Simulation of Land Use Dynamic Change using selected driving factors based on the method of Feature Selection

Yao-xing Wang [1,2,a], Ting-ting Xu[1,2,b] xuan-Lv[1,2,c]Zhi-gangZhang[3,d]

[1] Chongqing Land Resource and Housing Surveying and Planning Institute,Chongqing 400020,China;

[2] Chongqing Land Use and Remote Sensing Engineering Research Center, Chongqing 400020, Chongqing, China;

[3] School of Geographical Sciences,SouthwestUniversity,Chongqing,400715,China.

[a]cs1cs2@126.com, [b] Ting-ting Xu @163.com , [c] 84130866 @qq.com,[d]823628332@qq.com.

Keywords: Feature Selection,CHAID, Dirving Factors,Land Use Siumulaiton Chongqing.

Abstract: Method of feature selection was used in this paper to select dirving factors before models of land use change carried out simulation.Two different scenarios were set inthe study area- main districts of Chongqing。Scenario 1) carried out land use simulation involed all the factors 2）carried out land use simulation involed the factors after select by the means of feature selection. Compared to the simulation results without feature selection，the accuracy of the simulation in scenario 2,has improved 3.51%.Finally on the basis of feature selection ,thispaper carried out simulation of urban land use dynamic change in the next 10 years with the selected factors.

1. Introduction

Land use simulation models, based on a clear spatial orientation and a comprehensive integration of multi-scale,are provides an important technical means for the study of urban land use change driven mechanism, the law and trends of urban development. To simulate the change of urban land use, the main driving factors that affect the dynamic changes of urban land use are the key factors. Factor selection, spatial data acquisition、collection, processing, and series of parameters of the model are likely to affect the accuracy of the simulation. Due to the limitation of data, or other reasons,when using models simulates changes of land use , It is commonly believed that the more driving factor, the more the bette.However, 1) Models, as a help to analysis the quantitative relationship between land use change and driving factors and important tools to understand and predict patterns and process of land use[2],have limitations as a result of abstraction and simplification to the real world.Right now, there has no a comprehensive interdisciplinary model, uniting the social economy, culture, environment and so on various aspects factor, to simulating land use change on multi-scale simulation；2) Due to the presence of spatial autocorrelation, driving factors relate to each other in different scales.The more selected driving factors, driving factors are the more likely to expressed or approximate expressed by other remaining factors ,increased conversion intensity the land grid unit,and caused the land type conversion is not clear. In addition, social economy and natural conditions of study area actual also a certain influence on generality of the model; 3) Duringf the extraction of land use transformation rules, most models will take random factor into consideration，which will lead the converation probability are ultimately affected by random factors ,and increase the contingency and randomness of the simulation results;4) During process of simulation modes models gieve the same weight driving factor when , as to the key factors, without highlighting.

As a common method of dimension reduction,feature selection refers to according to a certain evaluation criteria, choose the optimal feature subset of output performance, from the original characteristics Its purpose is to make choose the optimal feature subset of the classification or regression model is built to meet and feature selection before the approximate even better prediction precision. Feature selection as a hot spot in the field of pattern recognition and artificial intelligence, has been widely used in biological, chemical engineering, financial credit information, and other fields . Towards above problems of actuation factor selection, this paper using feature selection method to extract the key factor of land use change to simulate changes of land use , in order to achieve the purpose of dimension reduction,and improvement of simulation effects.

Purpose of this article: 1) test data processing and the choice of factors how to influence the simulation and improve the simulation effect. 2) according to a new round of urban development and planning in chongqing, on the basis of urban development, predict the the spatial pattern and environmental effect of the city under different land use patterns changed.

2. Study area

Chongqing is located in the southwest of China, the upriver of Yangtze River and the southeast of Sichuan Basin, between 105°11' to 110°11' east longitude and 28°10' to 32°13' north latitude. Chongqing has an undulating terrain and various physiognomies which mainly include mountains and hills. So, Chongqing is also called “Mountain City”. Chongqing has various kind of physiognomy distinguish from other region obviously. In Chongqing, staggered distribution between urban and rural areas, which finally lead to forming a "multi-center, group-style" layout structure. The metropolitan area is about 5473 square kilometers, accounting for less than 9% of the whole city. Its altitude ranges from 141 to 1343 m. In the region， "plain", "hill", "valley" and "flat" staggered dispersedly . The permanent resident population in metropolitan area is about 7.458 million, accounting for one quarter of the whole Chongqing. Metropolitan area is the center of the political, economic and transportation in whole Chongqing. After Chongqing became a municipality， economic develops fast, income gap between urban and rural increases. Statistical data shows the GDP of metropolitan area accounted for 92% of the whole Chongqing in2008， with less than 9% of the total area of Chongqing, Which not only indicates that there is a huge wealth gap between urban and rural, also reflects that the metropolitan area is still key areas of urban development in the furture.Chongqing City Master Plan (1996-2020) puts forward that new space for urban development will be expanded, new transport system will be established to adapt mountain city’s character, and continue to controlling the population size in the old city and to promote the population growth in the new district and small towns. Urban land use patterns in the region will show different patterns in time and space as results of urban space expansions, improvements of transport facilities, demographic and other factors change in metropolitan area， etc. Unique topography and geomorphology in the study area are natural elements to determine urban form in the future, while rapid economic development, transportation and some other factors are social and economic elements to determine urban form in the future. Metropolitan area in Chongqing is an ideal zone to study the land use transformation.

3. Flowchart and methods

3.1 Data sources and data preprocessing

Dyna-CLUE model considered the role of socio-economic and biophysical driving factors for land conversion at a multi-scale and calculated competition in the relationship between various types of land use at a multi-scale.

According to the data availability and correlation, land-use data from the year of 2002, 2009 was used. Eight driving factors data were chosen, including: 1) Topographical factors such as elevation, slope, waviness and roughness, etc.2) Accessibility factors such as the distance to roads, railways, water bodies, rivers, and the center of markets, etc.3) Socio-economic factors, mainly population density .

Firstly, land use data for 2009 was reclassified to six categories, namely urban land, arable land, forest land, pasture land, water body and other land as land use types studied in the paper according to the amount and distribution characteristics of land use type in the study area refer to the National Land Classification (Trial) (2002) on the platform of ArcGIS9.3.Secondly, land-use data in the year of 2002 and 2006 was obtained after a series operation such as reclassification, unified geographic coordinates, map projection, unified scope of the study and resolution， etc.Topographical factors such as elevation, slope, waviness and roughness, were made by spatial analysis function of ArcGIS.Location factors such as distance to the road, water, the distance to the markets were made by buffer analysis.

3.2 Flowchart

This paper designed two different simulation programs: scheme 1) the above preparation of all the driving factors and do for the needs of the Dyna-CLUE model runs data regression analysis, on the land use change in the study area deployment simulation; 2) using chi square self mutual detection decision tree (CHAID) Chi-squared Automatic Interaction Detector feature selection method based on the above preparation of the driving factor for extraction, key factors of land use change in the study area were selected, the data required for Dyna-CLUE model operation, the regression analysis, the land use change in the study area deployment simulation

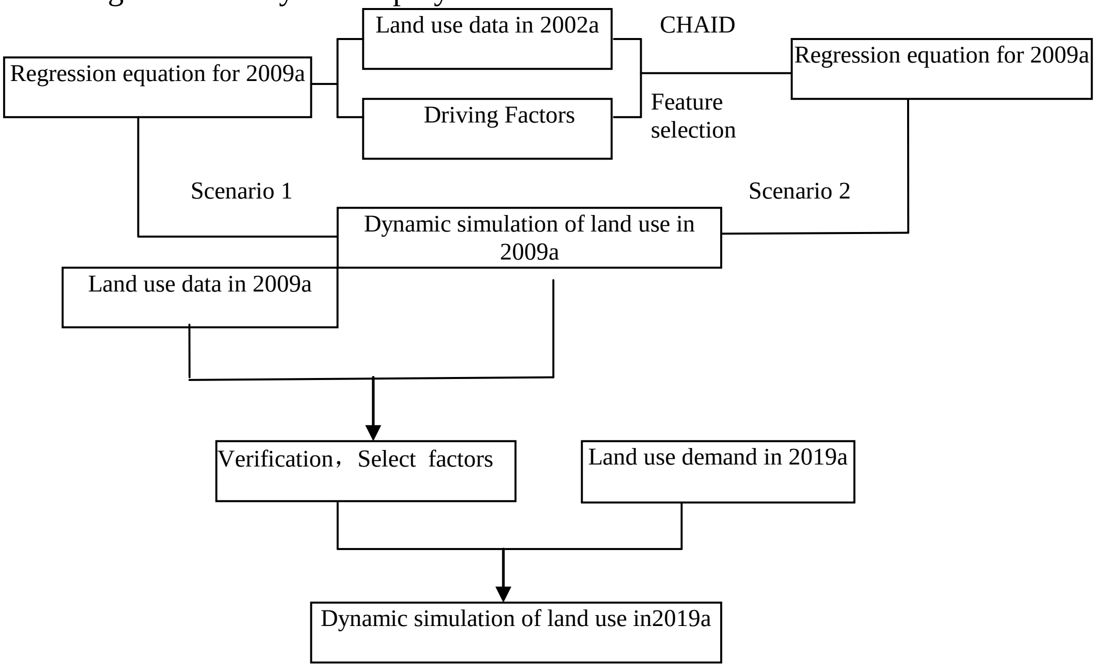

Fig.1. Flowchart of land use change in main district of Chongqing

4. Results analysis

4.1 CHAID analysis

Tab 1 AID Results of driving factors

	Urban land	Arable land	Forest land	Pasture land	Water body	Other land use
roughness of surface	0.03	0.03	0.33	-	0.37	0.02
DEM	0.01	0.01	0.01	1	0.03	-
Distance to roads	0.32	0.03	0.01	-	0.03	-
waviness of surface	0.11	0.01	0.01	-	0.03	-
Distance to rails	0.52	0.78	-	-	0.06	0.75
Population density	0.00	0.01	0.03	-	0.4	0.00
Distance to markets	0.1	0.03	0.01	-	0.06	-
Distance to mwater	0.01	0.14	0.60	-	0.09	0.22
slope	0.03	0.23	0.02	-	0.12	0.11

From results of CHAID analysis ,not all the factors forced land use changed.For example, the distance form railways on woodland, elevation, the distance from roads, and amplitude of landforms for other land use. Whilde elevation is the right factor that forced garden plots changed. Driving ability of different driving factors to the same land use were different ,such as the distance from water to the wood land use change was 0.60, but the was 0.01.

Two scenarios were set up in this paper: scenarios 1) Factors made up above were all involved in simulation scenarios;2)On the basis of CHAID screening and several attempts, the distance to roads, rail, market, water ,population density were selected to simulate land use change.

4.2 Simulation results analysis

Simulation results show that the number of correct simulation grids were 558417 and Kappa index was 79.22% when factors are involved, while the number of correct simulation grids were 558417 and Kappa index was 79.22% in Scenario 2. Among them, the simulation accuracy for forest increased from the original 83.26% to 97.79%, increased by 14.53%. The improvement in water body is not obvious, only increased by 0.02%. The overall simulation accuracy was improved from 82.68% to 89.30%, increased by 6.62%.

The essence of feature selection is to remove part of the driving factors and kepp key factors that driving land use change which will inevitably exclude some valid information. Therefore accuracy of a certain land use maybe reduced, but the overall simulation accuracy is improved just like in scenario 2.

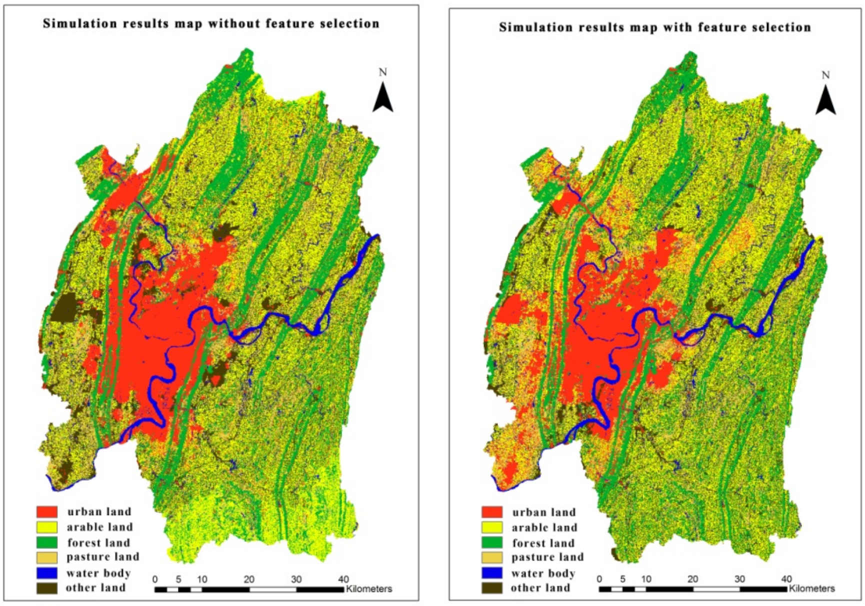

Fig 2 Difference of simulation results in the spatial distribution

（A_1 is simulation results without feature selection,

A_2 is the simulation results with feature selection ）

Tab2 Comparison of accuracy simulation results

Land use type	Amount of 2009 actual land grids	Correctly simulated amount of land grids without feature selection	Correctly simulated amount of land grids with feature selection l	accuracy of simulation without feature selection	accuracy of simulation with feature selection
Urban land	95202	56969	66177	59.84%	69.51%
Arable	166892	134259	134224	80.45%	80.43%
Forest	112074	92943	92941	82.93%	82.93%
Pasture	49055	36761	36726	74.94%	74.87%
Water bodies	22090	17084	17192	77.34%	77.83%
Other land use	96626	42417	48740	43.90%	50.44%
Total	541939	380433	396000	70.20%	73.07%

4.3 Land use change simulation in 2019

The factors after feature selection simulated land use change for 2009 in Chongqing Metropolitan Area well. Parameters obtained by using existing data were used to simulate land use change in future was reasonable. Selected factors were used to simulated and predict the land use pattern in the study area the next 10 years holding the hypothesis that to the study area maintain the current development.

According to the simulation results, Urban land expand rapidly. Arable,Pasture land that surroud Urban land are replaced by Urban land. The expansion trend of urban land is the most obvious to northward and westward, while southward and eastward of the area were followed.

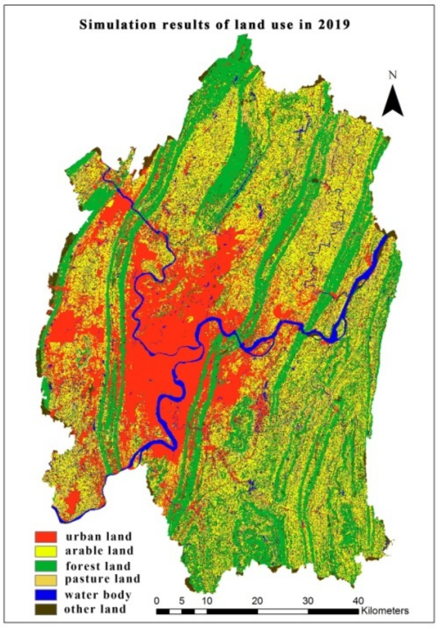

Fig.3. Simulation of land-use of main district of Chongqing city in 2019

Land use conversion matrix was established by overlay analysis of land use data in 2009 and 2019(table 6). As table 6 showed the number of arable land, garden and pasture land use and other land use significantly reduced in this period, and the number of urban land and forest land increased, while the number of water body unchanged. Source of construction land is the garden and pasture land and the other land, followed by woodland.

Tab3. Transition matrix of land use between 2009a and 2019a (unit: hectares)

	Urban land	Arable land	Forest	Pasture	Water bodies	Other land use	Total account in 2019a	Total account of Turn in in 2019a
Urban land	95202	—	—	—	906	90846	186954	91752
Arable land	—	166892	—	—	—	1263	168155	1263
Forest	—	—	112074	—	12	1450	113536	1462
Pasture	—	—	—	49055	—	389	49444	389
Water bodies	—	—	—	—	21172	—	21172	—
Other land use	—	—	—	—	—	2678	2678	—
Total account in 2009a	95202	166892	112074	49055	22090	96626	541939	—
Total account of Turn out in 2009a	—	—	—	—	918	93948	—	94866

5. Summary

Feature selection method of CHAID was used in this paper to optimal driving factors before simulation of land use change.Simulation results showed that Simulation accuracy of individual land use was lower, but overall simulation accuracy was improved 3.51%.The model used in the simulation,data selection and processing and other operation carried out during the simulaiton had a certain impact on the simulaiton result].Given the limitations of driving factors and selected simulation area, the next step is to prepare more driving factors and to expand simulation in multiple topography different area, in order to confirm the validity of the method.

Acknowledgment

This work was financially supported by Image Recognition Based High Resolution Remote Sensing Data Analysis Research (Grant NO. CQGT-KJ-2014032).

References

[1] Veldkamp A, Lambin E F. Predicting land-use change[J]. Agriculture, ecosystems & environment, 2001, 85(1): 1-6.

[2] Zhou Rui,Su Hailong,Hu Yuanman,Li Yuehui, Liu Miao.Scenarios simulation of town land use change under different spatial constraints.[J]TRANSACTIONS OF THE CHINESE SOCIETY OF AGRICULTURAL ENGINEERING:2011, 27(3)300-3008.

[3] JI Zhi-wei ， HU Min ， YIN Jian-xin 。 A survey of feature selection algorithm[J]. ELECTRONIC DESIGN ENGINEERING,2011, 19(9):46-51.

[4] Guyon I,Elisseeff A.An introduction to variable and feature selection[J].journal of Machine Learning Research,2003,3:1157-1182.

[5] Yvan Saeys, Inaki Inza and Pedro Larranaga. A review of feature selection techniques in bioinformatics. Bioinformatics, 23, 2507-2517.

[6] YIN Jian-qin,TIAN Guo-hui,WEI Jun, LI Jin-ping .Letc.Research on the Relationship of the Support and the Discriminative Ability for Classification of Features.Acta Electronica Sinica 2015, 43(2):248-254.

[7] Shen Sheng , Yang Honghui,Wang Yun,Pan Yue,etc.Joint Mutual Information Feature Selection for Underwater Acoustic Targets.Journal of Northwestern Polytechnical University[J].2015, 33(4):639-643

[8] Verburg P H, Overmars K P. Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model[J]. Landscape ecology, 2009, 24(9):1167-1181.

[9] Lima M L, Zelaya K, Massone H. Groundwater vulnerability assessment combining the drastic and Dyna-CLUE Model in the Argentine pampas[J]. Environmental management, 2011, 47(5):828-839.

[10] Shrestha R P. Monitoring farmland loss and projecting the future land use of an urbanized watershed in Yogyakarta, Indonesia[J]. Journal of Land Use Science, 2013, 8(1):59-84.