

Wang, Jinfeng; Liu, Xuhua; Chen, Hongyan

Conference Paper

Construction of a Cities Evolution Tree, with Applications

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Wang, Jinfeng; Liu, Xuhua; Chen, Hongyan (2010) : Construction of a Cities Evolution Tree, with Applications, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/119179>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Construction of a Cities Evolution Tree, with Applications

Jin-Feng Wang^{1*}, Xu-Hua Liu¹, Hong-Yan Chen²

¹State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

²Institute for Sustainable Water Integrated Management & Ecosystem Research, University of Liverpool, England L69 3GP, UK

Abstract

Cartesian coordinate system helps to look insight of the geometric relationship between scattered data, we proposed a new reference by which the spatial data is rearranged and presented according to their evolution in time from the underling mechanistic process, that allows the spatial pattern in the next time is predicable along the evolution perspective based on the observed data in the present time. We test the novel approach by reconstructing a cities evolution tree and the tree is later used to analyze the impact of urbanization on land occupation. Cities are clustered by their functions and development stages, which is illustrated by a cluster tree, a dynamic tree that depicts the evolution of cities. The evolution tree in one year is used to predict the state of a city in a future time period. Another application of the evolution tree is to predict urban-type relevant phenomena, such as urban occupation. It is found that comprehensive cities, business cities, and manufacturing cities have higher urban expansion rates than tourist cities, with a few exceptions that focus on both industry and tourism. Meanwhile, the speed and extent of city land growth are dominated by industrialization stages and economic patterns, as well as leap-development. The methodology presented in this study is especially suitable for identifying transition paths of a stochastic process in a complex dataset of 253 cities in China.

Keywords Economic Type; Development Stage; Cluster Tree; Markov Chain; Land Growth

Land is the interface between population and environment ([Chen et al., 2007](#)), and cities are the very places where land use changes most rapidly ([Guy and Henneberry, 2000](#)). Human activities prompt land-use change. Land use refines the population distribution and economic activities, and impacts the physical environment of cities. As a primary source of soil degradation ([Tolba et al., 1992](#)), land-use changes affect biodiversity ([Tolba et al., 1992](#); [Sala et al., 2000](#)), lead to local and regional climate change ([Chase et al., 1999](#)), and exert enormous overall socioeconomic impacts ([Wicke, 2006](#)). Although exogenous, land-use changes are ecosystem responses to urbanization across climatic and societal gradients ([Grimml et al., 2008](#)). The exponential increase in urbanization has not only brought fundamental change in the industrial structure, employment structure, and consumption pattern, but also promoted transformation from agricultural economies to urban economies ([Krugman, 1991](#)). These transformations have fueled significant changes to the size of urban areas.

To achieve “environmentally non-degrading, technologically appropriate, economically viable and socially acceptable” (FAO Fisheries Department 1997) objectives, a series of interdisciplinary / multidisciplinary research projects have been conducted in the fields of LUCC (land-use and land-cover change) and environmental sciences. Urbanization has been the most important determinant of land occupation in the last decades (Liu *et al.*, 2005). The existing studies on this topic focus on two aspects: identifying driving forces of land occupation and developing theory and models. The following factors have been recognized as determinants of land occupation: physical geography (Chen *et al.*, 2007), population (Xie *et al.*, 2007), and income (Deng *et al.*, 2008), along with other population variables, values of agricultural land, and transportation costs. As for theory and modeling, earlier studies have focused on quantitative and aggregate measurements and adopted mainly spatial metrics and urban models; the recent core-periphery model (Krugman, 1991; Fujita *et al.*, 1999) treats the geographic agglomeration of regions as a result of economic returns. Another major issue is statistical estimation in quantifying spatiotemporal development. Non-parametric methods (Capilla, 2008) such as local polynomial kernel regression and wavelet threshold have been developed to analyze trends and effects in urban areas, and there has also been an investigation of Bayesian maximum entropy (Christakos, 2005; Lee, 2008; Choi, 2008) in processing both continuous and categorical variables. Most of the models originate from the IPAT equation (Ehrlich and Holdren, 1971), which models a linear relationship between land use and income, population, agricultural land values, and transportation costs.

In this study, we developed a novel framework to investigate urbanization and land occupation. Land occupation is suspected to be related with the city type and developmental stage, so the cities are clustered according to their function and stage, called state, and illustrated in a tree-like evolution structure. The suspected association between city state and land occupation was confirmed by a correlation test in the framework of the city evolution tree.

The remainder of the paper is organized into five sections. We first introduce our data and methodology. This is then followed by a description of how to build the city evolution tree. Two applications of the evolution tree to predicting city evolution and interpreting urban land occupation are presented in two separate sections. We conclude the paper with a summary of our main findings and a discussion of future research directions.

Data and methodology

Systematic studies of the determinants of urbanization have rarely been done in developing countries due to data limitations. Large, nationwide, econometric studies using high quality data are rare or expensive (Deng 2008). The latest exhaustive surveys of the whole of China were in 1990 and 2000 and the next one is scheduled in 2010. In this study, a database composed of 253 cities has been built based on the Chinese Statistical Yearbooks for 1990 and 2000, the national population census 2000, and the China Development Zones Yearbook 2003. The observations for cluster analysis are 16 industries specified by the 2000 national population census, namely, “Farming, forestry, Animal Husbandry and Fishery”, “Mining and Quarrying”, “Manufacturing”, “Production and Supply of Electric Power, Gas and Water”, “Construction”, “Geological Prospecting and Water”, “Transport, Storage, Post and Telecommunications”, “Wholesale and Retail Trade &

Catering Services”, “Banking and Insurance”, “Real Estate”, “Social Services”, “Health Care, Sporting and Social Welfare”, “Education, Culture and Arts, Radio, Film and TV”, “Scientific Research and Polytechnical Services”, “Governments, Parties and Social Organization”, and “Others”.

The land use database was obtained from the Chinese Academy of Sciences. The data are from satellite remote sensing data provided by the US Landsat TM images with a spatial resolution of 30 by 30 meters. The database includes time-series data for 1990 and 2000. The data is stored in a GIS map at a 1:100,000 scale.

Figure 1 illustrates the methodology of the study. The 253 cities are categorized into 8 functional groups using a clustering algorithm and each of the groups is divided into six stages according to economic level. The hierarchical clusters are displayed as an evolution tree. The tree-like structure maps the functions and stages of the cities on 8 branches representing urban types. The leaves stand for cities, and the position and shade of the leaves denote the economic developmental stage. The tree grows over time with the dynamics of the cities (see Fig. 2 and detail in the next section). The evolution tree of cities can be used for predicting and interpreting urban state-related phenomena, such as the evolution of a city, interpretation of urban sprawl, and a city’s disease spectrum. A Markov chain is constructed by comparing the two evolution trees for 1990 and 2000 and then used to test the prediction of city stages in 2000 using the 1990 evolution tree. Furthermore, a correlation analysis between the land occupation and the city states in the framework of the city evolution tree found that land expansion was highly associated with the developmental stages and with functions of cities.

Fig. 1 here

Evolution tree of cities

As urbanization and growth proceed, both the absolute sizes and number of cities are unprecedented. Metro areas are expanding and satellite centers are springing up in the pursuit of economic development and the improvement of living conditions. The process of economic growth is very imbalanced in scale and pace between regions and presents a very diverse, complicated picture. In the following section we will examine the states of 253 Chinese cities during 1990 and 2000, aiming to explore characteristics of land change in different phases of urbanization and economic transformation.

Economic stage of cities

Referring to the structural change model (Chenery, 1979) and Northam Curve (Northam, 1979), we adopted such indicators as GDP per capita, industrial structure, employment structure, and urbanization degree to analyze structural change at different stages of economic development and urbanization. A classification algorithm to the four indicators results in the six stages of evolution, summarized in Table 1.

Tab. 1 here

Table 1 shows that GDP per capita is highly associated with the composition of the industrial and

employment structures and the urbanization level. The latter is measured by the percentage of non-farm population in cities. Furthermore, according to the composition features, we propose six city stages in the last column of [Table 1](#). [Table 1](#) illustrates the following points: (1) tertiary industry develops rapidly from elementary stages to the middle stage of industrialization, but slows down from middle stage to first stage of an advanced economy; (2) employment structure maintains steady, fast growth at both the middle stage and maturity stage; (3) and unlike traditional industrial structures in other countries, China has exhibited a unique development pattern. High priority was given to the development of heavy industry for a long time, but secondary and tertiary industries have been delayed. As a result, unbalanced development and serious polarization have hindered economic growth and impaired the function of cities. This implies that the pattern and extent of urbanization can be quantified in different economic stages by measuring GDP per capita, industrial and employment structure, and urbanization degree.

Function types of cities

Sixteen sectors and eight factors from government statistics that are connected to or overlap with one another were checked with a multi-linearity test and categorized into ten subsets: “Mining and Quarrying”, “Manufacturing”, “Production and Supply of Electric Power, Gas and Water”, “Construction”, “Geological Prospecting”, “Transport, Storage, Post and Telecommunications”, “Business”, “Governments, Parties and Social Organization”, “Weight of Mining and Quarrying”, and “Functional Index Of Tourism”. Cities are hierarchically clustered by three standards. The first criterion is to choose results with high pseudo F values, a small number of clusters, and small pseudo t^2 values, then classify the observations into a finite number of groups, i.e., 5, 8, 10, 12, and 14. Second, larger R^2 values and a smaller number of groups is preferable, thus observations are placed into 8, 9, 10, 11, and 12 groups. Third is a cubic clustering criterion ([Milligan et al 1985](#)) recommended by SAS, the result of which is 8, 9, 10, 11, 12, and 13 groups. Clustering algorithms in accordance with the three criteria yield three groupings. Intersections of the three criteria are 8, 10, and 12 groups. Apart from general indicators, cities that belong to the same category and grow at similar exponential rates at similar economic development stages should have small standard deviations. Having compared the standard deviation of cities based on 8-, 10-, and 12-group methods, it was found that the 8-group method had the smallest standard deviation and is the optimum size for city classification.

Considering its limitations, conventional cluster analysis is supplemented with the Nelson Classification method to identify the prominent function of city groups. Nelson Classification uses the arithmetical mean and standard deviation of employees to determine the superiority functions of the cities. The prominent functions of a city are identified as being at least one standard deviation above the average, while the function intensity can be measured by the number of standard deviations above average. [Table 2](#) lists the arithmetical mean and standard deviation of every industry of every type. Nelson Classification, however, is insufficient for discriminating the average function of city groups. To supplement its weakness, the weight ratio of employees is introduced as an indicator to assess the function of industries. Based on the results of the K-means partition and city function estimated by Nelson classification, the 253 cities are divided into eight categories ([Table 2](#)): I. small comprehensive cities without prominent industries; II. comprehensive cities with advantages in transportation, construction, electric power, gas, and water; III. mining cities; IV. middle-size cities with advantages in business; V. comprehensive cities with advantages in administration;

VI. tourism cities with advantages in business; VII commercial cities with advantages in manufacturing; and VIII. manufacturing cities.

Tab. 2 here

City evolution tree

We developed a highly specialized evolution tree to display the above clustering of function types and development stages of the 253 cities (Fig. 2), where the synthesis stage is derived from GDP per capita, industrial structure, employment structure, and urbanization degree. A label attached on every leaf represents city name and city classification code. Roman numerals in the stem (Fig. 2) are urban function types (Table 1). Generally, a higher position on the branch indicates a more advanced economic development stage. Moreover, cities in branches and stems are arranged in accordance with GDP per capita. As GDP per capita increases, cities will progress: the closer to stem, the higher GDP, and vice versa.

Fig. 2 here

The evolution tree provides a new coordinate system to display a tree structure of the urban system from the systematic evolution perspective. Various cities may seem to develop in a complicated, nonlinear, random way (Harte, 2006) but are actually connected and evolve quite regularly when viewed on the evolution tree. Each city in the evolution tree undergoes development along the distinct path of its type. As the tree grows, leaves move onward to the next position, and some may jump from one branch to another. In the novel coordinate system of the evolution tree, the future of a city is highly predictable.

There are many issues in population and environment that are related to city state evolution: pollution composition (Cole, 2005), disease spectrum (Mead and Brajer, 2005) and care systems, migration patterns (Neil et al., 2005; Gutman et al., 2006), and urban spawn, to name a few. These issues should be investigated from this new perspective.

Application I: prediction of city evolution

A difference between the two evolution trees of the 253 cities discloses a switch of city states (types and stages) from 1990 to 2000. The results are partly listed in Table 3 and expressed by the Markov chain in Figure 3.

Tab. 3 here

Fig. 3 here

Explanation of variables in Figure 3:

- Bracket: (number of cities in which stages transform, average city expansion rate)
- Roman numerals in the squares: urban function type (Table 2). That is: I = Small comprehensive cities without prominent industries; II = Comprehensive cities advantageous in the Transport, Construction,

Electric Power, Gas and Water; III = Mining cities; IV = Middle-size cities advantageous in business; V = Comprehensive cities advantageous in administration; VI = Tourism cities advantageous in business; VII = Commercial City advantageous in manufacturing; VIII = Manufacturing cities

- Arabic numerals in circles: economic stages, as [Table 1](#).
- Arrows begin at economic stages at 1990 and points to economic stages at 2000
- Dashed line “a” represents examples of regime switches and the dynamics of city evolution. It represents the switch from type I to type IV. Both types belong to comprehensive cities; the only difference lies in the intensity of function. As cities grow, small-sized comprehensive cities without a dominant function will gradually developed into middle-sized comprehensive cities.
- Dashed lines “b” and “c”: when mining cities age, they will transform, either changing from chemical industry cities to manufacturing cities or from communicative cities into business cities.
- Dashed line “d”: transformation from tourist cities into business cities.

The evolution tree of cities in 1990 ([Fig. 2](#)) is used to predict the direction and path of the evolution of cities along the branches. Most of the cities will go forward, although a few may go backwards or be stalled. The expectation is approximately confirmed by the Markov chain in [Figure 3](#) and [Table 3](#). For example, Beijing advanced from stage 4 in 1991 to stage 5 in 2000, and a few cities jumped from one type (branch) onto another (a, b, c, d in [Fig. 3](#), for example).

Application II: understanding land occupation growth

Although there is no shortage of documentation on the process of land use transformation as a consequence of economic growth and urbanization ([Healey et al., 1990, 1991](#); [Heilig, 1994](#), [Guy and Henneberry, 2000](#)), the evolution tree developed in this study implicates the rules underlying urban land utilization and causes of land expansion.

The expansion of cities is an endogenous growth along with urbanization and industrialization, and demonstrates different features in different development phases. City expansion is controlled by a balance between two forces ([Krugman, 1991](#); [Wang, 1993](#)):

$$P_v - P_l > 2\sigma$$

where σ is force that attracts inward resources, P_v corresponds to an outward pushing force, and P_l is inward resistant power. In more detail, (i) manufacturing cities and business cities develop at a faster pace at elementary stages, and constantly make breakthroughs when the balance is broken by interior pressure and exterior concussion. As long as exterior factors (government policies) are permitted, these cities expand rapidly even in higher stages of economic development; (ii) cities with a weak administration function indicate a small outward pushing force P_v ; (iii) single-function cities such as tourism cities are fairly insensitive to the economy stages and maintain low expansion rates even in the advanced economic stage; (iv) land in mining cities does not change according to economic development, but is restricted by the natural resources (e.g., recoverable reserves of coal). As a result, the dominant industry of mining cities is considered to experience the process of “prime development, maturity, and decline”. To solve this problem, mining cities usually support the second dominant industry when the dominant industry declines.

The density and complexity in metropolitan areas are measured with remote sensing data, while land area change is computed by an index (Liu et al., 2002). Figure 4 maps the land expansion rate observed. The sporadically dispersed spots are the 253 cities, while the grey scale represents the expansion rate.

Fig. 4 Here

The statistical results have confirmed the impact of urban state on land changes. The urban expansion rates of the different city types are listed in Table 4. There is a large gap in land expansion patterns between highly urbanized cities and less urbanized cities. There are also large differences with respect to land expansion rates among different urban types.

Tab. 4 here

Explanation of variables in Table 4:

- “Expansion Rate” = $(\text{area in 2000} - \text{area in 1990}) / \text{area in 1990}$;
- “Subset Code”: the three digit code is (type code, phrase 1990, phrase 2000). The 253 cities are categorized into 60 subsets, which are counted as “Number of cities”.
- “min”, “max”, “average” and “Std”: minimum, maximum, average and standard deviation values of city expansion, respectively.

Shenzhen, Beijing, and Shanghai are shaded dark grey because they have undergone economic growth of unprecedented scale and speed. There is a positive correlation between land expansion, city function, and economic phases (Table 5a; explanation of parameters in Table 5b). The correlation between city expansion rate and city type is 0.4 and the significance level is below 0.01. Except at certain phases of economic development, tourist cities, commercial city, and manufacturing cities display high variation in city expansion rate estimation. The city expansion rate of most cities at higher stages is acceptable. The large fluctuation is due to a high average mean and complex structure, as well as government policies. Cultivated land in the eastern cities of China decrease sharply, and the correlation between town expansion and land decrease is as high as 0.88. This is mainly because the increasingly optimized Chinese macro-economic environment has attracted international investment in China, especially in eastern cities that have solid economic foundations. In short, the land expansion rate in business-oriented cities strongly correlates with economic policies. Exceptions do exist, however, with particular cities and particular types. Specifically, among tourism-oriented cities, the city expansion rate of Yangzhou is as high as 22.97%, and that of Suzhou has reached 38.8%.

Tab. 5a here

Tab. 5b here

The above observation can be interpreted under the framework of the evolution tree of cities in Figure 2 and Markov chain of state transfer in Figure 3. The advanced economic stage is usually positively related to

the city expansion rate. Although type II “Transport, construction comprehensive cities” and type III “mining cities” have entered into the higher stage, urban construction land did not increase proportionally to economic stages. This indicates that, compared with economic stages, city type is more closely related to city expansion. Expansion of cities has different driving forces but is determined by economy development and is an endogenous process of industrialization and urbanization. From the Markov chain (Fig. 3) and evolution tree (Fig. 2) we find that the urban types of I (small comprehensive cities), III (mining cities), and IV (middle-size comprehensive cities) remain at lower stages of economic development; and VII (commercial cities) and VIII (manufacturing cities) have evolved into a higher stage of industrialization. Land expansion of administrative cities, commercial cities, and manufacturing cities is relatively high compared with the small comprehensive cities, mining cities, and tourist cities. All city types follow common trends. First, the more advanced they are in the stage of industrialization, the higher their city expansion rate. Second, cities with a faster pace in economic development have higher expansion rates. The Markov chain plays an auxiliary role in diagnosing the urban development trends by quantifying land, economic stages, and urban functions. It was found that when reaching higher stages of economic development, business cities demonstrate obvious land expansion. In the long term, there is a type switch among the eight types of cities. We propose some assumptions prompted by the Markov chain (Fig. 3) and evolution tree (Fig. 2). First, cities that belong to same function group and similar levels of economic development are inclined to follow the same routes. Second, within the same function catalog, cities that are at the lower stage of development will follow the poorer cities and the less-urbanized cities will follow the footprint of urbanized cities. Finally, due to economic integration and fluctuation, some cities’ cluster groups will switch economic stages and function types during their rapid catch-up growth and spark a region-wide growth cluster.

Conclusion and discussion

The interaction between population and environment always occurs at a space, and cities provide the most unique environment where people reside at the highest density. So the evolution of the cities and their sprawling size are the most concerns of humans.

Either traditional or spatial econometric models (Gujarati, 1995; Anselin, 1988; Wang, 2006) have been used to investigate urbanization and land occupation. The “IPAT equation” (Ehrlich & Holdren, 1971) is a widely used prototype to estimate environmental impact and land use from their determinants, e.g., population size, per-capita affluence level, and technologies. The model is good in its simplicity, but it ignores that the impact is nonlinear and varies with the types of industrial activities.

In this study, we proposed a novel framework in which the complicated city system has been systematically and harmoniously organized into a tree-like structure, with the branches representing the types and the positions of cities and the leaves on the branches denoting the developmental stage of the cities. The tree unveils city relationship evolution: connections between the cities, their neighbors, and the evolutionary path. The evolution tree, in part at least, explains the dynamics of urbanization and its related phenomena, and overcomes the difficulty in handling nonlinear processes using global statistic and pooled data (Harte, 2007). The evolution tree contrasts with Cartesian coordinators and maps, which are good at

coordinating and illustrating relatively simple phenomena and relationships. The validation of the new approach to handling the sophisticated urbanization data has been tested with the available data. The city evolution tree can be used to investigate urban state-related phenomena, such as prediction of city evolution, urban land occupation, and spatial distribution of pollution.

One application of the evolution tree is to assess land occupation by the different city types and possible future occupation. The assessment found that (1) urban construction land is significantly associated with economic stages; (2) urban construction land area increases sharply as urbanization enters into the mature stage and the economy accelerates its pace of development; (3) there is an association between city type and urban land expansion; (4) administrative cities, business cities, and manufacturing cities display higher land expansion rates, while small-sized comprehensive cities, mining cities, and tourist cities indicate lower land expansion rates; and (5) the Markov chain and evolution tree are effective in predicting the stationary process, where economic stage jumps from the initial stage to the next along with the conditional probability of transition.

Despite increasing returns to scale (Krugman, 1991), rapid development of cities has outpaced the available supply of natural resources. The connection between larger and better has been broken. Many successful cities suffer from overcrowding, overexploitation, and housing shortages. The solution lies in the equilibration and optimization of city scale. It is helpful to use the evolution tree and GIS to support intelligent city administration.

Acknowledgements: This study was supported by CAS (KZCX2-YW-308) and MOST (2007AA12Z233; 2007DFC20180). We also thank Ma Aihua in preparing the figures.

References

- Capilla, C. (2008). Time series analysis and identification of trends in a Mediterranean urban area. *Global and Planetary Change*, 63, 275-261
- Chase, T.N., Pielke, R.A. Sr., Kittel, T.G.F., Baron, J.S., and Stohlgren, T.J. (1999). Potential impacts on Colorado Rocky Mountain weather due to land use changes on the adjacent Great Plains. *J. Geophys. Res.* 104, 16673-16690.
- Chen, M., Xu, C., Wang, R. (2007). Key natural impacting factors of China's human population distribution. *Popul Environ*, 28, 187-200.
- Chenery, H. and Syrquin, M. (1979). *Structural Change and Development Policy*. Oxford University Press.
- Choi, K. M., Yu H. L., Wilson, M. L. (2008). Spatiotemporal statistical analysis of influenza mortality in the State of California during the period 1997-2001, *Stochastic Environmental Research and Risk Assessment*, 22 (Supplement 1).
- Christakos, G. (2005). *Random Field Models in Earth Sciences*. NY: Dover Publ.
- Cole, M., Neumayer, E. (2004). Examining the impact of demographic factors on air pollution. *Population and Environment*, 26, 5-21.
- Deng, X., Huang, J., Rozelle, S., Uchida, E. (2008). Growth, population and industrialization, and urban land expansion of China. *Journal of Urban Economics*, 63, 96-115
- Ehrlich, P., & Holdren, J. (1971). The impact of population growth. *Science*, 171, 1212-1217.
- Fujita, M, Krugman, P., Venables, A. (1999). *The Spatial Economics: Cities, Regions, and International Trade*, Cambridge, Mass: The MIT Press.
- Grimml, N. (2008). The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. *Front Ecol Environ* 6, 264-272.
- Gujarati, D. (1995). *Basic Econometrics (Third Edition)*. NY: McGRAW-HILL, Inc.
- Gutmann, M., Deane, G., Lauster, N., Peri, A. (2006). Two population-environment regimes in the great plains of the United States, 1930-1990. *Population and Environment*, 27, 191-225.
- Guy, S.; Henneberry, J. (2000). Understanding urban development processes: Integrating the economic and the social in property research. *Urban Studies*, 37, 2399-2416.
- Harte, J. (2007). Human population as a dynamic factor in environmental degradation. *Popul Environ*, 28, 223-236.
- Healey, P.; Barrett, S.M. (1990). Structure and agency in land and property development processes: some ideas for research. *Urban Studies*, 27, 89-104.

- Healey, P. (1991). Models of the development process: a review. *Journal of Property Research*, 8, 219-238.
- Heilig, G. (1994). Neglected dimensions of global land-use change: reflections and data. *Population and Development Review*, 20, 831-859.
- Herold, M. (2003). The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment*, 86, 286-302
- Hong, Y. (1996). Consistent testing for serial correlation of unknown form, *Econometrica*, 64, 837-864.
- Krugman, P. (1991). Increasing returns and economic geography. *The Journal of Political Economy*. 99, 483-499.
- Lee, S. J., Balling, R., Gober, P. (2008). Bayesian maximum entropy mapping and the soft data problem in urban climate research. *Annals of the Association of American Geographers*, 98, 309 – 322.
- Liu, J. Y., Liu, M. L., Zhuang, D. F., et al. (2002). Study on spatial patterns analysis of recent land-use change in China. *Science in China (Series D)*, 2, 1031-1040.
- Mead, R., Brajer, V. (2005). Protecting China's children: valuing the health impacts of reduced air pollution in Chinese cities. *Environment and Development Economics*, 10, 745-768.
- Nelson, R. R. (2006). *Economic Development from the Perspective of Evolutionary Economic Theory*. Columbia University Press.
- Northam, R. M. (1979). *Urban Geography*. New York: John Wiley & Sons.
- Sala, O. E., Chapin, F. S., Armesto, J. J., Berlow, E. et al. (2000). Global biodiversity scenarios for the year 2100. *Science*, 287, 1770—1774.
- Taubenböck, H., Wegmann, M., Roth, A., Mehl, H., Dech, S. (2008). Urbanization in India – Spatiotemporal analysis using remote sensing data. *Computers, Environment and Urban Systems*, doi:10.1016/j.compenvurbsys.2008.09.003.
- Tolba, M. K., El-Kholy, O.A., El-Hinnawi, E., Holdgate, M. W., McMichael, D. F., Munn, R. E. (1992). *The World Environment 1972-1992: Two Decades of Challenge*. Chapman and Hall, London.
- Wang J.F. (1993). *Regional Economics Modeling*, Beijing: Science Press.
- Wang J.F. (2006). *Spatial Analysis*, Beijing: Science Press.
- Wicke, B., Faaij, A., Smeets, E. (2007). The socio-economic impacts of large scale land use change and export-oriented bio-energy production in Argentina. *The 15th European Biomass Conference & Exhibition*, 7-11 May 2007, Berlin, Germany. 3031-3035.
- Xie, Y., Fang C., Lin G., Gong, H., Qiao, B. (2007). Tempo-spatial patterns of land use changes and urban development in globalizing China: a study of Beijing. *Sensors*, 7, 2881-2907.

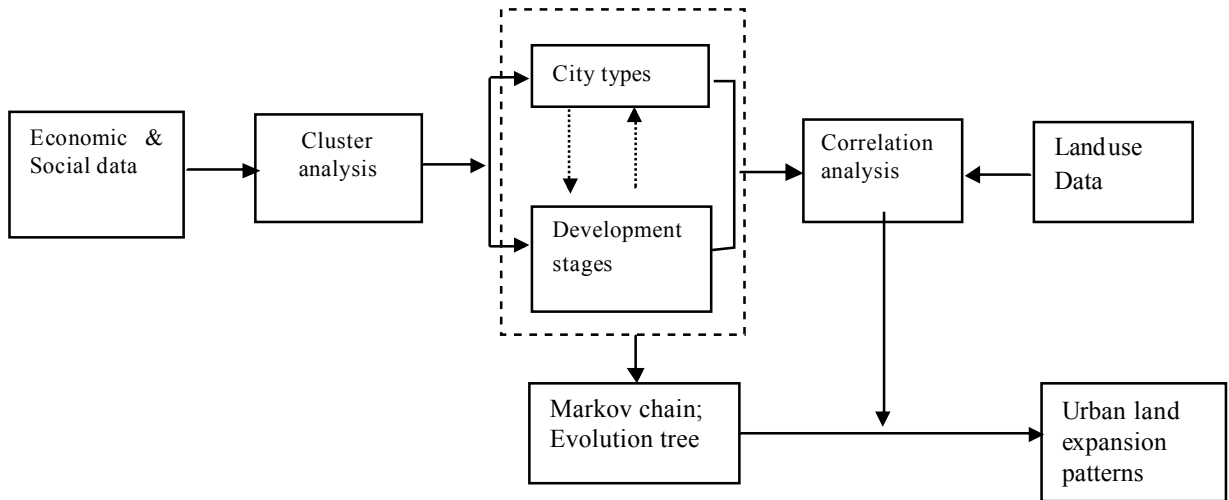


Figure 1 Methodology of the study

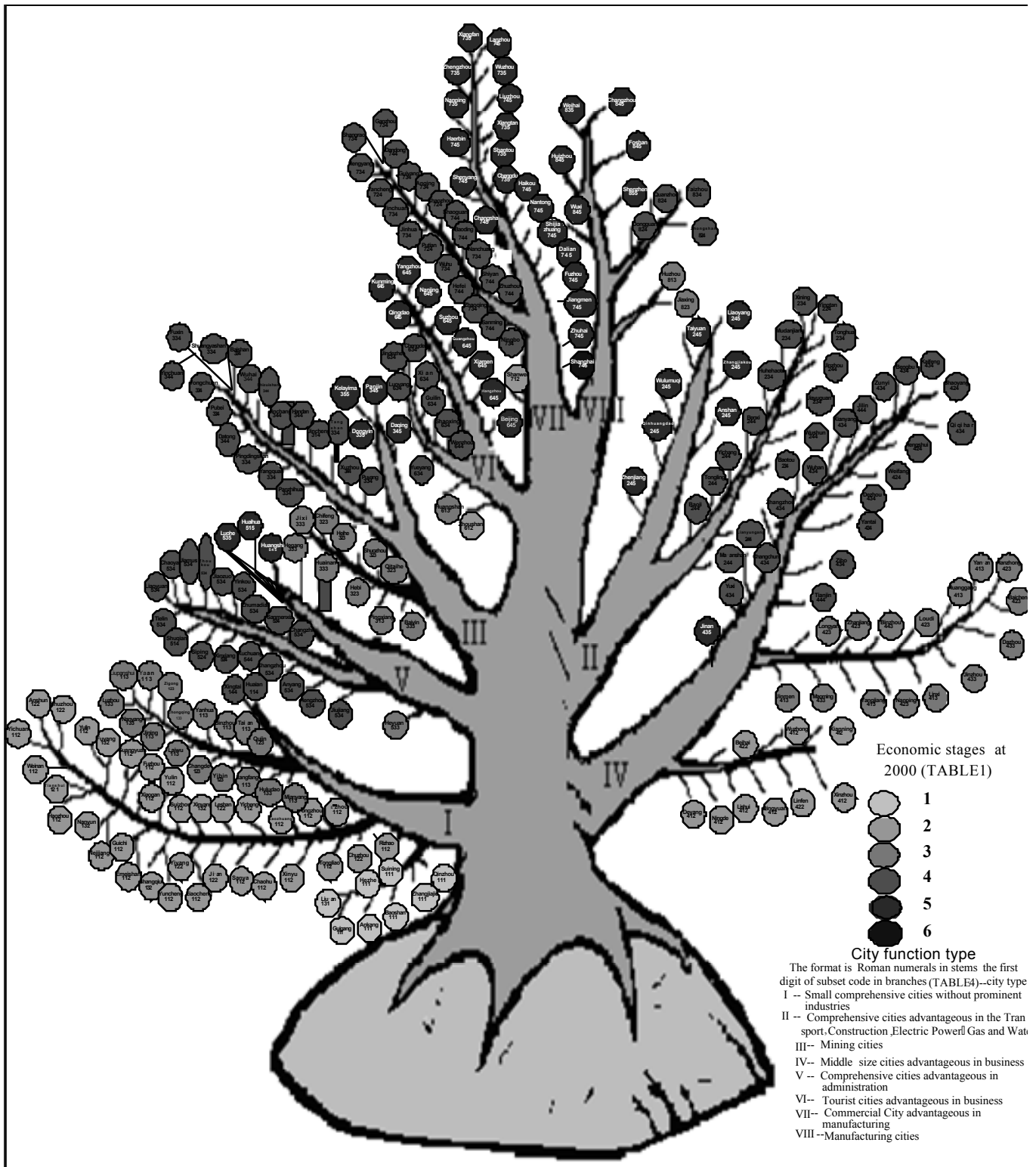


Figure 2a City evolution tree

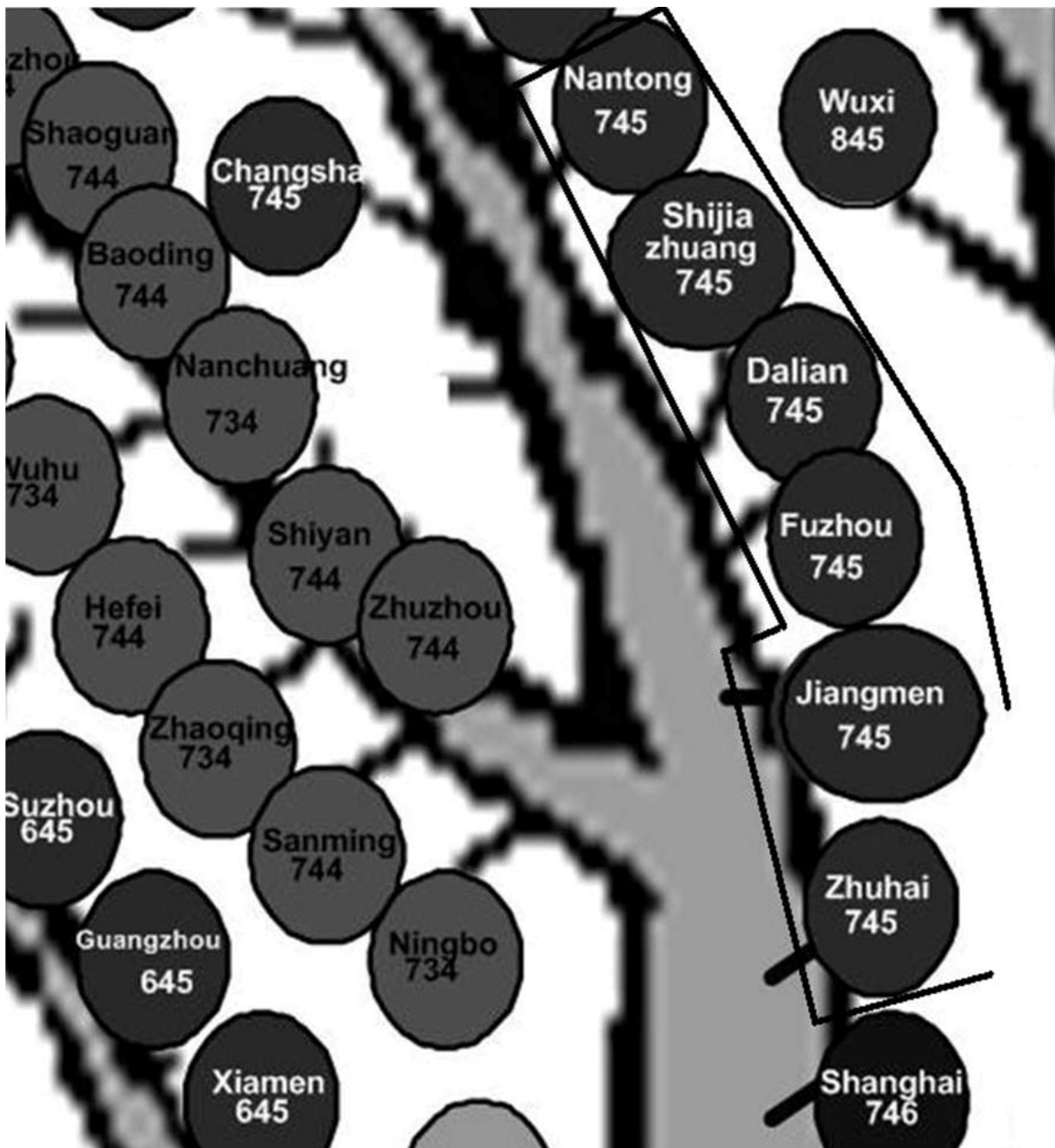


Figure 2b City evolution tree (part enlarged)

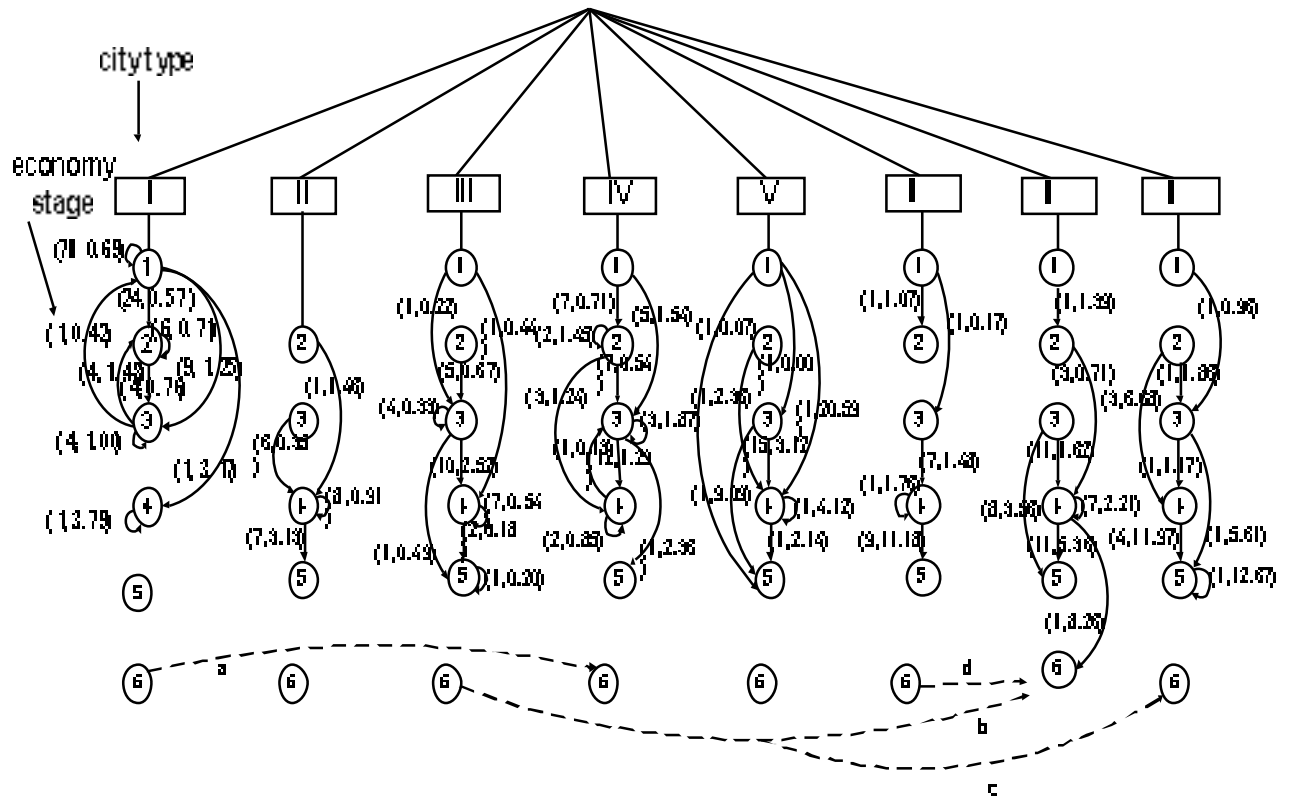


Figure 3 Markov chain of city economy stages, city expansion, and stage transformation

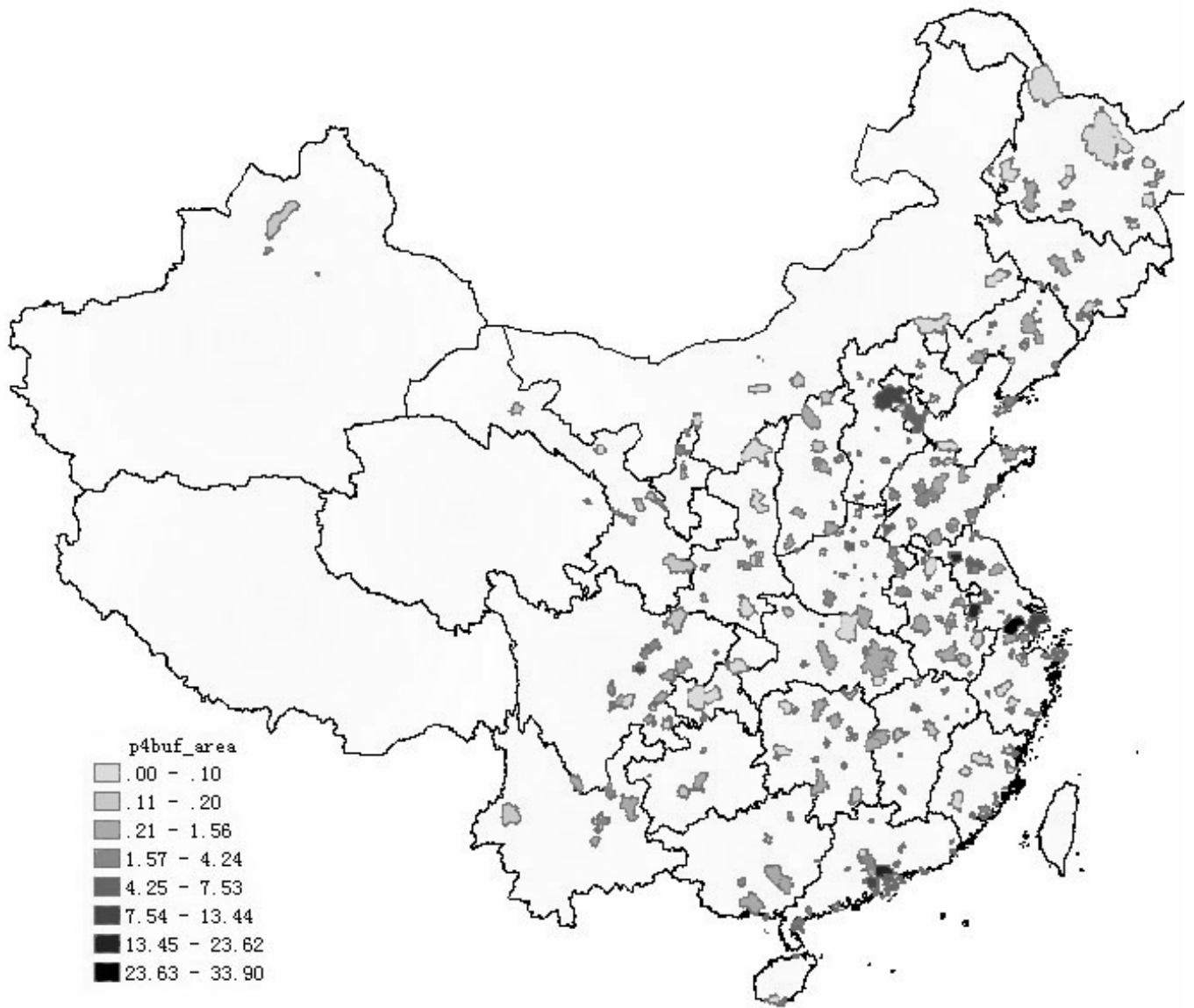


Figure 4 Observed Chinese urban dynamic expansion rate from 1990 to 2000

Table 1 Standards for classifying economic stages

Stages	GDP per capital in 1990 (USD)	Industrial Structure (%)			Employment Structure (%)			Urbanization Level (%)	Economy developmental stages	
		primary industry	secondary industry	tertiary industry	primary industry	secondary industry	tertiary industry			
1	300-600	38	26	36	65	17	18	5	Elementary products	
2	600-1200	29	32	39	57	20	23	30	elementary stage	stage of industrialization
3	1200-2400	20	40	40	50	22	28	40	middle stage	
4	2400-4500	13.5	46	40.5	36.5	25.5	38	54	advanced stage	
5	4500-7200	8	51	40	20	30	50	70	elementary stage	developed stages
6	7200-10800	3	47	50	8	30	62	80	elementary stage	

Table 2 Mean, standard deviation, and sample size of industries in China

City type	number of cities	Sectors	Manufacture	Electric Power, Gas and Water	Construction	Geology	Transport, Storage, Post and Telecommunications	Business	Governments and Social Organization	Third industry	Mining and Quarrying	Weight of Coal Mining	Weight of Tourism
		Statistics											
□	61	mean	9.03	0.66	2.47	0.16	2.73	6.77	2.60	6.94	0.91	3.99	3.44
		standard deviation	4.04	0.32	1.20	0.11	0.90	2.02	0.81	1.97	1.29	6.09	7.04
□	22	mean	26.94	2.84	7.57	0.42	7.96	15.57	5.18	17.42	2.71	2.31	1.82
		standard deviation	6.17	0.67	1.20	0.30	2.10	2.70	1.03	2.94	2.76	3.35	5.88
□	32	mean	16.20	3.07	4.60	0.48	6.49	11.42	5.06	14.18	13.66	42.23	1.56
		standard deviation	6.20	0.97	1.61	0.47	2.15	2.88	1.68	3.19	6.62	24.39	5.15
□	43	mean	18.24	1.27	4.19	0.33	4.71	11.68	4.31	12.03	0.98	2.22	0.93
		standard deviation	6.81	0.49	1.21	0.25	1.23	1.99	0.94	2.67	1.54	3.54	3.66
□	22	mean	24.63	2.62	4.07	0.44	6.63	14.20	6.99	15.74	1.49	3.61	0.45
		standard deviation	7.02	0.84	1.16	0.22	1.35	2.64	1.44	1.65	2.29	5.95	2.13
□	19	mean	30.65	1.16	6.07	0.18	5.19	16.73	4.13	17.50	0.70	1.50	28.42
		standard deviation	10.60	0.53	1.74	0.11	1.07	3.26	1.20	4.26	1.18	2.72	8.98
□	42	mean	28.93	1.45	5.94	0.26	5.73	17.93	5.01	17.86	0.53	1.25	0.95
		standard deviation	7.49	0.40	1.43	0.14	1.13	2.27	0.98	3.71	0.76	2.61	2.97
□	12	mean	48.03	0.76	5.43	0.12	3.28	13.95	2.76	10.71	0.26	0.24	0.83
		standard deviation	12.56	0.34	2.08	0.08	0.86	3.32	1.00	3.62	0.39	0.42	2.89

Table 3 Economic stages and urban function of Chinese cities in 1990 and 2000

Stages& types City Name	Developmental stage at 1990					Developmental stage at 2000					City Type	
	GDP per capita	industrial structure	employment structure	urbanization degree	synthesis stage	GDP per capita	industrial structure	employment structure	urbanization degree	synthesis stage	city type	city subset
Beijing	3	5	4	4	4	5	6	6	4	5	6	645
Tianjin	3	4	4	4	4	4	5	4	4	4	4	444
Shijiazhuang	3	4	4	5	4	5	6	4	5	5	7	745
Tangshan	2	4	4	4	3	4	4	4	4	4	3	334
Qinhuangdao	3	5	4	4	4	5	6	5	4	5	2	245
Handan	3	4	4	4	4	4	4	4	4	4	3	344
Xintai	3	4	4	4	4	3	4	4	5	4	1	144
Baoding	3	4	4	4	4	4	5	4	4	4	7	744
Zhangjiakou	3	4	4	4	4	4	5	4	5	5	2	245
.....												

Note: the number corresponds to city type. Arabic numerals are identical to Roman numerals. The city subset has same format as Tab 4 (please refer to explanation in session “Markov Chain”. We have computed two year value in 253 cities of China for all parameters. Any questions or comments on the data contents, please contact Jinfeng Wang: wangjf@Lreis.ac.cn

Table 4 Urban expansion rate and regression results of subset cities (by P4-area)

subset Code	number of cities (frequency)	min	max	average	Std
I-11	7	0.03	2.33	0.69	0.84
I-12	24	0.00	1.57	0.57	0.55
I-13	9	0.07	3.23	1.25	1.15
I-14	1	3.17	3.17	3.17	.
I-22	6	0.08	1.64	0.71	0.62
I-23	4	0.29	1.07	0.76	0.35
I-31	1	0.42	0.42	0.42	.
I-32	4	0.47	2.66	1.48	1.03
I-33	4	0.42	2.52	1.00	1.02
I-44	1	3.79	3.79	3.79	.
II-24	1	1.46	1.46	1.46	.
II-34	6	0.00	1.71	0.35	0.67
II-44	8	0.42	2.01	0.91	0.55
II-45	7	0.19	8.06	3.13	3.00
III-13	1	0.22	0.22	0.22	.
III-14	1	0.44	0.44	0.44	.
III-23	5	0.06	1.64	0.67	0.63
III-33	4	0.04	1.10	0.33	0.51
III-34	10	0.00	9.80	2.52	3.21
III-35	1	0.49	0.49	0.49	.
III-44	7	0.03	1.64	0.54	0.69
III-45	2	0.00	0.35	0.18	0.25
III-55	1	0.20	0.20	0.20	.
IV-12	7	0.00	3.79	0.71	1.39
IV-13	5	0.08	3.60	1.54	1.49
IV-22	2	0.38	2.51	1.45	1.51
IV-23	7	0.00	1.88	0.54	0.65
IV-24	3	0.49	2.37	1.24	1.00
IV-33	3	0.19	4.52	1.87	2.32
IV-34	12	0.07	3.84	1.29	1.18
IV-35	1	2.36	2.36	2.36	.
IV-43	1	0.13	0.13	0.13	.
IV-44	2	0.74	0.95	0.85	0.15
V-13	1	0.00	0.00	0.00	.
V-14	1	20.59	20.59	20.59	.
V-15	1	0.07	0.07	0.07	.
V-24	1	2.36	2.36	2.36	.
V-34	15	0.05	12.02	3.12	3.20
V-35	1	9.09	9.09	9.09	.
V-44	1	4.12	4.12	4.12	.
V-45	1	2.14	2.14	2.14	.
VI-12	1	1.07	1.07	1.07	.
VI-13	1	0.17	0.17	0.17	.
VI-34	7	0.14	3.68	1.48	1.35
VI-44	1	1.76	1.76	1.76	.
VI-45	9	1.76	38.78	11.18	12.97
VII-12	1	1.39	1.39	1.39	.
VII-24	3	0.00	1.44	0.71	0.72
VII-34	11	0.00	5.23	1.62	1.53

VII-35	8	0.00	8.09	3.56	3.27
VII-44	7	0.00	8.01	2.21	3.22
VII-45	11	0.00	21.19	5.36	6.76
VII-46	1	8.26	8.26	8.26	.
VIII-13	1	0.96	0.96	0.96	.
VIII-23	1	1.86	1.86	1.86	.
VIII-24	3	3.21	11.28	6.68	4.15
VIII-34	1	1.17	1.17	1.17	.
VIII-35	1	5.61	5.61	5.61	.
VIII-45	4	0.00	25.97	11.97	11.35
VIII-55	1	12.67	12.67	12.67	.

Table 5a Correlation matrix of city land expansion, developmental scale and developmental stages in 1990-2000

	p4	p4_area	tpop2k	agdp90stg	gdp90stg	lbr90stg	urb90stg	agdp2kstg	gdp2kstg	lbr2kstg	urb2kstg	chagdp	chnagrpopr	chnagrlbrr	chnagrgdpr
p4	1														
p4_area	.41**	1													
tpop2k	.69**	.09	1												
agdp90stg	.29**	.33**	.18**	1											
gdp90stg	.14*	.19**	.19**	.66**	1										
lbr90stg	.12	.18**	.14*	.63**	.81**	1									
urb90stg	.14*	.22**	.21**	.60**	.77**	.77**	1								
agdp2kstg	.34**	.39**	.21**	.75**	.52**	.47**	.42**	1							
gdp2kstg	.20**	.22**	.23**	.47**	.67**	.63**	.53**	.56**	1						
lbr2kstg	.13*	-.04	.09	-.34**	-.37**	-.34**	-.3**	-.3**	-.2**	1					
urb2kstg	.07	.28**	.06	.59**	.71**	.69**	.81**	.53**	.64**	-.37**	1				
chagdp	.36**	.32**	.14*	.66**	.32**	.26**	.23**	.79**	.33**	-.15*	.32**	1			
chnagrpopr	-.15*	.08	-.31*	-.21**	-.33**	-.38**	-.45**	.07	.02	-.03	.06	.08	1		
chnagrlbrr	-.14*	-.2**	-.18*	-.61**	-.76**	-.86**	-.84**	-.46**	-.54**	.5**	-.75**	-.25**	.38**	1	
chnagrgdpr	-.08	-.05	-.18*	-.45**	-.69**	-.64**	-.62**	-.13*	-.19**	.2**	-.32**	-.05	.7**	.62**	1

Confidence level: * .05, ** .01, *** .001

Table 5b Explanation of variables for Table 5a

definition	variables	
	year 1990	year 2000
increased city construction land	p4	
proportion of increase in city construction land with gross city administration area (1987 - 2000)	p4_area	
total urban population in 2000	tpop2k	
GDP per capital	agdp90stg	agdp2kstg
industrial structure	gdp90stg	gdp2kstg
employment structure	lbr90stg	lbr2kstg
urbanization degree	urb90stg	urb90stg
changes in GDP per capita (1990 - 2000)	chagdp	
changes in non-farm population(1990 - 2000)	chnagrpoprte	
proportion of non-farm employment (1990 - 2000)	chgnagrlbrrate	
proportion of the growth of non-agricultural GDP to total GDP	chgnagrgdprate	