

GeoCA Version 2.4

软件使用手册

文档编写：姚尧

高性能空间计算智能实验室（HPSCIL@CUG）

中国地质大学（武汉）地理与信息工程学院

目录

总体功能介绍	3
功能模块介绍	3
输入模拟数据设置模块（Simulation Data）	3
功能说明	3
界面说明	4
输出模拟结果设置模块（Output Settings）	6
功能说明	6
界面说明	6
模拟参数设置模块（Simulation Settings）	9
功能说明	9
界面说明	9
迭代参数（CA Iteration Parameters）设置	11
规则挖掘参数设置模块（Rule-mining Settings）	13
功能说明	13
界面说明	13
模拟结果检验模块（Validation Settings）	16
功能说明	16
界面说明	16
用地侵蚀制图模块（Land Erosion Mapping）	18
功能说明	18
界面说明	19
参数载入、保存和模型执行模块（Load and Export Params and Run）	21
功能说明	21
界面说明	21
处理日志输出模块（Process Log）	22
功能说明	22
界面说明	22
测试数据、参数和输出结果	23
相关文件说明	25
输入文件	25
土地利用数据文件	25
土地利用数据代码文件	25
土地利用重分类数据文件	25
空间辅助数据文件	26
附加概率数据文件	26
输出文件	26
城市转换文件	26
重分类土地利用文件	27
发展适宜性文件	27
用于回归分析（数据挖掘）的随机点文件	28
城市土地利用转换概率文件（非城市转向城市）	28
9 波段转换概率文件	28

模拟结果文件集	29
规则挖掘精度报告文件	29
模型验证文件	30
源/目标城市重分类文件	30
模拟结果文件集	30
输出精度文件集	30
用地侵蚀制图文件	32
土地利用 LULC 文件	32
模拟结果文件集	32
输出制图文件集	32
日志文件	33
参数文件	35
联系我们	35
参考文献	35

总体功能介绍

地理过程，尤其是城市发展过程，是极为复杂的。为了准确的模拟城市发展过程和预测未来城市形态，我们提供了一套基于元胞自动机的地理模拟软件（Geographical Simulation App via Cellular Automata, 以下简称 GeoCA）。

GeoCA 是一套用于模拟大规模像素级（栅格级）城市发展过程的软件。GeoCA 具备一套完备的高性能海量空间数据读写底层，可支持多线程读写、内存自动分配和空间数据坐标自动对齐特性。GeoCA 内封装的耦合元胞自动机和多机器学习模型支持的城市模拟模型，已被广泛的应用于城市发展过程分析、用地侵蚀分析、生态环境分析和城市规划相关领域。

GeoCA 包括以下功能模块：输入模拟数据、输出模拟结果、模拟参数设置、规则挖掘参数设置、模拟结果检验、用地侵蚀制图、模型执行、处理日志输出。模块和模块间采用都是数据松耦合的设计模式。

GeoCA 软件由中国地质大学（武汉）高性能空间计算智能实验室（HPSCIL@CUG）发布，版权为中国地质大学（武汉）高性能空间计算智能实验室（HPSCIL@CUG）所有。

基于元胞自动机的城市模拟模型理论和原理在文末附有参考文献，亦可于此下载相关资料：<http://www.urbancomp.net/2020/10/01/traditional-ca/>

以下将详细介绍每一个模块的功能、算法和操作步骤。

功能模块介绍

输入模拟数据设置模块（Simulation Data）

功能说明

该模块的主要功能为设置模型的输入数据。

输入数据包括三部分：

1. 输入数据的文件夹路径

- 2. 土地利用文件
- 3. 辅助空间变量数据
- 4. 附加概率文件

界面说明

界面如图 1 所示。输入分为四大部分：输入数据的文件夹路径（Source Filepath，必填）、土地利用文件（Landcover Files，必填）、辅助空间变量数据（Auxiliary Files，必填）、附加概率文件（Additional Probability File，选填）。

GeoCA V2.4 by HPSCIL@CUG

Simulation Data Output Settings Simulation Settings Rule-Mining Settings Validation Settings Land Erosion Mapping Process Log

Source Filepath

Landcover Filepath

.\WUHAN_DATA\WUHAN_LANDUSE

...

Auxiliary Filepath

.\WUHAN_DATA\WUHAN_AUXVAR

...

Output Filepath

.\WUHAN_DATA\WUHAN_OUTPUT

...

Landcover Files

Classification Code File

land_type_code.csv

...

Landcover File Name (Simulation Start Year)

2000/1/1

...

WUHAN_GLC30_2000.tif

...

Landcover File Name (Simulation End Year)

2010/1/1

...

WUHAN_GLC30_2010.tif

...

Auxiliary Files

	nFID	File Name	Minimum Value	Maximum Value	Average Value	Near Better (1=Yes, 0=No)	Comment
1	0	WUHAN_capital_city.tif	0.000000	130032.0...	60000.0...	1	WUHAN_capital_city
2	1	WUHAN_china_lspop2010_float.tif	0.000000	69483.00...	102.860...	1	WUHAN_china_lspop2
3	2	WUHAN_China_Slope.tif	0.000000	27.394400	1.000000	1	WUHAN_China_Slope
4	3	WUHAN_city.tif	0.000000	69498.60...	30000.0...	1	WUHAN_DISTRICT

NOTE: The maximum, minimum and average values of each spatial auxiliary variable must be entered manually!

Additional Probability File

☐ Do you use additional probability data multiplied by Pg? Plz put it in Aux_Filepath if use. (DATATYPE=FLOAT32)

Additional Probability File

...

Load Params

Export Params

Run

图 1 输入模拟数据设置界面

输入数据的文件夹路径（Source Filepath）

Landcover Filepath: 土地利用/土地覆被文件的文件夹路径。建议将所有土地利用和土地覆被文件放在该文件夹内。

Auxiliary Filepath: 辅助空间变量数据文件的文件夹路径。建议将所有辅助空间变量文件放在该文件夹内。

Output Filepath: 输出数据的文件夹路径。程序将会把所有输出文件放置于该文件夹内。

土地利用文件（Landcover Files）

Classification Code File: 放置于 Landcover Filepath 内的土地利用重分类文件，文件类型为 txt 或 csv 文本文件。每一类存储每一土地利用和土地覆被类型的 DN, NAME, TYPE_ID, RE_TYPE, RE_TYPE_ID, COMMENTS。

Landcover File Name（Simulation Start/End Year）: 模拟起始和终止年的土地利用文件，Geotiff 格式。需要输入年份和文件名。

辅助空间变量数据（Auxiliary Files）

该部分设置的按钮可批量打开辅助空间变量数据集。

每一个数据为 Geotiff 格式，建议数据类型为 FLOAT32 浮点型。

每一个辅助空间变量数据以行的形式显示，nFID 为辅助数据的唯一标识码。

为了数据归一化进入规则挖掘，需要手动修改每一个辅助空间变量数据的最小值（Minimum Value）、最大值（Maximum Value）和均值（Mean Value），默认最小值为 0、最大值为 1、均值为 0.5，可以按需修改。

是否越近越好（Better）：“是”填 1，“否”填 0。若填 0，归一化将采用逆序方式。

备注（Comments）：便于理解，可选填。

附加概率文件（Additionally Probability File）

该文件用于挖掘城市增长规则概率基础上，相乘另一个概率数值。

该文件可选是否使用，数据类型为 FLOAT32，数据范围不做限定。如果使用了该项概率，Simulation Settings >> CA Iteration Settings >> Minimum Transition Probability Threshold 也建议根据该文件数值范围修改。

提示：该文件可用于控制是否允许增长（数据为 0 或 1），或者用于控制（增加或减少）某种地物被侵蚀的概率。

输出模拟结果设置模块（Output Settings）

功能说明

该模块为设置输出模拟结果文件。

界面说明

图 2 为输出结果设置模块界面。

GeoCA 用于城市模拟输出的结果主要包括：

1. 土地利用转换图（Geotiff 格式，BYTE 类型）
2. 起始年、终止年土地利用重分类图（Geotiff 格式，BYTE 类型）
3. 土地发展适宜性数据（Geotiff 数据，BYTE 类型）
4. 用于规则挖掘的随机点和转换类型数据（文本数据）
5. 土地利用类型转换（非城市-城市）的单波段概率数据（Geotiff 格式，FLOAT32 类型）
6. CA 迭代结果文件集（Geotiff 格式，BYTE 格式）

7. 9 种土地利用转换类型（非城市-城市-保护区）相互转换概率文件（可选，Geotiff 格式，FLOAT32 类型）

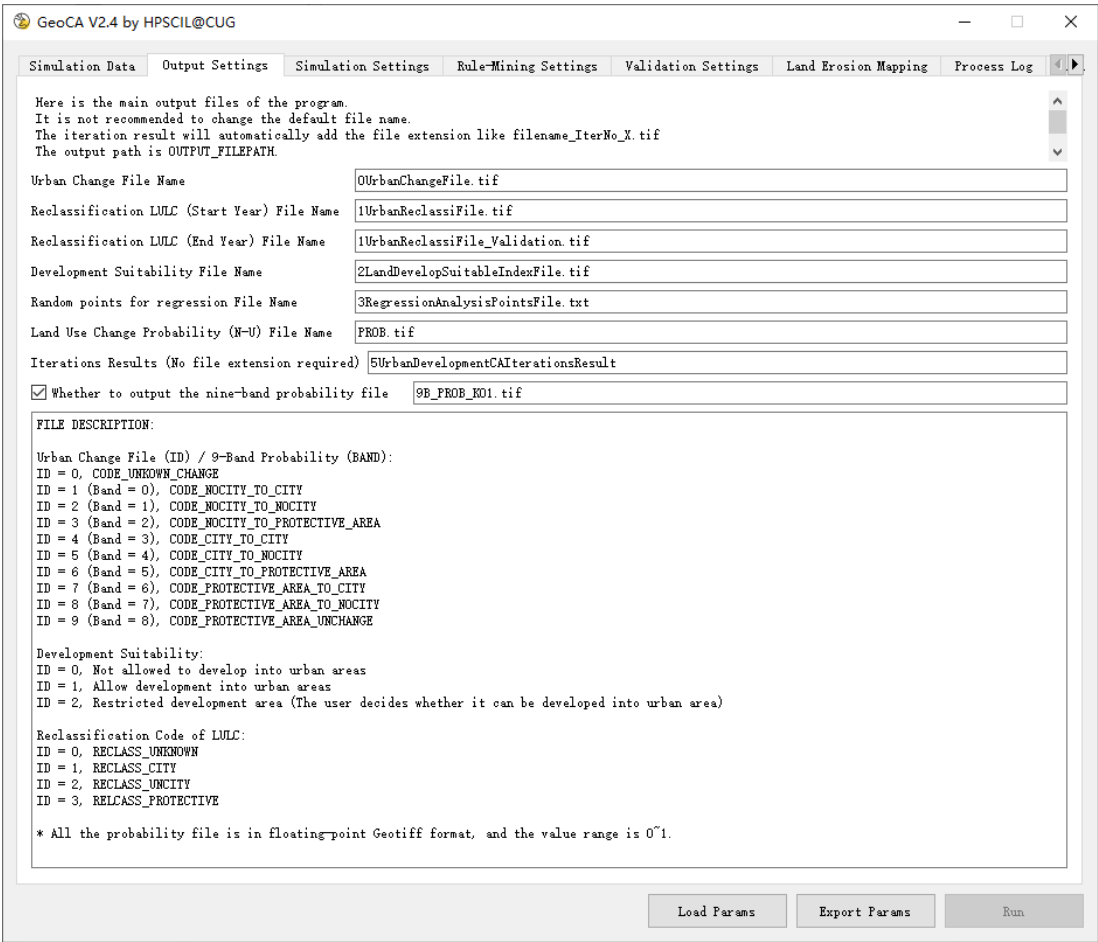


图 2 输出结果数据文件名设置界面

输出文件路径

所有数据将会输出于 Source Filepath >> Output Filepath 文件夹中。

Urban Change File Name（土地利用转换图）

统计的起始年和终止年土地利用类型转换差异。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE。

Reclassification LULC (Start/End Year) File Name （起始年/终止年的重分类文件）

根据重分类代码文件内设置，将起始年和终止年土地利用数据的重分类结果。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE。

Development Suitability File Name （土地发展适宜性结果）

根据重分类代码文件内设置，得到研究区每一个栅格的发展为城市的适宜性（允许发展、不允许发展、受限发展区域）。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE。

Random Points for Regression File Name （用于规则挖掘的随机点数据）

以 9 种不同的转换类型，在土地利用转换和辅助空间数据上进行随机取样，存储为随机点-转换类型数据。可用于后期检查数据和测试数据挖掘模型精度。

该文件为文本文件，存储了每一个随机点的经纬度、辅助数据集和转换类型。

Land-use Change Probability (N-U) File Name （非城市向城市转换的概率文件）

根据获取的随机点和用户选择的规则挖掘模型，得到的非城市向城市转换的 Pg 概率图文件。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE，数值范围为 0~1。

Iteration Results (No File Extension required) （CA 迭代结果文件集）

根据 CA 模型和设置的参数（Simulation Settings），得到的城市发展模拟迭代结果文件。

注意这里不需要设置文件扩展名，默认文件名为
5UrbanDevelopmentCAIterationsResult。

迭代结果会根据迭代次数自动命名为：
5UrbanDevelopmentCAIterationsResult_IterNo_*.tif，其中*为迭代次数。

Nine-band Probability File （土地利用类型 9 种转换的概率文件）

该文件较大，是否输出可选，默认不输出。

根据获取的随机点和用户选择的规则挖掘模型，得到的 9 种类型转换的 Pg 概率图文件。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE，数值范围为 0~1，具有 9 个波段。

模拟参数设置模块（Simulation Settings）

功能说明

该模块用于设置模拟参数和 CA 迭代参数。

界面说明

图 3 为模拟和迭代过程的参数设置界面。

主要包括两大设置：模拟参数（Simulation Parameters）设置和迭代参数（CA Iteration Parameters）设置。

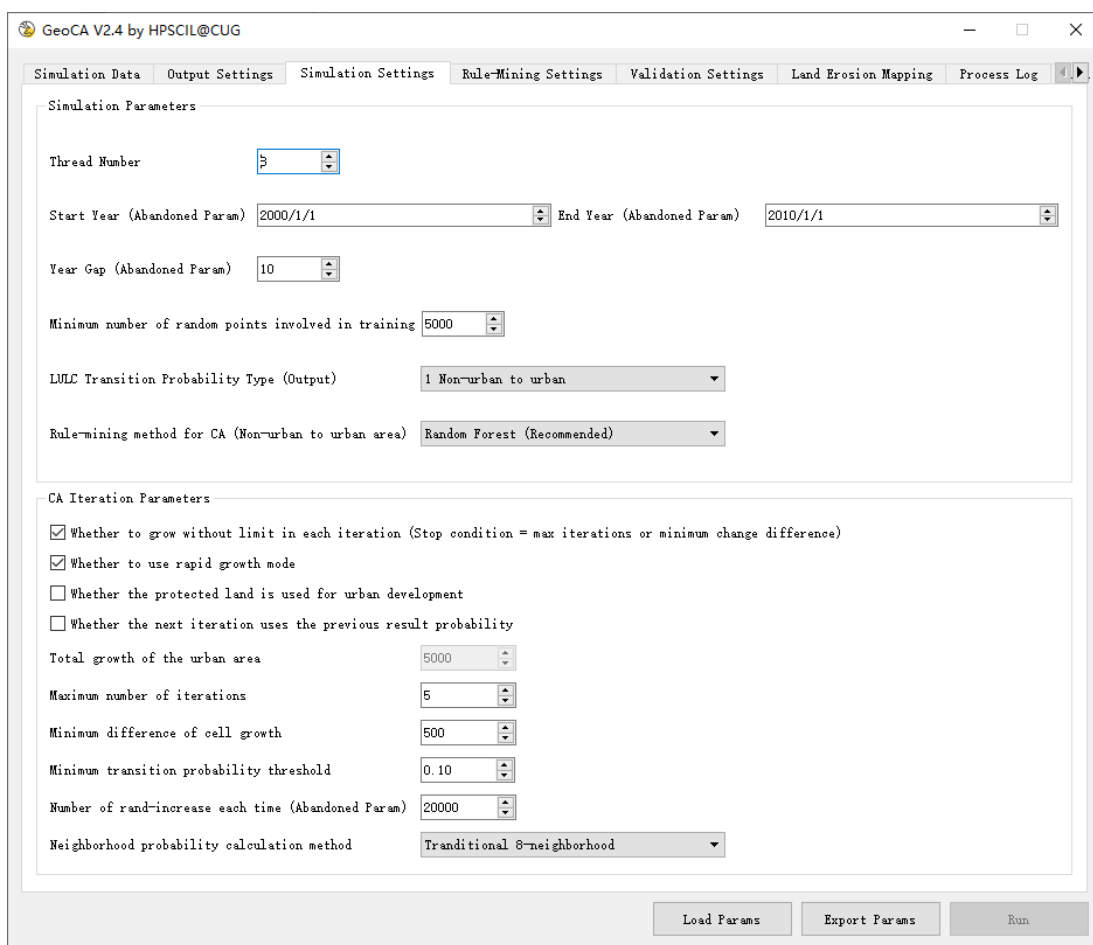


图 3 模拟和迭代过程参数设置界面

模拟参数（Simulation Parameters）设置

线程数目（Thread Number）：设置多线程数目。程序会自动获取电脑可使用的多线程数目，是否使用多线程由用户决定。若不使用多线程，该项设置为 1。

起始年（Start Year）、终止年（End Year）、年间隔（Year Gap）：这三项为废弃参数，不需要设置。

训练使用的最少随机点数（Minimum number of random points involved in training）：程序自动采集的最少随机点数目。一般情况，程序会采集用户设置数目的 2 倍随机点进入规则挖掘过程。

输出的土地利用转换概率类型 (LULC Transition Probability Type) : 输出的 Pg 概率类型, 默认为“1 非城市到城市”。如无必要, 不要修改为其他类型的输出。

规则挖掘模型选择 (Rule-mining method for CA) : 用于计算 Pg 概率的规则挖掘方法。GeoCA 目前提供 5 种模型: 随机森林模型 (Random Forest, 推荐使用)、神经网络模型 (Neural Network, 推荐使用)、逻辑模型 (Logit Model, 推荐使用)、复杂逻辑回归 (Logic Regression, 测试中, 不推荐使用) 和免疫算法 (Immune Algorithm, 测试中, 不推荐使用)。

迭代参数 (CA Iteration Parameters) 设置

城市总增长量 (Total growth of the urban area) : 设置城市增长的元胞数目。

最大迭代次数 (Maximum number of iterations) : 若迭代次数超过该值, 则停止迭代。

最小增长间隔差异 (Minimum difference of cell growth) : 前后迭代元胞增长数目如果小于该值, 则停止迭代。

最小转换概率 (Minimum Transition Probability Threshold) : 非城市发展为城市的最小发展概率。概率计算方法 $P = P_g * \Omega * RA$, 其中 Ω 为邻域概率, RA 为正态随机数。

每次增长的种子点数目 (Number of rand-increase each time) : 当前已废弃参数, 不需要设置。

邻域概率计算方法 (Neighborhood Probability Calculation Method) : 计算邻域概率 Ω 的方法。

GeoCA 提供了三种邻域概率计算模型:

Tranditional 8-neighborhood: 传统八邻域概率 (城市=1, 非城市=0)

8-neighborhood with Probability: 考虑到非城市发展概率的改进 8 邻域模型 (城市=1, 非城市= P_g)

Tan-curve model: 基于正切模型的邻域概率 $(\tan(0.25 * \text{邻域城市数量} - 1.2) + 3.0) / 4.0 - 0.09$

是否每次无限制发展（Whether to grow without limit in each iterations）：

√ 如果框选此项：

城市总增长量将不受限制（在参数文件中此项为 0）；

GeoCA 将会把每次能增长为城市的元胞都进行转变；

迭代停止条件为达到最大迭代次数或最小增长间隔差异。

○ 如果不框选此项：

城市总增长量为用户设置。

每次增长的数量 = 城市总增长量 / 最大迭代次数。

最小增长间隔差异非迭代停止条件。

是否采用快速发展模式（Whether to use rapid growth mode）：

√ 如果框选此项：

数据量较小情况下，直接统计可以增长的城市数据。

一般和无限限制发展配合使用。

○ 如果不框选此项：

根据城市总增长量和模拟城市扩张总数进行模拟。

一般和非无限限制发展配合使用。

是否保护用地发展为城市用地（Whether the protected land is used for urban development）：

√ 如果框选此项：保护用地在迭代过程中会发展为城市用地。

○ 如果不框选此项：保护用地保持不变。

是否下次迭代采用上一次的概率结果（Whether the next iterations uses the previous result probability）：

不建议框选。

√ 如果框选此项：每次迭代后总体发展概率 P_g 将会被 P 代替， P 的计算收敛速度将加快。

○ 如果不框选此项：总体发展概率 P_g 保持不变。

规则挖掘参数设置模块（Rule-mining Settings）

功能说明

该模块用于设置计算总体发展概率 P_g 的机器学习模型参数。

界面说明

图 3 为规则挖掘模型参数设置界面。

根据选择的机器学习模型，GeoCA 将自动开放相应的参数设置区域。

The screenshot shows the 'Rule-Mining Settings' tab in the GeoCA V2.4 software. The interface is divided into several sections for different machine learning models:

- Random Forest:**
 - Total Trees: 50
 - Training ratio: 0.80
 - Number of variables for split: 1000
 - Output Accuracy and Weight File Name: DecisionForestAccuracyReport.txt
- Neural Network:**
 - Hidden Layers Number: 3
 - Restart Times for Validation: 10
 - Output Accuracy File Name: NeuralNetworkAccuracyReport.txt
- Logit Model:**
 - Method:
$$P(class=i) = \frac{S(i)}{S(0) + S(1) + \dots + S(M-1)}$$
$$S(i) = \text{Exp}(A[i,0]*X[0] + \dots + A[i,N-1]*X[N-1] + A[i,N]), \text{ when } i < M-1$$
$$S(M-1) = 1$$
 - Output Coefficients File Name: AdvancedLinearRegCoefficients.txt
 - Output Accuracy File Name: AdvancedLinearRegAccuracyReport.txt
- Logic Regression:**
 - Saved Model File Name: LogicRegressionTrainingFile.mod
 - Update Model Every Iteration: ☐
 - Minimal Accepted Accuracy: 0.30
 - Max Iteration Loops: 100
 - Minimal Loss for Decrease: 0.001000
 - Learning Rate: 1.000000
 - Lambda of Gaussian Prior: 0.000000
 - Average weights of all iteration loops: ☐
- Immune Algorithm:**
 - Model Order: 1
 - Antibody Number: 200
 - Variation Coefficient: 0.80
 - Exchange Probability: 0.60
 - Max Training Iterations: 300000
 - Output Accuracy File Name: ImmuneAlgorithmAccuracyReport.txt

At the bottom, there are buttons for 'Load Params', 'Export Params', and 'Run'.

图 4 规则挖掘参数设置模块

随机森林（Random Forest）

决策树总数目（Toal trees）： 构建随机森林采用的决策树总数目；

训练比例（Training ratio）： 训练数据集（袋内数据）比例；

最佳分裂时使用的变量数（Number of variables for split）： 用于决策树分割时使用的变量数目

输出精度文件（Output Accuracy and Weight File Name）： 输出的随机森林精度和变量权重文件，存储于 Simulation Data >> Output Filepath 内。

神经网络（Neural Network）

隐藏层数目 (Hidden Layers Num) : 神经网络隐藏层数目;

训练重启次数 (Restart times for validation) : 训练时重启次数。设置为 10 时, 则为 10-fold 交叉验证;

输出精度文件 (Output Accuracy File Name) : 输出的神经网络交叉验证精度文件, 存储于 Simulation Data >> Output Filepath 内。

逻辑模型 (Logit Model)

输出参数因子文件 (Output Ccoefficients File Name) : 输出每一个转换类别的特征变量权重因子;

输出精度文件 (Output Accuracy File Name) : 输出的逻辑模型训练精度文件, 存储于 Simulation Data >> Output Filepath 内。

复杂逻辑回归 (Logic Regression)

模型文件 (Saved model file name) : 逻辑回归保存的模型文件。存储于 Simulation Data >> Output Filepath 内;

最小接受精度 (Minimal accepted accuracy) : 此为迭代停止条件, 当模型精度大于该数值时, 停止训练模型;

最大迭代次数 (Max iteration loops) : 此为迭代停止条件, 但迭代次数达到该数值时, 停止训练模型;

最小损失率 (Minimal loss for decrease) : 模型迭代时的损失函数参数;

高斯先验 λ (Lambda of Gaussian Prior) : 模型的高斯核参数;

学习率 (Learning rate) : 梯度下降时的学习率设置。

是否更新模型 (Update Model Every Iteration) :

√ 如果框选此项: 启动时载入已有模型进行更新。

○ 如果不框选此项: 将重新训练模型。

变量权重是否采用所有迭代均值 (**Average weights of all iteration loops**) :

- √ 如果框选此项: 变量权重将采用所有迭代的均值更新。
- 如果不框选此项: 变量权重将采用最后一次迭代的结果。

免疫算法 (**Immune Algorithm**)

模型阶数 (**Model order**) : 免疫模型模型的算法阶数;

抗体数目 (**Antibody number**) : 抗体数目;

方差参数 (**Variation Coefficient**) : 差异水平;

交换概率 (**Exchange Probability**) : 抗体交换基因的概率;

最大迭代次数 (**Max Training Iterations**) : 模型训练的最大迭代次数;

输出精度文件 (**Output Accuracy File Name**) : 输出的免疫算法模型训练精度文件, 存储于 Simulation Data >> Output Filepath 内。

模拟结果检验模块 (**Validation Settings**)

功能说明

该模块用于验证 CA 模拟结果, 支持对每一次迭代结果计算其精度。

该功能为可选功能。

界面说明

图 5 为 CA 模拟结果验证模块。

根据输入的验证数据, 支持对 CA 模拟迭代结果进行精度评价。

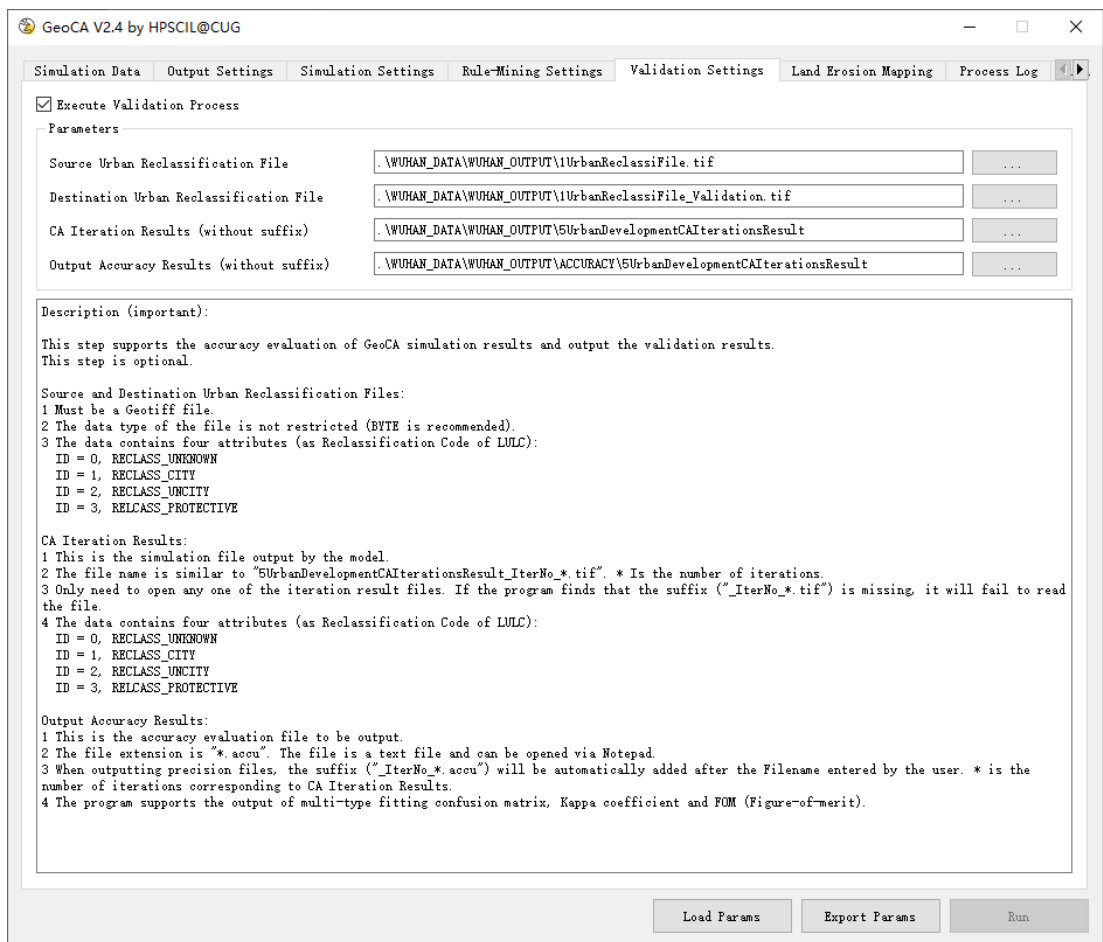


图 5 CA 模拟验证模块

源土地利用重分类文件（Source Urban Reclassification File）

模拟起始年的土地利用重分类文件。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE。

此处应该为文件的绝对路径。

目的土地利用重分类文件（Destination Urban Reclassification File）

验证年份（模拟目标年份）的土地利用重分类文件。

扩展名为 tif，文件类型为 Geotiff，数据类型为 BYTE。

此处应该为文件的绝对路径。

CA 模拟结果文件集 (CA Iteration Results without Suffix)

城市发展模拟迭代结果文件，为绝对路径。

注意这里不需要设置文件扩展名，默认设置文件路径是 Simulation Data >> Output Filepath，文件名为 Output Settings >> Iteration Results。

CA 迭代结果命名应为该类型：filename_IterNo_*.tif，其中*为迭代次数。

如果手动选择的文件缺少"_IterNo_"，则 GeoCA 会拒绝载入。

输出精度评价结果文件集 (Output Accuracy Results without Suffix)

城市发展模拟迭代的精度评价文件，为绝对路径。

注意这里不需要设置文件扩展名，默认设置文件路径是 Simulation Data >> Output Filepath，文件名为 Output Settings >> Iteration Results。

迭代结果会根据迭代次数自动命名为：

5UrbanDevelopmentCAIterationsResult_IterNo_*.accu，其中*为迭代次数。

*.accu 为文本文件，可以用记事本打开，GeoCA 自动输出混淆矩阵、kappa 系数、各类的错分和误分误差、各类的用户和生产者精度。以及 Figure-of-Merit (FoM)，FoM 的用户精度和生产者精度。

用地侵蚀制图模块 (Land Erosion Mapping)

功能说明

该模块为一个 CA 的典型应用，用于分析某类用地在城市发展过程中被侵蚀的比例，及被侵蚀的分布制图。

该功能为可选功能。

界面说明

图 6 为基于 CA 模拟结果的用地侵蚀制图和分析模块。

根据输入的 CA 模拟迭代结果和第三方土地利用数据，进行用户选择的用地开展侵蚀分析和制图。

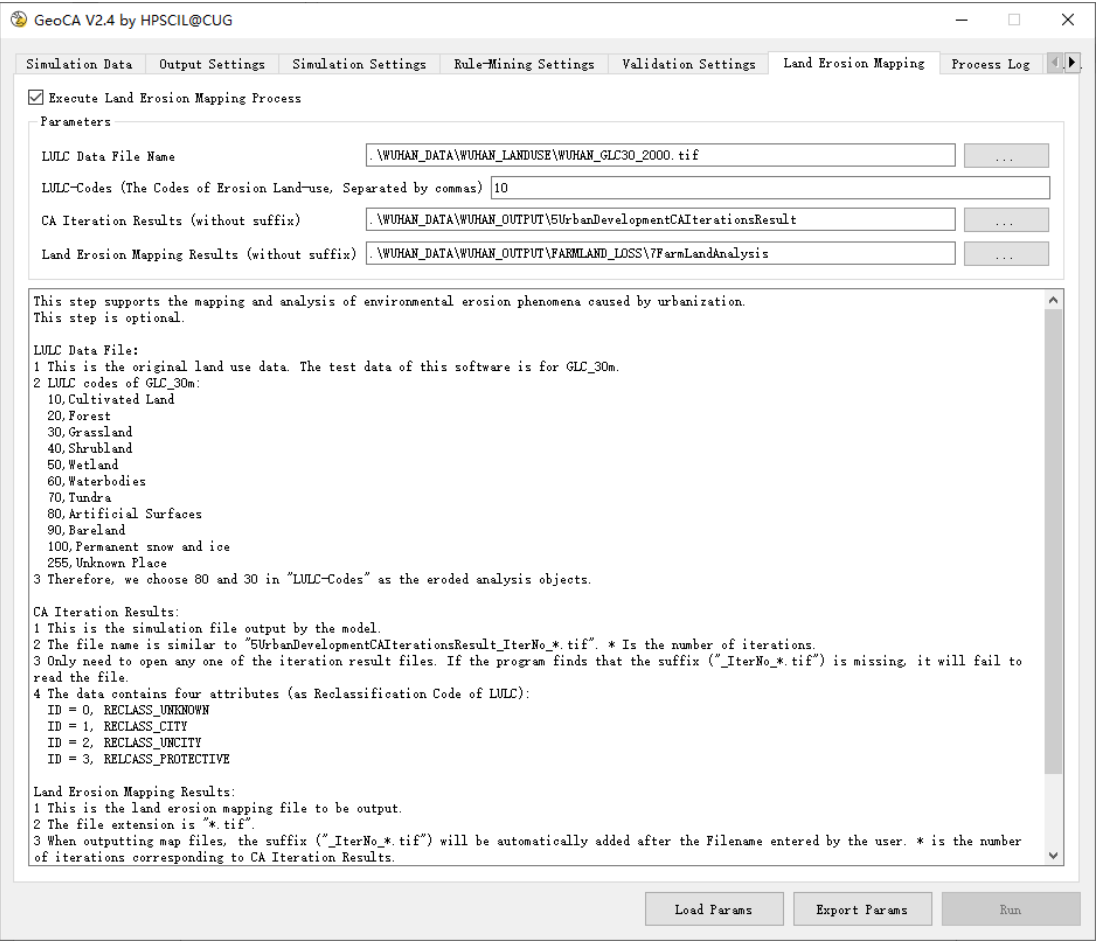


图 6 用地侵蚀制图和分析模块

第三方土地利用文件（LULC Data File Name）

输入的第三方土地利用文件，为绝对路径。

文件类型为 **Geotiff**，数据类型建议为 **BYTE**。

需要统计的被侵蚀用地类型代码（LULC-Code）

在第三方土地利用文件中，需要被分析的用地侵蚀情况。

可以输入多个用地代码用于分析，每个代码用英文逗号隔开。

CA 模拟结果文件集（CA Iteration Results without Suffix）

城市发展模拟迭代结果文件，为绝对路径。

注意这里不需要设置文件扩展名，默认设置文件路径是 **Simulation Data >> Output Filepath**，文件名为 **Output Settings >> Iteration Results**。

CA 迭代结果命名应为该类型：**filename_IterNo_*.tif**，其中*为迭代次数。

如果手动选择的文件缺少"**_IterNo_**"，则 **GeoCA** 会拒绝载入。

输出用地侵蚀结果文件集（Land Erosion Mapping Results without Suffix）

输出的用地侵蚀结果文件集，为绝对路径。

注意这里不需要设置文件扩展名，默认设置文件路径是 **Simulation Data >> Output Filepath**，文件名为"**7FarmlandAnalysis**"。

迭代结果会根据迭代次数自动命名为：**7FarmLandAnalysis_IterNo_*.tif**，其中*为迭代次数。

侵蚀结果文件为 **Geotiff**，数据类型为 **BYTE**。

统计的用地侵蚀比例可以在 **log** 文件中查询。

参数载入、保存和模型执行模块（Load and Export Params and Run）

功能说明

该模块为参数的载入、输出和执行。

界面说明

图 7 为模型的执行按钮集。

支持载入、保存 GeoCA 参数文件和运行模拟流程。

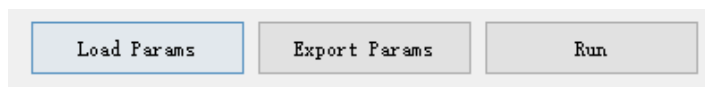


图 7 模型执行按钮集

参数载入功能（Load Params）

支持载入已有的参数文件。

参数文件为 xml 格式。

参数保存功能（Export Params）

支持保存用户设置的参数。

参数文件为 xml 格式。

运行模型功能（Run）

在用户保存设置参数后，点击该按钮会执行整个模拟过程。

模型执行时，将会自动跳转到处理日志输出模块。

注意：只有将参数保存后，该按钮才会成为有效状态。

处理日志输出模块（Process Log）

功能说明

该模块为模型执行时的日志输出和显示。

界面说明

图 8 为模型执行时的日志输出模块。

日志文件将会自动保存在输出路径（Simulation Parameters >> Output Filepath）的 logs 文件夹中。

*.log 文件为文本文件，可以用记事本打开查看。

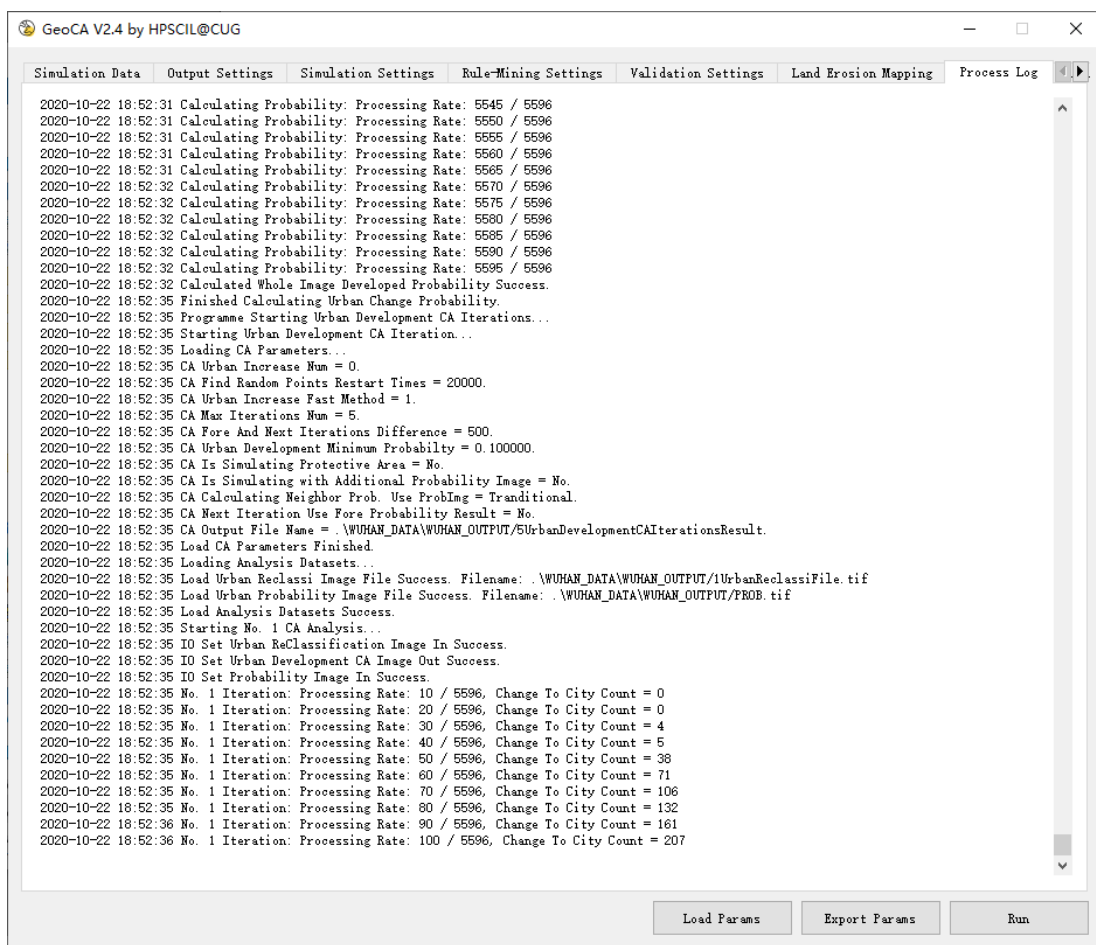


图 8 模型执行时的日志输出模块

测试数据、参数和输出结果

为了让用户更快的了解 GeoCA 的功能和文件需求，GeoCA 提供了一套测试数据集和测试参数文件。

测试数据集位置位于程序目录下的 ./WUHAN_DATA 文件夹中（图 9）。

对应的测试参数文件是程序目录下的 ./sim_params.xml（图 10）。

可以采用 GeoCA 提供的参数载入、保存和模型执行模块，载入参数并保存参数（到第三方文件）后可以直接执行。

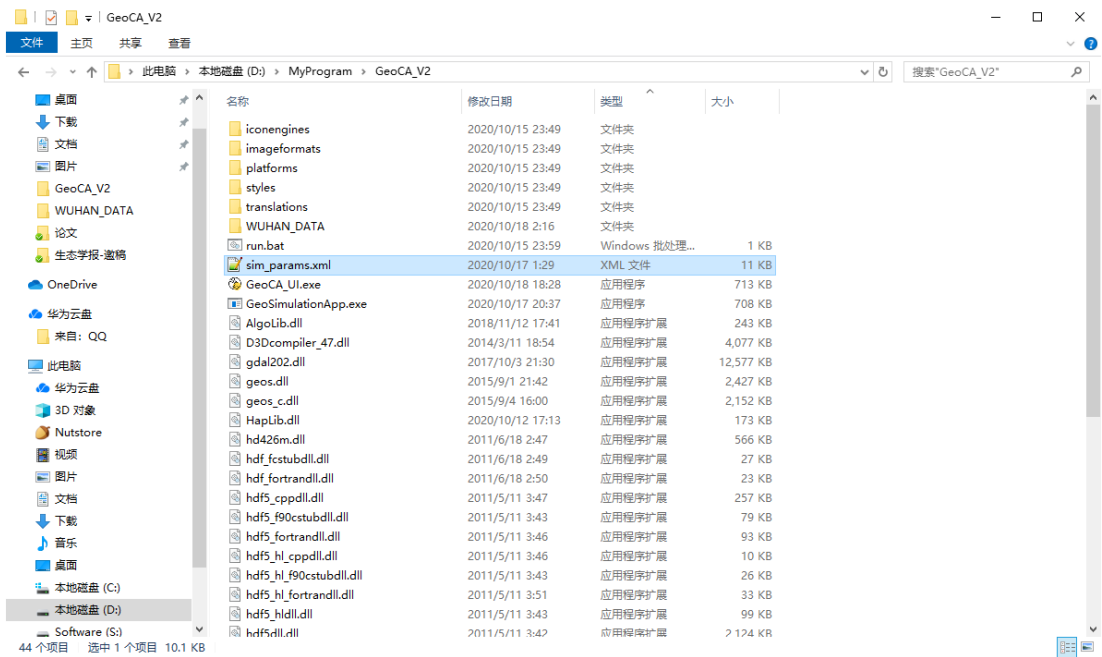


图 9 测试数据和参数所在位置

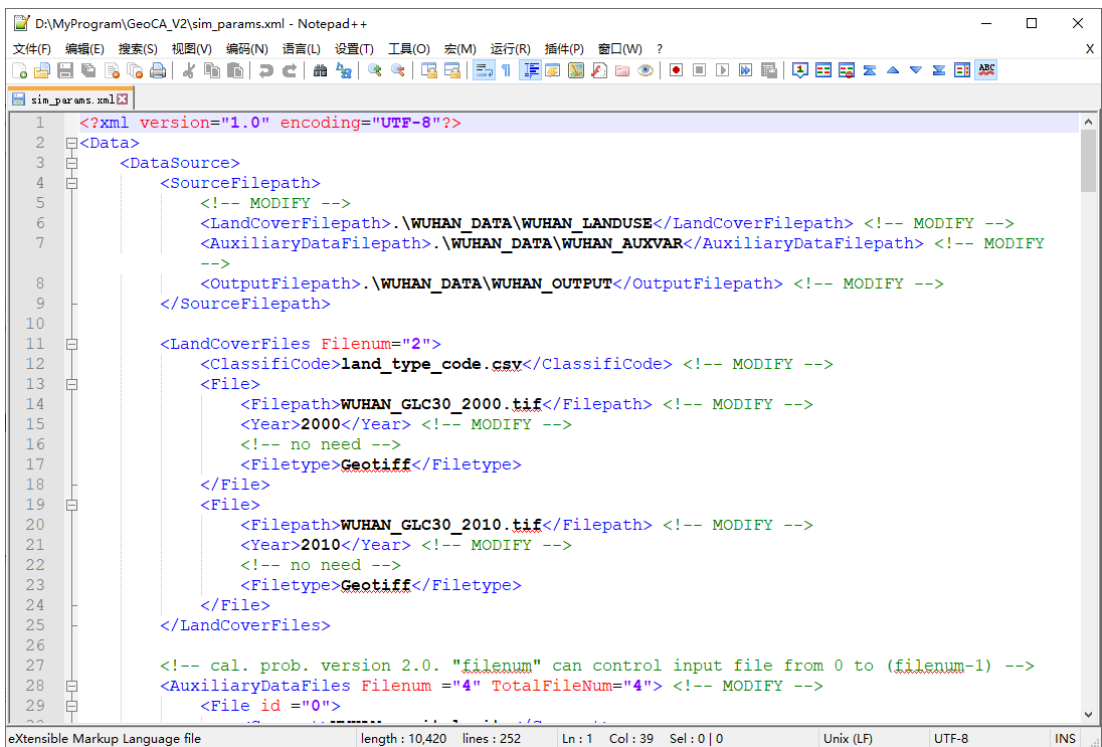


图 10 Notepad++打开测试参数文件

相关文件说明

输入文件

土地利用数据文件

文件格式：Geotiff 带投影坐标信息

数据类型：BYTE、UINT16、FLOAT32，单波段

文件说明：无

土地利用数据代码文件

文件格式：文本文件

数据类型：txt 或 csv

文件说明：

以逗号分隔的文本文件，建议为英文或数字，不要出现中文。

第一行为：DN,NAME,TYPE_ID,RE_TYPE,RE_TYPE_ID,COMMENTS

DN：土地利用数据中对应的数值（DN 值）

NAME：土地利用/土地覆被名称

TYPE_ID：1~N，根据土地利用类型数量编号

RE_TYPE：NON_URBAN_AREA（非城市用地）、URBAN_AREA（城市用地）、
PROTECTED_AREA（保护用地）

RE_TYPE_ID：NON_URBAN_AREA（0）、URBAN_AREA（1）、
PROTECTED_AREA（-1）

COMMENTS：备注，可无

土地利用重分类数据文件

文件格式：Geotiff 带投影坐标信息

数据类型：BYTE，数值范围 0-3，单波段

文件说明：根据土地利用数据代码文件将土地利用数据文件重分类。

数据代码：

0 未知用地

1 城市用地

2 非城市用地

3 保护用地

空间辅助数据文件

文件格式：Geotiff 带投影坐标信息

数据类型：FLOAT32，单波段

文件说明：空间辅助数据文件，如距离城市中心距离、距离主要公路距离、路网密度等。

附加概率数据文件

文件格式：Geotiff 带投影坐标信息

数据类型：FLOAT32，单波段

文件说明：乘在总体概率文件 P_g 上的附加概率文件。

输出文件

城市转换文件

文件格式：Geotiff 带投影坐标信息

数据类型：BYTE，数值范围 0-9，单波段

文件说明：根据两期土地利用重分类文件制作的转换文件。

数据代码：城市转换代码

0 未知转变

1 非城市用地转换为城市用地

2 非城市用地未转变

3 非城市用地转换为保护用地

4 城市用地未转变

5 城市用地转换为非城市用地

6 城市用地转换为保护用地

7 保护用地转换为城市用地

8 保护用地转换为非城市用地

9 保护用地未转变

重分类土地利用文件

参考“输入数据 >> 土地利用重分类数据文件”。

发展适宜性文件

文件格式：Geotiff 带投影坐标信息

数据类型：BYTE，数值范围 0-2，单波段

文件说明：根据土地利用数据代码文件制作的适宜性转变文件。

数据代码：

0 不允许转变为城市用地

1 允许转变为城市用地

2 保护区（用户决定是否可以转变）

用于回归分析（数据挖掘）的随机点文件

文件格式：文本文件

数据类型：txt

文件说明：

第一行为随机点数量，变量数量，转变类型数量

Longitude 经度

Latitude 纬度

Feature-(nFID+1) 第 nFID 个特征值

ClassID 城市转换代码（详见“城市转换文件”）

城市土地利用转换概率文件（非城市转向城市）

文件格式：Geotiff 带投影坐标信息

数据类型：FLOAT32，数值范围 0-1，单波段

文件说明：非城市用地、保护用地转化为城市用地的概率。

9 波段转换概率文件

文件格式：Geotiff 带投影坐标信息

数据类型：FLOAT32，数值范围 0-1，9 波段

文件说明：非城市用地、保护用地和城市用地相互转换的概率。

波段说明：

- 1 非城市用地转换为城市用地
- 2 非城市用地未转变
- 3 非城市用地转换为保护用地
- 4 城市用地未转变
- 5 城市用地转换为非城市用地
- 6 城市用地转换为保护用地
- 7 保护用地转换为城市用地
- 8 保护用地转换为非城市用地
- 9 保护用地未转变

模拟结果文件集

文件格式：Geotiff 带投影坐标信息

数据类型：BYTE，数值范围 0-3，单波段

文件说明：模拟结果文件集。文件名中有“_IterNo_*.tif”，*为迭代次数。

数据代码：

- 0 未知用地
- 1 城市用地
- 2 非城市用地
- 3 保护用地

规则挖掘精度报告文件

文件格式：文本文件

数据类型：txt

文件说明：规则挖掘模型精度报告。第一行为报告生成时间。

指标说明：

rel cls error 相对分类精度
avgce 平均交叉熵
rms error 均方根误差 (RMSE)
avg error 平均误差
avg rel error 平均相对误差
oob-* 袋外误差 (随机森林)
Top vars 变量重要性排序 (随机森林)
Var importances 变量重要性值 (随机森林)
Coefficients 变量权重参数 (逻辑回归)
Number of independent variables 独立变量个数 (逻辑回归)
Number of classes 类别个数 (逻辑回归)

模型验证文件

源/目标城市重分类文件

参考“输入数据 >> 土地利用重分类数据文件”。

模拟结果文件集

参考“输出数据 >> 模拟结果文件集”

输出精度文件集

文件格式：文本文件 (扩展名为*.accu)

数据类型: accu

文件说明: 模拟结果精度评价文件集。文件名中有“_IterNo_*.accu”, *为迭代次数(图 11)。第一行为评价的模拟结果文件名, 第二行为报告生成时间。

指标说明:

Confusion Matrix 混淆矩阵

Overall Accuracy 总体精度

Kappa Kappa 系数

Commission Error 错分误差

Omission Error 漏分误差

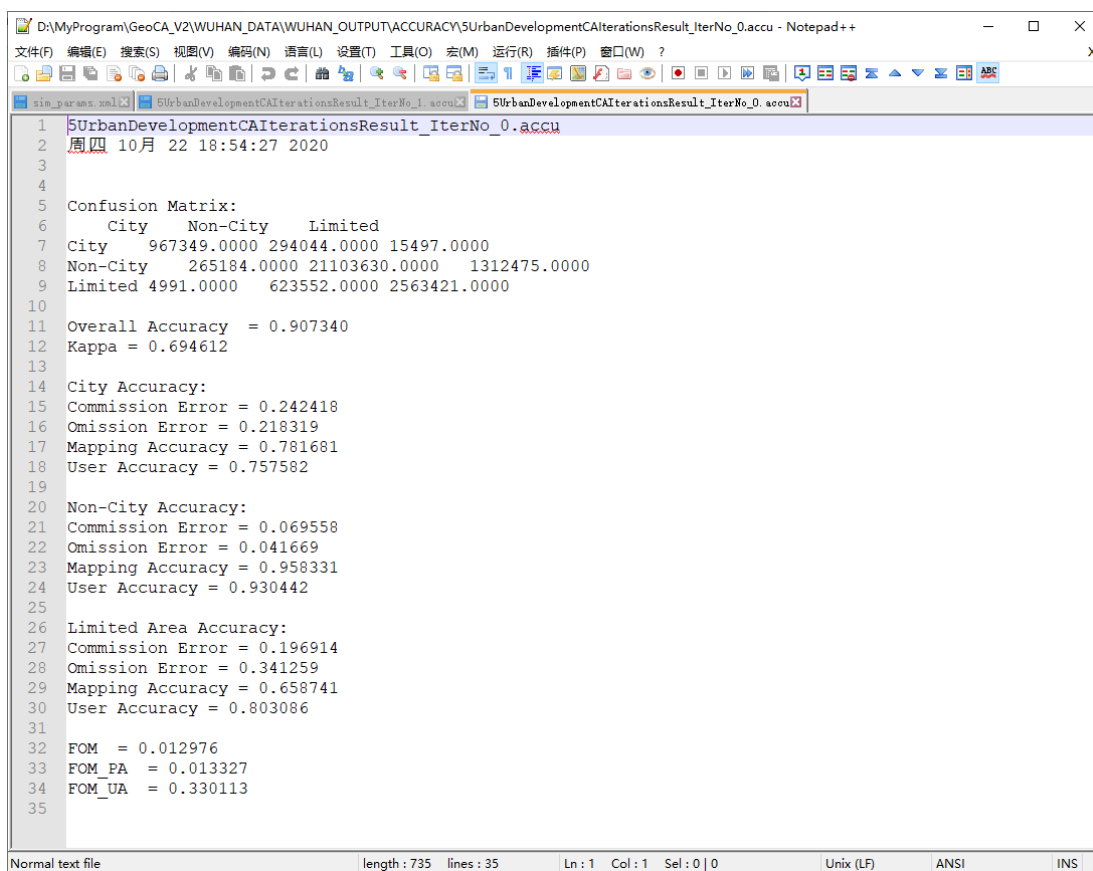
Mapping Accuracy 制图精度

User Accuracy 用户精度

FOM 品质因子 Figure-of-Merit

FOM_PA 基于 FOM 的生产者精度

FOM-UA 基于 FOM 的用户精度



```
1 5UrbanDevelopmentCAIterationsResult_IterNo_0.accu
2 周四 10月 22 18:54:27 2020
3
4
5 Confusion Matrix:
6   City   Non-City   Limited
7 City   967349.0000 294044.0000 15497.0000
8 Non-City 265184.0000 21103630.0000 1312475.0000
9 Limited 4991.0000 623552.0000 2563421.0000
10
11 Overall Accuracy = 0.907340
12 Kappa = 0.694612
13
14 City Accuracy:
15 Commission Error = 0.242418
16 Omission Error = 0.218319
17 Mapping Accuracy = 0.781681
18 User Accuracy = 0.757582
19
20 Non-City Accuracy:
21 Commission Error = 0.069558
22 Omission Error = 0.041669
23 Mapping Accuracy = 0.958331
24 User Accuracy = 0.930442
25
26 Limited Area Accuracy:
27 Commission Error = 0.196914
28 Omission Error = 0.341259
29 Mapping Accuracy = 0.658741
30 User Accuracy = 0.803086
31
32 FOM = 0.012976
33 FOM_PA = 0.013327
34 FOM-UA = 0.330113
35
```

图 11 模拟精度文件

用地侵蚀制图文件

土地利用 LULC 文件

参考“输入文件 >> 土地利用数据文件”。

模拟结果文件集

参考“输出文件 >> 模拟结果文件集”。

输出制图文件集

文件格式：Geotiff 带投影坐标信息

数据类型：BYTE，数值范围 0-2，单波段

文件说明：用地在未来的侵蚀情况空间制图。文件名中有“_IterNo_*.tif”，*为迭代次数（图 11）。侵蚀比例可以在日志文件中查看（图 12）。

数据代码：

0 未知

1 已被城市用地侵蚀

2 未被城市用地侵蚀

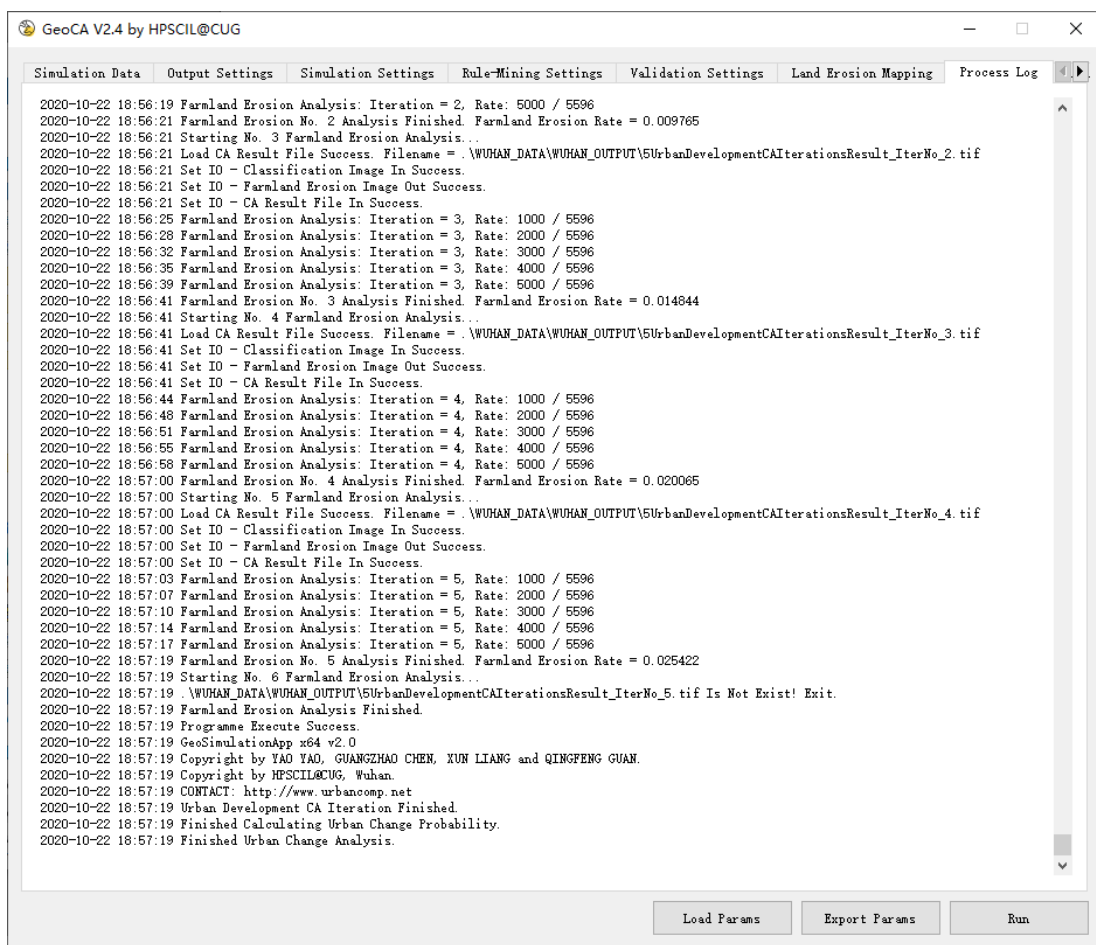


图 12 日志文件中的用地侵蚀分析输出

日志文件

日志文件为程序运行时的输出文件，可以在 GeoCA 的处理日志输出模块实时查看。

日志文件默认保存在 Output Filepath 的 logs 文件夹中（图 13）。

日志文件的命名格式为“HPSCIL_SIM_GEOCA_V2_小时_分钟_秒_日期.log”，文件名中的时间为程序执行的时间。

日志文件可通过记事本打开（图 14）。

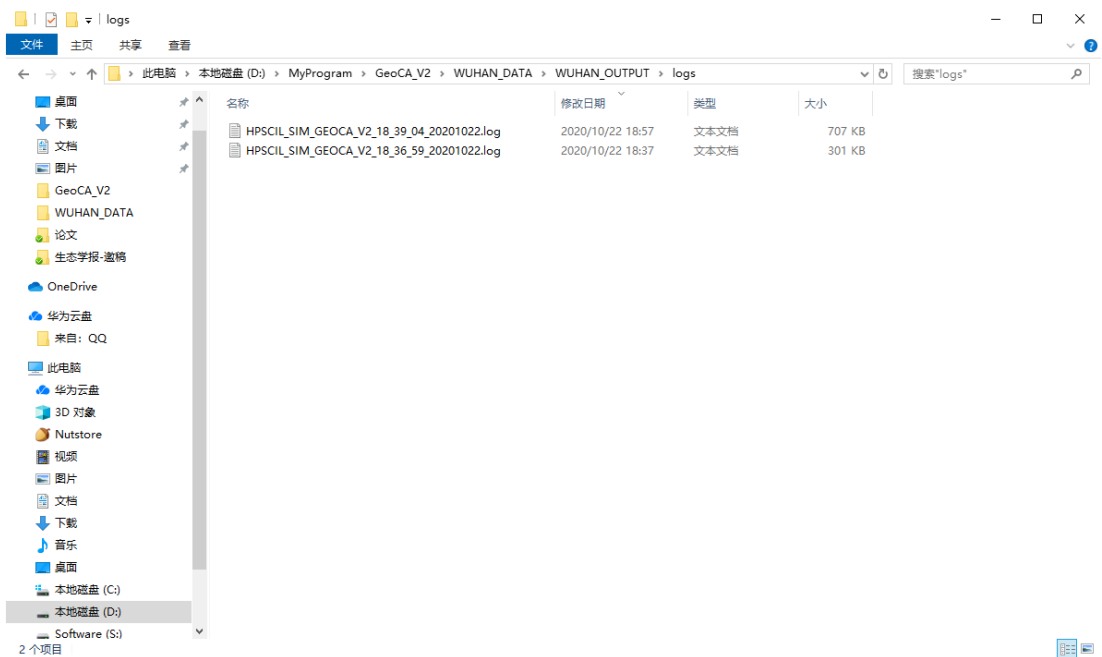


图 13 日志文件夹

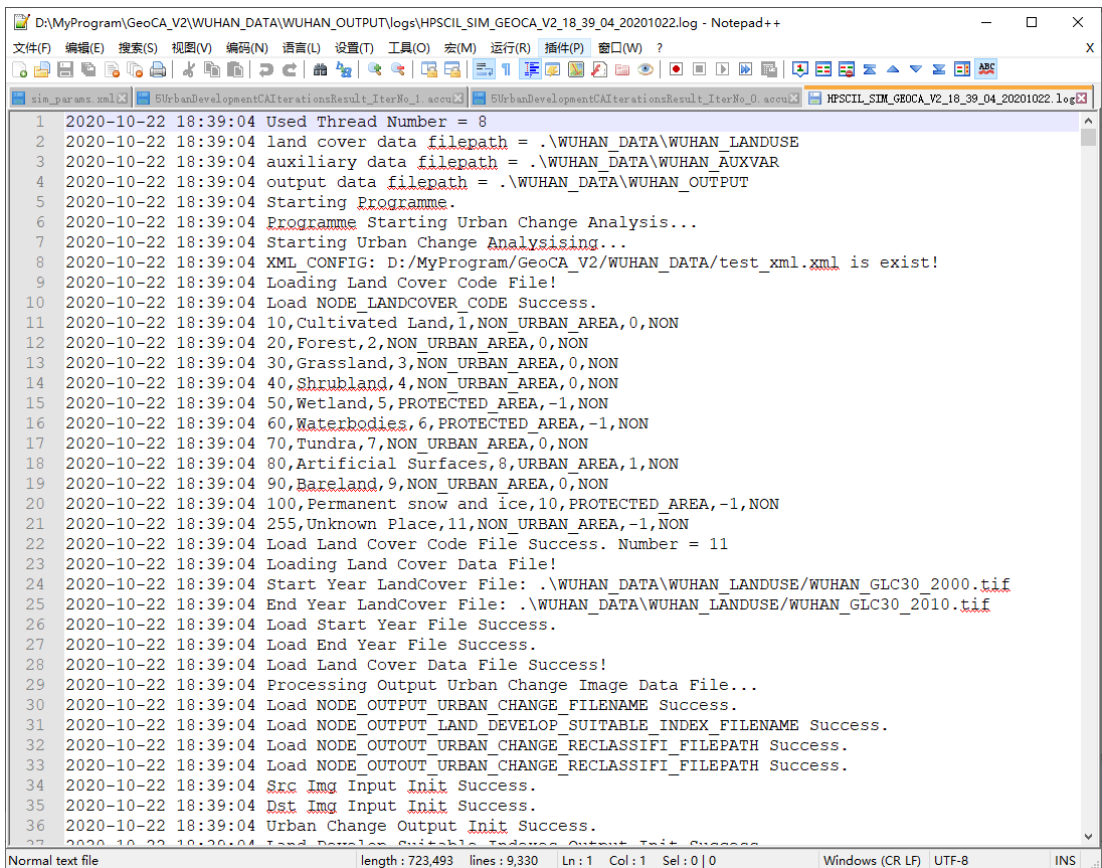


图 14 日志文件内容

参数文件

参数文件为 XML 格式。

XML 文件注释可查看程序目录下的 `.\sim_params.xml` 内的注释。

可以直接修改 XML 文件，修改程序目录下的 `.\run.bat` 直接执行模型。

联系我们

如果你在使用中遇到问题，请及时和我们联系。

请在邮件中附上模型导出的参数文件（*.xml）和模型执行错误时的日志文件（*.log）。

参考文献

- [1] Chen, D., Zhang, Y., Yao, Y., Hong, Y., Guan, Q., & Tu, W. (2019). Exploring the spatial differentiation of urbanization on two sides of the Hu Huanyong Line--based on nighttime light data and cellular automata. *Applied Geography*, 112, 102081.
- [2] Zhang, D., Liu, X., Wu, X., Yao, Y., Wu, X., & Chen, Y. (2019). Multiple intra-urban land use simulations and driving factors analysis: a case study in Huicheng, China. *GIScience & Remote Sensing*, 56(2), 282-308.
- [3] Yao, Y., Liu, X., Zhang, D., Liang, Z., & Zhang, Y. (2017). Simulation of Urban Expansion and Farmland Loss in China by Integrating Cellular Automata and Random Forest. *arXiv preprint arXiv:1705.05651*.
- [4] He, Y., Ai, B., Yao, Y., & Zhong, F. (2015). Deriving urban dynamic evolution rules from self-adaptive cellular automata with multi-temporal remote sensing images. *International Journal of Applied Earth Observation and Geoinformation*, 38, 164-174.
- [5] Li, X., & Yeh, A. G. O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16(4), 323-343.