiki, Nishi-Akashi, Kōriyama, Kitakami, Shin-Onomichi, Kumagaya, Morioka, Shin-Shimonoseki, Oyama, Hiroshima, Kurikoma-Kōgen, Toyama, Shin-Kōbe, Maibara, Urasa, Shizuoka, Nagano, Shin-Ōmuta, Shiroishizaō, Mikawa-Anjō, Kurume, Shin-Shirakawa, Shinagawa, Ueno, Himeji.
Commerce Growth (13)	—	+	Tokyo, Jōmō-Kōgen, Atami, Shin-Minamata, Nagoya, Toyohashi, Nagaoka, Chikugo-Funagoya, Mishima, Ninohe, Shin-Takaoka, Furukawa, Shin-Yamaguchi
Integrated Developing (29)	+	+	Sendai(K), Ueda, Aioi, Tsubame-Sanjō, Shin-Hanamaki, Niigata, Takasaki, Okayama, Fukuyama, Izumi, Shin-Yokohama, Higashihiroshima, Kagoshima-Chūō, Hakata, Shin-Tosu, Sakudaira, Kanazawa, Sendai, Shin-Aomori, Hamamatsu, Gifu-Hashima, Honjōwaseda, Shin-Tamana, Kokura, Ōmiya, Shichinohe-Towada, Shin-Yatsushiro, Annakaharuna, Kumamoto
*P: the growth ratio of population *E: the growth ratio of employments			

The spatial distribution of development tendency is shown in Fig.5-10. It was observed that the major stations were able to keep the commerce vitality, while many intermediate station areas, especially stations between Tokyo and Sendai station, were transforming to residential-oriented areas. Besides, it is also notable that along the recently built Shinkansen lines, many stations along Kyushu Line had their surrounding area integrated development, while many station areas along Hokuriku Line and Hokkaido Line kept population loss and commerce recession. An interpretation is that station areas along Kyushu Line gained population from the remoter area, while the population of some station areas along Hokkaido Line and Hokuriku Line migrated to greater cities along the Line. Since these areas have the similar configuration of circumstances for development, the disparity of the development tendency must closely relate to the planning strategies of local government, which can be a task for further study.

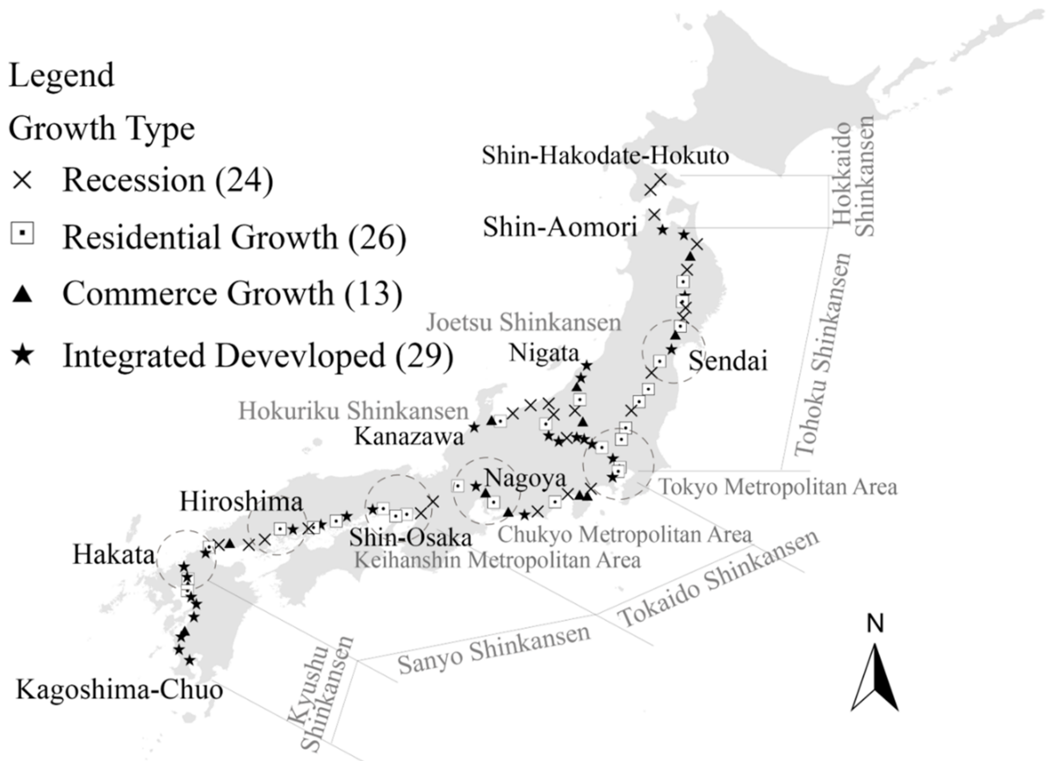


Fig. 5-10 The Spatial Distribution of Stations and the Functional Transformation Types

5.5.2 The Dynamic Transformation Trends of the Functional Composition Pattern of Core Shinkansen Station Areas

Through analyzing the development tendency of station areas in each land development typology group, the transformation relation between land development typologies became clear. As Fig.5-11 shows, integrated core areas were able to retain the economic vitality, while the dense commerce-oriented areas and the low-dense commerce-oriented areas, where there were lack of multi-functional amenity facilities and residential, were transforming to residential-oriented areas. And stations of the primitive state and low-dense residential-oriented typology were more likely to have population loss and commerce recession compared with other typology of station areas.

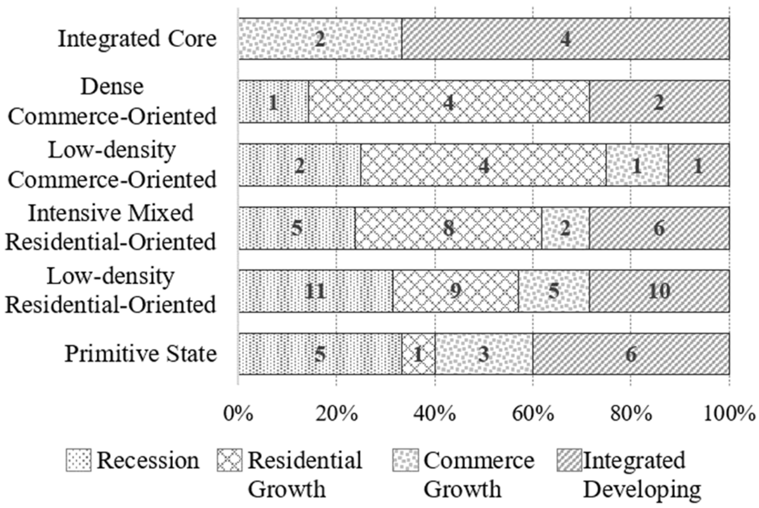


Fig. 5-11 Functional Transformation Situation in Station Areas of Different Composition Patterns

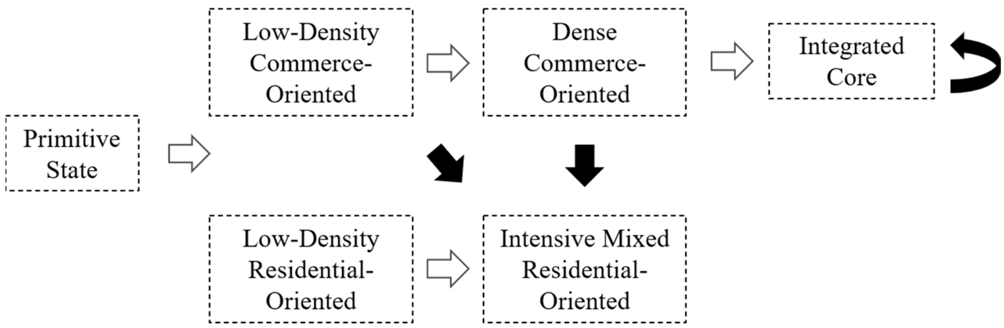


Fig. 5-12 The Dynamic Function Transformation Relation Between Composition Patterns of Core Shinkansen Station Areas

These phenomena seem inconsistent with the land-use effects of high-speed rail developments that existing research has demonstrated, and also failed to match up with the general anticipation of a boost for commercial and business activities around stations. Usually, for station area, we expect that it will go through the development process from its primitive state to an integrated core or to an intensive mixed residential-oriented area. However, the development of Shinkansen station areas during 2010 and 2015 suggested that although the integrated core station areas were able to maintain their commerce vitality, and to aggregate more population and economic activities, most of other land development typology station areas were failed to retain their economic vitality and were transforming to residential-oriented areas. (Fig.5-12)

To consider the reason, the influence of depopulation can be one factor. Because the economic activities always correspond to demand, the decrease in total population made it hard to have significant growth in the consumption and investment in domestic demand[24)], thus there are fewer chance to have large scale new development compared with the rapid economic development period. Instead, the main source of the development opportunities these days may mainly come from transfers from other locations along the line where the configuration of development conditions is poorer.

As to those intensively developed station areas with decades of development, their superiority in the regional location and transportation network always provides them priority development opportunities to enhance their commerce and business vitality. For some commerce-oriented intermediate stations, the depopulation and aging problem of the whole city resulted in the shrinking of local demand as well as the labor market, some enterprises may start to consider pulling out from local market and seek for relocation towards center cities along the line with more opportunities and larger market. When the enterprises started to leave, the decreased land price at many of these station areas[*11)], the superiority of accessibility and location made these areas attractive to residential, which explains the reason many station areas were transforming into residential-oriented typology. Besides, for newly built stations located in the suburb of local cities, although the efficient linkage to the regional central cities through Shinkansen removed the transportation barriers of cost and time consumption, which allow them to share the same market with the major cities, their configuration of development conditions may not be able to provide them much competitiveness in the competition for the development opportunities with regional center cities. and in most cases, better life supported services in some periphery areas of regional center cities supported by conventional rail may reduce the attractiveness of these intermediate station area served by Shinkansen. As a result, the economic activities became more concentrated at a few regional center cities, while many intermediate station areas started functional transformation towards residential oriented.

5.6 Summary

This chapter analyzed the current land development around 92 Shinkansen station areas and classified the land development into 6 typologies: 1). Primitive state, 2). Low-dense residential-oriented, 3). Intensively mixed residential-oriented, 4). Low-dense commerce-oriented, 5). Dense commerce-oriented 6). Integrated core. However, since the purpose of this study is to comprehensively elucidate the functional composition around a station, the heterogeneity of the development situation inside the 400-meter radius circular area was not discussed in this study.

Then, by analyzing the location condition of stations in each typology group, we found that station areas with higher development intensity always located in the regional central cities, and the stations are transportation hubs which served the station area with efficient interregional and intra-regional connection. However, the surrounding area of intermediate stations located in the local cities did not get as much development opportunities as expected. This suggested that Shinkansen's impact on promoting the development around the stations were under specific configuration of circumstances, the success at those regional center cities were not necessarily transferable to local cities.

Besides, according to the discussion about dynamic transformation trends of functional composition in the core Shinkansen station areas in recent years, the core Shinkansen station areas did concentrate more population, however, we also observed the recession of commerce vitality around more than half of the stations. Over half century's experience of Shinkansen development made us believe the construction of Shinkansen will solve both transport and regional problems and bring about the development opportunity to the surrounding area. A remarkable thing we found is that only those integrated developed station areas were able to retain their commerce vitality, while most of the intermediate station areas were transforming to residential-oriented areas. Therefore, it has to be recognized that the impact of Shinkansen on urban development of surrounding area under the depopulation background was no longer as it was in the rapid economic growth period. Faced with the unprecedent shrinking of population and aggregate demand, it may be no longer appropriate to follow the previous development theories. Through analyzing the current development circumstances, this research holds that intensively developed station areas in regional central cities still hold the superior development condition in the competition for the development opportunity, while the intermediate stations in the local cities should focus more on the intra-regional demand, and more effort should be put in improving the intra-regional connection.

Notes

*1) The Akita Shinkansen and Yamagata Shinkansen were not discussed in this research for their maximum running speed is 130Km/h, while it is generally accepted that the maximum running speed of high-speed rail should be in excess of 200 Km/h.

*2) All the data from the newest available government open source, many investigations such as national census are conducted every 5 years, because next time will be in 2020, this research focused on the year during 2010 and 2015.

*3) Road length is estimated from the length of road edge line data of Digital Map (Basic Geospatial Information).

*4) The land use data were extracted from the 100-meter land use subdivision mesh data provided by The Policy Bureau of the Ministry of Land, Infrastructure, Transport and Tourism. The land use feature type of each mesh unit was assessed based on the satellite data, the land use of each mesh unit was assessed as one of 17 land use feature types including farmland, other agricultural land, forest, wasteland, Tall Building area, factory, Low-rise Building area, Dense Low-rise Building area, road, railway, public facility land, space land, park and green land, river and lake, seaside area, sea, and golf land.

*5) Shopping center refers to integrated multi-functional commercial facilities whose retail shop area is over 1500 ㎡, or the number of tenant is over 10 besides the key tenant.

*6) Amenity facilities refer to public facilities, and customer attract facilities. Specifically, public facilities refer to welfare, post office, hospital, school, fire station, police station, local public organization, national administration facility, and other public service facilities. Customer attract facilities refer to movie theatre, assembly hall, playhouse, exhibition hall, sports center, dome, stadium, and other facilities with leisure and recreation function.

*7) The estimate of the optimal number (k) of clusters is the value that maximizes the gap statistic, however, as the gap statistic figure at the upper left of Fig.4 shows, when the number of clusters is greater than 6, the gap statistic did not have significant growth until k=10. Since 10 clusters will make the interpretation too complicated, 6 clusters were considered to be more appropriate.

*8) Sendai Station in Kagoshima Ken.

*9) The population refers to city population according to 2015 Population Census. For Shinagawa station, Tokyo station, and Ueno stations, city population refers to the total population of 23 special wards

*10) The bus route and conventional railway information was from the *National Land Numerical Information* published by Ministry of Land, Infrastructure, transport and tourism, and schedule information was from Ekitan.com

*11) Based on the public announcement of land price in 2015 published by Ministry of Land, Infrastructure, transport, and tourism, among 57 stations where there are standard lands in the station area, land price around 37 stations were observed decreased.

Chapter 6

Summary

In concern of the unprecedented nationwide depopulation, the Ministry of Land, Infrastructure, Transport, and Tourism (MLIT) has highlighted that construction of regional high-speed connections such as the Shinkansen network work is the key to achieve agglomeration economies and support the effective services delivery and regional innovation. As the transportation hubs and city gateway, the development of a Shinkansen station area is crucial to the comprehensive development of the whole city. However, the majority of the studies about the Shinkansen station area development are based on the successful cases along the early constructed line, particularly the Tokaido Shinkansen and Sanyo Shinkansen, which were initially constructed to meet the increasing traffic demands of on the conventional railway route. In contrast, the later constructed Shinkansen were mainly planned to reduce the regional imbalance, although later studies have demonstrated the implementation of Shinkansen may have filed to meet this expectation, and researchers urge that the development of Shinkansen not naturally bring positive benefits to the urban development of the city, especially to the intermediate stations in the small cities, a careful planning and appropriate policy intervention is important. Faced with the new reality of nationwide depopulation, a study on the current land development situation around Shinkansen stations will help us better understand the major issues in the Shinkansen station area development thus providing reference to the decision-making for the future development plan.

Accordingly, this study analyzed the current land development situation in the 92 existing Shinkansen station areas, aimed to identify the patterns of the current land development situation and the dynamic land use transformation situation. Specifically, this study defines the 1000-meter radius range area around a Shinkansen station as the station area and the 400-meter radius range area as the core station area. For the station area, this study focused on the construction intensity and analyzed the land-use pattern features including the composition and the spatial relationship of land uses. For the core station area, this study considers it as a spatial indifference zone and analyzes the urban functional composition of the station areas. On the whole, this study aimed to answer the questions of **what** land uses compose the station areas, **where** they are located around stations, and **how** they have changed.

To answer the research questions, this study started with developing a land-use pattern feature-based land-use pattern comparison method based on the image matching algorithm in Chapter 2. In contrast to the previous land-use pattern comparison approaches which are based on some specific indices that measure certain aspects of the land-use pattern feature, the image matching approach allows us to integratively extract the land-use pattern features including the composition and spatial relationships of land uses and numerically represented it as a land-use pattern descriptor. The similarity computed based on the land-use pattern descriptors will be able to reflect the similarity level between land-use maps with integrative consideration of the compositional and configurational characteristics, which will become the source for distance calculation in the further classification. Specifically, we treat the land-use map as an image, initialize it as a numeric array with the feature value of each unit representing the land-use category information of the corresponding land-use cell. Then, we use the image smoothing algorithm to convert the feature value of each land-use cell to a feature value that represents the land use composition and distribution situation in the defined neighborhood, thereby, the overall land-use pattern can be represented by the counts of the unique output feature values, which we call as the land-use pattern descriptor. Accordingly, the similarity between 2 land-use patterns can be calculated by the histogram correlation of the output image, which will evaluate the similarity level between the land-use patterns with a value in [-1,1], with the value closer to 1 indicates higher similarity. Chapter 2 explained the algorithm of the proposed method, then showed the effectiveness of the method by experiments of the land-use pattern comparison of the simulated random maps generated with specified compositional and configurational. As a result, the land-use pattern descriptors we extracted by the proposed method successfully identified the land-use patterns of the same group and distinguished the land-use patterns of different groups as the experiments have expected. Then Chapter 2 showed the applicability of the proposed method in the urban study by a comparative study of 6 sample Shinkansen station areas, the proposed method not only identified the Shinkansen station areas with similar land use composition proportions but also helped distinguish the distribution characteristics between the sample station areas. Finally, Chapter 2 explicated the influence of the parameters on the land-use pattern feature extraction and specified the limitation of the method due to the trade-off between computation efficiency and accuracy, thus clarified the principles to determine the parameters.

Then, in Chapter 3, we applied the method developed in Chapter 2 to compare the land-use patterns of 91 existing Shinkansen stations areas based on the 100-meter land-use map of 2016 provided by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism. According to the pairwise similarities computed by the proposed method, the land-use patterns of the 91-existing land-use patterns were classified into 6 groups with the hierarchical clustering analysis. Based on the histograms of the maps after smoothing, we analyzed the global and local land-use pattern features of station areas in each group. The global land-use pattern features include the areal proportions, aggregation level and spatial variations of land uses, while the local land-use pattern features are the major contiguity relations between land uses in the defined neighborhood. Also, we discussed the relationship between the land-use pattern and the population scales of the station cities, thus elucidating the trend of the land-use patterns in cities with different scales. Consequently, we identified that the land-use pattern of existing 91 Shinkansen station areas includes the following types:

1. **Land-use pattern type-I** is characterized by the dominated Non-construction land aggregated in the station areas, forming a large continuous unconstructed area, while other land uses scatter in the station areas. This type of land-use

pattern is usually observed in station areas of cities whose population is 300 thousand, particularly cities with a population under 100 thousand, or station areas near a mountain or a great river.

2. **Land-use pattern type-II** is distinct for the dominated Low-rise Building land that aggregates in the station area, forming a large continuous area covered by low-rise buildings, while other land uses have much fewer composition proportions and disperse in the station area. This is another typical land-use pattern for station areas of cities whose population is less than 300 thousand, but the majority are in mid-size cities that have a population between 100-300 thousand.

3. **Land-use pattern type-III** is characterized by the dominated Tall Building land mixing with the transportation land and covering the whole station area. This is a typical land-use pattern of station areas in the cities of the metropolitan centers.

4**. Land-use pattern type-IV** is marked for the Dense Low-rise Building land dominating and aggregating in the station area, forming large continuous areas covered by densely distributed low-rise buildings that are equal to or less than 3 stories. This land-use pattern corresponds to the station areas of many regional center cities.

5. **Land-use pattern type-V** is characterized by the fragmented pattern formed by the Non-constructed land and Low-rise Building land mixing with other land use categories. There is no single land-use category dominated in the station area and the average areal proportions between the Non-construction land, Low-rise Building land, and other land use categories approximate to 1:1:1. This land-use pattern is also mainly observed in station areas of small cities whose population is less than 100 thousand, particularly the recently constructed stations in the suburban areas.

6. **Land-use pattern type-VI** is characterized by a mixed pattern formed by small patches of Tall Building land, Dense Low-rise Building land, Low-rise Building land, and Transportation land. There is no single land-use category either, on average, the Dense Low-rise Building land and Low-rise Building land together take about 50% of the total area, while the Tall Building land and

the Transportation land take about 15% of the total area respectively. In the station area, the land uses formed some discernible land-use zone. This type is identified as a typical land-use pattern for station areas of regional center cities with a population of over 300 thousand.

Accordingly, we identified the current land-use patterns of 91 existing Shinkansen station areas and clarified the trend of the land-use pattern of a station area in cities of different population scales.

In chapter 4, we compared the land-use patterns of Shinkansen station areas over time, to make clear the dynamic land use transformation situation in the station areas. With the land-use maps of 2009 and 2006 provided by the Policy Bureau of the Ministry of Land, Infrastructure, Transport, and Tourism, we analyzed the land use transformation situation from the perspective of the national level, regional level, and the local level. At the national level, we summarized the overall composition proportion changes of each land use category, accordingly, we testified that the Tall Building land and Transportation land have significantly increased, the Dense Low-rise Building land, Low-rise building land, and the Non-construction land have significantly decreased, indicating that comprehensively the land development around Shinkansen stations has been to improve the intensity and traffic quality. On the regional level, according to the dissimilarity level between land-use patterns of 2009 and 2016 calculated by the method developed in Chapter 2, we identified the station cities whose Shinkansen station areas have been dynamically developed and station cities whose Shinkansen station areas' land development have stagnated. Finally, at the local level, based on geographic distribution analysis including mean center, standard deviational ellipse, and density analysis, we identified the location trend of the development of Tall Building land, Dense Low-rise Building land, Low-rise Building land, and Transportation land in station areas of cities with different population scales. In consequence, we found that in small cities and mid-size cities whose population is less than 300 thousand, the land-use transformation is characterized by the land-use transition to the low-rise building land, while the station areas of core cities and designated cities with a population of over 300 thousand are characterized by the increment of Tall Building land and Transportation land. As to the location trends, the development of Tall Building land in small cities, mid-size cities, and core cities are generally perpendicular to the direction of the railway extension and mostly aggregated in the 500-meter radius range of the station area, in contrast, in station areas designated cities, it is uniformly spread in the station area. For the newly increased Dense Low-rise Building land, it is significantly more aggregated in the central city side of the station in the small cities, while in other cities whose population is greater than 100 thousand, there is no significant direction trend, but a radial distribution trend of deviating from the station. As to the newly increased Low-rise Building land, there is no clear location trend in station areas of small cities, mid-size cities, and designated cities, however, it is observed in station areas of the core cities that it is more likely to locate in the area more than 500 meters away from the station. As to the newly increased transportation land, it is generally more likely to concentrate along the railway, however, in the core cities and designated cities, it is more uniformly distributed in the whole station area than in the station areas of small cities and mid-size cities.

The Chapter 3 and Chapter 4 have focused on the construction intensity of the 1000-meter radius range of Shinkansen station areas without discussing the urban functions including residential, commercial, business and amenity facilities. Thus, in Chapter 5, we focused on the function composition of the 400-meter radius core Shinkansen station areas. With 16 variables reflecting the development situation of the residential, commercial, business and amenity facilities in the station area, we implemented the principle component analysis and hierarchical analysis, clarified the functional composition of the core Shinkansen station areas into 6 typologies, including the primitive state, low-density residential-oriented typology, intensive mixed residential-oriented typology, low-density commerce-oriented typology, dense commerce-oriented typology, and integrated core typology. Then by analyzing the location condition of stations in each typology group, we clarified the relationship between the functional composition trend and the location condition of the station. Besides, we also discussed the dynamic transformation trends of the functional composition of the core Shinkansen station area according to the change of the population and employment. As a result, we identified that although the core station area did concentrate more population, however, we also observed the recession of commerce vitality around more than half of the stations. For the integratively developed station areas in the major cities, they are able to retain their commerce vitality, while most of the intermediate station areas have a transformation trend to residential-oriented areas.

In conclusion, this study has systematically summarized the land-use situation and the dynamic transformation situation of the 92 existing Shinkansen station areas and identified the trend of the land development situation in Shinkansen station areas of cities with different population scales. This study provided a comprehensive perspective to understand **what** land use is in the Shinkansen station areas, **how** they are arranged, and **how** they are changing. this study contributes to the existing study about the land development of Shinkansen station areas by providing a comprehensive insight into what is happening in the Shinkansen station areas, which will help better understand the current issues of the station areas development and provide reference to the decision-making of the future development plan. However, a limitation of this study is we have not answered why the land development pattern is as it is, the influential factors of the land development are beyond the research scope of the study. Because the influential factors for the station area development can be complicated, and they usually vary in different cities, further study is needed to specify the influential factors in different station areas.

Besides, this study has also presented a novel approach to identify the land-use pattern of certain typical urban areas. With the increase of accessible imagery data in the urban study field, interpreting spatial patterns in maps with image processing techniques will become an inevitable trend. In contrast to previous approaches which are either based on visual comparison of land-use maps or some quantifiable landscape metrics, the image matching approach compares land-use patterns based on integrative land-use pattern features, and the proposed approach identifies land-use patterns not only based on the composition proportions, but also the spatial relationships between land-use units, and it allows us to process a large number of object areas efficiently. Apart from the application in this study, the proposed method can also be applied to identify the land-use patterns of other typical urban areas, such as areas around large-scale commercial facilities, academic cities, medical cities, etc., in this sense, the proposed method is also a meaningful attempt to introduce the image processing techniques to geospatial analysis.

References

1. Garmendia M, Ribalaygua C, Ureña JM. High speed rail: implication for cities. Cities. 2012. pp. S26–S31. doi:10.1016/j.cities.2012.06.005

2. Straszak A. The Shinkansen as Part of the Management and Technology Area Studies on Large-Scale Planning Projects. The Shinkansen High-Speed Rail Network of Japan (Proceedings of an International Institute for Applied Systems Analysis Conference, June 27-30,1977). 1977;7: 5–8.

3. Kamada M. Achievements and Future Problems of the Shinkansen. The Shinkansen High-Speed Rail Network of Japan (Proceedings of an International Institute for Applied Systems Analysis Conference, June 27-30,1977). 1977;7: 41–56.

4. Blum U, Haynes KE, Karlsson C. Introduction to the special issue: The regional and urban effects of high-speed trains. The Annals of Regional Science. 1997. pp. 1–20. doi:10.1007/s001680050036

5. Taniguchi M. High Speed Rail in Japan: A Review and Evaluation of the Shinkansen Train. 1992 [cited 20 Sep 2021]. Available: https://escholarship.org/uc/item/5s48m11f

6. Sands B. The Development Effects of High-Speed Rail Stations and Implications for California. Built Environ. 1993;19: 257–284.

7. Ueda T, Nakamura H. The Impact of Shinkansen on Regional Development. Proceedings of Infrastructure Planning. 1989. p. 597Y604.

8. Cao J, Liu XC, Wang Y, Li Q. Accessibility impacts of China’s high-speed rail network. J Transp Geogr. 2013;28: 12–21.

9. Chen C-L. Reshaping Chinese space-economy through high-speed trains: opportunities and challenges. J Transp Geogr. 2012;22: 312–316.

10. Deng T. Impacts of Transport Infrastructure on Productivity and Economic Growth: Recent Advances and Research Challenges. Transp Rev. 2013;33: 686–699.

11. Hall P. Magic Carpets and Seamless Webs: Opportunities and Constraints for High-Speed Trains in Europe. Built Environ. 2009;35: 59–69.

12. Vickerman R. High-speed rail and regional development: the case of intermediate stations. J Transp Geogr. 2015;42: 157–165.

13. Ureña JM, Menerault P, Garmendia M. The high-speed rail challenge for big intermediate cities: A national, regional and local perspective. Cities. 2009. doi:10.1016/j.cities.2009.07.001

14. Monzón A, Ortega E, López E. Efficiency and spatial equity impacts of high-speed rail extensions in urban areas. Cities. 2013. pp. 18–30. doi:10.1016/j.cities.2011.11.002

15. Vickerman R. High-speed rail in Europe: experience and issues for future development. Ann Reg Sci. 1997;31: 21–38.

16. Bharule S, Kidokoro T, Seta F. Evolution of high-speed rail and its development effects: Stylized facts and review of relationships. Tokyo: Asian Development Bank Institute (ADBI); 2019. Available: http://hdl.handle.net/10419/222807

17. Ministry of Land, Infrastructure, Transport and Tourism (MLIT). Grand Design of National Spatial Development towards 2050. 4 Jul 2014 [cited 11 Dec 2020]. Available: https://www.mlit.go.jp/common/001088248.pdf

18. Statistics Bureau, Ministry of Internal Affairs, Communications. Final report of 2015 "POPULATION AND HOUSEHOLDS OF JAPAN." [cited 22 Sep 2021]. Available: https://www.stat.go.jp/english/data/kokusei/2015/final_en/final_en.html#Summary

19. Nishida M. History of the Shinkansen. In: Straszak Andrzej RT, editor. The Shinkansen High-Speed Rail Network of Japan; Proceedings of an IIASA Conference. International Institution for Applied Systems Analysis Conference; 1980. pp. 11–20.

20. Sasaki K, Ohashi T, Ando A. High-speed rail transit impact on regional systems: does the Shinkansen contribute to dispersion? The Annals of Regional Science. 1997. pp. 77–98. doi:10.1007/s001680050040

21. Kurihara T, Wu L. The Impact of High Speed Rail on Tourism Development: A Case Study of Japan. The Open Transportation Journal. 2016;10: 35–44.

22. Anastasia Loukaitou-Sideris, Deike Peters, Paige Colton, and Eric Eidlin. A Comparative Analysis of High-Speed Rail Station Development into Destination and Multi-Use Facilities: The Case of San Jose Diridon. Mineta

Transportation Institute Publications; 2017 Feb. Available: https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?article=1223&context=mti_publications

23. Boyce RR, Clark WAV. The Concept of Shape in Geography. Geogr Rev. 1964;54: 561–572.

24. Batty M. Classifying urban models. Environ Plann B Plann Des. 2016;43: 251–256.

25. Olmedo MTC, Camacho Olmedo MT, Paegelow M, Mas JF, Escobar F. Geomatic Approaches for Modeling Land Change Scenarios. An Introduction. Geomatic Approaches for Modeling Land Change Scenarios. 2018. pp. 1–8. doi:10.1007/978-3-319-60801-3_1

26. Izumiya, Taku. The recent trend in the process of consensus building of urban redevelopment:A case study of east entrance area of Omiya station. Proceedings of the General Meeting of the Association of Japanese Geographers. 2014;2014: 100200.

27. Arimura, Tomohide. Distribution and Change of CBD Functions in the Central Area of Kagoshima City:Focusing on the Redevelopment of Station Area in relation to the Implementation of Kyushu Shinkansen. Geospace space. 2019;12: 21–35.

28. Kikuchi, Yoshiyuki. The Changing of Business Office Location in Regional Center City:A Case Study of Tertiary Industry in Maebashi and Takasaki Region. ABSTRACTS of the Annual Meeting, The Human Geographical Society of Japan. 2003;2003: 44–44.

29. 大橋, 美幸. Case Studies of the Effects of Opening the Hokkaido Shinkansen Line(Part 5). 函館大学論究 = The review of Hakodate University. 2018;50: 59–92.

30. TSUJIMURA, Masashi. The Change of Sphere of Station's Influence and the Transformation in Surrounding Areas of the Stations Brought by the Opening of the Shinkansen Line : A Case Study of Yamaguchi Prefecture. GEOGRAPHICAL SCIENCES. 1982;37: 42–56.

31. Sawada, Kiyoshi. Urban Renewal in the Surrounding Area of Hakata Station at Fukuoka. Proceedings of the Institute of Natural Sciences, Nihon University. 1978; 13–25.

32. Calthorpe P. The Next American Metropolis: Ecology, Community, and the American Dream. Princeton Architectural Press; 1993.

33. Cervero R, Ferrell C, Murphy S. Transit-oriented development and joint development in the United States: A literature review. TCRP Research Results Digest. 2002 [cited 27 Sep 2021]. Available: https://trid.trb.org/view.aspx?id=726711

34. Southworth Michael. Designing the Walkable City. Journal of Urban Planning and Development. 2005;131: 246–257.

35. Schutz E. Stadtentwicklung durch Hochgeschwindigkeitsverkehr (Urban development by High-Speed Railways). Heft; 1998.

36. Moon C, Sato S, Tonuma K. A STUDY ON THE CHARACTERISTICS OF URBAN IMPROVEMENT AROUND THE SHINKANSEN STATION IN LOCAL CITIES : An analysis of real conditions and development process of project. J Archit Plan . 1997;62: 137–145.

37. Matsumoto E, Ubaura M. A Study on Changes and Problems of Land Use around non-central Shinkansen Stations in the Suburbs. Journal of the City Planning Institute of Japan. 2013;48: 627–632.

38. Long J, Robertson C. Comparing spatial patterns. Geogr compass. 2018;12: e12356.

39. McGarigal K. Landscape Pattern Metrics. Wiley StatsRef: Statistics Reference Online. Chichester, UK: John Wiley & Sons, Ltd; 2014. doi:10.1002/9781118445112.stat07723

40. Li H, Reynolds JF. A Simulation Experiment to Quantify Spatial Heterogeneity in Categorical Maps. Ecology. 1994;75: 2446–2455.

41. Li H, Reynolds JF. A new contagion index to quantify spatial patterns of landscapes. Landsc Ecol. 1993;8: 155–162.

42. Getis A, Paelinck JHP. An Analytical Description of Spatial Patterns. Espace Geogr. 2004;tome 33: 61–68.

43. Wentz EA. A shape definition for geographic applications based on edge, elongation, and perforation. Geogr Anal. 2010;32: 95–112.

44. Uuemaa E, Antrop M, Roosaare J, Marja R, Mander Ü. Landscape metrics and indices: An overview of their use in landscape research. Living Rev Landsc Res. 2009;3. doi:10.12942/lrlr-2009-1

45. Sankey TT. Statistical Descriptions of Spatial Patterns. In: Shekhar S, Xiong H, Zhou X, editors. Encyclopedia of GIS. Cham: Springer International Publishing; 2015. pp. 1–8.

46. O'Neill RV, Krummel JR, Gardner RH, Sugihara G, Jackson B, DeAngelis DL, et al. Indices of landscape pattern. Landsc Ecol. 1988;1: 153–162.

47. Wong DWS. The Modifiable Areal Unit Problem (MAUP). In: Janelle DG, Warf B, Hansen K, editors. WorldMinds: Geographical Perspectives on 100 Problems: Commemorating the 100th Anniversary of the Association of American Geographers 1904–2004. Dordrecht: Springer Netherlands; 2004. pp. 571–575.

48. Jelinski DE, Wu J. The modifiable areal unit problem and implications for landscape ecology. Landsc Ecol. 1996;11: 129–140.

49. Openshaw S. The modifiable areal unit problem. Quantitative geography: A British view. 1981; 60–69.

50. Hargis CD, Bissonette JA, David JL. The behavior of landscape metrics commonly used in the study of habitat fragmentation. Landscape Ecology. 1998;13: 167–186.

51. Boots BN. Comments on the use of eigenfunctions to measure structural properties of geographic networks. Environ Plan A. 1982;14: 1063–1072.

52. Gustafson EJ. Quantifying Landscape Spatial Pattern: What Is the State of the Art? Ecosystems. 1998;1: 143–156.

53. TK Remmel, F Csillag, SW Mitchell, B Boots. EMPIRICAL DISTRIBUTIONS OF LANDSCAPE PATTERN INDICES AS FUNCTIONS OF CLASSIFIED IMAGE COMPOSITION AND SPATIAL STRUCTURE. In: Costas Armenakis YCL, editor. Geospatial Theory, Processing and Applications. the International Society of Photogrammetry and Remote Sensing (ISPRS); 2002. Available: https://www.isprs.org/proceedings/XXXIV/part4/pdfpapers/083.pdf

54. Remmel TK, Csillag F. When are two landscape pattern indices significantly different? J Geogr Syst. 2003;5: 331–351.

55. Bishop CM. Pattern recognition and machine learning. springer; 2006.

56. Helber P, Bischke B, Dengel A, Borth D. EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2019;12: 2217–2226.

57. Xu Y, Du B, Zhang L, Cerra D, Pato M, Carmona E, et al. Advanced Multi-Sensor Optical Remote Sensing for Urban Land Use and Land Cover Classification: Outcome of the 2018 IEEE GRSS Data Fusion Contest. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2019;12: 1709–1724.

58. Weigand M, Staab J, Wurm M, Taubenböck H. Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data. Int J Appl Earth Obs Geoinf. 2020;88: 102065.

59. Zhang P, Ke Y, Zhang Z, Wang M, Li P, Zhang S. Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery. Sensors. 2018;18. doi:10.3390/s18113717

60. Wijnands JS, Nice KA, Thompson J, Zhao H, Stevenson M. Streetscape augmentation using generative adversarial networks: Insights related to health and wellbeing. Sustainable Cities and Society. 2019;49: 101602.

61. Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, et al. The cityscapes dataset for semantic urban scene understanding. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016. doi:10.1109/cvpr.2016.350

62. Xia Y, Yabuki N, Fukuda T. Development of a system for assessing the quality of urban street-level greenery using street view images and deep learning. Urban For Urban Greening. 2021;59: 126995.

63. Bohannon RW. Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants. Age Ageing. 1997;26: 15–19.

64. Ma J, Jiang X, Fan A, Jiang J, Yan J. Image Matching from Handcrafted to Deep Features: A Survey. Int J Comput Vis. 2021;129: 23–79.

65. Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision. Cengage Learning; 2014.

66. Brunelli R. Template Matching Techniques in Computer Vision: Theory and Practice. John Wiley & Sons; 2009.

67. Simonyan K, Vedaldi A, Zisserman A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. arXiv [cs.CV]. 2013. Available: http://arxiv.org/abs/1312.6034

68. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. Computer Vision – ECCV 2014. Springer International Publishing; 2014. pp. 818–833.

69. Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H. Understanding Neural Networks Through Deep Visualization. arXiv [cs.CV]. 2015. Available: http://arxiv.org/abs/1506.06579

70. Nanni L, Ghidoni S, Brahnam S. Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recognit. 2017;71: 158–172.

71. Kabbai L, Abdellaoui M, Douik A. Image classification by combining local and global features. Vis Comput. 2019;35: 679–693.

72. Wang S, Han K, Jin J. Review of image low-level feature extraction methods for content-based image retrieval. Sens Rev. 2019. Available: https://www.emerald.com/insight/content/doi/10.1108/SR-04-2019-0092/full/html

73. Karami E, Prasad S, Shehata M. Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images. arXiv [cs.CV]. 2017. Available: http://arxiv.org/abs/1710.02726

74. Cao B, Araujo A, Sim J. Unifying Deep Local and Global Features for Image Search. Computer Vision – ECCV 2020. Springer International Publishing; 2020. pp. 726–743.

75. Monica G. Turner, Robert H. Gardner, Robert V. O'Neill. Neutral Landscape Models. 1st ed. In: Monica G. Turner, Robert H. Gardner, Robert V. O'Neill, editor. Landscape Ecology in Theory and Practice: Pattern and Process. 1st ed. Springer-Verlag New York; 2001. pp. 135–156.

76. Etherington TR, Holland EP, O'Sullivan D. NLM py: a python software package for the creation of neutral landscape models within a general numerical framework. Methods Ecol Evol. 2015;6: 164–168.

77. Defays D. An efficient algorithm for a complete link method. Comput J. 1977;20: 364–366.

78. Nielsen F. Hierarchical Clustering. In: Nielsen F, editor. Introduction to HPC with MPI for Data Science. Cham: Springer International Publishing; 2016. pp. 195–211.

79. Sokal RR, Rohlf FJ. The Comparison of Dendrograms by Objective Methods. Taxon. 1962;11: 33–40.

80. Kassambara A. Machine Learning Essentials: Practical Guide in R. STHDA; 2018.

81. Shapiro SS, Wilk MB. An Analysis of Variance Test for Normality (Complete Samples). Biometrika. 1965;52: 591–611.

82. Conover WJ. Practical Nonparametric Statistics. John Wiley & Sons; 1999.

83. 総務省統計局. 平成 27 年国勢調査 人口等基本集計結果. Available: https://www.stat.go.jp/data/kokusei/2015/kekka/kihon1/pdf/gaiyou1.pdf

84. Kaiser HF. An index of factorial simplicity. Psychometrika. 1974;39: 31–36.

85. Bro R, Smilde AK. Principal component analysis. Anal Methods. 2014;6: 2812–2831.

86. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat. 2010;2: 433–459.