GeoTree Manual

Version 1.0

http://www.sssampling.cn/GeoTree/

201 Spatial Analysis Group[1]

[1] State Key Laboratory of Resources and Environmental Information System

Institute of Geographical Sciences and Natural Resources Research

Chinese Academy of Sciences, Beijing 100101, China

中国科学院地理科学与资源研究所
Institute of Geographic Sciences and
Natural Resources Research, CAS

May 2023

Legal statement

Please cite the reference when using this software:

1. Wang JF, Liu XH, Peng L, Chen HY, Driskell L, Zheng XY. 2012. Cities evolution tree and applications to predicting urban growth. **Population and Environment** 33(2-3): 186-201.
2. Wang JF, Zhang TL, Fu BJ. 2016. A measure of spatial stratified heterogeneity. **Ecological Indicators** 67: 250-256.

Content

1 Welcome to GeoTree

1.1 Introduction

GeoTree is a computer visualization exploratory analysis tool developed by **201 Spatial Analysis Group** of Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. It uses a "tree" structure to explore the spatiotemporal classification and evolution of phenomena. For spatiotemporal processes with evolution mechanisms, cross-sectional data can be used to reconstruct their classification and temporal evolution processes. The hidden mechanisms and evolutionary changes in multidimensional data are presented in a clear "tree" structure, which can be used for predicting and forecasting regional evolution, and for further understanding the relevant environmental and socio-economic phenomena and their changing trends.

1.2 Function

It mainly includes five modules: file, model, analysis, visualization, and help.

1.2.1 File

The file module includes the functions of importing, exporting, and editing files.

The file types used in the system mainly include two parts, cluster data and map data. Clustering data allows data in.csv,.xls and xlsx formats. Clustering analysis will cluster according to these data and build a GeoTree with hierarchical structure. The map data is mainly used to display the SHP data of the research unit for interactive query with the tree model.

1.2.1 Model

The Model section includes multi-level model analysis and Geodetector q statistical analysis.

Multi-level model analysis: The system provides the functionality of multi-level model analysis for two-level branch clustering data. Based on R statistical software, it performs statistical analysis on the first-level branch (representing classification) and secondary branch (representing evolutionary stages) and displays the statistical results.

Geodetector q statistical analysis analyzes the explanatory power of a given factor X on the spatial differentiation of research object Y.

$$q = \left(1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N\sigma^2}\right) \times 100\%$$

where h = 1, ..., L represents the strata of variable Y or factor X, such as classification, zoning, grouping, or some other type of partitioning. N_h and N are the number of units in stratum h and the total population, respectively. σ_h^2 and σ^2 are the variances of Y in stratum h and the total population, respectively. The value of q ranges from 0 to 1. For more information on the Geodetector q statistical analysis, please visit www.geodetector.cn.

1.2.3 Analysis

The analysis module consists of two sub-modules: primary branch clustering and secondary branch clustering. Multiple attribute fields imported into the system are used for clustering, and the clustering results can be used as the basis for building the Geotree.

The clustering methods included in the system are Simple K-means, EM, Filtered, Hierarchical, Farthest First and Wards. In addition, the secondary branch clustering module provides a method of clustering based on the level of social development of the study unit.

1.2.4 Visualization

The Visualization module mainly includes two sub-modules: Evolutionary tree construction and Markov Chain. This module provides the construction of a tree model for the results of clustering analysis, showing the evolutionary patterns of research objects in a hierarchical tree form, and supports linkage display with geographic data.

The Markov Chain is used to analyze the transition between different states.

1.2.5 Help

Contains the manual, cases, and software copyright information.

1.3 Installing GeoTree 1.4.1 in Windows

To download GeoTree V1.41, please visit http://www.sssampling.cn/GeoTree.

Unzip this file into the place where you want to install it. This can be anywhere, for example, your Program Files directory.

1.3.1 Hardware

Table 1.1 Hardware configuration

Category	Configuration
Hardware Configuration	CPU: dual-core, Memory: 4GB or higher

1.3.2 Software

Table 1.2 Software configuration

Category	Configuration
Operating System and Version	Windows7 and above
Other Software and Version	JDK 1.8 (32-bit) with environment variables configured, AccessDatabaseEngine

2 System usage instructions

2.1 Functional framework

Double-click the desktop icon to enter the main page of the system, as shown in Figure 2.1.

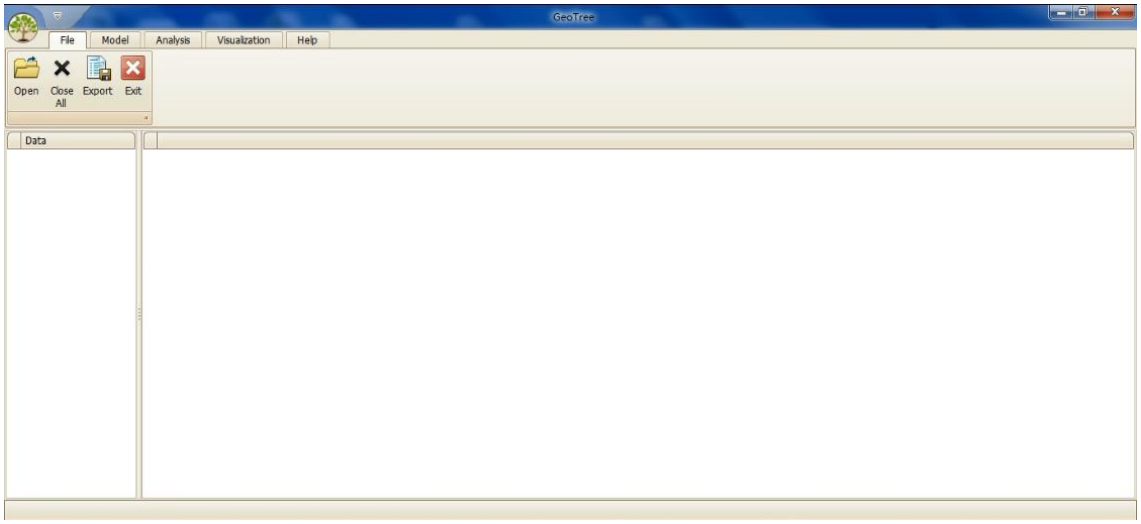


Figure 2.1 Main interface

The main interface of system is the Ribbon style, which includes the following parts:

Menu bar: contains five modules including file, model, analysis, visualization, help.

Data list: Displays a list of open data files.

View window: Display the selected data information in the data list.

2.2 File

The module includes open data, export data, close data, exit. The main function is to import the data to be analyzed into the system. The system supports simple editing of data (the edited data will not be saved directly to the source data file) and can use the function of exporting data to save the data.

2.2.1 Open data

2.2.1.1 Function description

Open data in.csv,.xls and xlsx formats.

2.2.1.2 Operation

Click "Open", and the dialog box for selecting data will pop up as shown in Figure 2.2. Select the data to be opened and click "Open" to load the selected data file. After data loading is complete, the "Data List" window displays the list of data files.

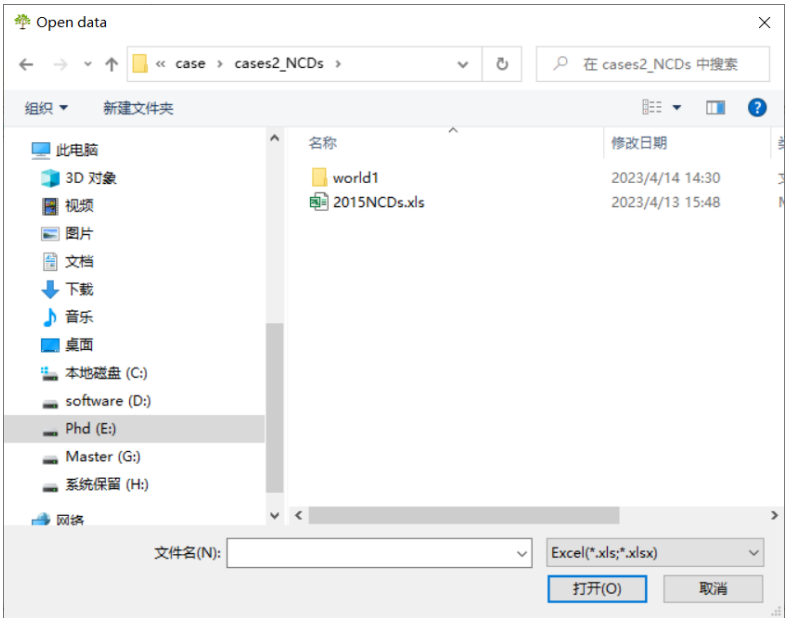
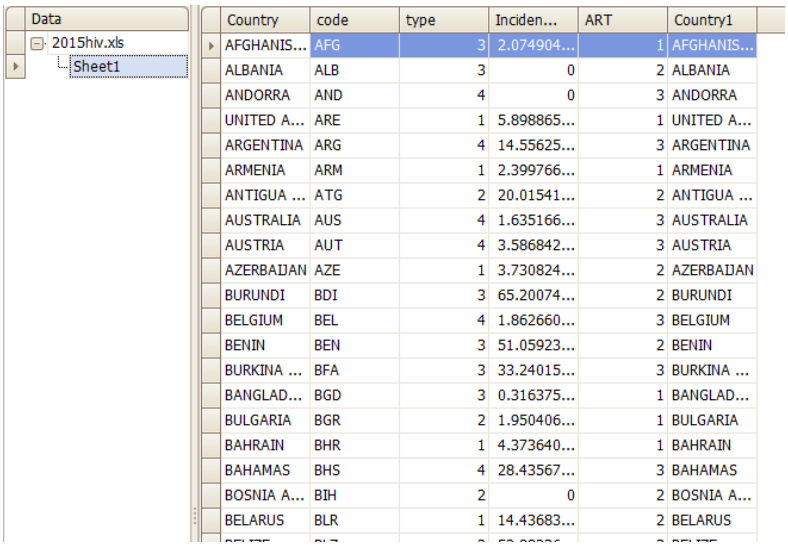


Figure 2.2 Open data



Country	code	type	Inciden...	ART	Country1
AFGHANIS...	AFG	3	2.074904...	1	AFGHANIS...
ALBANIA	ALB	3	0	2	ALBANIA
ANDORRA	AND	4	0	3	ANDORRA
UNITED A...	ARE	1	5.898865...	1	UNITED A...
ARGENTINA	ARG	4	14.55625...	3	ARGENTINA
ARMENIA	ARM	1	2.399766...	1	ARMENIA
ANTIGUA ...	ATG	2	20.01541...	2	ANTIGUA ...
AUSTRALIA	AUS	4	1.635166...	3	AUSTRALIA
AUSTRIA	AUT	4	3.586842...	3	AUSTRIA
AZERBAIJAN	AZE	1	3.730824...	2	AZERBAIJAN
BURUNDI	BDI	3	65.20074...	2	BURUNDI
BELGIUM	BEL	4	1.862660...	3	BELGIUM
BENIN	BEN	3	51.05923...	2	BENIN
BURKINA ...	BFA	3	33.24015...	3	BURKINA ...
BANGLAD...	BGD	3	0.316375...	1	BANGLAD...
BULGARIA	BGR	2	1.950406...	1	BULGARIA
BAHRAIN	BHR	1	4.373640...	1	BAHRAIN
BAHAMAS	BHS	4	28.43567...	3	BAHAMAS
BOSNIA A...	BIH	2	0	2	BOSNIA A...
BELARUS	BLR	1	14.43683...	2	BELARUS

Figure 2.3 Browse data

Data can be simply modified in the view window, and the modified data can be exported using the function of "Export Data" after modification.

Select the data in the data list and right-click "Remove" to remove the selected data.

2.2.2 Export data

2.2.2.1 Function description

Export data in csv, xls and xlsx formats.

2.2.2.2 Operation

Click "Export", and the export data dialog box pops up, as shown in Figure 2.4. The data currently open in the system is displayed in the list. Select the data to be exported, click "Save" to set the output path and output format of the file, and click "Export" to output the selected data to the data file in the specified format. Click "Cancel" to cancel the data export and close the form.

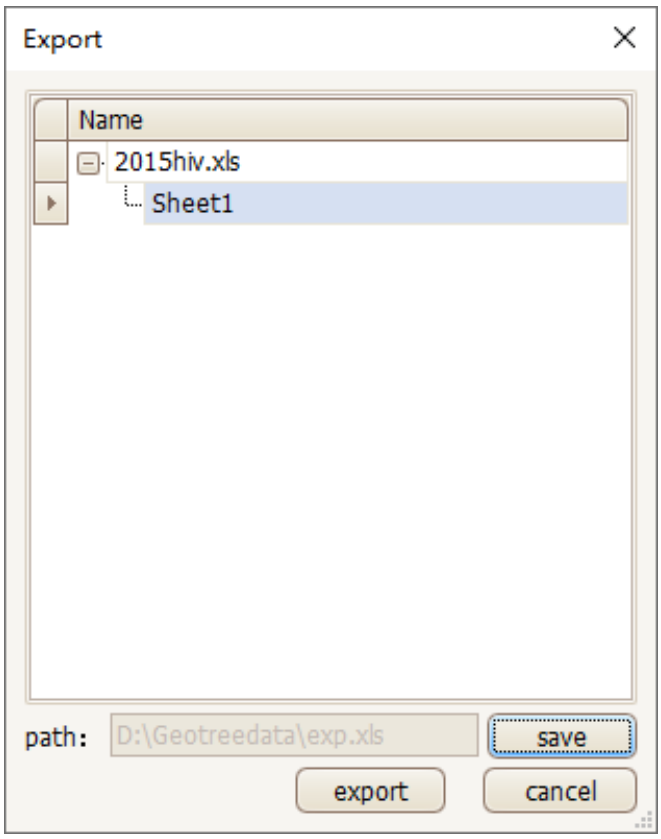


Figure 2.4 Export data

2.2.3 Close data

2.2.3.1 Function description

Close the data file and restore the form to its initial state.

2.2.3.2 Operation

Click the Close to close all the opened data and restore the form to the initial state.

2.2.4 Exit

2.2.4.1 Function description

Exit: shuts down the system.

2.2.4.2 Operation

Click "Exit" to close the system. Confirm with the user before disabling.

2.3 Model

This module includes geodetector q statistical analysis and multi-level analysis.

2.3.1 Statistical analysis of Geodetector q

2.3.1.1 Function description

The analysis results were obtained through geodetector q statistical analysis of the data.

2.3.1.2 Operation

Click "q-statistic" and the q statistical analysis dialog box will pop up, as shown in Figure 2.5.

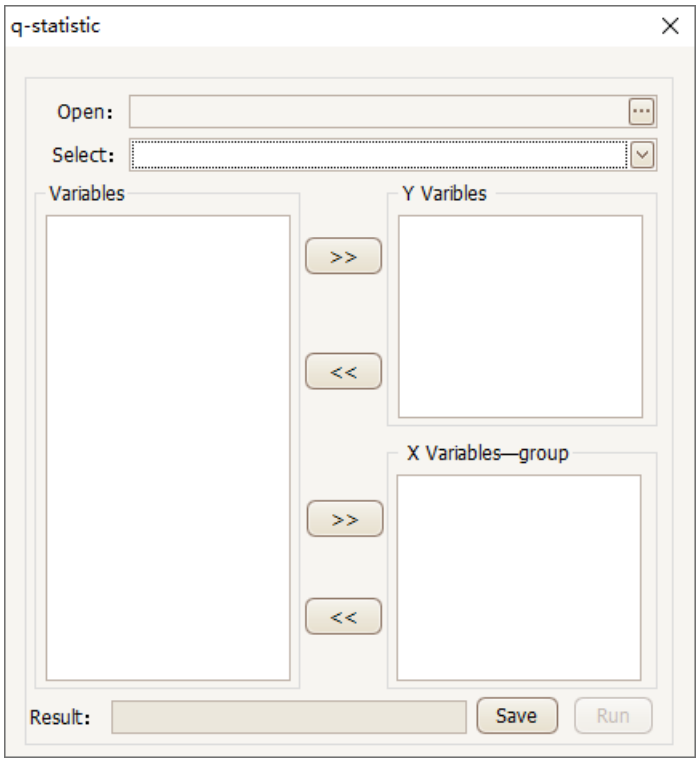


Figure 2.5 Statistical analysis of geodetector q

The analysis data of this function can be imported directly from the external data and selected from the current open data of the system.

Open external data: Click the "Open" ··· and the open data dialog box will pop up. The system supports opening .xls, ".xlsx " and ".csv ",files.

Open the imported data: If the system has used the Open Data function to import data, you can directly select the target data from the Select Data drop-down list.

Parameter setting: After data is selected, all attribute fields in the analysis variable list are displayed. Select the target field and click >> to add the field to the list of dependent variable or group variable. Click << to remove the selected field from the list of dependent variable or group variable.

Set the analysis result path: click Save to set the analysis result output path. The file format of result is.txt.

Data analysis: Click Run to conduct the geodetector q statistical analysis. After the analysis is completed, the system will pop up a message indicating the

completion of analysis and save the analysis results of geodetector q to the specified path.

2.3.2 Multi-level model analysis

2.3.2.1 Function description

Multi-level model analysis was conducted on the data.

2.3.2.2 Operation

Click "Multi-level simulation Analysis", and the multi-level simulation dialog box will pop up, as shown in Figure 2.6

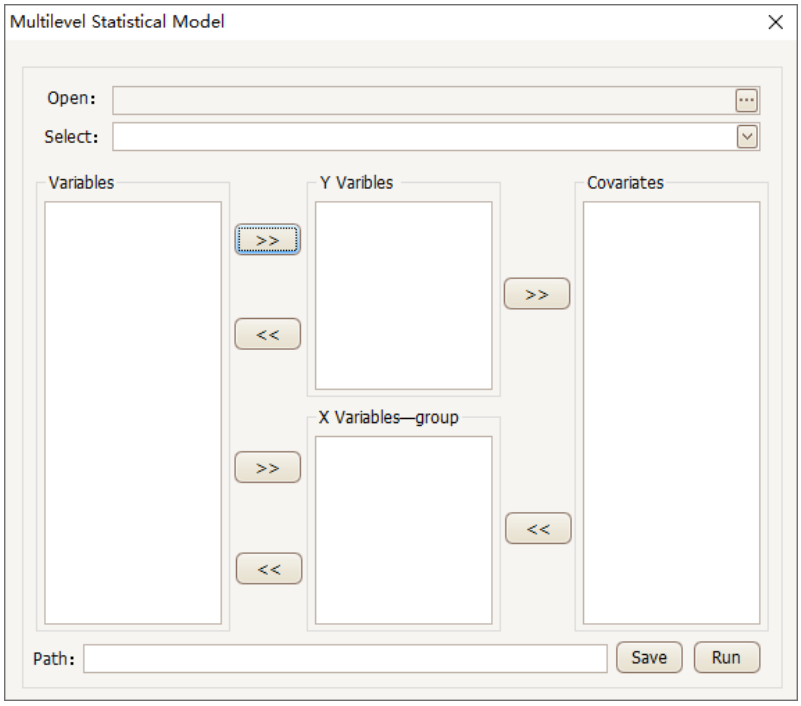


Figure 2.6 Multi-level simulation analysis

The analysis data of this function can be imported directly from the external data and selected from the current open data of the system.

Open external data: Click the "Open" and the open data dialog box will pop up. The system supports opening ".xls ", ".xlsx ", ".csv ", and ".mdb "files.

Open the imported data: If the system has used the Open Data function to import data, you can directly select the target data from the Select Data drop-down list.

Parameter setting: After data is selected, all attribute fields in the variable list are

displayed. Select the target field and click >> to add the field to the list of dependent variables, covariable and group variable. Click << to remove the selected field from the corresponding list of variables.

Set the analysis result path: click Save to set the analysis result output path. The file format is.txt.

Data analysis: Click Run to conduct multi-level simulation statistical analysis. After the analysis is completed, the system will pop up a message indicating the completion of analysis and save the analysis results to the specified path.

2.4 Cluster analysis and processing

This functional module includes primary branch clustering and secondary branch clustering.

2.4.1 The primary branches

2.4.1.1 Function description

According to the needs of users, cluster analysis is carried out to obtain the primary branches.

2.4.1.2 Operation

Click the function of "The Primary Branches" in the module of "Analysis", and the following dialog box pops up. Users can perform the primary clustering analysis on data. As shown in Figure 2.7.

The analysis data of this function can be imported directly from the external data and selected from the current open data of the system.

Open external data: Click the "Open" ... and the open data dialog box will pop up. The system supports opening ".xls" files.

Open the imported data: If the system has used the Open Data function to import data, you can directly select the target data from the Select Data drop-down list.

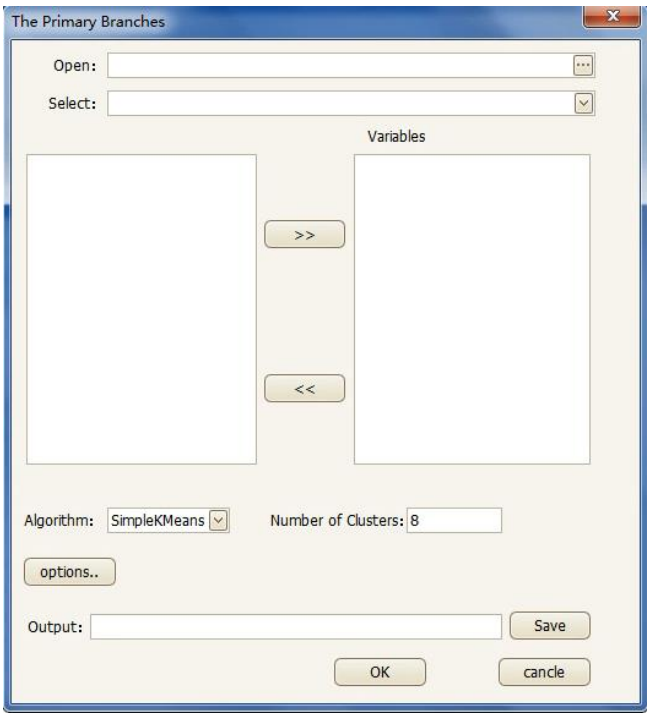


Figure 2.7 Clustering of first-order branches

Clustering variable selection: After data is selected, all attribute fields in the variable list are displayed.

Select the target field and click >> to add the field to the clustering list. Click << to remove the selected field from the clustering list.

Parameter setting: Set clustering list, clusterer and other parameters.

Set the analysis result path: click Save to set the saving path for the clustering result, and the saving file format is.xls.

After the clustering is complete, the system will automatically load the clustering results, and users can browse the analysis results directly.

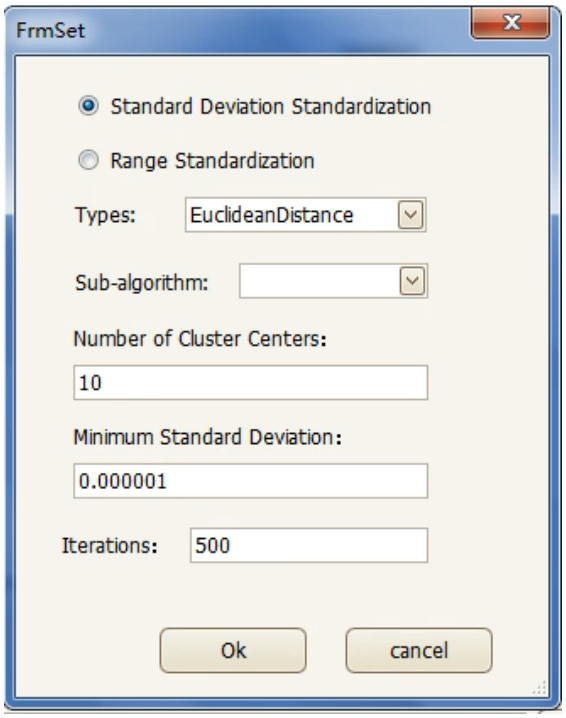


Figure 2.8 Parameter setting of first-level branch clustering.

2.4.2 Secondary branch clustering

2.4.2.1 Function description

The data were used for secondary branch clustering analysis.

2.4.2.2 Operation

Click "The Secondary Branches", the secondary branch clustering analysis dialog box will pop up, as shown in Figure 2.9. The secondary clustering includes two kinds of clustering methods. The first is to use the primary clustering method for clustering again, as shown in Figure 2.9, and the second is to use the socio-economic development data of the study unit for clustering, as shown in Figure 2.10.

Clustering method 1: The operation refers to the primary branch clustering.

Clustering Method 2: The analysis data of this function can be imported directly from the external data and selected from the current open data of the system (Figure 2.10).

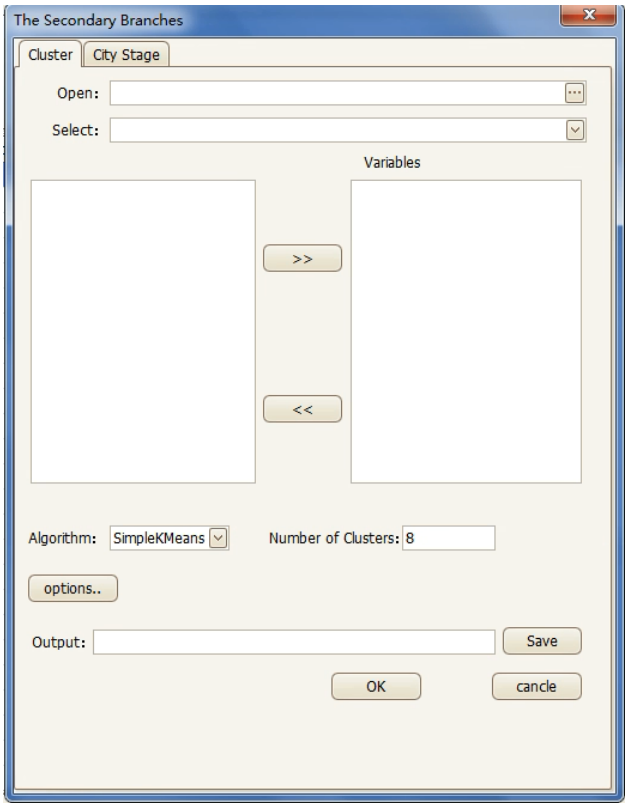


Figure 2.9 Secondary branch clustering

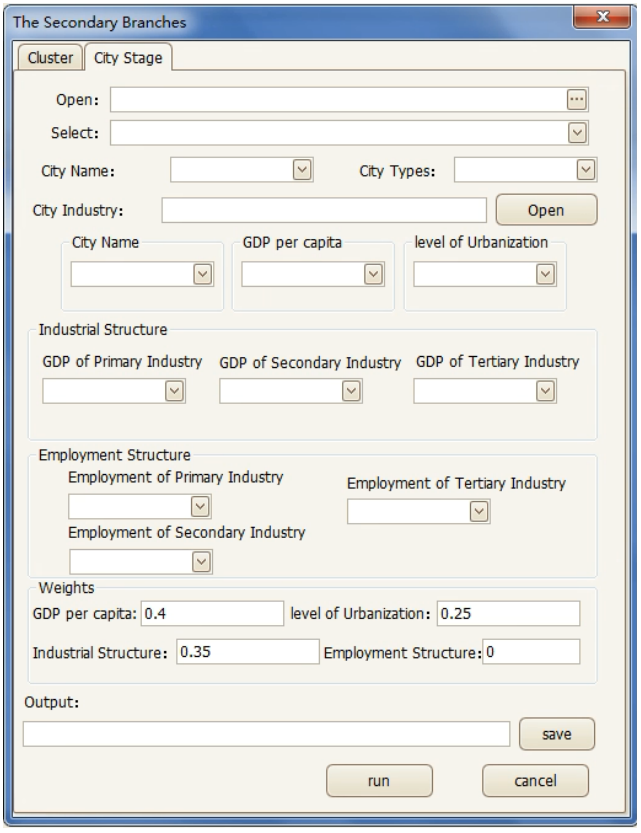


Figure 2.10 Secondary branch clustering using city industrial data.

Open the socio-economic development data: click Open, select the file in.xls

format, and import the city industry data.

Parameter setting: select the corresponding field according to the name of each data box. The name of the study unit should be consistent with the name of the study unit in the primary branches. Finally, set the weights in context.

Set the analysis result path: Click 

to set the analysis result saving path and file name. The file format is .xls.

Secondary branch clustering analysis: click 

to run the secondary cluster analysis. After the clustering is complete, the system will automatically load the clustering results, and users can browse the analysis results directly.

2.5 Visualization

2.5.1 Generating GeoTree

2.5.1.1 Function description

The user can import data or select data already imported into the system to generate the GeoTree.

2.5.1.2 Operation

Click "Geotree" 

to display the Geotree import interface, as shown in Figure 2.11.

Import data to generate Geotree: Click "Import data generation"  , users can import external clustering data for generating Geotree.

Clustering result generate Geotree: The system can directly use the result data of its own clustering analysis to generate Geotree.

Select data: Choose the data to generate Geotree from a list of opened data.

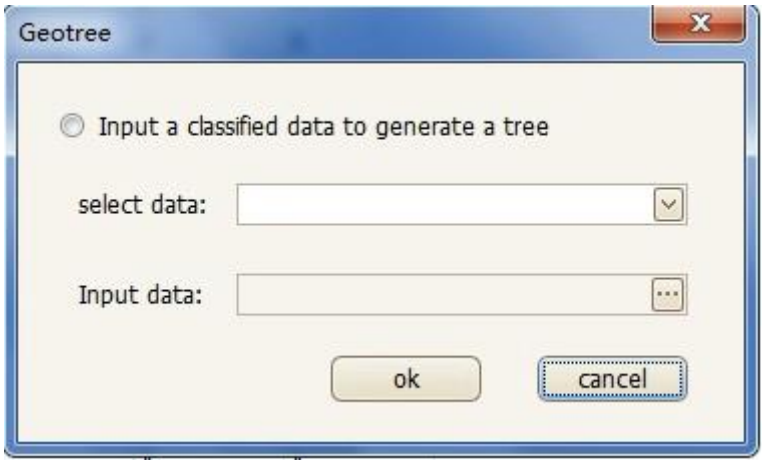


Figure 2.11 Loading data

Click "OK", and the parameter setting interface pops up, as shown in Figure 2.12.

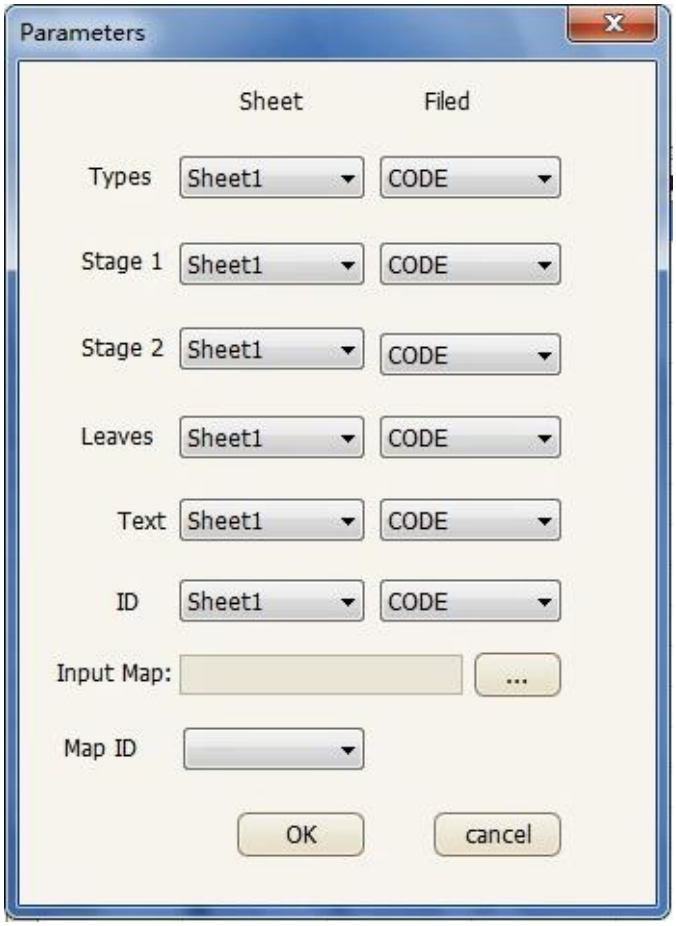
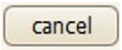


Figure 2.12 Parameter Settings

Parameter Settings: On the left side, the open data table is displayed (currently only single table is supported), and on the right side are the corresponding data fields. Users select fields to set Types (primary branch type), Stage 1 (secondary branch type), Stage 2 (secondary branch type at different time points, can be empty), Leaves (leaf

color), and Text (text content on leaves). Map data can be imported according to user needs. If map data is imported, the data for generating Geotree and the map data must be associated with the field "ID", and the data in the ID field cannot be duplicated. Note that field names in the data cannot contain spaces.

Click "OK" 

to generate the Geotree and click "Cancel" 

to cancel the the tree generation. Figure 4.5.3 shows the generated. Figure 2.13 shows the generated Geotree and map. and map.

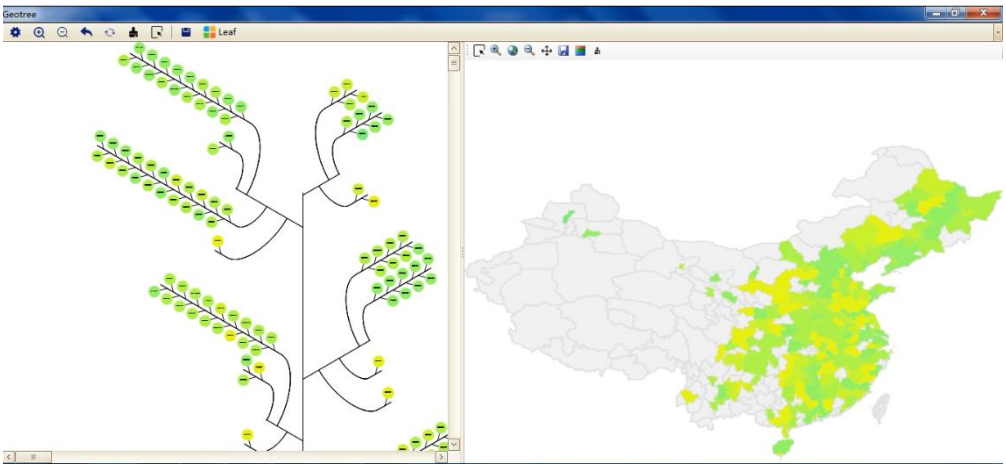


Figure 2.13 Geotree and map

On the left is the generated Geotree, and on the right is the map data associated with the tree through ID. The color of the research unit in the map is consistent with the color of the corresponding leaf node in the evolution tree, and when the color of the leaf node changes, the color of the corresponding unit in the map will also be adjusted accordingly. Selecting a leaf node will highlight and center the corresponding unit in the map. When using the selection tool to select a unit on the map, the corresponding leaf node will also be selected.

Geotree module tool: Tree style setting, tree browsing, image export, reclassification.

Tree style setting: Click  and the Tree style setting form pops up, as shown in Figure 2.14.

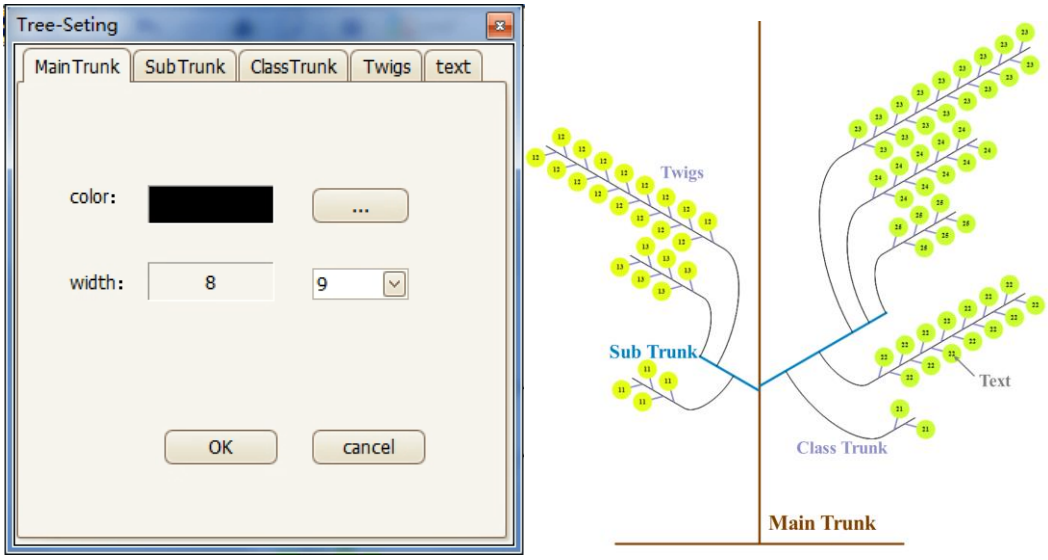


Figure 2.14 Tree style Settings

Main Trunk: sets the color and thickness of the main trunk.

Sub Trunk: sets the color and thickness of sub trunk.

Class Trunk: sets the class trunk of stem color and thickness.

Twigs: sets the color and thickness of the leaf stem.

text: sets the font color. As shown in Figure 2.15, the interface contains two radio options to display numbers and text, three color setting buttons, and one text size setting button. When selecting 'text', the unit's name is displayed on the leaf. When selecting the numerical option, the unit's corresponding type and development stage are displayed on the leaf. The color and size of the text on the leaf can be set below.

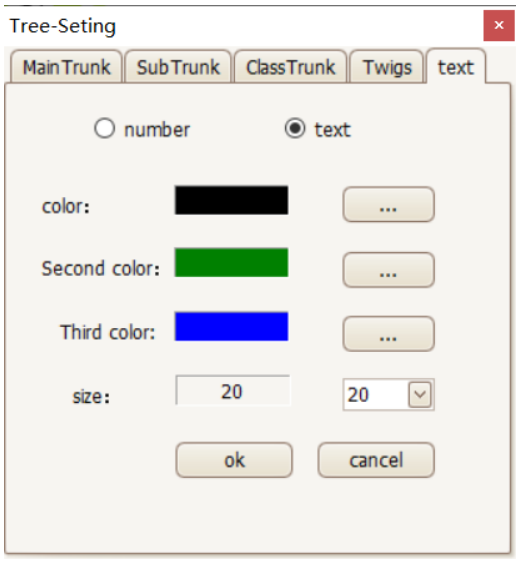


Figure 2.15 Leaf text style Settings

Click "OK"  to apply the settings. The system will redraw the tree and the associated map according to the selected style.

Zoom in: click  , to zoom in on the tree to a certain scale.

Zoom out: click  , to zoom out the tree to a certain scale.

Export image: click  , to open the "Save Image" dialog box, set the file path and name to save the generated tree as an image in the corresponding path."

Leaf rendering Settings: Click  , and the leaf classification rendering Settings dialog box pops up, as shown in Figure 2.16.

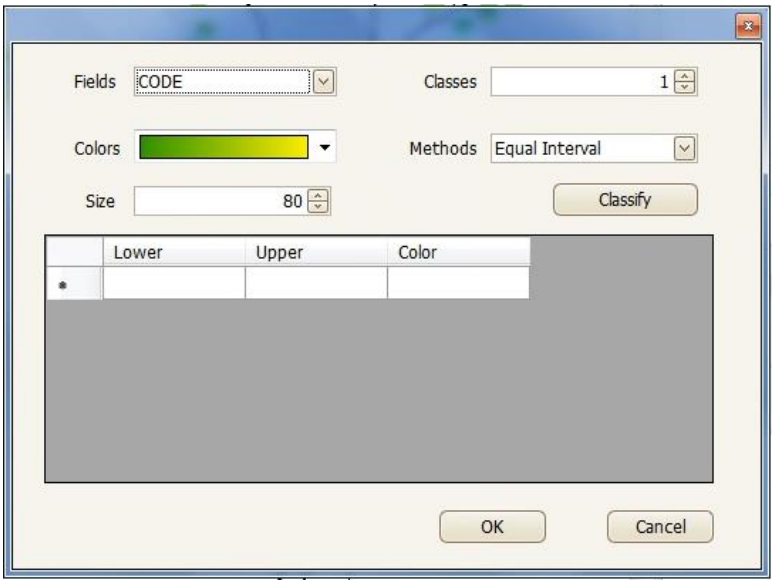


Figure 2.16 Tree leaf classification rendering

Fields: Select a numeric field.

Classes: Set the number of classes.

Colors: Select category render color band.

Methods: Set the classification method.

Size: Sets the leaf size.

Click "Classify"  and obtain the classification result, as shown in Figure 2.17. You can modify the maximum, minimum, and classification color separately for each classification. After Remark is set, the legend will be displayed according to Remark. Finally, click "OK"  to apply the style settings.

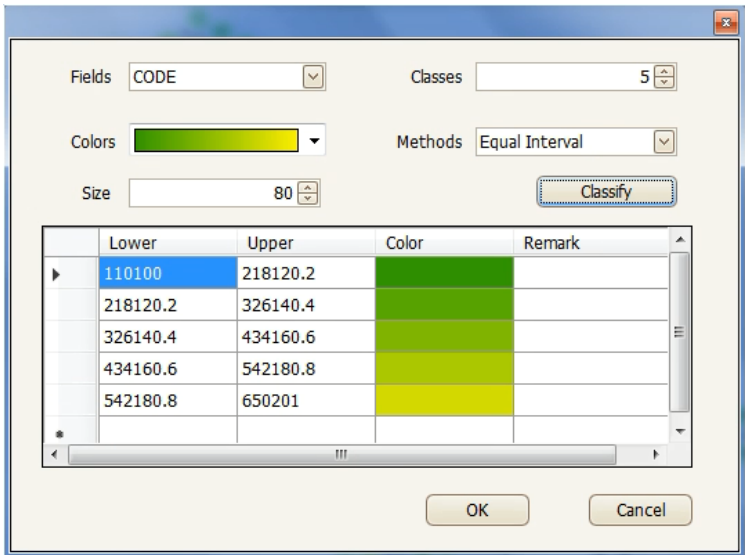


Figure 2.17 Tree classification rendering

2.5.2 Maps

2.5.2.1 Function description

Supports zooming in and out of maps, linkage between maps and trees, clicking to query, and exporting maps as image formats.

2.5.2.2 Operation

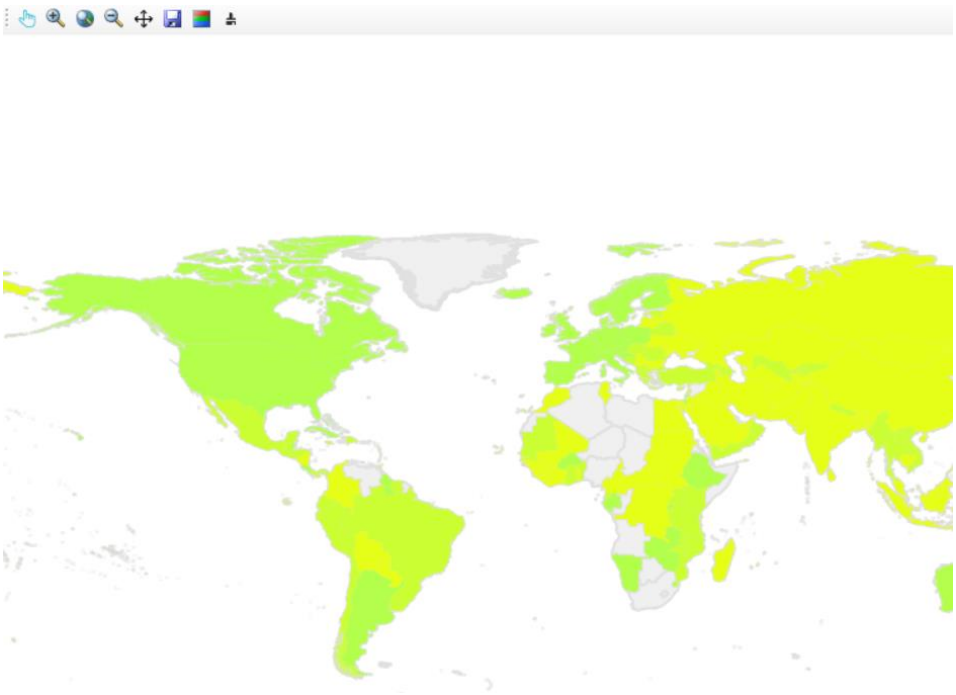
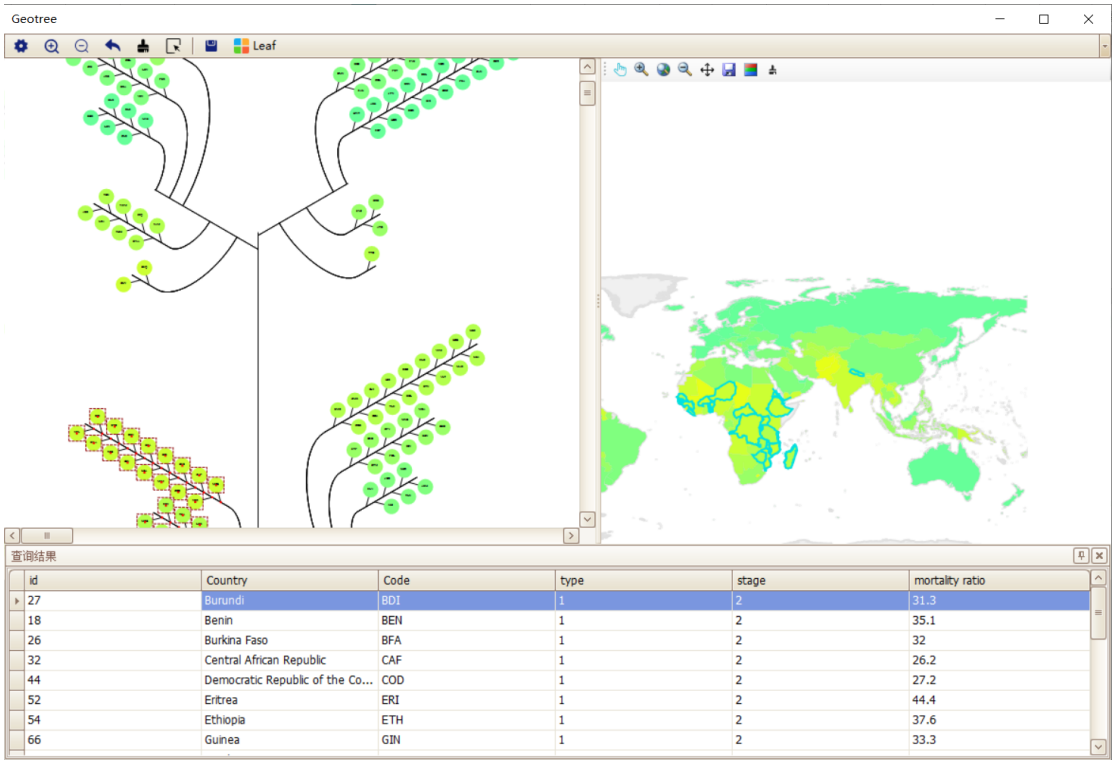


Figure 2.18 Map

Map Query: Click , and then select a research unit you want to query. The selected unit will be highlighted, and the corresponding leaf node in the linked Geotree on the left will also be selected. A data panel for querying unit will appear at the bottom of the interface. See Figure 2.19.



id	Country	Code	type	stage	mortality ratio
27	Burundi	BDI	1	2	31.3
18	Benin	BEN	1	2	35.1
26	Burkina Faso	BFA	1	2	32
32	Central African Republic	CAF	1	2	26.2
44	Democratic Republic of the Co...	COD	1	2	27.2
52	Eritrea	ERI	1	2	44.4
54	Ethiopia	ETH	1	2	37.6
66	Guinea	GIN	1	2	33.3

Figure 2.19 Map Data Query

Map Zoom In: Click to zoom in on the map.

Full Map: Click to show the full map.

Map Zoom Out: Click to zoom out of the map.

Data Roaming: Click to pan the map.

Export Image: Click to save the map as an image.

Export Map with Legend: Click 地图图例 to pop up the map-making dialog box, as shown in Figure 2.20. Click on the image to select the position where the legend is to be placed, and the system will automatically generate a legend at the mouse-clicking position. Click "save" to save the map, pop up the map save dialog box, select the save path and name, and save the map with the legend as an image.

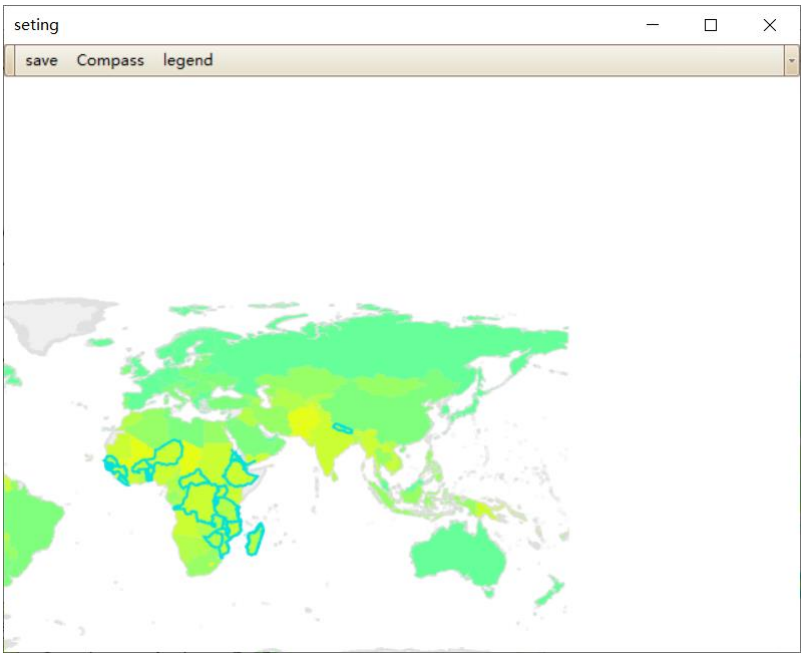


Figure 2.20 Map-making

Compass: Click on "Compass" to open the compass style selection box. Select the targeted compass and click on "OK". Then, click on the targeted location on the image to add the compass.

Legend: Add a legend. Click to add a legend to the map, as shown in Figure 2.21.

Save: Save the map.

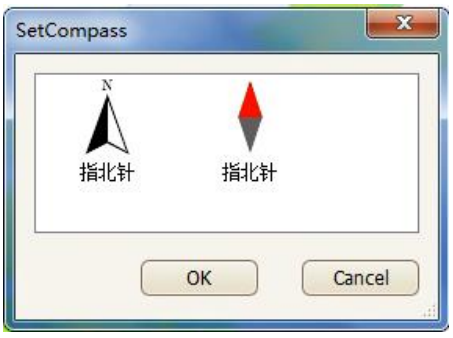


Figure 2.21 Select Compass

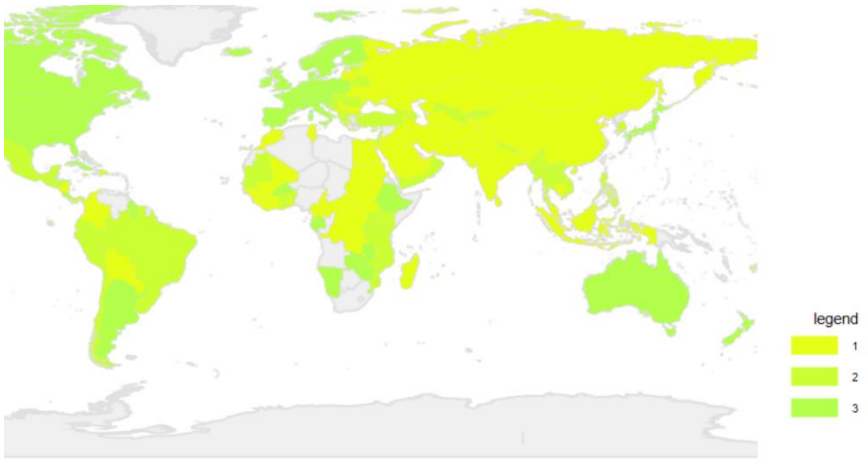


Figure 2.22 Add Legend

2.5.1 Markov chain

2.5.1.1 Function description

Based on the developmental stage changes of the research units at two different time points, use Markov chain analysis to analyze the transition patterns and probability matrix between each developmental stage.

2.5.1.2 Operation

First, import the data to be analyzed, click on "Markov Chain", and the Markov Chain settings interface will pop up, as shown in Figure 2.23.

Set parameters: In the Table column, select the data table to be analyzed (single table only); in the Field column, set Branch (primary branch type), Twig-1 (secondary branch type at T_1), Twig-2 (secondary branch type at T_2), and Statistics (variables of interest). Click "OK"  to run the Markov Chain.

The first page of the result shows a diagram of the number of research units that underwent transitions between secondary branch types and the mean value of the variable of interest, as shown in Figure 2.24. Click  to remove color settings.

The second page shows the probability transition matrix between secondary branch types, as shown in Figure 2.25.

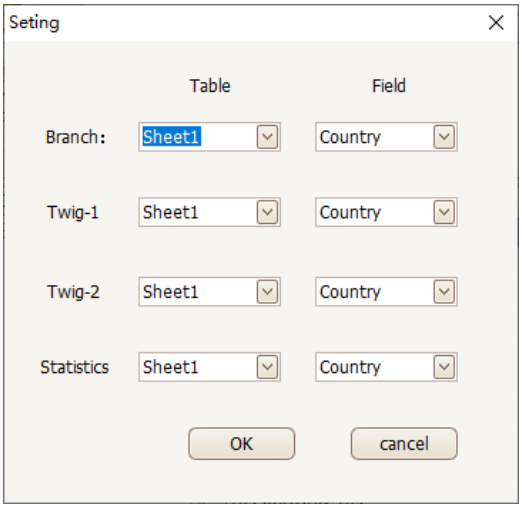


Figure 2.23 Markov Chain Parameter Settings

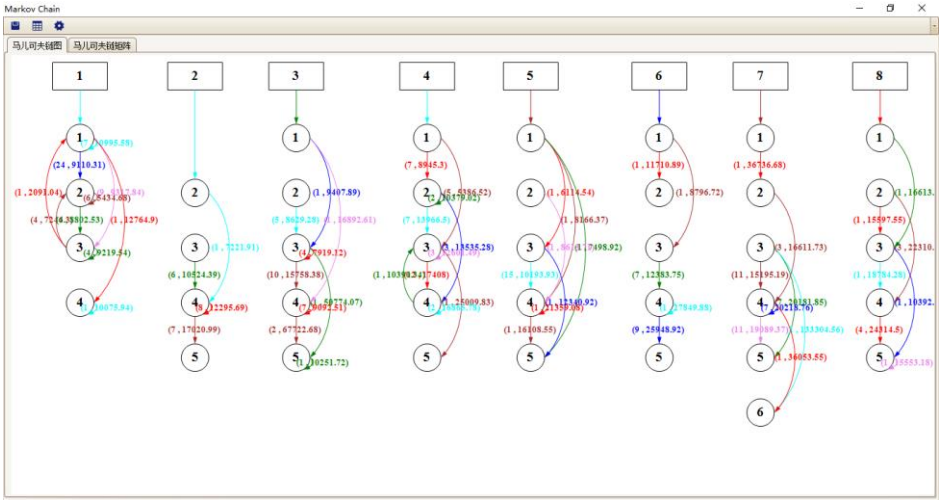


Figure 2.24 Markov Chain

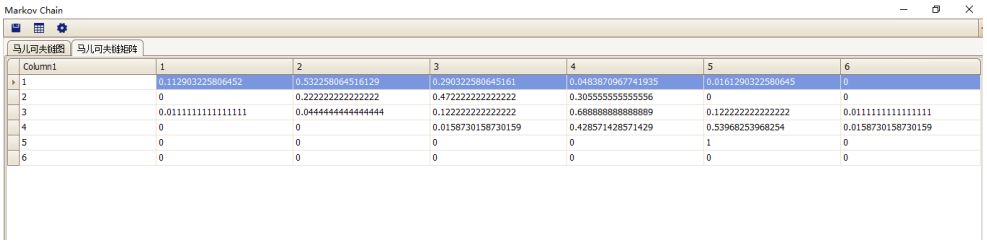


Figure 2.25 Probability Transition Matrix

Save the result as an image: Click "Save" to bring up the "Save As" dialog box, set the image saving path and file name, and save the result as an image.

Export the table: Click "Export" to bring up the "Save As" dialog box, set the file saving path and name, and export the probability transition matrix result as an .xls file.

3 Case 1: Geography evolution of Chinese cities

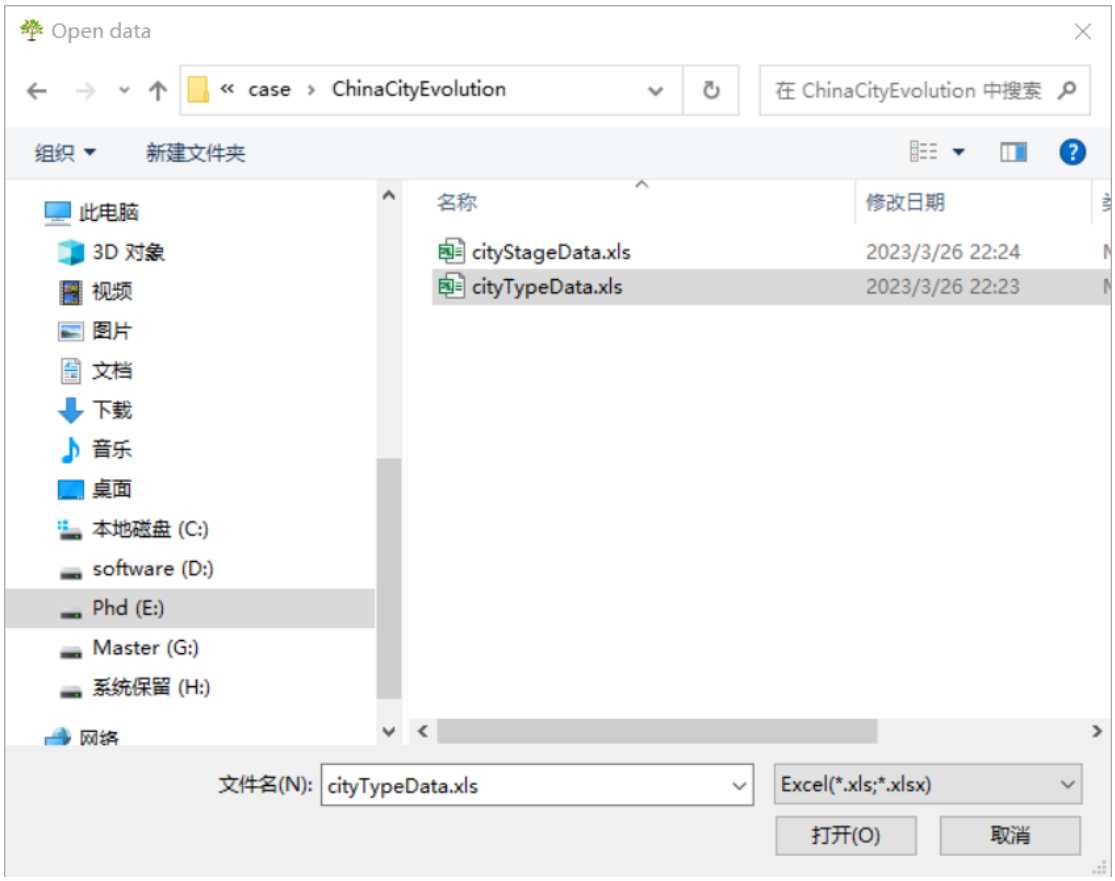
This case study uses the Geotree method to explore the relationship between land expansion of Chinese cities and their types and development stages. The primary branch represents city types, and the secondary branch represents the development stages of cities (Jinfeng Wang et al., 2012). For more details, please refer to Chapter 19 of the 2nd edition of the *Spatial Data Analysis Tutorial* (《空间数据分析教程》).

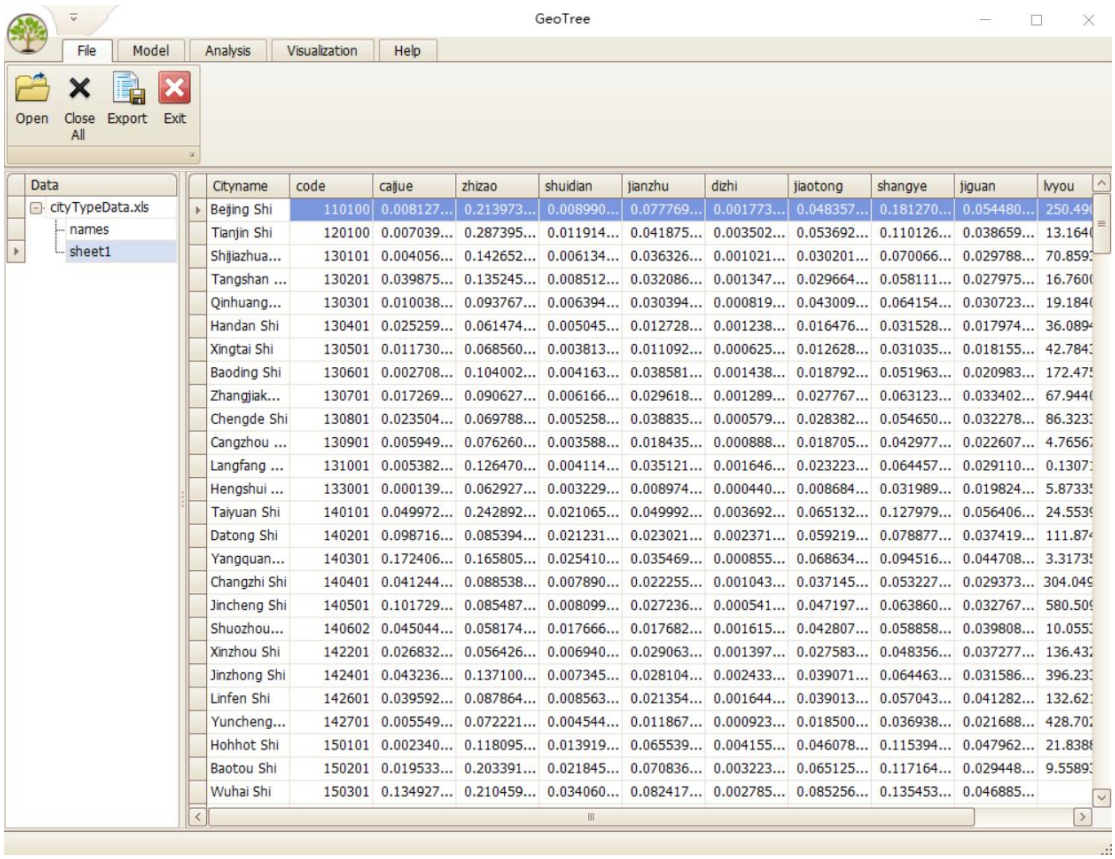
3.1 The primary branch: classification of city types

Urban land expansion is related to the functional types of cities. For example, industrial cities with the same level of development should have a higher urban expansion rate than mining cities. Therefore, cities are first classified by their functional types.

The data from the fifth national census in 2000 divided the economic activities of cities into 16 industries. This case uses this data to classify the types of cities, but due to collinearity issues and other problems, some data were removed and a tourism functional index was added. Finally, "mining industry", "manufacturing industry", "water, electricity and gas industry", "construction industry", "geological exploration industry", "transportation and postal industry", "commercial industry", "government agencies and organizations", and "tourism functional index" were used for multivariate clustering. K-means method was used to cluster and analyze cities based on population data by industry. The specific steps are as follows:

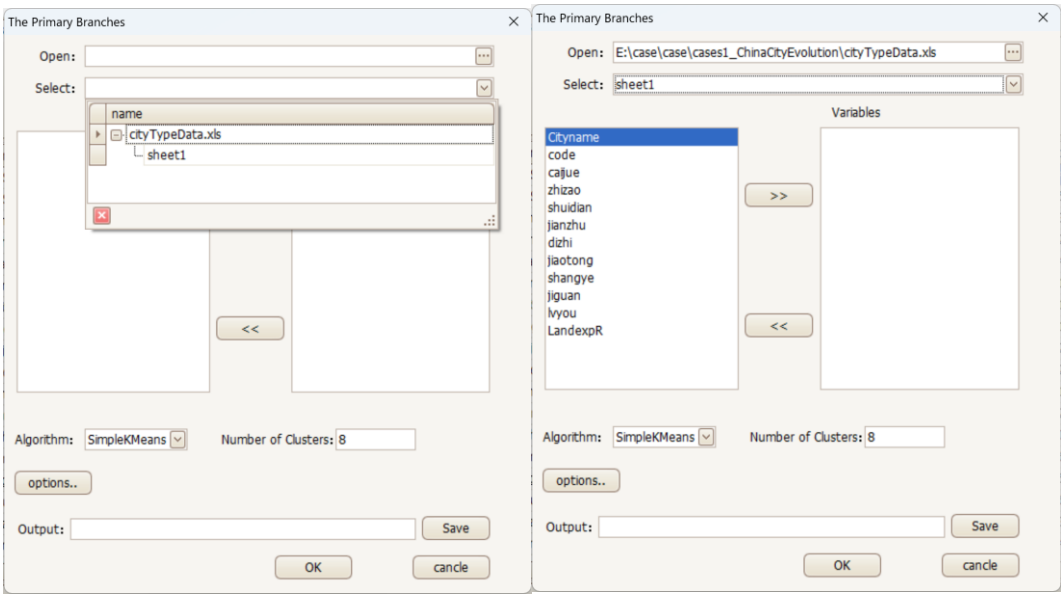
First, switch to the "File" module and click "Open" to import the industry data.



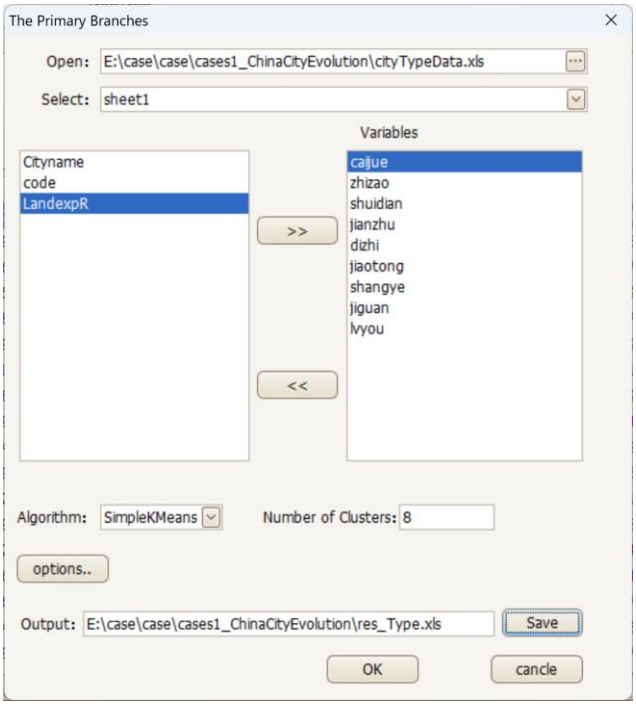


Select the "Analysis" module, click "The Primary Branches", and the following

dialog box will pop up. Select the industry data imported in the previous step from the "Select" dropdown menu.

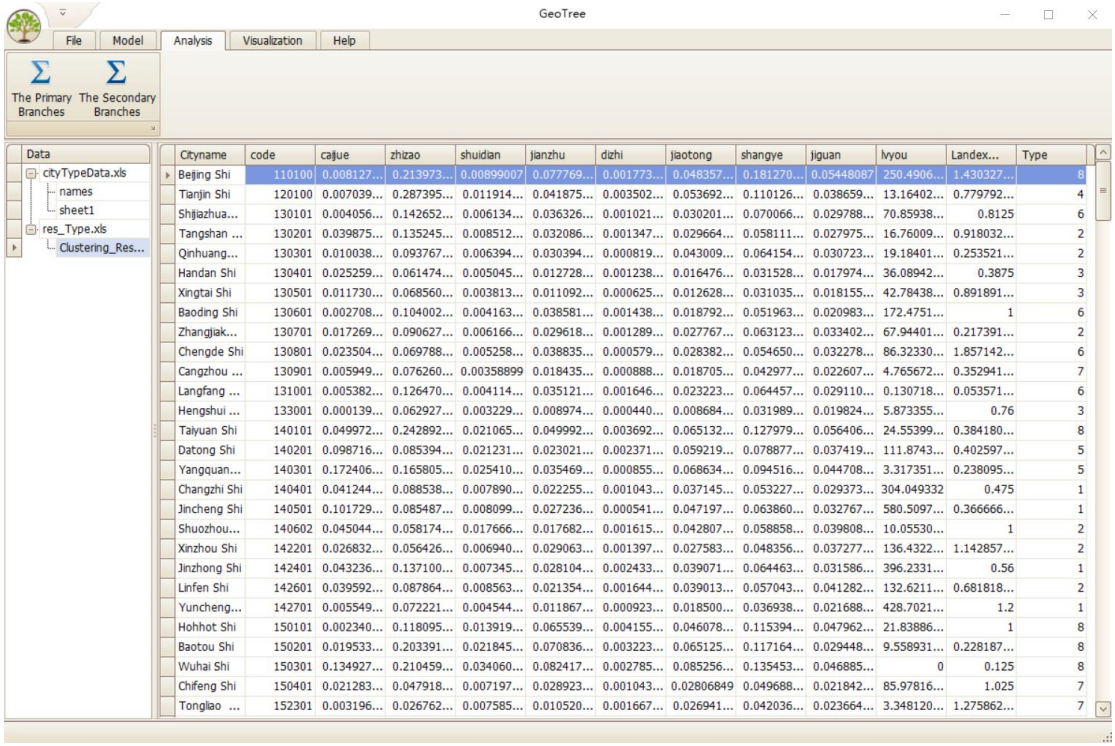


Select the variables to be included in the clustering analysis on the left side and click >> to add them to the right side. Choose the clustering method from the Algorithm dropdown menu, input the number of clusters in the Number of Cluster, set the output path and filename in the Output section. Finally, cclick "OK" OK to obtain the clustering result of city types.



The clustering result will be automatically loaded into the system, and a "Type"

column will be added based on the clustering data to represent the type of city.



Cityname	code	caljue	zhizao	shuidian	jianzhu	dizhi	jiaotong	shangye	jiguan	lvyou	Landex...	Type
Beijing Shi	110100	0.008127...	0.213973...	0.00899007	0.077769...	0.001773...	0.048357...	0.181270...	0.05448087	250.4906...	1.430327...	8
Tianjin Shi	120100	0.007039...	0.287395...	0.011914...	0.041875...	0.003502...	0.053692...	0.110126...	0.038659...	13.16402...	0.779792...	4
Shijiazhua...	130101	0.004056...	0.142652...	0.006134...	0.036326...	0.001021...	0.030201...	0.070066...	0.029788...	70.85938...	0.8125	6
Tangshan ...	130201	0.039875...	0.135245...	0.008512...	0.032086...	0.001347...	0.029664...	0.058111...	0.027975...	16.76009...	0.918032...	2
Qinhuang...	130301	0.010038...	0.093767...	0.006394...	0.030394...	0.000819...	0.043009...	0.064154...	0.030723...	19.18401...	0.253521...	2
Handan Shi	130401	0.025259...	0.061474...	0.005045...	0.012728...	0.001238...	0.016476...	0.031528...	0.017974...	36.08942...	0.3875	3
Xingtai Shi	130501	0.011730...	0.068560...	0.003813...	0.011092...	0.000625...	0.012628...	0.031035...	0.018155...	42.78438...	0.891891...	3
Baoding Shi	130601	0.002708...	0.104002...	0.004163...	0.038581...	0.001438...	0.018792...	0.051963...	0.020983...	172.4751...	1	6
Zhangjiak...	130701	0.017269...	0.090627...	0.006166...	0.029618...	0.001289...	0.027767...	0.063123...	0.033402...	67.94401...	0.217391...	2
Chengde Shi	130801	0.023504...	0.069788...	0.005258...	0.038835...	0.000579...	0.028382...	0.054650...	0.032278...	86.32330...	1.857142...	6
Cangzhou ...	130901	0.005949...	0.076260...	0.00358899	0.018435...	0.000888...	0.018705...	0.042977...	0.022607...	4.765672...	0.352941...	7
Langfang ...	131001	0.005382...	0.126470...	0.004114...	0.035121...	0.001646...	0.023223...	0.064457...	0.029110...	0.130718...	0.053571...	6
Hengshui ...	133001	0.000139...	0.062927...	0.003229...	0.008974...	0.000440...	0.008684...	0.031989...	0.019824...	5.873355...	0.76	3
Taiyuan Shi	140101	0.049972...	0.242892...	0.021065...	0.049992...	0.003692...	0.065132...	0.127979...	0.056406...	24.55399...	0.384180...	8
Datong Shi	140201	0.098716...	0.085394...	0.021231...	0.023021...	0.002371...	0.059219...	0.078877...	0.037419...	111.8743...	0.402597...	5
Yangquan...	140301	0.172406...	0.165805...	0.025410...	0.035469...	0.000855...	0.068634...	0.094516...	0.044708...	3.317351...	0.238095...	5
Changzhi Shi	140401	0.041244...	0.088538...	0.007890...	0.022255...	0.001043...	0.037145...	0.053227...	0.029373...	304.049332	0.475	1
Jincheng Shi	140501	0.101729...	0.085487...	0.008099...	0.027236...	0.000541...	0.047197...	0.063860...	0.032767...	580.5097...	0.366666...	1
Shuozhou...	140602	0.045044...	0.058174...	0.017666...	0.017682...	0.001615...	0.042807...	0.058858...	0.039808...	10.05530...	1	2
Xinzhou Shi	142201	0.026832...	0.056426...	0.006940...	0.029063...	0.001397...	0.027583...	0.048356...	0.037277...	136.4322...	1.142857...	2
Jinzhong Shi	142401	0.043236...	0.137100...	0.007345...	0.028104...	0.002433...	0.039071...	0.064463...	0.031586...	396.2331...	0.56	1
Linfen Shi	142601	0.039592...	0.087864...	0.008563...	0.021354...	0.001644...	0.039013...	0.057043...	0.041282...	132.6211...	0.681818...	2
Yuncheng...	142701	0.005549...	0.072221...	0.004544...	0.011867...	0.000923...	0.018500...	0.036938...	0.021688...	428.7021...	1.2	1
Hohhot Shi	150101	0.002340...	0.118095...	0.013919...	0.065539...	0.004155...	0.046078...	0.115394...	0.047962...	21.83886...	1	8
Baotou Shi	150201	0.019533...	0.203391...	0.021845...	0.070836...	0.003223...	0.065125...	0.117164...	0.029448...	9.558931...	0.228187...	8
Wuhai Shi	150301	0.134927...	0.210459...	0.034060...	0.082417...	0.002785...	0.085256...	0.135453...	0.046885...	0	0.125	8
Chifeng Shi	150401	0.021283...	0.047918...	0.007197...	0.028923...	0.001043...	0.02806849	0.049688...	0.021842...	85.97816...	1.025	7
Tongliao ...	152301	0.003196...	0.026762...	0.007585...	0.010520...	0.001667...	0.026941...	0.042036...	0.023664...	3.348120...	1.275862...	7

3.2 Secondary branch: classification of city development stages

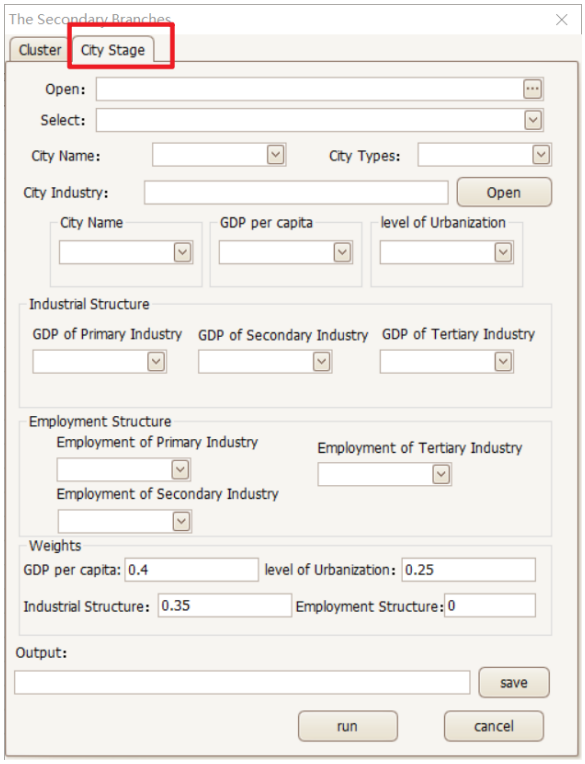
The transition from agricultural economy to industrial economy and then to service economy is a general law of economic development. Along with the process of industrialization, the socio-economic structure exhibits certain stage characteristics, and urban expansion should also exhibit stage characteristics in accordance with the economic development stages.

The indicators for determining the stages of industrialization are based on Chenery's and relevant urbanization stage theories (Table 3.1). Based on the per capita GDP, industrial structure, employment structure, and urbanization level data of 242 cities in China in 2000 and 2010, the economic stages of these cities were calculated.

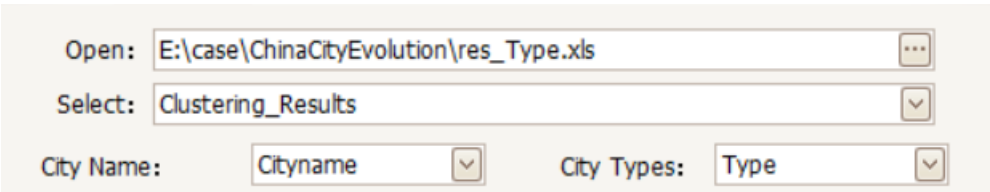
Table 3.1 Indicators of stage of city development, by GDP range

Satge	Per capita GDP in 1980 （USD）	Industrial structure/%			Employment structure/%			Urbanization level/%	Economic stages of development	
		Primary industry	Secondary industry	Tertiary industry	Primary industry	Secondary industry	Tertiary industry			
1	300~600	39	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	Intermediate stage	
4	2400~4500	13.5	46	40.5	36.5	25.5	38	54	Advanced stage	
5	4500~7200	9	51	40	20	30	50	70	Elementary stage	Developed stages
6	7200~10800	3	47	50	8	30	62	80	Advanced stage	

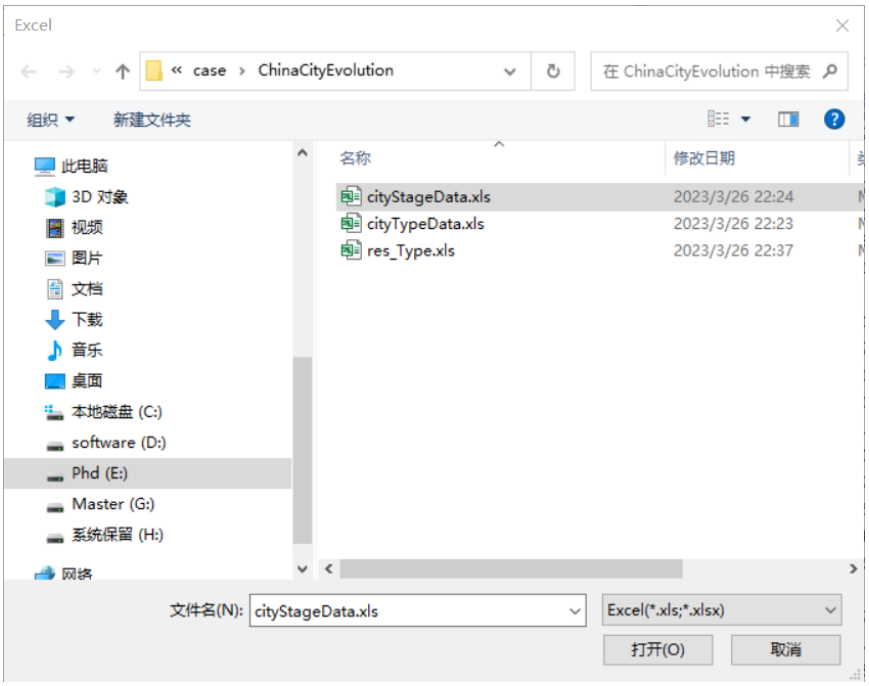
Firstly, select the "Analysis" module, click on "The Secondary Branches", and then choose the "City Stage" interface.

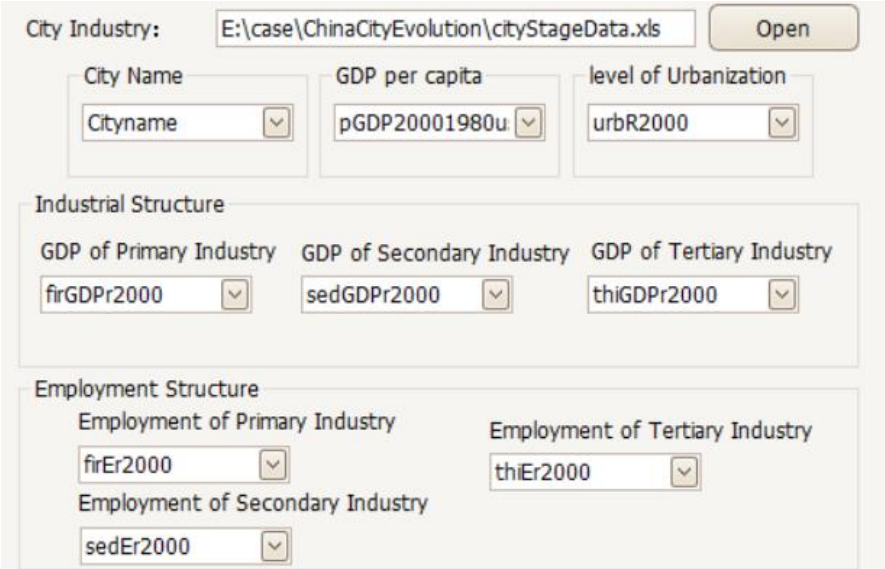


Click on the "Select" dropdown menu, choose the clustering result from the previous step, and set the city name and type fields.

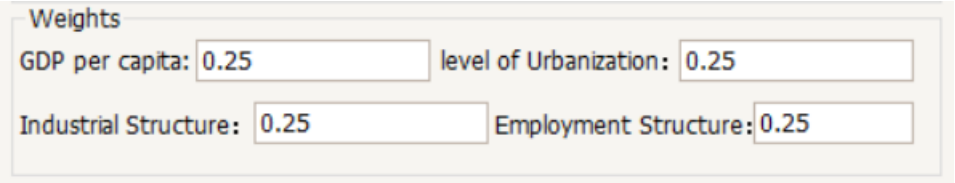


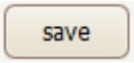
Click on "Open" Open to import the data for determining the development stages of cities, and set the corresponding fields according to the following figure.





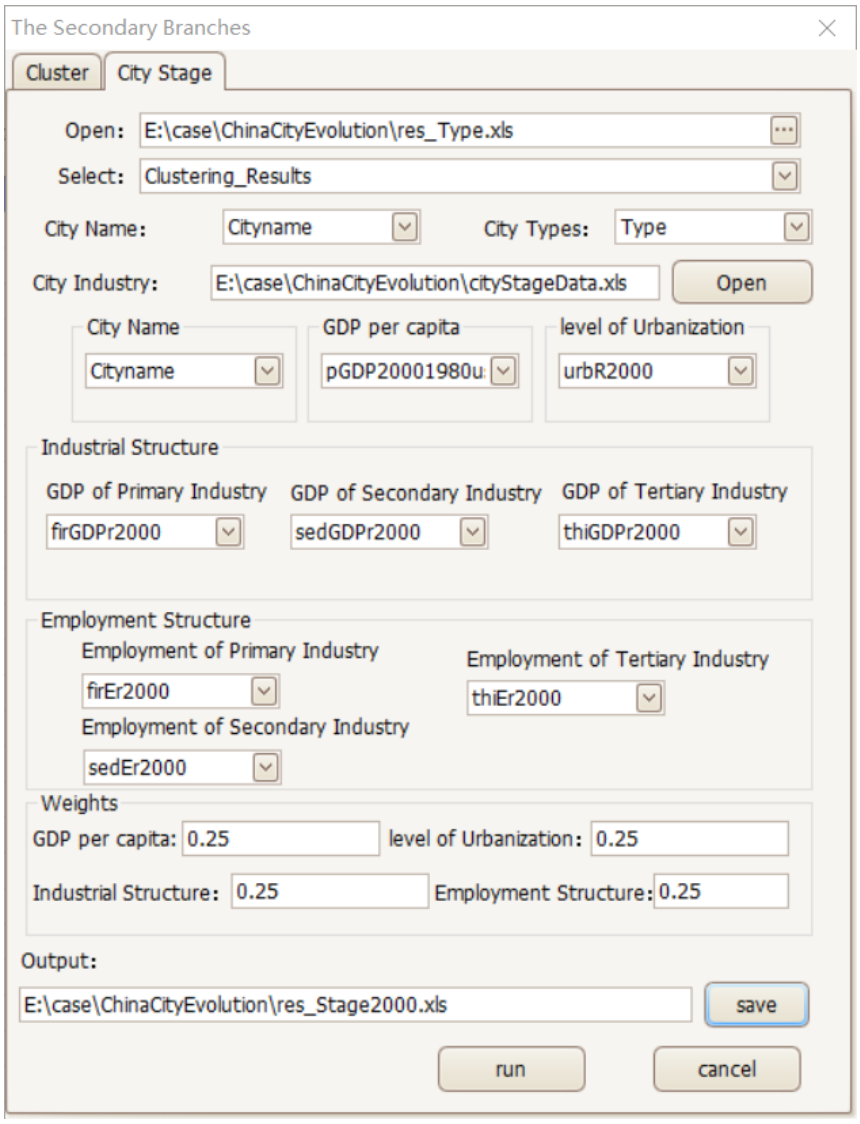
Based on the specific situation, weights are assigned to calculate the final development stage of the cities. The weight configuration for this case study is as follows:

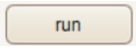


Click "Save" 

to set the output path and file name for the result.

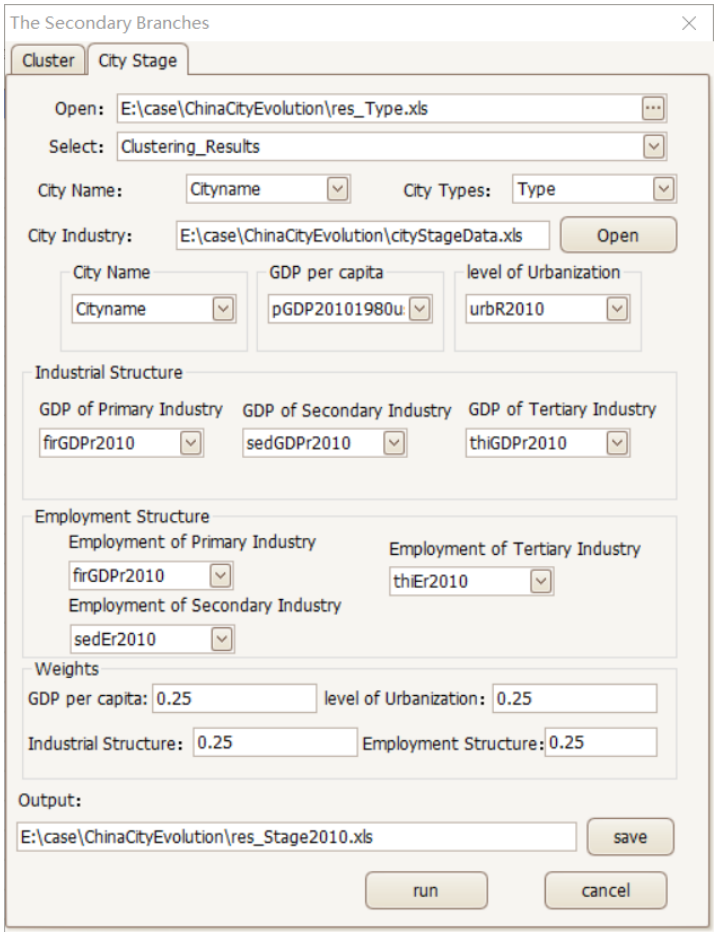
All parameter settings are shown in the following figure.

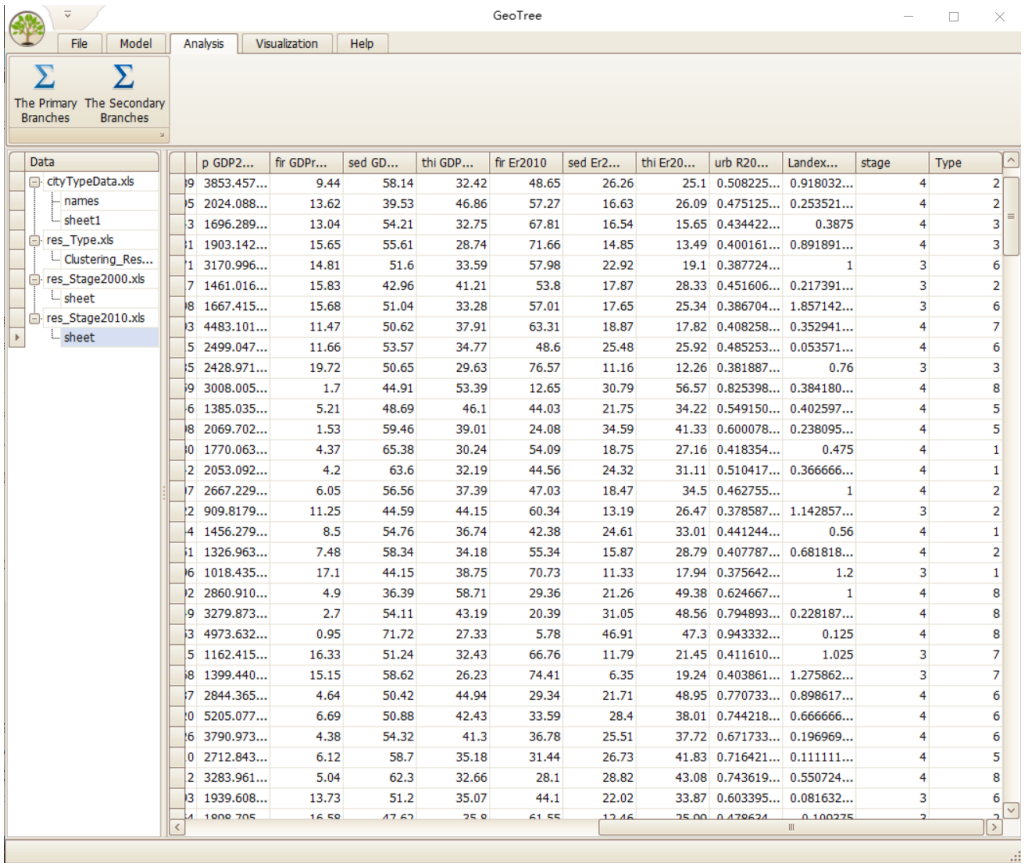


Finally, click "Run" 

to start the process and obtain the results of the city development stages. The results will be automatically loaded into the system, with a new column called "Stage" added to the original data, indicating the development stage of each city.

Using the same steps as described above, the city development stages for 2010 were obtained, with the following parameter settings and results as shown in the figure below.

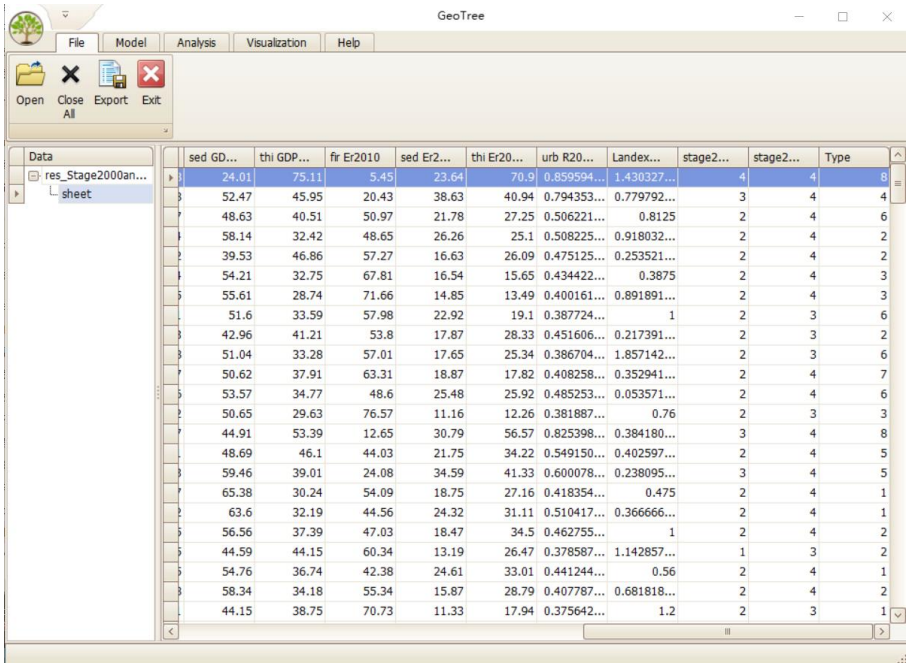

The Secondary Branches
Cluster
City Stage
Open: E:\case\ChinaCityEvolution\res_Type.xls
Select: Clustering_Results
City Name: Cityname
City Types: Type
City Industry: E:\case\ChinaCityEvolution\cityStageData.xls
Open
City Name
Cityname
GDP per capita
pGDP20101980u
level of Urbanization
urbR2010
Industrial Structure
GDP of Primary Industry
firGDPr2010
GDP of Secondary Industry
sedGDPr2010
GDP of Tertiary Industry
thiGDPr2010
Employment Structure
Employment of Primary Industry
firGDPr2010
Employment of Tertiary Industry
thiEr2010
Employment of Secondary Industry
sedEr2010
Weights
GDP per capita: 0.25
level of Urbanization: 0.25
Industrial Structure: 0.25
Employment Structure: 0.25
Output:
E:\case\ChinaCityEvolution\res_Stage2010.xls
save
run
cancel


GeoTree
File
Model
Analysis
Visualization
Help
The Primary Branches
The Secondary Branches
Data
cityTypeData.xls
names
sheet1
res_Type.xls
Clustering_Res...
res_Stage2000.xls
sheet
res_Stage2010.xls
sheet
p GDP2...
fir GDPr...
sed GD...
thi GDP...
fir Er2010
sed Er2...
thi Er20...
urb R20...
Landex...
stage
Type

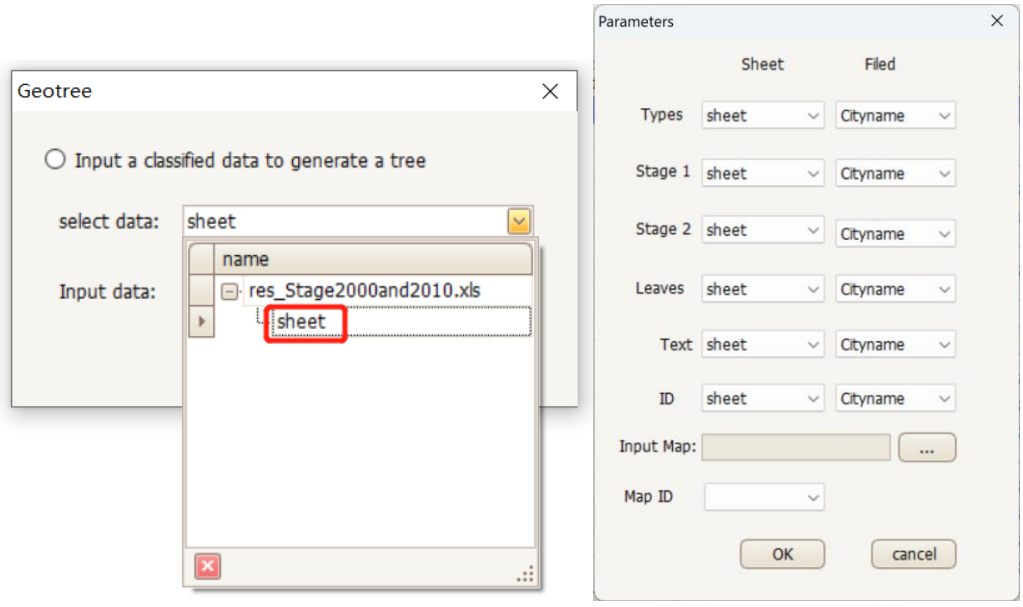
3.3 Geotree construction

Based on the above steps to obtain the development types and stages of cities, construct a Geotree. The specific steps are as follows:

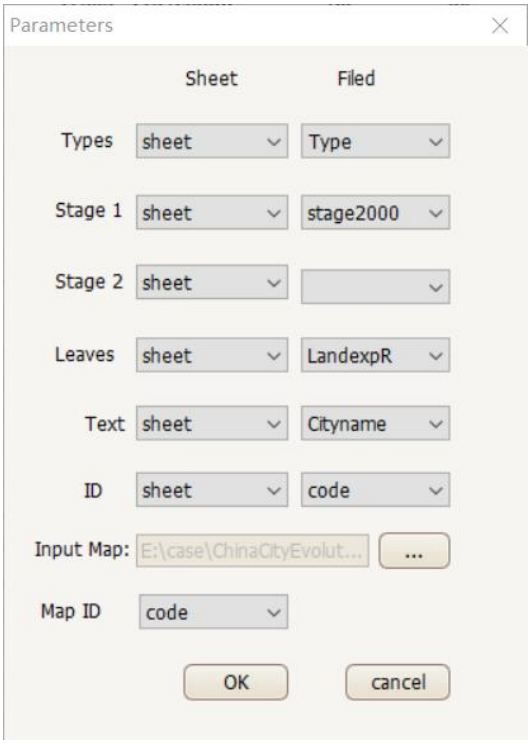
First, import the obtained city type and development stage in 2000 into the system.



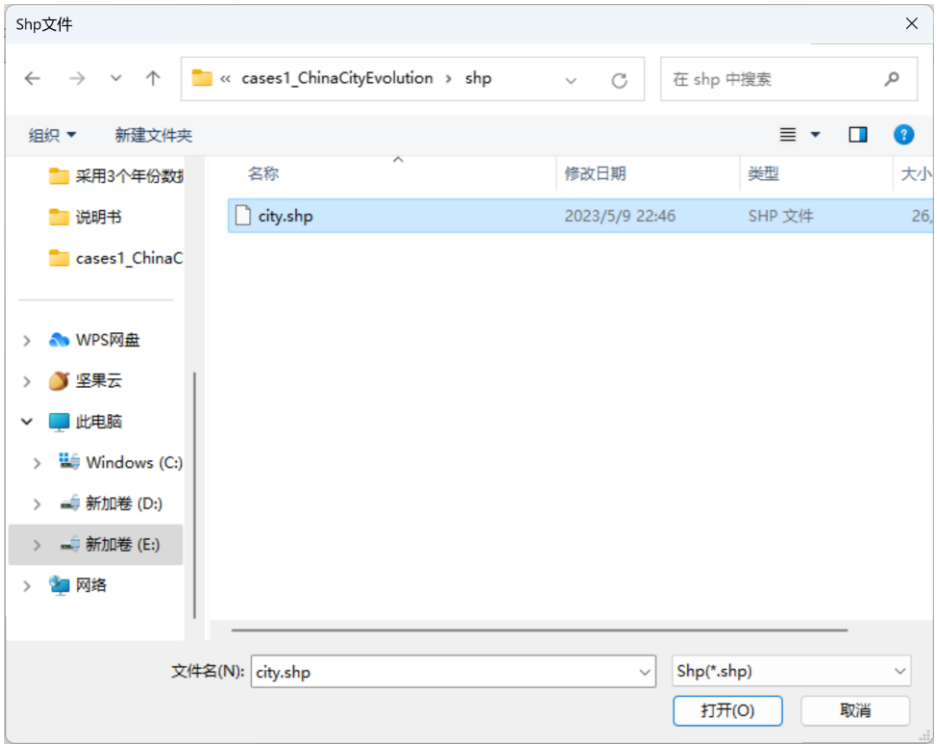
Select the "Visualization" module, click on "Geotree" and the following interface will pop up. Select the data that includes the city types and development stages that were imported in the previous step, then click "OK". This will open the tree model parameter settings box.



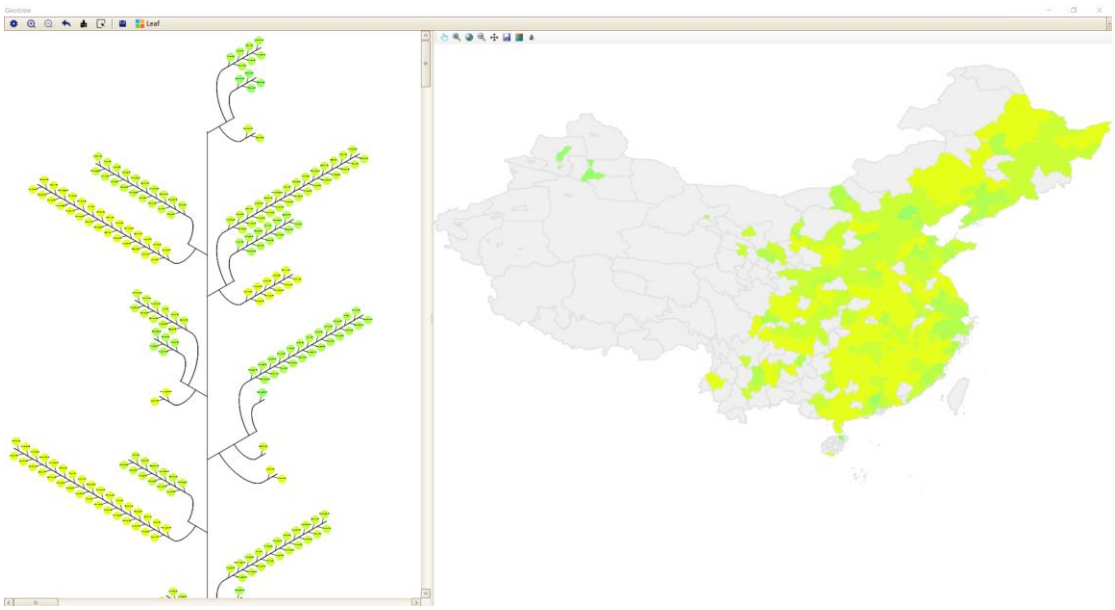
Set the parameters for the Geotree: "Type" is city type, "Stage1" is the 2000 city development stage, "Stage2" is set to empty, "Leaves" is city expansion rate, " Text " is city name, and ID is the unique identifier for the city.



Click  after the "Input Map" button to import the vector layer of the cities and set the "Map ID" to the field that corresponds to the city's unique identifier (ID).



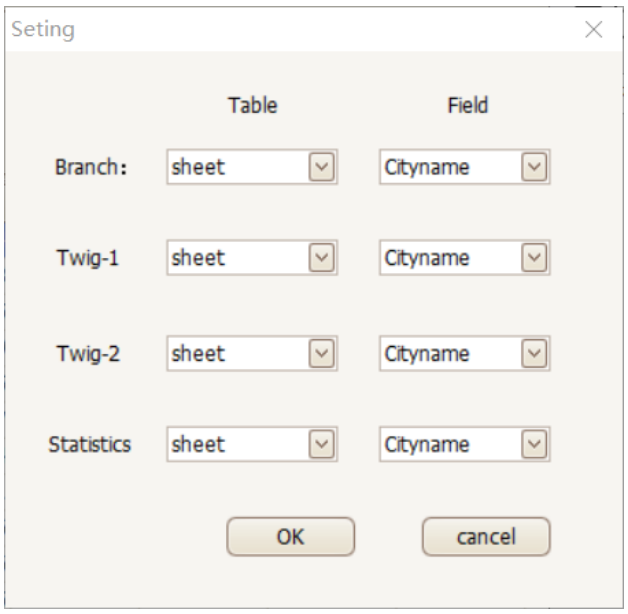
Finally, click "OK" to generate the Geotree and corresponding map.



3.4 Development stage transitions

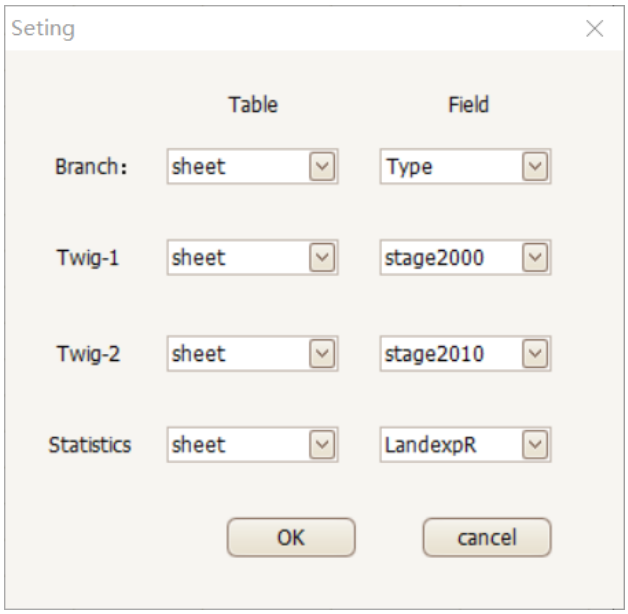
Assuming that the city types remained unchanged between 2000 and 2010, we use Markov Chain to study the transitions between different development stages of cities during this period.

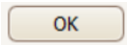
In the "Visualization" module, click "Markov Chain" to open the following window.

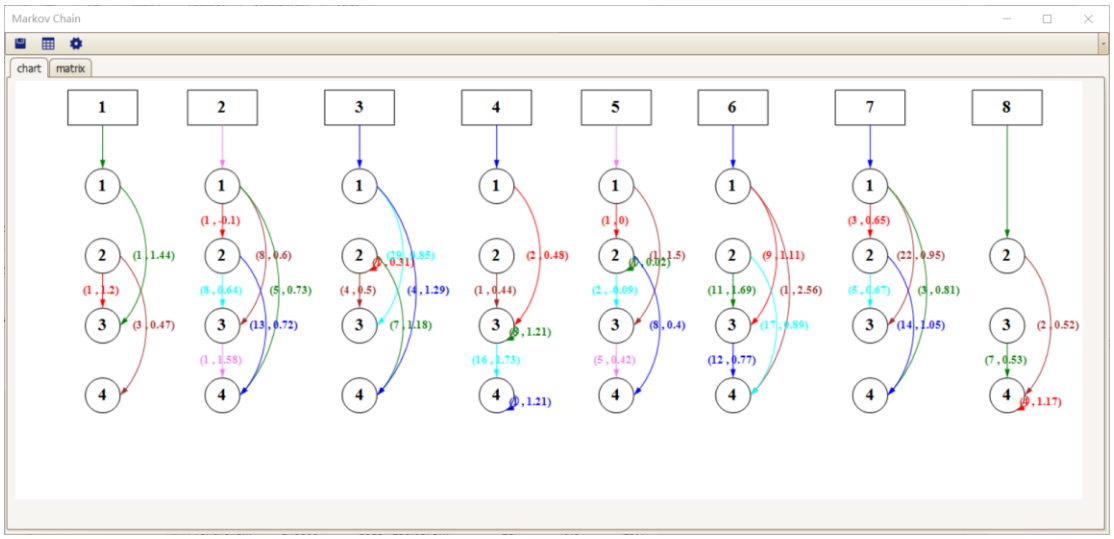


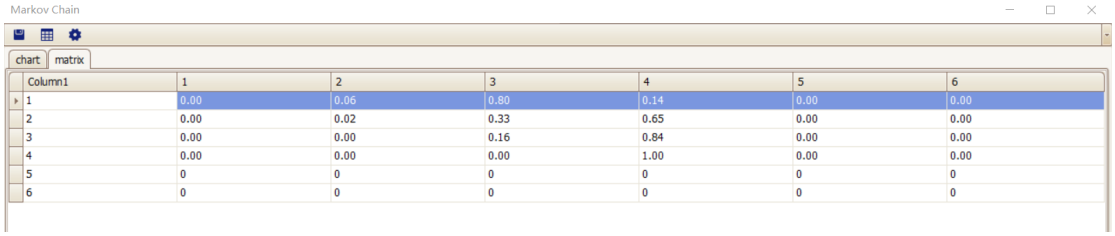
Set the parameters for Markov Chain: "Branch" is set as the field of city types,

"Twig-1" is set as the city development stage in 2000, "Twig-2" is set as the city development stage in 2010, and "Statistics" is set as the urban land expansion rate.



After setting the above parameters, click "OK"  to obtain the Markov Chain results. The first figure shows the number of cities that underwent development stage transitions and the average urban land expansion rate for each type of city. The second figure is the probability transition matrix for urban economic development.





4 Case 2: Global Geotree of non-communicable disease incidence rate

Non-communicable diseases are a major health and development challenge for humans and are closely related to socio-economic development. Therefore, the deaths caused by non-communicable diseases should vary in different socio-economic development stages. In this case, a global non-communicable disease evolution tree was constructed based on socio-economic types and development stages (Yang Wang et al., 2020).

4.1 The primary branch: classification of country types

According to the World Bank's income classification standard in 2018, economies with a per capita Gross National Income (GNI) of 1025 USD or less were classified as low-income economies, those with a GNI between 1026 and 3995 USD were classified as lower-middle-income economies, those with a GNI between 3996 and 12375 USD were classified as upper-middle-income economies, and those with a GNI of 12376 USD or above were classified as high-income economies. Based on the per capita GDP level, 176 countries were divided into four different country types: (I) low-income economies (26 countries), (II) lower-middle-income economies (50 countries), (III) upper-middle-income economies (51 countries), and (IV) high-income economies (49 countries).

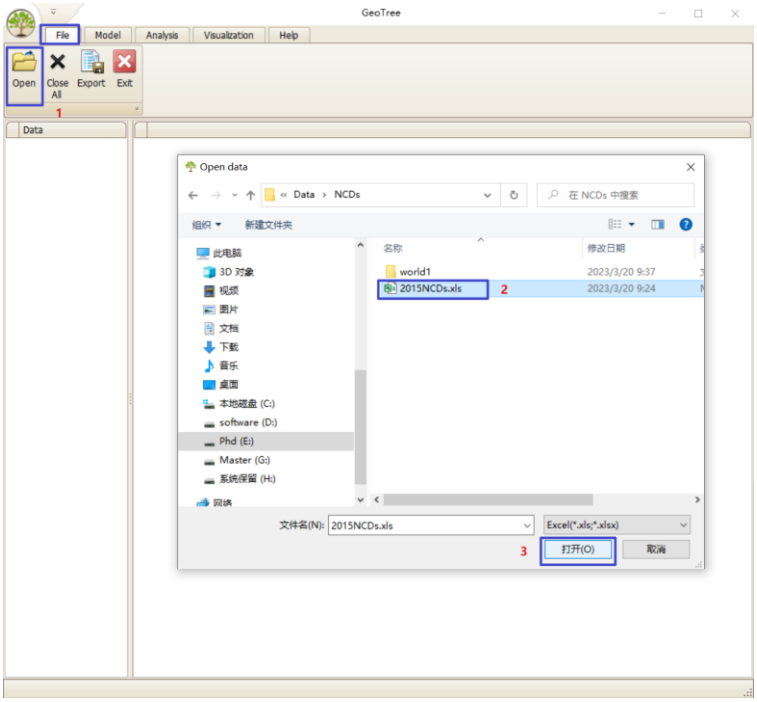
4.2 Secondary branch: classification of country development stages

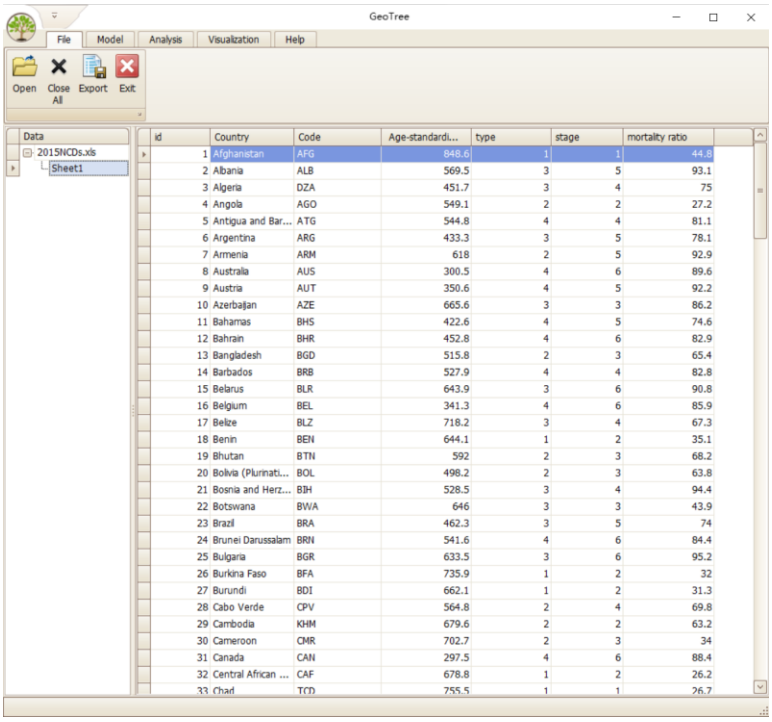
The development stages of countries are measured by urbanization and neonatal mortality rates, which are indicators of social development level and healthcare level. Each indicator is divided into six levels using the natural discontinuity method. The

weights are calculated by the explanatory power of the influencing factors on NCD mortality and are finally synthesized into the overall development stage.

4.3 Geotree construction

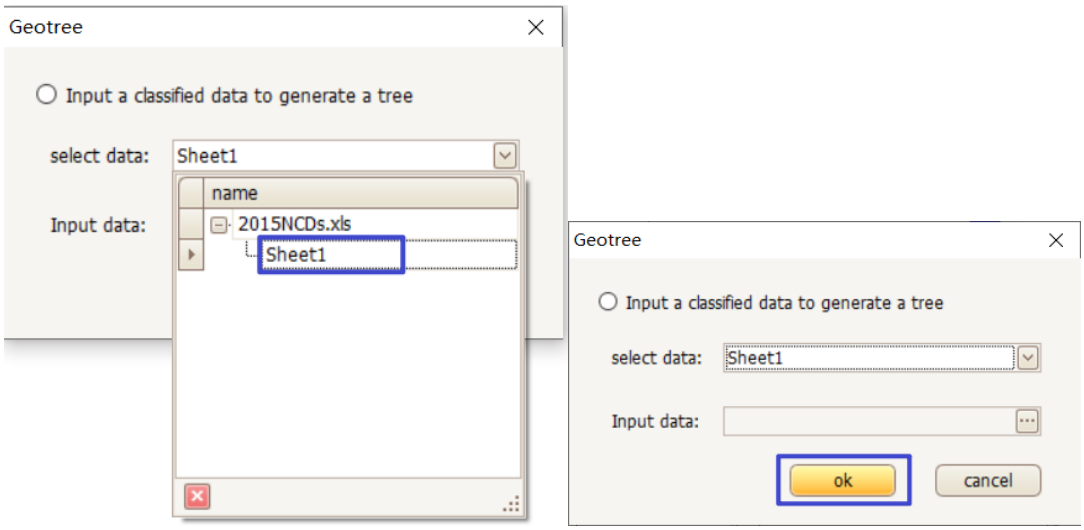
First, open the GeoTree software. The software opens by default in the "File" module. Click "Open" to import data.



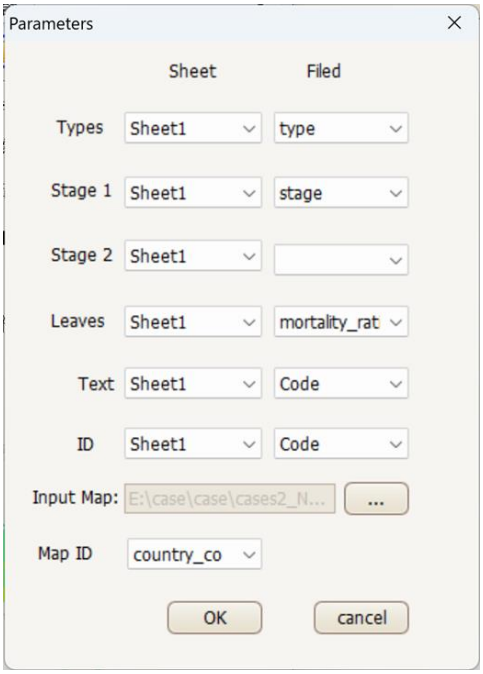


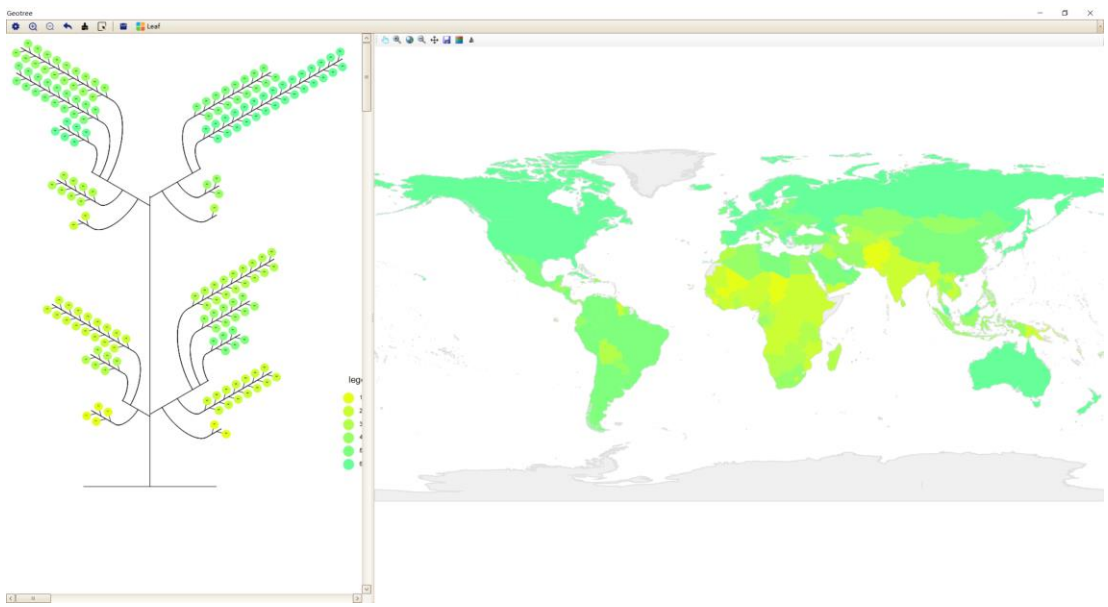
id	Country	Code	Age-standardi...	type	stage	mortality ratio
1	Afghanistan	AFG	848.6	1	1	44.8
2	Albania	ALB	569.5	3	5	93.1
3	Algeria	DZA	451.7	3	4	75
4	Angola	AGO	549.1	2	2	27.2
5	Antigua and Bar...	ATG	544.8	4	4	81.1
6	Argentina	ARG	433.3	3	5	78.1
7	Armenia	ARM	618	2	5	92.9
8	Australia	AUS	300.5	4	6	89.6
9	Austria	AUT	350.6	4	5	92.2
10	Azerbaijan	AZE	665.6	3	3	86.2
11	Bahamas	BHS	422.6	4	5	74.6
12	Bahrain	BHR	452.8	4	6	82.9
13	Bangladesh	BGD	515.8	2	3	65.4
14	Barbados	BRB	527.9	4	4	82.8
15	Belarus	BLR	643.9	3	6	90.8
16	Belgium	BEL	341.3	4	6	85.9
17	Belize	BLZ	718.2	3	4	67.3
18	Benin	BEN	644.1	1	2	35.1
19	Bhutan	BTN	592	2	3	68.2
20	Bolivia (Plurinati...	BOL	498.2	2	3	63.8
21	Bosnia and Herz...	BIH	528.5	3	4	94.4
22	Botswana	BWA	646	3	3	43.9
23	Brazil	BRA	462.3	3	5	74
24	Brunei Darussalam	BRN	541.6	4	6	84.4
25	Bulgaria	BGR	633.5	3	6	95.2
26	Burkina Faso	BFA	735.9	1	2	32
27	Burundi	BDI	662.1	1	2	31.3
28	Cabo Verde	CPV	564.8	2	4	69.8
29	Cambodia	KHM	679.6	2	2	63.2
30	Cameroon	CMR	702.7	2	3	34
31	Canada	CAN	297.5	4	6	88.4
32	Central African ...	CAF	678.8	1	2	26.2
33	Chad	TCD	755.5	1	1	26.7

Switch to the "Visualization" module and click on "Geotree" to start building the evolution tree. Click the dropdown menu to select the data imported in the previous step, and then click "OK".



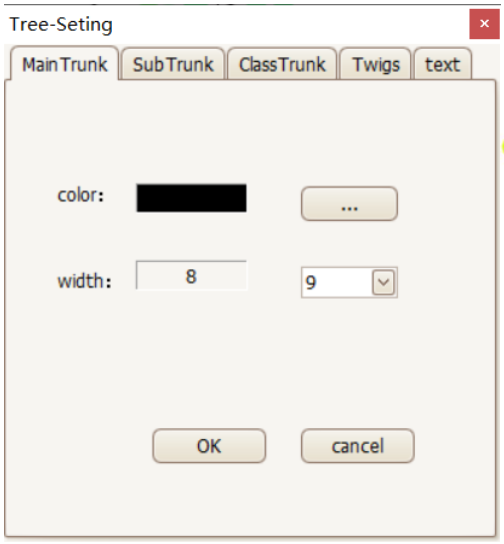
In the Parameters box, Set "Type" as the country type, "Stage1" as the development stage of the country, and "Stage2" as the development stage of other years. In this case, "Stage2" is left empty. Set "Text" as the name of the country and "ID" as the unique identifier for each country. Click ... to import map data, and then set the Map ID that corresponds to the above ID. Otherwise, the generated tree cannot interact with the map. Finally, click OK to generate Geotree. The left side of the result page is the tree view, and the right side is the map view.

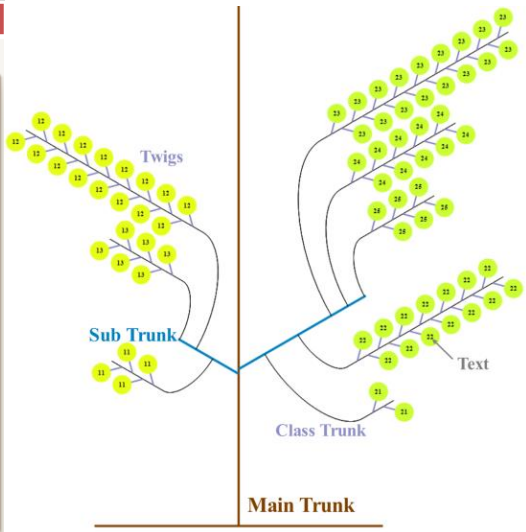




In the menu bar above the tree view, you can click "Setting"  and "Leaf"  to set the tree style.

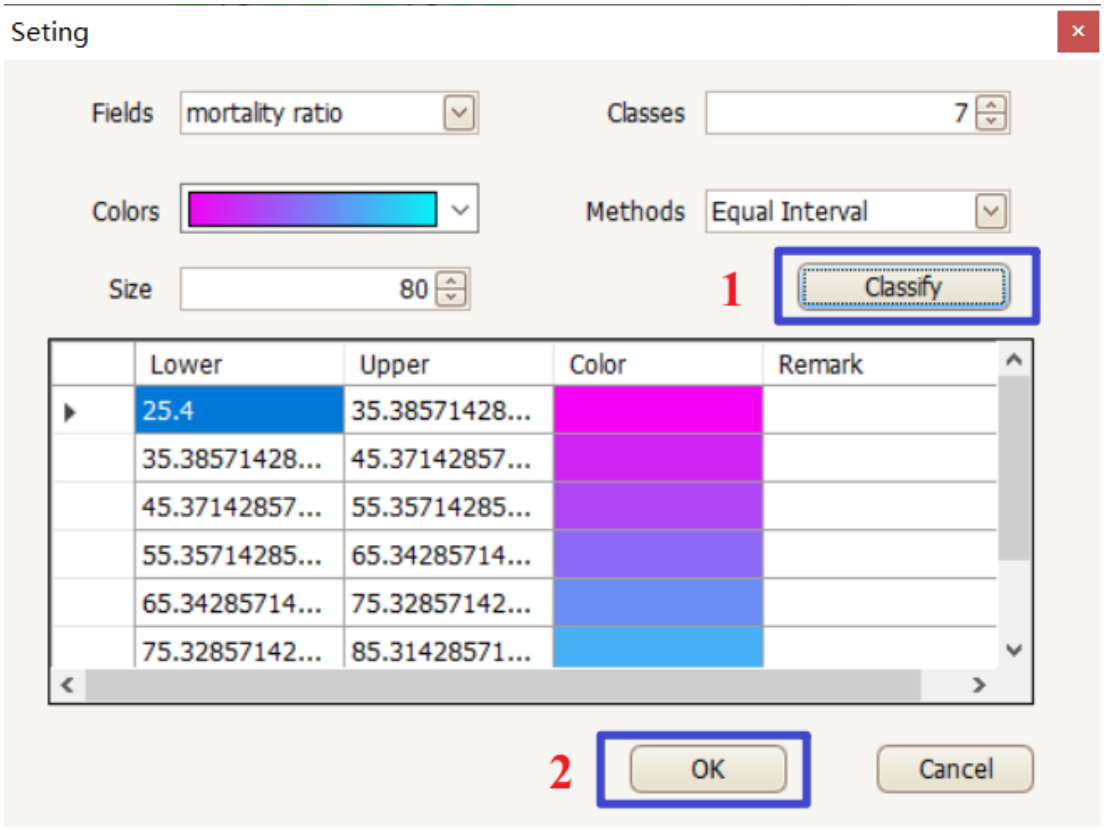
In the leaf settings  , you can change the color and thickness of the branches at each level.

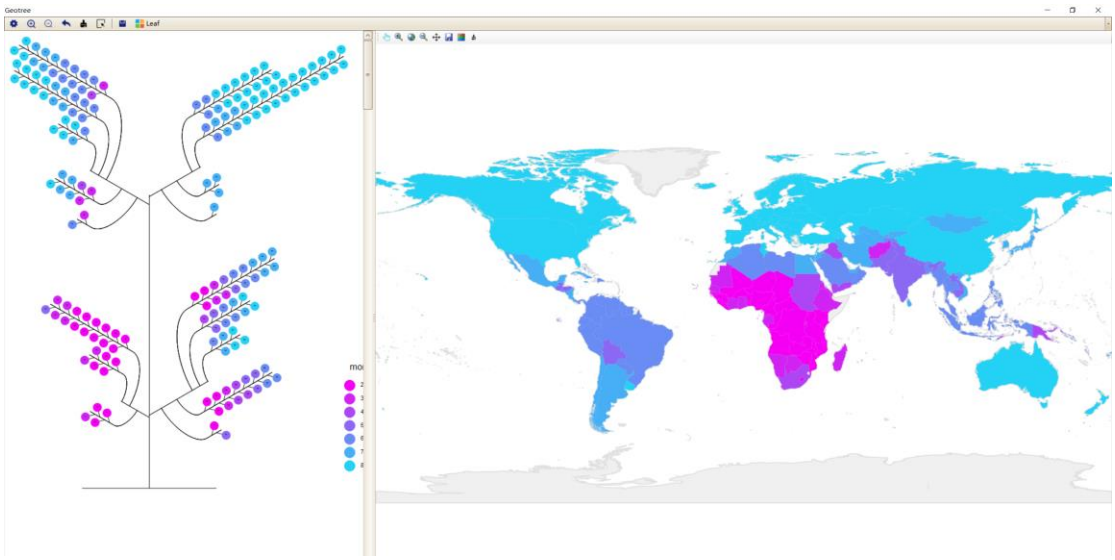




In the leaf settings  , select the field for color classification in the "Fields" drop-down menu. Set the number of categories in "Classes," choose a color scheme in "Colors", set the classification method in "Methods," and adjust the size of the leaves in "Sizes." After setting these parameters, click "Classify" to generate the upper and lower limits and colors for each interval of the current classification method in the lower box. You can directly modify the upper and lower limits and the color of the generated

interval. Finally, click "OK" to apply the current classification to the Geotree.





After setting the tree style, click "Export" to output the current tree as a JPG image. In the map view on the right side, click "Export" to output the current map as a JPG image as well.

5 Case 3: Tree-like evolution of global urban land expansion

Urban land expansion impacts global sustainable development and correlates closely with socioeconomic growth. Industrialization and urbanization are the main features of global socioeconomic growth. The increase in industries and the urban population requires sufficient urban land to provide hosting space, so global urban land has been expanding dramatically (Jing et al., 2022).

This case combines socioeconomic data from 162 countries in 2005 and 2015 and uses the Geotree model to explore the relationship between global urban land expansion and country type and development stage. The primary branch is the type of industrial structure, and the secondary branch is the development stage of the country.

5.1 Branch division

Based on the proportion of primary and tertiary industries, 162 countries is divided into four types: (I) Initial, (II) Accelerated, (III) Mature, and (IV) Post-Industrialization. Referring to the urbanization process and the World Bank's classification of economies, the 162 countries are divided into four development stages based on urban population and per capita GDP: low, middle, upper-middle, and high.

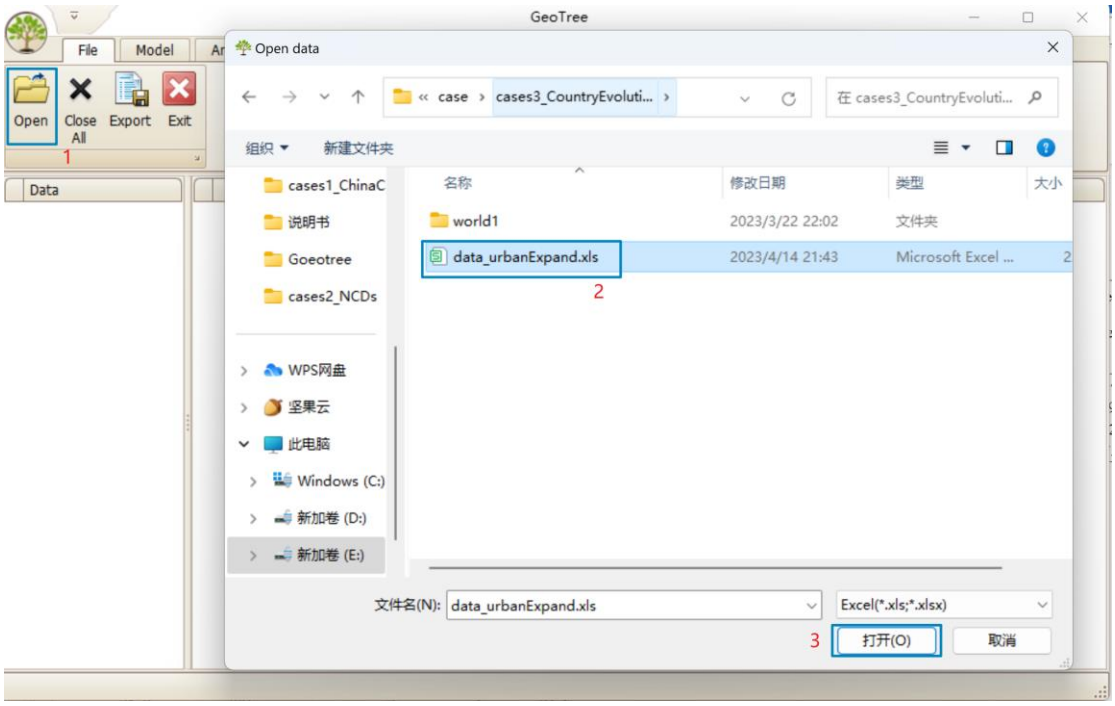
Table 5.1 f Indicators of industrial structure types and development stages.

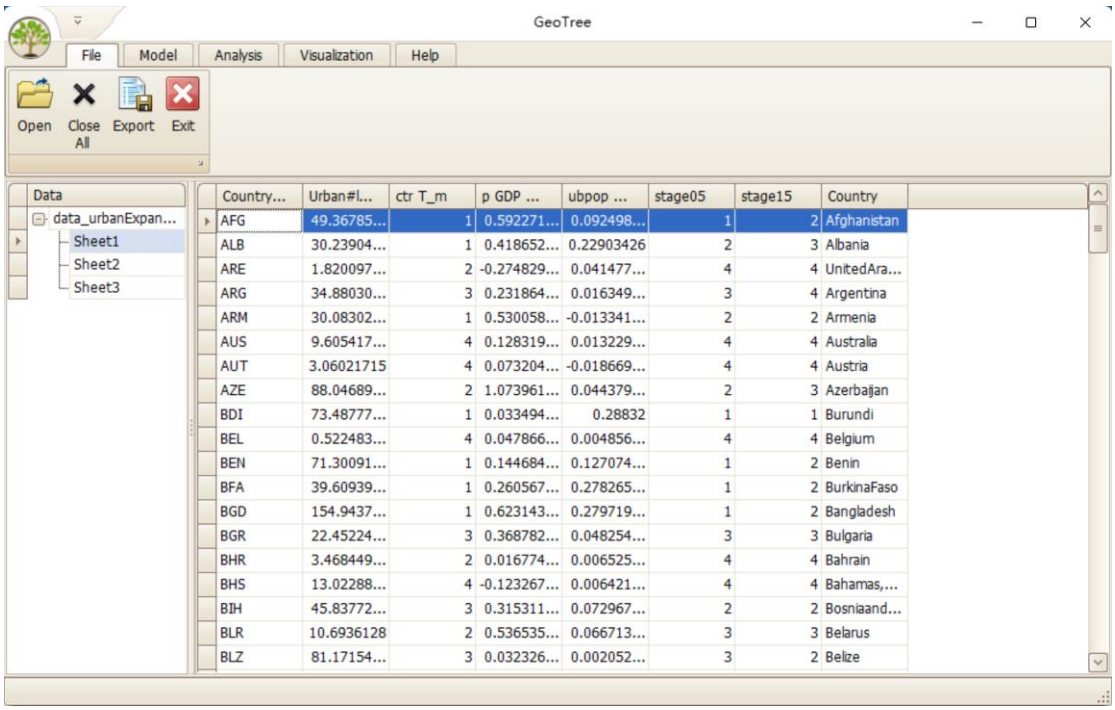
Industrial structure types (Development stages)	Industrial structure types		Development stages	
	primary industry (%)	tertiary industry (%)	urban population (%)	GDP per capita ($)
I (1)	＞15	＜62	＜19	＜1025
II (2)	＜15	＜62	19-50	1025-3995
III (3)	＞3	＞62	50-81	3995-12375
IV (4)	＜3	＞62	＞81	＞12375

The GDP per capita in this study is used in constant 2005 USD.

5.2 Geotree construction

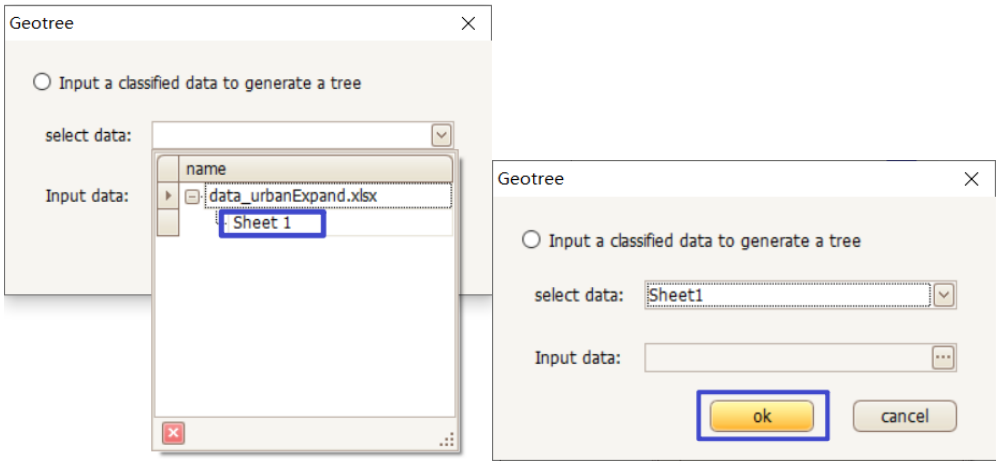
To import data, first open the "File" module and click "Open". This will bring up a file selection window. Select the prepared data file and click "Open" to import the data into the system.



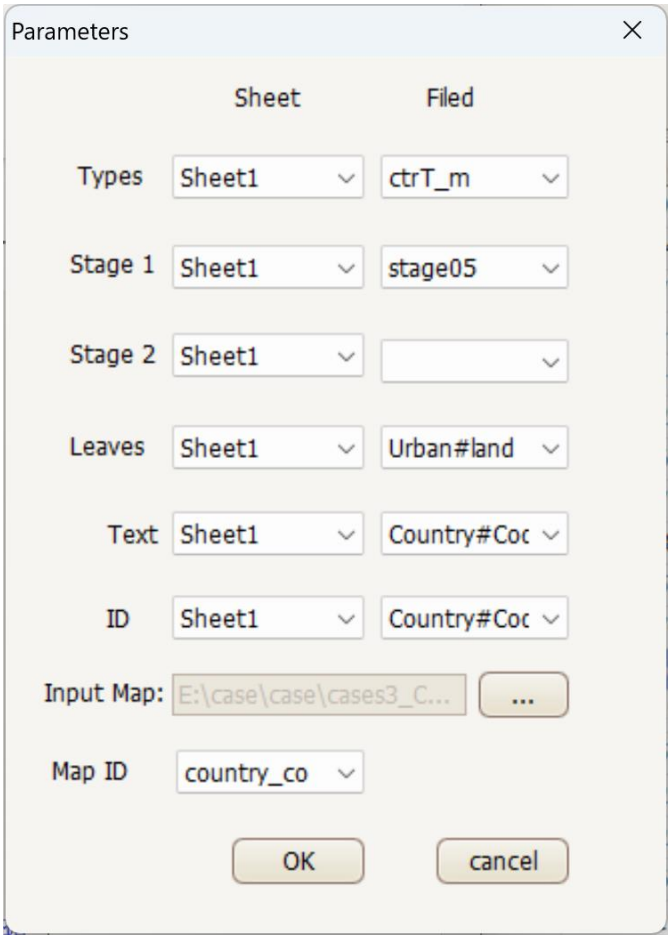


Country...	Urban#l...	ctr T_m	p GDP ...	ubpop ...	stage05	stage15	Country
AFG	49.36785...	1	0.592271...	0.092498...	1	2	Afghanistan
ALB	30.23904...	1	0.418652...	0.22903426	2	3	Albania
ARE	1.820097...	2	-0.274829...	0.041477...	4	4	UnitedAra...
ARG	34.88030...	3	0.231864...	0.016349...	3	4	Argentina
ARM	30.08302...	1	0.530058...	-0.013341...	2	2	Armenia
AUS	9.605417...	4	0.128319...	0.013229...	4	4	Australia
AUT	3.06021715	4	0.073204...	-0.018669...	4	4	Austria
AZE	88.04689...	2	1.073961...	0.044379...	2	3	Azerbaijan
BDI	73.48777...	1	0.033494...	0.28832	1	1	Burundi
BEL	0.522483...	4	0.047866...	0.004856...	4	4	Belgium
BEN	71.30091...	1	0.144684...	0.127074...	1	2	Benin
BFA	39.60939...	1	0.260567...	0.278265...	1	2	BurkinaFaso
BGD	154.9437...	1	0.623143...	0.279719...	1	2	Bangladesh
BGR	22.45224...	3	0.368782...	0.048254...	3	3	Bulgaria
BHR	3.468449...	2	0.016774...	0.006525...	4	4	Bahrain
BHS	13.02288...	4	-0.123267...	0.006421...	4	4	Bahamas,...
BIH	45.83772...	3	0.315311...	0.072967...	2	2	Bosniaand...
BLR	10.6936128	2	0.536535...	0.066713...	3	3	Belarus
BLZ	81.17154...	3	0.032326...	0.002052...	3	2	Belize

Switch to the "Visualization" module and click "Geotree" to open the following window. Select the data you imported in the previous step from the dropdown menu and click "OK".

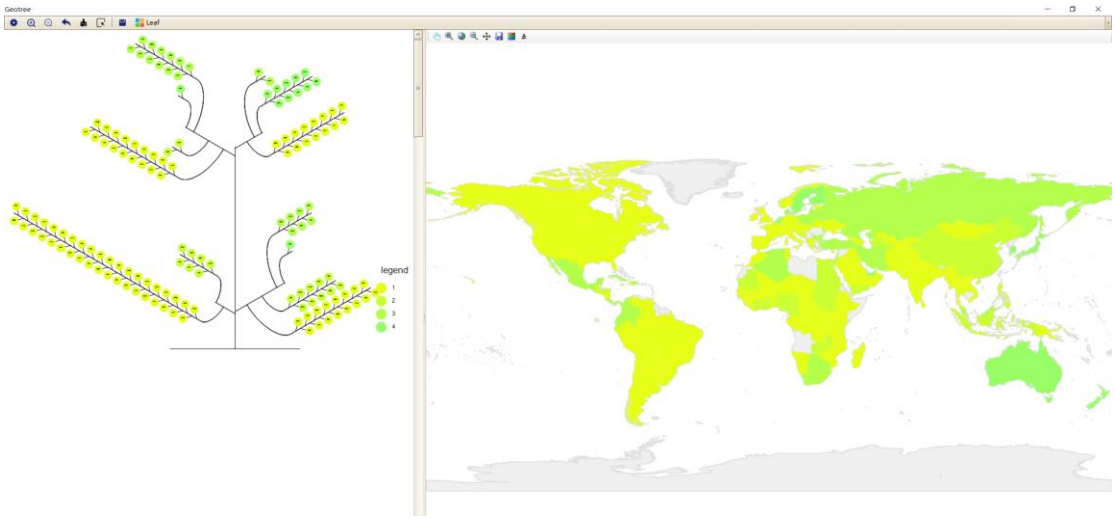


In the Parameters box, Set "Type" as the country type, ""Stage1" as the development stage of each country in 2005, leave "Stage2" empty. Set "Text" as the country code and "ID" as the unique identifier for each country. Click ... to import map data, and then set the Map ID that corresponds to the above ID. Otherwise, the generated tree cannot interact with the map.



Finally, click OK to generate Geotree. The left side of the result page is the

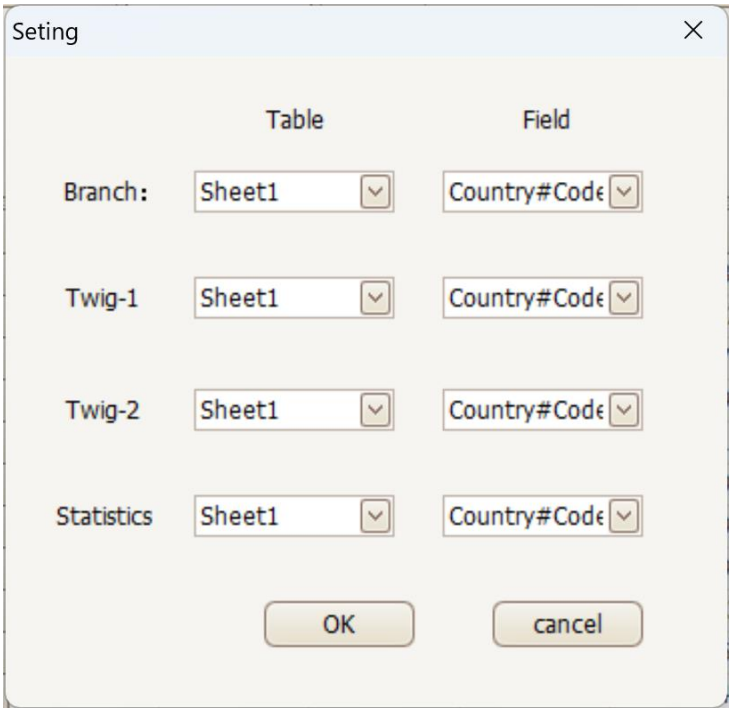
tree view, and the right side is the map view.



5.3 Development stage transitions

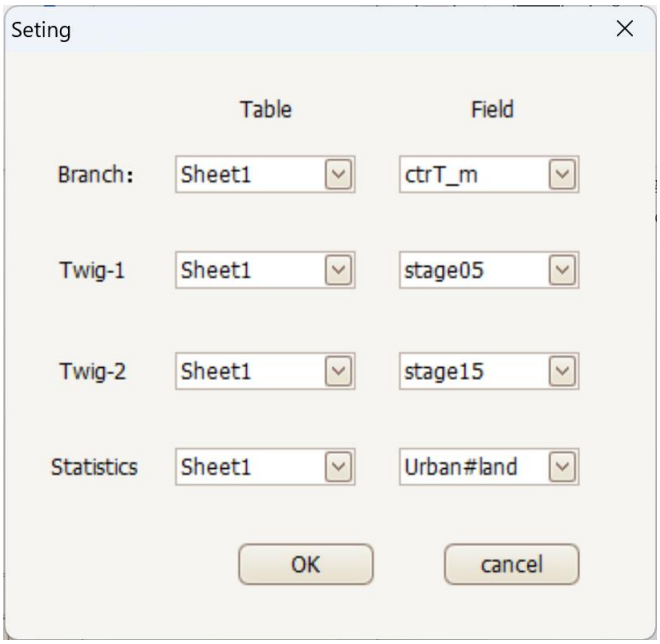
We use Markov Chain to study the transitions between different development stages of counties from 2005 to 2015.

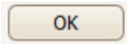
In the "Visualization" module, click "Markov Chain" to open the following window.

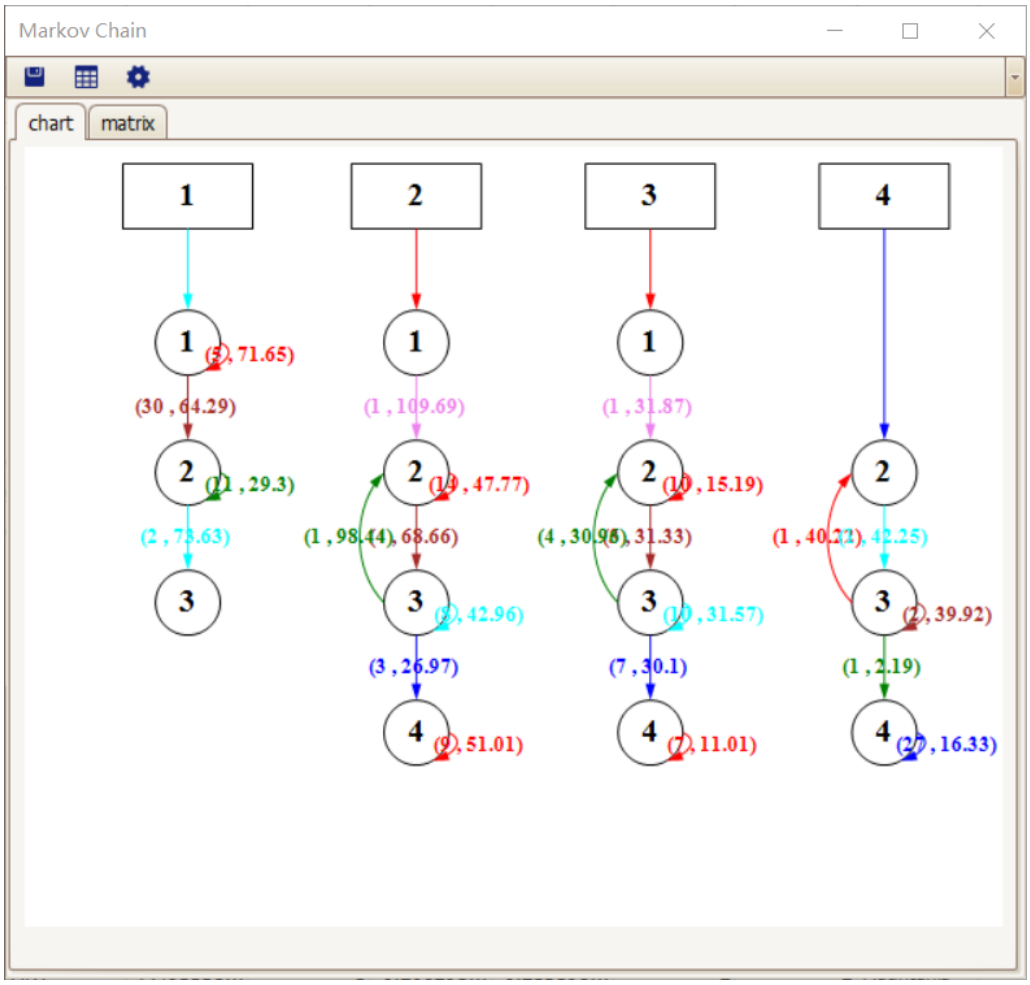


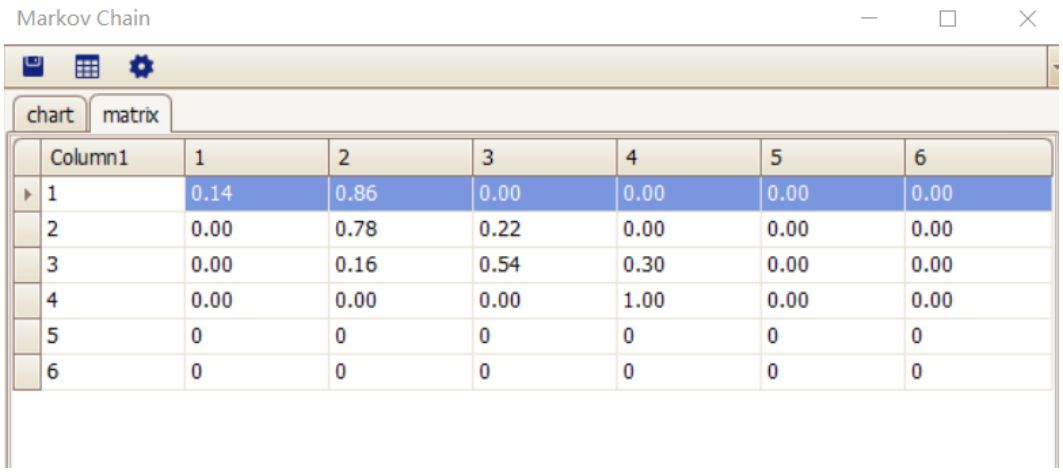
Set the parameters for Markov Chain: "Branch" is set as the country's industry structure type, "Twig-1" is set as the country development stage in 2005, "Twig-2" is

set as the country development stage in 2015, and "Statistics" is set as the urban land expansion rate.



After setting the above parameters, click "OK"  to obtain the Markov Chain results. The first figure shows the number of countries that underwent development stage transitions and the average urban land expansion rate for each type of country. The second figure is the probability transition matrix for urban economic development.





Column1	1	2	3	4	5	6
1	0.14	0.86	0.00	0.00	0.00	0.00
2	0.00	0.78	0.22	0.00	0.00	0.00
3	0.00	0.16	0.54	0.30	0.00	0.00
4	0.00	0.00	0.00	1.00	0.00	0.00
5	0	0	0	0	0	0
6	0	0	0	0	0	0

6 References

2012 Wang JF, Liu XH, Peng L, Chen HY, Driskell L, Zheng XY, 2012. Cities evolution tree and applications to predicting urban growth. **Population and Environment**, 33(2): 186-201.

2020 Wang Y, Wang JF, 2020. Modelling and prediction of global non-communicable diseases. **BMC Public Health**, 20(1): 822.

2022 Jing SQ, Wang JF, Xu CD, Yang JT, 2022. Tree-like evolution pathways of global urban land expansion. **Journal of Cleaner Production**, 378: 134562.

2023 Lei YH, Wang JF, Wang Y, Xu CD. 2023. Geographical evolutionary pathway of global tuberculosis incidence trends. **BMC Public Health**, 23:755

7 Developers

Academic Advisor: Jinfeng Wang (wangjf@Lreis.ac.cn)

Developers: Yanhui Lei (leiyh@Lreis.ac.cn);

Yuqing Feng (fengyq.19s@igsnrr.ac.cn);

Yuguang Wei (weiyugaung20@mails.ucas.ac.cn);

Yang Wang (wangyang622@lreis.ac.cn).

Website Design and Maintenance: Yuguang Wei (weiyugaung20@mails.ucas.ac.cn); Yuqing Feng (fengyq.19s@igsnrr.ac.cn);