MOLUSCE
Modules for Land Use Change Evaluation

Quick Help

(MOLUSCE ver. 3.)*

ASIA AIR SURVEY

NEXTGIS OPENSOURCE GEOSPATIAL SOLUTIONS

Table of Contents

1. Plug-in overview

1.1. Introduction

Open source software platforms are progressively becoming widely used in the public and private sectors. In Geographical Information System (GIS), open source software packages such as QGIS are actively being developed. More importantly, customization and further development is possible since developers create specific plug-ins with flexibility.

Asia Air Survey Co., Ltd. (AAS) started to move towards open source software since 2012 becoming the first QGIS gold sponsor worldwide. Furthermore, open source software started to be used more extensively for internal use and in international project. Alongside with these recent changes AAS also started to develop open source solutions aiming to further extend its market.

1.2. What is MOLUSCE?

AAS released MOLUSCE (Modules for Land Use Change Evaluation) at FOSS4G 2013. MOLUSCE is a user-friendly plug-in for QGIS 2.0 and above. MOLUSE is designed to analyse, model and simulate land use/cover changes. The plug-in incorporates well-known algorithms, which can be used in land use/cover change analysis, urban analysis as well as forestry applications and projects.

MOLUSCE is well suited to:

- analyse land use and forest cover changes between different time periods;
- model land use/cover transition potential or areas at risk of deforestation; and
- simulate future land use and forest cover changes

1.3. Functions

MOLUSCE user interface offers an easy-to-use interface with specific modules and functions. Following is a brief description of basic modules in MOLUSCE.

Input module

Land use/cover maps from different epochs, biophysical and socio-economic driving factor data such as road network, rivers, topography, population *etc.,* are loaded in the input module.

Area change analysis

Computes land use/cover changes between two time periods (T1 and T2). Land use/cover change transition matrices as well as land use change maps are produced.

Modelling methods

Four methods, namely Artificial Neural Networks (ANN), Logistic Regression (LR), Multi-Criteria Evaluation (MCE) and Weights of Evidence (WoE) are used for modelling land use/cover change transition potential.

Simulation

Displays transition potential maps, certainty function (experimental) and simulation results. A simulated (projected) land use/cover map is produced based on a Monte Carlo Cellular-automata modelling approach.

Validation

This sub-module incorporates kappa statistics (standard kappa, kappa histogram and kappa location), which will be used to validate the accuracy of the simulated land use/cover maps.

2. How to use MOLUSCE

2.1. *Inputs*

Data can be loaded using the inputs tab.

❶ All raster files are loaded in the inputs tab

Load the initial and final land use/cover maps as shown in steps ❷ and ❸. The base map determines the geometry of all the output files, pixel size, scaling and projection.

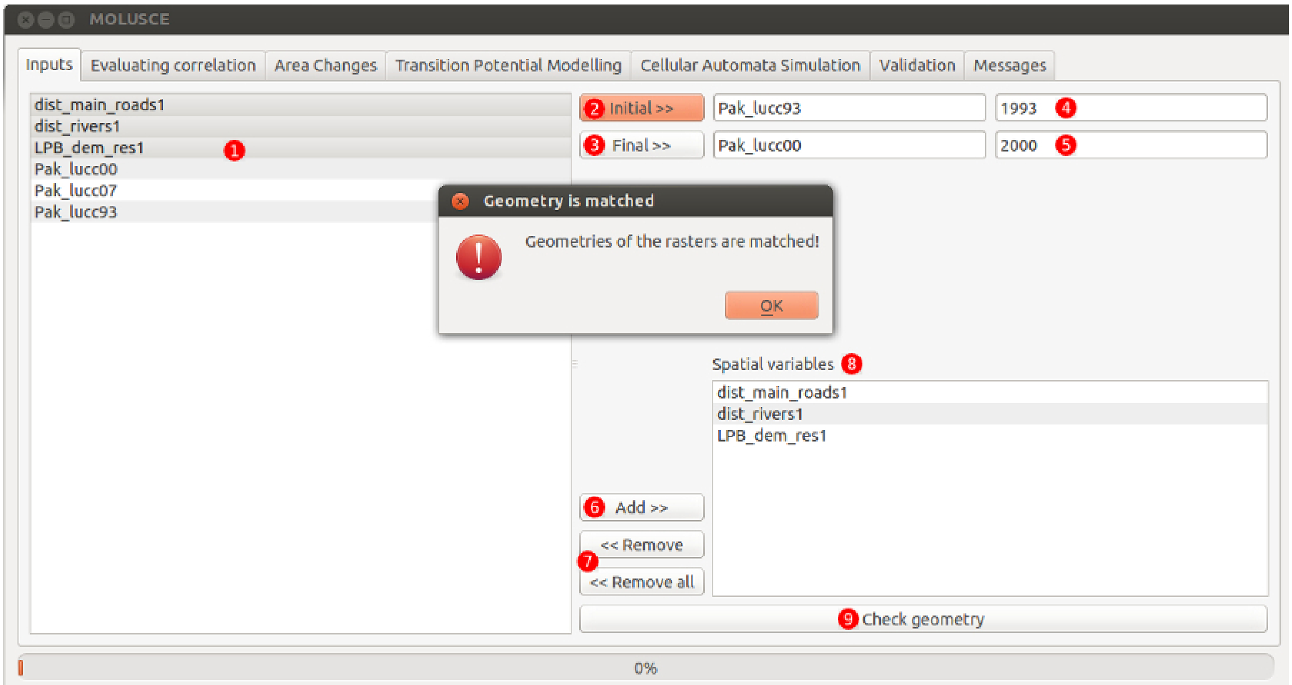
Steps ❹ and ❺ show the corresponding years. (the user can type here If the corresponding years do not appear automatically).

Steps ❻ and ❼ show the add/remove buttons used to add or remove spatial variables.

The selected spatial variables are shown in step ❽

The check geometry button is a mandatory step to check if the geometry of the selected raster is matched (step ❾).

Note: It is important when layers are added to QGIS that the No Data Value (NDV) is set. If this is not done, MOLUSCE will process NDV areas as land use/cover classes, increasing processing time and confusing the model calibration. MOLUSCE picks up the NDV of the input (base) layer and propagates it to any output maps generated, along with the geometry of the base layer.



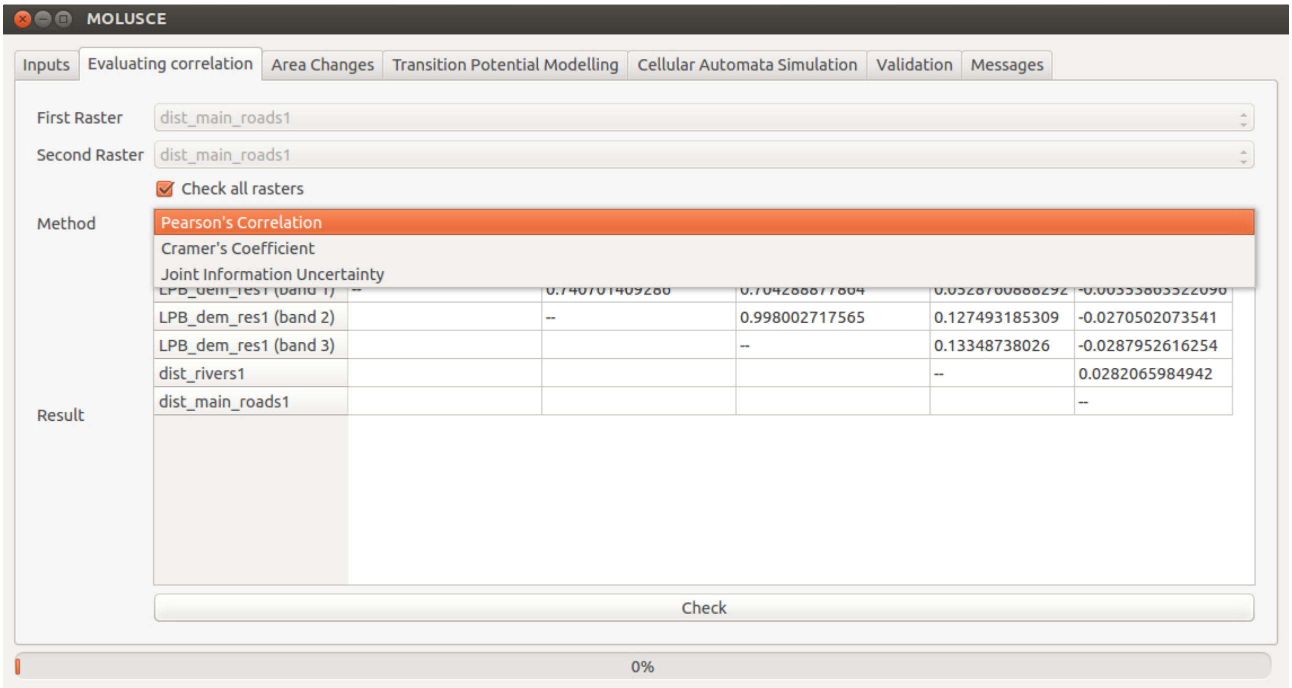
2.2. *Evaluating correlation*

The evaluating correlation module contains three techniques for performing correlation analysis:

1. Pearson's correlation
2. Cramer's coefficient
3. Joint information uncertainty

The user can choose between a two-way raster comparison by selecting first raster and second raster or check all rasters loaded into MOLUSCE.
The user can run correlation by pressing the check button located at the bottom of the window.

Note: The Cramer's coefficient and joint information uncertainty work only with categorical data. The data should be converted to categorical data (eg. using GRASS).



2.3. *Area Changes*

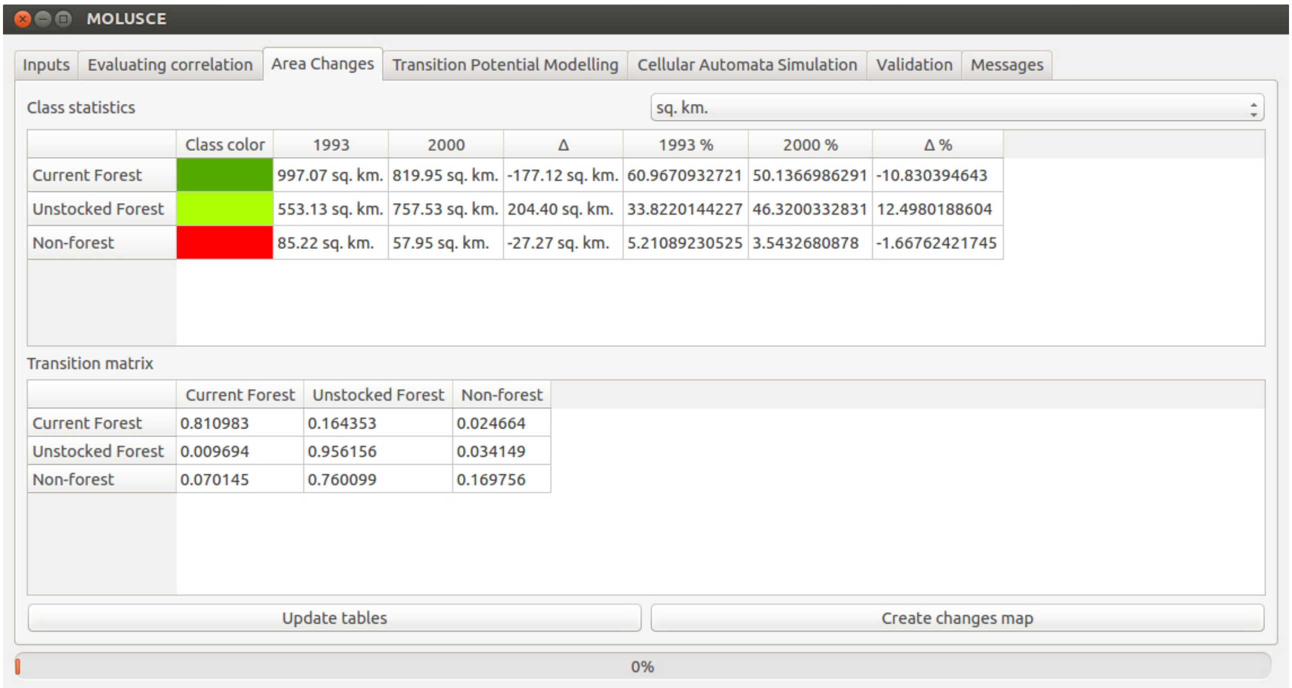
The update tables button produces class statistics and transition matrix tables.

The class statistics table shows the initial and final land use/cover change (LUC) areas.

The transition matrix shows the proportions of pixels changing from one land use/cover to another.

The create change map button will generate a map of change classes. This will be added automatically to QGIS and saved as a GeoTiff.

Note: Data from tables can be copied and pasted directly into spreadsheet programs, simply by selecting the desired rows/columns and by pressing the "Ctrl + C" keyboard combination.



2.4. Transition potential modelling

MOLUSCE uses Artificial Neural Network (ANN), Multi Criteria Evaluation (MCE), Weights of Evidence (WoE) and Logistic Regression (LR) methods to model land use/cover transition potential. The user can select a method from the drop down menu.

2.4.1. Artificial Neural Network (ANN)

The define samples function specify number of samples and sampling mode. In addition, the sampling points created can be saved and displayed.

Five inputs are used to customize the ANN modelling:

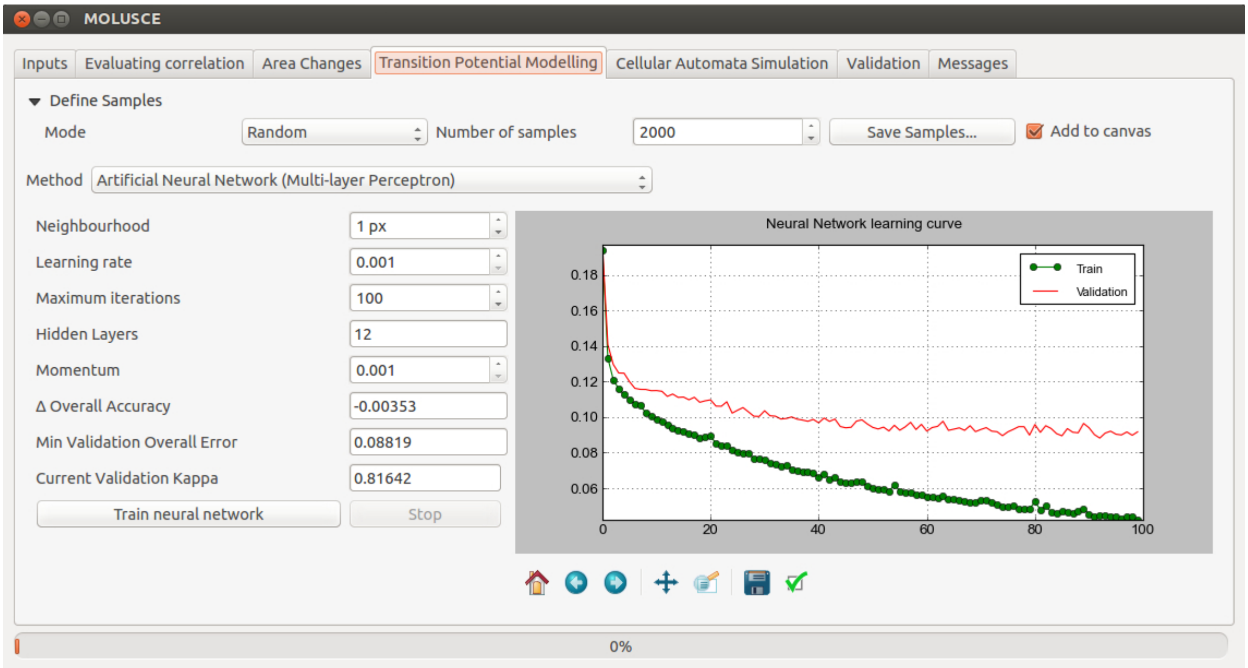
- The neighbourhood defines count of neighbour pixels around current pixel. Size=1 means 9 pixels (3x3 region), size=2 means 25 pixels (5x5), etc.
- Learning rate, momentum and max iterations number define parameters of learning. Big learning rate and momentum allow fast learning, but the learning process can be unstable (spikes on the graph). Small learning rate and momentum means stable but slow learning.
- Hidden layers input string takes a list of numbers: N^1 N^2 … N^k, where N^1 is number of neurons in 1st hidden layer, N^2 is number of neurons in 2nd hidden layer and so on, N^k is the number of neurons of the last hidden layer (k^{th} layer). For example if the user types in the input string "2" then a network with 1 hidden layer and 2 neurons will be created. In order to create a network with 2 hidden layers the user should insert 2 numbers, such as "10 2" which will create a network with 10 neurons in the first hidden layer and 2 neurons in the second.

The following outputs are proposed (for the current learning iteration):

- The graph area. Contains errors of training and validation sets. It is the main information about learning process. The graph can be edited and saved as image.
- The min validation overall error contains information about min reached error on validation set of samples.
- The delta overall accuracy contains difference between min reached error and current error.
- The current validation kappa shows the kappa value.

The process can be started by pressing on the train neural network button and stopped at any time using the stop button.

Learning algorithm analyses the reached accuracy on training and validation sets of samples and stores the best neural net in memory. The training process finishes when the best accuracy is reached.



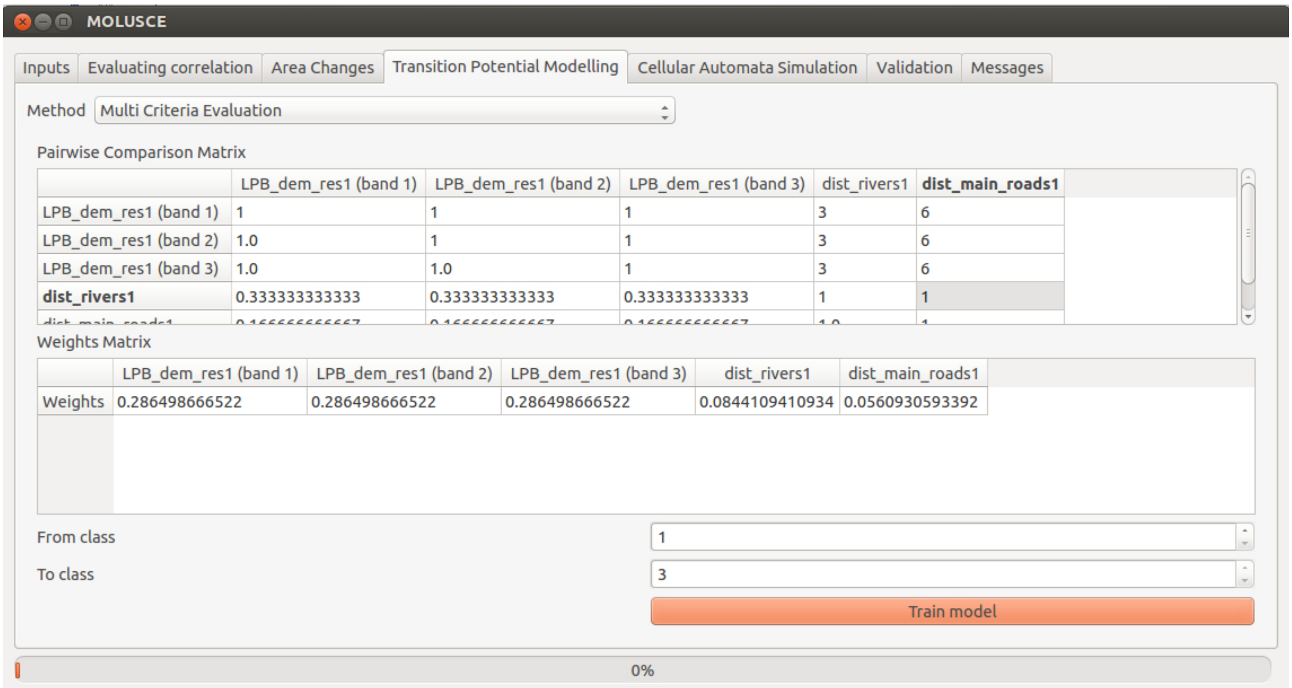
2.4.2. Multi Criteria Evaluation (MCE)

The user can set the values inside the pairwise comparison matrix.

The user can also select which classes to use to train the model by changing the from class and to class values located at the bottom of the window.

The model can be started by pressing the train model button.

The weights of the spatial variable will then appear in the weights matrix.

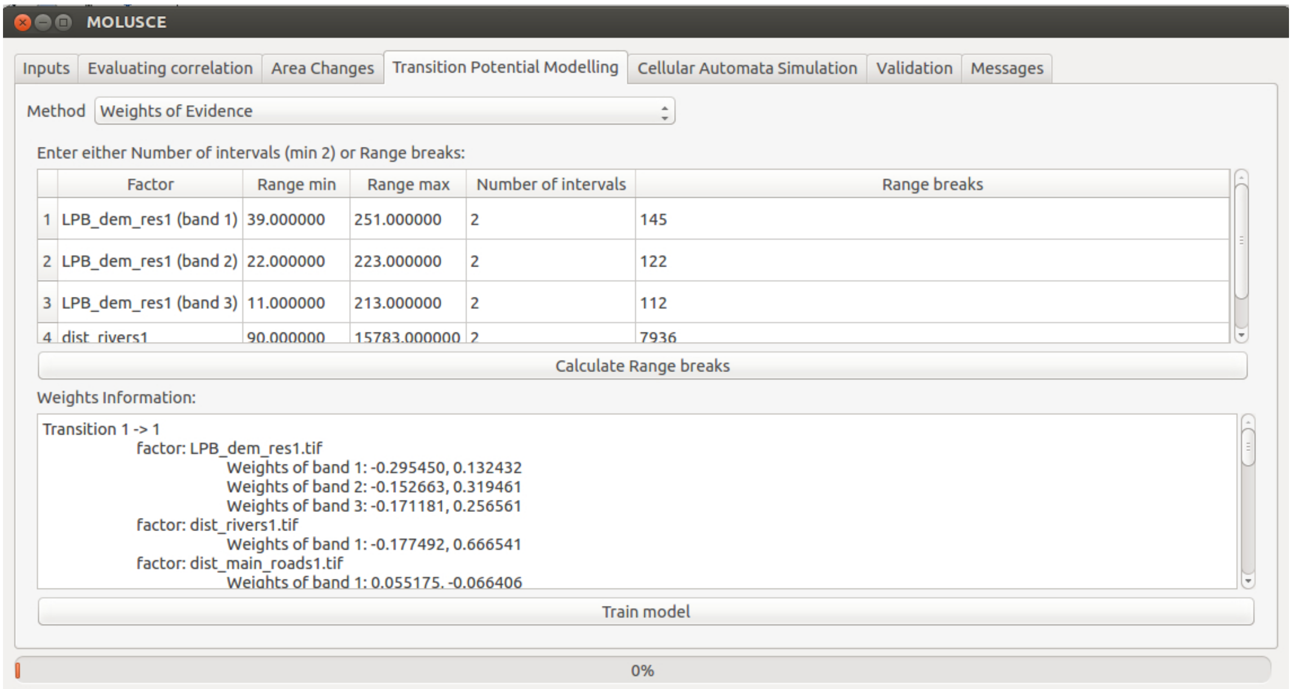


2.4.3. Weights of Evidence (WoE)

The WoE method proposes two ways of defining the range breaks. The user can define a <u>number of intervals</u> or specify the <u>range breaks</u> values.

When the <u>calculate range breaks</u> button is pressed the weights information for each transition are produced.

Once the user is satisfied with the weights the <u>train model</u> button can then be pressed.



2.4.4. Logistic Regression (LR)

The LR method offers the possibility to define samples. (number of samples and sampling mode) as well as save and display the sampling points created.

Two inputs are used to customize the LR modelling:

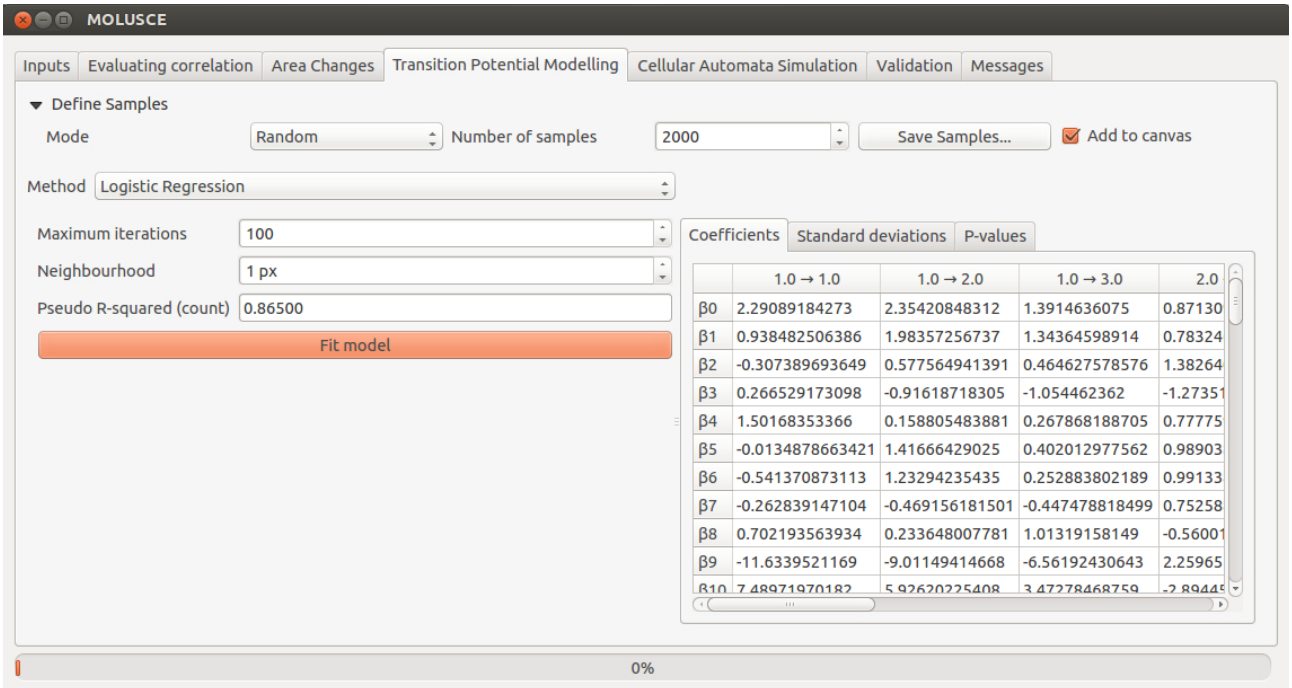
- The maximum iteration defines the total number of iterations.
- The neighbourhood defines count of neighbour pixels around current pixel. Size=1 means 9 pixels (3x3 region), size=2 means 25 pixels (5x5), etc.

The following outputs are proposed (for the current learning iteration):

- The pseudo R-squared shows the goodness-of-fit
- The coefficients tab
- The standard deviations tab
- The p-values tab

The user can run the model by pressing on the fit model button.

Note: For additional information on the LR outputs please consult the “Technical information: Methods and Algorithms” manual.

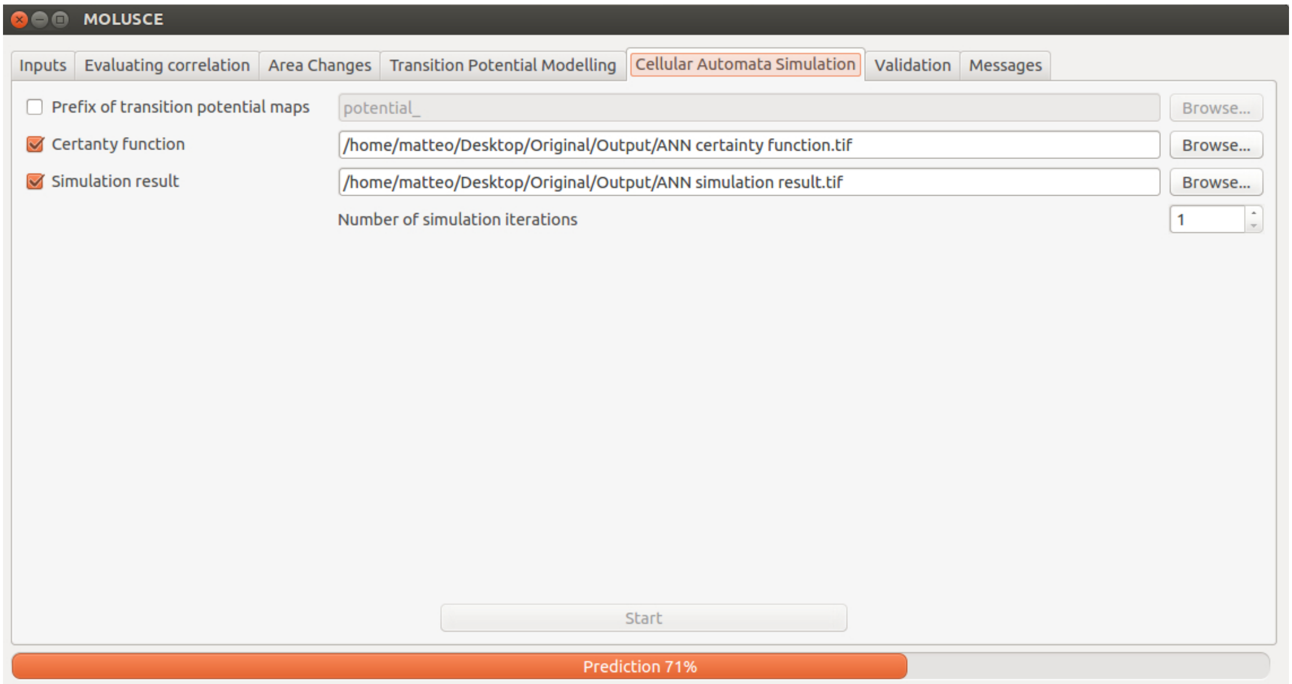


2.5. Cellular Automata Simulation

Once one method has been chosen from the transition potential modelling tab, the user can then access to the cellular automata simulation tab. Be aware that MOLUSCE will keep in memory the latest method processed, if for example the user runs first the ANN and then the LR methods, the cellular automata simulation tab will retain the results from the LR.

Three type of output maps are produced. A check box at the beginning of each output is provided to allow the user to enable only what it is necessary. A browse... button located at the end of each output allows to save each map.

- The prefix of transition potential maps button allows to select prefix of the names of transition potential maps. Transition potential map shows the probability or potential to change from one land use/cover class to another. Transition potential values range from 0 (low transition potential of change) 100 (high transition potential). Transition potential maps will be produced from the corresponding land use/cover changes (e.g., "forest to unstocked forest" transition potential, "forest to non-forest transition potential).
- The certainty function (As mentioned earlier, this is an experimental function).
- The simulation result produces a simulated land use/cover map.



2.6. Validation

The validation tab allows the user to check, validate and compare the simulation results. Reference and simulated land use/cover maps must be loaded in order to start the validation process. The former indicates a land use/cover map (T^3). T^1 refers to the initial land use/cover, while T^2 refers to the final land use/cover used in the model). A browse. button located at the end of each output allows to load the desired map. A two way map comparison is performed from reference land use/cover (T3) and simulated land use/cover maps.

In order to perform a three way map comparison, the user can check the risk class validation map check-box. Although it is not shown explicitly, the three-way map comparison uses the initial land use/cover map (T1), the reference land use/cover map (T3) and the simulated land use/cover map.

The multiple-resolution budget method is also used in MOLUSCE. The user can decide the number of validation iterations and can start the process by pressing the start validation button. The graph can be edited and saved as image.

the overall accuracy (% of correctness), kappa (overall), kappa (histo) and kappa (loc) can be executed by clicking the calculate kappa button.