UrbanM2M 城市扩张模拟软件 V1.0

用户手册

1 软件介绍

1.1 软件介绍

城市扩张是城市化的直接结果。城市扩张具有包括生态效应、经济社会效应等效应在内的多重效应,受到城市规划、生态环境等领域的关注。对未来城市扩张情形的模拟,有助于专业人员了解一个地区在未来一段时间内可能发生的城市扩张情况,进而促进对该地区未来城市发展、生态保护等情形的理解。因此,对未来城市扩张情景的模拟十分重要。

元胞自动机可以很好地模拟未来城市扩张情景。以往的元胞自动机模型多以像元为单位模拟城市扩张情形,且仅利用一期或两期的城市用地图进行模拟。这类模拟方法缺乏对城市扩张时空动态规律的深入挖掘能力。

就此,我们开发了UrbanM2M软件,以供用户进行更高可靠性的未来城市扩张情景模拟。UrbanM2M基于ConvLSTM(卷积长短时记忆网络)深度学习模型,使用时序城市用地栅格数据及空间变量,模拟未来城市扩张情景。UrbanM2M同时集成了数据归一化、精度评估等数据处理模块。UrbanM2M的前端界面基于Gradio开发,易于用户上手。

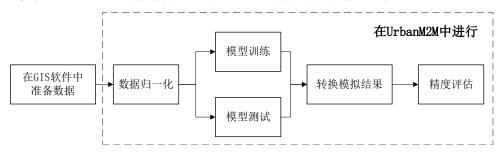


图 1-1 UrbanM2M 城市扩张模拟过程

2 软件运行环境

2.1 硬件环境

UrbanM2M 对计算机硬件要求较高,如下:

硬件	最低配置	推荐配置
内存容量	>=8GB	>=16GB
硬盘剩余容量	软件本体≈7GB+数据	>=20GB
显卡	英伟达显卡,显存>=4GB, 驱动版本>=452.39(不支持A卡)	英伟达显卡,显存>=8GB, 驱动版本>=452.39 (不支持 A 卡)

2.2 系统支持

UrbanM2M 支持在 Windows 上完美运行(包括核心算法包、Web GUI 界面);核心算法可以在 Linux 上完美运行,但 Web GUI 界面在 Linux 上不支持部分功能(文件、文件夹选择功能)。未曾在 Mac 系统上测试,不保证能运行成功。

2.3 UrbanM2M 下载

从 Github 下载项目 zip 压缩包或克隆 UrbanM2M 项目

2.4 运行环境准备

运行 UrbanM2M 要求的环境为:

Python==3.10.x PyTorch==1.10.0 GDAL==3.x CUDA==11.3 cuDNN=8.2.1 gradio==3.x

如果用户的设备中有满足要求的 Python 环境且 CUDA/cuDNN 版本符合要求,可以跳过以下步骤。否则建议下载环境压缩包。

(链接: https://pan.baidu.com/s/1YzYKjb3hXWjbVtwn0Vpesw?pwd=0td9) 将 pym2m.tar.gz 解压至 Urbanm2m 中。解压后的文件夹结构如下。

名称	~ 修改日期	类型	大小
igit .git	2023/11/07 10:29	文件夹	
idea .idea	2023/11/07 15:06	文件夹	
config	2023/11/07 15:05	文件夹	
gr_utils	2023/11/07 15:05	文件夹	
m2m_core	2023/11/06 16:03	文件夹	
pym2m	2023/08/28 15:43	文件夹	
trained_models	2023/10/30 14:31	文件夹	
gitignore	2023/08/29 15:19	Git Ignore 源文件	1 KB
config_py.py	2023/11/07 10:36	Python 源文件	4 KB
fom_gui.py	2023/10/18 13:58	Python 源文件	1 KB
init.bat	2023/11/07 10:23	Windows 批处理	1 KB
prob2sim_gui.py	2023/09/06 21:19	Python 源文件	1 KB
🌌 pym2m.rar	2023/08/28 15:58	WinRAR 压缩文件	2,254,223
readme.md	2023/10/30 13:35	Markdown 源文件	4 KB
split_gui.py	2023/10/30 15:08	Python 源文件	1 KB
e test_gui.py	2023/10/30 15:02	Python 源文件	1 KB
train_gui.py	2023/10/30 13:57	Python 源文件	2 KB
webUI.py	2023/11/07 10:41	Python 源文件	18 KB

图 2-1 Urbanm2m 文件夹结构

3 数据准备

3.1 概述

与大部分 CA 模型不同, UrbanM2M 使用时间序列城市用地栅格及空间变量预测未来城市扩张情形,这要求用户准备连续多年的城市用地栅格。且 UrbanM2M 以栅格切片(而不是大部分模型使用的像元)为单位预测转换概率。这要求用户准备的城市用地栅格、空间变量等比研究区域略大一圈,以保证研究区边缘区域栅格切片的完整。因此, UrbanM2M 对数据的空间一致性要求较高。

本文档的第三、第四章将以模拟长三角太湖流域的城市扩张为例,介绍数据**准备流程及** UrbanM2M 使用流程。其中本章介绍在 GIS 软件中准备 UrbanM2M 输入数据的过程。

用户也可以下载已经准备好的示例数据:

https://pan.baidu.com/s/17Rj-qi28VVWd3HTqhqwheQ?pwd=81g9

3.2 数据基本要求

- (1) 所有数据坐标系一致。
- (2) 强烈建议将所有数据转换至投影坐标系。
- (3) 所有栅格数据行列数目一致。

3.3 预备

在开始之前,用户需要创建一个文件夹,储存数据。并创建两个空的文件夹, vars 及 year。创建好后,文件夹结构如下。



图 3-1 数据根目录结构准备

以下步骤中, 我们将以长三角为例进行说明。

3.4 创建研究区缓冲区矢量

在这一步, 用户需要准备的数据如下:

表 1 创建研究区缓冲区矢量——数据准备

数据	说明
研究区域矢量文件	/



图 3-2 研究区域矢量文件

使用 GIS 软件的【缓冲区】功能创建缓冲了的研究区矢量(后文简称**缓冲区矢量**)。缓冲区矢量用于提取比研究区域更大一圈的用地、变量栅格等。在大多数 GIS 软件中,缓冲区的参数要求如下:

- i. 输入数据。请根据用户的数据进行确定。
- ii. 输出数据。可以自行确定。
- iii. 缓冲半径单位与数值。请保证单位为米 (Meters)。请保证缓冲半径至少是用户将要使用的城市用地栅格分辨率的 70 倍。如:使用 30 米分辨率的栅格时,至少设为 2100 米;使用 100 米分辨率栅格时,至少设为 7000 米。
- iv. 融合类型。选择融合为一个要素。

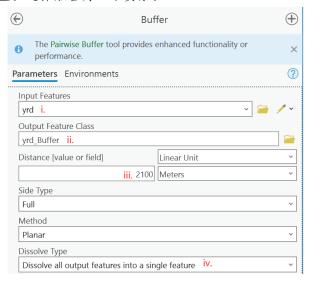


图 3-3 缓冲区创建参数



图 3-4 缓冲区创建结果

3.5 准备用地栅格

这一步用户需要用到的数据如下:

表 2 准备用地栅格——数据准备

数据	
时间序列城市用地栅格(更大范围)	/
5.4 中得到的缓冲区矢量	/

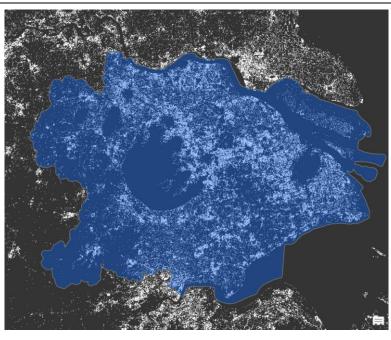


图 3-5 研究区城市用地栅格示意

在开始这一步前,请确保用户的用地栅格的范围至少和上一步中得到的缓冲区矢量一样大(而不是和研究区一样大)。

这一步使用缓冲区矢量,得到缓冲区范围内的多幅时间序列的城市用地栅格。用户需要 重复进行以下步骤多次。 使用【按掩膜提取工具】,设置如下参数,逐个提取每一年的城市用地栅格。(也可以用 arcpy 等方法批量处理)

- i. 输入栅格: 当年的城市用地栅格
- ii. 掩膜:缓冲区矢量。
- **iii.** 输出:必须保存在数据根目录的 year 文件夹中,且必须命名为 land_{year}. tif。如将 2011 年的矢量命名为 land_2011. tif。**如果命名不正确,后续程序将无法**识别相关内容。
- iv. 处理范围:选择为缓冲区矢量。

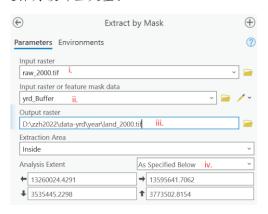


图 3-6 准备城市用地——参数设置

3.6 准备研究区范围栅格 range. tif

这一步需要准备的数据如下:

表 3 准备研究区范围栅格——数据准备

数据	说明
任意一年的研究区城市用地栅格	作为模板栅格
研究区矢量	/
缓冲区矢量	/

(1) 为研究区矢量添加一个长整形字段,使用字段计算器,设置其值为1。



图 3-7 为研究区矢量添加长整形字段

- (2)使用【面转栅格】工具,将研究区矢量转为值为1的栅格。参数及环境设置如下:
 - i. 输入:研究区矢量(而不是缓冲区矢量)。
 - ii. 值字段:在上一小步中设为1的字段。

- iii. 输出:必须保存在5.3 创建的数据文件夹根目录,命名必须为 range. tif。
- iv. 栅格大小:和用户的用地栅格保持一致。
- v. 环境处理-坐标系:和用户的用地栅格保持一致。
- vi. 环境设置-处理范围:设为任意一年的研究区城市用地栅格。
- vii. 环境设置-捕捉栅格:设为任意一年的研究区城市用地栅格。

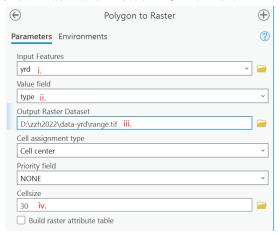


图 3-8 准备研究区范围栅格——参数设置

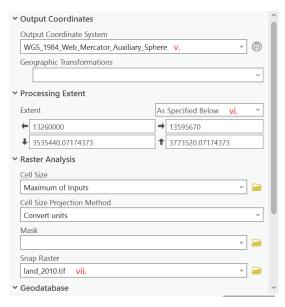


图 3-9 准备研究区范围栅格——环境设置

如果创建无误,在 3.5 中创建的用地栅格和本步中创建的范围栅格应当具有相同的行列数目。

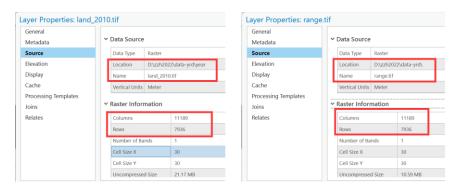


图 3-10 成功的创建

3.7 准备约束栅格 restriction. tif

约束栅格用于确定哪些区域是不可用于发展城市的区域。在约束栅格中 1 值代表不可发展,其他值表示不做约束。如果用户不打算进行约束,可以跳过这一步。在后续步骤,如果缺失 restriction. tif,程序会自动生成全零栅格,表示不做任何约束。

这一步需要准备的数据如下:

表 4 准备约束栅格——数据准备

数据	说明
限制区域矢量	如水体、保护区等; 空间坐标系与城市用地栅格相同

在准备好限制区域矢量后,首先为矢量添加一个长整形字段,使用字段计算器,设置其值为1。



图 3-11 添加长整形字段

随后使用【面转栅格】工具,得到约束栅格,参数如下:

- i. 输入:约束区域矢量。
- ii. 值字段:在上一小步中设为1的字段。
- iii. 输出:必须保存在5.3 创建的数据根目录,命名必须为 restriction. tif。
- iv. 栅格大小:和用户的用地栅格保持一致。
- v. 环境处理-坐标系:和用户的用地栅格保持一致。

vi. 环境设置-处理范围:设为任意一年的研究区城市用地栅格。 vii. 环境设置-捕捉栅格:设为任意一年的研究区城市用地栅格。

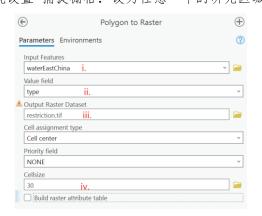


图 3-12 准备约束栅格——参数设置

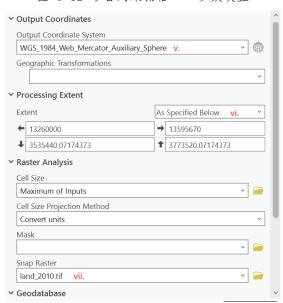


图 3-13 准备约束栅格——环境设置

处理后得到的栅格示意如下:

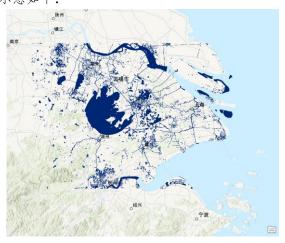


图 3-14 约束栅格示意

成功创建的约束栅格的行列数也应当与用地栅格一致。

3.8 准备空间变量栅格

这一步需要准备的数据如下:

表 5 准备空间变量栅格——数据准备

数据	说明
用于生成空间变量的数据	如 DEM,城市中心点
缓冲区矢量	/

用户可以根据自己的需求选择空间变量。但如果用户准备使用我们提供的训练好了的模型,用户需要准备的空间变量必须是:坡度、距离县(县级市、区)中心距离及距离镇中心距离。

以下以准备坡度、距离镇中心距离为例,说明空间变量栅格准备过程。

3.8.1 准备坡度栅格

首先准备 DEM 文件。用户可以从网络上下载分辨率合适的 DEM 文件。 使用 GIS 软件中的【坡度】工具,参数设置如下:

- i. 输入:用户准备的DEM。
- ii. 输出: 自由设定,不要放在 vars 文件夹下即可。
- iii. 输出测量单位: DEGREE。
- iv. 环境处理-坐标系:和用户的用地栅格保持一致。
- v. 环境设置-处理范围:选择5.5中得到的任意一年的城市用地栅格即可。
- vi. 环境设置-捕捉栅格:选择 5.5 中得到的任意一年的城市用地栅格即可。

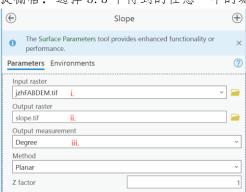


图 3-15 准备坡度变量——参数设置

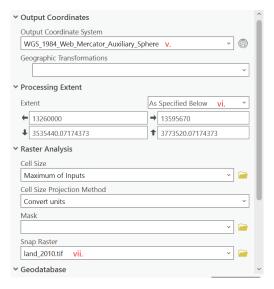


图 3-16 准备坡度变量——环境设置

3.8.2 准备距离镇中心距离栅格

以准备距离镇中心距离栅格为例进行说明。准备其他距离栅格的步骤是类似的。(注意,在非正版 ArcGIS Pro 中计算较大范围的欧式距离可能输出无效栅格。如遇这种情况,建议适用 ArcMap 或 QGIS 计算距离栅格)

首先准备镇中心点矢量文件。

使用 GIS 软件中的【欧式距离工具】计算距离镇中心距离,参数及环境设置如下:

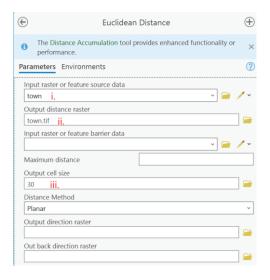


图 3-17 准备距离变量——环境设置

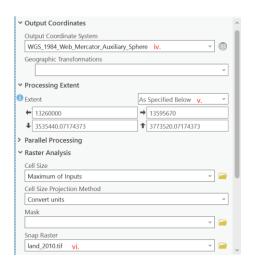


图 3-18 准备距离变量——环境设置

- i. 输入:用户准备的点矢量。
- ii. 输出: 自由设定,不要放在 vars 文件夹下即可。
- iii. 输出栅格大小:与用地栅格保持一致。
- iv. 环境处理-坐标系:与用地栅格保持一致。
- v. 环境设置-处理范围:选择 5.5 中得到的任意一年的城市用地栅格即可。
- vi. 环境设置-捕捉栅格:选择5.5中得到的任意一年的城市用地栅格即可。

3.9 数据文件夹结构

在完成上述步骤后,数据根目录中应当有 vars/year 两个文件夹及 range. tif, restriction. tif 两个栅格。year 文件夹中存有用地栅格,vars 文件夹尚空(将用于在后续步骤中归一化得到的结果)。

4 启动 UrbanM2M WebUI

4.1 使用 pym2m 环境:

如果选择了下载提供的 Python 环境 (pym2m), 右击 init. bat, 以管理员身份运行。在浏览器中访问 https://127.0.0.1/7860。

4.2 使用自带的环境:

如果选择使用自带的 Python 环境:

- (1) 右击 init. bat, 编辑。
- (2) 将第 4 行的./pym2m/python. exe 替换为用户的 python. exe 路径。
- (3) 保存。

右击 init. bat,以管理员身份运行。在浏览器中访问 https://127.0.0.1/7860。

4.3 UrbanM2M WebUI 界面介绍

UrbanM2M WebUI 包含数据路径设置、主题处理流程选项卡及辅助工具三部分。在启动后, Data dir 是必须设置的。而后用户可根据需求选用流程选项卡中的功能。

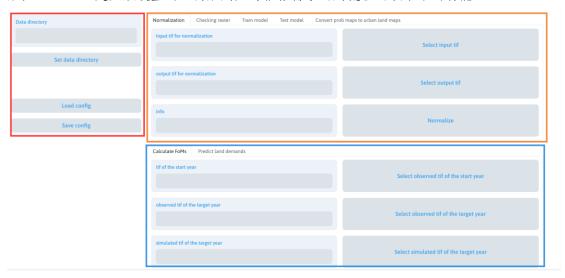


图 4-1 UrbanM2M WebUI 界面

5 UrbanM2M 核心流程

5.1 UrbanM2M 使用注意事项

- (1) 请确保第三章中生成的所有栅格都具有相同的空间范围、行列数目、坐标系。 否则可能出现意外的错误结果。
- (2) 如果希望进行跨区域模型迁移,训练模型区域与迁移目标区域的范围不应相差 太大,否则效果可能不甚理想。
- (3) 如果希望进行跨区域模型迁移,但训练模型区域远大于迁移目标区域时,较大

的用地需求量可能得到较为混乱的结果。

5.2 设置数据根目录

在 Data directory 文本框中输入用户的数据根目录(或点击 Set data directory 设置)即可。如果没有设置数据根目录,几乎所有功能都会报错,因此请务必先设置数据根目录,再使用其他功能。

后文以\$data dir 表示数据根目录。

5.3 数据归一化(Normalization)

5.3.1 说明

数据归一化用于保证输入的栅格符合 ConvLSTM 的输入要求。输入 ConvLSTM 的空间变量应当处于 0-1 的范围内。

在【Normalization】选项卡中,在【input tif】中选择需要归一化的栅格,在【output tif】选择需要保存路径。

请注意**,请确保保存的结果位于\$data_dir/vars 文件夹中。**这样才能保证程序正确读取空间变量。

如果用户希望使用解压后就位于 urbanm2m/trained_models 当中的,提前训练好的模型,应当对坡度、距离镇中心距离、距离县中心距离三个变量进行归一化,并命名为 slope. tif, town. tif, county. tif。



图 5-1 归一化界面

5.3.2 案例

第四章中准备的空间变量栅格为 county.tif, town.tif, slope.tif。分别对它们进行归一化,并保存在数据文件夹的 vars 路径下。

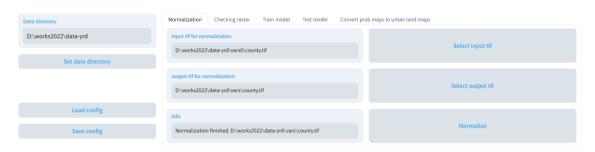


图 5-2 归一化案例

5.4 检查栅格可用性

5.4.1 说明

UrbanM2M 对栅格数据的空间一致性要求较高,如果输入不符合要求的栅格可能得到意想不到的错误结果。因此强烈建议在训练或测试模型前使用【Checking rasters】模块检查\$data_dir下的栅格数据是否符合要求。该模块会读取\$data_dir,并【Start year】和【End year】间的用地栅格是否符合要求。在检查(无误)后,会在 range.tif 的图片框中展示研究区域的二值图像。

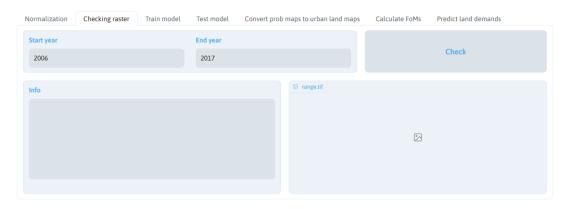


图 5-3 检查栅格可用性模块界面

该模块检查的内容包括:

- (1) 范围栅格(range.tif)是否存在,vars 文件夹下是否至少有一个空间变量栅格, years 文件夹下是否有【Start year】和【End year】之间的所有用地栅格。
- (2) 范围栅格(range.tif),约束栅格(restriction.tif,如果有),vars 文件夹下的空间变量栅格,years 文件夹下的用地栅格是否具有完全相同的空间参考及栅格行列数。
- (3) 范围栅格和空间变量栅格的有效范围是否大于用地栅格。(必须大于,以保证边缘切片能被无误地训练与预测)。
- (4) 空间变量栅格是否被归一化; 用地栅格是否由 0,1 值组成。

5.4.2 案例

设置起始年份为 2000,终止年份为 2017,运行。检查无误后如下:

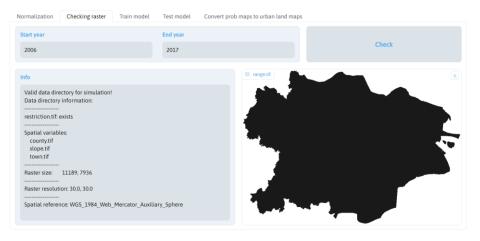


图 5-4 检查栅格可用性案例

5.5 模型训练 (可选)

5.5.1 概述

训练 ConvLSTM 模型用于预测用地转换概率图。模型的输入为:空间变量+多年城市用地栅格切片;输入为:未来多年用地转换概率图。

5.5.2 训练数据设置

确定用于训练模型的数据,包括四个参数:训练起始年份、训练输入长度、训练输出长度及空间变量。

空间变量(Spatial variables): 训练模型使用的空间变量。请直接选择\$data_dir/vars下需要作为输入的空间变量切片。

训练起始年份 (First observed year): 训练数据中的第一年。如希望使用 2000-2005 年作为输入、2006-2011 作为输出训练模型,训练起始年份为 2000。

训练输入长度(Count of years to input): 训练时的输入年份数目。如希望使用 2000-2005 年作为输入、2006-2011 作为输出训练模型,训练输入长度为 6。

训练输出长度(Count of years to output): 训练时的预测年份数目。如希望使用 2000-2005 年作为输入、2006-2011 作为输出训练模型,训练输出长度为 6。

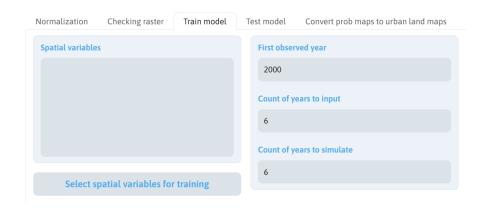


图 5-5 模型训练数据设置界面

5.5.3 模型参数设置

模型参数设置说明如下:

Batch size	Learning rate
8	0.00001
Count of sampled tiles	Validation proportion
Count of sampled tiles for training	Validation proportion 0.1
•	

图 5-6 模型训练参数设置界面

批量大小 (Batch size): 影响模型训练效果。批量大小应为整数,默认且建议为 8。 批量大小越大,训练时占用的显存越大。每 1 批量大小占用的显存约为 600MB。请根据用户的 GPU 显存调整合适的大小。

学习率 (Learning rate): 影响模型训练速度。学习率越大,训练越快。但过大的学习率可能导致模型无法收敛。默认为 0.00001,非常不建议将学习率调整为大于 0.0001 的数。

训练切片数目(Count of sampled tiles for training): 用于训练的样本数目。这个值应当小于在tile_train文件夹中的文件夹数目(即可用的切片数目)。**建议至少为2000**。使用过少的训练样本可能无法得到效果较好的模型。

验证比例 (Validation proportion): 训练样本中用于验证的比例。默认为 0.1。

5.5.4 训练过程

点击【Train】后,会出现一个 cmd 窗口,其中将显示训练过程信息。

在训练过程中,请观察每一轮训练完成后进度条后方显示的"avg_loss"(即平均损失函数)。当该值在多轮训练中不再明显下降,甚至略有上升时,可以按下ctrl+z终止训练。(建议至少训练20轮)

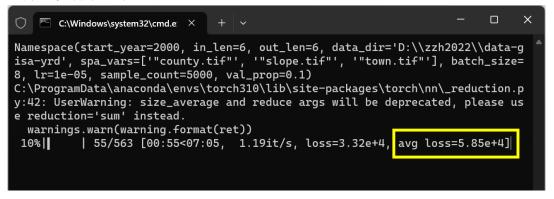


图 5-7 模型训练过程中出现的 cmd

训练得到的模型将被储存在 trained_models 文件夹中。模型的命名规则如下



图 5-8 模型名称解释

5.5.5 案例

选择准备好的三个空间变量,设置训练起始年份为 2000 年,输入 6 年,输出 6 年。进行训练。训练 40 轮后停止训练。

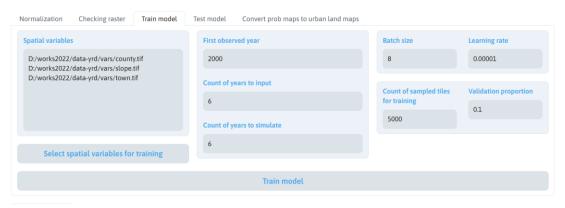


图 5-9 训练模型案例

5.6 模型测试

5.6.1 概述

使用训练好的模型,预测用地转换概率图。模型每次预测一个切片的转换概率图,在完成所有切片的预测后,程序自动将预测好的转换概率图合并为研究区域的完整的转换概率图。

5.6.2 测试数据设置

阶段需要设置五个参数。

起始输入年份 (First observed year): 测试数据中的第一年。如希望使用 2006-2011 年作为输入、预测 2012-2017 的转换概率,测试起始年份为 2006。

预测起始年份 (First year to simulate): 首个预测的年份。如希望使用 2006-2011 年作为输入、预测 2012-2017 的转换概率,预测起始年份为 2012。

测试输出长度 (Count of years to output): 测试年份数目。如希望使用 2006-2011 年作为输入、预测 2012-2017 的转换概率,测试年份数目为 6。

转换概率图保存文件夹 (Dir for probability maps): 保存转换概率图的文件夹。建议放在\$data_dir 下。

模型路径 (pth model file): 训练好的模型路径。

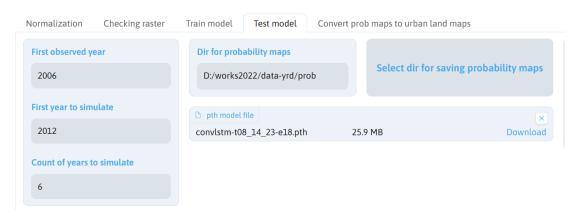


图 5-10 模型测试数据设置

5.6.3 参数设置

预测阶段, 可以设置两个影响预测速度的参数。

批量大小 (Batch size): 批量大小越大, 预测速度越快。但越大的批量大小要求更大的 GPU 显存。每 1 批量大小占用约 40Mb 显存。请根据用户的 GPU 显存调整合适的大小。

线程数目(num_worker for torch. loader):如果用户不了解 PyTorch 的 torch. loader 的机制,将参数保持为 0 即可,否则可能导致预测失败。如果了解 PyTorch 的 torch. loader 的机制,可以根据需求调整。



图 5-11 模型测试参数设置

5.6.4 案例

以 2006 年为首个输入年份,2012 年为预测起始年份,预测六年,选用训练好的模型进行测试,将概率上颚保存在\$data_dir/prob 下。点击 Test model 开始预测。进度条会显示在info 文本框中。

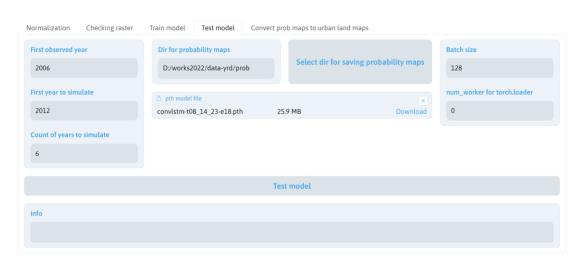


图 5-12 测试模型案例

5.7 将概率图转为用地预测结果

5.7.1 概述

本步骤根据用地需求增量,将转换概率图中转换概率最大的若干像元变为城市用地。

5.7.2 参数设置

需要设置四个参数。

最后已知用地图 (Final observed landmap):最后一年的已知用地图。如在上一步【测试】中使用了 2006-2011 年作为输入,则应选择\$data_dir/year/land_2011.tif 作为输入。

转换概率图保存文件夹 (Dir for probability maps): 上一步中设置的保存转换概率 图的文件夹。

结果保存文件夹 (Dir for simulated maps): 保存预测结果的文件夹。建议放在 \$data dir 下。

用地需求增量(Land demands): 每一个预测年份的城市用地需求增量(栅格数量)。每年的需求量间以逗号划分。请保证用地需求量的数目等于预测年份的数目。(请注意逗号必须是英文逗号)

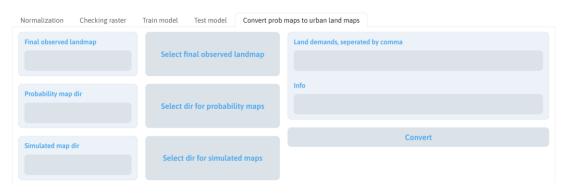


图 5-13 结果转换界面

5.7.3 案例

因预测起始年份为 2012 年,选择最后已知用地图\$data dir/year/land 2011.tif,选择概

率图文件夹为上一步中的设置,将模拟结果保存在\$data_dir/sim下。

案例为测试模型效果,使用真实用地数据进行结果验证,故设置用地需求量为 2012-2017 年间实际城市用地增量。设置用地增量为 584425,661210,431138,491400,449350,790081。

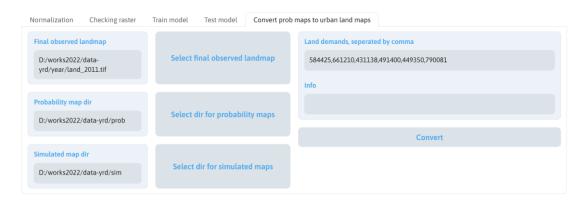


图 5-14 结果转换案例

6 工具模块介绍

6.1 概述

7 模型效果验证(可选)

如果有预测年份的真实用地数据,可以使用模型效果验证模块计算模拟结果的 FoM (Figure of Merit,命中率)。计算结果换显示在选项卡右侧的【info】中。 参数包括:

最后已知用地图 (Final observed landmap):最后一年的已知用地图。如在【测试】中使用了 2006-2011 年作为输入,则应选择\$data_dir/year/land_2011.tif 作为输入。

目标年份真实用地图 (target observed map): 目标年份的真实用地图。目标年份预测结果 (simulated map): 目标年份的预测结果。

□ final_observed_map	拖放文件至此处 - 或 - 点击上传
□ target_observed_map	拖放文件至此处 -或- 点击上传
□ simulated_map	the thereto (the CT the lot)
	拖放文件至此处
	- 或 -
	点击上传

图 7-1 模型效果验证界面