Mix-cell Cellular Automata V1.0 (MCCA) User's Manual

Authors:

Xun Liang, Qingfeng Guan*, Keith C. Clarke, Guangzhao Chen, Song Guo, Yao Yao

* Corresponding author. E-mail address: guanqf@cug.edu.cn (Qingfeng Guan).

Consultation

If you have technical questions regarding MCCA software, please contact Dr. Xun Liang (liangxun@cug.edu.cn),

Homepage: https://www.researchgate.net/profile/Xun_Liang3

Contact info:

High-performance Spatial Computational Intelligence Lab (HPSCIL) (https://github.com/HPSCIL)

School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei 430078, China

For any possible research collaboration, please contact Prof. Qingfeng Guan (guanqf@cug.edu.cn)

Article DOI: https://doi.org/10.1016/j.landurbplan.2020.103960

Update date: 2020/10/16, 2022/10/16

Contents

Abstract

Cellular Automata (CA) for land use change model traditionally assume each cell to be one specific landuse type at each time step, ignoring the mixed land use structures that are often found in the land units. Mixed cells, composed of multiple land-use types with cover proportions, can better express the continuous changes of land use, and provide a new perspective for modeling the spatio-temporal dynamics of mixed land use structures.

Simulating land use change with mixed cells is challenging because mixed-cell CAs are fundamentally different from conventional CAs. We present the first attempt at developing a mixed-cell CA (MCCA). The structure of CA is re-designed based on the characteristics of mixed cell representatiing cell state, lattice, and neighborhood. The transition rules are automatically constructed by random-forest regression over historical data and a competition mechanism selects among multiple land use types at the sub-cell scale. Besides, evaluation methods for both simulation accuracy and similarity of land use structure are proposed for MCCA.

The MCCA represents a new breed of geospatial CA models for spatio-temporal dynamics of mixed land use structures. It provides a new approach to enable more dynamic mixed land use modeling to move away from the analysis of static patterns.

One of the advantages of MCCA models is the capability of simulating the quantitative and continuous changes in multi -components cells, while pure-cell CA models can only simulate the qualitative and discrete change of land use at the cell level. Therefore, MCCA models are able to simulate subtle changes in land use structures caused by minor variations of socio-economic, eco-environmental and political driving factors,which can provide a detailed perspective for understanding land use change process.

Also, the simulation results generated by MCCAmodels that contain the information for land use structure in each cell have the potential to help researchers more precisely evaluate the impacts of land use change on many environment variables, such as air quality, the urban heat island, landscape connectivity, net primary production (NPP) lost, ecological service value, energy consumption and more. MCCA models may better support space-time continuous analysis and the quantitative calculation of environment variables. In addition, MCCA models provide an enabling approach to the simulation of structural changes of mixed land use, as most previous studies focused on the measurement and static analysis of mixed land use structures and ignored their dynamic evolution. The MCCA model can simulate the gradual change in land use structures and help the researchers understand how the multiple driving factors interact to effect the future distribution of mixed land use.

MCCA was developed purely in the C++ language. The parallel technology of MCCA software is from High-performance Spatial Computational Intelligence Lab @ China University of Geosciences (Wuhan) (https://github.com/HPSCIL). The Random forest technique in MCCA is from a powerful opensource library called Alglib 3.9.2 (http://www.alglib.net/). The UI of the software is built using a famous opensource library Qt 5 (https://www.qt.io/download/). This UI provides a real-time display of dynamic changes of land use in simulation process. Moreover, the using of open source library GDAL 2.0.2 (http://www.gdal.org/) allows our model to directly read and write raster data (.tif, .img, .txt files) that includes geographical coordinate information.

This user's manual intends to provide a "quick start"for the users of MCCA. All of the necessary data and files for the tutorial have been provided as templates to format your own files later on.

1. Download and start MCCA

The latest version of MCCA software and the user's manual can be downloadin the link (https://github.com/HPSCIL/Mixed_Cell_Cellullar_Automata). Click the executable files "MCCA VX.0.exe" to start the software.

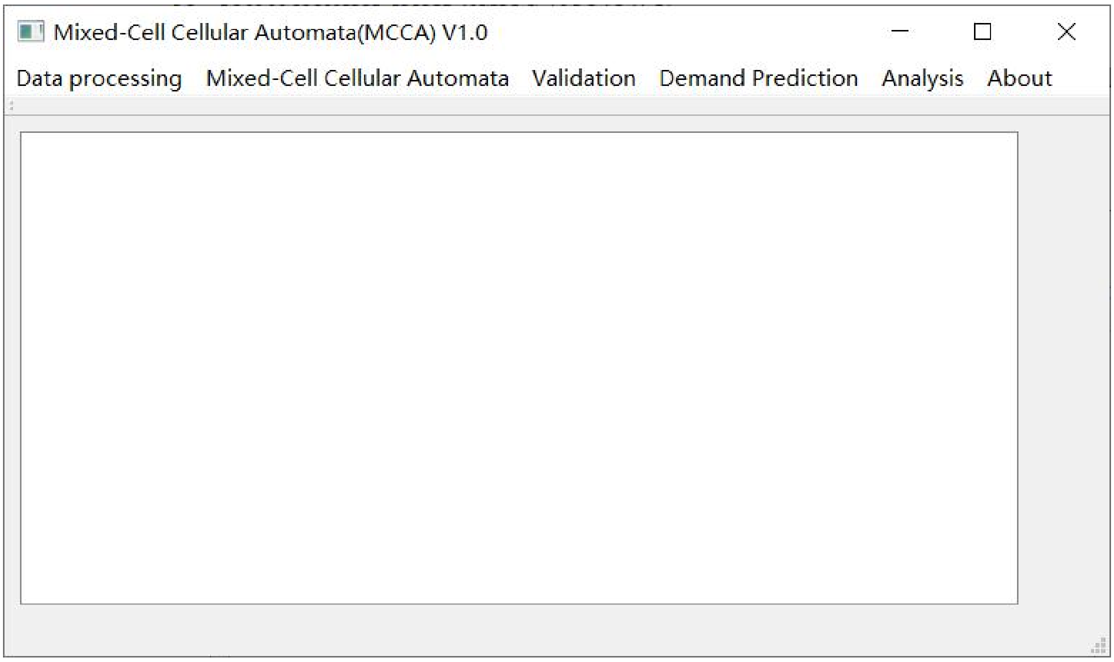


Fig. 1. The main interface of the MCCA model.

The main interface is a brief image browser which provides dynamic display function for the simulation process.

2. Example data description

Tab. 1 List of test data

Category	**Data**	**Filename**	**Description**
Mixed Land use data	Land use structure data	Xianning_2000_landuse_1~7.tif Xianning_2005_landuse_1~7.tif Xianning_2010_landuse_1~7.tif Xianning_2015_landuse_1~7.tif	Each land use structure data has 7 layers, each layer represents the cover proportion of a land use type: Layer1: Cropland, Layer2: Woodland, Layer3: Glassland, Layer4: Waterbody, Layer5: Urban land, Layer6: Rural Settlements, Layer7: Other
Constraint data	Land use constraint	OpenWater.tif	The cover proportion of open water
Socioeconomic data	Population GDP	population.tif gdp2010.tif	http://www.geodoi.ac.cn/WebCn/Default.aspx
	Proximity to Town center	dst_town.tif	World Urbanization Prospects: The 2014 Revision, CD-ROM Edition
	Proximity to highway Proximity to the arterial road Proximity to the primary road Proximity to the secondary road Proximity to the tertiary road	dst_highway.tif dst_arterialroad.tif dst_primaryroad.tif dst_secondaryroad.tif dst_tetriary.tif	OpenStreetMap (https://www.openstreetmap.org/)
	Proximity to high-speed railway stations	dst_highspeedrailwaystation.tif	http://lbsyun.baidu.com/
Climatic and environmental data	Soil type	soiltype.tif	HWSD v 1.2 (http://westdc.westgis.ac.cn/data/844010ba-d359-4020-bf76-2b58806f9205)

Annual Mean Temperature	annualTemperature.tif	WorldClim v2.0 (http://www.worldclim.org/)
Annual Precipitation	annualPrecipitation.tif	
DEM	DEM.tif	NASA SRTM1 v3.0
Slope	slope.tif	

The test dataset can be found in the ***TestData*** folder. Use this dataset to run an example model or compared it with your simulation results. In practical applications, more driving forces regarding human -natural effects can beinvolved. The study region in the example dataset is in Xianning city, located in the south of Hubei provincein central China. **Note that cover proportions of all land use components represent the land use structure of a cell.**

3. Some key notes in using MCCA model

1) The input land use structure data must be in a 'double' format.

2) In the step of mining of quantitative transition rules, the coordinate or projection system must be supported by GDAL. We recommend the World Geodetic System (WGS 84). This step does not require the input image data have the same number of rows and columns. An coordinates alignment mechanism is available.

3) In the step of simulating dynamics of land use structures, the input image data need to have the same number of rows and columns, including the land use pattern, restricted area, and change probability data.

4) We strongly recommend the users do not include non-English characters or sign in the file paths of the MCCA model, including the spaces. The file names and folder names must begin with a letter, not a number.

4. User interfaces of the MCCA model

The MCCA consists of two main parts, 1) Mining of quantitative transition rules and 2) Simulating dynamic of landuse structures. Users are allowed to start the corresponding module by selecting an option from the “Mix-cell Cellular Automata Model” button.

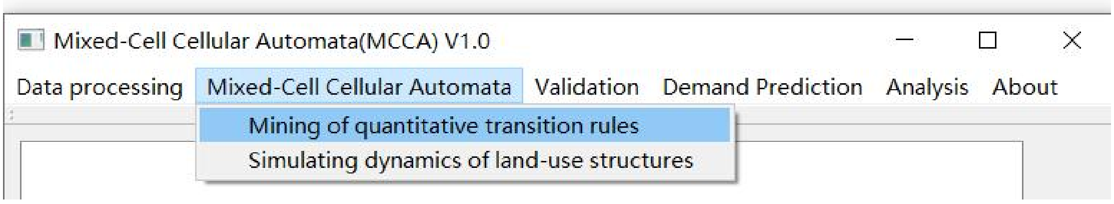


Fig. 2. Starting the ‘Mining of quantitative transition rules’ module.

4.1. Mining of quantitative transition rules

4.1.1. Start the module

Click the “Mining of quantitative transition rules” item to start the module. Then input the amount of land use type inthe simulation . amount of land use types in this experiment is 7. Note that the amount of land use can not be less than 2 (e.g., urban land and non-urban land).

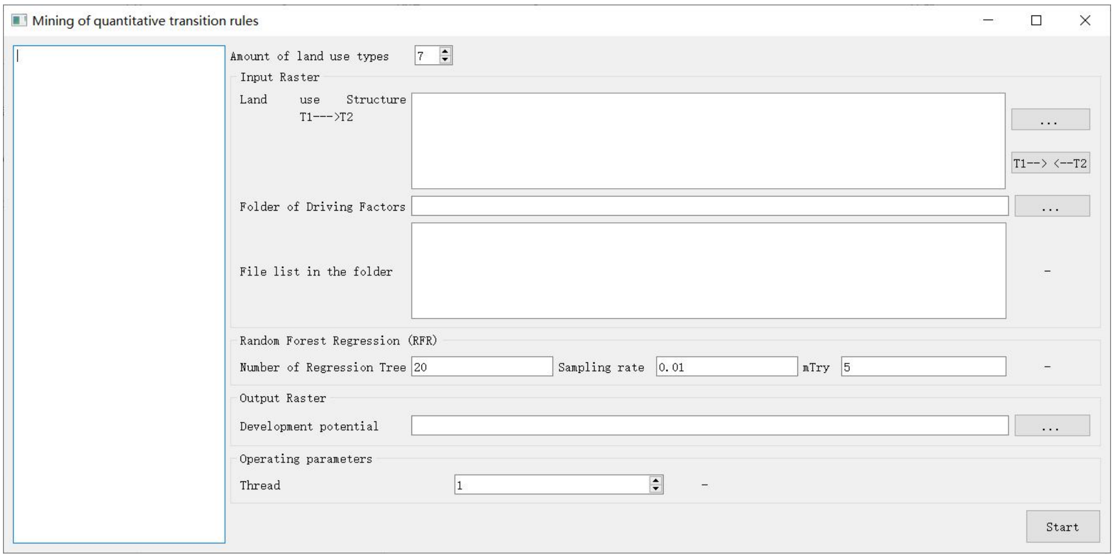


Fig. 3. UI of the ‘Mining of quantitative transition rules’ module.

4.1.2. Input land use data

Click the  button in the “**Land use Structure**” group box and select the two periods of land use data: ‘Xianning_2000_landuse’ and ‘Xianning_2015_landuse’ in the pop-up dialog. Note that in this experiment, 7 land use types mean 7 land use components in each land use structured data. You only need to select the first band of the two land use structure data (7 bands in total) in this step.

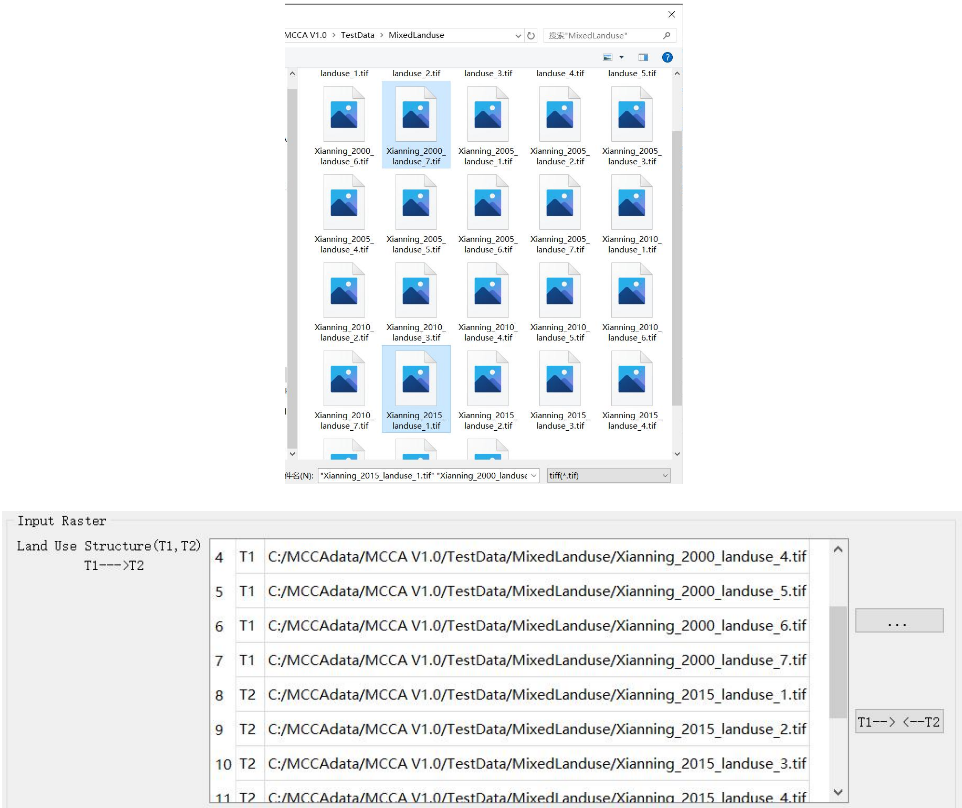


Fig. 4. Select two periods of land use structured data.

Please ensure that the latter land use structure data (e.g., T2, Xianning_2015_landuse) rank behind the first land use structure data (e.g., T1, Xianning_2000_landuse). The users can click the 

button to transpose their sort.

4.1.3. Input driving factors

Click the  button in the "**Folder of driving factors**" group box and select the folder of driving factors in the pop-up dialog. The MCCA software would automatically load the "tiff" File in the folder. Our example data provide 14 driving factors forsimulation.

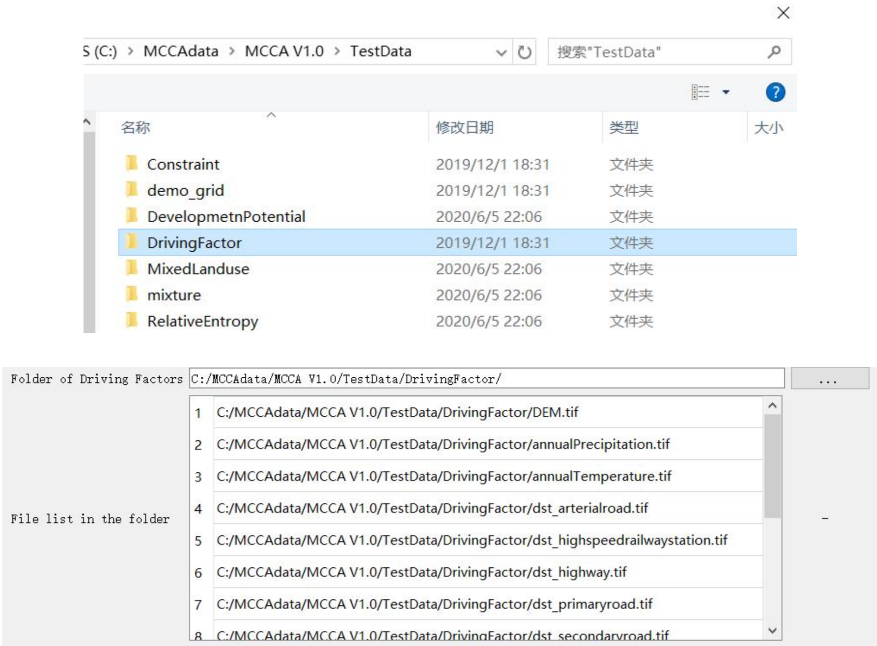


Fig. 5. Input all the driving factors by selecting one folder.

4.1.4. Setting the training parameters and output path

Users are allowed to set the parameters of **Random Forest Regression (RFR)**. The sampling rate is set to 0.1 by defaultwhich means about one percent of pixels will beselected for training. The number of regression tree is set to 100 in this experiment. mTry means the feature number used to train the RFR model. The max feature number is 14 in this experiment, equal to the number of driving factors. Then, click the  button in the “**Output Raster**” group box and input the save path of change probability in the pop-up dialog. The generated change probability is named “result_changeProbability.tif”. Users can increase the number of parallel internal threads to accelerate the running speed.

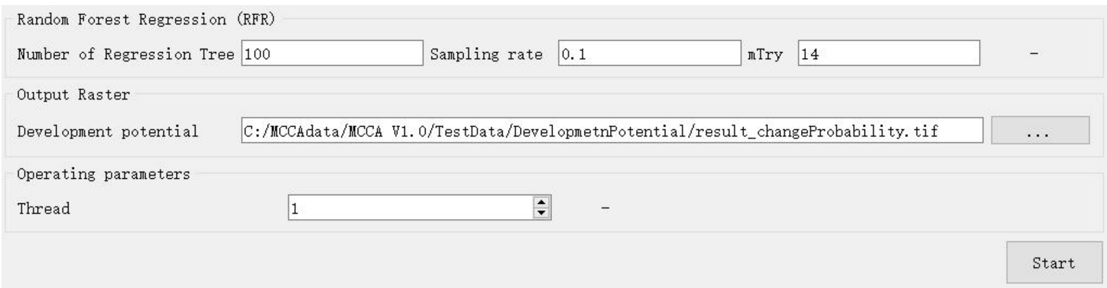


Fig. 6. Determining the training parameters and output path of the RFR.

4.1.5. Running program and examine the result

Once the setup mentioned above has completed, click the button 

for running the module. A message box will pop up when the model process run is complete as below.

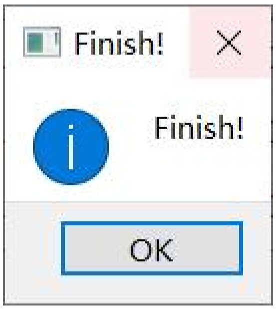


Fig. 7. A message that shows the program run successfully.

The text panelat the left of the user interface will output the training velocity and accuracies of the RFRs for each land use component. There will be three accuracies indicatorsas RMSE and OOB-RMSE for each land use type.

```
Type 0
make df-trees success = 1
RMSE =  0.017783
OOB RMSE =  0.0510151
Run time: 32.203

Type 1
make df-trees success = 1
RMSE =  inf
OOB RMSE =  inf
Run time: 40.594

Type 2
make df-trees success = 1
RMSE =  0.0102598
OOB RMSE =  0.0302431
Run time: 33.266

Type 3
make df-trees success = 1
RMSE =  0.0283541
OOB RMSE =  0.0823347
Run time: 32.75

Type 4
make df-trees success = 1
RMSE =  0.0318527
OOB RMSE =  0.0923667
Run time: 33.328

Type 5
make df-trees success = 1
RMSE =  0.00904057
OOB RMSE =  0.0261192
Run time: 32.64

Type 6
make df-trees success = 1
RMSE =  0.00938683
OOB RMSE =  0.0272855
Run time: 33.157
```

Fig. 8. Accuracy for each land use type.

Then the model will generate 7 sets of change probability maps for each land use type.

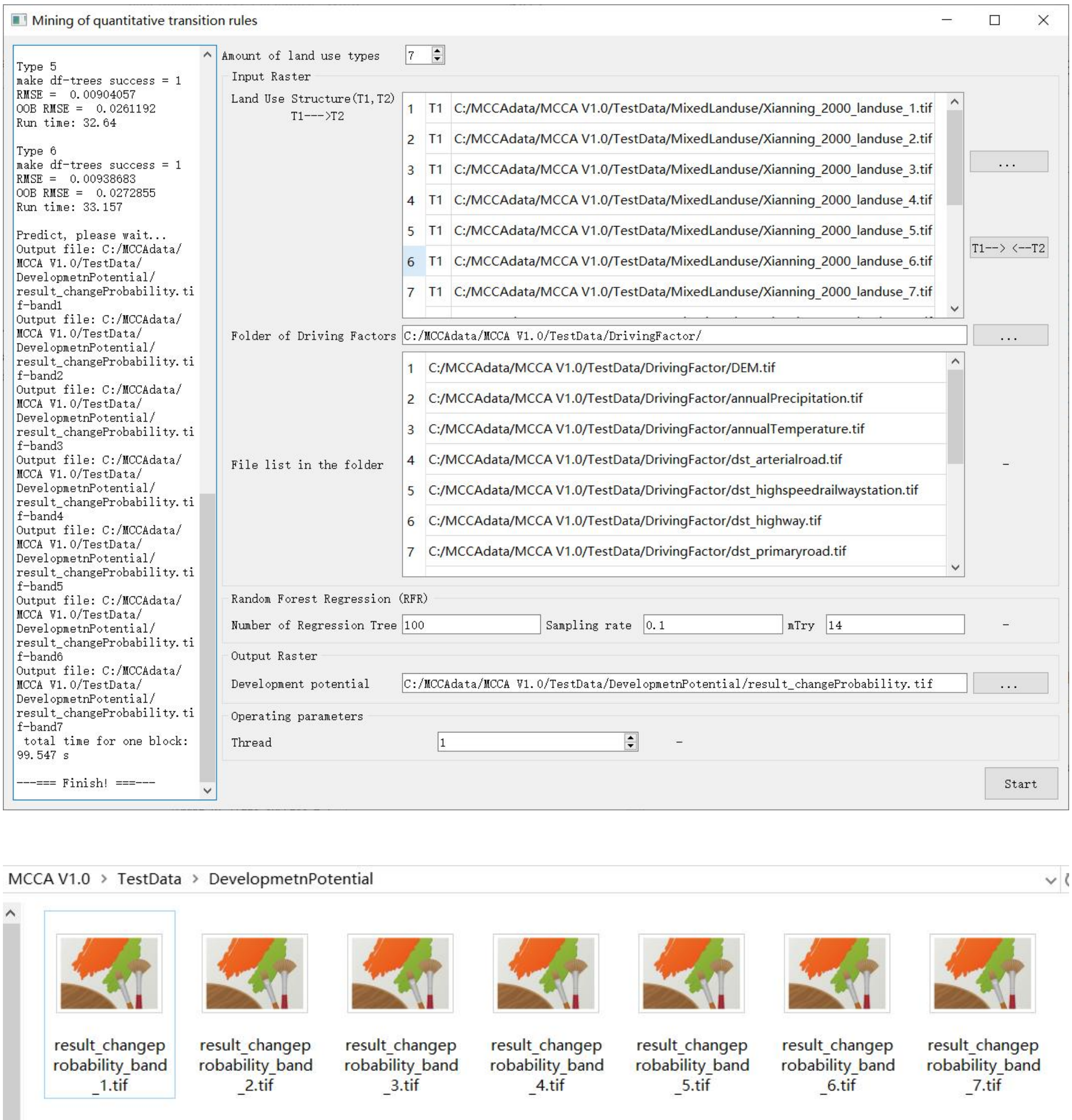


Fig. 9. Output of the ‘Mining of quantitative transition rules’ module.

4.2. Obtaining historical and future land use demands

Before the simulation, the historical land use amounts (e.g., 2000-2015) and future land use demands(e.g., 2015-2035) should be calculated for model validation and future projection respectively. However, the acquisition of historical land use amounts of mixed-cell land use data is not as easy as that of the pure-cell land use data. Therefore, we provide two tools based on linear regression and Markov Chain respectively, which can project future land use demands in a specific period while output the historical land use amounts.

4.2.1. Linear regression method

4.2.1.1. Starting the land use demand projection module

Future land use demands can be determined by many methods, such as expert experience, linear regression, Markov chains, system dynamics models, or an integrated assessment model. This step provides a linear regression method. This experiment projects future land use demands based on the historical data of 2000, 2005, 2010 and 2015. Click the “Linear regression” item to start the simulation module.

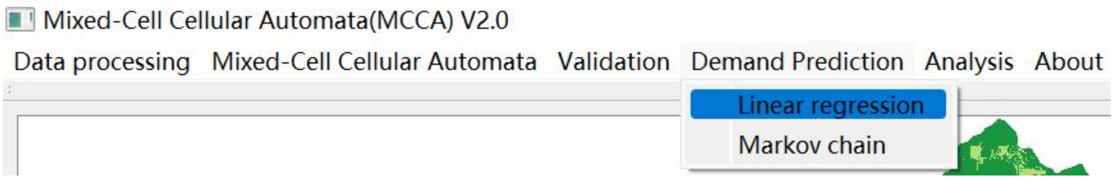


Fig. 10. Activating the land use demand projection module.

Then input the amount of land use type in the simulation, the amount of land use types in this experiment is 7. Set the predict year amount as 20. Then the module will predict the future land use demand from 2015 to 2035 (2015+20).

4.2.1.2. Input historical land use structure data

Click the  button, select the historical land use structure data in 2000, 2005, 2010 and 2015. Users only need to select the first band of the four land use structure data in this step. The MCCA software would automatically load the other bands of these land use structured data.

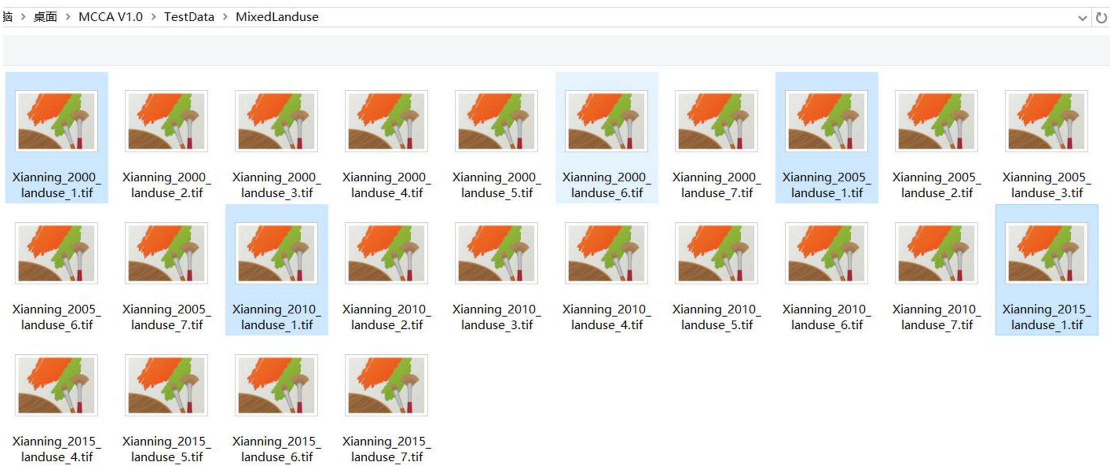


Fig. 11. Select the historical land use structured data.

What is more, the program can automatically read the year on the filename. Uses can also manually input the year to the 'Year' column on the chart of data.

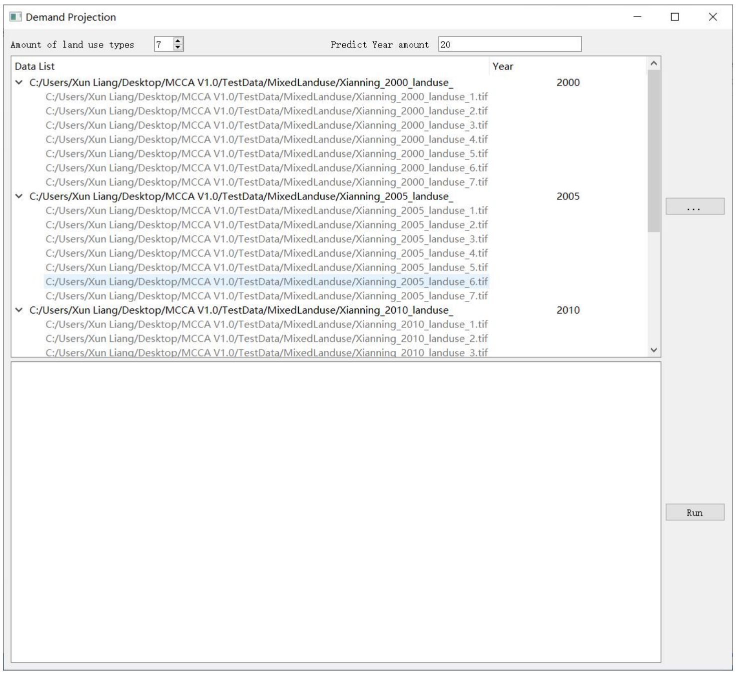


Fig. 12. Inputting the historical land use structure data to the Demand projection module.

4.2.1.3. Running linear regression and predicting future demands

Click the 

button to start calculating future land use demand. A message box will pop up when the program completes the calculation. The results will be showed on the interface and saved in a file named "PredictDemand.csv" in the folder 'Parameterfile'. Noted that this module also can generate historical land use cover proportions for model validation (Section 4.4).

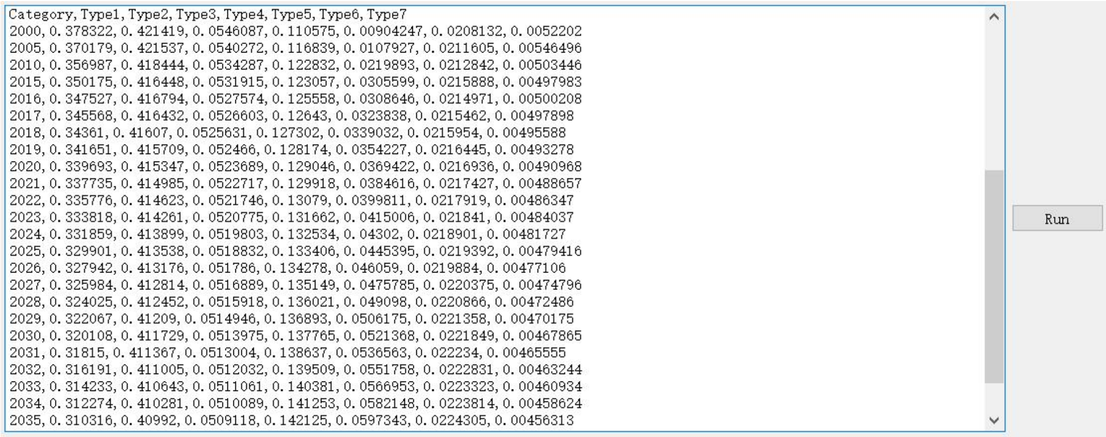


Fig. 13. Running the linear regression and predicting future land use demands.

4.2.2. Markov chain method

4.2.2.1. Starting the land use demand projection module

This step provides a Markov chain method to predict future land use demand based on historical data in 2010 and 2015. Click the “Markov chain” item to start the prediction module.

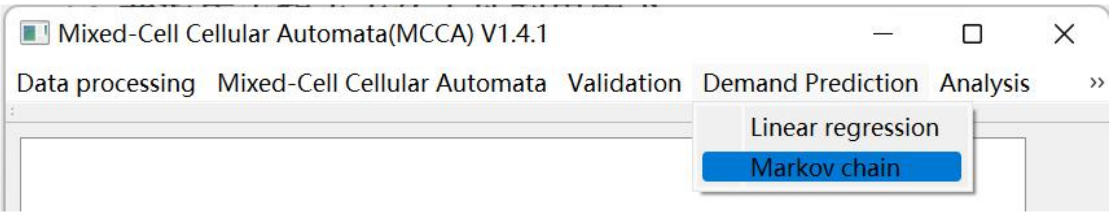


Fig. 14. Activating the Markov chain method

4.2.2.2. Input historical land use structure data

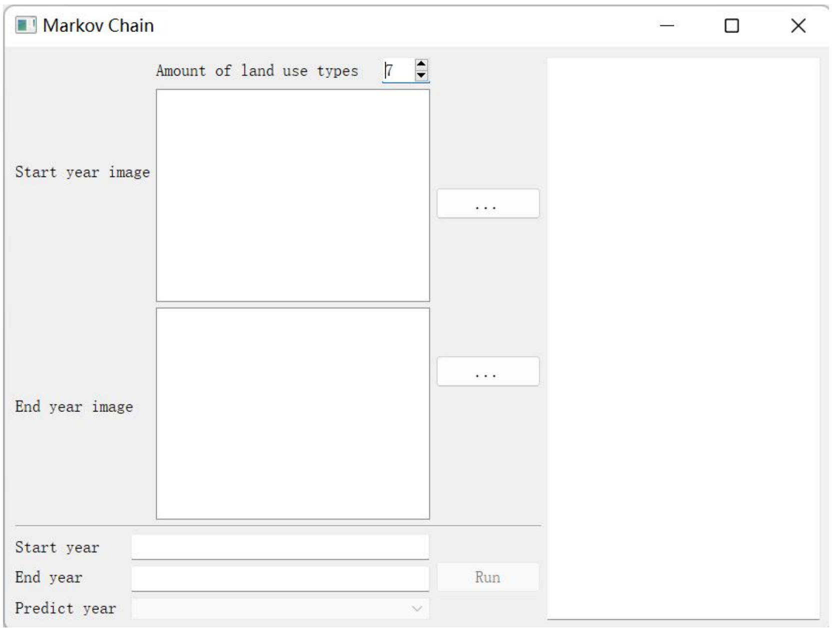


Fig. 15. Select the historical land use structured data.

Then input the amount of land use type in the simulation, the amount of land use types in this experiment is 7. Click the  button in the "**Start year image**" row to input the land use data of start year. Then click the  button in the "**End year image**" row to input the land use data of end year. Users only need to select the first band of the two land use structure data in this step. The MCCA software would automatically load the other bands of these land use structured data. Then fill in the text plain below and activated the "**Predict year**" pull-down menus. Users can choose the future year they want to predict such as 2035 year.

4.2.2.3. <u>Predicting future land demands with Markov chain</u>

Click the 

button to start calculating future land use demand. A message box will pop up when the program completes the calculation. The results will be showed on the interface and saved in a file named "MakovChain.csv" in the folder 'Parameterfile'. Noted that this module also can generate historical land use cover proportions for model validation (Section 4.4).

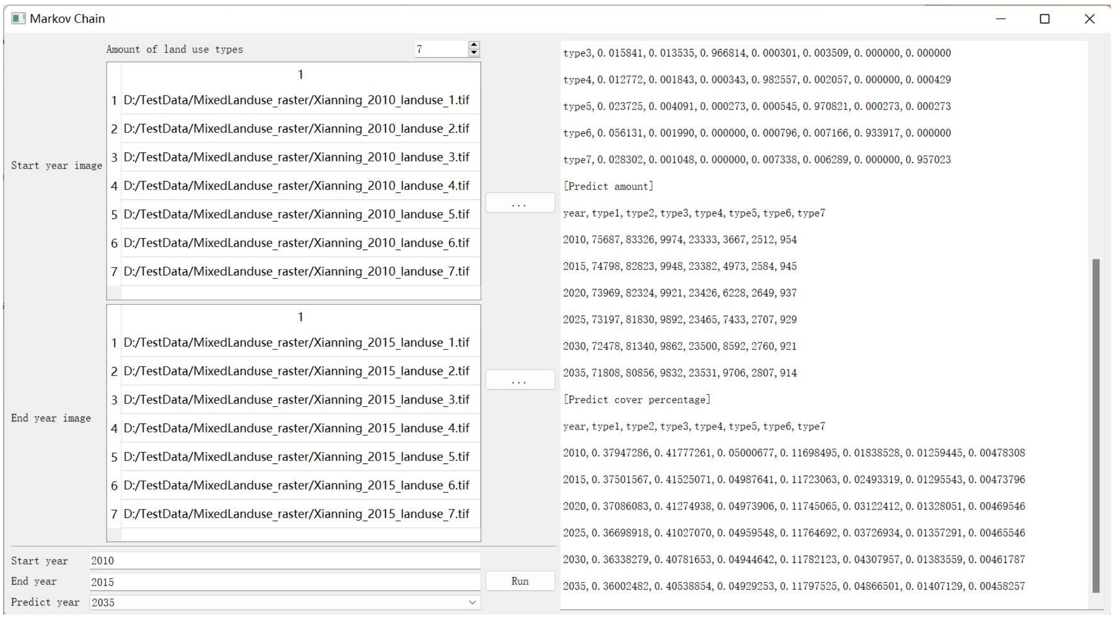


Fig. 16. Predict future land use amount with Markov chain

4.3. Simulating dynamics of land use structures

4.3.1. Start the module

Click the "**Simulating dynamics of land use structures**" item to start the simulation module. Then input the amount of land use type in the simulation. The **amount of land use types** in this experiment is 7. The users can increase the number of parallel internal **threads** to accelerate the running speed. The **neighborhood** effect of MCCA is set to 3 in this experiment.

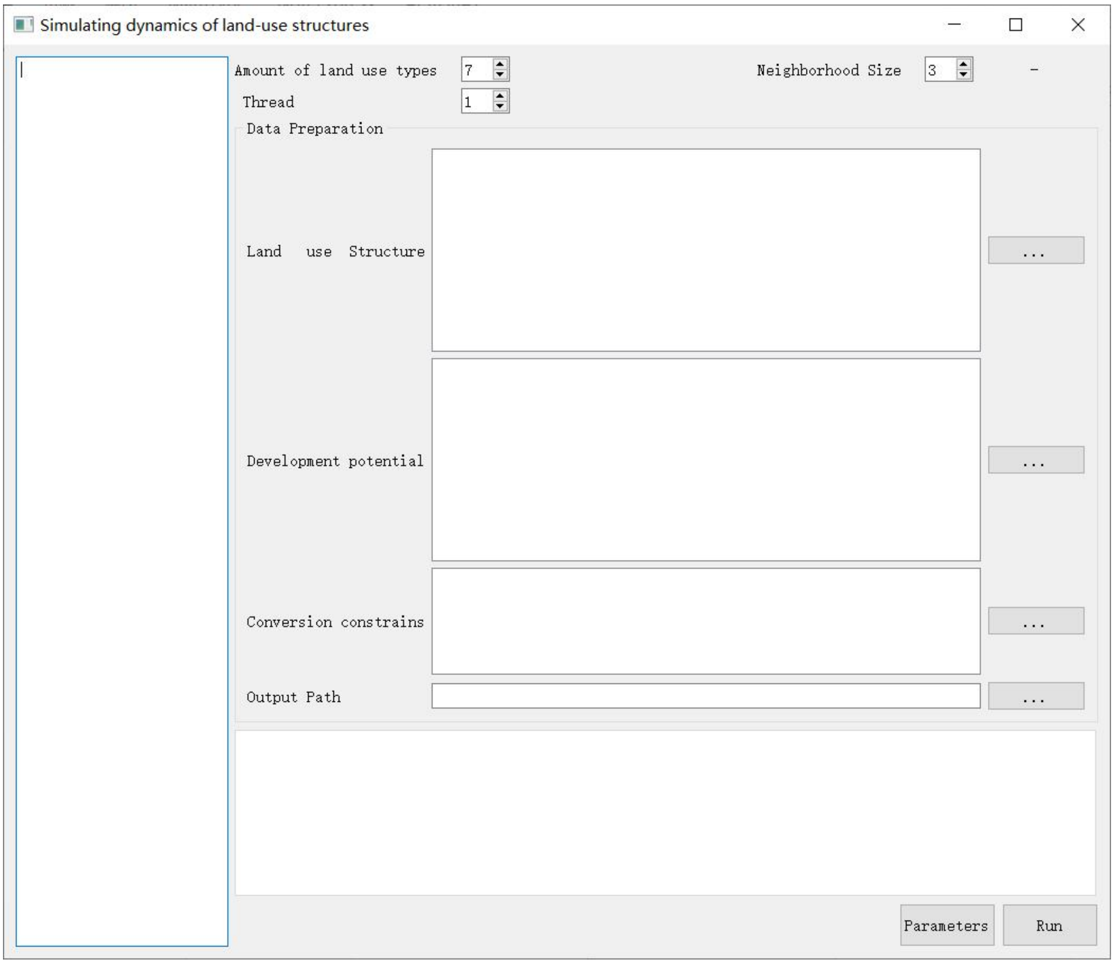


Fig. 17. The UI of ‘Simulating dynamics of land use structures’ module.

4.3.2. Input land use data

Click the  button in the “**Land Use Structure**” group box and select the start year of land use data “Xianning_2000_landuse_1.tif” in the pop-up dialog. The model will automatically load all bands of land use structured data.

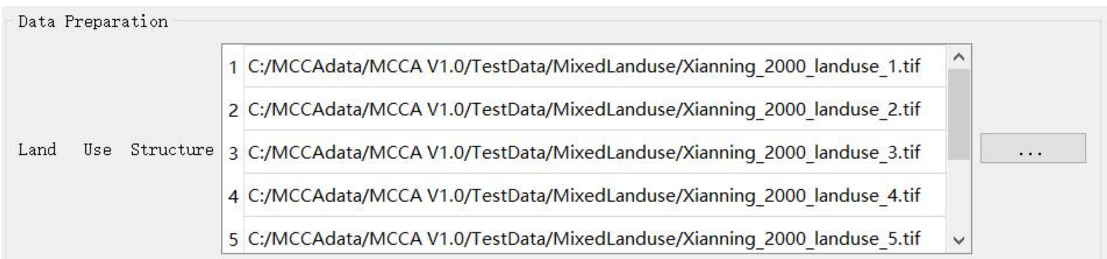


Fig. 18. Select the land use structure data at start year.

4.3.3. Input change probability data

Input the change probability file output from the ‘Mining of quantitative transition rules’ module in the “**Development Potential**” group box.

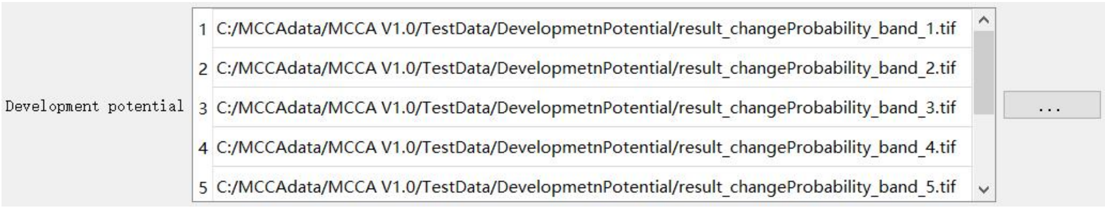


Fig. 19. Select the change probability data.

4.3.4. The constraint of open water (Spatial policies)

We assume that open water (a sub-category of waterbody) is not allowed to be converted to other land use components. Therefore, a distribution map of the cover proportion of open water was used to provide the minimum quantity of waterbody of each cell, which means that the cover proportion of waterbody within each cell is not less that the cover proportion of open water. Meanwhile, the mutual conversions between other pairs of land use components are allowed in these cells.

Then we input the land use cover proportion of the open water layer in the “**Conversion constraints**” group box. The second parameter is used to specify the land use type that the open water is about to restrain. In this experiment, the open water is used to constrain change of water body, and the code of water body is 4. Thus the parameter is set to 4.

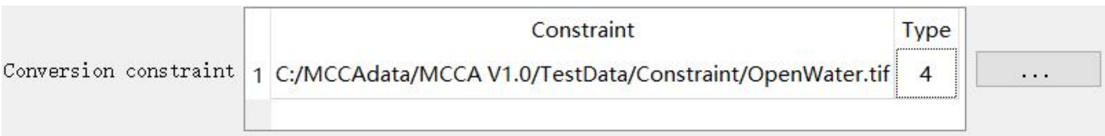


Fig. 20. Select the change probability data.

4.3.5. Set the path of saving simulation result

Input the save path of the simulation result in the “**Output path**” group box.

Fig. 21. Select the output path.

4.3.6. Setting the simulation parameters

Simulation parameters include the land use demands, transition matrix and step size. The land use demands need to be firstly set according to the actual situation of the study region.Users need to firstly use the external model to determine the land use demands under future scenarios. **For our example run, the time span of the land use change simulation is from 2000 to 2015, so the land use demands are set to be the same as the actual area of various land use types in 2015.** <u>We have obtained the historical land demand and predict future land demand in section 4.2.1 or 4.2.2.</u>

```
Category, Type1, Type2, Type3, Type4, Type5, Type6, Type7
2000, 0. 378322, 0. 421419, 0. 0546087, 0. 110575, 0. 00904247, 0. 0208132, 0. 0052202
2005, 0. 370179, 0. 421537, 0. 0540272, 0. 116839, 0. 0107927, 0. 0211605, 0. 00546496
2010, 0. 356987, 0. 418444, 0. 0534287, 0. 122832, 0. 0219893, 0. 0212842, 0. 00503446
2015, 0. 350175, 0. 416448, 0. 0531915, 0. 123057, 0. 0305599, 0. 0215888, 0. 00497983
2016, 0. 347527, 0. 416794, 0. 0527574, 0. 125558, 0. 0308646, 0. 0214971, 0. 00500208
2017, 0. 345568, 0. 416432, 0. 0526603, 0. 12643, 0. 0323838, 0. 0215462, 0. 00497898
2018, 0. 34361, 0. 41607, 0. 0525631, 0. 127302, 0. 0339032, 0. 0215954, 0. 00495588
2019, 0. 341651, 0. 415709, 0. 052466, 0. 128174, 0. 0354227, 0. 0216445, 0. 00493278
2020, 0. 339693, 0. 415347, 0. 0523689, 0. 129046, 0. 0369422, 0. 0216936, 0. 00490968
2021, 0. 337735, 0. 414985, 0. 0522717, 0. 129918, 0. 0384616, 0. 0217427, 0. 00488657
2022, 0. 335776, 0. 414623, 0. 0521746, 0. 13079, 0. 0399811, 0. 0217919, 0. 00486347
2023, 0. 333818, 0. 414261, 0. 0520775, 0. 131662, 0. 0415006, 0. 021841, 0. 00484037
2024, 0. 331859, 0. 413899, 0. 0519803, 0. 132534, 0. 04302, 0. 0218901, 0. 00481727
2025, 0. 329901, 0. 413538, 0. 0518832, 0. 133406, 0. 0445395, 0. 0219392, 0. 00479416
2026, 0. 327942, 0. 413176, 0. 051786, 0. 134278, 0. 046059, 0. 0219884, 0. 00477106
2027, 0. 325984, 0. 412814, 0. 0516889, 0. 135149, 0. 0475785, 0. 0220375, 0. 00474796
2028, 0. 324025, 0. 412452, 0. 0515918, 0. 136021, 0. 049098, 0. 0220866, 0. 00472486
2029, 0. 322067, 0. 41209, 0. 0514946, 0. 136893, 0. 0506175, 0. 0221358, 0. 00470175
2030, 0. 320108, 0. 411729, 0. 0513975, 0. 137765, 0. 0521368, 0. 0221849, 0. 00467865
2031, 0. 31815, 0. 411367, 0. 0513004, 0. 138637, 0. 0536563, 0. 022234, 0. 00465555
2032, 0. 316191, 0. 411005, 0. 0512032, 0. 139509, 0. 0551758, 0. 0222831, 0. 00463244
2033, 0. 314233, 0. 410643, 0. 0511061, 0. 140381, 0. 0566953, 0. 0223323, 0. 00460934
2034, 0. 312274, 0. 410281, 0. 0510089, 0. 141253, 0. 0582148, 0. 0223814, 0. 00458624
2035, 0. 310316, 0. 40992, 0. 0509118, 0. 142125, 0. 0597343, 0. 0224305, 0. 00456313
```

Fig. 22Historical land use amounts in 2015.

The model will stop when the allocated area equals the demanded area for all land use types. In the tab widget, the second row "Future Amounts" in the "Land Demand" page is for inputting the future land use demand of each land use type. The initial pixel number will be automatically loaded after running the program. You need to input the cover proportion of each land use type in the predicted year as shown in figure 19 (e.g., the 2015 land use amount).

Step Size | Transtion Matrix | Land Demands

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
Start Amounts	0	0	0	0	0	0	0
Future Amounts 1	0.350175	0.416448	0.0531915	0.123057	0.0305599	0.0215888	0.00497983

Fig. 23. Input the targeted land use demand.

Then, switch the tab widget to the "**Transition Matrix**" page. The columns of this matrix indicate the current land use types, and the rows indicate the future land use types. A value of 1 means the conversion is allowed while a value of 0 indicates that the conversion is not possible. A **Transition Matrix** is a collection of a series of prior knowledge of experts about the study area. An example of a cost matrix is shown below:

Step Size | Transtion Matrix | Land Demands

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
Type 1	1	0	1	1	1	1	0
Type 2	1	1	1	0	1	1	0
Type 3	1	1	1	0	1	1	0
Type 4	1	0	0	1	1	1	0
Type 5	0	0	0	0	1	0	0
Type 6	0	0	0	0	1	1	0
Type 7	1	1	1	1	1	1	1

Fig. 24. Transition Matrix.

And in the "**Step Size**" page, users need to set the Step Size of different land use types, because the neighborhood effects may be different for each land use types. The step size for each land use type is determined according to expert knowledge and a series of model tests, ranging from 0-1. At this example, we set the parameter as below:

Step Size | Transtion Matrix | Land Demands

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7
Step Size	1	1	1	1	1	1	1

Fig. 25. Step Size, the value of 1 for each land use type is recommended.

4.3.7. Running the multiple CA allocation model

Finally, click the button 

to start the simulation. The information of each iteration is shown on the left of the interface of 'Simulating dynamics of land use structures' module. Also, the MCCA is able to show the process of land use change dynamically, including the spatial change, value change and change curve of each land use type on the main interface.

The visualization panel allowes users to zoom in and observe the land use dynamics of local regions. Users can checkany land use type on the visualization panel by activating the corresponding radio buttons on the right. The horizontal axis of the dynamic chart is the number of iteration, and the vertical axis represents the product of the cover proportion and amount of cell.

The model will stop and save the simulation result in 2015 when the allocated area equals the demanded area for all land use types. The MCCA software will save all the parameters to two configuration files: “SCAparameters.tmp” in the ‘Parameterfile’ folder. The module will load the parameters that are recorded in the configuration files when clicking the 

button, which is convenient for users to repeat their experiments.

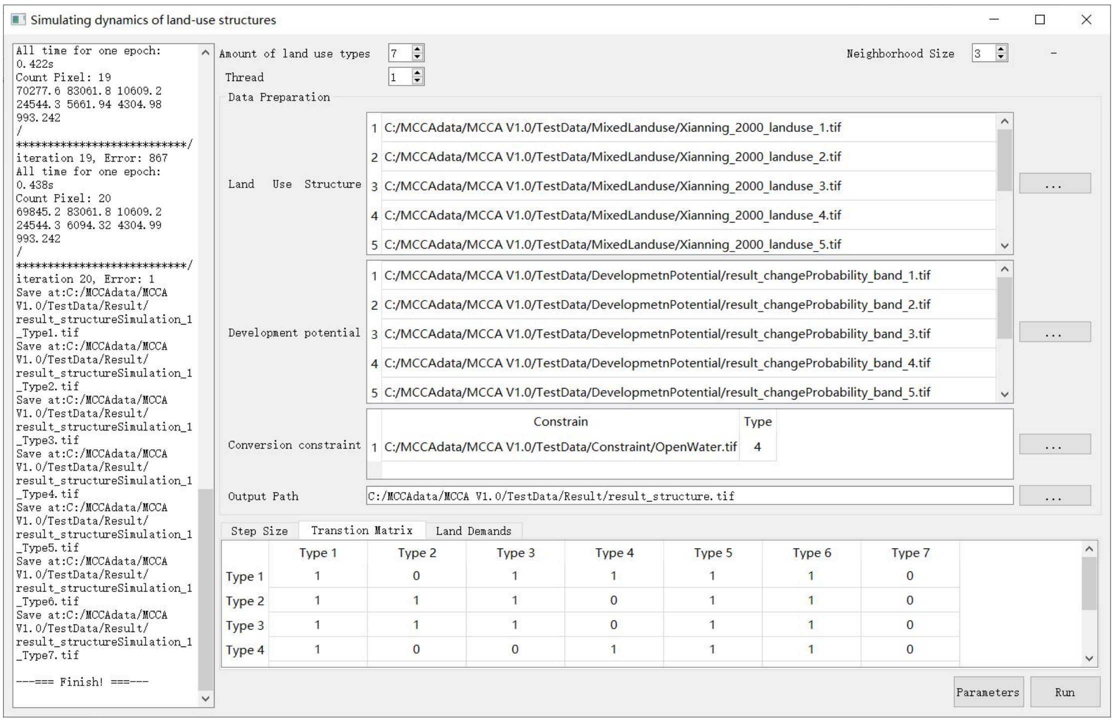


Fig. 26. The dynamic information of each iteration is show on left.

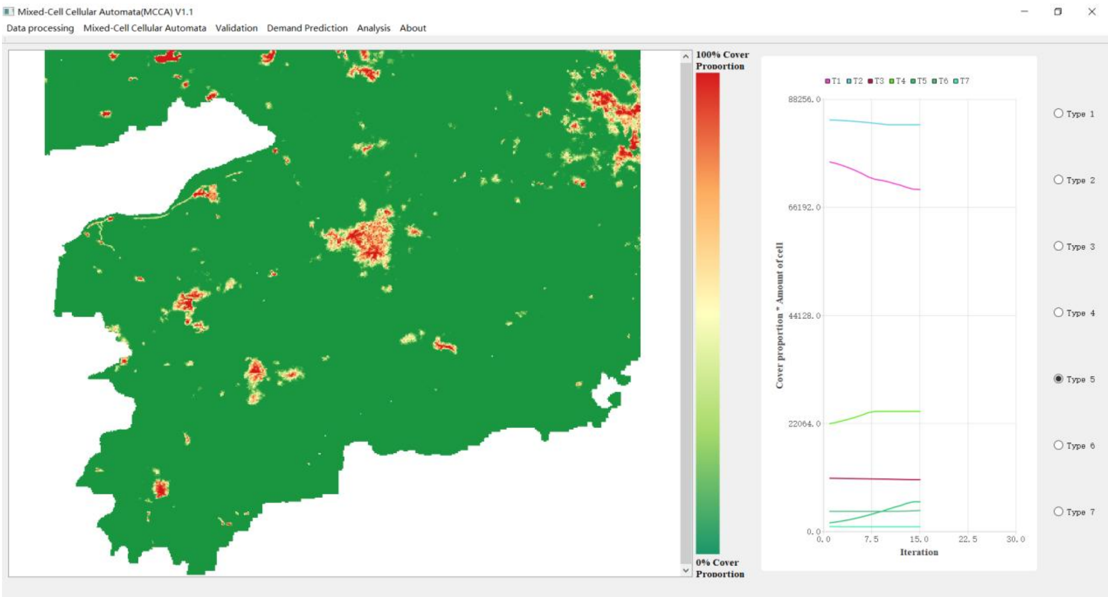


Fig. 27. The dynamic display of the simulation process on the main interface.

4.4. Accuracy validation of simulation

MCCA provides a sub-pixel confusion matrix (SCM) statistic tool for calculating the kappa coefficient for land use structured data. The conventional confusion matrix is not available for evaluating the simulation result of mixed-cell land use data. Then a Relative Entropy (RE) statistic tool for measuring the similarity between simulation land use structure and actual land use structure. Moreover, a mixed-cell figure of merit (mcFoM) is proposed to validate the accuracy of the change cells simulated by the MCCA model.

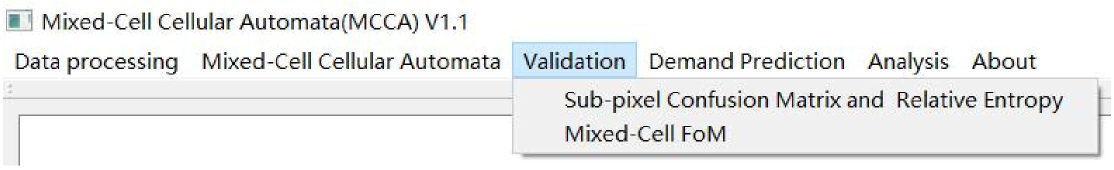


Fig. 28. Activate the validation module.

4.4.1. Load data

Input the amount of land use type in the simulation, the amount of land use types in this experiment is 7. Click the  button in the "**Ground Truth**" row, select the start year of land use data "Xianning_2015_landuse_1.tif" in the pop-up dialog.

The program will automatically load all the 7 bands of the land use structured data. Then click the  button in the "**Simulation Result**" row to input the simulated land use structure data 'result_structureSimulation_1_Type_1.tif'.

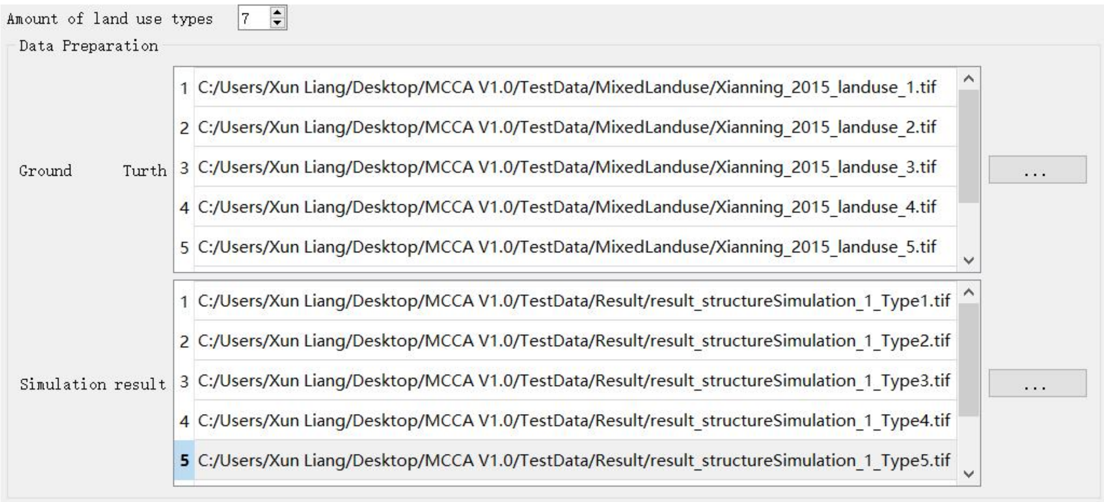


Fig. 29. Input the actual and simulation land use structured data.

4.4.2. SCM statistic tool

Set the sampling rate to 5%. Click the  button to start calculating the Kappa coefficient. A message box will pop up when calculation completes. The results include the Kappa coefficient, overall accuracy and sub-pixel confusion matrix they will be saved in a file named "SCM.csv" in the folder 'Parameterfile'.

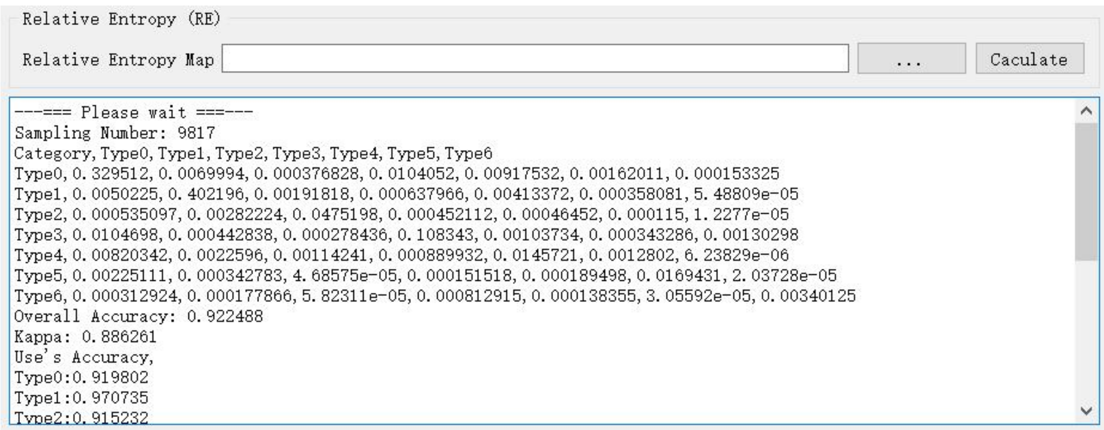


Fig. 30. Sub-pixel confusion matrix, Kappa coefficient, and overall accuracy.

4.4.3. Relative Entropy statistic tool

We computed the Relative Entropy (RE) as an indicator to evaluate the similarity of land use structure, which can represent the information decay of the simulation process. When the actual and simulated land use structures of mixed cell i are identical, RE is 0. A larger RE value indicates a more significant difference between the actual and simulated land use structure. The mean RE of the whole region will be calculated sequentlyas the measurement of similarity of land use structure at the regional level.

Click the 

button in the “**Relative Entropy Map**” row, then choose a path for the calculation of relative entropy map. Click the 

button to start calculating RE. A message box will pop up when the program completes the calculation. The similarity between simulation structure and the actual structure of one cell are shown in the image “RE.tif”.

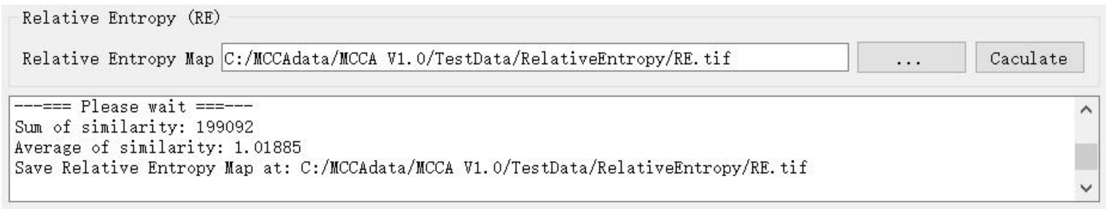


Fig. 31. Output the mean Relative Entropy of the region and Relative Entropy on each cell.

4.4.4. Mixed-cell Figure of Merit

Click the “Mixed cell FoM” item to start the simulation module. Input the amount of land use type in the simulation, the amount of land use types in this experiment is 7. Click the 

button in the “**Ground Truth**” row， select the start year of land use data “Xianning_2015_landuse_1.tif” in the pop-up dialog. The program will automatically load all the 7 bands of the land use structured data. Then click the

 button in the “**Simulation Result**” row to input the simulated land use structure data ‘result_structureSimulation_1_Type_1.tif’. Click the  button in the “ **Initial map** ” row to input the simulated land use structure data ‘Xianning_2015_landuse_1.tif’. Click the 

button to start calculating mcFoM indicator. A message box will pop up when the program completes the calculation.

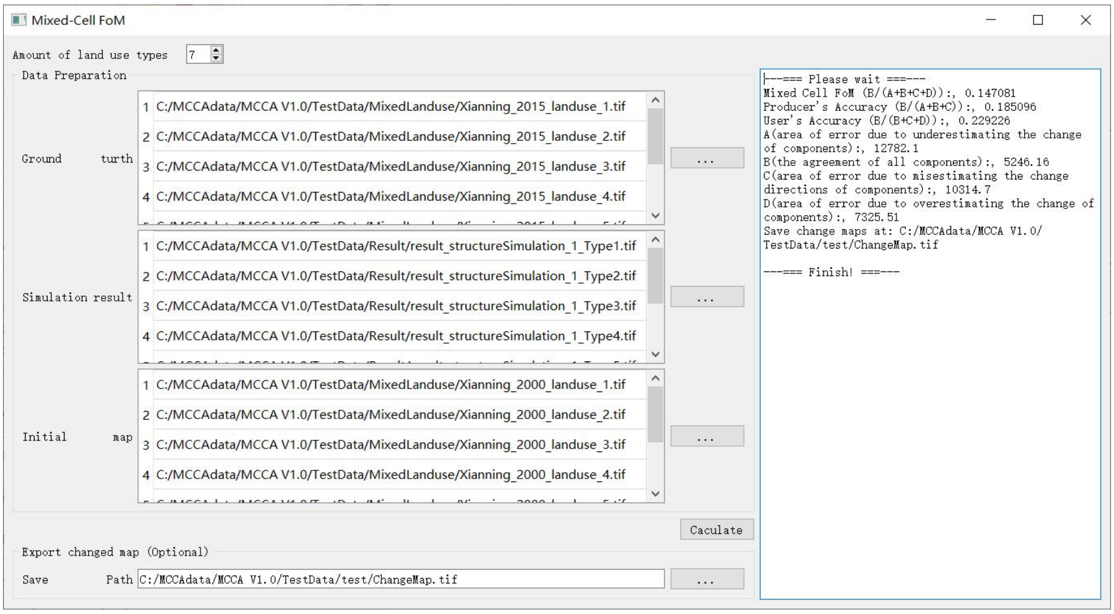
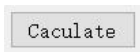


Fig. 32. Calculate the mcFoM indicator and the simulated and actual changed maps of all land use components.

The users can also export both the maps of simulated changed cells and actual changed cells. Click the  button in the “**Export changed Map**” row, then choose a path for the calculation of relative entropy map. Click the 

button to start calculation. Then the simulated and actual changed maps of all land use components can be saved.

4.5. Simulating future land use structure from 2015 to 2035

4.5.1. Obtain change probability form 2010 to 2015

Repeat the same procedures in section 4.1 using historical land use structure data: ‘Xianning_2010_landuse’ and ‘Xianning_2015_landuse’ in the ‘MixedLanduse’ folder. Then obtain the change probability form 2010 to 2015.

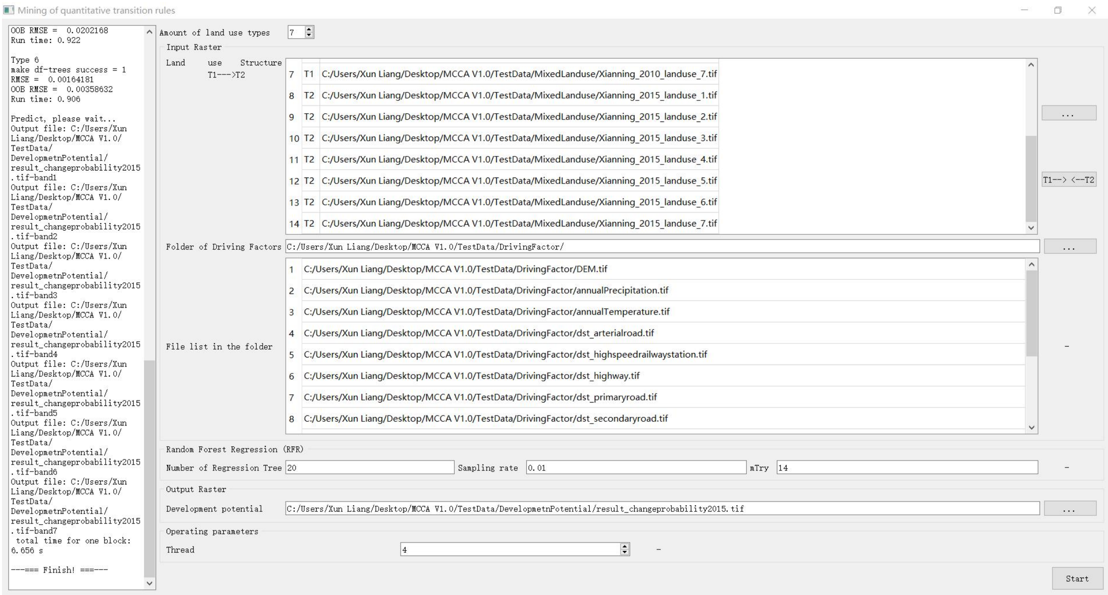


Fig. 33. Running the ‘Mining of quantitative transition rules’ module for future simulation.

4.5.2. Simulating land use structure change from 2015 to 2035

Repeat the same procedures in section 4.2 using historical land use structure data: ‘Xianning_2015_landuse’ in the ‘MixedLanduse’ folder and the change probability exported by ‘Mining of quantitative transition rules’ module in section 4.1.5. Import the predicted land use demands in 2035, which have been obtained in section 4.2.1. or 4.2.2. Other simulation parameters that have been well-calibrated remain the same (section 4.3).

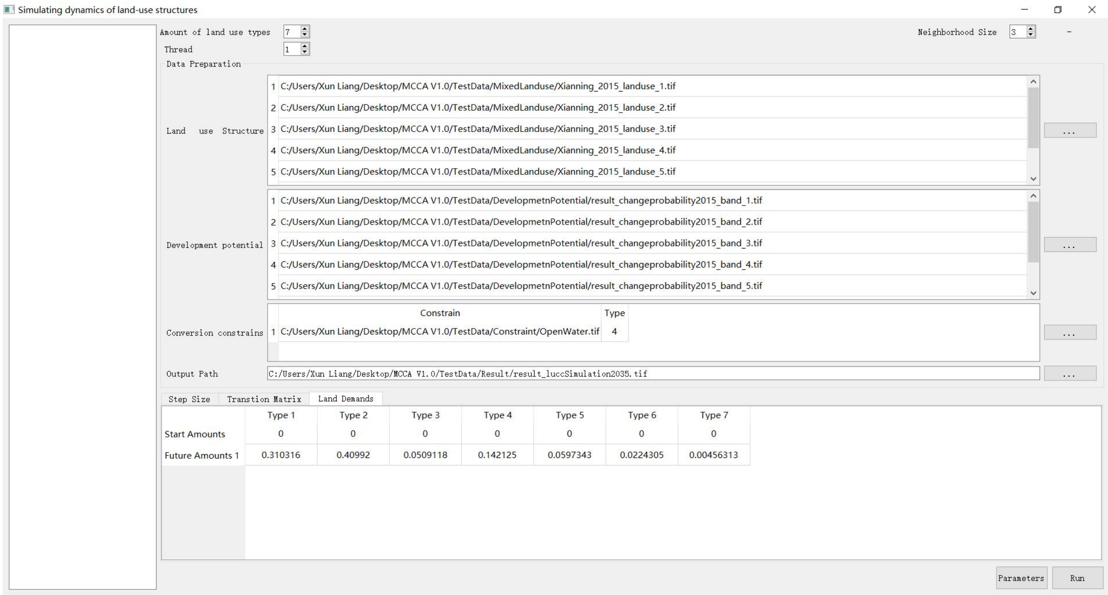


Fig. 34. Parameter setting for simulation from 2015 to 2035.

The MCCA software will save all the parameters to two configuration files: “SCAparameters.tmp” in the ‘Parameterfile’ folder. The module will load the parameters that are recorded in the configuration files when clicking the 

button, which is convenient for the users to repeat their experiments. Finally, click the 

button to generate the simulation result of the land use structure in 2035.

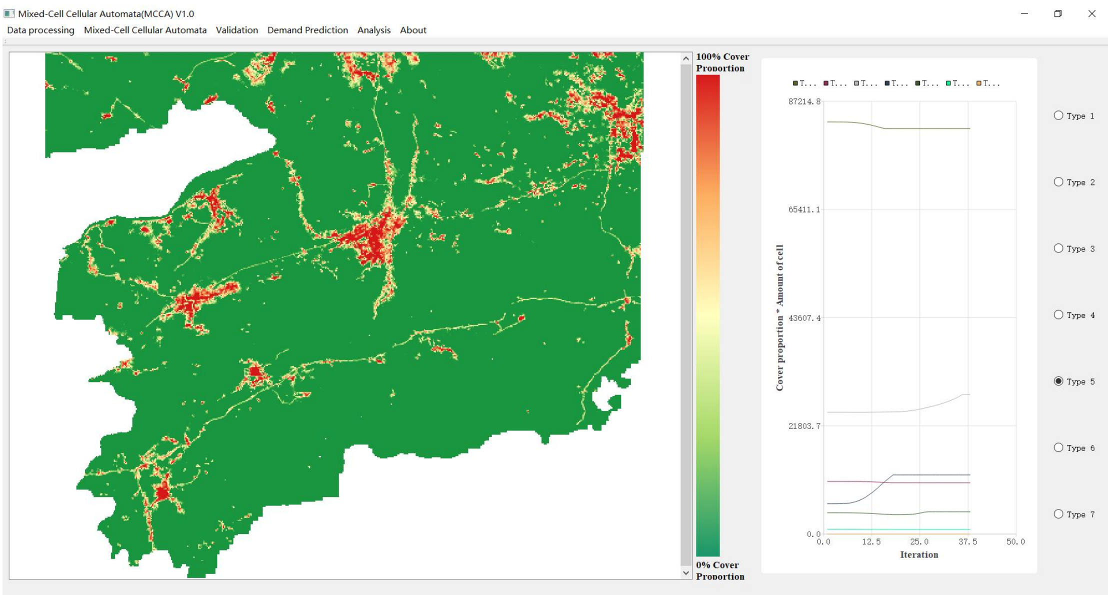


Fig. 35. Dynamic display of the simulation from 2015 to 2035.

4.6. Calculation of land use mixure

The mixture of land use components in land units is closely related to the socio-economic activities, environmental functions and landscape amenities which is very important for regional sustainable development. Different from pure-cell CA models, mixed cell CA models have the advantage of simulating the changes of land use structure within individual cells. Their entropy can directly measure the cell-level mixture according to the cover proportions of land use components. Higher entropy means a higher degree of mixture or diversity of land use in an individual cell.

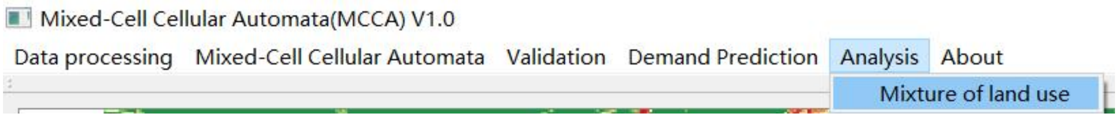


Fig. 36. Activating the ‘mixture of land use’ module.

4.6.1. Setting the input and output paths

Input the amount of land use type in the simulation, the amount of land use types in this experiment is 7. Click the  button in the “**Land use structure**” row, select the simulation result in 2035 “result_luccSimulation2035Simulation _1_Type1.tif” in the pop-up dialog. The program will automatically load all the 7 bands of the simulated land use structured data. Then click the  button in the “**Output**” row to set the output path of the mixture image of land use.

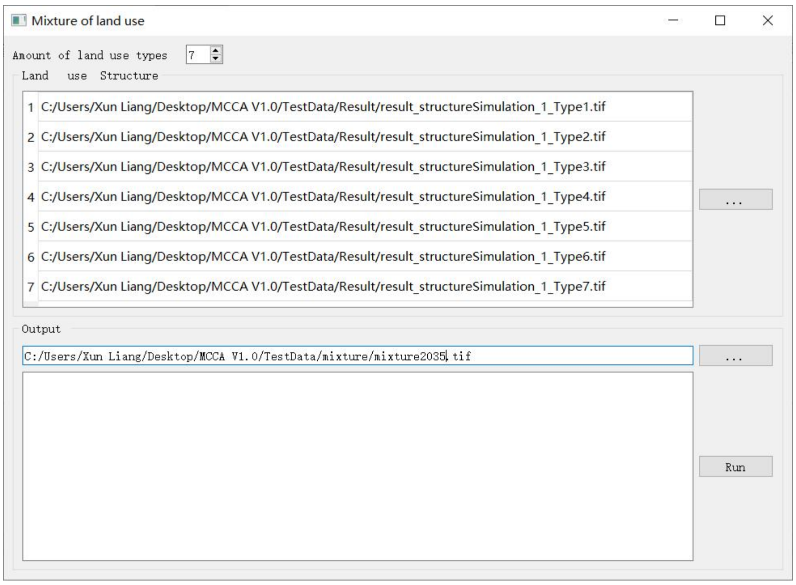


Fig. 37. Setting the input and output paths for the 'mixture of land use' module.

4.6.2. Calculating the land use mixture of each cell

Click the  button to generate the predicted land-use mixture of each cell in 2035.

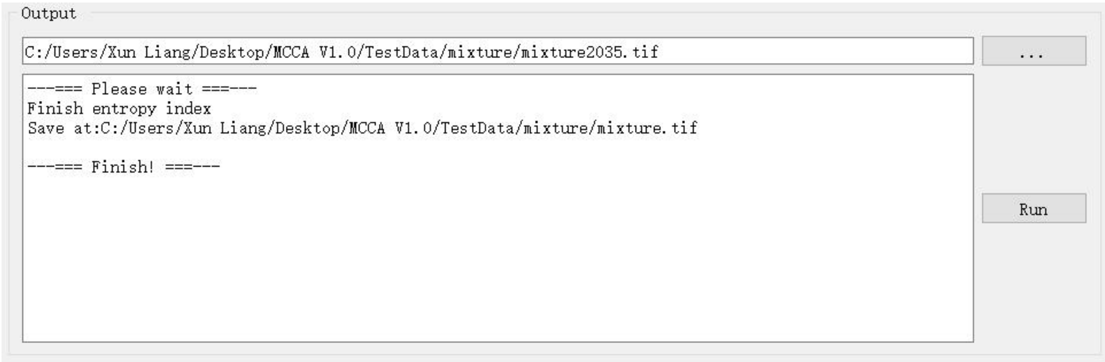


Fig. 38. Generating the image of land use mixture in 2035.

5. Preparing for historical land use structure data

In practical application, the mixed-cell land use data can be obtained using at least four approaches: (1) remote sensing inversion method; (2) decomposition of mixed pixels; (3) extracting the land use mix structure from multi-source big data (e.g., social network data, taxi trajectories, points of interest) ; and (4) aggregating fine-resolution land use data into coarse-resolution data. We used the fourth approach——the most convenient method to obtain the mixed-cell land use data to illustrate the MCCA model. The method can be handled in two ways: by using the fishing net conversion or by generating it using the aggregation tool.

5.1. Fishnet

To aggregate fine-resolution land use data into coarse-resolution data and calculate the cover proportion of each land use type for each cell. The users first need to generate a fishnet for the study region, just like the 'demogrid' in 'demo_grid' folder provided in our test data and calculate the cover proportions with zonal statistics tool in ArcMap or QGIS. Please don't clip the finshnet grid with region boundary!

However, the grids with the area of each land use type should be converted to the raster format so that they can be as the input of MCCA software. The MCCA model provides two data processing tools to help users prepare the land use structure data of raster format.

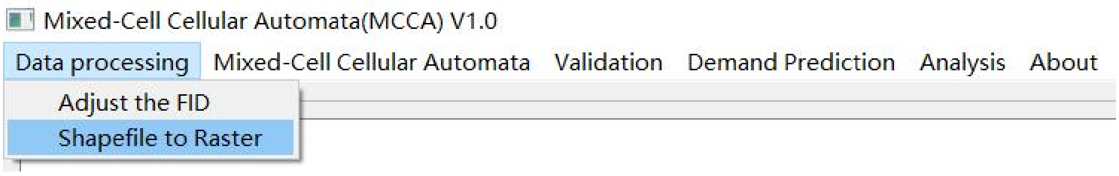


Fig. 39. Data processing tools for preparing the historical land use structured data.

The grid generated by ArcMap or QGIS cannot be directly used to obtain the area information of multiple land use. Because the FID (serial number) of the ArcMap grid starts from the bottom left corner and increase from the bottom row to the top row, while the FID (serial number) of the QGIS grid starts from the up left corner but increase from the left column to the right column. These two arrangement is inconsistent with the raster format data that start from the up left corner and increase from the top row to the bottom row. Therefore, users can firstly apply the 'Adjust the FID' module to modify to fishnet grid.

5.1.1. Adjusting the FID for grid shapefile

Click the "**Adjusting the FID**" item to start the simulation module. Select the fishnet grid generated by the ArcMap or QGIS grid, and set the output path of adjusting fishnet grid.

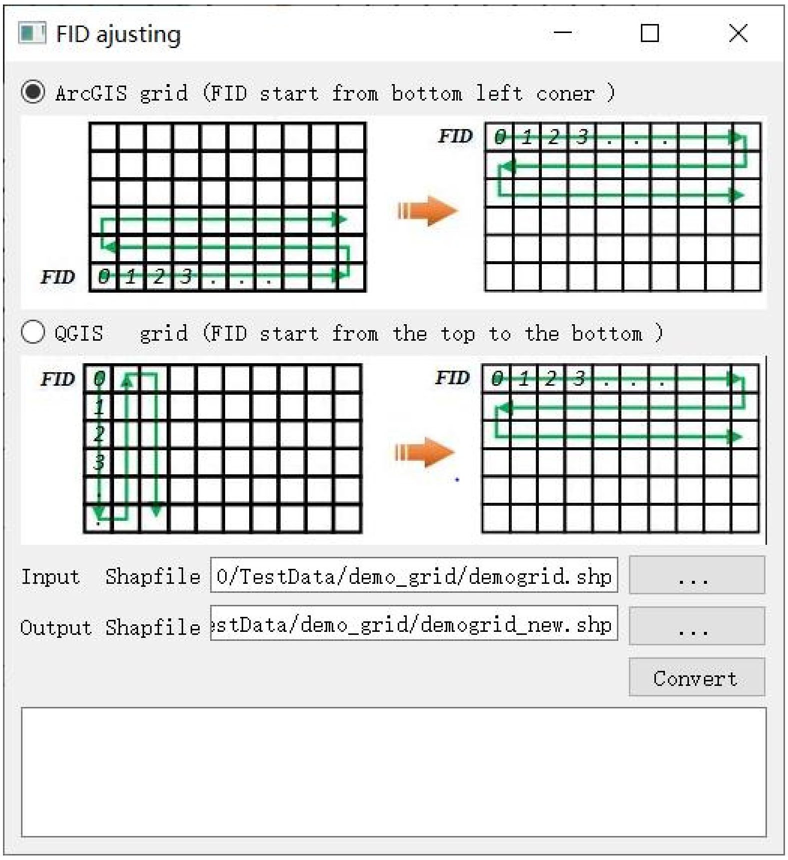


Fig. 40. The interface of the 'FID adjusting' module.

Then click the  button to obtain the new fishnet grid. If the original grid is generated by QGIS, just choose the 'QGIS' radio button. There will come out a message box when the model process run is complete.

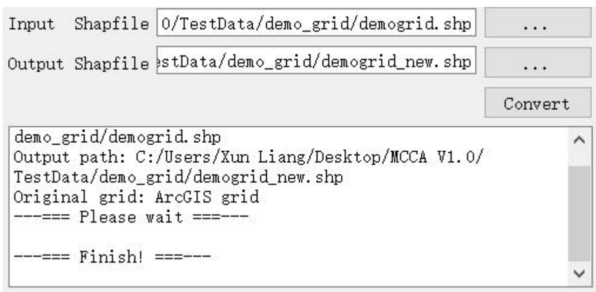


Fig. 41. The output information of the 'FID adjusting' module.

After this step, the users can used the adjusted grid to obtain the area of each land use in each cell, by aggregating fine-resolution land use data into the coarse-resolution grid using zonal statistics tools in ArcMap (Area Tabulate) or QGIS. Then the users can ouput a new adjusted grid (like Xianning_2015.shp in the test data) with area information of multiple land use components with GIS software.

5.1.2. Shapefile land use structure data to raster format

Click the "**Shapefile to Raster**" item to start the simulation module. Select the fishnet grid exported by the last step. We provide the four land use structure data in shapefile format: 'Xianning_2000.shp', 'Xianning_2005.shp', 'Xianning_2010.shp', 'Xianning_2015.shp', which have been already converted.

Select the four shapefile grid with land use area of each land use type on each cell, and choose an output folder.

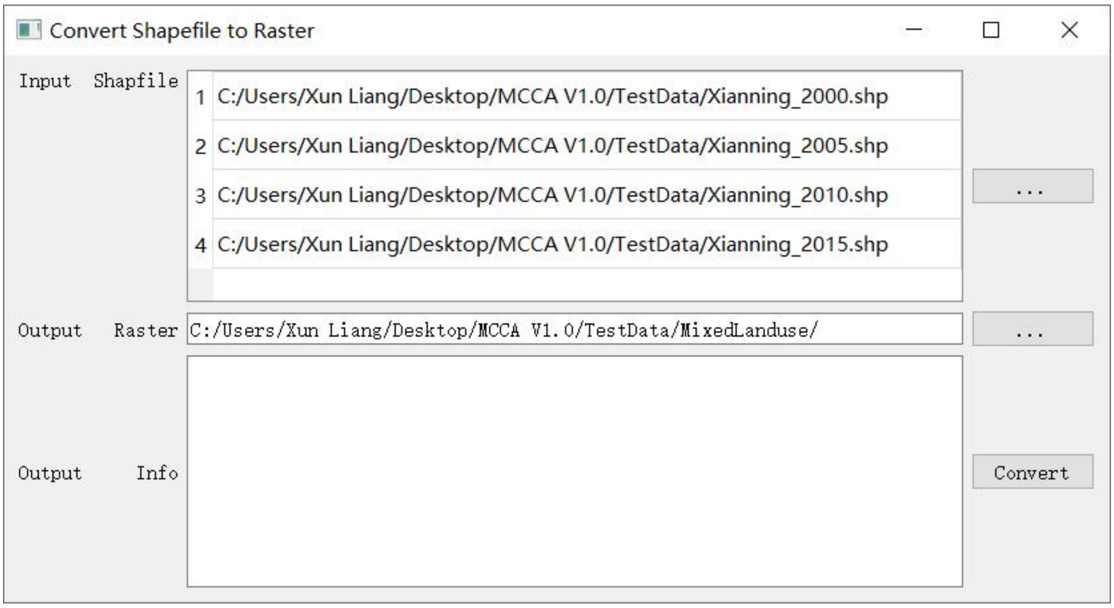


Fig. 42. The interface of the 'Convert Shapefile to Raster' module.

Then click the 

button to generate land use structure data of 7 layers. There will come out a message box when the model process run is complete. **Note that the program only identifies the fields in the shapefile with the format of 'VALUE_X(1~7)'.**

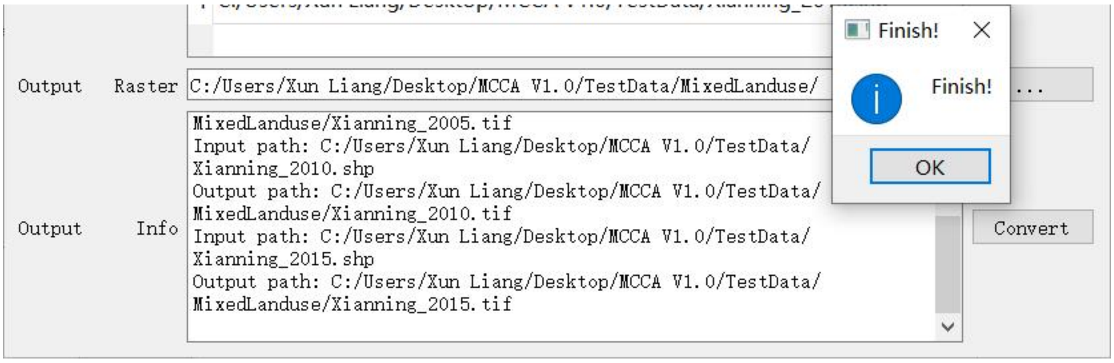


Fig. 43. The output information of the ‘Convert Shapefile to Raster’ module.

5.2. Aggregate

5.2.1. Extract each land class

The user first needs to start the Arcmap software, open the ArcToolbox menu, and use the Raster Calculator tool to extract each land class separately. The path to the "Raster Calculator" tool is: Spatial Analyst Tools - Map Algebra - Raster Calculator. -Raster Calculator. In the map algebra expression using the Con function, enter the expression Con("land use data" == value,1,0), land use data for a certain year, value is set to the value of a certain class of land, the formula means that the land use data, equal to the value of the set In this example, there are 7 land use data types, so we need to set the value to 1-7 in order to run the raster calculator 7 times to get the results of 7 land use types in the same year.

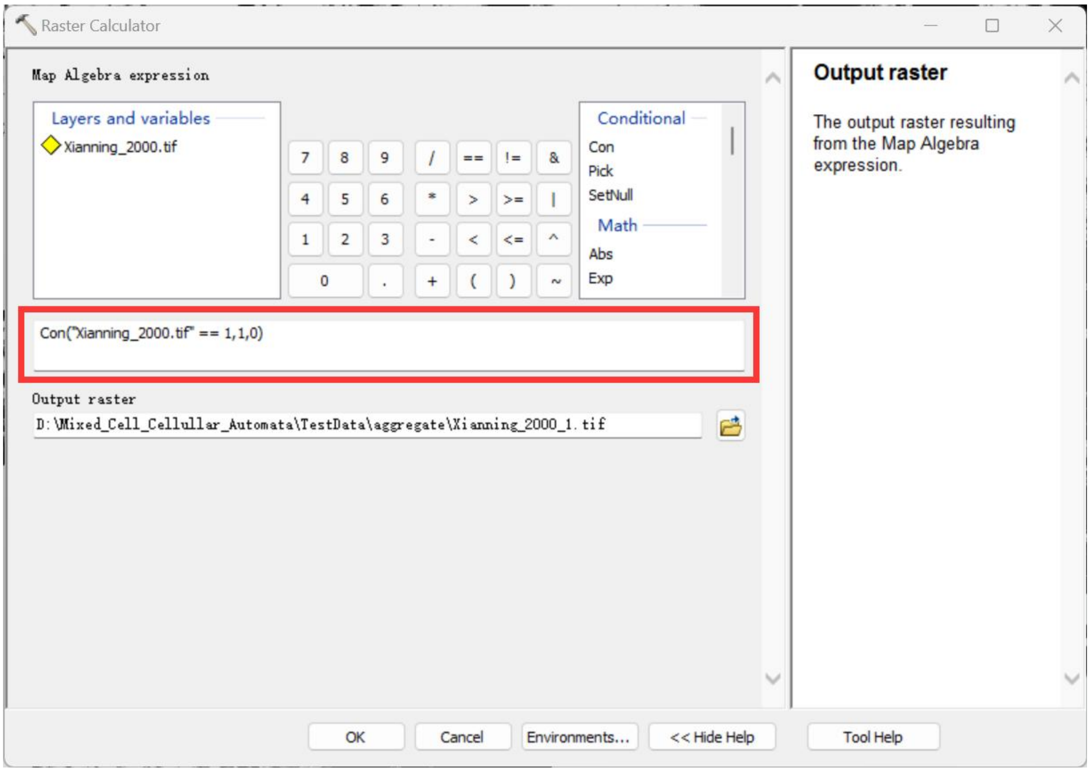


Figure 44. Using the raster calculator to extract a certain land class

5.2.2. Aggregate analysis

Use the "Aggregate" tools in Arcmap to generate structural data for each land use type. The path to the "Aggregate" tool is: Spatial Analyst Tools - Generation - Aggregate. In the "Input raster" line, enter the land use structure data obtained in the previous step, set the output raster path, set the image element coefficients according to the experimental requirements, the coefficients must be integers greater than 1, and choose "MEAN" for the aggregation technique, each output image element contains Each output image contains the average value of the input images covered in the range of this image.

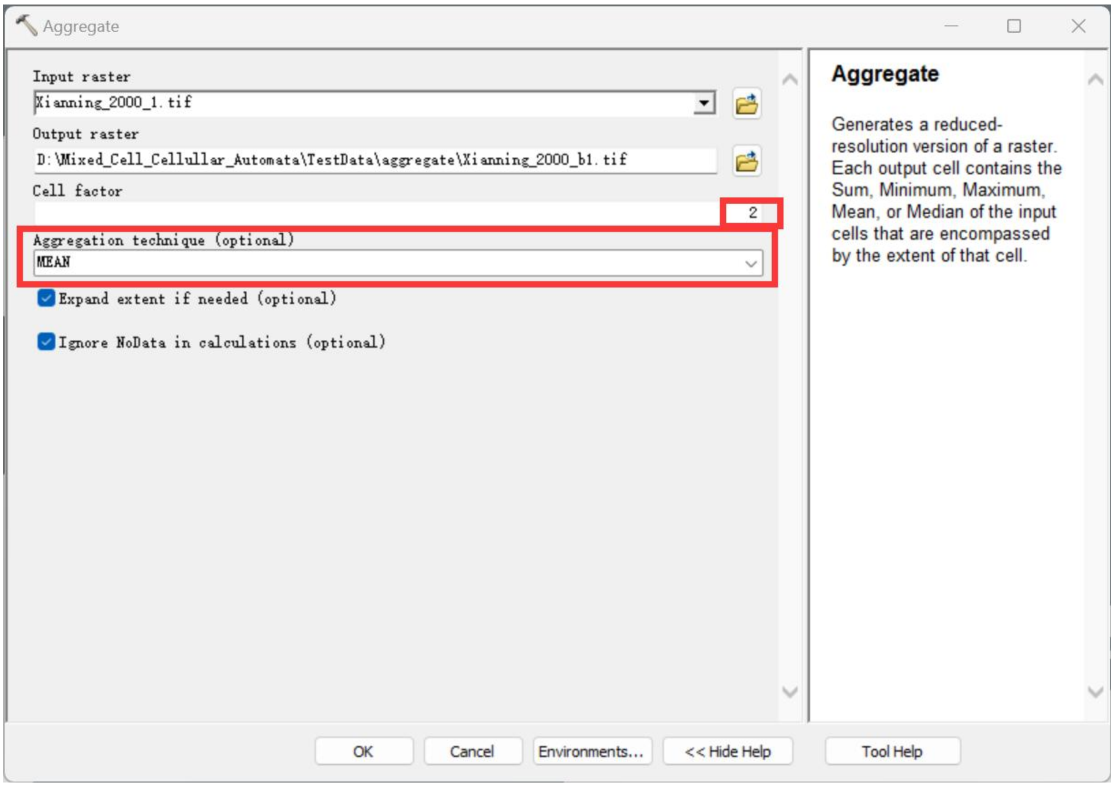


Figure 45. Using the aggregation tool to generate coarse-resolution data

5.3. Convert data to 'double' format

MCCA software requires data in a 'double' format. This step provides a method to convert data to 'double' format using ArcMap software. Start the ArcMap software, open the ArcToolbox menu, and the tool path of the **Copy Raster** we used is: Data Management Tools - Raster - Raster Dataset - Copy Raster. Input a raster data, select an output file, and choose **64_BIT** in the Pixel Type drop-down box to export data in a 'double' format.

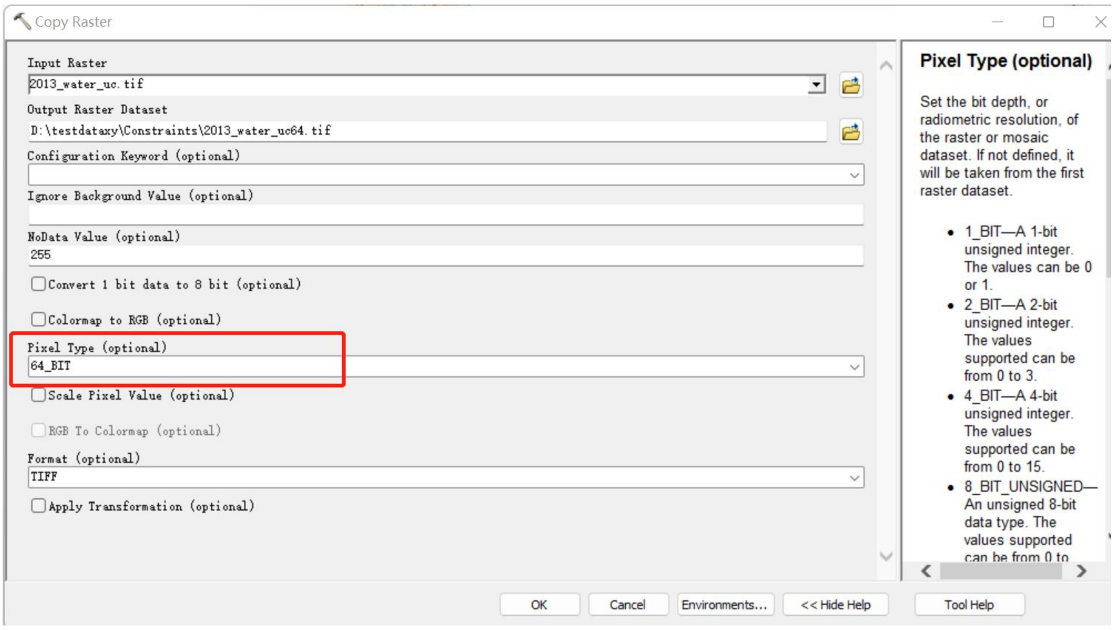


Fig. 44. The Copy Raster Tool of Converting data to 'double' format