Open space Cellular Automaton V1.0 (OS-CA)

User's Manual

Authors:

Xun Liang[#], He Tian[#], Xia Li* , Jun-Long Huang, Keith C. Clarke, Yao Yao, Qingfeng Guan， Guohua Hu*

* Corresponding author. E-mail address: lixia@geo.ecnu.edu.cn (X. Li); ghhu@geo.ecnu.edu.cn (G. Hu)

Consultation

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

Organization:

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

Key Lab of Geographic Information Science (Ministry of Education), School of Geographic Sciences, East China Normal University, Shanghai 200241, China

Contents

Abstract 1

1. Download and start OS-CA 2
2. Example data description 3
3. User interfaces of OS-CA model 4
   - 3.1. Probability surfaces caculation 4
     - 3.1.1. Start the module 4
     - 3.1.2. Input land expansion data 5
     - 3.1.3. Input driving factors 5
     - 3.1.4. Setting the training parameters and output path 6
     - 3.1.5. Running program and examine the result 7
   - 3.2. A cellular-automaton sub-model for OS 9
     - 3.2.1. Start the module 9
     - 3.2.2. Input initial open space 9
     - 3.2.3. Input probability surfaces 10
     - 3.2.4. The constraint of open water (Spatial policies) 10
     - 3.2.5. Set the path of saving simulation result 11
     - 3.2.6. Simulation parameters 11
     - 3.2.7. Setting the colors of all land use types for dynamic display 13
     - 3.2.8. Running the A cellular-automaton sub-model for OS 14
4. Some notes applying the OS-CA model 17

Abstract

We proposed a spatio-temporal approach based on Cellular Automata (CA) for simulating the spatial dynamics of open spaces (include urban green spaces, parks, squares, trails, courtyards, and other natural spaces), by considering a set of spatial data that represents the infrastructural and socio-economic factors, namely the OS-CA (Open Space Future Land Use Simulation) model. The OS-CA can accurately simulate the emergence of OS by considering inter-attraction between OSs and urban land. And the proposed model can simulate the pattern of OS under the scenarios of building different miean size of OS, which is of importance for assisting urban designers in making suitable plans for urban open space.

OS-CA was developed purely in the C++ language. The parallel technology of OS-CA software is from High-performance Spatial Computational Intelligence Lab @ China University of Geosciences (Wuhan) (https://github.com/HPSCIL). The Artificial Neural Network technique in our model is from a powerful open -source library called Alglib 3.9.2 (http://www.alglib.net/). The UI of the software is built using a famous open -source library Qt 5.13 (https://www.qt.io/download/). This UI provides a real-time display of dynamic changes of land use in the 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 users a "quick start" on how to use the OS-CA software. All of the necessary data and files for the tutorial have been provided and can be used as templates forpractices.

1. Download and start OS-CA

The latest version of OS-CA software and the user's manual are provided at a download link (https://github.com/HPSCIL/Open-Space-Cellular_Automaton). Click the executable files "**OS-CA VX.0.exe**" to start the software. The users should download the whole "Zip" files. Unpack it and double-click the executable files "**OS-CA VX.0.exe**" to start the software. Note that the OS-CA model will not work if the users only download a single "exe" file.

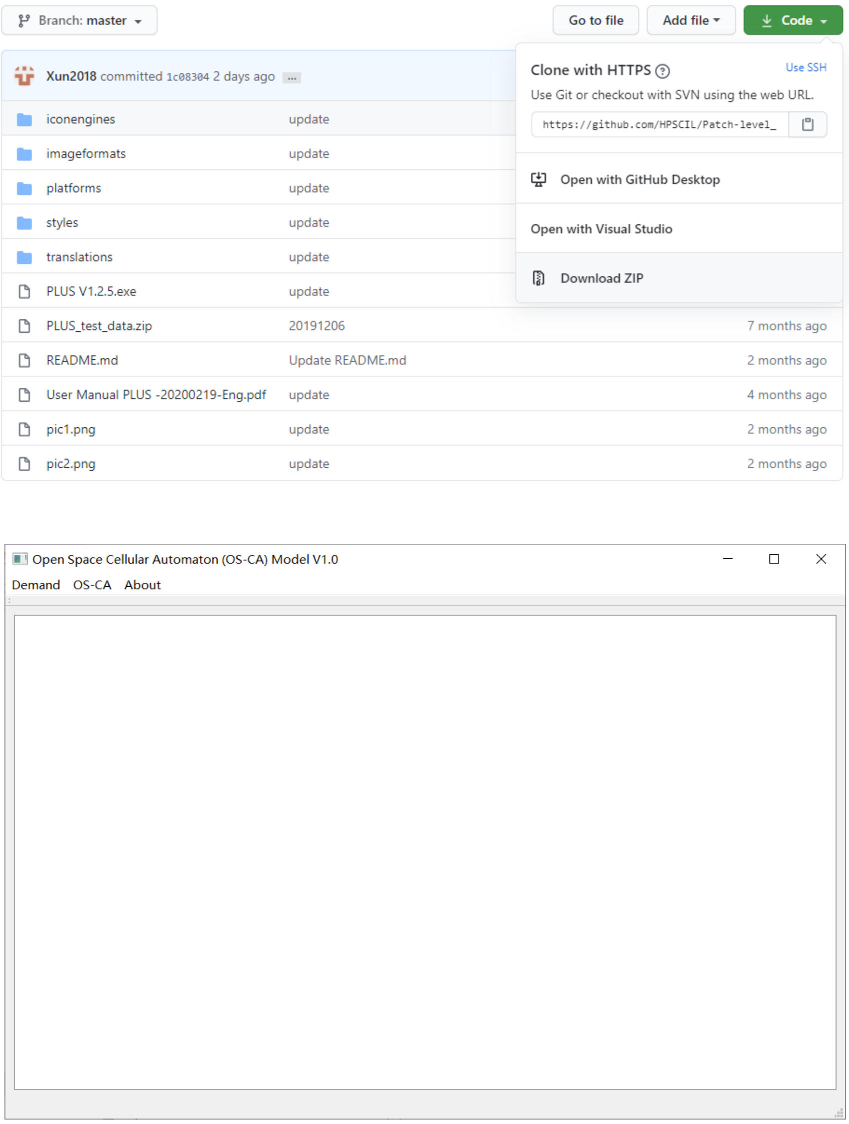


Fig. 1. The main interface of the OS-CA model.

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

2. Example data description

Tab. 1 List of test data

Category	Data	Filename	Description
Open space data	Land use data	dgnewlanuse30nomoun.tif	1: Non-urban land, 2: Urban land, 3: Open spaces (urban green spaces, parks, squares, trails, courtyards, and other natural spaces)
Constraint data	Land use constraint	dgrestriction1.tif	
Infrastructure	Density of bus stations	dgbusstationdis.tif	Point of interest (POI)
	Density of entertainment	dgentertaindis.tif	
	Proximity to hospitals	dghospitaldis.tif	
	Density of restarants	dgrestaurant.tif	
	Density of shops	dgshopdensity.tif	
	Proximity to restrooms	Dgtoiltdis.tif	
	Proximity to factories	dgfactorydis.tif	
Terrain	Slop	dgslope.tif	Terrain data
	DEM	dgdem.tif	
Data products	Population	dgpop2015v3.tif	Soio-economic driving factors
	GDP per capita	dgrenjunGDP.tif	
Location factors	Other roads	dgqitadaoludis.tif	Traditional driving factor for CA simulation
	Proximity to railway	dgrailway.tif	
	Proximity to rivers	dgriverdis.tif	
	Proximity to towncenters	dgzhencenter.tif	
	Proximity to mainroads	dgzhugandaodis.tif	
	Density of intersetions	Roadpointdis.tif	
	Proximity to forest	dgforestdis.tif	

Proximity to highway	dggaosudis.tif
Proximity to open spaces	dgosdis.tif
Other roads	dgqitadaoludis.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, more human and natural environment effects can be taken into consideration. The study region in this example is Wuhan city, located in Hubei province, central China.

3. User interfaces of OS-CA model

First of all, data processing is critical! The OS-CA model only supports the raster format land use/land cover (LULC) data, whose land use types only includes 3 land use types: non-urban land, urban land and open space, which are encoded as 1, 2, 3 in this example respectively. Also, the OS-CA model only supports the raster-format land use constraint data, which is a binary value image that ranges from 0 to 1. **All the input image data need to have the same number of rows and columns, including the land use pattern, restricted area, driving factors, and probability-ofoccurrence data.**

3.1. Probability surfaces caculation

3.1.1. Start the module

Click the "**Probability surfaces caculation**" item to start the module.

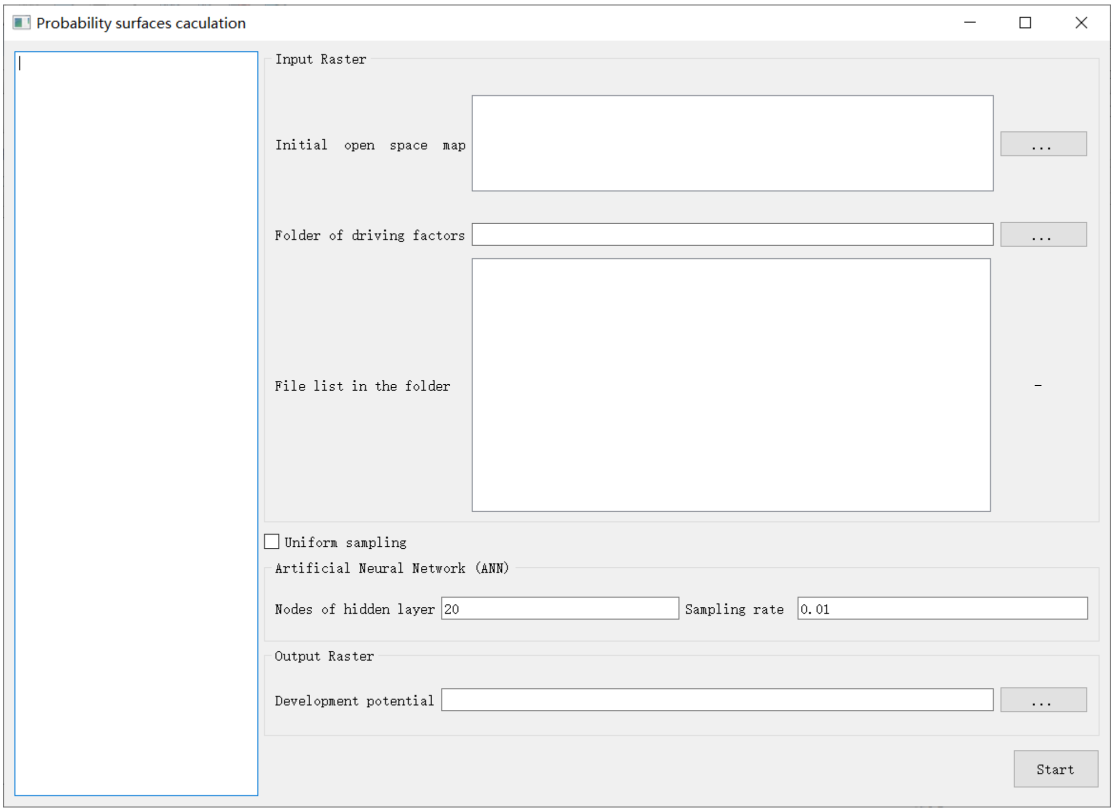


Fig. 5. UI of the ‘Probability surfaces caculation’ module.

3.1.2. Input land expansion data

Click the  button of the “Initial open space map” text box and select the land expansion map “dgnewlanuse30nomoun.tif” outputted in the previous step in the pop-up dialog.

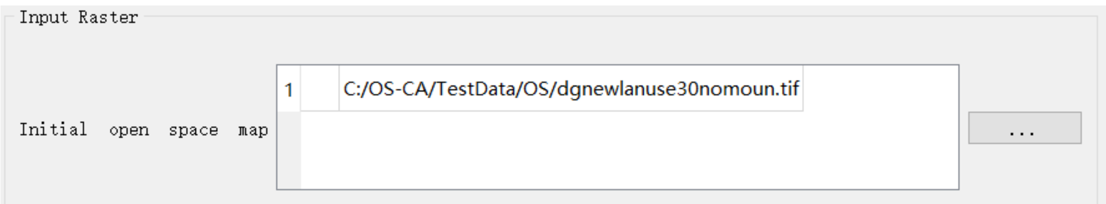


Fig. 6. Select the land use map outputted in the last step.

3.1.3. 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 OS-CA software would automatically load the "tiff" File in the folder. We provide 16 driving factorsimplementing the simulation.

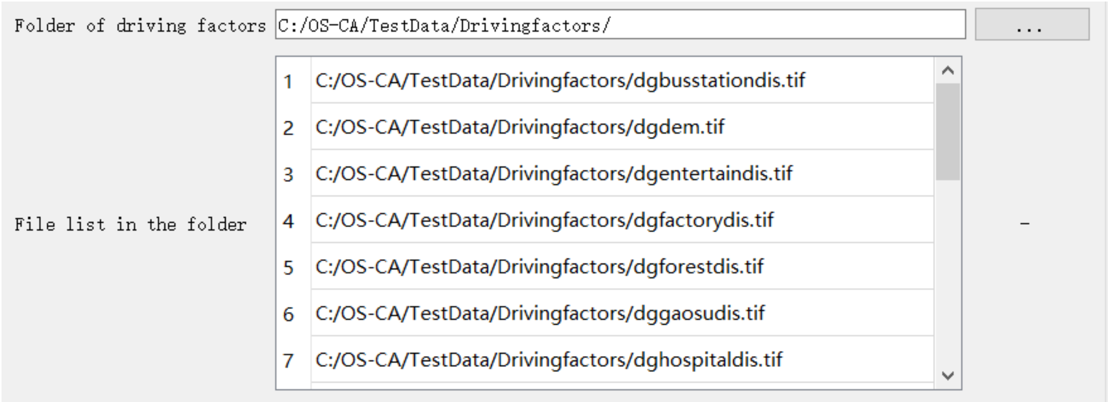
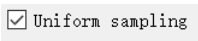


Fig. 7. Input the driving factors.

3.1.4. Setting the training parameters and output path

Users can set the parameters of **Artificial Neural Network (ANN)**. The sampling rate is set to 0.01 by defaultwhich means about one percent of pixels is selected for training. The number of hidden nodes is set to 10 in this experiment. If the checkbox 

is checked, the sampling points for each land use type will be the same. If not, the sampling points will randomly scatter across the study area. Then, click the  button in the "**Output Raster**" group box and select the change probability in the pop-up dialog. In the test data, the generated development potential of each land use type is named "osprobability.tif". Users can increase the number of parallel internal threads to accelerate the running speed.

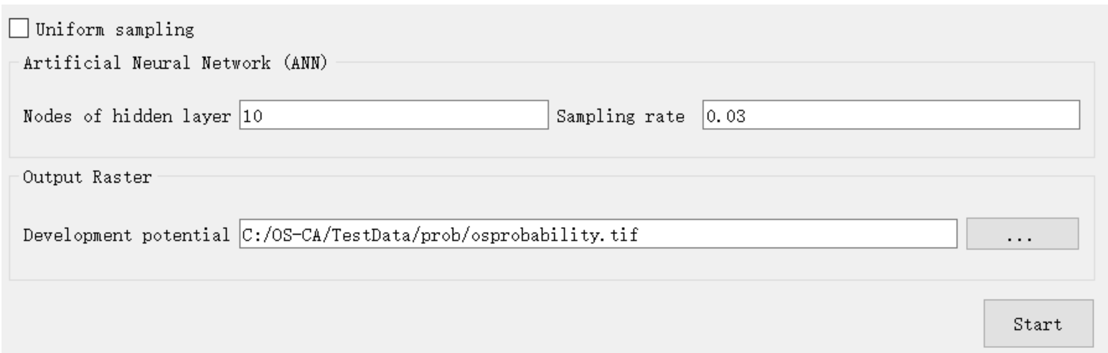
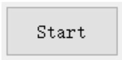


Fig. 8. Determining the training parameters and output path of the ANN.

3.1.5. Running program and examine the result

Once the setting-ups complet, click the button 

for running the module. A message box will come out as below when the model process run isfinished.

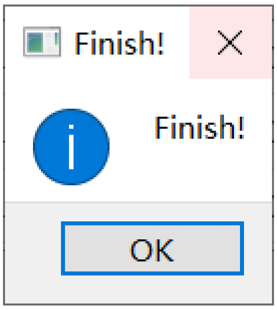


Fig. 9. A message that shows the program run successfully.

The text panel at the left of the interface will output the training accuracies of the RFCs for each land use component. The ANN outputs two accuracy indicators: RMSE, average error and average relative error for each land use type. And a probability surface data that include three bands (corresponding to non-urban ladn, urban land and open space) will be generated.

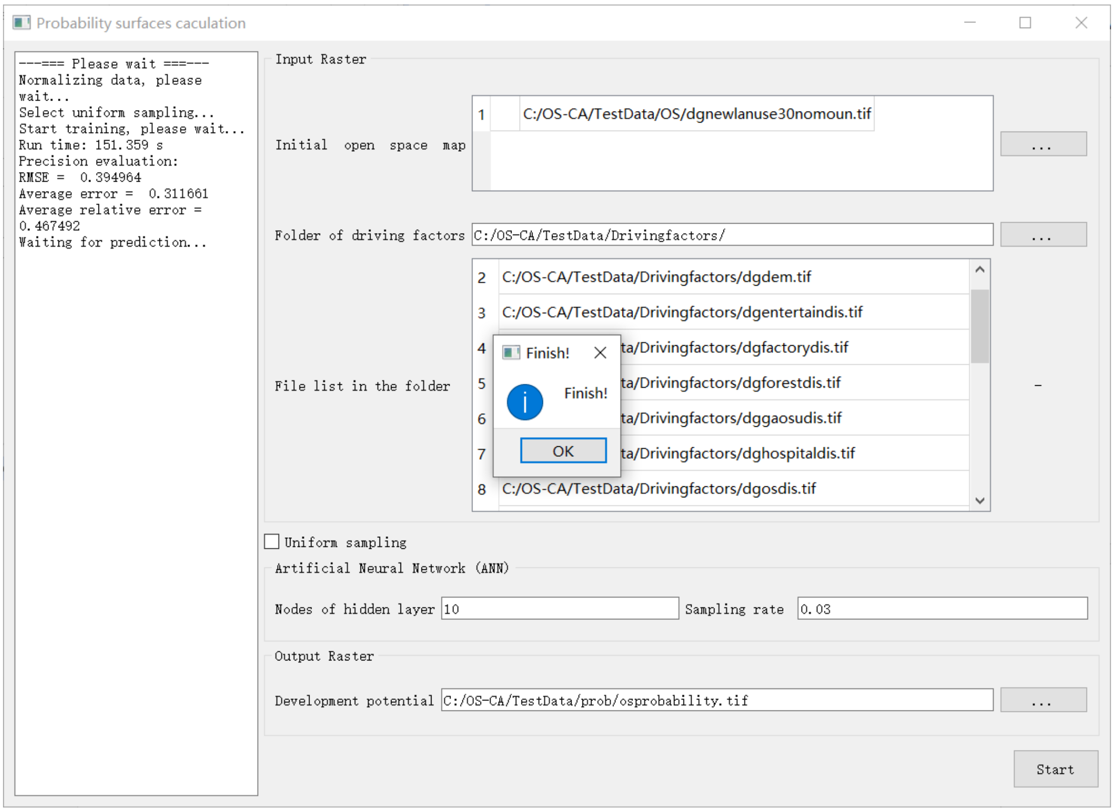


Fig. 10. Accuracy for each land use type.

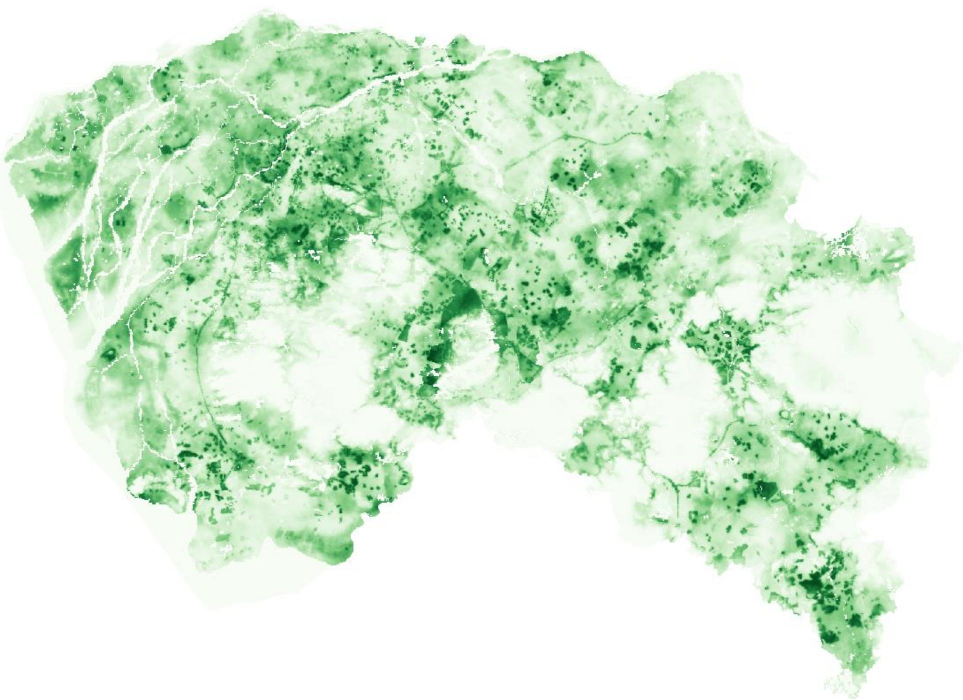
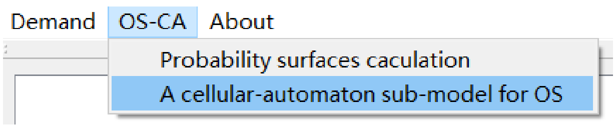


Fig. 11. Probability surface for open space (the third band).

3.2. A cellular-automaton sub-model for OS

3.2.1. Start the module

Click the "**A cellular-automaton sub-model for OS**" item to start the simulation module.



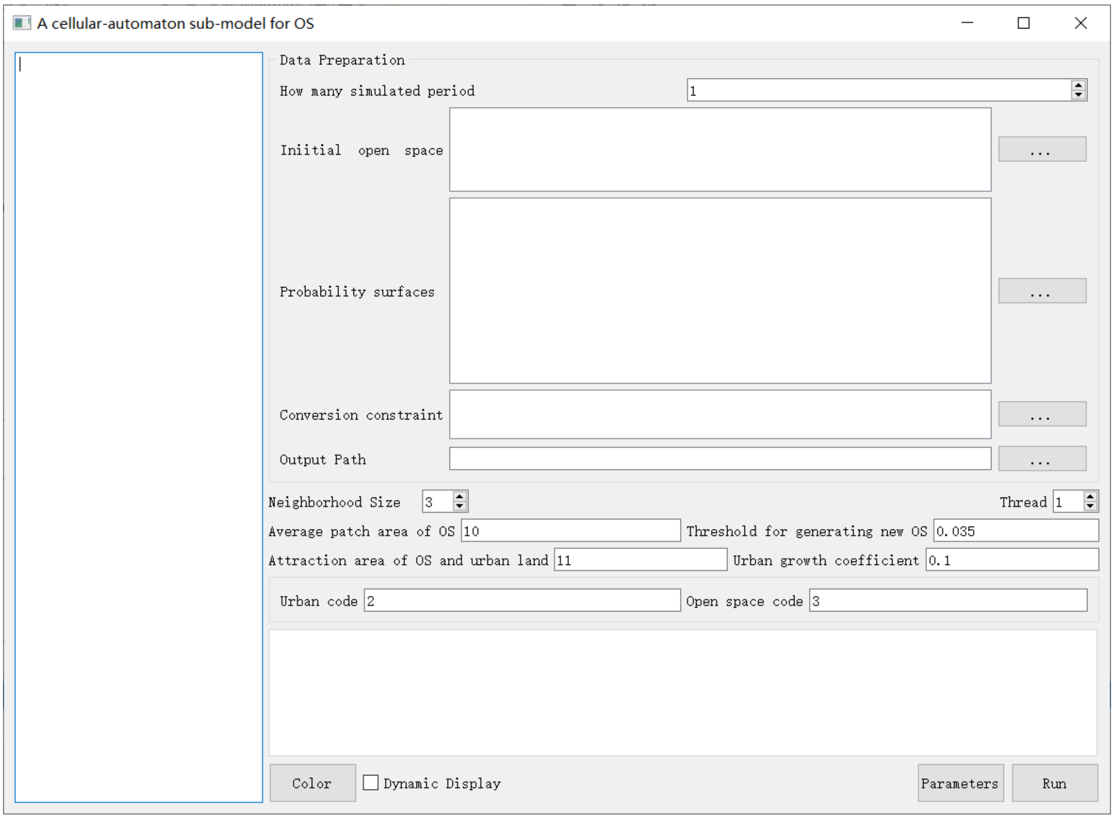


Fig. 12. The UI of the 'A cellular-automaton sub-model for OS' module.

3.2.2. Input initial open space

Click the  button in the "**Initial open space**" text line and select the land use data of the start year "dgnewlanuse30nomoun.tif" in the pop-up dialog. We aims to simulate the development of open space from 2015-2020, 2020-2025 and 2025-2030. So the simulated period is set to 3.

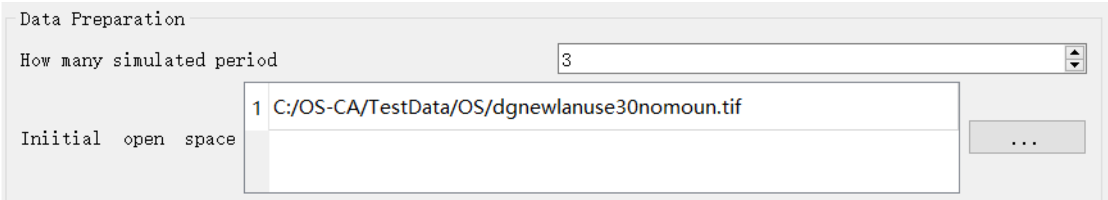


Fig. 13. Select the land use structure data at the start year.

3.2.3. Input probability surfaces

Input the development potential file generated from the ‘Probability surfaces calculation’ module in the “**Probability surfaces**” group box. After that, three lines of blank future demands will be automatically loaded in the tab widget and waiting for the users to fill in.

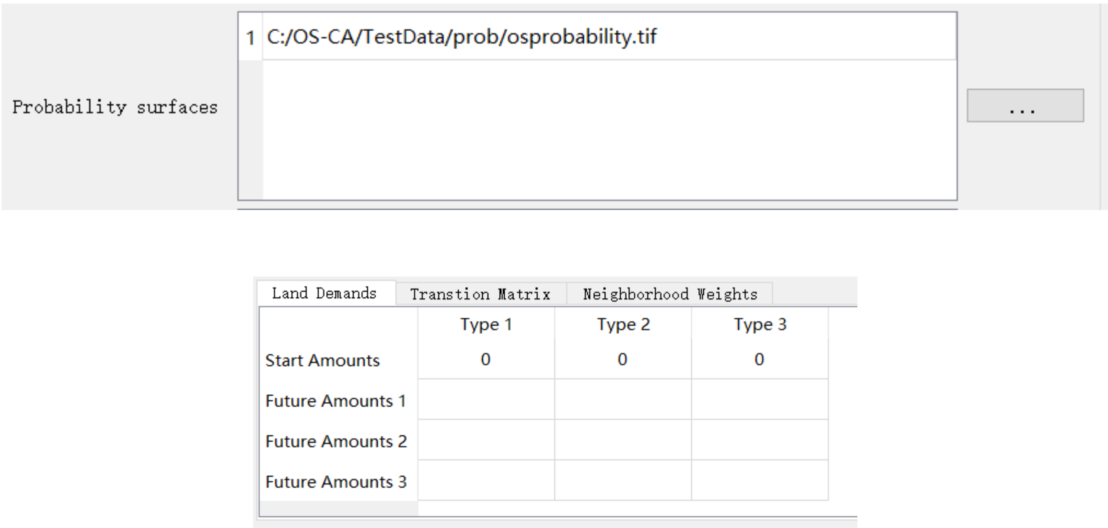


Fig. 14. Select the potential development maps generated by the Probability surfaces calculation module.

3.2.4. The constraint of open water (Spatial policies)

Some policies restrict all types of land change in designated areas, such as open water and some nature conserved areas. Users should prepare a binary image of restricted areas that each pixel only contains value of 0 or 1. The value 0 means no conversion while the value 1 meansconvertible. In this case, we assume that open water (a sub-category of waterbody) is not allowed to be converted to other land use types.

Fig. 15. Select the constraint data.

3.2.5. Set the path of saving simulation result

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

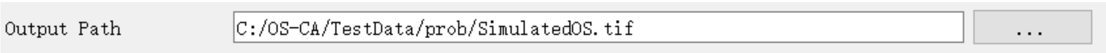


Fig. 16. Select the output path.

3.2.6. Simulation parameters

Simulation parameters include the land use demand, transition matrix, and neighborhood weights. The land use demand is an vital parameter of the OS-CA modelthat needs to be firstly set according to the actual situation of the study region. We suggest users to apply external models to determine the land use demand for future scenarios. <u>We provide TWO simple tools for calculating the historical land demand and predict future land demand (markov chain and linear regression)</u>. However, both the two methods need at least 2 terms of land use data. But we don’t have two terms of spatial open space data in this study. So we don’t introduce the two prediction tools in this user’s manual, detailed description of these tools is elaborated in https://github.com/HPSCIL/Patch-level_Land_Use_Simulation_Model.

Fig. 17. Two predicton tools of future demand of open space and urban land.

Therefore, we directly used the demand of open space in 2020, 2025 and 2030 predicted by System Dynamic (SD) model in our article to illustrate the OS-CA software. In the tab widget, users can input the future land use demands of each land use type of different years to the rows “Future Amounts” of the “Land Demand” page. The OS-CA model will automatically stop when the allocated area equals to the demanded area for all land use types in each year.

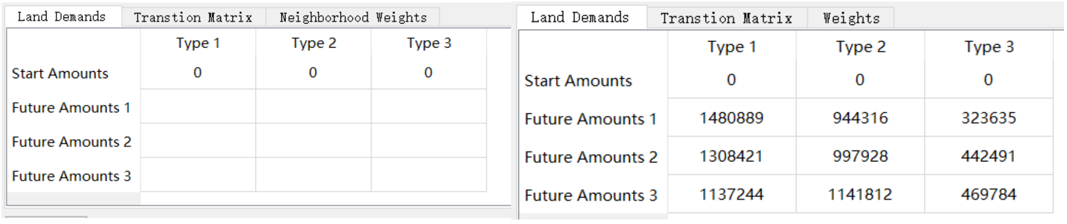


Fig. 18. Input the future land use demand.

In addition, the neighborhood effect of OS-CA is similar to the traditional CA model. Here the default value of the **neighborhood effect** is 3. The **Average patch area of OS** is a parameter for generating new patches under different scenarios. The **threshold for generating new OS** rangs from 0 and 1.0. It only allows the cells with the OS probability larger than the threshold to have the chance to emerge new OS. The **Attraction area of OS and urban land** is set to 11 in this example. The urban growth coefficient ranges from 0 to 1. A higher value will result in a fragmentized pattern of urban land, the default value is 0.1. What's more, the

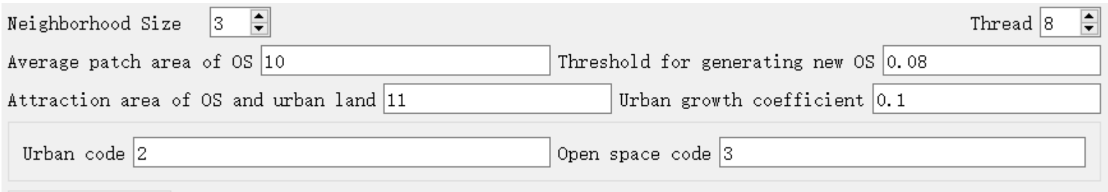


Fig. 19. The simulation parameters of the OS-CA model.

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 the expertise about the study area. An example of the cost matrix is shown below:

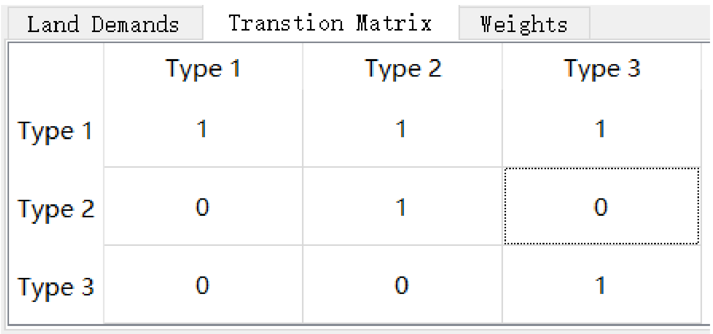


Fig. 20. Transition Matrix.

In the “**Neighborhood weights**” page, users need to set the neighborhood weights of different land use types. The neighborhood effects may be different for different land use types in a unique study region. The value of the neighborhood weight for each land use type can be determined according to expert knowledge and a series of model tests, ranging from 0-1. At this example, we set the parameter as below:

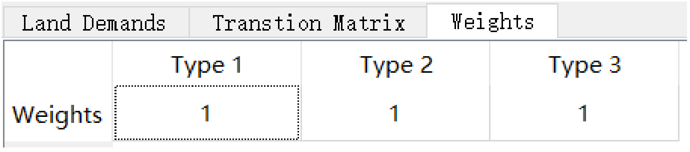


Fig. 21. Neighborhood weight for each land use type.

3.2.7. Setting the colors of all land use types for dynamic display

Fig. 22. The position of the ‘Color’ button on the UI.

Click the “**Color**”button to activate the ‘Set Color Display’ module.

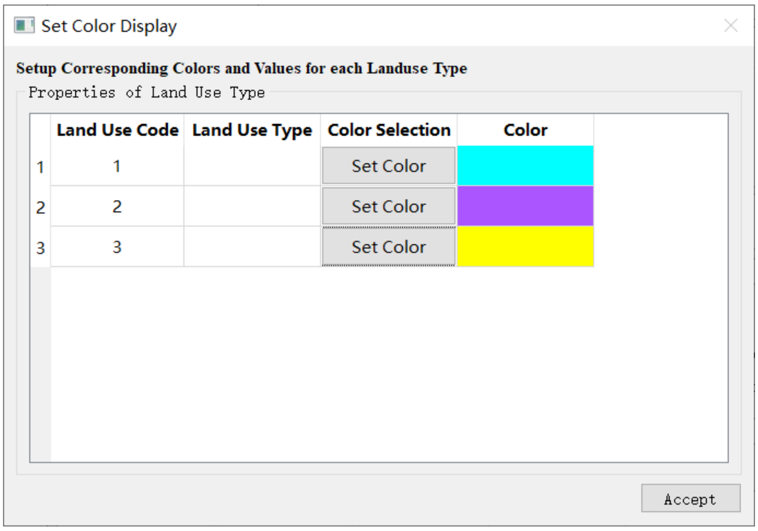


Fig. 23. The ‘Set Color Display’ interface.

Click the “Set Color” button of each row in the list box to set an RGB value (such as rgb(150,200,50)) for all land use types. Then, input the name of each land use type in the ‘Land Use Type’ column. Finally, click the “**Accept**” button, confirm the above settings.

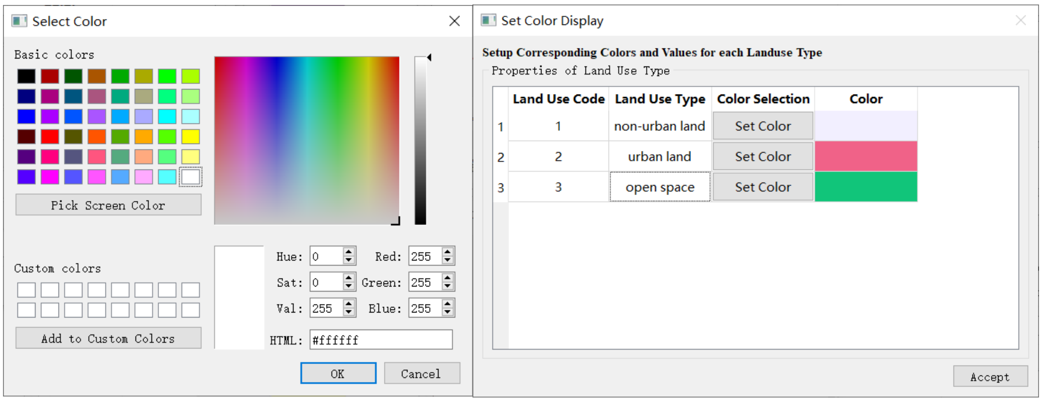
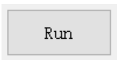


Fig. 24. Setting color for all land use types.

Note that the OS-CA model allowed the users to not activated the dynamic display function by unchecking the ‘Dynamic Display’ checkbox. Because the display function can cause software crashes in some situations. Running the OS-CA model without a dynamic display can also make the model run faster.

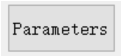
Fig. 25. Cancel the dynamic display by unchecking the ‘Dynamic Display’ checkbox.

3.2.8. Running the A cellular-automaton sub-model for OS

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 OS-CA model can presentthe process of land use change dynamically, including the land cover spatial change, value change and change curve of each land use type on the main interface.

The visualization panel allows the users to zoom in and observe the land use dynamics processes of study regions. Users can check any land use types by activating the corresponding radio buttons on the right. The horizontal axis of the dynamic chart represents the number of iteration, and the vertical axis represents the product of the cover proportion and the total amount of cell.

The model will stop and save the simulation result when the allocated area equals to the demanded area for all land use types. The OS-CA software will save all the parameters to two configuration files: "SimulationLog.txt" 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 user to use to repeat their experiments.

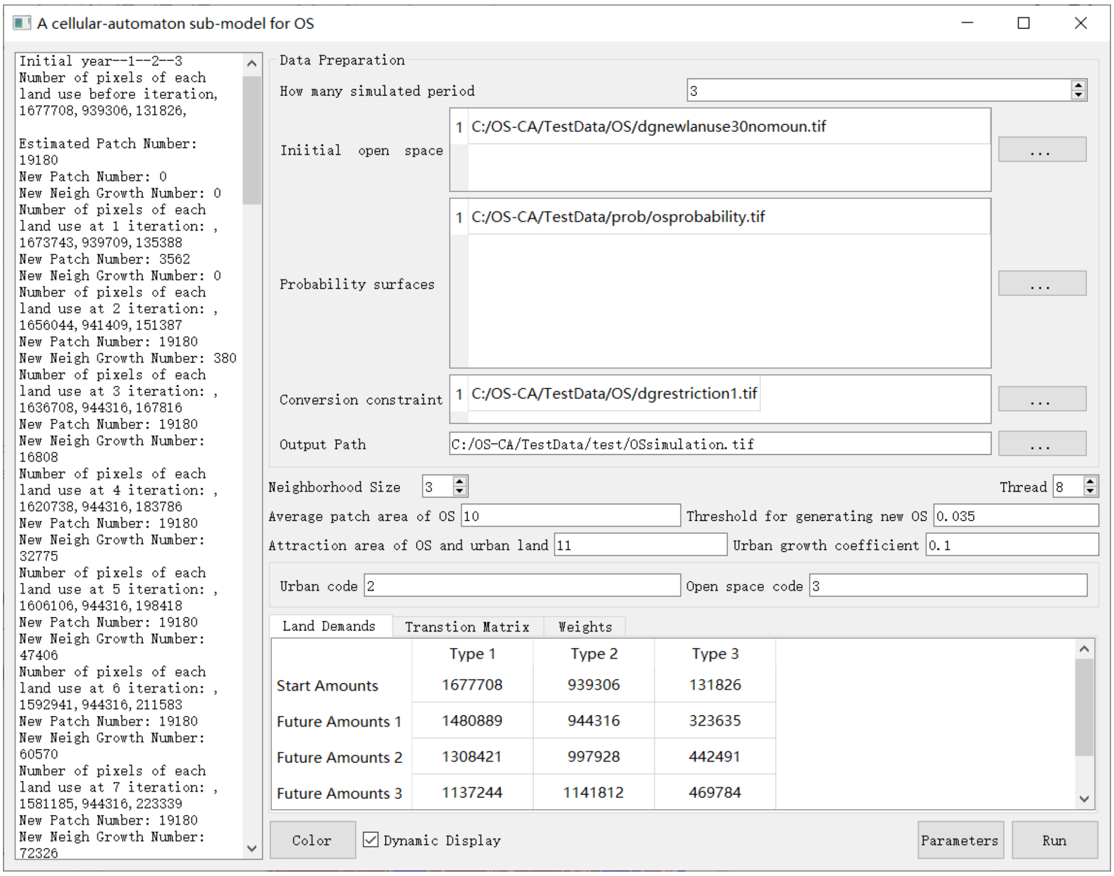
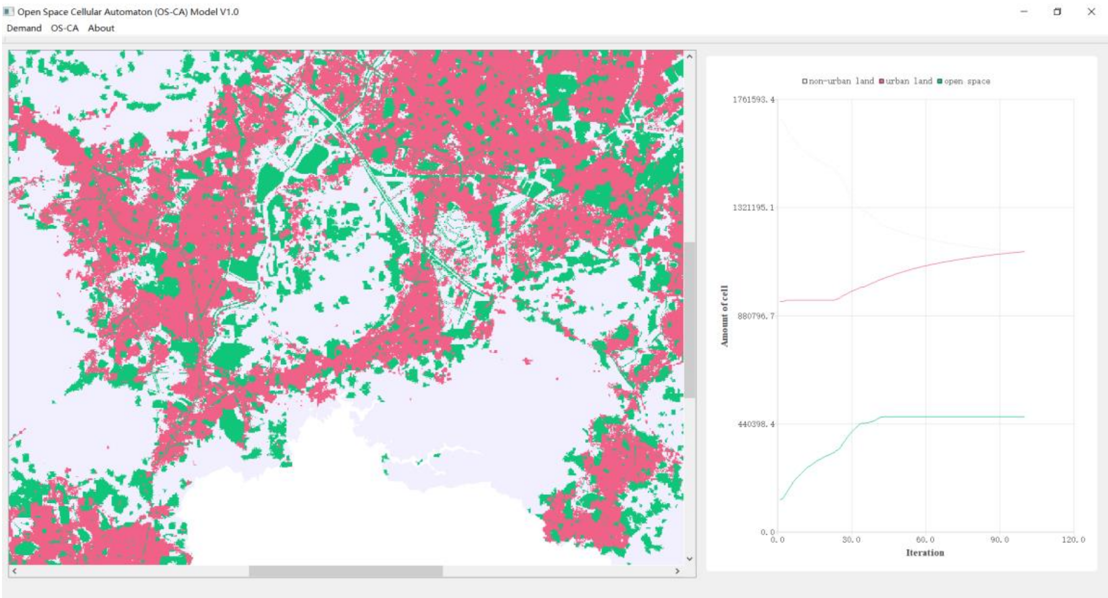


Fig. 22. The dynamic information of each iteration is shown on the left.



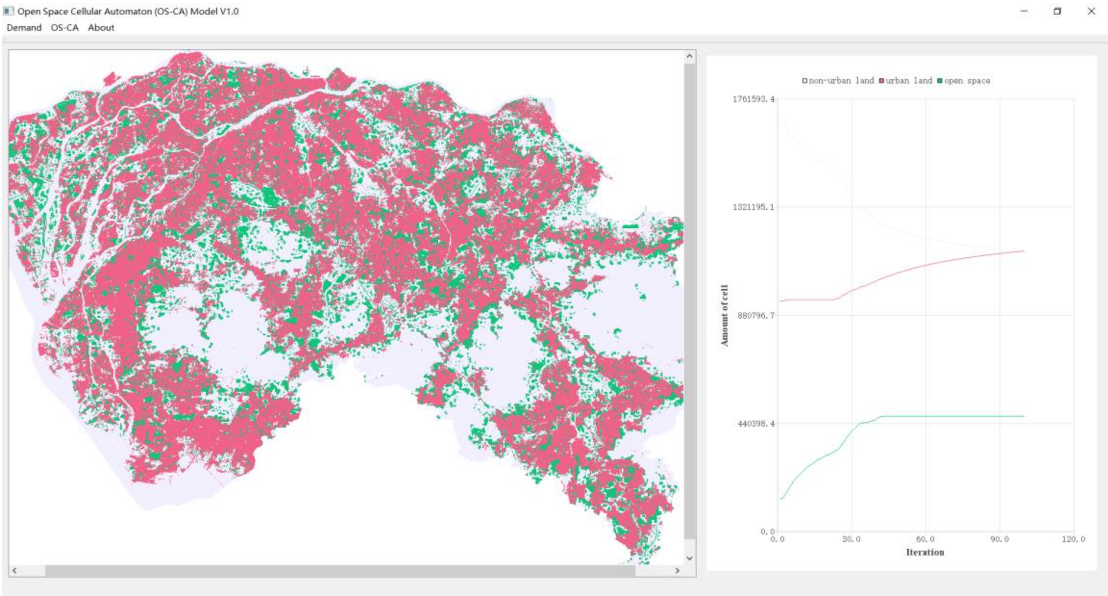


Fig. 23. The dynamic display of the simulation process on the main interface.

4. Some notes applying the OS-CA model

1) Mutiply **land use maps in all the modules** should have the same number of rows and columns.

2) In the step of simulating dynamics of land use change (CARS), 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.

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