GeoSOS-FLUS User's Manual:

A Future Land Use Simulation Model by coupling Human and

Natural Effects

Authors:

Xiaoping Liu

Xia Li

Xun Liang

Contact info:

Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou 510275, PR China

http://www.geosimulation.cn/GeoSOS/

Funding/Disclaimer:

Citation:

Abstract

Our newly proposed FLUS model is an integrated model for multi-type land use scenario simulation by coupling human and natural effects. At the same time, the spatial simulation module of FLUS model is made into GUI software named GeoSOS-FLUS. The GeoSOS-FLUS is developed as an extension of the previous GeoSOS software, to facilitate the multiple land use change simulation. The software provides a multiple CA allocation model for simulating land use change and scenario analysis.

The multiple CA allocation model is developed based on the theory of Cellular Automata (CA), but with several improvements over the traditional CA:

- The multiple CA allocation model is based on the analysis of the pattern of latest period land use instead of the change in two terms land use, this improvement not only avoid the effect of error accumulation from the disagreement of two terms land use data, but also make the model have much wider range of applications.

- An artificial neural network (ANN) is used for taking both human activities and natural ecological effects into consideration by finding the complex relationships between land use pattern and various human and natural driving forces.

- An elaborate self-adaptive inertia and competition mechanism is designed to address the competition and interactions among different land use types. The stochastic characteristics of this mechanism enable the model to reflect the uncertainty of real-world LUCC dynamics.

Besides, In order to ensure the speed of model, GeoSOS-FLUS was developed purely in the C++ language. The ANN technique in our model is from a powerful open source library called Shark 3.1.0 (http://image.diku.dk/shark/). The UI of the software is built using a famous open source library Qt 4.8.5 (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 1.9.2 (http://www.gdal.org/) allows our model to directly read and write raster data (.tif, .img, .txt files) that includes geographical coordinate information.

Therefore, the GeoSOS-FLUS is a powerful tool for making land use change simulation more convenience and efficient, which can be easily used for several purposes including 1) the establishment of urban construction boundaries (UCBs); 2) the high-resolution simulation of the change of internal land use change of the city; 3) environmental management and urban planning; 4) large-scale land use change and its effect on climate; 5) Regional land suitability analysis; 6) early warning for the lost of natural and agricultural land cover types; 7) hotspot recognition for land-use change and etc. When simulate future land use pattern, users need other external models, such as SD model or Markov chain, to project future land use demands as inputs of GeoSOS-FLUS.

This user's manual is intended to provide users of GeoSOS-FLUS a "quick start" on how to use the software. All of the necessary data and files for the tutorial have been provided and these can be used as templates for how to format your own files latter on.

1. Download and start GeoSOS-FLUS

The latest version of FLUS software and the user's manual are provided at a download link (http://www.geosimulation.cn/flus.html) from GeoSOS website. After you download the compressed file, unpack it and click the executable files "GeoSOS-FLUS VX.0.exe" to start the software.
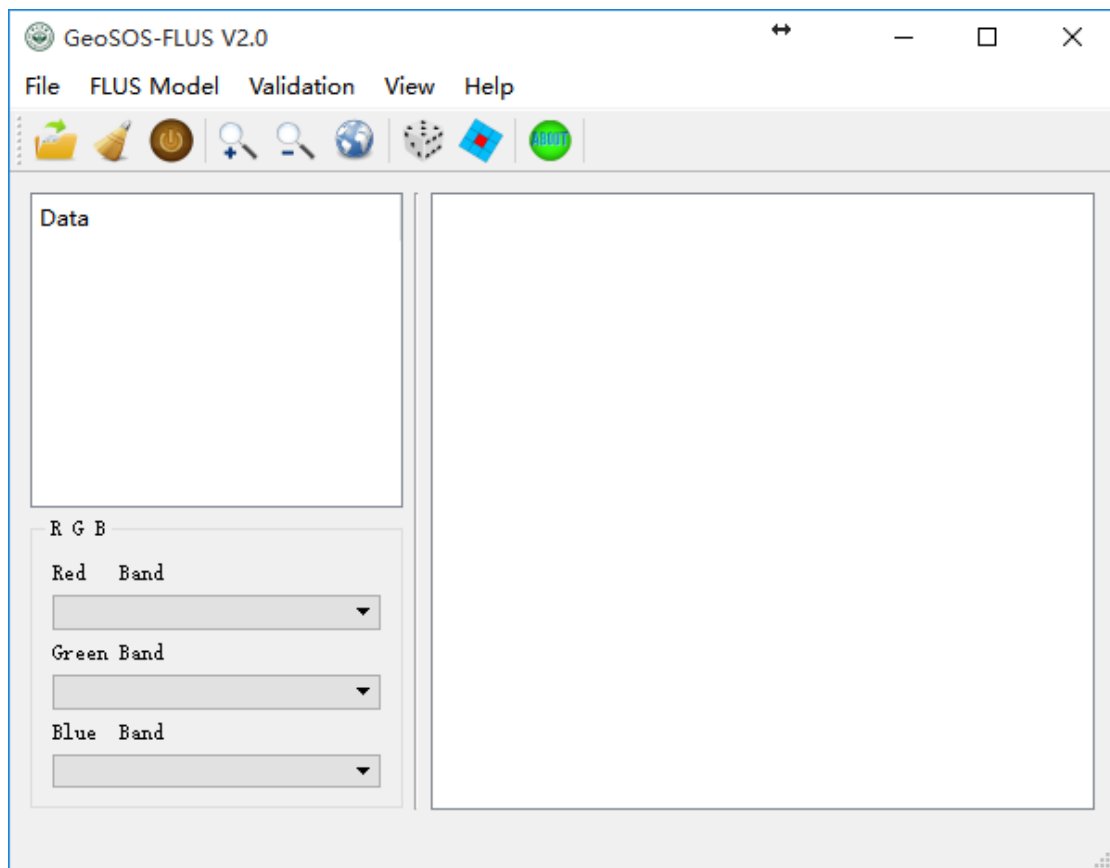
The main interface is a brief image browser which provide some basic function for image browse, including showing image in in grey scale, zoom in and zoom out, viewing image's properties.

2. Example data description

Once the compressed file "GeoSOS-FLUS VX.0.rar" has been inflated, the test dataset are found in the *test data* folder. Use this dataset to run example model or compared with your simulation results.

Category	File Name	Definition	Usage

Land use pattern	dg2001coor.tif	Land use classification data of 2001 in Dongguan, including five type of land use	Model input
	dg2006true.tif	Land use classification data of 2006 in Dongguan	Precision validation
Maps of restricted areas	restrictedarea.tif	Open water	Designating the areas where land use change is restricted
Driving factors	dem_dg.tif	Elevation data	Natural factors for calculating probability-of- occurrence
	slop.tif	Slopes	
	Aspect.tif	Slope directions	
	tocity_dg.tif	The distance to center of the city	Human factors for calculating probability-of- occurrence
	distotown.tif	The distance to center of the town	
	distohighway.tif	The distance to highway	
	distoroad.tif	The distance to road	
	distorailway.tif	The distance to railway	

In practical applications, more human-related and natural environment effects can be taken into consideration, such as：GDP、population，temperature、precipitation、soil types and etc.

3. User interfaces of FLUS model

The GeoSOS-FLUS consist of two main parts, 1) ANN-based probability-of-occurrence estimation module and 2) self-adaptive inertia and competition mechanism CA module. Users are allowed to start corresponding module by selecting option from the "FLUS Model" button's pull-down menu on the main menu.
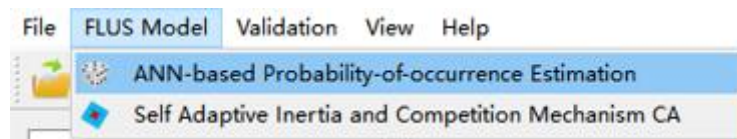
3.1. ANN-based Probability-of- occurrence Estimation

3.1.1. Start the module

Click the "ANN-based Probability-of- occurrence Estimation" item to start the module that has the same name as the item.
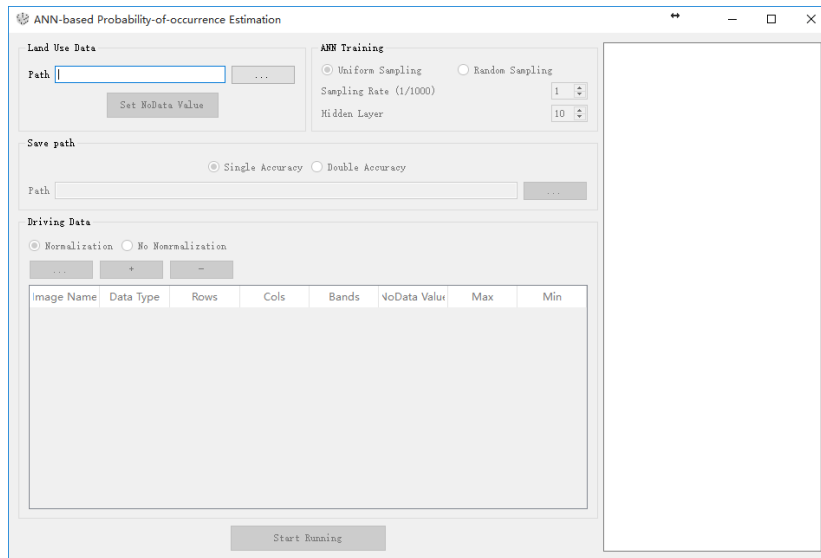
3.1.2. Input land use data

Click the button in the "Land Use Data" group box and select the start year land use data "dg2001coor.tif" in the pop-up dialog。


Then，Click the "**Set NoData Value**" button，and the software will open a window for defining the no data value of land use data. Under the second column, click on the drop-down menu to change it to "Valid Data" or "NoData Value". Users can only select one no data value at a simulation. Click "Accept" to accept the settings.
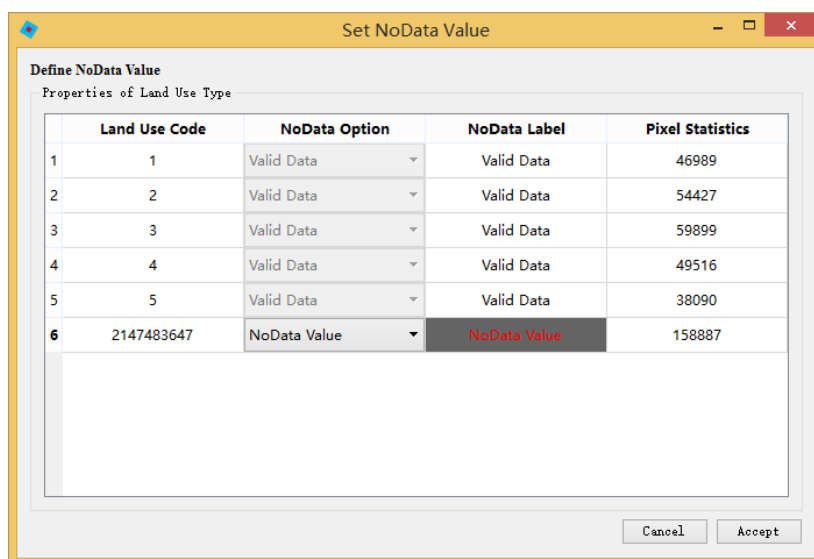
3.1.3. Setting the training parameters and output path

Users are allowed to set sampling strategy in the "A**NN Training**" group box. The "**Uniform Sampling**" option means the same numbers of sampling points for each land use type. The "**Random Sampling**" option means that the sampling points will randomly scatter across the study area. Then setting the sampling rate and the number of hidden layers of ANN in the following two spin boxes, respectively. The unit of sampling number is thousandths of the total number of valid pixels. When Sampling rate is set to 20, it means about two percent of valid pixels is selected for training. The number of hidden layer is set to 12 according to the expertise.
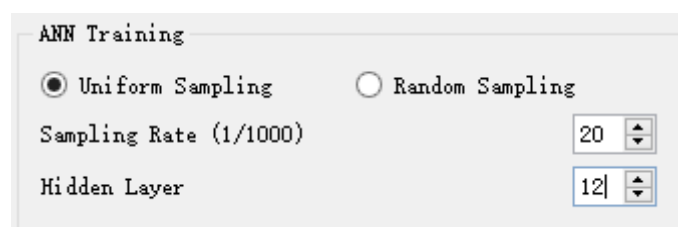
Then, click the [...   ] button in the "**Save path**" group box and input the save path

of probability-of- occurrence image in the pop-up dialog. The generated probability-of-occurrence file is named "Probability-of-occurrence.tif".
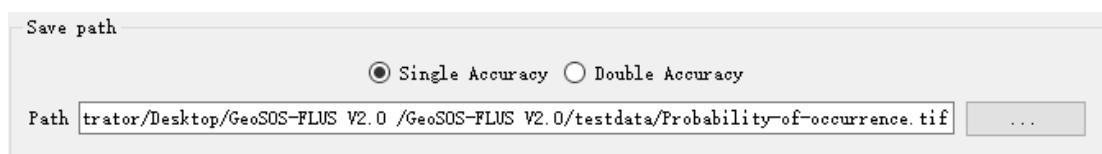
The"Save path" group box gives users the options to select the output data type. If users choose the "Single Accuracy" option (the default), the module would finally generate the FLOAT type probability-of- occurrence from TIF file. Similarly, another "Double

Accuracy" option will generate a DOUBLE type probability-of- occurrence TIF file. We usually adopt the"Single Accuracy" option, since it is able to achieve the ordinary precision requirement of simulation and saves lots of memory spaces, although the "Double Accuracy" option can generate higher accuracy data.

3.1.4. Input driving factors

Click the [...] button in the "**Driving Data**" group box and select multiple driving factors in the pop-up dialog. Our example data provide eight driving factors for example run of FLUS model.
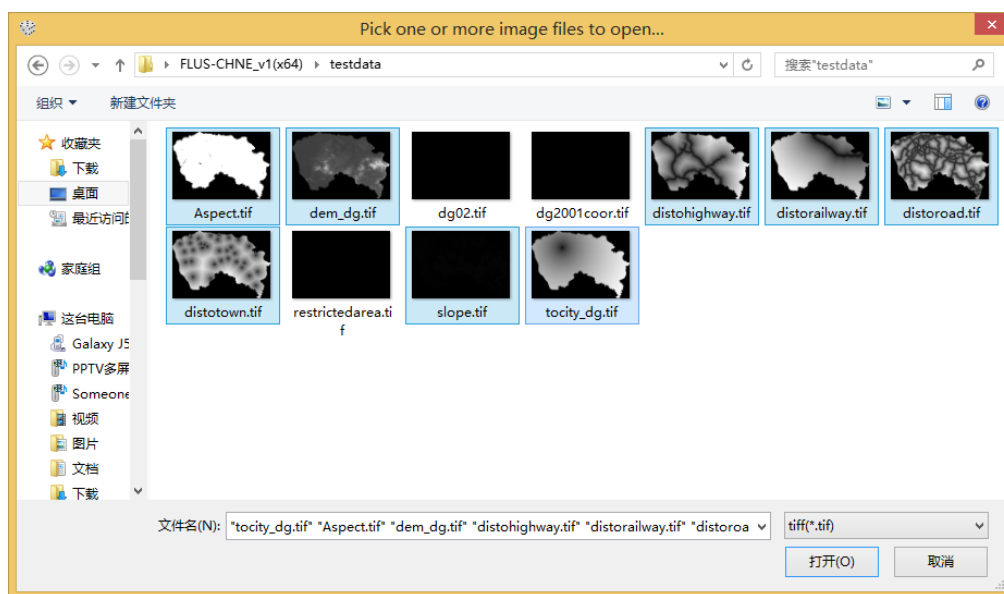
Attribute information of driving factors are displayed in a list box, including file name, data type, maximum value and maximum value, etc. Click button [+] for adding one new driving factor to list box, or click button [-] for removing one driving factor from list box.
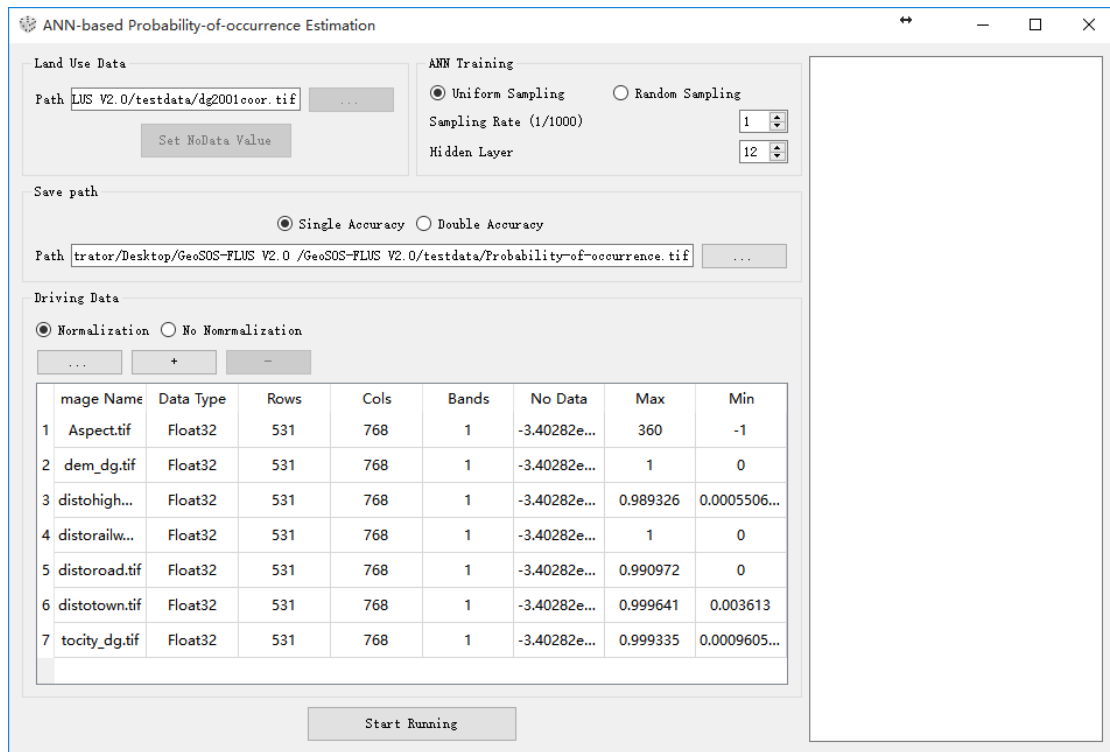
Besides，if users choose the "Normalization" option (the default)，all the driving factors will be normalized to a range of 0-1.0. If you driving factors have previously been normalized, you can choose "No Normalization" option to skip this step.

3.1.5. Running program and examine the result

Once the setup mentioned above has completed, click the button [Start Running] for running the module. There will come out a message box as below when the model process run is complete.
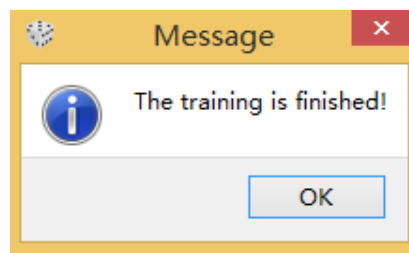

The text box at the right of the interface will output the training velocity and accuracy of the network after a successful run of ANN-based probability-of- occurrence estimation module. Three indicators are the RMSE, average error and average relative error, for measuring the training accuracy of the ANN, respectively.

```
Normalizing data, please wait...
Select uniform sampling...
Start training, please wait...
Run time: 54.984 s
Precision evaluation:
RMSE = 0.325524
Average error = 0.206084
Average relative error = 0.51521
Waiting for prediction...
The training is finished!
```

Users are able to use the image browser provided by GeoSOS-FLUS to examine the output result of ANN-based probability-of- occurrence estimation module. You can see a multi-band image, which each band corresponds to the probability-of- occurrence of one land use type.
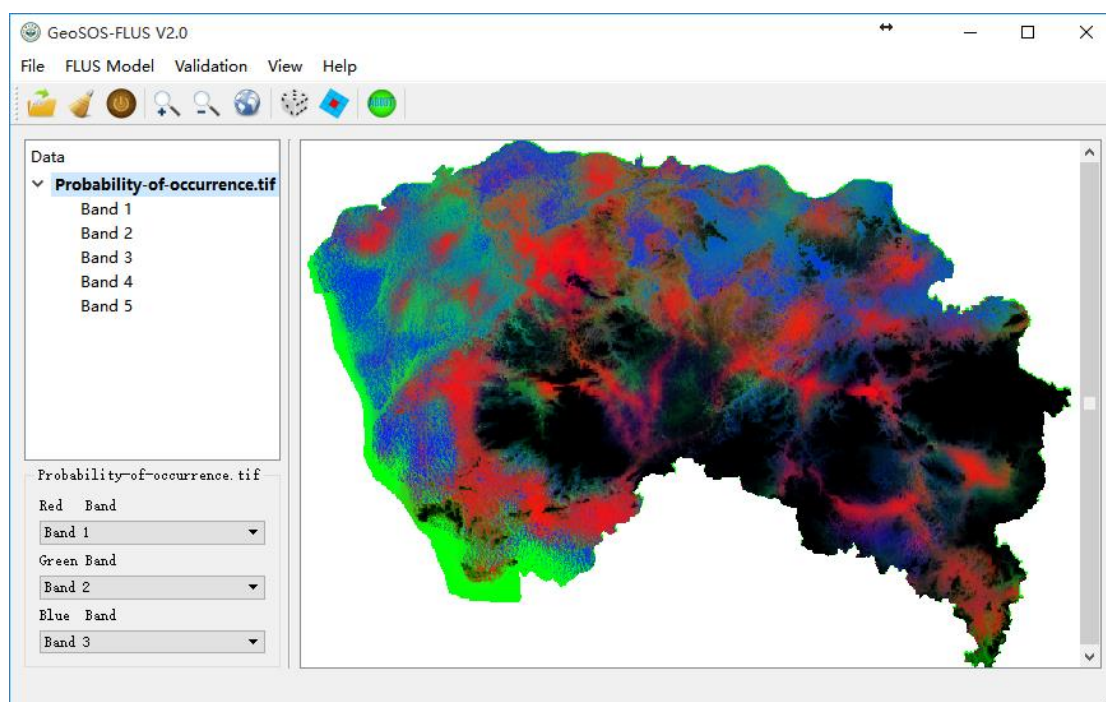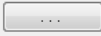
3.2. Self-adaptive inertia and competition mechanism CA

3.2.1. Start the module

Click the "Self-adaptive inertia and competition mechanism CA" item to start the self-adaptive inertia and competition mechanism CA for simulation.
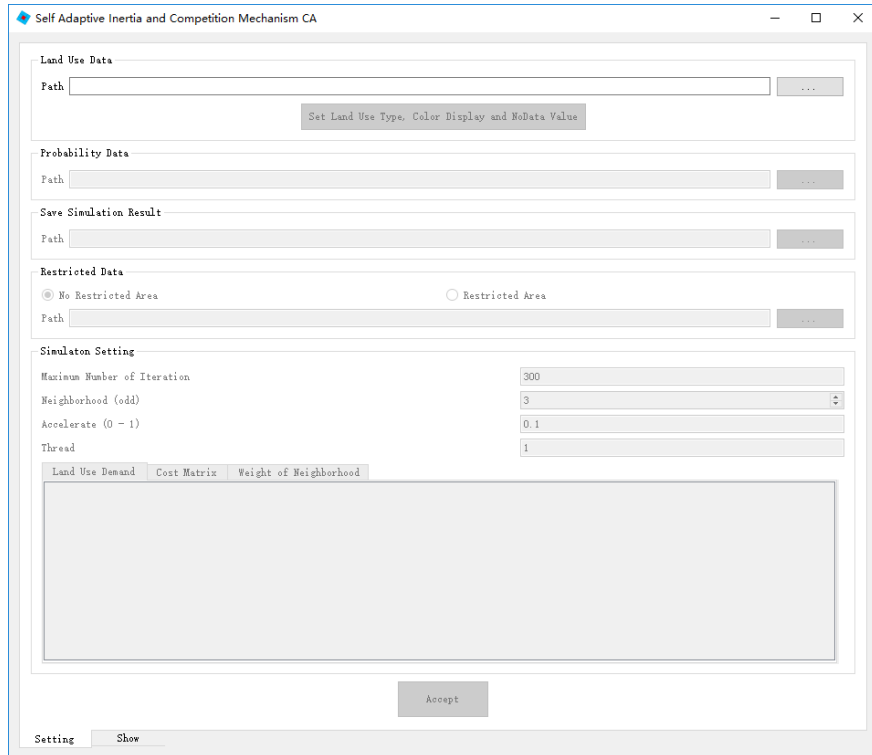
3.2.2. Input land use data

Click the [...] button in the "Land Use Data" group box and select the start year land use data "dg2001coor.tif" in the pop-up dialog。


Then，Click the "**Set Land Use Type, Color Display and NoData Value**" button，and the software will open a window for defining the no data value of land use data and setting display Color . The first column of window list shows the land use code of each land use type. Under the second column "**NoData Option**", users need to click on the drop-down menu of one of the rows to change it to "Valid Data" or "NoData Value" to define it as "NoData Value", only one no data value is allowed. The third column "**Land Use Type**" is for users to name the tag of each land use type. The GeoSOS-FLUS software will default the names of various land use types as "Landuse1", "Landuse2", etc.
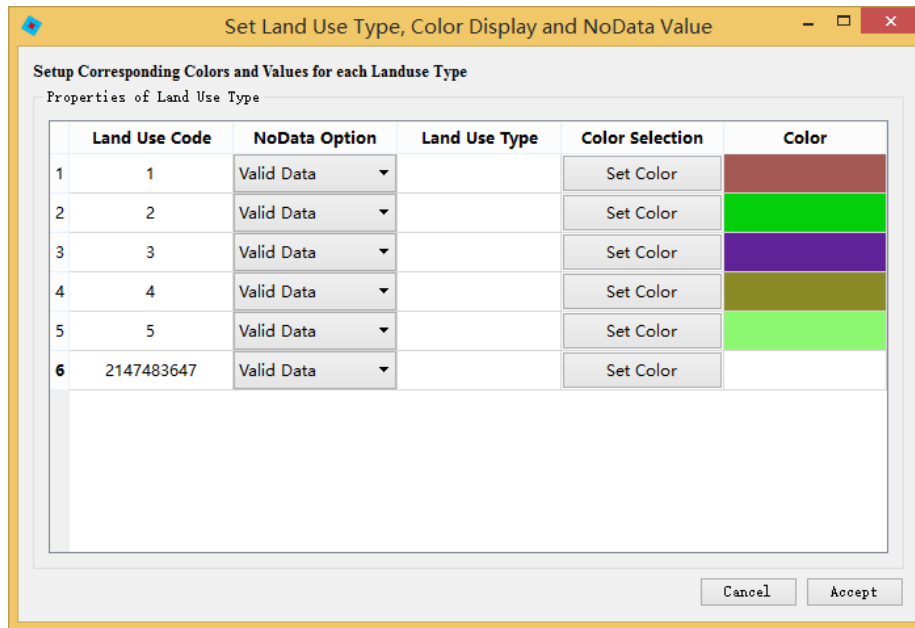
Click"**Set Color**"button of each row in the list box to set an RGB value (such as rgb(150,200,50)) for all land use types.
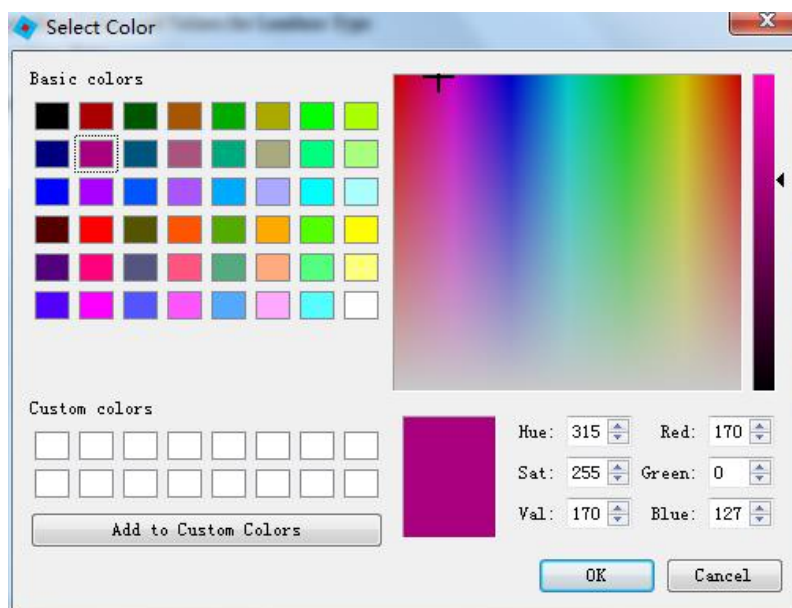
Finally, click the "**Accept**" button, confirm the above settings.
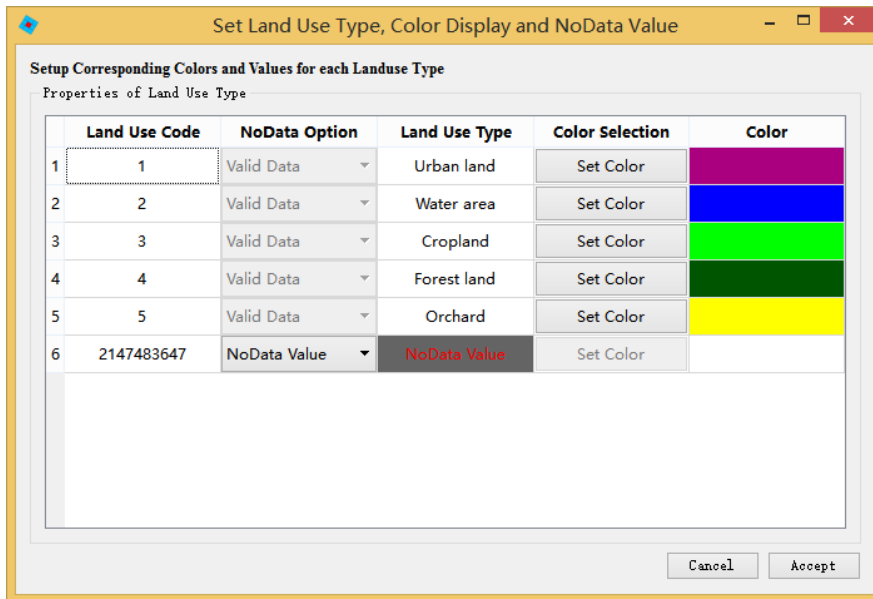
3.2.3. Input probability-of- occurrence data

Input the probability-of- occurrence file output from the ANN-based Probability-of-occurrence Estimation module in the "Probability Data" group box.

3.2.4. Set the path of saving simulation result

Input the save path of simulation result in the "Save Simulation Result" group box.

3.2.5. Area restrictions (Spatial policies)

Some spatial policies restrict all land use change in designated areas, such as open water and some strictly protected nature reserves area. When spatial policies need to be considered into simulation, users should prepare a binary restricted area image that all the pixel are either 0 or 1. The value 0 means that grid cells with one land use type can never be converted to another type; 1 means that land use change is allowed in this region.

3.2.6. Setting the simulation parameters

Simulation parameters mainly include the land use demands, number of iterations, neighborhood influence, neighborhood weights and acceleration factor. The land use demands are parameters of our model, which need to be firstly set according to the actual situation of study region. So users need to firstly use external model to determine the land use demands of future scenario. For our example run, the time span of the land use change simulation is from 2001 to 2006, so the land use demands are set to be the same as the actual area of various land use types in 2006. The number of iterations is set to a large value (maximum number of iteration, default is 300), the model will stop when the allocated area equal the demanded area for all land use types or reach the maximum number of iteration. The neighborhood effect of GeoSOS-FLUS is similar to traditional CA, the default value is 3. The acceleration factor is a parameter for improving the speed of reaching land use demands, ranging from 0 and 1.0. When simulating land use change for large area, users may need to set a larger value of acceleration factor for shortening the time consumption in simulation process.

Simulaton Setting

Maximum Number of Iteration	300
Neighborhood (odd)	3
Accelerate (0 – 1)	0.1
Thread	8

Land Use Demand | Cost Matrix | Weight of Neighborhood

	Urban land	Water area	Cropland	Forest land	Orchard
Initial Pixel Number	46989	54427	59899	49516	38090
Future Pixel Number	80016	54427	43599	42433	28446

In the tab widget, the second row "Future Pixel Number" in the "Land Use Demand" page is for inputting the future land use demand of each land use type. The initial pixel number will be automatically loaded after land use pattern data is input. In FLUS 2.4, the users can set the thread count to speed up the FLUS model. We set the thread count to 8 in this tutorial.

Land Use Demand | Cost Matrix | Weight of Neighborhood

	Urban land	Water area	Cropland	Forest land	Orchard
Initial Pixel Number	46989	54427	59899	49516	38090
Future Pixel Number	80016	54427	43599	42433	28446

Then, switch the tab widget to the "Cost 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 Cost Matrix is a collection of a series of prior knowledge of experimenter about study area. An example of cost matrix is provided below:

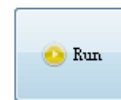|	Urban land	Water area	Cropland	Forest land	Orchard
Urban land	1	0	0	0	0
Water area	0	1	0	0	0
Cropland	1	1	1	1	0
Forest land	1	0	1	1	0
Orchard	1	0	1	0	1

And in the "Weight of Neighborhood" page, users need to set the neighborhood weight of different land use types, because the neighborhood effects may be different for different land use types in a study region. The value of the neighborhood weight 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:

	Urban land	Water area	Cropland	Forest land	Orchard
Weight of neighborhood	1	0.9	0.5	1	0.1

Finally, click "Accept" for confirming what you have set up before.

3.2.7. Running the multiple CA allocation model
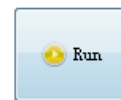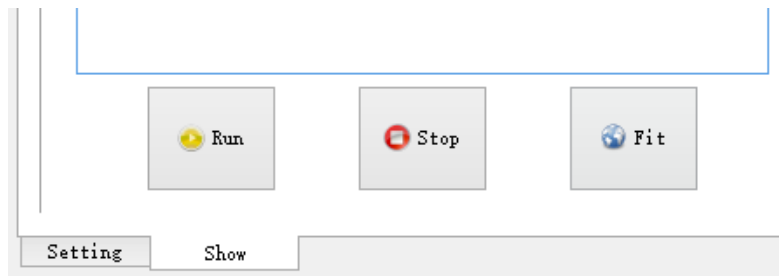
Switch the tab widget to the "Show" page, click button [Run] to start the simulation. The GeoSOS-FLUS is able to show the process of land use change dynamically, including the spatial change, change in value and change curve of each land use type.
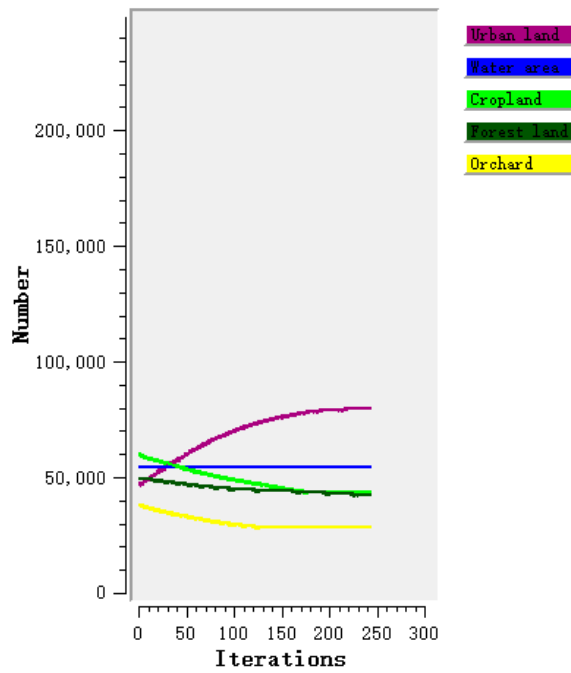
```
Maximum Number Of Iterations: 300
Neighborhood influence: 3
Acceleration for iterate: 0.1
Cost Matrix
1 0 0 0 0
0 1 0 0 0
1 1 1 1 0
1 0 1 1 0
1 0 1 0 1
Future year


————————————Start to iterate————————————

Land use type, Urban land, Water area, Cropland, Forest
land, Orchard
Number of pixels of each land use before
iteration, 46989, 54427, 59899, 49516, 38090
Number of pixels of each land use at 1
iteration, 47468, 54427, 59532, 49533, 37961
Number of pixels of each land use at 2
iteration, 47915, 54427, 59261, 49506, 37812
Number of pixels of each land use at 3
iteration, 48316, 54427, 59058, 49468, 37652
```

The model will stop and save the simulation result when the allocated area equal the demanded area for all land use types or reach the maximum number of iteration. Besides, user can manually stop the simulation process and save the result by clicking the

 button.

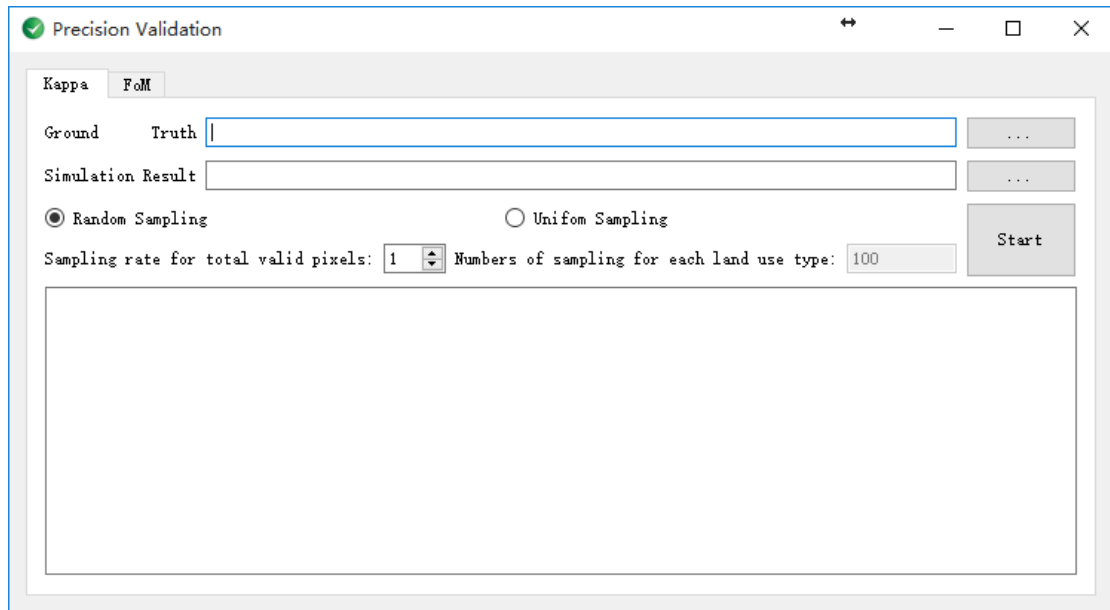## 3.3. Precision validation of simulation

GeoSOS-FLUS provides a kappa statistic tool and a FoM statistic tool for measuring agreement between simulation result and actual land use pattern.
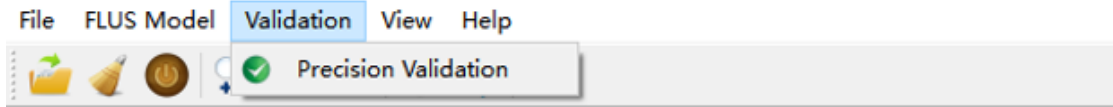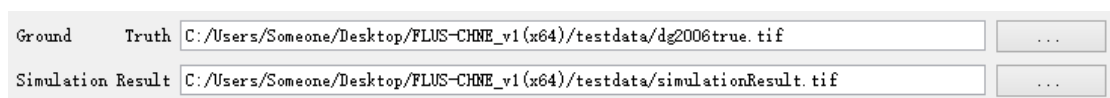

3.3.1. Kappa statistic tool

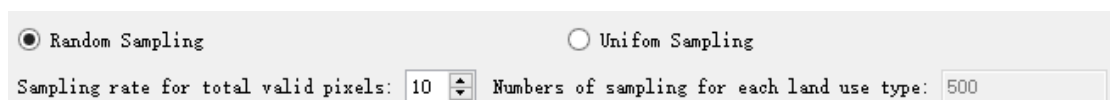Click the **Validation → Precision Validation→Kappa** item on the main menu to start the kappa statistic tool.

3.3.1.1. Load data

Click the [...] button in the "Ground Truth" row，to input the actual land use pattern named "dg2006ture.tif" in 2006. Then click the [...] button in the"Simulation Result"row to input the simulated land use pattern.
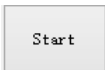

3.3.1.2. Select a strategy for sampling

GeoSOS-FLUS supports two strategies for sampling: Random Sampling and Uniform Sampling. The "**Uniform Sampling**" option means the same numbers of sampling points for each land use type. The "**Random Sampling**" option means that the sampling points will randomly scatter across the study area. At this example run, we choose "**Random Sampling**" strategy and set the sampling rate to 10%.
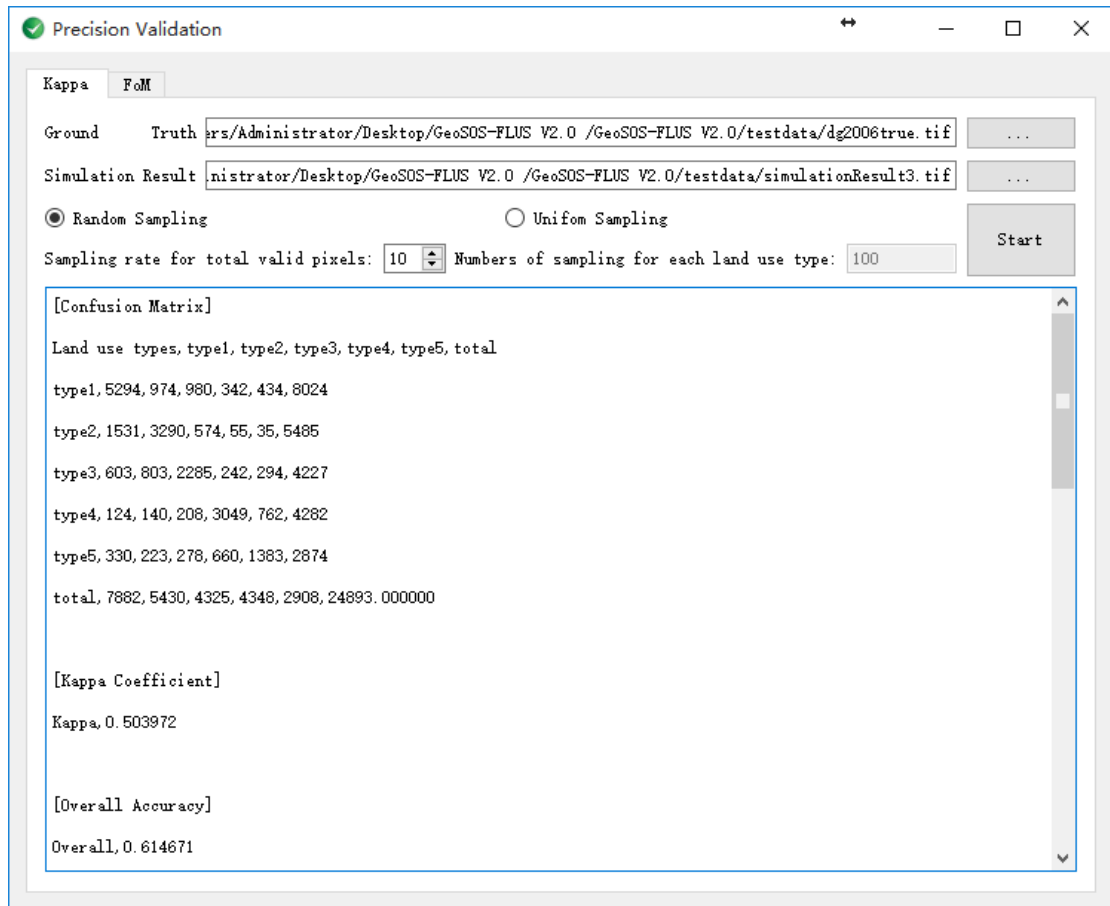

3.3.1.3. Start calculating and save the result

Click the [Start] button to start calculating Kappa coefficient. A message box will pop up when the program completes the calculation.
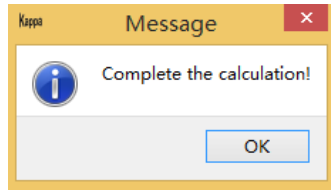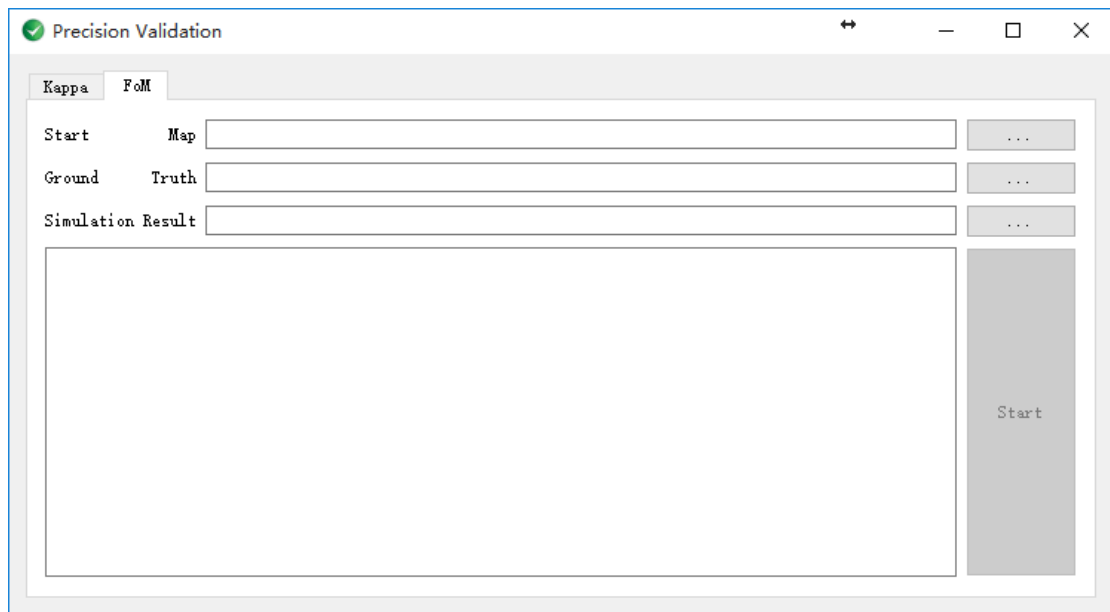
The results include Kappa coefficient, overall accuracy and confusion matrix, which will be saved in a file named "Kappa.csv".

3.3.2. FoM statistic tool

Click the **Validation → Precision Validation→FoM** item on the main menu to start the kappa statistic tool.

Click the [...] button in the "Start Map" row，to input the start land use pattern named "dg2001coor.tif" in 2001. Click the [...] button in the "Ground Truth" row，to input the actual land use pattern named "dg2006ture.tif" in 2006. Then click the [...] button in the"Simulation Result"row to input the simulated land use pattern.

Click the [Start] button to start calculating FoM coefficient. A message box will pop up when the program completes the calculation.
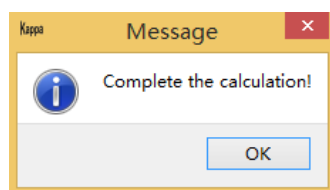

The results include FoM statistics, producer's accuracy, which will be saved in a file named "FoM.csv".
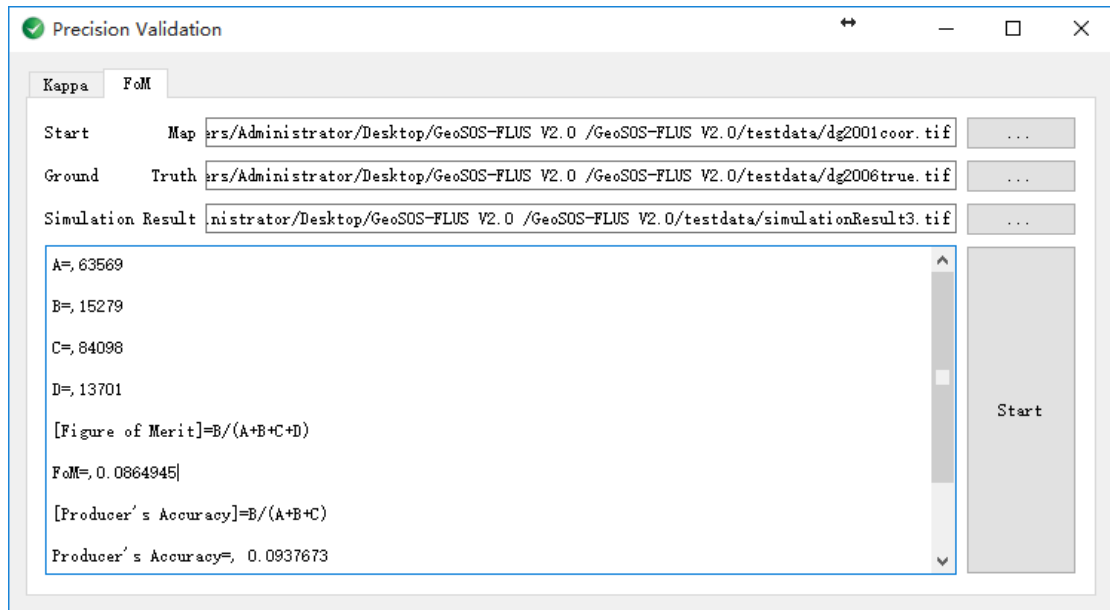
4. Some notes in using GeoSOS-FLUS

1) The valid land use codes of land use data must be a tolerance of 1 arithmetic progression beginning with 1(e.g. 1,2,3,4,5,…). NoData Value can be any other value except these valid land use codes.

2) 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-of-occurrence data.

3) The Self-adaptive inertia and competition mechanism CA module will save all the parameters to two configuration files: "config_color.log", "config_mp.log". The module will automatically loads the parameters that are recorded in the configuration files including the information about show colors and names of land use types, which is convenient for the user to use to repeat their experiments. But users' must firstly delete these two configuration files when changing to a new study region, for resetting simulation parameters.

5. Generated files

After simulation, GeoSOS-FLUS creates two kinds of files: configuration files and recorded files.

5.1. Configuration files

config_color.log: Record the RGB value, name of land use types, number of pixels of each land sue types in base year in last simulation.

```
1   [Index, Count, Land Use Type, R, G, B]
2   1,46989,Urban land,170,0,127
3   2,54427,Water area,0,0,255
4   3,59899,Cropland,0,255,0
5   4,49516,Forest land,0,85,0
6   5,38090,Orchard,255,255,0
```

config_mp.log: Record the land use demands, cost matrix, maximum number of iteration, acceleration factor, neighborhood weights of last simulation.

```
1    [Number of types]
2    5
3    [Future Pixels]
4    80016
5    54427
6    43599
7    42433
8    28446
9    [Cost Matrix]
10   1,0,0,0,0
11   0,1,0,0,0
12   1,1,1,1,0
13   1,0,1,1,0
14   1,0,1,0,1
15   [Intensity of neighborhood]
16   1
17   0.9
18   0.5
19   1
20   0.1
21   [Maximum Number Of Iterations]
22   300
23   [Size of neighborhood]
24   3
25   [Accelerated factor]
26   0.1
```

5.2. Recorded files

NetworkInput.csv: Training data of ANN.

ANNoutput.log: Record the training accuracy of ANN.

logFileTrain.log: The input and output data paths of ANN-based Probability-of-occurrence Estimation module.

output.log: Save the number of pixels of each land use type in every iteration.

logFileSimulation.log: The input and output data paths of self-adaptive inertia and competition mechanism CA module.

Kappa.csv: Save the results of accuracy evaluation.

FoM.csv: Save the results of accuracy evaluation.